Pertanika Journal of

SCIENCE & TECHNOLOGY

JST

VOL. 22 (2) JUL. 2014

PERTANIKA JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

About the Journal

Pertanika is an international peer-reviewed journal devoted to the publication of original papers, and it serves as a forum for practical approaches to improving quality in issues pertaining to tropical agriculture and its related fields. *Pertanika* began publication in 1978 as the Journal of Tropical Agricultural Science. In 1992, a decision was made to streamline *Pertanika* into three journals to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university. The revamped Journal of Science & Technology (JST) aims to develop as a pioneer journal focusing on research in science and engineering, and its related fields. Other *Pertanika* series include Journal of Tropical Agricultural Science (JTAS); and Journal of Social Sciences and Humanities (JSSH).

JST is published in **English** and it is open to authors around the world regardless of the nationality. It is currently published two times a year, i.e. in **January** and **July**.

Goal of Pertanika

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 12 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

Indexing of Pertanika

Pertanika is now over 33 years old; this accumulated knowledge has resulted in Pertanika JST being indexed in SCOPUS (Elsevier), EBSCO, Thomson (ISI) Web of Knowledge [CAB Abstracts], DOAJ, Google Scholar, ERA, ISC, Citefactor, Rubriq and MyAIS.

Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

We also have views on the future of our journals. The emergence of the online medium as the predominant vehicle for the 'consumption' and distribution of much academic research will be the ultimate instrument in the dissemination of research news to our scientists and readers.

Aims and scope

Pertanika Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: *bioinformatics, bioscience, biotechnology and biomolecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management,* and related fields of study.

Editorial Statement

Pertanika is the official journal of Universiti Putra Malaysia. The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. Technol.*

JST
Journal of Science & Technology
AN INTERNATIONAL PEER-REVIEWED JOURNAL

EDITORIAL OFFICE
JOURNAL DIVISION
Office of the Deputy Vice Chancellor (R&I)
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia.
Gen Enq.: +603 8947 1622 | 1619 | 1616
E-mail: executive_editor.pertanika@upm.my
URL: www.journals-jd.upm.edu.my

PUBLISHER
Kamariah Mohd Saidin
UPM Press
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.
Tel: +603 8946 8855, 8946 8854
Fax: +603 8941 6172
E-mail: penerbit@putra.upm.edu.my
URL: http://penerbit.upm.edu.my

ABSTRACTING/INDEXING

Pertanika is now over 35 years old; this accumulated knowledge has resulted the journals being indexed in SCOPUS (Elsevier), Thomson (ISI) Web of Knowledge [BIOSIS & CAB Abstracts], EBSCO, DOAJ, Google Scholar, AGRICOLA, ISC, Citefactor, Rubriq and MyAIS. JST is also indexed in ERA.

Pertanika Journal of Science & Technology
Vol. 22 (2) Jul. 2014

Contents

Foreword

Welcome to the **Second Issue 2014** of the Journal of Science and Technology (JST)!

JST is an open-access journal for studies in science and technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university and is run on a non-profit basis for the benefit of the world-wide science community.

In this issue, **23 articles** are published, out of which **one** is an invited review, **two** are review articles, **13** are regular articles and **one** is a case report. **Six articles** are from Curtin University's Technology, Science and Engineering International Conference "Engineering Goes Green" (CUTSE 2012). The authors of these articles vary in country of origin, coming from **Malaysia**, **Japan**, **Nigeria**, **India**, **Singapore** and **Ireland**.

The invited review is an in-depth discussion of bifurcation/chaos and its practical relevance to chemical and biological systems (*Said S. E. H. Elnashaie*). The first review article in this issue discusses plate solar collectors and applications (*Bande, Y. M.* and *Mariah, N. A.*) while the second review article is on the development of vaccination strategies from BCG to new vaccine candidates (*Nadiya T. Al-alusi* and *Mahmood A. Abdullah*).

The regular articles cover a wide range of study. The first paper is on comparison of physical activity prevalence among government employees in Kangar, Perlis, Malaysia using an international physical activity questionnaire (IPAQ) that measures steps/day based on an accelerometer(*Hazizi, A. S., Zahratul Nur, K., Mohd Nasir, M. T., Zaitun, Y.* and *Tabata, I.*). The following articles look at: hypothesis tests of goodness-of-fit for Fréchet distribution (*Abidin, N. Z., Adam, M. B.* and *Midi, H.*); distribution of recent Ostracoda in offshore sediment of the South China Sea (*Ramlan, O.* and *Noraswana, N. F.*); empirical correlation of refrigerant HC290/HC600a/HFC407C mixture in adiabatic capillary tube using statistical experimental design (*Shodiya, S., Azhar, A. A.* and *Darus, A. N.*); identification of hot spots in proteins using modified gabor wavelet transform (*D. K. Shakya, Rajiv Saxena* and *S.N. Sharma*); self-care behaviour among type 2 diabetes patients (*Siti Khuzaimah, A. S., Aini, A., Surindar Kaur, S. S., Hayati Adilin, M. A. M.* and *Padma, A. R.*); a reduced τ-adic Naf (RTNAF) representation for an efficient scalar multiplication on anomalous binary curves (ABC) (*Faridah Yunos, Kamel Ariffin Mohd Atan, Muhammad Rezal Kamel Ariffin* and *Mohamad Rushdan Md Said*); relevance of integrated geophysical methods for site characterisation in the construction industry – a case of Apa in Badagry, Lagos State, Nigeria (*K. S. Ishola, L. Adeoti, F. Sawyerr* and *K. A. N Adiat*); antimicrobial activities of three different seed extracts of *Lansium* varieties (*H. Alimon, A. Abdullah Sani, S. S. Syed Abdul Azziz, N. Daud, N. Mohd Arriffin* and *Y. Mhd Bakri*); a study on the performance of multivariate exponentially weighted moving average (MEWMA) and multivariate synthetic charts (*Ellappan, S.* and *Khoo Michael, B. C.*); an experimental and modelling study of selected heavy metal removal from aqueous

solution using *Scylla serrata* as biosorbent (*Aris A. Z., Ismail F. A., Ng, H. Y.* and *Praveena, S. M.*); Bayesian network classification of gastrointestinal bleeding (*Nazziwa Aisha, Mohd Bakri Adam, Shamarina Shohaimi* and *Aida Mustapha*); removal of toluene from aqueous solutions using oil palm shell based activated carbon: equilibrium and kinetics study (*Kwong, W. Z., Tan, I. A. W., Rosli, N. A.* and *Lim, L. L. P.*) and one case report on the delay in diagnosis of lung cancer (*Ching, S. M., Chia, Y. C.* and *Cheong, A. T.*).

I conclude this issue with six articles arising from the CUTSE 2012 international conference: intelligent monitoring interfaces for coal-fired power plant boiler trips: a review (*Nistah, N. N. M., Motalebi, F., Samyudia, Y.* and *Alnaimi, F. B. I.*); storm runoff pollution from a residential catchment in Miri, Sarawak (*Ho, C. L. I.* and *Choo, B. Q.*); septage treatment using pilot vertical-flow engineered wetlands system (*Jong, V. S. W.* and *Tang, F. E.*); improvement of engineering properties of peat with palm oil clinker (*M. E. Rahman, M. Leblouba* and *V. Pakrashi*); optimisation and analysis of bioethanol production from cassava starch hydrolysis (*Liew, E. W. T.*); and numerical modelling of molten carbonate fuel cell: effects of gas flow direction in anode and cathode (*Tay C. L.* and *Law M. C.*).

I anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking, and, hopefully, useful in setting up new milestones. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

I would also like to express my gratitude to all the contributors, namely, the authors, reviewers and editors for their professional contribution towards making this issue feasible. Last but not least, the editorial assistance of the journal division staff is fully appreciated.

JST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

Chief Executive Editor
Nayan Deep S. KANWAL, FRSA, ABIM, AMIS, Ph.D.
nayan@upm.my

Pertanika J. Sci. & Technol. 22 (2): 337 - 364 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Invited Review

Bifurcation/Chaos and Their Practical Relevance to Chemical and Biological Systems

Said S. E. H. Elnashaie[1,2]

[1]Chemical and Environmental Engineering Department, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Chemical and Biological Engineering Department, University of British Columbia (UBC), Vancouver, Canada

ABSTRACT

Bifurcation and chaos are important phenomena affecting many physical and chemical systems. They are also related to the stability/instability and multiplicity phenomena associated with these systems. The phenomena are not only of theoretical/mathematical interest but are also important for laboratory, pilot plant and commercial units. This paper concentrates on 3 systems:

1. The novel auto-thermic Circulating Fluidized Membrane Steam Reformer (CFBMSR) for the efficient production of the clean fuel hydrogen and which shows multiplicity of the steady state (static bifurcation)
2. A novel fermentor for the efficient production of bio-ethanol that shows static/dynamic bifurcation as well as chaotic behaviour
3. The neurocycle of the acetylcholine transmitter in the brain using diffusion-reaction models in order to gain insight into their possible connection to Alzheimer and Parkinson Diseases (AD/PD); these are preliminary efforts to investigate the bifurcation and chaotic behaviour of this neurocycle.

Keywords: Bifurcation, multiplicity, chaos, CFBMSR, chemical/biological reaction engineering, neuro-cycle, Alzheimer and Parkinson Diseases (AD/PD)

INTRODUCTION

Multiplicity of steady states in chemically reactive systems was first observed by Liljernoth (1919). However, it was not until the 1950s that great interest in this

E-mail address:
selnashaie@gmail.com (Said S. E. H. Elnashaie)

ISSN: 0128-7702

phenomenon was generated, inspired by the Minnesota school of Aris, Amundson and others (1958) and their students (e.g. Balakotaiah & Luss, 1981; 1983). Uppal, Ray and Poore (1974), Ipsen and Schreiber (2000) and Melo *et al.* (2001) added on to the basic knowledge and understanding regarding these phenomena. The bifurcation behaviour of non-isothermal bubbling fluidized bed catalytic reactors has been extensively investigated by Elnashaie and co-workers (1973; 1977; 1989; 1980; 1996). Earnshaw and Keener (2010) investigated in a more general manner the global asymptotic stability of non-autonomous master equations, while Danforth (2013) investigated chaos of atmosphere hanging in a wall. Dingqi *et al*. (2011) studied in some detail the combustion of coal based on the chaos theory. For more general applications, Werndl (2009) investigated the implications of chaos on unpredictability and Jua'rez (2011) investigated health and prediction applications.

Excellent reviews on the subject were published by Ray (1977), Bailey (1977, 1998), Gray and Scott (1994), Elnashaie and Elshishini (1996) and Epstein and Pojman (1998). Important mathematical principles are given by Golubitsky and Schaeffer in their book published in 1985. These important phenomena do not receive proper attention in chemical and biological engineering education and research. This state of affairs is perhaps due to lack of understanding regarding the practical importance and implications of these phenomena (Elnashaie & Elshishini, 1996). In this review, three chemical and biological reaction engineering systems are presented which are strongly affected by these phenomena. There are other systems also affected by these phenomena, such as Fluid Catalytic Cracking (FCC) for the production of high-octane-number gasoline (Elnashaie & Elshishini, 1996), UNIPOL process for the production of polyethylene and polypropylene (e.g. Choi & Ray, 1985) as well as other polymerisation reaction engineering systems (e.g. Teymour & Ray, 1992) as well as catalytic carbon mono-oxide oxidation (e.g. Razon & Schmitz, 1986), just to mention a few. This paper concentrates on only three main systems; two are related to clean fuel production and the third is related to modeling of the brain's neurocycle transmission process. The first mathematical reporting of period doubling to chaos was by Feigenbaum on 1980.

The first system considered in this paper is efficient production of hydrogen in the novel auto-thermic CFBMSR; the second is efficient production of the clean fuel ethanol by fermentation; and the third is the reaction-diffusion modeling of the behaviour of the acetylcholine neurotransmitter in the brain.

Hydrogen is becoming a promising clean energy source for the future (Goltsov & Vezirohlu, 2002; Ohi, 2002; Elnashaie, 2005); it is mainly produced by the steam reforming of hydrocarbons (Tottrup, 1982; Elnashaie & Elshishini, 1993; Christensen, 1996). However, the conventional fixed bed steam reforming process is inefficient, highly polluting and suffers from catalyst deactivation (Rostrup-Nielsen, 1977; Elnashaie et. al., 1988; Twigg, 1989; Elnashaie & Elshishini, 1993). One of the main bottlenecks in this classical configuration is the extremely low catalyst effectiveness factor in the range of 0.001-0.01 (Elnashaie & Elshishini, 1993). The catalyst effectiveness factor has been considerably improved using fine catalyst particles in the bubbling fluidized bed (Elnashaie & Adris, 1989; Adris *et al*., 1991, 1994a, b, 1997, 2002). The second main bottleneck is the thermodynamic equilibrium limitation associated with the reversibility of the reactions. This limitation has been "broken" using permeable hydrogen membranes (Elnashaie & Adris, 1989; Adris *et al*., 1991; 1994a, b;

1997; 2002). Recently, a novel CFBMSR has been shown to be more efficient and more flexible for pure hydrogen production, through which a wide range of hydrocarbons such as natural gas, naphtha, gasoline, diesel and bio-oils can be processed for use as feedstock (Chen & Elnashaie, 2002; Prasad & Elnashaie, 2002; Chen *et al*., 2003a & b). The third bottleneck is associated with the inefficiency of heat supply to the highly endothermic reforming reactions through the huge top/side-fired furnaces (Twigg, 1989; Elnashaie & Elshishini, 1993). This limitation is "broken" using auto-thermal operation in a reaction-regeneration process, which "breaks" not only the heat transfer limitation, but also the carbon formation/catalyst regeneration limitation (Chen & Elnashaie, 2005; Elnashaie *et al*., 2005). However, auto-thermal operation is associated with complex static and dynamic bifurcation behaviour (Elnashaie & Elshishini, 1996). Deep understanding of this complex phenomenon is essential for efficient design and safe operation/control of this process. In the hydrogen production part of this process, the static bifurcation behaviour and its practical implications are explored for two auto-thermal reforming configurations, as shown in Fig.1a and Fig.1b.

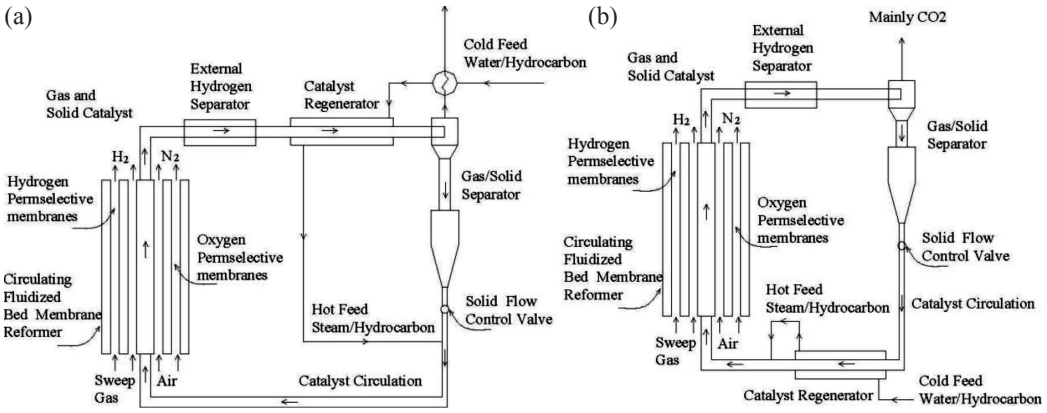


Fig.1: The novel auto-thermal CFBMSR with continuous catalyst regeneration (a) before (b) and after the gas-solid separation

Kari *et al*., (2009) investigated the effect of chaos on the control of the nonlinear chaotic reactor. Pyror *et al*. (2008) used the chaos theory to investigate into the probabilities and possibilities that may have their implications on all types of system.

In the first configuration the catalyst regeneration is carried out before the gas-solid separation (Fig.1a), while in the second configuration it is carried out after the gas-solid separation (Fig.1b). Heptane is used as a model component of heavy hydrocarbons. A number of palladium-based hydrogen permselective membranes and dense perovskite oxygen permselective membranes are used inside the reformer for hydrogen removal and oxygen supply respectively. For the efficient production of hydrogen, the remaining hydrogen exiting from the riser reformer is further removed in an external hydrogen separator. In the first configuration, the carbon deposited on the nickel reforming catalyst is burned off by excess air in the catalyst regenerator as well as the combustion of flammable gases such as un-reacted heptane and the by-products, methane and carbon monoxide. The regenerated catalyst is separated from the gases

using a gas-solid separator and then recycled to the riser reformer. In the second configuration, the solid catalyst is first separated from the gases and then the carbon deposited on the nickel catalyst is burned off in the regenerator. This is followed by separating the regenerated catalyst from the carbon dioxide. Exit hydrocarbons and carbon monoxide are used together with carbon dioxide from the catalyst regenerator as part of the feed to a novel CO_2 dry reformer (Prasad & Elnashaie, 2002). The regenerated catalyst is recycled to the riser reformer.

Both nature and modern industry, as well as its associated trends, generate huge amounts of lignocelluloses that can be useful as raw materials for a wide spectrum of products (Kaylen *et al*., 1999; So & Brown, 1999). The most important among the spectrum of useful products to be produced from this abundant lignocellulose raw material is ethanol, a clean fuel, an important solvent and a source of other important chemicals and fine chemicals. The production of ethanol from lignocellulose can be a difficult task, with the level of difficulty depending on the choice of raw materials. The primary step in the production of ethanol from the lignocellulose is the separation of lignin from lignocellulose, a process now called lignocellulose fractionation that is seeing extensive research and development. Sugars are produced by the hydrolysis of biomass and other cellulosic sources. Various hydrolysis technologies for cellulosic feedstock always produce a mixture of sugars (glucose, xylose, arabinose etc.). These two steps of the process have been the subjects of extensive research (Sun & Cheng, 2002; Mais *et al*., 2002; Iranmahboob *et al*., 2002; Krishnan *et al*., 1997; Ho *et al*., 2000). In this part of the review the focus is on the process of bioconversion i.e. fermentation, with the aim of efficiently fermenting the sugars produced by hydrolysis and obtaining better productivity and yield of ethanol. Ethanol derived from renewable sources such as lignocellulose wastes/materials is an attractive clean fuel to limit air pollution and reduce dependence on fossil fuels (Kralik *et al*., 2001; Newton, 2003). Interest in biofuels has been regained since the Kyoto Protocol which binds industrialised countries that signed this treaty to reduce their carbon dioxide emissions and other greenhouse gas production. Improvements over conventional fermentation processes have been proposed to increase the sugar conversion, ethanol yield and productivity of the process. Continuous ethanol removal using different methods are useful in reducing the ethanol inhibition on the biocatalyst (Jeong *et al*., 1991; O'Brien & Craig, 1996; Ikegami *et al*., 1997; Nomura & Nakao, 2002). Another method of reducing the initial inoculation amount of microorganisms and of increasing the rate of fermentation is to recycle the microorganisms back to the fermentor (Roca & Olsson, 2003; Lang *et al*., 2001; Giridhar & Srivastava, 2001; Giorno *et al*., 2002; Laluce *et al*., 2002). For the fermentor it is important to notice that stability and consistency of bioprocesses has become more important with the development of new biological products and the tighter regulations on product quality. This in turn requires a quantitative knowledge of the bio-culture stability and dynamics to understand, control and optimise the process (Daveyet *et al*., 1996). In fermentation processes, many investigators have reported the presence of sustained oscillations in experimental fermentors and they have developed suitable mathematical relations to model these fermentors (Ghommidh *et al.,* 1989; Daugulis *et al*., 1997; McLellan *et al*., 1999; Jobses *et al.,* 1985; 1986a; Jobses, 1986b).

This review summarises the rich static/dynamic bifurcation behaviour of the ethanol fermentor system (Garhyan *et al*., 2003; Garhyan & Elnashaie, 2004a, 2004b, 2004c, 2004d,

2005; Mahecha-Botero *et al*., 2005). It also shows that these oscillations can be complex, leading to chaotic behaviour, and that these periodic and chaotic attractors can be useful. Using an experimentally verified mathematical model, it is shown that the average conversion of sugar and average yield/productivity of ethanol is sometimes higher for periodic and chaotic attractors than for the corresponding steady states. The results have been confirmed experimentally (Garhyan & Elnashaie, 2005).

In the case of the fluidized bed membrane reactor, bifurcation studies are limited to "static bifurcation" because it is formed of coupled distribute systems, with complicated dynamic characteristics which have not been studied yet, while in the case of the fermentor both "static and dynamic" are studied since the fermentor is described by lumped system, and this has been studied. The dynamics of the fluidized bed membrane reformer is not any less important than that of the fermentor, and a dynamic model should be developed and dynamic experiments carried out.

From the environmental point of view, efficient production of the clean fuels, hydrogen and ethanol represents an important contribution towards achieving a cleaner environment and radically reducing carbon footprints of modern societies. From a sustainable-development point of view these advanced technologies (together with the production of biodiesel) that have adopted the use of renewable raw materials such as biomass represent an important step towards freeing modern societies of over-dependence on non-renewable fossil fuels with all their economical, environmental and political disadvantages.

The third example is dealing with the acetylcholine neurotransmitter in the brain. The chemical synapse is a highly specialised structure that has evolved for exquisitely controlled voltage-dependent secretion. The chemical messengers, stored in vesicles, are released from the pre-synaptic cell following the arrival of an action potential that triggers the vesicular release into the pre-synaptic terminal (Quinn *et al*., 1995). Once released from the vesicles, the transmitter diffuses across a narrow synaptic cleft and binds to specific receptors on the postsynaptic cell. This initiates an action potential event on the nerve-muscle cell membrane triggering muscle contractions (Quinn *et al*., 1995; Llinas, 1999; Paez & Fayad, 1999; Guyton & Hall, 2000; Tucek, 1978). Acetylcholine (ACh*)* plays a well-recognised role in nerve excitation (Quinn, 1995; Llinas, 1999; Guyton and Hall, 2000; Tucek, 1978; Barman, 1969). It is found in cholinergic synapses that provide a stimulatory transmission in the nervous system. Its complete neurocycle constitutes a coupled two-enzyme system. The two-enzyme/ two-compartment model shown in this review paper differs from other membrane excitation models such as the Rose-Hindmarsh model of action potential (Hindmarsh & Rose, 1982; 1984) which is a modification of the Fitzhugh model (Fitzhugh, 1985). This later model was developed to simulate the repetitive, patterned and irregular activity seen in molluscan neurons.

With the above motivations and extending the earlier investigation completed by Elnashaie and coworkers (El-Rifaie *et al*., 1980, Elnashaie *et al*., 1983a; Elnashaie *et al*., 1983b; Elnashaie *et al*., 1984; Elnashaie *et al*., 1995; Ibrahim & Elnashaie, 1997; Ibrahim *et al*., 1995;), a novel diffusion-reaction model is formulated using the non-linear dynamic approach to simulate and study the synaptic gap.

HYDROGEN PRODUCTION

The reactions and kinetics for the riser reformer are summarised in earlier work by many researchers e.g. Elnashaie and Elshishini (1996), Paradeep and Elnashaie (2002) and Chen and Elnashaie (2005). For the first configuration, because the gases exiting the external hydrogen separator contains some flammable gases such as un-reacted heptane, the by-products methane and carbon monoxide, and with the excess air feed in the catalyst regenerator, combustion of these gases take place.

For the second configuration, however, only the burning of carbon takes place. For modeling, only the heat produced in the catalyst regenerator and its thermal effects on the entire system are considered. For the sake of simplicity, we assume plug flow in both the reaction side and permeable membrane tubes. Then the riser reformer is modeled as a plug-flow reactor with no gas-solid slip because of high gas velocity (~3 m/s) and the use of fine catalyst particles (186 μm) (Patience *et al*., 1992). The other main assumptions are as follows:

1. Steady-state operation
2. There is no heat loss for the entire adiabatic process.
3. There is no oxidation of hydrogen or carbon monoxide over the nickel-reforming catalyst, which is based on reported experimental investigation that oxygen can be safely and successfully introduced into a catalytic steam reformer (Roy *et al*., 1999).
4. The palladium based hydrogen membranes and dense perovskite oxygen membranes are 100% selective for the permeation of hydrogen and oxygen, respectively (Shu *et al*., 1994; Barbieri & Di Maio, 1997; Tsai, 1997).
5. The heat capacities of the components and the heat of reactions are constant.
6. The pressures in the riser reformer, hydrogen and oxygen membrane tubes are constant.
7. The performances of hydrogen and oxygen permselective membranes are not affected by carbon deposition.
8. The heat of reactions in the catalyst regenerator are used to preheat the cold feed water to generate steam, the heptane and the regenerated catalyst before recycling to the riser reformer, in which the necessary heat for the endothermic steam reforming reactions is supplied.

The steady state model equations for this CFBMSR form a set of initial value non-linear ordinary differential equations. The reported maximum operating temperatures of current palladium-based hydrogen membranes are in the order of 900 K (Shu *et al*., 1994; Barbieri & Di Maio, 1997). Unless otherwise specified, the results presented are under a set of standard parameters and reaction conditions, e.g. Elnashaie & Elshishini (1996), Paradeep and Elnashaie (2002) and Chen and Elnashaie (2005). It is important to notice that under auto-thermal operation, the feed temperature to the riser reformer is a system variable rather than a feed parameter. Heptane conversion is defined as the total moles of heptane converted per mole of heptane fed. The hydrogen yield is defined as the total moles of hydrogen produced per mole

of heptane fed. Under auto-thermal operation, there is no external heat supply and therefore it is also the same as the net hydrogen yield.

RESULTS AND DISCUSSION

Detailed bifurcation results for the two configurations using different parameters as bifurcation parameters are presented. and they show very clearly the importance of the bifurcation phenomena on design, optimisation and operation of these novel reformers (Chen & Elnashaie, 2005). In this paper we give only a sample of these results for the case where the steam-to-carbon (of heptane) feed ratio is used as the bifurcation parameter.

Bifurcation behaviour of Configuration I of CFBMSR: Catalyst is regenerated before the gas-solid separation

Fig.2 (a-d) shows the bifurcation behaviour when steam to carbon (S/C) feed ratio is used as the bifurcation parameter. The base value of S/C= 2.0 mol/mol. When the S/C feed ratio is lower than 1.44 mol/mol, only one steady state exists in the auto-thermal CFB membrane reformer. In this region low S/C feed ratio is used and thus, the extent of steam reforming (i.e. the fraction of heptane which reacts with steam) is small. On the other hand, for this low S/C, the carbon formation competes for the reactant heptane. The higher the reformer exit carbon flow rate, the higher the amount of carbon deposited on the catalyst and the higher the heat generation in the catalyst regenerator. As a result, the auto-thermal circulating feed temperature to the reformer i.e. the temperature of the stream fed to the reformer is high. For S/C feed ratios below 1.15 mol/mol, the auto-thermal circulating feed temperature increases sharply, as shown in Fig. 2a. Although the bifurcation behaviour discussed in this review is limited to static bifurcation, however, based on the bifurcation theory (Elnashaie & Elshishini, 1996), any decrease in the steam-to-carbon ratio (or increase in the temperature) will lead to thermal "runaway". When the S/C feed ratio is between 1.44 and 2.25 mol/mol, multiplicity with three steady states occurs, which are usually classified into lower, middle and upper steady states according to the feed temperature to the reformer, respectively (Fig. 2a). For example, at the S/C feed ratio of 2.0 mol/mol, three steady states exist with circulating reformer feed temperatures of 688, 702 and 740 K, respectively. Fig.2c shows that at the middle and upper steady states the conversion of heptane is 100%, while at the lower steady state the conversion of heptane is between 81.6 and 95.8%. The conversion of heptane on this lower steady state branch decreases with the increases in the S/C feed ratio. This is the opposite of the usual operation where the reformer performance increases when the S/C feed ratio increases. The reason for this behaviour is that the reformer exit carbon flow rate shown in Fig. 2b is low on this lower steady state branch, making the heat generation from the burning of carbon in the regenerator low. On the other hand, the heat consumption for preheating the cold feed water and heptane increases when the S/C feed ratio increases. In order to balance the heat supply and heat consumption in the reformer-regenerator system, the conversion of heptane under auto-thermal operation has to decrease with the increase in the S/C feed ratio. In the multiplicity region, the higher the reformer exit carbon flow rate (net carbon deposition on the catalyst),

the higher the auto-thermal feed temperature to the reformer. But the trend of net hydrogen yield is totally reversed. As shown in Fig. 2d, the higher the auto-thermal feed temperature to the reformer, the lower the net hydrogen yield. For example, at the S/C feed ratio of 2 mol/mol, the net hydrogen yield is 14.4 at the upper-feed temperature 740 K, while it is 14.6 at the middle-feed temperature of 702 K and 14.66 at the lower-feed temperature of 688K. This reverse relationship between the net hydrogen yield and the auto-thermal circulating feed temperature is due to the carbon formation-burning process in the auto-thermal system. Steam reforming does not only extract hydrogen from hydrocarbons but also extracts hydrogen from steam, while carbon formation only extracts hydrogen from hydrocarbons. Therefore, the larger the fraction of heptane cracking for carbon formation, the more the carbon generation, the higher the auto-thermal circulating feed temperature and the lower the net hydrogen yield. According to the bifurcation theory (Elnashaie & Elshishini, 1996), only the lower and upper steady states are stable, while the middle steady state is an unstable saddle-type state. When the S/C feed ratio is higher than 2.25 mol/mol, the auto-thermal circulating feed temperatures and the reformer exit carbon flow rates are almost constant while the conversion of heptane is decreasing. This is due to the fact that the higher the steam feed, the higher the reaction extents of endothermic steam reforming (of heptane and methane) and the mildly exothermic water gas shift reaction. In order to keep the system operating auto-thermally, the conversion of heptane has to be relatively low in order to provide enough "fuel" to generate heat in the catalyst regenerator to supply the necessary heat for the endothermic steam reforming in the riser reformer. As a result, the auto-thermal circulating feed temperature to the riser reformer is low and the conversion of heptane is low. For example, at the steam-to-carbon feed ratio of 2.5mol/mol, the conversion of heptane is 78.1%, the reformer exit carbon flow rate is 0.9 kg/h and the auto-thermal circulating feed temperature is 694 K. Because the net hydrogen yield is defined as the amount of hydrogen produced per mole of heptane fed, the longer the feed heptane is burned in the regenerator, the smaller the net hydrogen yield (Fig. 2d). For example, at a high S/C feed ratio of 2.5 mol/mol, the net hydrogen yield is 13.8 moles of hydrogen per mole of heptane fed. In this set of results the maximum net hydrogen yield is about 15.6 moles of hydrogen per mole of heptane fed at the lower steady state when the steam-to-carbon feed ratio is close to 1.44 mol/mol, as shown in the left-hand bifurcation point in Fig. 2d. It is about 70.8% of the theoretical maximum hydrogen yield of 22 when the final reforming products are carbon dioxide and hydrogen.

Bifurcation Behavior of Configuration II: Catalyst Regenerated After the Gas-solid Separation

In the second configuration only the burning of carbon deposited on the nickel catalyst supplies the heat for the endothermic reforming reactions and for the vaporisation of cold feed water and heptane. Because steam reforming of heptane is a highly endothermic reaction, the heat supply requirement from the burning of carbon in the catalyst regenerator is high. In order to generate enough carbon on the catalyst as an energy source for an auto-thermal operation, the steam-to-carbon (S/C) feed ratio is much lower than in the usual processes. The standard S/C feed ratio is in mol/mol. In order to keep the same total feed rate to the riser reformer, the base

feed flow rates of heptane and steam for configuration 2 (Fig. 1b) is S/C=1. Similar behaviour to this is observed in Figs. 2 (a-d). In this case the S/C feed ratio is the bifurcation parameter. In this configuration the multiplicity is over a very narrow region from the S/C feed ratio of 0.994 to 1.023 mol/mol. For a practical operation, the control for the accurate S/C feed ratio is very important. In this case the auto-thermal circulating feed temperature to the reformer increases with the increase in the S/C feed ratio on the lower and upper steady state branches, while there is a decrease on the middle steady state branch. However, this case shows that the reformer exit carbon flow rate increases slightly with the increase in the S/C feed ratio on the lower and middle steady state branches, but a decrease on the upper steady state branch. The conversion of heptane is always 100%; the yield of by-product methane decreases on the lower and upper steady state branches but increases on the middle steady state branch with an increase in the S/C feed ratio. The net hydrogen yield in Fig. 3d increases to the multiplicity region from the S/C feed ratio of 0.975 mol/mol. Then it is almost constant when S/C feed ratio is in the range of 1.023 to 1.05 mol/mol. In the multiplicity region the order of reformer exit carbon flow rate or the order of methane yield from high to low is the lower, middle and upper steady states, while the order of net hydrogen yield from high to low is the middle, upper and lower steady states, respectively. For example, at the S/C feed ratio of 1.02 mol/mol, the reformer exit carbon flow rates are 42.1, 41.5 and 39.1 kg/h at the lower, middle and upper steady states, respectively. The methane yields are 0.15, 0.080 and 0.008 moles of methane per mole of heptane fed. The net hydrogen yields are 14.26, 14.41 and 14.35 moles of

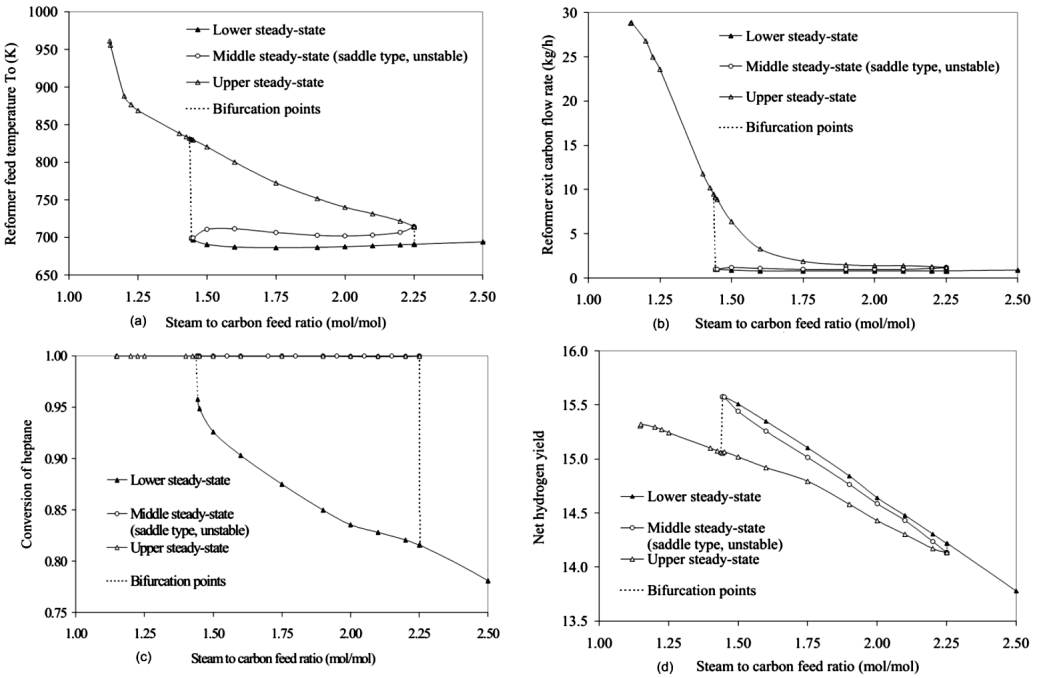


Fig.2: Bifurcation behaviour with steam-to-carbon feed ratio as the bifurcation parameter in the auto-thermic CFBMSR when catalyst is regenerated before gas-solid separation (Fig.1a) and after gas-solid separation (Fig. 1b)

hydrogen per mole of heptane fed at these three steady states. The order of net hydrogen yield is middle>upper>lower steady states, and this yields a curve that bends and crosses itself at the right bifurcation point close to the S/C ratio of 1.023 mol/mol.

BIOETHANOL PRODUCTION

In the preceding case of hydrogen production, the novelty is the configuration itself; bifurcation occurs when the process is thermally intensified to get the highest hydrogen yield and productivity with minimum energy consumption. In the following case of fermentor for the production of bio-ethanol, the process configuration remains simple but novel modes of operation are explored: periodic and chaotic. This approach for the bio-ethanol production using an unstable mode of operation of a CSTR fermentor exploits the bifurcation and chaos theories to maximise ethanol yield and productivity. The research work reviewed in the following sections consists of: developing a reliable and relatively simple model to describe the fermentation process, verifying the model used in bifurcation and chaos analysis work and then using these results to guide experimental investigation of bifurcation and chaos and their implications on improving ethanol yield and productivity. Verification of the model results against experimental results (Garhyan *et al*., 2003; Garhyan & Elnashaie, 2004a, 2004b, 2004c, 2004d, 2005; Mahecha-Botero *et al*., 2005) is essential to validate the model. Bifurcation analysis was carried out for configuration where continuous in-situ ethanol removal using an ethanol selective membrane was reported by Garhyan and co-workers (2004a; 2005). Detailed bifurcation analysis was also reported for more complicated configurations consisting of multiple fermentors with continuous ethanol removal and cell recycle (Garhyan & Elnashaie, 2004c).

MODEL DEVELOPMENT AND DISCUSSION

Biochemical reactors can be viewed as highly complex dynamical systems in which various chemical components are present in the intracellular and extracellular spaces and each cell has unique properties. A rigorous model accounting for the above complexities is called a structured-segregated model (Cazzador *et al*., 1987; Eakman *et al*., 1966; Srienc & Dien, 1992); this model is difficult to formulate and analyse. The presence of oscillations within cellular systems has been observed and mathematically described, such as glycolytic oscillations and oscillations in intracellular calcium concentrations in different types of cells (Goldbeter, 1996). More simplified models can be formulated either by neglecting the intracellular chemical variations (unstructured models) (Bellgardt, 1994; Duboc & von Stockar, 2000; Hjortso & Nielsen, 1994) or by neglecting the heterogeneity of the cell population (un-segregated models) (Cazzador, 1991; Jones & Kompala, 1999; Ramkrishna *et al*., 1987). The simplest models are unstructured-un-segregated models (Daugulis *et al*., 1997). The structured-un-segregated bio-kinetic model describing the continuous fermentation of glucose to ethanol in a continuously stirred fermentor can be described by a set of four differential equations (Garhyan *et al*., 2003). This model has four state variables, namely, the concentrations of sugar (C_S), ethanol (C_P) and microorganism (C_X), and an internal key component (C_e). This cellular component "e " is essential for both growth and product formation and has a non-linear dependence on

ethanol concentration. Hence the inhibition by ethanol does not directly influence the specific growth rate of the culture, but the effect is indirect (Jobses *et al*., 1985, 1986a, b). The term E is the ratio of the internal key component concentration to the microorganism concentration ($E = C_e/C_X$). The model is expressed in the form of four coupled non-linear differential equations (Garhyan *et al*., 2003; Garhyan & Elnashaie, 2004b). For the steady state, the set of four differential equations reduces to a set of four coupled non-linear algebraic equations, which can only be solved simultaneously. Jobses *et al*. (1985; 1986a, b) used successfully this four-dimensional model to simulate the oscillatory behaviour of an experimental continuous fermentor in the high-feed sugar concentration region. This model has been used to explore the different possible complex static/dynamic bifurcation behaviours of this system in the two-dimensional [D (dilution rate = (1/residence time)) -C_{SO} (feed concentration of substrate, sugar)] parameter space, and to review the investigations into the implications of these phenomena on substrate conversion and ethanol yield and productivity. The system parameters for one of the experimental runs of Jobses *et al.* (1986a, b) showing oscillatory behaviour is used as the base set of parameters (Garhyan *et al*., 2003).

PRESENTATION TECHNIQUES AND NUMERICAL TOOLS

The bifurcation diagrams are obtained using the software package AUTO97 (Doedel *et al*., 1997). This package is able to perform both steady state and dynamic bifurcation analysis, including the determination of entire periodic solution branches using the efficient continuation techniques (Kubaiecek & Marek, 1983).The DIVPAG subroutine available with IMSL Libraries for FORTRAN with automatic step size to ensure accuracy for stiff differential equations is usually used for numerical simulation of periodic as well as chaotic attractors. FORTRAN programmes are written for plotting the Poincare plots. MATLAB can also be used instead of FORTRAN. The classical time trace and phase plane for the dynamics are used. However, for high periodicity and chaotic attractors these techniques are not sufficient. Therefore, other presentation techniques are used. These techniques are based upon the plotting of discrete points of intersection (return points) between the trajectories and a hyper-surface (Poincare surface) chosen at a constant value of the state variable (e.g. $C_X = 1.55$ kg/m^3 in fermentation processes, where C_X is the concentration of the microorganism). These discrete points of intersection are taken such that the trajectories intersect the hyper-plane transversally and cross it in the same direction. The return points are used to construct a number of important diagrams namely: Poincare one parameter bifurcation diagram, which is a plot of one of the co-ordinates of the return points (e.g. C_S) versus a bifurcation parameter (e.g. D) and return point histogram, which is a plot of one of the co-ordinates of the return points [e.g. C_S, substrate (sugar concentration)] versus time.

RESULTS AND DISCUSSION

A detailed static/dynamic and chaotic analysis has been carried out for this system and will be reviewed. However, in this review we give only samples of these non-linear dynamics analysis results. Fig.3 is a two-parameter continuation diagram of D vs. C_{SO} showing the loci of Static Limit Points (SLPs) and Hopf Bifurcation (HB**)** points. Two-parameter bifurcation diagrams are

usually constructed by taking a fixed value of all parameters and having thebifurcation diagram between two parameters, e.g. Fig.3 shows the two parameters, D and C_{SO} bifurcation diagrams. The one-parameter bifurcation diagram keeps one of the two parameters constant and takes a state-variable-vs.-the-other-bifurcation parameter. In order to evaluate the performance of the fermentor as an alcohol producer, the following variables are calculated from state variables and feed conditions: conversion of substrate, the product (ethanol) yield and its productivity according to the simple relations incorporated in the computer programme. Substrate (sugar conversion), $X_s = (C_{s0}–C_s)/C_{s0}$; Ethanol Yield = $Y_p = (C_p\text{-}C_{p0})/C_{s0}$ or $\overline{Y}p = (C_p\text{-}C_{p0})/(C_{s0}.X_s)$.

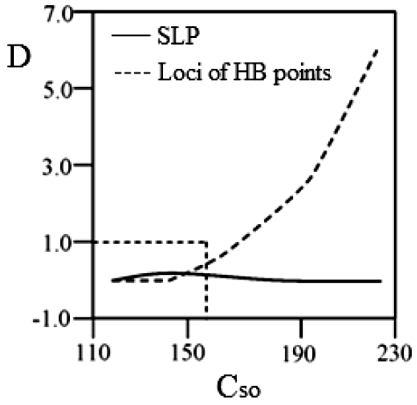


Fig.3: Two parameters continuation diagram of C_{SO} vs. D

Ethanol productivity (production rate per unit volume, kg/m^3.hr) of the fermentor is $P_P = CP.D$. For the oscillatory and chaotic cases, the average conversion $\overline{X}_s$, average yield $\overline{Y}_s$ and the average production rate $\overline{P}_P$ as well as the average ethanol concentration $\overline{C}_P$ $\overline{C}_P$ are computed. They are defined as, $\overline{X}_S = \left[\int_0^\tau X_S.\mathrm{d}t/\tau\right]$, $\overline{Y}_P = \left[\int_0^\tau Y_P.dt/\tau\right]$, $\overline{P}_P = \left[\int_0^\tau P_P.dt/\tau\right]$, $\overline{C}_P = \left[\int_0^\tau C_P.dt/\tau\right]$,where the τ values in the periodic cases represent one period of the oscillations, and in the chaotic cases, they are taken long enough to be reasonable representation of the "average" behaviour of the chaotic attractor.

Sample case, dilution rate D as the bifurcation parameter with C_{SO} = 140 kg/m^3. Jobses *et al*. (1986a, b) use this value of C_{SO} in their experiments together with a dilution rate of D = 0.022 hr^{-1}. Their model results represent the experimental results very well. Details of static and dynamic bifurcation behaviour for this case are shown in Fig.4(A-D), with the dilution rate D as the bifurcation parameter. Fig.4A shows the bifurcation diagram for substrate concentration (C_S) with clear demarcations between the different regions using dotted vertical lines. It is clear that the static bifurcation diagram is an incomplete S-shape hysteresis type with a Static Limit Point (SLP) at very low value of D = 0.0035 hr^{-1}. The dynamic bifurcation shows a Hopf Bifurcation (HB) at D_{HB} = 0.05 hr^{-1} with a periodic branch emanating from it. The region in the neighbourhood of the SLP is enlarged in Fig.6b. The periodic branch emanating from HB terminates homoclinically (with infinite period) when it touches the saddle point very close to the SLP at D_{HT} = 0.0035 hr^{-1}. Fig.4c is the bifurcation diagram for the ethanol concentration *(*C_P). It is clear from Fig. 6c that the average ethanol concentrations for the periodic attractors are

higher than those corresponding to the unstable steady states. Figures 6d show the bifurcation diagrams for ethanol yield (Y_P) where the average yield (Y_P) for the periodic branch are shown as diamond-shaped points and it is clear that the average yield of the periodic attractor is higher than the unsteady state. The same applies for productivity which is not shown here. The period of oscillations increases as the periodic branch approaches the homoclinical bifurcation point; the period tends to infinity indicating homoclinical termination of the periodic attractor at D_{HT} = 0.0035hr^{-1}. The region (Region1) with three point attractors ($D<D_{HT}$) is characterised by the fact that two of them are unstable and only the steady state with the highest conversion is stable. The highest conversion (almost complete conversion) occurs in this region for the upper stable steady state. This steady state also gives the highest ethanol yield which is equal to 0.51 (Fig. 4d). On the other hand, this region has the lowest ethanol production rate due to the low values of the D (a given fermentor active volume corresponds to a very low flow rate). Region 2 of $D_{HB}>D>D_{HT}$ is characterised by a unique periodic attractor (surrounding the unstable steady state), which starts at the HB point and terminates homoclinically at a point very close to SLP as shown in Figs 4a and 4b. It is clear that in this region, the average of the oscillations for the periodic attractor gives (as shown in Figs 4c, d) a higher C_P, Y_P and P_P than that of the corresponding steady states, which means that the operation of the fermentor under periodic conditions is not only more productive but will also give higher ethanol concentrations by achieving a higher sugar conversion. A comparison between the static branch and the average of the periodic branch in this region (e.g. at D = 0.045 hr^{-1}) shows that the percentage improvements are as follows: for C_P it is 9.34%, for X_S it is 9.66%, for Y_P it is 8.67% and for P_P it is 9.84%. Therefore, the best production policy for ethanol concentration, yield and productivity for this case is a periodic attractor.

It is fundamentally and practically important to notice that conversion, yield and productivity are very sensitive to D changes in the neighbourhood of the HB point. This sensitivity is not only qualitative regarding the birth of oscillations for $D<D_{HB}$ but also quantitative comparing the conversion, yield and productivity for $D>D_{HB}$ and their average values for $D<D_{HB}$. The further decrease in D beyond D_{HB} causes the average values of conversion, yield and productivity to increase. Region 3 is characterised by the existence of a unique stable steady state having the conversion, yield and productivity characteristics very close to those of the unstable steady state in region 2. In general, there is a trade-off between concentration and productivity, which requires economic optimisation to determine the optimum D. The phenomenon of possible increase in conversion, yield and productivity through deliberate unsteady state operation has been known for some time (Douglas, 1972). Deliberate unsteady operation is associated with non-autonomous (externally forced) systems. In the present work, the unsteady state operation of the system (periodic operation) is an intrinsic characteristic of the system in certain regions of the parameters. Moreover, this system intrinsically shows not only periodic attractors but also chaotic attractors. Static and dynamic bifurcation and chaotic behaviour are due to the non-linear coupling of the system (Elnashaie & Elshishini, 1996). This non-linear coupling is the cause of all the phenomena including the possibility of higher conversion, yield and productivity. Physically it is associated with the unequal excursion of the dynamic trajectory (periodic or chaotic) above and below the unstable steady state as shown in Fig. 5 (Garhyan *et al*., 2003).

CONCLUSIONS AND RECOMMENDATIONS FOR THE FERMENTATION PROCESS

Extensive non-linear analysis of an ethanol fermentor is investigated and the published results show the complex static and dynamic bifurcation. The mathematical investigations have been used as a guide to carry out experimental work to substantiate the analysis. The investigations reveal rich static and dynamic bifurcation behaviour of the fermentation system, which includes bi-stability, incomplete period doubling cascade, period doubling to banded chaos and homoclinical (infinite period) bifurcation (Keener, 1981) for periodic as well as chaotic attractors. The investigations concentrate on the effect of the different values of the dilution rate and substrate feed concentration on the bifurcation/chaotic behaviour of the system. Special emphasis is given to the implication of these phenomena on the sugar conversion, ethanol yield and productivity of the fermentation process. It is well known from the dynamical system theory that these experimentally observed and mathematically simulated oscillations must start and end at certain critical points (Elnashaie & Elshishini, 1996). Therefore, the models are used to investigate the rich static and dynamic bifurcation behaviour of such experimental fermentors over a wide range of parameters. Bifurcation analysis provides insight into the possible utilisation of periodic attractors to enhance the conversion, yield and productivity of the fermentation process. It is seen in the continuous experiments that the experimental values of the state variables closely match the simulated values, thus confirming that the simplified structured-un-segregated models are suitable for the description of the fermentation process. Different configurations are also explored with in-situ ethanol removal with continuous recycle of the product stream and microorganisms (Garhyan & Elnashaie, 2004a-c). It is clear from the literature that the continuous in-situ removal of ethanol stabilises the oscillations and increases the sugar conversion due to alleviation of product inhibition. Experiments were carried out to show that a change in bifurcation parameter (dilution rate, Dhr^{-1}) results in sustained oscillations. Moreover, when the dilution rate is above the HB value, the oscillations disappear to give a steady-state value. It has been clearly established that periodic and chaotic attractors can give higher average ethanol yield and productivity than the corresponding steady state. Some earlier work in chemical synthesis report the conversion and yield improvements by operating the chemical reactor under unstable oscillating conditions (Douglas & Rippin, 1966; Douglas, 1972).

BIFURCATION AND CHAOTIC BEHAVIOUR OF ACETYLCHOLINESTERASE AND CHOLINE ACETYLTRANSFERASE ENZYMES SYSTEMS: MODELING THE CHEMICAL SYNAPSE TOWARDS AN UNDERSTANDING OF ALZHEIMER AND PARKINSON DISEASES

Although the causes of Alzheimer and Parkinson Diseases are still unknown, extensive multidisciplinary research is still being carried out in order to uncover the causes. Chemical reaction engineering tools can be used to simulate the chemical synapse as an enzymatic diffusion-reaction system. This approach may be useful in obtaining some insight into the dynamic behaviour of neural transmission and contribute to the efforts in treating these diseases.

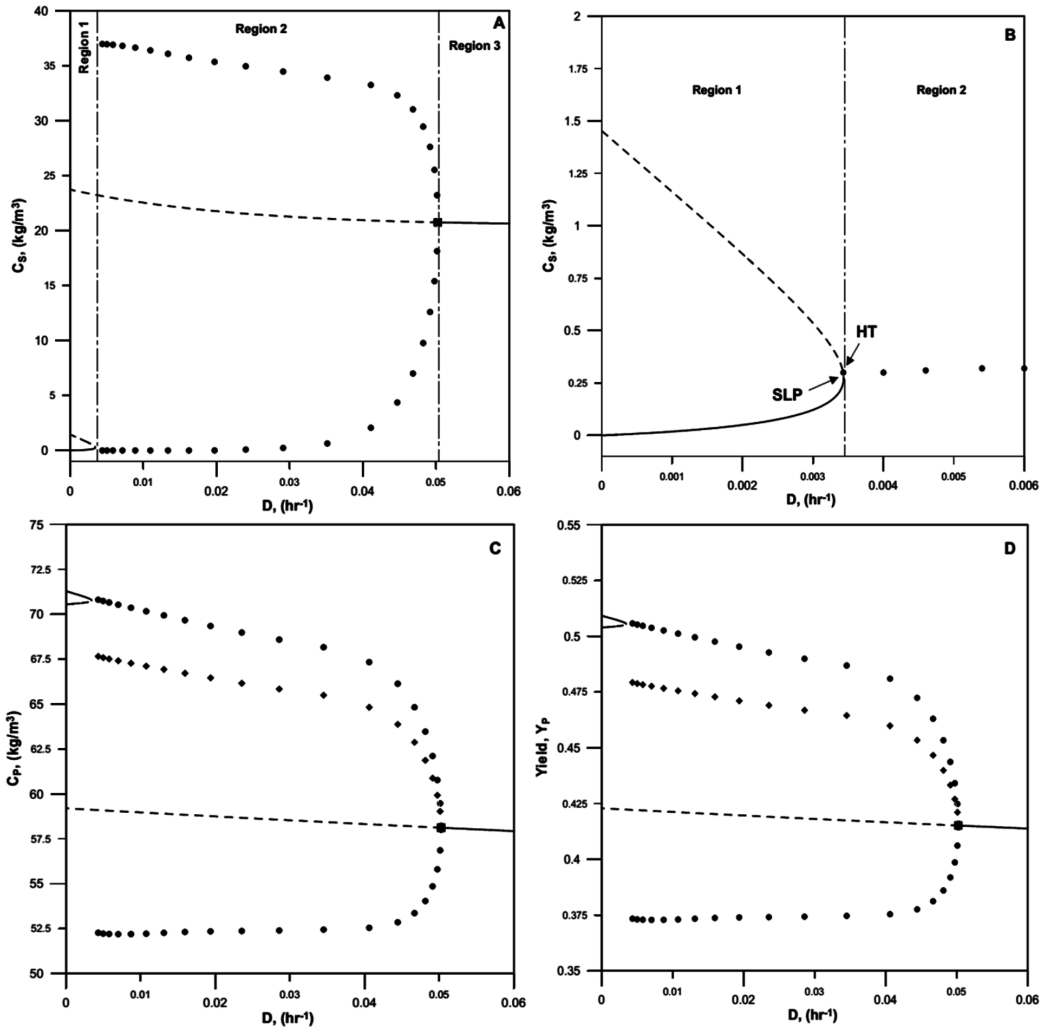


Fig.4: Bifurcation diagrams at C_{SO} = 140 kg/m^3 with D as bifurcation parameter. Steady state branch (_____stable, ----- unstable), Periodic branch (••••• stable, ♦♦♦♦♦ unstable, average of oscillations)

Alzheimer Disease (AD)

AD is a degenerative disease of nerve cells on the cerebral cortex that progressively deteriorates memory and ability to learn, reason, communicate and perform daily activities and, at some point, results in failure to recognise loved ones. The accumulation of protein clusters in the brain is associated with AD, but their role is still unclear. The clumps of protein are composed of beta-amyloid proteins and form what are called neuritic plaques and neurofibrillary tangles, taking place outside and inside neurons respectively. AD is also associated with the poor performance of the neurotransmitter acetylcholine (Braunwald *et al*., 2001; Alzheimer's Association, 2006; Alzheimer's d isease Education and Referral Center, 2006).

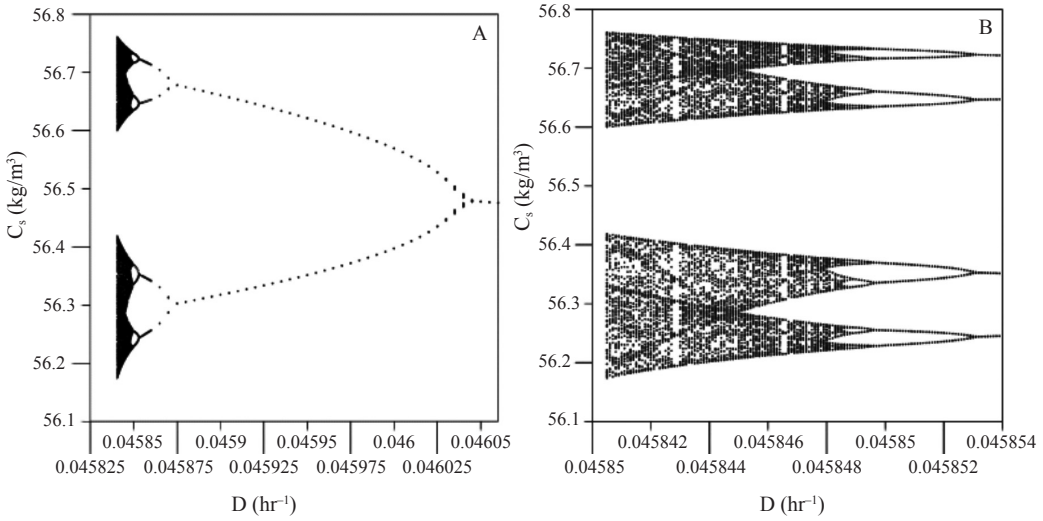


Fig.5. Dynamic characteristics at C_{SO} = 200 kg/m^3 and D = 0.04584 hr^{-1}, A) One dimensional Poincare bifurcation diagram, B) Enlargement of (A)

Parkinson Disease (PD)

PD is a neurodegenerative disorder of the central nervous system. It causes a progressive loss of nerve cell function in the part of the brain that controls muscle movement. As a part of the brain called the substantia nigra becomes impaired, tremors, rigidity, slow movement, stiffness in arms and legs and balance problems, and difficulty walking are seen to manifest; these characteristics are the common symptoms of this disease. PD is also associated with the poor performance of the neurotransmitter dopamine and an imbalance between dopaminergic and cholinergic systems (National Parkinson Foundation, 2006; Parkinson's Information, 2006).

The Chemical Synapse

The chemical synapse is a highly specialised structure that evolved for exquisitely controlled voltage-dependent secretion. Different chemical messengers, stored in vesicles, are released from the presynaptic cell following the arrival of an action potential that triggers the vesicular release into the presynaptic terminal (Llinas, 1999). Once released from the vesicles the transmitter diffuses across a narrow synaptic cleft (Paez & Fayad, 1999) and binds to specific receptors in the postsynaptic cell and initiates an action potential event in the nerve-muscle cell membrane triggering muscle contractions (Llinas, 1999). Acetylcholine plays a recognised role in nerve excitation (Guyton & Hall, 2000). It is found in cholinergic synapses that provide a stimulatory transmission in the nervous system. Its complete neurocycle implies a coupled two-enzyme/two-compartment model with two strongly coupled events as follows:

The activation event. Acetylcholine (S) is synthesised from Choline and Acetyl Coenzyme A (Acetyl-CoA) by the enzyme Choline Acetyl-transferase (ChAT) (Barman, 1969; Tucek, 1978) and is immediately stored in small vesicular compartments closely attached to the cytoplasmic side of the presynaptic membranes.

The degradation event. Once Acetylcholine has completed its activation duty, the synaptic cleft degradation begins removing the remaining Acetylcholine (Soreq & Xakut, 1993). This occurs trough the destruction (hydrolysis) of Acetylcholine by the Acetyl-cholinesterase enzyme (AchE) to form Choline and acetic acid (Barman, 1969). Diseases such as Alzheimer (Alzheimer, 1906) and Parkinson (Parkinson, 1817) are the result of an imbalance of the cholinergic system considered above, with devastating consequences on human health. The above simplified sequence of events suggests that one can obtain some insight into such a cycle using a simple diffusion-reaction model (Naka *et al*., 1997; Naka & Sakamoto, 2001) simulating the non-linear interaction between these events. Other modeling details can be found elsewhere (Mahecha-Botero *et al*., 2004, 2005).

SIMPLIFIED DIFFUSION-REACTION MODELING FOR THE ACETYL-CHOLINE NEUROCYCLE

The complete neurocycle of the Acetylcholine as a neurotransmitter may be simulated as a simplified two-enzyme/two-compartment system. Each compartment may be described as a constant flow, constant volume, isothermal, continuous stirred tank reactor (CSTR). The presynaptic and postsynaptic cells are represented by these two compartments separated by a permeable membrane as shown in Fig.6. Assuming that all the events are homogeneous in all the vesicles, and using the proper dimensionless state variables and parameters, the behaviour is considered for a single synaptic vesicle as described by this simple two-compartment system. Using dynamic mole balances for the chemical species involved in the enzymatic neurocycle of the neurotransmitter obtained is represented by a set of differential equations describing the system. Nonlinear analysis of this set of highly non-linear balance equations gives preliminary insight into the dynamic/static, bifurcation and chaotic behaviour of this complex biological system. From an enzyme kinetics point of view, the most general case is usually considered, where both enzymes have non-monotonic dependence upon substrate and hydrogen ions concentrations. It is usually assumed that Acetylcholine is synthesised in the presynaptic cell by the enzyme Choline Acetyl-transferase due to an activation reaction (where the stimulatory neurotransmitter Acetylcholine is synthesised) as follows:

$$\textbf{R(1): Choline + Acetyl-CoA} \xrightarrow{\textbf{CHOLINE-ACETYLTRANSFERASE}} \textbf{Acetylcholine + CoA}$$

Acetylcholine is destroyed (hydrolysed) in the postsynaptic cell by the enzyme Acetyl cholinesterase through a degradation reaction (where the stimulatory neurotransmitter Acetylcholine is degraded) as follows:

$$\underset{\textbf{(A)}}{\textbf{R(2): Acetylcholine}} + \underset{\textbf{(B)}}{\textbf{Water}} \xrightarrow{\textbf{Acetyl-cholinesterase}} \textbf{Choline + Acetate + H+}$$

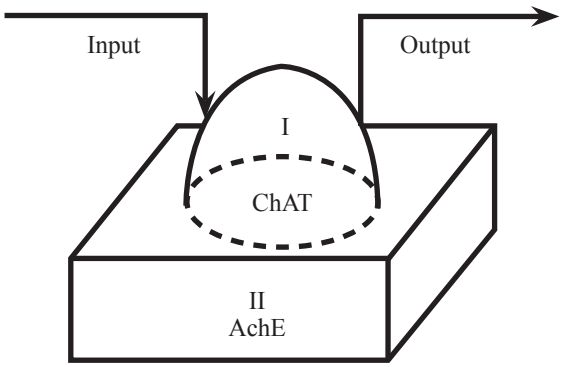


Fig.6: Two enzyme/two-compartment model

Using the approach described above, static and dynamic simulations are usually carried out for different parameter values, obtaining many static and dynamic characteristics of the system. Some of results are shown in Figs 7-9. Fig.7 shows static bifurcation behaviour with the Acetyl-cholinesterase enzyme activity as the bifurcation parameter. S is the dimensionless concentration of the substrate Acetylcholine and SB_1 and SB_2 are the two static bifurcation points. It is clear that the behaviour is strongly dominated by hysteresis and multiplicity phenomena. This means that if for any reason (any external disturbance as an inhibitor, denaturalisation of the enzyme or a cholinergic disease) the Acetyl-cholinesterase enzyme activity is affected, then the 'normal' concentrations of all components could change dramatically. For example, if we observe the transmitted Acetylcholine at the postsynaptic cell and assume that the Acetyl-cholinesterase enzyme activity increases, the neurotransmitter should show a change in its concentration from $2.4*10^{-7}$ (kmol/m3) to $1.2*\ 10^{-5}$ (kmol/m3). This means that its concentration increases by a factor bigger than 50. It is possible that some phenomena like the one described here could be the cause of Alzheimer Disease.

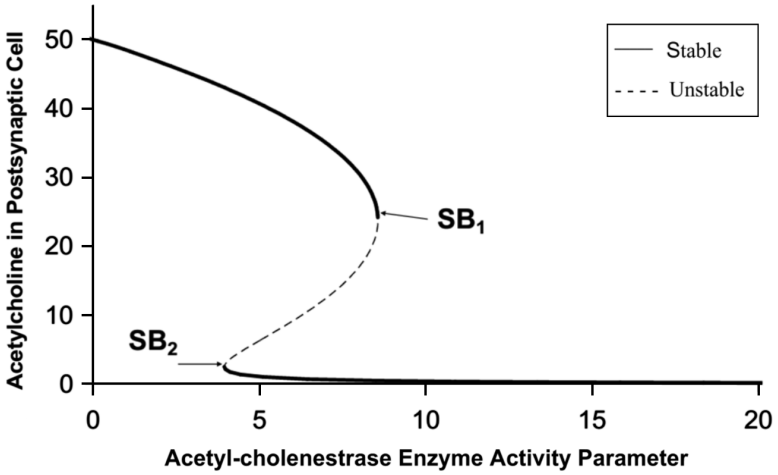


Fig.7: Bifurcation diagram as the Acetyl-cholinesterase enzyme activity is varied. Effect on the neuro-transmitter in the postsynaptic cell. Steady state branches. SB: Static Bifurcation.

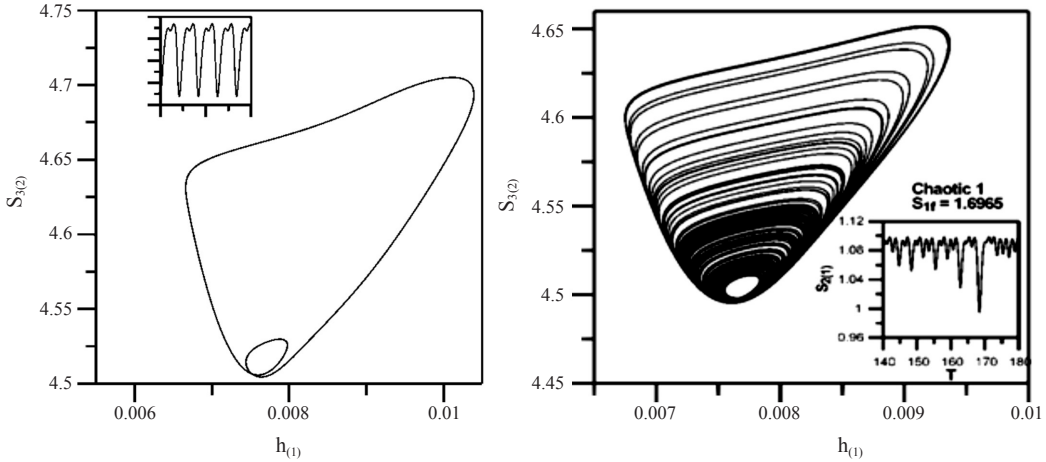


Fig.8. Dynamic simulations show region of periodic/chaotic behaviour. (A) Periodic, and (B) Chaotic

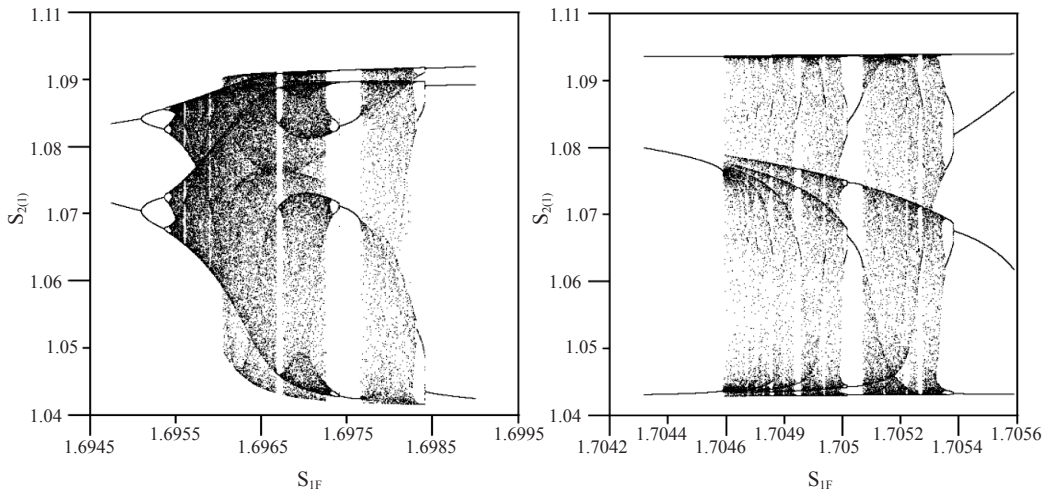


Fig.9: Poincare maps show period doubling to chaos and enlargement of region of chaotic behaviour.

Fig.8 shows the strong sensitivity of the behaviour to a very small change in $\mathbf{S}_{1f}$ (dimensionless acetylcholine concentration in the feed), changing the behaviour from periodic to chaotic. Simulation results of the concentrations of chemical species in both the presynaptic and postsynaptic cells show a period doubling mechanism to chaos (Fig.8, Fig.9); S_1 is dimensionless acetylcholine concentration; S_2 is dimensionless choline concentration and S_3 is dimensionless acetate concentration; h is the dimensionless hydrogen concentration. Between brackets 1 is compartment 1 and between brackets 2 is compartment 2. Subscript f is the feed. Oscillating values from periodic and chaotic attractors of chemical species concentrations fairly agree with measured data from rat and guinea pig data found in the literature. It is interesting to notice that some state variables correspond to the physiological expected values in some regions. This may open the door for future experimental research in order to use the model for simulating real physiological behaviour.

CONCLUSION: THE NEUROCYCLE

A review of the modeling of an acetylcholine neurocycle is presented. Bifurcation and chaotic dynamics dominate the behaviour of the system. Extensive multidisciplinary research in neurosciences is required to develop a deeper understanding of the human brain. Furthermore, chemical reaction engineering modeling techniques may be useful to unveil the behaviour of complex systems such as the chemical synapses. The type of work presented in this review leads towards a better understanding of the complex behaviour of the human brain. This will help to direct new research on neurodegenerative diseases such as Alzheimer and Parkinson.

NUMERICAL TECHNIQUES

Almost all bifurcation and chaos phenomena are associated with non-linear systems, two point boundary value differential equations and non-monotonic behaviour of the non-linear functions. Therefore, the investigation of these phenomena strongly depends upon the use of numerical techniques to solve these non-linear equations (Elnashaie & Elshishi, 1996; Elnashaie & Garhyan, 2003; Elnashaie & Uhlig, 2007). Because bifurcation and chaos do not only involve solving the non-linear equations but also present it in the form of bifurcation and chaos diagrams, special techniques are needed to do them (Elnashaie & Elshishini, 1996; Elanashaie & Uhlig, 2007). Such models describing realistic chemical engineering systems exhibit as explained above very complex bifurcation and chaos behaviour due to the strong coupling between the transport processes and the nonlinear dependence of the reaction rates on temperature and concentration. During the past 50 years, they have proved to be an inexhaustible source for the development and testing of various local bifurcation and chaos techniques. Chemical reacting systems are the richest among these systems, unlike the Navier-Stokes equations, which are partial differential equations in time and two/three spatial coordinates, models of chemical reactors and reacting flows vary from a pair of ordinary differential equations describing the behaviour of a continuous flow stirred tank reactor with a single exothermic reaction to several PDEs, which describe combustion problems in which the fluid flow is strongly coupled to heat and mass diffusion and complex chemistry. Another distinguishing feature of reactor models is that the number of dimensionless parameters that appear is usually large, typically four to 15 for a single reaction. Hence, a comprehensive numerical study of the behaviour of a given model is impractical without some theoretical guidance. Balakotaiah and Khinast (2000) present an excellent such guide of numerical techniques for bifurcation analysis. Other numerical techniques for bifurcation and chaos analysis are presented by excellent authors in books and papers (e.g. Wiley, 2000; Nikolay *et al*., 2002; Eugene, 2003)

GENERAL CONCLUSIONS: BIFURCATION AND CHAOS IN CHEMICAL AND BIOLOGICAL SYSTEMS

Static/dynamic bifurcation and chaos are fundamentally very important phenomena that every chemical and biological engineering engineer should know about as an essential part of their training (Elnashaie & Garhyan, 2003). These phenomena are widespread in many chemical and biological engineering systems and have important practical implications on the behaviour of those systems. Understanding and analysing these phenomena is essential

for the rational design, optimisation, operation and control of these chemical and biological engineering processes. These phenomena can sometimes be dangerous and harmful; in such cases engineers and researchers should have the techniques to avoid them (e.g. the relatively modern field of Chaos Control), in some other cases they can be beneficial and it should be possible to exploit them (e.g. the frementor presented in this review). In other cases they are neither harmful nor beneficial but they need to be analysed and utilised to understand certain complex behaviour especially in biological systems. The source of these phenomena is mainly non-linearity and synergetic non-linear coupling. However non-linearity is not sufficient; it seems that at least one of the processes dependent upon at least one of the state variable needs to be no-monotonic. Coupling between reaction and diffusion seems also to be a necessary condition for the occurrence of these phenomena. For example non-monotonic processes when taking place in homogeneous plug flow reactors do not show any of these phenomena; however, as soon as diffusion comes into the picture, through axial dispersion in a homogeneous system and/or solid-gas interaction in a heterogeneous system these phenomena start to become a part of the characteristics of the system. Another important point is what may be called feedback of information, like counter-current operation, recycle or continuous circulation (e.g. the novel CFBMSR review), this can also give rise to bifurcation and chaos. These simple intuitive arguments apply mainly to chemical and biological reaction engineering systems. For other systems like fluid flow systems the source of bifurcation and chaos can be quite different. It is well established now that the transition from laminar flow to turbulent flow is actually a transition from non-chaotic to chaotic behaviour. The synergetic interaction between hydrodynamically induced bifurcation and chaos and that resulting from chemical and biological reaction/diffusion is not well understood and calls for extensive multidisciplinary research efforts.

ACKNOWLEDGEMENTS

The author is very thankful to Engineer Mohammad Rezazadeh Mehrjou from the Electrical Engineering Department, UPM for his excellent help with the editing of this review paper.

REFERENCES

Adris, A. M., Elnashaie, S. S. E. H., & Hughes, R. (1991). A Fluidized Bed Membrane Reactor for the Steam Reforming of Methane, *The Canadian Journal of Chemical Engineering, 69*(10), 1061.

Adris, A., Grace, J., Lim, C., & Elnashaie, S. S. E. H. (1994a). *Fluidized Bed Reaction System for Steam/Hydrocarbon Gas Reforming to Produce Hydrogen*. US Patent no 5,326,550.

Adris, A. M., Lim, C. J., & Grace, J. R. (1994b). The Fluidized Bed Membrane Reactor (FBMR) System: a Pilot Scale Experimental Study. *Chemical Engineering Science, 49*, 5833-5843.

Adris, A. M., Lim, C. J., & Grace, J. R. (1997). The fluidized-Bed Membrane Reactor for Steam Methane Reforming: Model Verification and Parametric Study. *Chemical Engineering Science, 52*(10), 1609-1616, (1997).

Adris, A., Grace, J., Lim, C., & Elnashaie, S. S. E. H. (2002). *Fluidized Bed Reaction System for Steam/Hydrocarbon Gas Reforming to Produce Hydrogen*. Canadian Patent no. 2,081,170, Date: December 24, 2002.

Alzheimer, A. (1906). *Uber einen eigenartigen schweren Erkrankungsprozeb der Hirnrinde. Neurologisches Central blatt* (Lecture presented at the 37th meeting of the southwest German psychiatrists in Tubingen), 23, pp.1129-1134.

Alzheimer's Association. (2006). *Alzheimer's Disease*. Retrieved from http://www.alz.org/.

Alzheimer's Disease Education and Referral Center (2006). *Alzheimer's Information*. Retrieved from http://www.nia.nih.gov/Alzheimers/.(2006).

Aris, R., & Amundson, N. R. (1958). An analysis of chemical reactor stability and control. Parts I-III. *Chemical Engineering Science, 7*(3), 121-155.

Bailey, J. E., & Ollis, D. R. (1977). *Biochemical Engineering Fundamentals*. McGraw Hill, New York.

Bailey, J. E. (1998). Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Chemical Engineering Communications, 13*(1-3), 111-132.

Balakotaiah, V., & Khinast, J. (2000). Numerical Bifurcation Techniques for Chemical Reactor Problems", Numerical Methods for Bifurcation Problems and Large-Scale Dynamical Systems. *The IMA Volumes in Mathematics and its Applications, 119*, 1-36.

Balakotaiah, V., & Luss, D. (1981). Analysis of multiplicity patterns of a CSTR. *Chemical Engineering Communications, 13*(1-3), 111-132.

Balakotaiah, V., & Luss, D. (1983). Multiplicity criteria for multiple-reaction networks. *AIChE Journal, 29*(4), 552-560.

Barbieri, G., & Francesco, P. Di Maio. (1997). Simulation of Methane Steam Reforming Process in a catalytic Pd-Membrane Reactor. *Ind. Eng. Chem. Res., 36*, 2121-2127.

Barman, T. E. (1969). *Enzyme Handbook-Volumes 1&2*. Springer-Verlag: New York.

Bellgardt, K.-H. (1994). Analysis of synchronous growth of baker's yeast: Part I: Development of a theoretical model for sustained oscillations. *J. Biotechnol., 35*, 19.

Braunwald, E., Fauci, A. S., Kasper, D. L., Hauser, S. L., Longo, D. L., & L. Jameson, J., (Eds.). (2001). *Alzheimer's disease and other primary dementias. Harrison's principles of internal medicine*. New York, McGraw-Hill (2001).

Cazzador, L. (1991). Analysis of oscillations in yeast continuous cultures by a new simplified model. *Bull. Math. Biol., 53*,685.

Cazzador, L. Alberghina, L. Martegani, E. Mariani, L. (1987). Bioreactor control and modeling: A simulation program based on a structured population model of budding yeast. In Moody, G. W., Baker, P. B. (Eds.) *Bioreactors and Biotransformations.* Elsevier: London, p.64.

Chen, Z., & Elnashaie, S. S. E. H. (2002). *Efficient Production of Hydrogen from Higher Hydrocarbons using Novel Membrane Reformer*. Proceeding of the 14th World Hydrogen Energy Conference, Montreal, Canada.

Chen, Z., & Elnashaie, S.S.E.H. (2005). Bifurcation and Its Implications for a Novel Auto-thermal Circulating Fluidized Bed Membrane Reformer for the Efficient Pure Hydrogen Production, *Chemical Engineering Science, 60*(15),p4287-4309.

Chen, Z., Yan, Y., & Elnashaie, S. S. (2003). Modeling and optimization of a novel membrane reformer for higher hydrocarbons. *AIChE journal, 49*(5), 1250-1265.

Chen, Z., Yan, Y., & Elnashaie, S. S. (2003). Novel circulating fast fluidized-bed membrane reformer for efficient production of hydrogen from steam reforming of methane. *Chemical Engineering Science, 58*(19), 4335-4349.

Chen, Z., & Elnashaie, S.S.E.H. (2005). Bifurcation and Its Implications for a Novel Auto-thermal Circulating Fluidized Bed Membrane Reformer for the Efficient Pure Hydrogen Production. *Chemical Engineering Science , 60*(15), 4287-4309.

Choi, K. Y., & Harmon Ray, W. (1985). The dynamic behaviour of fluidized bed reactors for solid catalysed gas phase olefin polymerization. *Chemical Engineering Science, 40*(12), 2261-2279.

Christensen, T. S. (1996). Adiabatic Pre-reforming of Hydrocarbons - an Important Step in Syngas Production. *Applied Catalysis A: General, 138*, 285-309.

Danforth, C. M. (2013, April). *Chaos in an Atmosphere Hanging on a Wall*. Mathematics of Planet Earth.

Daugulis, A. J., McLellan, P. J., & Li, J. (1997). Experimental investigation and modeling of oscillatory behavior in the continuous culture of Zymomonas mobilis. *Biotechnology and bioengineering, 56*(1), 99-105.

Davey, H. M., Davey, C. L., Woodward, A. M., Edmonds, A. N., Lee, A. W., & Kell, D. B. (1996). Oscillatory, stochastic and chaotic growth rate fluctuations in permittistatically controlled yeast cultures. *Biosystems, 39*(1), 43-61.

Doedel, E. J., Champneys, A. R., Fairgrieve, T. F., Kuznetsov, Y. A., Sandstede, B., & Wang, X. J. (1997) *AUT097: Continuation and bifurcation software for ordinary differential equations*. Department of Computer Science, Concordia University, Montreal, Canada.

Douglas, J. M. (1972). *Process Dynamics and Control, Volume 2, Control System Synthesis*. Prentice Hall, New Jersey, USA.

Douglas, J. M., & Rippin, D. W. T. (1966). Unsteady state process operation. *Chemical Engineering Science, 21*(4), 305-315.

Duboc, P., & von Stockar, U. (2000). Modeling of oscillating cultivations of *Saccharomyces cerevisiae*: Identification of population structure and expansion kinetics based on on-line measurements. *Chemical Engineering Science, 55*(1), 149-160.

Eakman, J. M., Fredrickson, A. G., & Tsuchiya, H. H. (1966). Statistics and dynamics of microbial cell populations. *Chemical Engineering Progress. Symp.Ser., 62*, 37.

Earnshaw, B. A., & Keener, J. P. (2010). Global asymptotic stability of solutions of nonautonomous master equations. *SIAM Journal on Applied Dynamical Systems, 9*(1), 220-237.

Elnashaie, S. S. E. H. (1977). Multiplicity of the steady in the fluidized bed reactors—III: Yield of the consecutive reaction A→ B→ C. *Chemical Engineering Science, 32*(3), 295-301.

Elnashaie, S. S. E. H., & Adris, A. (1989). *Fluidized Bed Steam Reformer for Methane*. Proceedings of the IV International Fluidization Conference, Banff, Canada, May.

Elnashaie, S. S. E. H., & Elshishini, S.S. (1996). *Dynamic Modeling, Bifurcation and Chaotic Behavior of Gas-Solid Catalytic Reactors*. Gordon and Breach Science Publishers, London, UK.

Elnashaie, S. S. E. H., & El-Bialy, S. (1980). Multiplicity of the steady state in fluidized bed reactors. V- the effect of catalyst decay. *Chemical Engineering Science, 35*(6), 1357-1368.

Elnashaie, S. S. E. H., & Yates, J. G. (1973). Multiplicity of the steady state in fluidized bed reactors. I- steady state considerations. *Chemical Engineering Science, 28*(7), 515-520.

Elnashaie, S. S. E. H. (2005). *Efficient Production and Economics of the Clean Fuel Hydrogen*. Keynote Lecture, First International Green Energy Conference, Waterloo , Ontario, Canada, 11-16 June, 2005.

Elnashaie, S. S., Abashar, M. E., & Al-Ubaid, A. S. (1988). Simulation and optimization of an industrial ammonia reactor. *Industrial & engineering chemistry research, 27*(11), 2015-2022.

Elnashaie, S. S. E. H., & Adris, A. (1989). *Fluidized Bed Steam Reformer for Methane*. Proceedings of the IV International Fluidization Conference, Banff, Canada, May, 1989.

Elnashaie, S. S. E. H., Elshishini, S. S. (1993). *Modelling, simulation and optimization of industrial fixed bed catalytic reactors*. Gordon and Breach Science Publishers: London, UK.

Elnashaie, S. S. E. H., & Garhyan, P. (2003). *Conservation Equations and Modeling of Chemical and Biochemical Processes*. Marcel Dekker, NY, USA.

Elnashaie, S. S. E. H., Prasad, P., & Chen, Z. (2005). Static Bifurcation Characteristics of an Auto-thermal Circulating-Fluidized Bed Hydrogen Generator for Fuel Cells. *Industrial & Engineering Chemistry Research, 44*(14), 4871-4883.

Elnashaie, S. S., Affane, C., & Uhlig, F. (2007). *Numerical Techniques for Chemical and Biological Engineers Using MATLAB®: A Simple Bifurcation Approach*. Springer.

El-Rifai, M. A., Elnashaie, S. S. E. H., Aboulfath, H. (1980). Effect of feedback inhibition on the behavior of multienzyme mono-linear reaction chains. *J. Solid-Phase Biochem., 5*, 235-244.

Epstein, I. R., & Pojman, J. A. (1998). *An introduction to Nonlinear Chemical Dynamics*. Oxford University Press, New York..

Eugene, L. A., & Kurt, G. (2003). Introduction to Numerical Continuation Methods. *SIAM Classics in Applied Mathematics*, 45.

Feigenbaum, M. J. (1980). Universal behaviour in nonlinear systems. *Los Alamos Science, 1*, 4-36.

Fitzhugh, R. "Impulses and physiological states in theoretical models of nerve membranes " Biophys. J.,1, 445-462(1985)

Garhyan, P., Elnashaie, S. S. E. H., Al-Haddad, S. M., Ibrahim, G., & Elshishini, S. S. (2003). Exploration and exploitation of bifurcation/chaotic behavior of a continuous fermentor for the production of ethanol. *Chemical engineering science, 58*(8), 1479-1496.

Garhyan, P., & Elnashaie, S. S. E. H. (2004a). Utilization of mathematical models to investigate the bifurcation and chaotic behavior of ethanol fermentors. *Mathematical and computer modelling, 39*(4), 381-427.

Garhyan, P., & Elnashaie, S. S. E. H. (2004b). Static/dynamic bifurcation and chaotic behavior of an ethanol fermentor. *Industrial & engineering chemistry research, 43*(5), 1260-1273.

Garhyan, P., & Elnashaie, S.S.E.H. (2004c). Bifurcation analysis of two continuous membrane fermentor configurations for producing ethanol. *Chemical Engineering Science, 59*, 3235-3268.

Garhyan, P., & Elnashaie, S. S. E. H. (2004d, October). Bio-ethanol Production-Solving the Efficiency Bottleneck, *The Chemical Engineer, 760*, 30-32.

Garhyan, P., & Elnashaie, S. S. E. H. (2005). Experimental investigation and confirmation of static/dynamic bifurcation behavior in a continuous ethanol fermentor. Practical relevance of bifurcation and the contribution of Harmon Ray. *Industrial Engineering and Chemistry Research, 44*, 2525-2531.

Ghommidh, C., Vaija, J., Bolarinwa, S., & Navarro, J. M. (1989). Oscillatory behavior of Zymomonas mobilis in continuous cultures: a simple stochastic model. *Biotechnology Letters, 2*(9), 659-664.

Giorno, L., Chojnacka, K., Donato, L., Drioli, E. (2002). Study of a Cell-Recycle Membrane Fermentor for the Production of Lactic Acid by Lactobacillus bulgaricus. *Industrial & Engineering Chemistry Research, 41*(3), 433-440.

Giridhar, R., Srivastava, A. K. (2001). L-sorbose production in a continuous fermentor with and without cell recycle. *World Journal of Microbiology & Biotechnology 17*(2), 185-189.

Goldbeter, A. (1996). *Biochemical 0scillations and Cellular Rhythms. The Molecular Bases of Periodic and Chaotic Behavior*. Cambridge University Press: Cambridge, U.K.

Goltsov, V. A., & Veziroglu, T. N. (2002). A step on the road to hydrogen civilization. International *Journal of Hydrogen Energy, 27*(7), 719-723..

Golubitsky, M., & Schaeffer, D. G. (1985). *Singularities and bifurcation theory, Vol. I. Applied Mathematical Science, Vol. V.* Springer, Berlin.

Govaerts, W. J. (2000). *Numerical methods for bifurcations of dynamical equilibria* (Vol. 66). Siam.

Gray, P., & Scott, S. K. (1994). *Chemical Oscillations and Instabilities*. Clarendon Press, Oxford.

Guyton, C. A. Hall, J. E. (2000). *Textbook of Medical Physiology*, 10th Ed..W. B. Saunders Company: Amsterdam.

Hindmarsh, J. L., & Rose, R. M. (1984). A model of neuronal bursting using three coupled first order differential equations. *Proc. Roy. Soc., B221*(1222), 87-94.

Hindmarsh, J. L., & Rose, R. M. (1982). A model of the nerve impulse using two first-order differential equations. *Nature, 296*(5853), 162-169.

Hjortso, M. A. Nielsen, J. A. (1994). Conceptual model of autonomous oscillations in microbial cultures. *Chem. Eng. Sci., 49*, 1083.

Ho, N. W. Y., Chen, Z., Brainard, A. P., Sedlak, M. (2000). Genetically engineered Saccharomyces yeasts for conversion of cellulosic biomass to environmentally friendly transportation fuel ethanol. *ACS Symposium Series, 767*, 143.

Ibrahim, G., & Elnashaie, S. S. E. H. (1997). Hyper-chaos in acetyl-cholinesterase enzyme systems. *Chaos, Solitons & Fractals, 8*, 1977-1988.

Ibrahim, G., Teymour, F. A., & Elnashaie, S. S. E. H. (1995). Periodic and chaotic behavior of substrate-inhibited enzymatic reactions with hydrogen ions production. *Appl. Biochem. Biotechnol., 55*, 175-190, (1995)

Ikegami, T., Yanagishita, H., Kitamoto, D., Karaya, K., Nakane, T., Matsuda, H., Koura, N., & Sano, T. (1997). Production of Highly Concentrated Ethanol in a Coupled Fermentation/Pervaporation Process using Silicate Membranes. *Biotechnology Techniques 11*(12), 921-924.

Ipsen, M., & Schreiber, I. (2000). Dynamical properties of chemical systems near Hopf bifurcation points. *Chaos, 10*(4),791-802.

Iranmahboob, J., Nadim, F., Monemi, S. (2002). Optimizing acid hydrolysis: A critical step for production of ethanol from-mixed wood chips. *Biomass and Bioenergy, 22*(5), 401.

Jeong, Y. S., Vieth, W. R., Matsuura, T. (1991). Transport and Kinetics in Sandwiched Membrane Bioreactors. *Biotechnology Progress, 7*, 130-139.

Jobses, I. M. L., Egberts, G. T. C., Ballen, A. V., & Roels, J. A. (1985). Mathematical modeling of growth and substrate conversion of Zymomonas mobilis at 30 and 35°C. *Biotechnology & Bioengineering, 27*(7), 984-995.

Jobses, I. M. L., Egberts, G. T. C., Luyben, K. C. A. M., & Roels, J. A. (1986a). Fermentation kinetics of Zymomonas mobilis at high ethanol concentrations: oscillations in continuous cultures. *Biotechnology & Bioengineering, 28*(6), 868-877.

Jobses, I. M. L. (1986b). *Modeling of anaerobic microbial fermentations: the production of alcohols by Zymomonas mobilis and Clostridium beijerincki*. (PhD Thesis). Delft University, Delft, Holland.

Jones, K. D., & Kompala, D. S. (1999). Cybernetic modeling of the growth dynamics of Saccharomyces cerevisiae in batch and continuous cultures. *J. Biotechnol., 71*, 105.

Juárez, F. (2011). Applying the theory of chaos and a complex model of health to establish relations among financial indicators. *Procedia Computer Science*, *3*, 982–986.

Karri, R. R., Damaraju, P. R., & Venkateswarlu, C. (2009). Soft sensor based nonlinear control of a chaotic reactor. *Intelligent Control Systems and Signal Processing, 2*, 537-543.

Kaylen, M., Van Dyne, D. L., Choi, Y. S., Blase, M. (1999). Economics , feasibility of producing ethanol from lignocellulosic feedstock. *Bio-resources Technology, 72*(1), 19.

Keener, J. P. (1981). Infinite period bifurcation and global bifurcation branches. *Journal of Applied Mathematics, 41,* 127-144.

Kralik, M., Macho, V., Brautbar, N., Vachalkova, A., & Mikulec, J. (2001). Engine fuels in the 21st century. *Petroleum and Coal, 43*(2), 72-79.

Krishnan, M. S., Xia, Y., Ho, N. W. Y., & Tsao, G. T. (1997). Fuel ethanol production from lignocellulosic sugars. Studies using a genetically engineered Saccharomyces yeast. *ACS Symposium Series, 666*, 74.

Kubaiecek, M., & Marek, M. (1983). *Computational methods in bifurcation theory and dissipative structures*. Springer Verlag, New York.

Laluce, C., Souza, C. S., Abud, C. L., Gattas, E. A. L., & Walker, G. M. (2002). Continuous ethanol production in a nonconventional five-stage system operating with yeast cell recycling at elevated temperatures. *Journal of Industrial Microbiology & Biotechnology, 29*(3), 140-144.

Lang, X., Macdonald, D. G., & Hill, G. A. (2001). Recycle bioreactor for bio-ethanol production from wheat starch II. Fermentation and economics. *Energy Sources, 23*(5), 427-436.

Li, D., Cheng, Y., Wang, L., Wang, H., Wang, L., & Zhou, H. (2011). Prediction method for risks of coal and gas outbursts based on spatial chaos theory using gas desorption index of drill cuttings. *Mining Science and Technology (China), 21*(3), 439-443.

Liljernoth, F. G. (1919). Starting and stability phenomenon of ammonia oxidation and similar reactions. *Chemical Metallurgical Engineering, 19*, 287-291.

Llinas, R. R. (1999). *The Squid Giant Synapse: A Model for Chemical Transmission*. Oxford University Press: London.

Mahecha-Botero, A., Garhyan, P., & Elnashaie, S. S. E. H. (2004). Bifurcation and chaotic behaviour of a coupled acetyl-cholinesterase/choline acetyl-transferase diffusion-reaction enzymes system. *Chemical Engineering Science, 59*, 581-597.

Mahecha-Botero, A., Garhyan, P. & Elnashaie, S. S. E. H. (2005). Modeling, bifurcation and chaotic behaviour of a coupled acetyl-cholinesterase/choline acetyl-transferase enzymes neurocycle. *Mathematical and Computer Modelling, 41*(6/7), 655-678.

Mahecha-Botero, A., Garhyan, P., & Elnashaie, S. S. E. H. (2005). Bifurcation, Stabilization and Ethanol Productivity Enhancement in a Membrane Fermentor. *Mathematical and Computer Modelling, 41*(4-5), 391-406.

Mais, U., Esteghlalian, A. R., Saddler, J. N., & Mansfield, S. D. (2002). Enhancing the enzymatic hydrolysis of cellulose materials using simultaneous ball milling. *Applied Biochemistry and Biotechnology*, 98-100, 815 (Biotechnology for Fuels and Chemicals).

McLellan, P. J., Daugulis, A. J., & Li, J. (1999). The incidence of oscillatory behavior in the continuous fermentation of Zymomonas mobiles. *Biotechnology Progress, 15*(4), 667-680.

Melo, P. A., Sampaio, J. G., Biscaisa Jr., E. C., & Pinto, J. C. (2001). Periodic oscillations in continuous free-radical solution polymerization reactors-a general approach. *Chemical Engineering Science, 56*(11), 3469-3482.

Naka, T., & Sakamoto, N. (2001). Two-dimensional compartment model for generation of miniature endplate current at the neuromuscular junction. *Comments on Theoretical Biology, 6*, 1-28.

Naka, T., Shiba, K., & Sakamoto, N. (1997). A two-dimensional compartment model for the reaction-diffusion system of acetylcholine in the synaptic cleft at the neuromuscular junction. *Bio-Systems, 41*, 17-27.

National Parkinson Foundation. (2006). *About Parkinson's Disease*. Retrieved from http://www.parkinson.org.

Newton, R. (2003). Energy: bio-fuels are the future. *Chemistry & Industry (London), 11*, 14-15.

Nomura, M., Bin, T., & Nakao, S. I. (2002). Selective ethanol extraction from fermentation broth using a silicalite membrane. Separation and purification technology, 27(1), 59-66.

Nikolay, S., Boris, L., Aleksandr, S., & Michail, F. (2002). *Lyapunov-Schmidt Methods in Nonlinear Analysis and Applications*. Kluwer Academic Publishers.

O'Brien, D. J., & Craig Jr., J. C. (1996). Ethanol production in a continuous fermentation/membrane pervaporation system. *Applied Microbiology and Biotechnology, 44*(6), 699-704.

Ohi J. (2002). *Hydrogen Energy Futures: Scenario Planning by the U.S. DOE Hydrogen Technical Advisory Panel*. 14thWorld Hydrogen Energy Conference, Montreal, Canada, June 9-13, 2002

Paez, M. C., & Fayad, R. (1999). Diffusion of the neurotransmitter cannot govern the rise time of miniature end-plate. *Revista Colombiana de Fisica, 31*, 163-168.

Parkinson, J. (1817). *Essay on Shaking Palsy.* London.

Parkinson's Information. (2006). *About Parkinson's Disease*. Retrieved from http://www.parkinsonsinfo.com/

Patience, G. S., Chaouki, J., Berruti, F., & Wong, R. (1992). Scaling considerations for circulating fluidized bed risers. *Powder Technology, 72*(1), 31-37.

Prasad, P., & Elnashaie, S.S.E.H. (2002). Novel Circulating Fluidized Bed Membrane Reformer for the Efficient Production of Ultra clean Fuels from Hydrocarbons. *Ind. Eng. Chem. Res., 41*, 6518-6527.

Pryor, R. G., Amundson, N. E., & Bright, J. E. (2008). Probabilities and possibilities: The strategic counseling implications of the chaos theory of careers. *The Career Development Quarterly, 56*(4), 309-318.

Quinn, D. M., Balasubremanian, A. S., Doctor, B. P., & Taylor, P. (1995). *Enzymes of the Cholinesterase Family*. Plenum Press: London.

Ramkrishna, D., Kompala, D. S., & Tsao, G. T. (1987). Are microbes optimal strategists ? *Biotechnology Progress, 3*, 121.

Ray, W. H. (1977). Bifurcation phenomena in chemically reacting systems. In Rabinowiz, P. H. (Ed.) *Applications of Bifurcation Theory*. Academic Press, New York, 285-315, (1977).

Razon, L. F., & Schmitz, R. A. (1986). Intrinsically Unstable Behavior During the Oxidation of Carbon Monoxide on Platinum. *Catal. Rev., Sci.Eng., 28*(1), 89-164.

Roca, C., Olsson, L. (2003). Increasing ethanol productivity during xylose fermentation by cell recycling of recombinant Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology, 60*, 560-563.

Rostrup-Nielsen, J. (1977). Hydrogen via Steam Reforming of Naphtha. *Chemical Engineering Progress, 9*, 87.

Shu, J., Grandjean, B., & Kaliaguine, S. (1994). Methane steam reforming in asymmetric Pd-and Pd-Ag/porous SS membrane reactors. *Applied Catalysis A: General, 119*(2), 305-325.

So, K. S., & Brown, R. C. (1999, January). Economic analysis of selected lignocellulose-to-ethanol conversion technologies. In *Twentieth Symposium on Biotechnology for Fuels and Chemicals* (pp. 633-640). Humana Press.

Soreq, H., & Xakut, H. (1993). *Human Cholinesterases and Anticholinestrases*. Academic Press: New York (1993)

Srienc, F., & Dien, B. S. (1992). *Kinetics of the cell cycle of Saccharomyces cerevisiae*. Ann. N.Y. Acad. Sci. p.59.

Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: A review. *Bio-resource Technology, 83*(1), 1.

Teymour, F., & Ray,W. H. (1992). The Dynamic Behavior of Continuous Polymerization Reactors. IV. Complex Dynamics in Full-scale Reactors. *Chemical Engineering , 47*(15-16),4133-4170.

Tøttrup, P. B. (1982). Evaluation of intrinsic steam reforming kinetic parameters from rate measurements on full particle size. *Applied Catalysis, 4*(4), 377-389.

Tsai, C. Y., Dixon, A. G., Moser, W. R., & Ma, Y. H. (1997). Dense perovskite membrane reactors for partial oxidation of methane to syngas. *AIChE Journal, 43*(S11), 2741-2750.

Tucek, S. (1987). *Acetylcholine Synthesis in Neurons*. John Wiley, New York.

Twigg M. V. (1989). *Catalyst Handbook, 2nd Edition*. Wolfe Publishing Ltd, London, England, pp.225-282.

Uppal, A., Ray, W. H., & Poore, A. B. (1974). On the dynamic behavior of continuous stirred tank reactor. *Chemical Engineering Science, 29*(4), 967-985.

Werndl, C. (2009). What are the new implications of chaos for unpredictability?. *The British Journal for the Philosophy of Science, 60*(1), 195-220.

Pertanika J. Sci. & Technol. 22 (2): 365 - 385 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

Flat Plate Solar Collectors and Applications: A Review

Bande, Y. M.* and Mariah, N. A.

Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

In this study, various methods and applications of flat plate solar collectors are discussed and pictorial representations are presented. Low temperature applications of flat plate collectors are identified in solar cooking, solar water heating, space and air heating, industrial heating plants and in agricultural produce drying processes. Basic equations, as presented by many researchers in the performances of flat plate collectors, are also presented. The review discusses the analysis of losses from flat plate collectors towards obtaining the overall heat loss coefficient which indicate the performance of flat plate collectors.

Keywords: Flat plate collectors, air heating, water heating, heating and cooling of buildings, solar drying

INTRODUCTION

Flat plate collectors are designed to collect solar radiation at high frequencies. This radiation falls on its black painted surface and converts the radiation into heat energy. They are placed in relation to the latitude and longitude of a location to obtain the best output from them. Flat plate collectors are described as the most important in solar energy applications that involve heating or cooling, irrespective of the end-use application. Sparrow *et al.* (1977) concluded that the performance of the flat plate collector is dependent on the losses from the bottom, top and sides of the collector. Therefore, to calculate heat loss from the collector to its surrounding is crucial for the design performance of the solar collectors as presented by Akhtar and Mullick (1999). Hence, the effectiveness of the collector is the difference between the incident radiation and the amount of energy loss from the collector from any of the exits since the heat loss can be evaluated with the useful energy, if the overall heat loss coefficient is known.

Article history:
Received: 8 December 2011
Accepted: 29 February 2012

Email addresses:
Bande, Y. M. (ymbande@yahoo.com),
Mariah, N. A. (mariah@upm.edu.my)
*Corresponding Author

ISSN: 0128-7702

In analyzing of performance of these collectors, the overall heat loss coefficient is assumed constant per location and its configuration. This explains why conventional collectors, where heat transfer coefficients are considered constant, the flat plate collectors are considered linear (Cooper *et al.,* 1981). Thus, flat plate collectors find applications most commonly in low temperature applications such as water heating, solar cooking, agricultural produce drying, solar distillation, desalination processes, domestic ovens, industrial process heating and even space heating or cooling and heating of buildings.

For high temperature applications, solar flat plate collectors are "modified" by enhancing "concentrated" solar radiation beams, and using additional reflectors directed towards the collectors to tremendously increase the temperature generation of collectors. This has found applications in solar reflectors and solar concentric furnaces. The applications of flat plate collectors are therefore many and the application of which is directly related to end use. Goyal and Tiwari (1999) experimented on the reverse flat plate absorber for the drying of agricultural produce. In their findings, the reverse flat plate absorber offered a better and more uniform drying since the produce was not exposed to direct contact with the sun.

Several authors have also reviewed works on flat plate collectors. However, no published work has been done on the applications of plate collectors so far. Thus, the present review is specifically on flat plate solar collectors and their applications. In particular, the review focuses on heating, cooling, drying, furnace, pumping, thermal power generation, cooking and solar pond applications.

FLAT PLATE COLLECTORS

The performance of solar energy flat plate collector is subject to the amount of heat energy losses from its surface as proposed by Francey and Papaionnou (1985). Computation of these losses is usually cumbersome. Hottel and Woertz (1942) suggested an empirical expression for calculating Ut, the top loss coefficient. This has undergone several modifications by researchers' worldwide. The useful energy possibly obtainable from a solar collector is the difference between the solar energy incident on the collector and the amount of heat lost from the collector. Garg *et al.* (1983) concluded that the overall heat-loss coefficient, U_l, is a function of many parameters such as material properties of the collector, ambient conditions, position of the collector and its configurations, wind speed and direction, and absorber temperature.

Calculation to obtain heat losses from collectors to the surrounding area is of vital importance for the design or simulation of their performance. It is evaluated by considering convection and radiation losses from the absorber plate in the upward direction (Soddah *et al*., 1982).

Two methods of calculating U_t are generally identified as:

- Approximate method
- Numerical methods

Approximate solution has however been accepted much by researchers in solar thermal systems. Hottel and Woertz (1942) first proposed the empirical equation estimating U_t of the flat plate collectors. Later, the empirical equation that gives good correlation with emmittance at absorber plate was modified by Klein (1975). Solar flat collectors are designed to gain useful heat energy from the incident solar insolation. Several types of solar flat collectors can be developed, and these may include flat-plate or concentrating types. The applications of the generated heat energy suggest the type to be used, whereby the concentrating type found to generate more heat energy than the flat plate type. In crop drying, for example, a simple flat-plate solar collector can generate just enough heat energy to attain equilibrium moisture content for drying.

A simple flat plate solar energy collector consists of an absorbing surface, painted usually in black to absorb insolation and transmit it to the working fluid (Ekechukwu & Norton, 1999, p. III). Meanwhile, Akhtar and Mullick (2007) concluded that solar energy collectors are special kinds of heat exchangers that transform solar radiation energy into internal energy of the transport medium, usually water or air. The solar energy collected is carried from the circulating fluid, either directly to the water or to the space conditioning equipment or to a thermal energy storage tank, which can be drawn for use at night or on cloudy days (Soteris, 2009). Generally, the flat-plate with more than one cover is used when high temperatures are required and a single cover is used when low temperatures are required (Akhtar & Mullick, 1999). Samdarshi and Mullick (1983, cited in Akhtar & Mullick, 1999) indicated that it was the best estimation of U_t, top loss coefficient in the literature, while Akhtar and Mullick (1999) compared their finding with that of Garg *et al.* (1983), where the percentage of error was found to be minimal. Akhtar and Mullick (1999) and Garg *et al.* (1983) concluded that the error in their findings was much lower than that of the previous findings, as a result of empirical relations use instead of numerical solutions.

Ho *et al.* (2009) investigated the effect of air recycling in double pass flat plate collectors with internal fins as heat sinks. An experimental investigation of three different types of solar air heaters adopting flat plates, two with fins attached, and others without fin was carried out by Alta et al. (2010). In their experiment, one of the heaters with fins had single glass cover, while the other with double. Based on the results, they concluded that the heater with double cover and fin was more efficient in terms of temperature difference of inlet and outlet air. Several other researchers concluded that flat plate collectors like V-grove, fin-air collectors, and chevron pattern absorbers as less efficient than the others. In their findings, El-sawi *et al.* (2010) concluded that putting the three patterns under the same conditions, the V-corrugated pattern was found to be more efficient than the others, with flat plate collector having the lowest performance.

The performances of flat plate collectors are affected by the climatic conditions of the test area or location. Sunshine time and intensity determine the efficiency of the plates. Other conditions such as dust, dirt, shadow, persistent rainfall are among the factors militating against their performances. The performances of the solar flat plate collectors are subject to location. In tropical regions where average daily sunshine can be up to 12 hours per day, better efficiencies are recorded compared to rain forests.

Types of Solar Energy Flat Plate Collectors

The solar energy flat plate collectors are basically divided into two broad classes, as follows:

- Bare-plate solar energy collectors
- Covered plate solar collectors

Bare-plate Solar Energy Collectors

These are the simplest of all solar collectors. The top-most surface is the absorber plate, with the rear insulated and an air duct, as presented by Ekechukwu and Norton (1999, p. III). Consider the following figure:

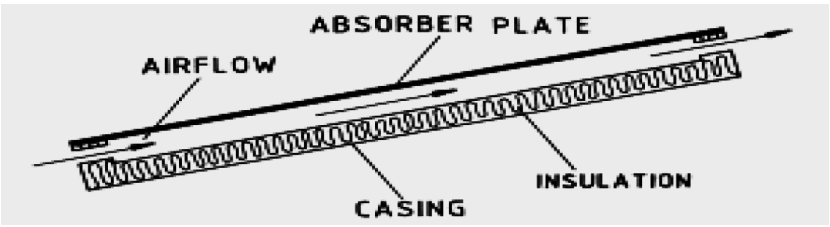


Fig.1: Bare plate solar collector (Ekechukwu & Norton, 1999, p. III)

They are commonly used in agricultural food drying. Optical loss which is caused by transmission reduction of incoming solar radiation, as a result of passing through the transparent cover, is about 10% of the insolation (Soddah *et al.,* 1982).

Covered Plate Solar Collector

Using one or more transparent cover materials minimizes the upward heat loss from the collectors (Ekechukwu & Norton, 1999, p. III). Glass, clear plastics and plexi-glass are identified as the most common materials used but glass material has gained more acceptance. The cover material, in addition to being transparent to allow direct falling of solar insolation to the plate, also prevents convective heat loss and plate against cooling especially during rain, and reduces long-wave radioactive heat losses. This is an advantage to the covered plate as it operates at higher efficiency than the bear plate (Ekechukwu & Norton, 1999, p. III). Ekechukwu and Norton (1999, p. III) describe four different types of covered plate solar heaters. The following schematic diagram is self-explanatory.

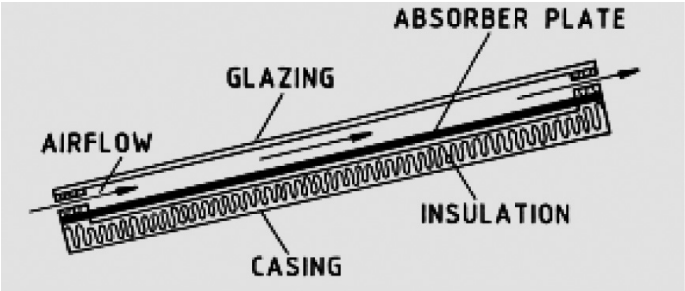


Fig.2: Front pass solar collector

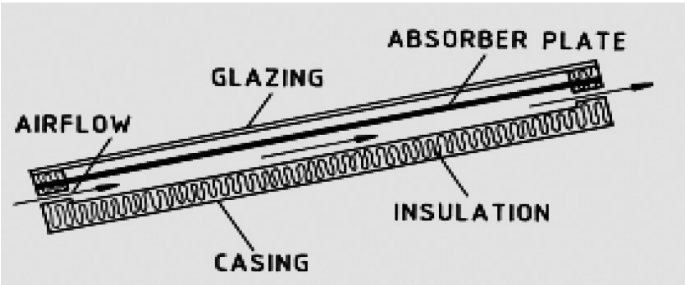


Fig.3: Back pass solar collector

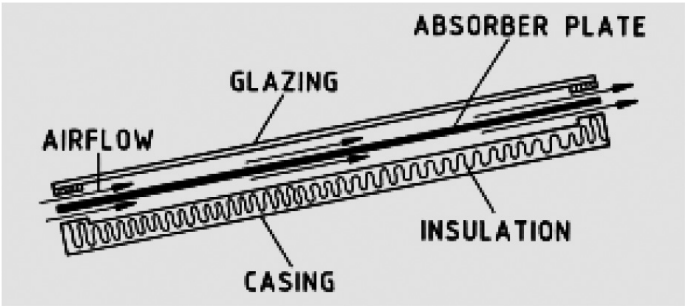


Fig.4: Parallel pass suspended solar plate collector (Ekechukwu & Norton, 1999, p. III)

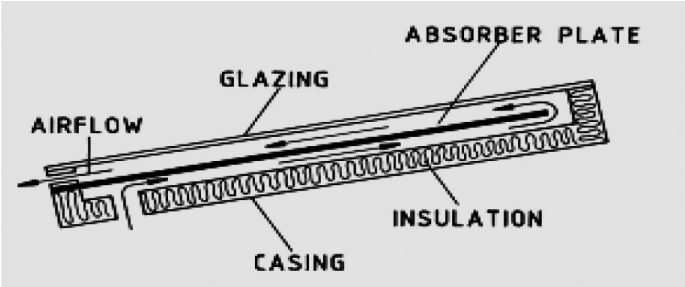


Fig.5: Double pass suspended solar collector

Solar Collectors in Cooling of Buildings

Many researchers have agreed that for solar cooling, the following guides are adoptable and acceptable (Rai, 1984; Tabor, 1966; Baker *et al.*, 1976):

1. Optimal collector area is greater than the area needed for heating only.
2. Two glass covers are optimum for sub-tropical climate; otherwise, three glass covers are desirable due to higher temperature requirement for the cooling system.
3. Optimal collector tilt, (O_T) is equal to latitude (φ), (except in south, where $O_T = \varphi - 10^0$)

Cold storage offers certain advantages over hot storage. The temperature difference ΔT between cold storage and the temperature of a building is less than the temperature difference ΔT between hot storage and building temperature. Thus, less cooling effect is lost from cold storage than from hot storage. This is adopted either to provide cooling for food preservation or for conform cooling. These can be done through refrigeration or evaporation processes.

Solar Collector in Heating of Buildings

The amount of heat to be supplied by solar conventional heater must equal the sum of heat loss through walls, the amount of heat required to warm ventilating air, and air entering through infiltration. If Q cubic meter per hour of air is introduced by ventilation and infiltration, the heat required to bring this air to room temperature is (Sodha *et al*., 1985; Malik *et al.,* 1982):

$$q_v = Q \cdot C_p \rho (T_{in} - T_{out}) \quad \text{Kcal/hr} \tag{1}$$

C_p = Specific heat of air Kcal/Kg^{o}C
ρ = Density of air Kg/m^3
T_{in} = Room temperature (^{o}C)
T_{out} = Outside temperature (^{o}C)

Heat losses through structures on hourly basis are usually based on steady heat conditions. The rate of heat loss through a wall, a window or a ceiling in a steady state is given by:

$$q_{wall} = UA(T_{in} - T_{out}) \tag{2}$$

As U is an overall heat transfer coefficient, the total heat loss on a house is:

$$q_{total} = q_v + q_{wall} \tag{3}$$

The important parameters to be considered in the design of solar collector are collector tilt, number of covers and area of the collectors. Studies have shown that the optimum collector tilt to deliver each unit of heating at a minimum cost is approximately latitude + 15^o for hot services water heating required throughout the year. The optimum tilt is equal to the latitude. The most favourable orientation of the collector is facing south. The optimum number of glass covers varies with climate.

Performance of Collectors in Solar Air-heaters

The performance of solar air heater is given by the overall efficiency. Derivation of performance equations based on heat transfer model is presented in this subsection. Useful energy gain can be given as follows (Rai, 1984):

$$Q_u = \dot{m} c_p (T_o - T_i) \tag{4}$$

Under steady-state conditions, heat balance equation thus;

$$Q_u = AF_R I(\tau\alpha)_e - AF_R U_L (T_i - T_a) \tag{5}$$

A combination of factor, effective transmissivity, absorptivity product (τα) defines the ratio of solar radiation absorbed by the absorber plate to the incident solar flux (Rai, 1984; Malik *et al.,* 1982).

$$(\tau\alpha)_e = \tau\alpha + a[F_1 + F_2 + F_3 + \ldots + F_n] \tag{6}$$

Where, τα is the fraction transmitted by the cover plates and absorbed by absorbing plate, while F_1 , F_2 and F_3 are the fractions absorbed by each glass cover.

Instantaneous efficiency of the collector can be obtained from the following (Rai, 1984; Malik *et al.,* 1982):

$$\eta = F_R\left[(\tau\alpha) - U_L \frac{(T_i - T_a)}{I}\right] \tag{7}$$

$$\eta = \frac{Q_u}{IA} \tag{8}$$

The efficiency, η, is plotted against $(T_i - T_a)/I$. A single generalized curve represents the collector performance. Thus, various cover arrangements and/or selective coatings could be compared, as shown in Fig.6 below (Ekechukwu & Norton, 1999, p. III).

It is generally assumed that U_L is constant and the plot above is linear, as confirmed by Ekechukwu and Norton (1999, p. III), Rai (1984) and Tawari *et al*. (1991). This is true for small values of $(T_i - T_a)$, when the radiation loss is a small fraction of the total heat loss. At high values of $(T_i - T_a)$, the radiation loss predominates over convection and conduction losses as it is proportional to T^4_1. U_L has a positive but minimal value at $T_1 = T_a$. It increases with the increase in the value of T_1 or plate temperature and has a maximum value at the equilibrium plate temperature [9].

As for the test procedures, the value of Q_U is calculated from equation 1, but solar insulation, I, ambient and inlet temperature of fluid T_a and T_i are measured under assumed steady state conditions (Tawari *et al*., 1991). Duffie and Beckman (1980) concluded that long-term experiences on the performance of solar energy collectors showed that intercepts and slopes are good enough to characterize them.

Works have been carried out by many researchers world over on single and double pass solar air heaters. Double pass solar air heaters have a single glass cover but a double air-way

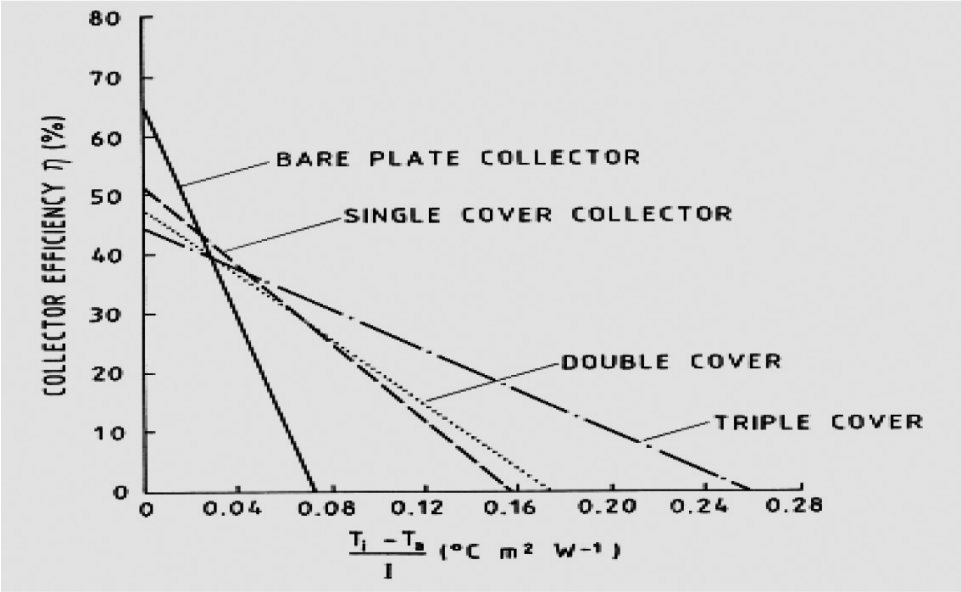


Fig.6: Performance curve for solar collectors

through which air passes at inlet at low temperature and at exit at high temperature. Air initially passes between the glass and absorber plate and then between the absorber plate and the bottom cover. This process has been found to enhance heat generation (see Ong, 1995; Hegazy, 2000; Naphon & Kongtragol, 2003; Aboul *et al.,* 2000; Hegazy, 1999). Other researchers like Choudhury *et al.* (1995) and Yeh *et al.* (2000) found that the increase in velocity of air increased the heat transfer coefficient, resulting in improved performance. Similarly, Sebaii (2011) also stated that to increase fluid velocity, recycling of air would produce the effect of remixing the inlet fluid with the outgoing fluid.

It is vital for all solar thermal systems to be able to store maximum energy for a working period so as to conserve energy in the flat plate collector box since most energy losses are through glass or plastic compared to walls, if good insulation is made to the walls and bottom of the collector (Abhishek *et al.*, 2011).

The performance of solar air heaters is given by overall efficiency as in the case of a flat-plate collector. The derivation of performance equations of the most popular solar air heater design, the flat-plate with air flow on one side based on the heat transfer model is presented by Rai (1984). Effective transmissivity determination for a given transparent cover combination is identical to that of a liquid heater. Effective transmissivity-absorptivity product (τα), which defines the ratio of solar radiation absorbed by the absorber plate to the incident solar flux $(\tau\alpha)_e$, can be calculated from:

$$(\tau\alpha) = \tau\alpha + a(A_1 + A_2 + A_3 + \ldots) \tag{9}$$

where τα is the fraction transmitted by cover plates and absorbed by the absorbing plate and A_1, A_2 the fraction absorbed by each glass cover.

The flux on the absorber can be computed by the term HR $(\tau\alpha)_e$ and heat losses by U_L (T_p-T_a). Heat balance on the absorber plate:

$$HR(\tau\alpha)_e(D\delta x) = U_L(D\delta x)(T_p + T_a) + h_c(D\delta x)(T_p - T) + h_r\varepsilon(D\delta x)(T_p + T_b) \tag{10}$$

D is the width, T_b is the bottom temperature
But ($U_L \sim U_{UP}$)
Heat balance on the air stream (control volume) is given by:

$$\dot{m}C_p\left(\frac{dT}{\delta X}\right)\delta x = h_c(D\delta x)(T_p - T) + h_c^I(B\delta x)(T_b - T) \tag{11}$$

where $\dot{m}$ is the mass flow rate.
Heat balance on the rear plate is obtained as:

$$h_r\varepsilon(D\delta x)(T_p + T_b) = h_c(D\delta x)(T_b - T) + U_bD\delta x(T_b - T_a) \tag{12}$$

The solution to these simultaneous equations yields a differential equation which defines the temperature variation of the air stream.

$$\dot{m}\frac{C_p}{D}\left(\frac{dT}{dx}\right) = \left[\frac{1}{1 + \frac{U_L}{h}}\right]\left[HR(\tau\alpha)_e - U_L(T - T_\alpha)\right] \tag{13}$$

$$h = h_C + \frac{1}{\frac{1}{h_c^I} + \frac{1}{\varepsilon h_r}} \qquad \text{as} \quad (U_b << U_L) \tag{14}$$

Letting $F^I = \frac{1}{1 + \frac{U_L}{h}}$ and solving the differential equation, the air temperature, T, at a distance x from the inlet end of the collector is found to be:

$$T = \frac{HR(\tau\alpha)_e}{U_L} - (T_1 - T_\alpha) + \frac{HR(\tau\alpha)_e}{U_L}\exp\left(\frac{-DF^IU_Lx}{\dot{m}C_p}\right) + T_a \tag{15}$$

T_1 is the air inlet temperature:

$$T = \frac{S}{U_L} - T_1 + T_a + S\frac{S}{U_L}\exp\left(\frac{-DF^IU_Lx}{\dot{m}C_p}\right) + T_a \tag{16}$$

The air temperature at the collector exists at a distance L is:

$$T_2 = T_1 + \left\{\frac{S}{U_L} - (T_1 - T_a)\left[1 - Exp\left(-F^I\frac{U_L}{GC_p}\right)\right]\right\} \tag{17}$$

G is the mass flow rate per unit area.

The overall heat transfer coefficient between the air stream and outside air through the transparent cover U_O is given by:

$$U_O = F^I U_L = \left(\frac{1}{\frac{1}{U_L} + \frac{1}{h}} \right) \tag{18}$$

Under ideal conditions, i.e. where the heat transfer from the absorber plate to the air stream is perfect;

$$h = \infty,\ F^I = 1, \text{ and } U_O = U_L \tag{19}$$

Substituting $F^I = U_o / U_L$, the heat picked up by the air stream is found to be:

$$\frac{Q_u}{A_c} = GC_V(T_2 - T_1) = \left(\frac{1}{1 + \frac{U_L}{h}} \right) \left[\frac{1 - Exp\left(-U_o / GC_p \right)}{U_o / GC_p} \right] x\left[S - U_L(T_1 - T_a) \right] \tag{20}$$

In terms of flow factors, F^{II}, the equation is reduced to

$$\frac{Q_U}{A_C} = F^I F^{II}\left[S - U_L(T_1 - T_o) \right] \tag{21}$$

$$q = F_R\left[S - U_L(T_1 - T_o) \right] \tag{22}$$

Thus, $F^I = \dfrac{F_R}{F^{II}}$

From equation 22 above, we see that efficiency is increased by increasing F' and F", $(\tau\alpha)_e$ and reducing U_L. U_L and $(\tau\alpha)_e$ are related to the characteristics of the upper surface of the absorber plate and the transparent cover system. F' and F" can be increased by increasing the mass flow rate per unit area G and in the convective heat transfer coefficient between the air stream and the absorber plate, as well as increasing surface to transfer heat from the absorber plate to the air stream using fins.

The threshold level is defined by

$$(H_C R) = U_L \frac{(T_1 - T_a)}{\tau\alpha_e} \tag{23}$$

(H_cR) denotes the critical radiation intensity.
The usual heat collected is therefore expressed with a linear relation of

$$\frac{Q_u}{A_c} = C(HR) - K(T_1 - T_a), \qquad HR \geq P(H_c R) \tag{24}$$

$$C = F'F''(\tau\alpha)_e$$

$$K = F'F''(U_L)$$

The solar air heaters efficiency must be defined as:

$$\eta = \frac{Q_u}{A_c(HR)} = \frac{q_u}{(HR)} \tag{25}$$

Performance of Collectors in Water Heating

Many different designs of collectors have been experimented on and are basically on insulated box with one or more transparent covers, containing a black painted metal absorber plate and some arrangements for circulating water pass the plate. The flat-plate collectors that are most commonly used are shown schematically below (Rai, 1984; Sheffer, 1994; Rosenbaum, 1991).

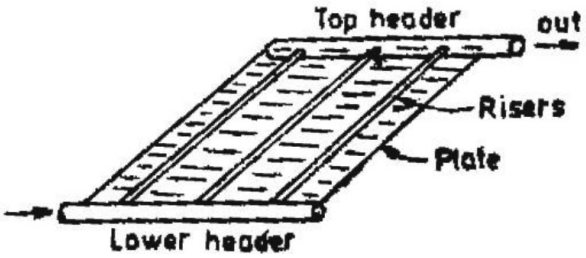


Fig.7: A schematic view of a 'conventional' solar water heater collector plate

The absorber typically uses parallel 1.2 to 1.5 cm diameter tubes, place about 12 to 15 cm apart, soldered or brazed into headers of about 2.5 cm in diameter at the top and the bottom and the tubes are soldered to the plates. Another type has a long continuous tube bent in a sinusoidal shape rather than parallel tubes, and plates formed of one flat and one corrugated sheet fastened together to form water passages (Lawand *et al.,* 1966).

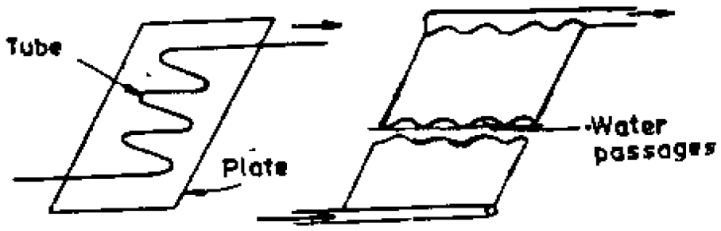


Fig.8: Alternative solar water heater collector plates

Several basic constructional differences are available for the absorber plates and the principal aim is to achieve the following:

- Inexpensive unit that is leak proof and reliable
- Efficient means of transferring heat quickly and evenly to water flow
- Steady and uniform distribution of water
- Maximum energy absorbing surface

Among the common materials used for absorber plates are copper and aluminium, apart from galvanized iron. The absorber plates are mounted in a metal or absorber plates with 5 to 10 cm of insulation behind the plates and 2.5 cm air gaps. It is usually coated with a substance that enhances its absorbing ability. The simplest and cheapest but ideal one is called selective black paint for higher absorption and low emissivity to minimize re-radiation of energy. The most commonly used transparent cover is glass. It is easily available and has good transmittance but it may be expensive, heavy and brittle.

SOLAR DRYING

Food is a basic need for all living things, in addition to air and water. Food shortages arise in most developing and under developed countries around the globe as a result mostly of their inability to preserve food rather than producing it. Agricultural produce are always in excess of immediate consumption, and therefore storage and/or processing becomes imperative. In developing countries, more than 80% of the production is by small scale farmers (Murphy, 2009).

Drying is quite a simple ancient skill. It is one of the easily accessible and widely used processing technologies (Jairaj *et al.,* 2009). Drying is a dual process of:

- Heat transfers to the product from the heating source
- Mass transfer of moisture from the interior of the product to its surface and from the surface to the surrounding air (Ekechukwu & Norton, 1999, p. II)

Solar energy is used either as the main source of heat or as a supplementary source. Agricultural produce are dried in different ways.

1. In direct type drying, produce are exposed to solar radiation or a combination of solar radiation and reflected radiation;
2. In indirect type of solar dryers, produce are NOT exposed directly but are air heated by solar radiation; the air heated by solar radiation is made to pass though them.
3. In the mixed type solar dryer, produce are exposed directly to solar radiation with pre-heated air flowing through.
4. In natural circulation mode, air is heated from different points and made to pass through the produce by buoyancy force or as a result of wind pressure or both.
5. In forced circulation mode, heated air is circulated through the produce by a motorized air blower that is powered by another source, which is probably by conventional electricity or by solar power.

All drying systems can be classified primarily according to operating temperature ranges in two main groups; high and low temperature dryers (Ekechukwu & Norton, 1999, p. I). It is generally accepted that high temperature dryers are powered by fossil fuels and low temperature by solar energy. Solar energy drying systems are schematically classified as shown in Fig.9 below (Ekechukwu & Norton, 1999, p. I).

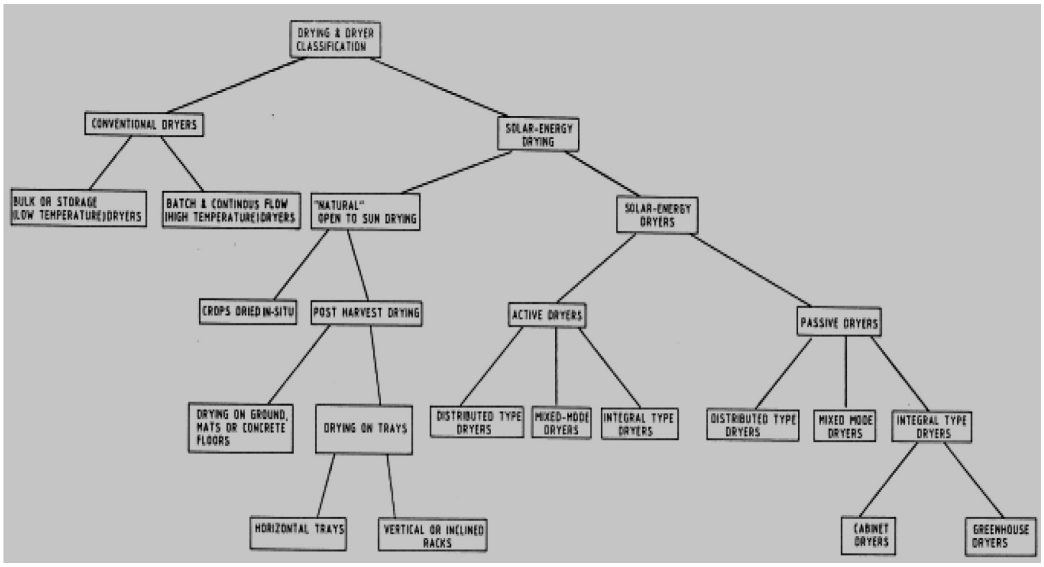


Fig.9: Classification of dryers and drying modes

Fig.9 illustrates a schematic classification of dryers. The concept of three similar classifications under active and passive methods is also schematically shown in Fig.10 (Ekechukwu & Norton, 1999, p. I). In order to further classify the drying methods, various types of solar dryers are identified and schematically presented by Jairaj *et al.* (2009). These include the following:

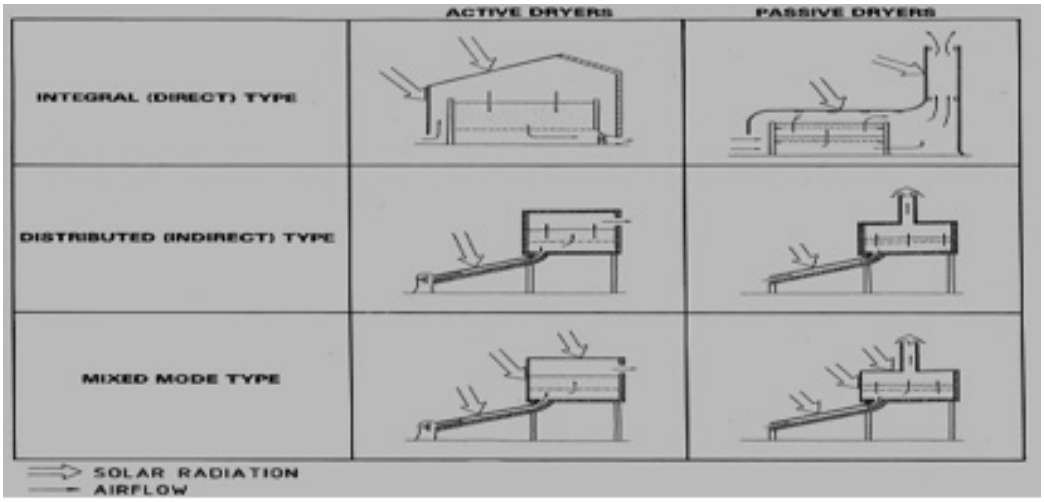


Fig.10: A typical solar energy dryer design

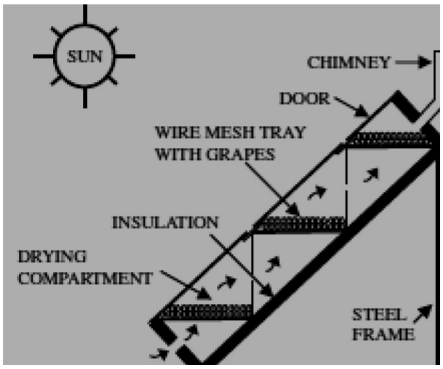


Fig.11: Stair case dryer

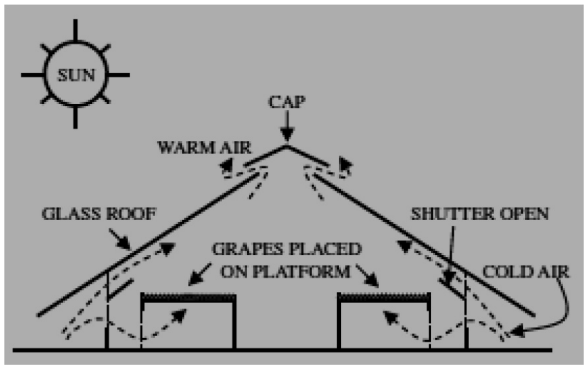


Fig.12: Glass roof solar dryer

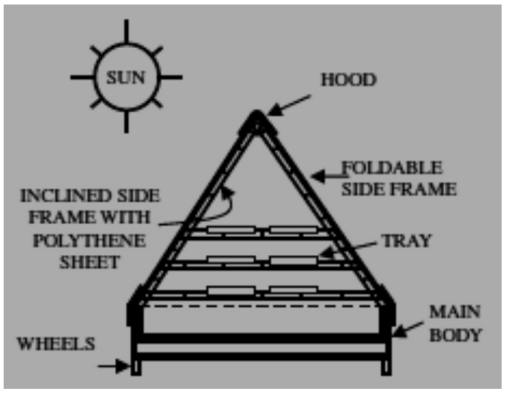


Fig.13: Foldable dryer

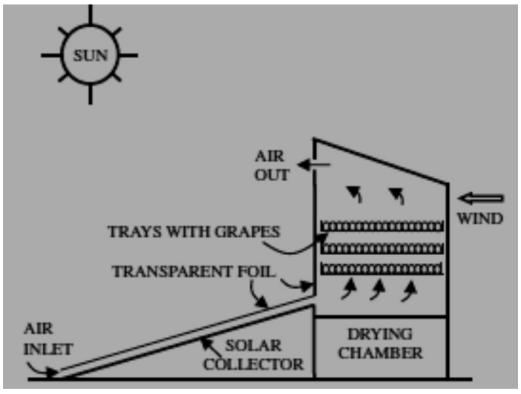


Fig.14: Indirect type of conventional solar dryer

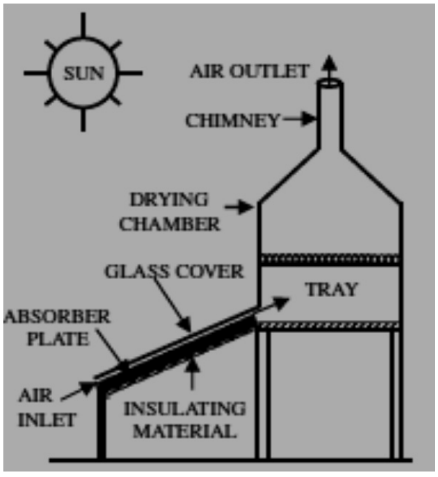


Fig.15: Indirect natural conventional dryer with chimney

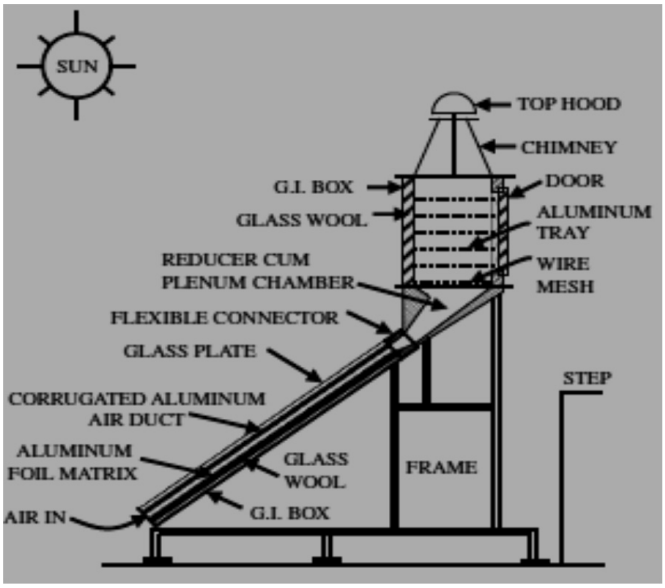


Fig.16: Multipurpose natural convention dryer

Karsli (2007, as cited in Sebali & Shalaby, 2011) made a comparative study of four types of air heating flat-plate collectors: a finned collector with an angle of 75°, a finned collector with an angle of 70°, a collector with tubes and a base collector. He concluded that their efficiency depended significantly on solar radiation, and surface geometry of the collectors and that the overall loss is lower at higher reduced temperature parameter.

Fudholi *et al.* (2010) carried out comparative studies between direct and indirect cabinet dryers. The result tilted in favour of the direct solar dryer. They found that 70% of humidity in pepper was removed within 72 hrs using a direct cabinet dryer compared to within 243 hrs using an indirect one. The mathematical modelling presented by the authors was to put models that describe drying kinetics, and thus:

The first exponential model was in the form of

$$Xr(t) = A_0 + A_1 \exp(-bt) \tag{26}$$

The A and b are functions of drying conditions.
The second proposed model was in the form;

$$Xr(t) = \exp(-\sqrt{kt}) \tag{27}$$

K is the coefficient, a function of drying conditions.

Drying helps to reduce moisture content to a level where deterioration does not occur and dried products can be stored for a definite length of time. Various researchers have worked on testing the performances of different designed dryers, and results have been published by authors such as Klein (1975), Sharma *et al.* (1993) and Tiwari and Ghosa (2005).

Methods of Drying

Drying is the most common form of food preservation and way of extending shelf life of food (Sarasavadha *et al*., 1999). Water, a major constituent of fruit and food, is important in controlling the rates of deteriorative reaction including those resulting in nutrient losses (Sayul & Karel, 1980). This is done by lowering the moisture to a level where microorganisms cannot grow and reaction rates are slowed down (Mahmutoghu *et al.*, 1996).

At the start of drying in the first instance, the food to be dried is very moist or damp, and the moisture content is the same for the inside and outside of the food produce. Then, as the heating continues, with the water particles being taken away by the heating process, the surface of the product begins to dry. This is followed by the migration of water particles from the inside of the product to the surface via diffusion and the drying process continues. At this stage, the energy required for this process is much more important than the first condition, and this stage of drying is subject to type of product. Temperature influence at this stage is critical, as the maximum allowable temperature is about 20^0C in excess of the ambient temperature.

In the open sun drying process, solar radiation heat is used to evaporate the moisture present in the product. Since sunshine is intermittent and varies, the product may become over- or under dried [13]. Hence, solar energy is used to heat a large volume of air and this air flows over the product to remove and take away the moisture. However, hot air can be circulated again to save energy (McDoom *et al.,* 1999). Lack of uniform drying and inaccurate prediction of drying times are among the possible problems that are usually encountered in solar drying (Scheonal *et al.*, 1996). Equilibrium moisture level is defined as the moisture level when in equilibrium with the relative humidity of the environment (Bala, 1998; Klein, 1975).

Open Sun Drying without Cover

In this case, part of the solar radiation falls, penetrates and be absorbed by the material itself resulting in heat generation in the interior of the product as well as its surface, and thus generating heat transfer. Solar radiation absorptance-transmittance product is an important factor here. Generally, the agricultural produce in the dryer are, at intervals, rotated to expose the other parts to sun's radiation and ensure a uniform drying. The drying time length varies according to drying materials pending on the amount of water content of the produce; however, the environmental condition of dryer usage may affect the performance. Consider Fig.17 below for grapes drying (Jairaj *et al.,* 2009).

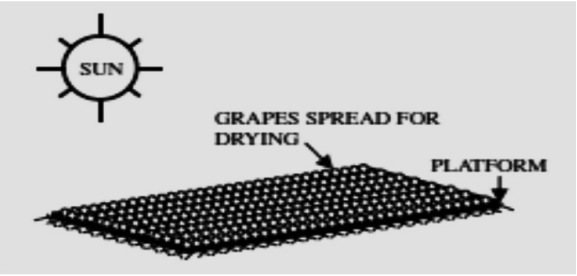


Fig.17: Open sun drying without cover

Open Sun Drying with Cover

A schematic view of open sun drying with plastic cover for grapes is shown in Fig.18 (Jairaj *et al.,* 2009). In this case, the plastic cover offers two services; as a heat trap and as a protective cover from contamination to a certain extent. Jairaj *et al.* (2009) found that the quality of produce from the grapes dried with cover is better than that from grapes with open sun drying.

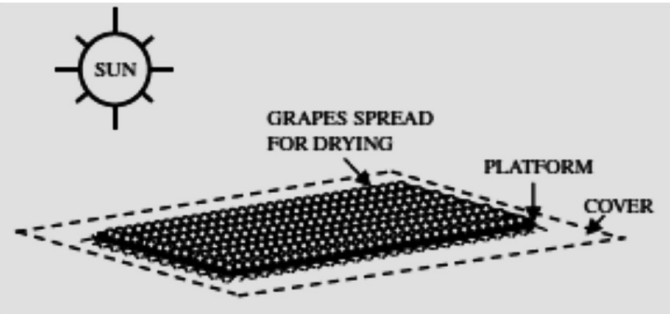


Fig.18: Open sun drying with cover

Natural Rack Drying

This is also another form of shade drying process. Fig.19 depicts a shade dryer. The dryer may consist of up to 10 shelves or more with a uniform spacing between each shelf. The orientation is such that the contents receive solar radiation in the morning and afternoon, while shaded at zenith angle with the shade provided by the roof. For grapes, the drying may take place in two weeks (Amba & Anand, 1972).

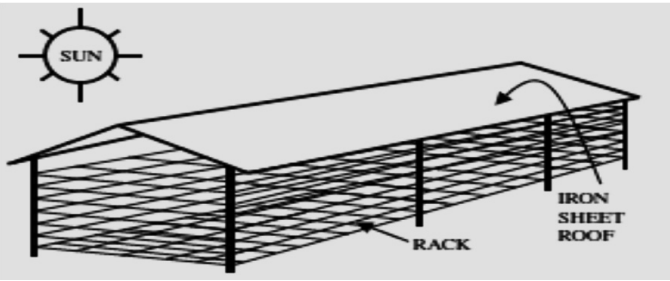


Fig.19: Natural rack dryer

OTHER APPLICATIONS OF FLAT PLATE COLLECTORS

Apart from the applications of the flat plate solar collectors discussed above, other applications include:

(a) Solar pond, in which a mass of shallow water about 1.5 – 2 metre deep, with a large collection area, acting as a heat trap. In such a pond, salt is added to maintain a stable density gradient. As solar incidence falls, it penetrates throughout the depth of the pond with its black bottom.

(b) Solar thermal power generation, in which solar energy is converted into thermal energy by flat plate or concentrating collectors. Thermal energy generated is used to drive heat engine.

(c) Solar furnaces, in which large amount of heat energy is generated by concentrating solar radiation onto a specimen. In this case, flat plate collectors are utilized to boost energy generation.

(d) Solar water pumping, which can be used for irrigation for drinking, either by animals or by humans. Flat plate is utilized to generate heat which is then used by a heat exchanger to power a heat engine that will in turn start a heat pump.

(e) Solar distillation, either for obtaining distilled water or for obtaining potable water from sea or salty water body. This is accomplished by exposing thin body of water to solar radiation and condensing the water vapour produced on a transparent cover so that it is collectable in a receiving trough. The thin layer of water lies on a flat plate collector.

(f) Solar cookers, in which the heat collected by the flat plate in the cooker box is transmitted to the food, with solar concentrator on the plate to boost heat generation.

CONCLUSION

A brief overview of the basic relations used in the analysis of solar collectors is presented in solar cooling and heating of buildings, air and water heating, and agricultural produce drying. Different types of flat-plate solar collector utilizations are enumerated. Meanwhile, types of agricultural product drying methods are presented in relation to specific products. Flat plate collectors are generally concluded to be less efficient than V-corrugated collectors, finned plate collectors, and chevron pattern collectors. However, flat plate collectors are suitable for low temperature applications. The overall heat loss coefficient determining the performance of flat-plate collectors and possible losses from top, bottom and sides have also been enumerated in the review.

REFERENCES

Abhishek, S., Varun, S. P., & Srivastav, G. (2011). A thermodynamic review of solar box cookers. *Renewable and sustainable energy review, 15,* 3301-3318.

Aboul-Ennei, S., El-Sebaii, A. A., & Ramadan, M. R. I. (2000). Parametric study of a solar air heater with and without thermal storage for solar drying applications. *Renewable energy, 21*, 505-522.

Akhtar, A., & Mullick, S. C. (2007). Computations of glass-cover temperatures and top heat loss coefficient of flat-plate solar collectors with double glazing. *Energy Journal, 32*, 1067-1074.

Akhtar, N., & Mullick, S. C. (1999). Approximate methods of computation of glass-cover temperatures and top loss coefficient of solar collectors with single glazing. *Solar Energy, 66(5*), 349-354.

Alta, D., Bilgili, E., Erketin, C., & Yalzid, O. (2010). Experimental investigation of three different solar air heaters: energy and exergy analyses. *Applied Energy, 87*, 2953-2973.

Amba, D., & Anand, J. C. (1972). Make raisin when the grape is plenty, *Indian Horticulture, 5,* 6-10.

Atul, S., Chen, C. N., & Nguyen, V. L. (2009). Solar energy drying systems: a review. *Renewable and sustainable energy reviews, 13,* 1185-1210.

Baker, S., McDaneils, D. K., Lowndes, D. H., Mathew, H. M., Reynolds, J., & Gray, R. (1978). Enhanced solar energy collection using reflector-solar thermal collector system. *Technical Notes, Solar energy, 20,* 415.

Bala, B. K. (1998). *Solar drying systems: simulation and optimization* (1st Edn.). Udaipur, India: Agrotech Publ.

Bejan, A., Tsatsaronis, G., & Moran, M. (1996). *Thermal design and optimization*. New York: John Willey & Sons.

Choudhury, C., Chauhaun, P. M., & Garg, H. P. (1995). Performance and cost analysis of two pass solar air heater. *Heat recovery systems and CHP, 15*(8)*,* 755-773.

Cooper, P. I., & Dunkle, R. V. (1981). A non-linear flat –plate collector model. *Solar Energy, 26*, 133-140.

Duffie, J. A., & Beckman, N. A. (1980). *Solar Engineering of thermal processes*. New York: Willey.

Ekechukwu, O. V., & Norton, B. (1999). Review of solar energy drying systems I: an overview of drying principles and theory. *Energy conservation and Management, 40*, 593-613.

Ekechukwu, O. V., & Norton, B. (1999). Review of solar energy drying systems II: an overview of drying technology. *Energy conservation and management, 40*, 615-655.

Ekechuwu, O. V., & Norton, B. (1999). Review of solar energy drying systems III: low temperature air-heating solar collectors for crop drying applications. *Energy Conservation and Management, 40*, 657-667.

El-sawi, A. M., Wifi, A. S., Younan, M. Y., Elsayed, E. A., & Basily, B. B. (2010). Application of folded sheet metal in flat bed air heaters. *Applied thermal Engineering, 30*, 864-871.

Francey, J. L. A., & Papaioannou, J. (1985). Wind-related heat losses of a flat-plate collector. *Solar Energy Journal, 35*(1)*,* 15-19.

Fudholi, A., Sophia, K., Rusia, A., & Sulaiman, M. Y. (2010). Review of solar dryers for agriculture and marine products. *Renewable and Sustainable Energy Reviews, 14*, 1-30.

Garg, H. P, Gouri, D., & Bandyopadhyay, B. (1983). The temperature dependence of the overall heat-loss coefficient for flat plate collectors. *Energy, 8*(6), 451-459.

Goyal, R. K., & Tiwari, G. N. (1999). Performance of a reverse flat-plate absorber cabinet dryer: a new concept. *Energy conversion and Management, 40,* 385-392.

Hegazy, A. A. (1999). Optimum channel geometry for solar air heaters of conventional design and constant flow operation. *Energy Conservation and Management, 40,* 757-774.

Hegazy, A. A. (2000). Performance of flat solar air heaters with optimum channel geometry for constant/variable flow operation. *Energy conservation and Management, 41*, 401-417.

Henry, Z. A., Bledsoe, B. L., & Edder, D. D. (1977). Drying of large hay packages with solar heated air. *ASAE Paper 77*, 3001.

Ho, C. D., Yeh, H. M., Cheng, T. W., Chen, T. C., & Wang, R. C. (2009). The influence of recycle on the performance of baffled double pass plate solar air heater with internal fins attached. *Applied Energy, 86,* 1470-1478.

Hotel, H. C., & Woertz, B. (1942). Performance of flat-plate solar heat collectors. *Trans ASME, 64*, 91.

Jairaj, K. S., Singh, S. P., & Srikant, K. (2009). A review of solar dryers developed for grapes drying. *Solar Energy Journal, 83*, 1698-1712.

Karsli, S. (2007). Performance analysis of new design solar air heater collectors for drying applications. *Renewable Energy, 32,* 1645-1660.

Kaygusuz, K., & Bilgen, S. (2009). Thermodynamic aspect of renewable and sustainable development. *Energy Sources, Part A, 31,* 287-298.

Klein, S. A. (1975). Calculations of flat-plate collector loss coefficient. *Solar Energy Journal, 17*, 79-81.

Lawand, T. A. (1966). *Solar Energy, 10*(4), 158-164.

Lyes, B. (2011). Reviewing the experience of solar drying in Algeria with presentation of the different design aspects of solar dryers. *Solar Energy 15*(2), 3371-3379.

Mahmutoghu, T., Felunde, E., & Birol, S. Y. (1996). Sun/ solar drying of different treated grape and storage stability of dried grape. *Journal of Food Engineering, 29*, 289-300.

Malik, M. S. A, Tiwari, G. N., Kumar, A., & Sodha, M. S. (1982). *Solar Distillation.* Oxford: Pergamon Press.

McDoom, I. A., Ramsaroop, R., Sunkers, R., & Tang, K. A. (1999). Optimization of solar crop drying. *Renewable Energy, 16,* 749-752.

Minuralini, T., Iniyan, S., & Ranko, G. (2010). A review of solar thermal technologies. *Renewable and Sustainable Energy Reviews, 14,* 312-322.

Murphy, M. V. R. (2009). A review of new technologies, models and experimental investigation of solar driers. *Renewable and Sustainable Energy Reviews*, *13*(4), 835-848.

Naphon, P., & Kongtragol, B. (2003). Theoretical study of heat transfer characteristics and performance of flat-plate solar air heaters. *International Communication in Heat and Mass Transfer, 30*(8), 1125-1136.

Ong, K. S. (1995). Thermal performance of solar air heaters, mathematical modeling and solution procedure. *Solar Energy, 55*(2)*,* 93-109.

Rai, G. D. (1984). Solar Energy utilization. India: Khanna Publishers.

Rosenbaum, M. (1991), *Solar Today* (5:5). In *Solar Hot Water for the 90s*, p. 20.

Sarasavadha, P. N., Sadhey, R. L., Pangadhave, D. R., & Singh, S. P. (1999). Drying behavior of brined onion slices. *Journal of food Engineering, 40,* 219-226.

Sayug, I., & Karel, M. (1980). Modeling of quality deterioration during food processing and storage. *J. of Food Tech., 78,* 84-86.

Scheonal, G. J., Arinze, E. A., Sokhansaj, S., & Trauttmansdoft, F. G. (1996). Evaluation of energy conservation potential by exhaust air recirculation of a commercial type heated air batch hay drier. *WREC*, 676-681.

Sebaii, A. A., & Shalaby, S. M. (2011). Solar drying of agricultural products: A review. *Renewable and Sustainable Energy Reviews, 15,* 348-359.

Sharma, A. K., & Sodha, M. S. (1986). Testing of solar collectors in solar water heating system. *Sol. Energy J., 34*(17), 383-397.

Sharma, V. K., Colnagole, A., & Spagna, G. (1993) Experimental performance of an indirect type solar food and vegetable dryer. *Energy conservation and Management, 34*(4), 293-298.

Sheffer, M. B. (1994). Solar Water Heating: A Viable Technology Alternative. *Energy User News, 19*(9), 44.

Soddah, M. S., Bansal, N. K., Kumar, K. B., Ansal, P. K., & Malik, M. A. S. (1982). *Solar crop drying*. Florida, USA: CRC Press.

Sodha, M. S., Dang, A., Bansal, P. K., & Sharma, S. B. (1985). Using solar for development. *Energy Conv. Man., 25*(3), 263-271.

Soteris, K. (2009). *Solar energy engineering processes and systems*. Academy Press.

Sparrow, E. M., Ramsey, J. W., & Mass, E. A. (1977). Effect of finite width on heat transfer and fluid flow about an inclined rectangular plate. *ASME Journal of Heat Transfer, 99*, 507- 511.

Taboh, H. (1966). Mirror boosters for solar collectors. *Solar Energy, 10*, 111-118.

Tawari, R. C., Ashvini, K., Gupta, K., & Sootha, G. D. (1991). Thermal performance of flat plate solar collectors manufactured in India. *Energy Conversion and Management, 31*(4), 309-313.

Tiwari, G. N., & Ghosa, M. K. (2005). *Renewable energy resources: basic principles and applications*. Narosa Publishing House.

Torres-Reyes, E., Navarrete-Gonzales, J. J., & Ibarra-Salazar, B. A. (2002). Thermodynamic method for designing dryers operated by flat-plate solar collectors. *Renewable energy, 26*, 649-660.

Yeh, H. M., Ho, C. D., & Sheu W. S. (2000). Double pass heat or mass transfer through a parallel channel with recycle. *International Journal of Heat and Mass Transfer, 43*, 487-491.

Pertanika J. Sci. & Technol. 22 (2): 387 - 400 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review Article

Development of Vaccination Strategies: From BCG to New Vaccine Candidates

Nadiya T. Al-alusi* and Mahmood A. Abdullah

Department of Molecular Medicine, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

ABSTRACT

During recent years, extensive development has been made to improving vaccines for tuberculosis. This is due to the presence of genome sequences of diverse mycobacterial species and *Mycobacteroum tuberculosis* (*M. tuberculosis*) isolates which has led to advances in the characterization of genes and antigens of *M. tb* and better realization of protective immune responses to the disease in both animals and humans. This review summarizes vaccine types, reasons for variable efficacy of BCG, latest advances in tuberculosis vaccine development and major vaccine design strategies.

Keywords: Tuberculosis, vaccines, bacille Calmette-Guerin (BCG), Region of difference (RD)

INTRODUCTION

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB). The proof for the presence of tuberculosis in ancient times was found when fragments of the spinal column from Egyptian mummies from 2400 B.C. showed definite signs of tuberculosis (Morens, 2002).

Article history:
Received: 26 March 2012
Accepted: 21 May 2012

Email addresses:
Nadiya T. Al-alusi (nadiadarweesh@yahoo.com),
Mahmood A. Abdullah (mahmood955@yahoo.com)
*Corresponding Author

ISSN: 0128-7702

Tuberculosis ranks second only to human immunodeficiency virus as a cause of death from an infectious agent, with 1.7 million deaths from TB in 2007 (Philippe *et al.*, 2009). The most important factor for TB prevalence is the incidence of HIV (Simonney *et al.*, 2007). MDR-TB is another incidence factor for TB disease, as TB bacilli resistant to at least isoniazid and rifampicin, the two most powerful anti-TB drugs (von der Lippe *et al.*, 2006).

Calmette and Camille Guerin is the current vaccine used against TB. BCG has been administered worldwide. BCG offers unique advantages as a vaccine because:

(1) it is unaffected by maternal antibodies and can therefore be given at any time after birth; (2) BCG is usually given as a single dose eliciting a long-lasting immunity; (3) it is stable and safe; (4) BCG can be administrated orally; and (5) it is inexpensive to produce compared to other live vaccines. In addition, the extraordinary adjuvant properties of mycobacteria make them an attractive vector for the development of recombinant vaccines (Reginaldo *et al*., 2009). However, doubts about its efficacy are increasing, and these have been particularly reflected by its highly variable protective efficacy in controlled clinical trials. A number of suggested reasons for the variation of BCG efficacy include the genetic variability amongst, and different age of, the vaccinated individuals and immunological cross reactivity between BCG and environmental mycobacterial strains prevalent in the different parts of the world (Brandt *et al.*, 2002).

Recently, there has been growing interest in improving vaccine as an alternative to BCG for controlling TB. Many vaccines have been presented, and they can be classified mainly into live-attenuated vaccine (Kamath *et al.*, 2005), subunit vaccine (such as Ag85 and ESAT-6) (Brooks *et al.*, 2001; Orme, 2006; Olsen *et al.*, 2001) and DNA vaccines (Nor & Musa, 2004; Kaerch *et al.*, 2002). Highlighting BCG defects and finding the correct strategies for vaccine design may possibly help to produce effective vaccines against tuberculosis.

TUBERCULOSIS VACCINES

Live Attenuated Vaccine

An effective cell mediated immunity (CMI) response is important in controlling tuberculosis because it is an intracellular pathogen. The fact that live-attenuated strains are generally more potent than non-living vaccines in stimulating CMI responses supports the use of live vaccines. Moreover, the immune responses elicited following vaccination highly resemble the responses that follow a natural infection, since most of the antigens are expressed *in vivo* (Sambandamurthy & Jacobs, 2005).

In spite of the fact that live vaccines are effective against intracellular pathogens, the development of these vaccines is difficult because of the difficulty in balancing between producing required level of attenuation and immunogenicity. The strength of the immune response is a function of the amount of antigen expressed and T-cell response against individual antigens. Obviously, the high attenuation of BCG strain for several decades of years has led to limited replication *in vivo* and subsequently to BCG failure (Sambandamurthy & Jacobs, 2005).

The first strategy of researchers, i.e. improving live-attenuated vaccine for tuberculosis, has focused on strengthening the immunogenicity of vaccine or producing genetically modified recombinant BCG strains by insertion or over-expression of immunodominant antigens or immunostimulatory cytokines of *M. tuberculosis*. BCG30 was the first recombinant BCG vaccine reported to induce higher protective immunity to TB than the standard BCG vaccine in animal models. This vaccine which is constructed by a plasmid expresses the 30 kDa major secreted proteins. Another example is Ag85B which is introduced into BCG Tice strain. The plasmid containing Ag85B gene has been shown to be stable for long periods of time without antibiotic selective pressure *in vitro* and *in vivo* (Kamath *et al.*, 2005; Sambandamurthy & Jacobs, 2005).

Extending the idea of inducing BCG vaccine to express additional *M. tuberculosis* antigens, more than one *M. tuberculosis* antigen was introduced to BCG Pasteur strain by introducing RD-1 (region of difference) locus of *M. tuberculosis*, which expresses immunoprotective proteins such as ESAT-6 and CFP-10. In spite of the capability of ESAT-6- to stimulate specific $CD4^+$ T cells following immunization, this vaccine candidate displayed increased virulence in the animal model compared to the parental strain (Kamath *et al.*, 2005; Sambandamurthy *et al.*, 2005).

To improve the immunogenicity of BCG vaccine, recombinant BCG has been modified to produce (rBCGΔUreC: Hly) vaccine through deletion and addition of genes into BCG pasture strain. The deletion was done for urease gene, which normally enables mycobacteria to block the acidification of the early phagosome of the macrophage. As a result, urease deletion can prevent neutralization of the acidic pH in phagosomes. The vaccine (rBCGΔUreC: Hly) has also been reengineered by over-expression of listeriolysin O, which increases MHC class I presentation and $CD8^+$ T-cell responses. This recombinant BCG (rBCGΔUreC: Hly) was found to be safe for mice and provide protection against aerosol TB in the mouse model (Kamath *et al.*, 2005; Tony *et al.,* 2011).

The second strategy for producing tuberculosis vaccine is by using attenuated mycobacteria such as auxotrophic *M. tb* strains like *Pho*P mutant of *M. tuberculosis*. This vaccine was produced by disrupting the gene of *phoP* in the MT103 strain of *M. tuberculosis*. *phoP* importance belongs to its involvement in regulating a number of genes in *M. tuberculosis*. PhoP has also been identified to be a virulence factor, as a highly virulent strain of multi-drug resistant tuberculosis was found to be over-expressing the gene of *phoP* during a major disease outbreak (Soto, 2004). Nonetheless, the attenuation of virulence by deletion of the *phoP* gene has not affected the induction of wild-type *M. tuberculosis*-like immune response (Kamath *et al.*, 2005).

On the same base of attenuation, mutants for *M. tuberculosis*, *M. tuberculosis* mc^2 6020 vaccine and *M. tuberculosis* mc^2 6030 were produced. In mc^2 6020 vaccine, the genes encoding LysA which produce non-replicating *M. tuberculosis* and panC, D genes involved in lipid metabolism were deleted from *M. tuberculosis* genome, while in *M. tuberculosis* mc^2 6030 vaccine, the RD-1 gene region was deleted, and this means inhibition for ESAT-6, CFP-10 and other proteins encoded by this region. In spite of the fact that both mutants have been shown to be severely attenuated, their protective efficacy was found to be similar to BCG immunization (Kamath *et al.*, 2005). However, their evaluation in humans still involves problems with safety and stability.

Less virulent mycobacteria such as *M. microti*, *M. vaccae* or *M. smegmatis* that overproduce immunogenic antigens of *M. tuberculosis* are also another choice for producing vaccine against tuberculosis (Girard *et al.*, 2005). The vaccination of BALB/c and C57BL/6 mice by *M. microti* orally or by aerosol and subsequently challenging them via respiratory route with virulent *M. tuberculosis* gave a slightly better protection than BCG at oral dose of 100 million CFU (Helke *et al.*, 2006). However, the authors' claim must be viewed with some scepticism due to the massive dose required and the absence of a BCG vaccinated group which received the same high dose (Gupta *et al.*, 2007). Heat-killed *M. vaccae* is another vaccine from environmental mycobacteria which has been reported to stimulate Th1 response.

However, a high dose can only stimulate Th2 response, which is required for fighting against tuberculosis (Stanford, 1991).

Using environmental mycobacterium is not expected to be more effective than BCG since the first factor in BCG failure is losing specific genes for *M. tuberculosis* and these strains may be genetically different and less resemble *to M. tuberculosis* antigens, while the induction of CD8+ T cells may not appear as vital as in virtual *M. tuberculosis* infection.

Although live vaccines can provide sufficient and effective level of protection, safety concerns still hinder broad public acceptance of attenuation. For *M. tuberculosis* mutants as vaccine candidate, additional safety criteria such as verification of stable attenuation are still required. At least two no reverting independent mutations are strongly recommended for vaccines derived from *M. tuberculosis* before they could be considered for human trials (David *et al*., 2011).

Both rBCG30 and BCG ÄureC::hly recombinant live attenuated vaccines have entered clinical trials in the last decade. The development of rBCG30 has been achieved by the University of California. Preclinical models have demonstrated a better protection of rBCG30 over its parental strain and a Phase I clinical trial has proven its safety in adults. Two variants of the candidate vaccine were generated to limit its replication *in vivo*, thereby striving for increased safety in immunocompromised patients: (1) deletion of the *mbtB* gene rendering the strain dependent on exogenous mycobactin for iron acquisition, and (2) deletion of *panCD* conferring auxotrophy for pantothenate, vitamin B5. Although elevated efficacy and safety of rBCG30 variants over parental BCG have been achieved, the development of this candidate is currently on hold (Kaufmann, 2012).

Meanwhile, the vaccine candidate VPM1002 (BCG ΔureC::hly) induced better protection against TB in a preclinical challenge model. Superior protection of this particular vaccine over parental BCG is related to increased stimulation of different T cell populations involved in protective immunity. This is probably due to cross-priming as a consequence of elevated apoptosis of antigen-presenting cells harboring VPM1002. This recombinant candidate completed safety evaluation in adults in two Phase I trials and it is currently undergoing Phase II assessment in South Africa to determine its safety and immunogenicity in the target population (Kaufmann, 2012).

Live-attenuated vaccines offer a very potent means to prevent several human diseases, and a number of them are in routine use such as vaccines for polio, measles, mumps, rubella, and varicella. In general, these vaccines are safe, efficacious, and induce both local and systemic immune responses. While stimulating both humoral and cell-mediated immune (CMI) responses, live-attenuated vaccines also activate the innate and adaptive branches of the immune system. Although live vaccines offer great promises against intracellular pathogens, one obstacle to the development of these vaccines is the difficulty in achieving a satisfactory level of attenuation without severely compromising immunogenicity. The strength of the immune response is a function of the amount of the antigen expressed (Bastos & Borsuk, 2009).

Subunit Vaccine

Total cell wall, protoplasm and killed whole cell preparations were tested as experimental vaccine candidates for tuberculosis, including polypeptides and carbohydrates and lipids. Although the subunit vaccine candidate can be extracted from culture filtrates of *M. tuberculosis,* the emphasis is given to the proteins of *M. tuberculosis* surface such as the Ag85 mycolyl transferase family, the RD1 gene region proteins, ESAT-6 and CFP-10. After all, the filtrate alone contains more than 100 proteins revealed by proteomic analysis and theoretically has multiple protective antigens (Orme, 2006; Nicholas *et al.,* 2011).

The cell-mediated immune response, but not the humeral immunity, is essential for controlling intracellular pathogens including *M. tuberculosis*. Subsequently, the molecules inducing dominant T cell response and no B cell (antibody) response in *in vitro* immunological assays are the references for vaccine designer. IFN-γ is considered to be the most potential factor in protective immunity, as the IFN-γ gene disruption has led to fatal growth of bacilli in infected mice (Casadevall, 2003).

The Ag85 complex (A–C) consists of 30–32 kDa proteins' family that acts as mycolyl transferase. Ag85 has been used in two ways; first, as main vaccine with delivery vehicles such as liposomes and microspheres and the result of these studies showed reasonable protection from a high dose intravenous challenge. Second, Ag85 has also been used in boosting young mice in middle age with Ag85A after BCG vaccination. The result showed restoring to the waning immunity for *M. tuberculosis* challenge in old mice (Brooks *et al.*, 2001; Orme, 2006).

ESAT-6 is another candidate of tuberculosis vaccine. The importance of this protein is contributed by its presence in *M. tuberculosis*, but not in BCG. Some studies have shown ESAT-6 immunogenicity and it was found to have produced IFNγ responses in mice when it was mixed with potent adjuvant as MPL/DDA. However, the protection values in those studies were relatively small (Brandt *et al.*, 2002).

Fusion proteins or polyproteins (recombinant engineering vaccine) were first applied by Olsen (Olsen *et al.*, 2001) when a fusion protein made up of Ag85A and ESAT-6, and used to protect mice from *M. tuberculosis* challenge. The most important finding in Oslen's fusion protein is that the immunity sustained for at least 30 weeks and the vaccine could be given orally (Orme, 2006).

Mtb72F is a fusion protein constructed by Skeiky and his colleagues. This molecule consists of a polyprotein made up of Mtb32[C-terminal]–Mtb39–Mtb32 [N-terminal] of molecular size 72 kDa. Mtb72F in its native form successes in boosting BCG and stimulating strong IFNγ and CD8 CTL responses (directed to an epitope in the Mtb32-C component). The guinea pigs showed protection for over 50 weeks in similarity to BCG controls. Moreover, in the surviving animals, there were very few obvious granulomatous lesions and most of the lung tissues were clear (Orme, 2006; Skeiky *et al.*, 2004).

In the absence of reliable methods to predict the protective molecules, individual mycobacterial components (especially large number of proteins) are primarily screened for their immunoreactivity in terms of recognition by T lymphocytes and subsequent induction of IFN-γ

in vitro cultures. *In vivo* 'protective immunity' method by low dose infection animal models, followed by challenge experiments, is the standard protocol to evaluate the immunobiological features for individual antigens (Sable *et al.*, 2007a).

There are some features for the ideal candidate as tuberculosis vaccine can be considered before vaccine designing. It is believed that low molecular polypeptides can induce dominant Th1 response *in vitro* and they are recognized by T lymphocytes of animal and different human populations (Sable *et al.*, 2005b). The potential antigenicity of low molecular polypeptides may attribute to their stability and resistance to proteolysis process or because their small size render them more susceptible to proteolytic degradation, processing, intracellular trafficking and presentation (Skjot *et al.*, 2000). On the other hand, molecular mass proteins have failed to produce significant protection as compared to BCG and more than 30% of subunit proteins stimulating strong IFN-γ response (such as ESAT-6 and CFP-10, TB 10.4, TB 10.3) are not able to protect immunized mice from bacilli challenge, unless injected with multiple strong adjutants (Brandt *et al*., 2002). Moreover, EAT-6 has been observed to develop an active disease within a period of 2–5 years (Doherty *et al.*, 2002; Sable *et al.*, 2007a).

The hydrophobic or hydrophilic nature of a ligand and its homology with a receptor on the cell surface are important conditions for immunobiological activity through the interaction with class-I and class-II MHC molecules and induction T-cell responses (Doytchinova & Flower, 2001).

Structural stability with respect to the ability to induce an immune response is a potential factor for subunit vaccine. Many mycobacterial secretory proteins were found to be prone to degradation making them unacceptable as subunit vaccines. The failure of HSP-70 as a vaccine candidate is due to its instability and role in autoimmune diseases and immunopathology. On the other hand, some secretory proteins like HSP-16 and 19 kDa lipoprotein are highly stable due to the post-translational modifications (Sable *et al.*, 2007a). However, a possible negative aspect of stability is immunopathology as immunization with some of these long-lived molecules has been found to be correlated with immunopathology when it was used as experimental vaccines (Edwards *et al.*, 2001).

The post-translational modifications of mycobacterial proteins can positively or negatively affect in vaccine candidature by altering their immunogenicity or pathogenicity and this theory appears clearly in 45/47 kDa alanine-proline-rich antigen because its immunogenicity decreases with the changes in the glycosylation pattern (Romain *et al.*, 1999).

DNA Vaccines

A DNA vaccine stimulates immunity against an infectious disease by presenting particular genes or epitopes from an organism to the immune system (Nor & Musa, 2004). Wolff has found that expression of plasmid DNA carrying foreign genes is possible in mouse skeletal muscle (Istvan & Wolff, 1994). The produced protein is exposed on class I MHC molecules by endogenous pathway, followed by presentation to CD8 T cells (Kamar & Sercarz, 1996). Transfection of APCs is the second mechanism in DNA vaccine; however, limited number of DCs involves in transfection and priming immune response (Fine, 2001). Cross priming is the third mechanism in DNA vaccine. APCs receive antigen from myocytes or apoptotic cells and is subsequently presented to CD8 T cells (Sanjay *et al.*, 2000).

The use of a DNA vaccine offers many advantages in terms of stability, safety, and easiness in production and storage. The most important point, as shown in animal studies, is that DNA vaccines are able to induce both humoral and cellular immune responses (Montgomery *et al.*, 1997).

Furthermore, one DNA plasmid can yield many copies of the antigen over a prolonged period since antigens are endogenously expressed and presented on the MHC class I and activation of cytotoxic T lymphocytes (CTL), which is required against an intracellular pathogen like *M. tuberculosis* (Nor & Musa, 2004). However, the rate of protection for DNA vaccines is lower than that conferred by the BCG vaccine, so a cocktail of *M. tuberculosis* antigens to sustain protective immunity has been presented by some researchers as a solution for this particular drawback (Delogu *et al.*, 2000).

DNA vaccine can be constructed together with immunostimulatory molecules such as cytokine genes and immunostimulatory DNA sequences like CpG motif (unmethylated DNA sequence for build -in adjuvant) to enhance NK cell cytolytic activity and stimulate APCs since these molecules can polarize the immune response to Th1 (Kaerch *et al.*, 2002) and provide an indubitable protective immunity against TB. It has been suggested to immunize mice with IL-12 DNA plus the antigen of interest, as this combination can stimulate the immune system towards TH 1 cellular response and controlling the infection for a long term (Gurunathn *et al.*, 1998).

The other proposed improvement for DNA vaccine is incorporating granulocyte-macrophage colony-stimulating factor (GM-CSF). Although the addition of (GM-CSF) did not develop protective efficacy, increases in lymphoproliferation and INF production were detected (Kamath *et al.*, 1999).

Many suggestions have been presented to improve the plasmid DNA uptakes by the cells such as epidermal gene gun with DNA–coated gold particles. Unfortunately, this method induced type 2 immunity more than type 1, which is required against *M. tuberculosis.* Electrical pulse is another attempt to increase cell uptake for plasmid DNA (Mir, 2001). However, it is still unknown whether this method can actually increase immune protection.

According to the source of antigen, most of the genes used to produce vaccine against tuberculosis belong to *M. tuberculosis* such as Ag 85, 38kDa, Mtb39A, Pst-3, ESAT-6, Mtb 8.4, 72f and MPT-64,63,83, series and Ag85B,MPT-64+Ag85B, MPT-64+Ag85B+GMC-SF cocktails have shown some degree of protection for *M. tuberculosis.* However, other genes such as Hsp70, Hsp65 and 6kDa belonging to *M. leprae* also provide immunogenic protection to different models of infection (Taylor, 2003; Okada *et al.,* 2011).

Heat shocked protein (Hsp60) and Ag85A as DNA vaccines are important and this may attribute to the capability of these genes to elicit protection equals to BCG via induction TH1-type CD4 T cells and cytokines (Lowrie *et al.*, 1997; Taylor, 2003).

However, there are many factors to be evaluated in DNA vaccine before applying it in human level, especially in term of its safety, detecting the best method of injection, number of doses and finding the most antigenic and immunogenic epitops or by using cocktail of several mycobcterial antigens.

FAILURE OF BCG

BCG was produced in 1921. Ever since its early uses, the efficacy of the BCG vaccine has been questioned and an extensive number of trials have been carried out to analyze its performance (Agger & Andersen, 2002). Hence, understanding the reasons for failure in BCG may lead to identifying steps necessary for designing more effective vaccine.

Using BCG in all parts of world has led to high attenuating and as a result, differences in both phenotypic and genotypic levels, not only comparison with the original parents strain of BCG but also between the various daughter strains of BCG. The first theory, which may explain such difference, is the effect of mutations by losing a number of genes with potential relevance to protective immunity and this theory is supported by the variable protection level found in clinical trials using different strains of BCG (Behr *et al.,* 2002). The second proposed reason for the failure of BCG is the absence of specific T-cell antigen or BCG stimulates non-optimal blend of T cells; in another words, the lack of a particular induction of CD8 cells and subsequently insufficient level of immunity (Kaufman, 2001).

The final explanation for an inadequate protective effect of BCG is the human population's exposure in the tropical regions to large quantities of environmental mycobacteria, which may interfere with the efficacy of the BCG vaccination. Almost all defects of BCG such as the interfering of environmental mycobacteria and the absence of suitable antigen to T cells subset may be solved by using DNA vaccine (Agger & Andersen, 2002).

TUBERCULOSIS VACCINE STRATEGIES

The optimal tuberculosis vaccine may be achieved by understanding the limitations of BCG and designing a plane to overcome them.

BCG Lacks Important Antigens—exploiting genomics for rational vaccine development

Comparative genomic has revealed the differences between BCG and virulent *M. tuberculosis* as a number of regions designated RD1–RD16 and encompassing 129 open reading frames have been identified as absent regions in BCG vaccine. These genes are present in virulent mycobacteria but deleted during the attenuation and repeated passages of BCG (Brosch *et al.*, 2000). Some of these genes are likely associated with a virulence character and could play an important role in the failure of BCG; therefore, reintroducing selected genes from RD1–RD16 to BCG has been suggested as a way to enhancing the protective efficacy of the existing BCG vaccine. Hence, this study is an attempt to explore the capability of the RD regions as a possible candidate for TB vaccine.

RD1 (region of difference) is the best characterized region. Reintroduction of the RD1 into BCG has been done by using ESAT-6 in both protein and DNA forms and by mixing the Tcell inducer antigen with Ag85B to form subunit vaccine (Olsen *et al.*, 2001).

Another example of introducing missing genes was done recently using Ag85B and it was found to promote a level of protection greater than conventional BCG (Horwitz *et al.*, 2000). DNA immunization for Ag85B had also been applied successfully with MPT64, although this gene was found to be absent from some BCG vaccines (Morris *et al.*, 2000).

The opposite approach, i.e. knocking out genes from *M. tuberculosis*, has also been suggested as a way to create attenuated live mycobacterial vaccines with all the relevant protective *M. tuberculosis* antigens. McKinney *et al.* designed an isocitrate lyase knockout strain which multiplies like wild-type strains during the initial phase of its infection but this is rapidly eliminated in the chronic phase of infection (McKinney *et al.*, 2000).

However, safety concerns with genetically modified live vaccines are still very significant problems before these vaccines could become a reality. So far, many of these modified strains have turned out be too virulent. Moreover, none of the evaluated modified live vaccines has shown superior efficacy to the BCG vaccines (McKinney *et al.*, 2000).

Promoting the Optimal Blend of the T-cell Subsets — Targeting the CD8 T-cell Subset

An efficient level of CD8 cells is required for tuberculosis infection control. Disability of BCG antigens to gain access to the cytoplasmic compartment of host cells leads to poor CD8 cells stimulation. Many researchers have worked on inventing efficient antigen delivery systems to facilitate MHC class I-restricted immune responses. The most successful vaccine which stimulates acceptable level of CD8 cells has been done by Hess *et al.* via constructing a recombinant BCG strain secreting a pore-forming sulfhydryl-activated cytolysin (listeriolysin) from *Listeria monocytogenes* which provides an escape route from the phagosome into the cytosol of infected host cells (Hess *et al.*, 1998). Although listeriolysin-secreting recombinant BCG did not release from phagosomal vacuoles, an increased MHC class I presentation of cophagocytosed ovalbumin was observed (Agger & Andersen, 2002).

DNA vaccines have advantage of stimulating strong CD8 T-cell responses with both secretion of IFN-γ and cytotoxic activities. Meanwhile, 38-kDa lipoglycoprotein of *M. tuberculosis* and Mtb 8.4 DNA vaccine were found to induce a prominent CD8 T-cell response whereas antibodies against these antigens were absent (Fonseca *et al.*, 2001).

Prime-boost vaccination is another strategy to improving vaccine immunogenicity by inducing MHC class I-restricted responses. These studies are based on initial priming by DNA vaccine, followed by a boost with the antigen expressed in a viral vector system such as vaccinia virus. The primary results of these prime-boost vaccination schemes include enhanced immunogenicity and protective levels equivalent to BCG vaccine (McShane *et al.*, 2001).

Environmental Mycobacteria Interaction with BCG—design of Vaccines to Overcome the Influence of Sensitisation by Environmental Mycobacteria

There are several factors that cause low efficacy of BCG vaccine; these include temperature, ultraviolet radiation, and prevalence of environmental mycobacteria in tropical regions. The most important factor which can affect the BCG activity is the presence of environmental mycobacteria. There are some explanations for these drawbacks that are caused by this environmental mycobacteria. Among other, the exposure to environmental mycobacteria may mask the protection of a subsequent BCG vaccination. In other words, BCG can be totally eliminated after vaccination. Some animal studies have suggested that infection with environmental mycobacteria changes the immune reaction towards a detrimental humoral response, which could not be overridden by a subsequent BCG vaccination (Agger & Andersen, 2002).

In a recent study, Brandt and colleagues have inoculated mice with mycobacteria isolated from soil the samples in Karonga, Malawi, whereby BCG vaccination has been demonstrated not to protect against TB in a number of studies. Only weak immune responses after BCG vaccination were induced and these could be attributed by the inhibition of the initial BCG multiplication (Brandt *et al.*, 2002).

This finding clearly demonstrates that subunit vaccines are not influenced by the sensitization and stimulation of protective T cell response, while these factors could affect initial multiplication of BCG. This result may support the use subunit vaccine or DNA vaccine, which can provide enough level of protection against tuberculosis.

Preventing the waning of BCG efficacy—the potential of booster vaccination

BCG vaccine promotes high levels of immunity against childhood TB. However, the prevalence of TB increases with time, which means waning of BCG protection. This finding is observed by monitoring the increased number of pulmonary cases of TB in adolescence, which is not prevented by a BCG revaccination strategy. As the majority of the human population is already BCG vaccinated, an alternative strategy would be by boosting this existing immune response. Ag85 antigen has been reported as a successful booster to avoid the waning of immune response after BCG vaccination in mice. As opposed to live vaccines such as BCG and based on few selected antigens, a subunit vaccine can probably induce strong immune responses, which are not influenced by previous mycobacterial exposure (Brooks *et al.*, 2001).

GENOMIC DIFFERENCES BETWEEN *M. TUBERCULOSIS* AND *M. BOVIS* BCG

Continuous *in vitro* passaging of BCG leads deletion of specific regions from the genome of BCG and causes over-attenuation of the vaccine (Mostowy *et al.*, 2004). Comparative genome analyses of the *M. tuberculosis* genome and *M. bovis* BCG have shown that 16 genomic regions of *M. tuberculosis* are deleted or lacking in some or all strains of *M. bovis* and/or *M. bovis* BCG (Brosch *et al.*, 2000). Meanwhile, the loss of genomic regions during *in vitro* passaging is believed to have deleted not only the virulence factors but also certain key protective antigens in BCG (Behr, 2002).

Immunological evaluation of the proteins encoded by these regions was predicted to identify the antigens of *M. tuberculosis* which are important for developing specific diagnostic reagents and vaccines to control TB (Brosch *et al.*, 2000; Mustafa, 2005). These DNA sequences, designated as the Regions of Difference (RD1–16), are known to encode many putative molecules relevant for designing improved diagnostic and prophylactic strategies (Andersen *et al.*, 2000).

A few well-characterized proteins like ESAT-6, CFP-10, CFP-21 and MPT-64 are encoded by RD1 and RD2 of *M. tuberculosis* genome. Meanwhile, the loss of RD1 has been implicated to be the primary deletion that led to the attenuation of *M. bovis* and generation of vaccine strain BCG. Therefore, using the missing immunodominant RD antigens seems to be a promising strategy to produce potential vaccine against this deadly disease, "tuberculosis".

REFERENCES

Agger, E. M., & Andersen, P. (2002). A novel TB vaccine towards a strategy based on our understanding of BCG failure. *Vaccine, 21,* 7-14.

Andersen, P., Munk, M. E., Pollock, J. M., & Doherty, T. M. (2000). Specific immune-based diagnosis of tuberculosis. *Lancet*, *356,* 1099–1104.

Bastos, R. G., & Borsuk, S. (2009). Recombinant Mycobacterium bovis BCG. *Vaccine*, *27*(47), 6495-6503.

Behr, M. A. (2002). BCG–different strains, different vaccines? *Lancet Infec., 2,* 86–92.

Brandt, L., Cunha, J. F., Olsen, A. W., Chilima, B., Hirsch, P., & Appelberg, R. (2002). Addressing the failure of the *Mycobacterium bovis* BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and the induction of protective immunity to tuberculosis. *Infect. Immun., 70*, 672–678.

Brooks, J. V., Frank, A., Keen, M. A., Belisle, J. T., & Orme, I. A. (2001). Boosting vaccine for tuberculosis, *Infect. Immun*., *69*, 2714–2717.

Brosch, R., Pym, A., Eiglmeier, K., Garnier, T., & Cole, S. T. (2000). Comparative genomics of the mycobacteria. *Int. J. Med. Microbiol*., *290,* 143-152.

Casadevall, A. (2003). Antibody-mediated immunity against intracellular pathogens: two-dimensional thinking comes full circle. *Infect Immun*; 2003; *8***:** 4225–4228.

David, A. H., & Amit, M. (2011). Aerosol vaccines for tuberculosis: A fine line between protection and pathology. Tuberculosis, 91, 82-85. Retrieved from http://ezproxy.um.edu.my:2095/science/article/pii/S1472979210001101 - aff2.

Delogu, G., Howard, A., Collins, F., & Morris, S. (2000). DNA vaccination against tuberculosis: expression of a ubiquirin-conjugared tuberculosis protein enhances antimycobacterial immunity. *Infect Immun*., *68,* 3097-3102.

Doherty, T. M., Demissie, A., Olobo, J., Wolday, D., Britton, S., & Eguale, T. (2002). Immune responses to the *Mycobacterium tuberculosis* - specific antigen ESAT-6 signal subclinical infection among contacts of tuberculosis patients. *J. Clin. Microbiol., 40*, 704–706.

Doytchinova, I. A., & Flower, D. R. (2001). Toward the quantitative prediction of T-cell epitopes: coMFA and coMSIA studies of peptides with affinity for the class I MHC molecule HLA. *J. Med. Chem., 44*, 3572–3581.

Edwards, K. M., Cynamon, M. H., Voladri, R. K., Hager, C. C., Destefano, M. S., & Tham, K. T. (2001). Iron-cofactored superoxide dismutase inhibits host responses to *Mycobacterium tuberculosis*. *Am. J. Respir. Crit. Care Med., 164*, 2213–2219.

Fine, P. E. M. (2001). BCG: the challenge continues. *Scan J. immunol., 4,* 243-245.

Fonseca, D. P. J., Benaissa-Trouw, B., Engelen, M. V., Kraaijeveld, C. A., Snippe, H., & Verheul, F. M. (2001). Induction of cell-mediated immunity against *Mycobacterium tuberculosis* using DNA vaccines encoding cytotoxic and helper T-cell epitopes of the 38-kilodalton protein. *Infect. Immun., 69*, 4839–4845.

Girard, M. P., Fruth, U., & Kieny, M. P. (2005). A review of vaccine research and development: Tuberculosis. *Vaccine*, *23,* 5725-5731.

Gupta, U. D., Katoch, V. M., & Mcmurray, D. N. (2007). Current status of TB vaccines. *Vaccine*, *25*, 3742-3751.

Gurunathn, S. C., Prussin, D. L., & Sacks-Seder, R. A. (1998). Vaccine requirments for sustain vaccine immunity to an intrcellular parasites infection. *Nat. Med., 4*, 1409-1415.

Helke, K. L., Mankowski, J. L., & Manabe, Y. C. (2006). Animal models of cavitation in pulmonary tuberculosis. *Tuberculosis, 86,* 337-348.

Hess, J., Miko, D., Catic, A., Lehmensiek, V., Russell, D. G., & Kaufmann, S. H. E. (1998). *Mycobacterium bovis* Bacille Calmette–Guérin strains secreting listeriolysin of Listeria monocytogenes. *Proc. Natl. Acad. Sci.* U.S.A., *95*, 5299–5304.

Horwitz, M. A., Harth, G., Dillon. B. J., & Maslesa-Galic, S. (2000). Recombinant bacillus Calmette–Guérin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc. Natl. Acad. Sci. U.S.A., 97,* 13853–13858.

Istvan, D., & Wolff, J. A. (1994). Direct gene transfer into muscle. *Vaccine, 12,* 1499-1502.

Kaerch, S., Herry, E., & Ahmed, R. (2002). Effector and memory T cells differentiation implication for vaccine development. *Nat. Rev. Immunal., 2*, 251-262.

Kaufmann, S. H. E. & Gengenbacher, M. (2012). Recombinant live vaccine candidates against tuberculosis, *Current Opinion in Biotechnology*.

Kamar, V., & Sercarz, E. (1996). Genetic vaccination the advantages of going naked. *Nat. Med., 2*, 857-859.

Kamath, T., Fruth, U., Brennan, M. J., Dobbelaer, R., & Hubrechts, P. (2005). New live mycobacterial vaccines: the Geneva consensus on essential steps towards clinical development. *Vaccine, 23,* 3753-3761.

Kamath, T., Hanke, T., Briscoe, H., & Britton, W. J. (1999). Co-immunization with DNA vaccines expressing granulocyte-macrophage colony-stimulating factor and mycobacterial secreted proteins enhances T-cell immunity, but not protective efficacy against *Mycobacterium tuberculosis*. *Immunology, 96*, 511-516.

Kaufman, S. H. (2001). How can immunology contribute to the control of tuberculosis? *Nat Rev. Immunol., 1*, 20-30.

Lowrie, D. B., Silva, C. L., Colston, M. J., Ragno, S., & Tascon, R. E. (1997). Protection against tuberculosis by a plasmid DNA vaccine. *Vaccine*, *15*(8)*,* 834-838.

Mckinney, J. D., Honer, Z., Bentrup, K., & Munoz-Elias, E. J. (2000). Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature*, *406*, 735–738.

Mcshane, H., Brookes, R., Gilbert, S. C., & Hill, V. S. (2001). Enhanced immunogenicity of CD4 T-cell responses and protective efficacy of a DNA-modified vaccinia virus Ankara prime-boost vaccination regimen for murine tuberculosis. *Infect. Immun., 69*, 681–686.

Mir, L. M. (2001). Therapeutic perspectives of *in vivo* cell electropermeabilization, *Bioelectrochemistery, 53*, 1-10.

Montgomery, D. L., Huygen, K., & Yawman, M. (1997). Induction of humoral and cellular immune responses by vaccination with *M. tuberculosis* antigen 85 DNA. *Cell. Mol. Biol., 43*, 285–292.

Morens, D. (2002). At the deathbed of consumptive art. *Emerg. Infect. Dis., 8*, 1353-1358.

Morris, S., Kelley, C., Howard, A., Li, Z., & Collins, F. (2000). The immunogenicity of single and combination DNA vaccines against tuberculosis. *Vaccine, 18,* 2155–2163.

Mostowy, S., Cleto, C., Sherman, D. R., & Behr, M. A. (2004). The *Mycobacterium tuberculosis* complex transcriptome of attenuation. *Tuberculosis*, *84*, 197-204.

Mustafa, A. S. (2005). Recombinant and synthetic peptides to identify *Mycobacterium tuberculosis* antigens and epitopes of diagnostic and vaccine relevance. *Tuberculosis, 85*, 367-376.

Nor, N. M., & Musa, M. (2004). Approaches towards the development of a vaccine against tuberculosis: recombinant BCG and DNA vaccine. *Tuberculosis, 84*, 102-109.

Okada, M., Kita, Y., Nakajima, K., Kaneda, Y., Saunderson, P., Tan, E. V., & McMurray, D. N. (2011). A Novel Therapeutic and Prophylactic Vaccine against Tuberculosis Using the Cynomolgus Monkey Model and Mouse Model. *Procedia in Vaccinology*, *4*, 42-49.

Olsen, A. W., Van P. L., Okkels, L. M., Rasmussen, P. B., & Andersen, P. (2001). Protection of mice with a tuberculosis subunit vaccine based on a fusion protein of Antigen 85B and ESAT-6. *Infect. Immun*., *69*, 2773–2778.

Orme, I. M. (2006). Preclinical testing of new vaccines for tuberculosis: A comprehensive review. *Vaccine, 24*, 2-19.

Philippe, G., Katherine, F., & Mario, R. (2009). Global Burden and Epidemiology of Tuberculosis. *Clin. Chest. Med*., *30*, 621-636.

Romain, F., Horn, C., Pescher, P., Namane, A., Riviere, M., & Puzo, G. (1999). Deglycosylation of the 45/47-kilodalton antigen complex of *Mycobacterium tuberculosis* decreases its capacity to elicit in vivo or in vitro cellular immune responses. *Infect. Immun*., *67*, 5567–5572.

Sable, S. B., Kalra, M., Indu, V., & Khuller, G. K. (2007a) Tuberculosis subunit vaccine design: The conflict of antigenicity and immunogenicity. *Clin. Immunol., 122*, 239-251.

Sable, S. B., Verma, I., & Khuller, G. K. (2005b). Multicomponent antituberculous subunit vaccine based on immunodominant antigens of *Mycobacterium tuberculosis*. *Vaccine*, *23,* 4175–4184.

Sambandamurthy, V. K., & Jacobs, J. W. R. (2005). Live attenuated mutants of *Mycobacterium tuberculosis* as candidate vaccines against tuberculosis. *Microbes Infect., 7*, 955-961.

Sanjay, G., Chang-Yu, W., Brenda, L. F., & Robert, A. S. (2000). DNA vaccines: a key for inducing long-term cellular immunity. *Curr. Opin. immunol., 12*, 442-447.

Simonney, N., Chavanet, P., Perronne, C., Leportier, M., Revol, F., & Herrmann, J-L. (2007). B-cell immune responses in HIV positive and HIV negative patients with tuberculosis evaluated with an ELISA using a glycolipid antigen. *Tuberculosis*, *87*, 109-122.

Skeiky, Y. A., Alderson, M. R., Ovendale, P. J., Guderian, J. A., Brandt, L., & Dillon, D. C. (2004). Differential immune responses and protective efficacy induced by components of a tuberculosis polyprotein vaccine, Mtb72F, delivered as naked DNA or recombinant protein. *J. Immunol*., *12,* 7618-7628.

Skjot, R. L., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., & Jacobsen, S. (2000). Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. *Infect. Immun., 68*, 214–220.

Stanford, J. L. (1991). Improving on BCG. *APMIS, 99,* 103–113.

Taylor, J. L. (2003). Novel vaccine strategies against aerosol infection with *Mycobacterium tuberculosis* in mice. Unpublished PhD Dissertation, Colorado State University, Colorado.

Tony, H., & Hassan, M. (2011). Prospects for a new, safer and more effective TB vaccine, *Paediatr. Res. Rev*., *12*, 46-51.

Von Der Lippe, B., Sandven, P., & Brubakk, O. (2006). Efficacy and safety of linezolid in multidrug resistant tuberculosis (MDR-TB): a report of ten cases. *Infection, 52,* 92-96.

Pertanika J. Sci. & Technol. 22 (2): 401 - 417 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Comparison of Physical Activity Prevalence among International Physical Activity Questionnaire (IPAQ), Steps/Day, and Accelerometer in a Sample of Government Employees in Kangar, Perlis, Malaysia

Hazizi, A. S.[1,3]*, Zahratul Nur, K.[1], Mohd Nasir, M. T.[1], Zaitun, Y.[1] and Tabata, I.[2]

[1]*Department of Nutrition and Dietetics, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

[2]*Faculty of Sport and Health Science, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan*

[3]*Sports Academy, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*

ABSTRACT

The aim of this study was to compare physical activity prevalence estimates among the International Physical Activity Questionnaire (IPAQ), steps/day, and accelerometer in a sample of government employees in Kangar, Perlis, Malaysia. Ten government agencies in Kangar were randomly chosen, and all employees were invited to participate. A self-administered questionnaire was employed to obtain information on socio-demographic characteristics and a physical activity assessment using the IPAQ. Anthropometric measurements, which include measurements of weight, height, body mass index, percent body fat, waist and hip circumference, were carried out. An accelerometer was used to assess total daily energy expenditure and the number of steps/day. A total of 272 respondents were involved in this study with a response rate of 83.2%. According to IPAQ, accelerometer and steps/day, the majority of the respondents (22.0%, 55.1%, and 77.6%, respectively) were classified as sedentary. The agreement between physical activity level as determined by the accelerometer vs. the IPAQ (Kappa=-0.46 {95% CI -0.384,-0.536}, p=0.238) and the IPAQ vs. steps/day (Kappa =0.037 {95% CI 0.090,-0.016}, p=0.175) was not significant, but the agreement between physical activity level as determined by the accelerometer vs. steps/day was classified as fair (Kappa=0.296 (95% CI 0.392, 0.200}, p<0.001). Our study highlights the need for a valid, accurate, and reliable self-report physical activity assessment tool for Malaysian adults.

Keywords: Prevalence, physical activity measurements, accelerometer, IPAQ, steps/day.

Article history:
Received: 3 January 2012
Accepted: 27 April 2012

Email addresses:
Hazizi, A. S. (hazizi@upm.edu.my), Zahratul Nur, K. (anakkalmi@yahoo.com), Mohd Nasir, M. T. (mnasirmt@upm.edu.my), Zaitun, Y. (zaitun@upm.edu.my), Tabata, I. (tabatai@fc.ritsumeikan.ac.jp)
*Corresponding Author

ISSN: 0128-7702

INTRODUCTION

Physical activity is defined as "any bodily movement produced by the contraction of skeletal muscle that increases energy expenditure above a basal level" (The U.S. Department of Health and Human Services, 2008). Physical activity has been shown to improve health, and the World Health Organization (WHO) recommends that adults aged 18–64 years engage in moderate intensity aerobic physical activity at least 150 minutes per week. Alternatively, they require at least 75 minutes of vigorous intensity aerobic physical activity per week or an equivalent combination of moderate to vigorous intensity activities (World Health Organization, 2011).

One can assess physical activity level using objective methods such as accelerometers, pedometers, and double labelled water, as well as subjective methods such as questionnaires or physical activity diaries. Epidemiological studies often use questionnaires simply because they are lower in cost, easy to administer, and present a smaller burden to respondents. Many studies have used accelerometers to validate data obtained through self-report. Craig *et al*. (2003) studied the use of The Computer Science Application (CSA) motion detectors (accelerometers) in 12 countries to validate the International Physical Activity Questionnaire (IPAQ). Bull *et al*. (2009) used an objective measure (a pedometer and an accelerometer) to validate the Global Physical Activity Questionnaire (GPAQ) over 7 days.

Self-report measures of physical activity usually rely on the participants' ability to recall, to be honest, and to estimate the time, frequency, and intensity of their activities over the past 7 days. Some findings from the self-report have proved to be overestimations, as compared to objective methods of physical activity assessment. Troiano *et al*. (2008) hypothesised that such overestimations might stem from respondents' misperceptions/misclassifying of activities (i.e., the type) and underestimations by accelerometers.

One of the self-report instruments used in physical activity assessments is IPAQ. According to Craig *et al*. (2003), IPAQ has reasonable measurement properties for monitoring a population's physical activity levels with a criterion validity median of about 0.30 among 18 to 65 year old adults from various countries. However, several studies have shown that IPAQ overestimates physical activities by 85% among Vietnamese, 100% among Americans, 170% among Hong Kongers (Lee *et al*., 2011), 165% among New Zealanders (Boon *et al*., 2008), and 247% among cancer patients (Johnson-Kozlow *et al*., 2008). Although IPAQ is regarded as a good measure of the physical activity levels among certain populations such as well-educated respondents, there are studies which have proven that it is less accurate when administered to the population of other groups.

To the authors' knowledge, no data have been published on the agreement between IPAQ and objective methods of physical activity assessment such as accelerometers among Malaysians in Malaysia, a developing country with an upper-middle income economy (The World Bank Group, 2011). Questionnaires such as IPAQ, GPAQ, and Physical Activity Questionnaire for Older Children (PAQ-C) have been used in population studies in Malaysia to determine the prevalence of physical activity (see for instance, Soo, Wan Abdul Manan & Wan Suriati, 20112; Siti Affira, Mohd Nasir, Hazizi, & Kandiah, 2011; Farah Wahida, Mohd Nasir, & Hazizi, 2011). The validity of these questionnaires has been established among developed nations. Due to the differences in terms of social, economic, and cultural backgrounds,

however, researchers need to study the comparability of the physical activity level estimates using IPAQ, steps/day, and accelerometers among Malaysians. Hence, the aim of this study is to compare physical activity level estimates using the IPAQ, steps/day, and accelerometers among a sample of government employees in Kangar, Perlis, Malaysia.

MATERIALS AND METHODS

This study employed a cross-sectional design involving employees from government agencies. A list of government agencies in Kangar, Perlis, Malaysia was obtained from the official website of the state government of Perlis. The sample size for this study was calculated based on the formula by World Health Organization (2008) and data from the National Health and Morbidity Survey III (Institute for Public Health, 2008). The minimum number of the respondents needed for this study was 96. Ten agencies were randomly chosen, and all the employees from the selected agencies were invited to participate in the study (n = 327). Ethical approval for this study was obtained from the Medical Research Ethics Committee of the Faculty of Medicine and Health Sciences, Universiti Putra Malaysia. All the respondents were briefed on the study using an information sheet and their written informed consent was obtained.

Inclusion criteria were employment in the selected government agencies in Kangar, Perlis, and age range of 18 to 65 years. Those who were pregnant or with physical disabilities (the ones that limited mobility such as use of a wheelchair, crutches, walking stick, arthritis, etc.) were excluded. Physical disabilities were identified based on the self-reported data obtained from the respondents.

Measures

Demographics and Self-reported Physical Activity

The questionnaire, which is written in the Malay language and self- administered, comprises of two sections. Section A contains socio-economic and demographic characteristics which include occupation, age, date of birth, sex, race, religion, marital status, educational level, monthly income, and number of household members. Section B includes an assessment of physical activity using IPAQ-Short in the Malay language downloaded from the IPAQ website (www.ipaq.ki.se). IPAQ has been validated in studies carried out in several countries (Craig *et al*., 2003). The classification of MET scores for the IPAQ and physical activity classification was based on Sjostrom *et al*. (2005). Total physical activity was calculated based on the total number of days, minutes and intensity of physical activity reported; these were then classified as ‘‘Low’’, ‘‘Moderate’’ or ‘‘High’’ levels of physical activity.

Anthropometric Measurements

Height was measured using a Body Meter (SECA 206, Body Meter, Germany) that was fixed on the wall. The respondents were asked to remove all footwear and head gear. Height was measured to the nearest 0.1 cm, with the respondents standing erect without shoes. Body weight and percentage of body fat were measured using a Tanita Body Composition Analyzer (Tanita

TBF-306, Body Composition Analyzer, Japan) with minimal clothing (and no shoes or socks). Weight was measured to the nearest 0.1 kg. The classification of % body fat was based on Lee and Nieman (2003). Body mass index (BMI) was calculated using a standard formula. The classification of BMI for adults was based on the guidelines by World Health Organization (1998). Waist circumference was obtained by measuring the distance around the smallest area below the rib cage and above the umbilicus (belly button) using a non-stretchable tape measure (World Health Organization, 2008). Waist circumference was measured in centimetres and classified as high risk if it is ≥ 90 cm for males or ≥ 80 cm for females (IOTF/ WHO/IASO, 2000). Respondents' blood pressure was measured by the researcher using Omron Blood Pressure Monitor (IA2, Japan) in a sitting position after 5 minutes resting. The American Heart Association classified hypertension as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90mmHg ratings on each of 2 or more office visits (Chobanian *et al*., 2003).

Accelerometer-Determined Physical Activity

The respondents were instructed to wear the accelerometer (Lifecorder, Suzuken, Japan) for 3 days, 2 weekdays and 1 weekend day. The accelerometer was attached vertically to the waistband of the clothing during waking hours, except when bathing or swimming. The respondents were encouraged not to alter their usual physical activity habits during the 3 days of measurement. The validity and reliability of the Lifecorder accelerometer have been discussed elsewhere (Total energy expenditure r=0.928 P<0.001 Kumahara *et al*., 2004; Total energy expenditure r=0.998 P<0.001 Heng & Hazizi, 2010). Steps/day was also determined by the accelerometer and classified based on Tudor-Locke and Bassett (2004) as follows: sedentary (<5000 steps/day), inactive (5000-7499 steps/day), somewhat active (7500-9999 steps/day), and active (≥ 10,000 steps/day). The accelerometer-determined physical activity level was calculated as the ratio of total energy expenditure (TEE) to basal metabolic rate (BMR). Meanwhile, physical activity level (PAL) was classified according to FAO/WHO/UNU (2004), where the respondents' activity levels were categorised as sedentary (PAL 1.40-1.69), active (PAL 1.70-1.99), or vigorous (PAL 2.00-2.40).

Data Analysis

Data analysis was performed using SPSS package for Windows, version 19. Descriptive statistics such as frequencies, means, percentages, and standard deviations, were used to describe variables like age, BMI, waist circumference, and physical activity level. Pearson's correlation coefficient and kappa statistics are presented in tables. The kappa statistics was used to determine the agreement between the methods used in the study. A statistical probability level of $p < 0.05$ was considered as significant.

RESULTS

The total number of employees from the 10 agencies chosen was 327. All 327 employees were invited to participate in the study, but only 272 agreed (83.2% response rate) to take part. Among the 272 participants, 55% (n=151) were males, and 32.4% (n=88) were between 41 and

50 years of age (mean age = 39±11 years). Table 1 shows the distribution of the respondents by their socio-demographic characteristics. Most of the respondents (46.7%) earned around RM1500-RM2500 monthly. The mean (±s.d.) income was RM1964.98 ± RM986.36 (1USD = RM3.2).

TABLE 1
Distribution of the respondents by selected socio-demographic characteristics

Characteristics	Male (n=151)	Female (n=121)	Total (n=272)
	n (%)	n (%)	n (%)
Sex	151 (55.5)	121 (44.5)	272 (100.0)
Age (years)			
19-30	39 (25.8)	41 (33.9)	80 (29.4)
31-40	29 (19.2)	44 (36.4)	73 (26.8)
41-50	56 (37.1)	32 (26.4)	88 (32.4)
51-62	27 (17.9)	4 (3.3)	31 (11.4)
Income (RM)			
< 1500	43 (28.5)	51 (42.1)	94 (34.6)
1500- 2500	66 (43.6)	61 (50.4)	127 (46.6)
2501- 3500	17 (11.3)	6(5.0)	23 (8.5)
> 3500	25 (16.6)	3 (2.5)	28 (10.3)
Occupation			
Professional	31 (20.5)	4 (3.3)	35 (12.8)
Administrative/clerical	67 (44.4)	110 (90.9)	177 (65.1)
General Assistant	53 (35.1)	7(5.8)	60 (22.1)
Marital Status			
Single	26 (17.2)	22 (18.2)	48 (17.7)
Married	122 (80.8)	94 (77.7)	216 (79.4)
Divorced/ Widowed	3 (2.0)	5 (4.1)	8 (2.9)
Education Level			
Primary School	6 (4.0)	1 (0.8)	7 (2.5)
SRP/ PMR/ LCE	29 (19.2)	15 (12.4)	44 (16.2)
SPM/ MCE/ O Level	36 (23.8)	41 (33.9)	77 (28.3)
STPM/ HSC/ A Level	20 (13.2)	30 (24.8)	50 (18.4)
Diploma/ Degree	60 (39.8)	34 (28.1)	94 (34.6)

SRP - Sijil Pelajaran Malaysia
PMR - Penilaian Menengah Rendah
LCE - Malaysia Lower Certificate of Education
SPM - Sijil Pelajaran Malaysia
MCE - Malaysian Certificate of Education
O Level - Ordinary Level General Certificate of Education (GCE)
STPM - Sijil Tinggi Pelajaran Malaysia
HSE - Higher Certificate of Education
A Level - Advanced Level General Certificate of Education (GCE)

In term of occupation, most of the respondents 65.1% (n=171) worked in administrative or clerical positions such as administration assistant. Only 12.8% (n=35) worked in professional positions. Among the respondents, 79.4% (n=216) were married, with 34.6% (n=94) holding either diploma or degree qualifications.

Table 2 shows the distribution of the respondents by BMI, waist circumference, waist-hip ratio, percentage of body fat and blood pressure. More respondents were classified as overweight (37.5%) than having a normal weight (34.9%). The overweight group contained more male (42.4%) than female respondents (31.4%), while 21.7% of the respondents were classified as obese. Only 5.9% of them were underweight. Meanwhile, 57.0% and 56.2% of the male and female respondents were classified as at risk based on their waist circumferences.

TABLE 2
Distribution of the respondents by BMI, waist circumference, waist-hip ratio, percentage of body fat and blood pressure

Indicators	Cut-off point	Male (n=151)	Female (n=121)	Total (n=272)
		n (%)	n (%)	n (%)
BMI (kg/m²)				
Underweight	<18.5	4 (2.7)	12 (9.9)	16 (5.9)
Normal	18.5–24.9	57(37.7)	38 (31.4)	95 (34.9)
Overweight	25.0–29.9	64 (42.4)	38 (31.4)	102 (37.5)
Obese	30.0–>40.0	26 (17.2)	33 (27.3)	59 (21.7)
Waist circumference (cm)				
Acceptable	Male:<90 Female:<80	65 (43.0)	53 (43.8)	118 (43.4)
At risk	Male:≥90 Female: ≥80	86 (57.0)	68 (56.2)	154 (56.6)
Waist-hip ratio				
Acceptable	Male: <0.9 Female:<0.8	70 (46.4)	58 (47.9)	128 (47.1)
At risk	Male: ≥0.9 Female: ≥0.8	81 (53.6)	63 (52.1)	144 (52.9)
Fat Percentage (%)				
Acceptable	Male: 6-24 Female: 9-31	58 (38.4)	42 (34.7)	100 (36.8)
Unhealthy	Male: ≥25 Female: ≥32	93 (61.6)	79 (65.3)	172 (63.2)
Systolic blood pressure (mmHg)				
Normal	<140	69 (45.7)	77 (63.6)	146 (53.7)
Elevated	≥140	82 (54.3)	44 (36.4)	126 (46.3)
Diastolic blood pressure (mmHg)				
Normal	<90	106 (70.2)	87 (71.9)	193 (71.0)
Elevated	≥90	45 (29.8)	34 (28.1)	79 (29.0)

In terms of waist-hip ratio, more male respondents (53.6%) were categorised as at risk than the female respondents (52.1%), although the difference between the genders was not significant. For the percentage of body fat, most respondents were classified as having an unhealthy percentage of body fat: 61.6% and 65.3% for the male and female respondents, respectively. The percentage of the respondents with elevated systolic blood pressure was also higher in males as compared to females. However, the percentage of elevated diastolic blood pressure was almost similar between both genders.

Table 3 shows the distribution of the respondents' physical activity level as measured by IPAQ, accelerometer, and steps/day. Using IPAQ, most of the respondents (47.7% of males and 66.9% of females) were found to having a moderate level of physical activity, whereas only 22.1% and 21.7% were categorised as having low and high levels of physical activity, respectively. Using the accelerometer, most of the respondents (55.1%) were classified as sedentary, and only 8.0% were classified as having a vigorous level of physical activity. Similarly, based on the number of steps/day, most of the respondents (77.6%) were classified as sedentary and low active. Only 7.7% were classified as leading active lifestyles.

TABLE 3
Distribution of the respondents by gender and physical activity level as assessed by accelerometer, IPAQ, and steps/day

Instruments	Cut off point	Male (n=151)	Female (n=121)	Total (n=272)
		n (%)	n (%)	n (%)
Accelerometer-determined physical activity level				
Sedentary	1.40-1.69	75 (49.7)	75 (62.0)	150 (55.1)
Active & Moderately active	1.70-1.99	61 (40.4)	39 (32.2)	100 (36.8)
Vigorous	2.00-2.40	15 (9.9)	7 (5.8)	22 (8.1)
IPAQ				
Low	<600 MET-min/week	26 (17.2)	34 (28.1)	60 (22.0)
Moderate	600-2999 MET-min/week	72 (47.7)	81 (66.9)	153 (56.3)
High	≥3000 MET-min/week	53 (35.1)	6(5.0)	59 (21.7)
Steps/day				
Sedentary & Low active	<7500 steps	105 (69.5)	106 (87.6)	211 (77.6)
Somewhat Active	≥7500-9999 steps	27 (17.9)	13 (10.7)	40 (14.7)
Active	≥10000 steps	19 (12.6)	2 (1.7)	21 (7.7)

Table 4 shows the distribution of physical activity levels as measured by steps/day, and IPAQ according to accelerometer measured PAL. The agreement between the activity levels, as determined by the accelerometer vs. IPAQ and IPAQ vs. steps/day, was not significant but the agreement between the activity levels as determined by the accelerometer vs. steps/day was significant and classified as fair (Kappa=0.296, $p<0.001$). The percentage of agreement between steps/day and accelerometer-determined physical activity level was 64% and those between physical activity level as determined by the accelerometer vs. IPAQ, and IPAQ vs. steps/day were around 30%.

TABLE 4
Distribution of physical activity levels as measured by steps/day and IPAQ according to accelerometer-measured PAL

	Accelerometer-determined physical activity level					
	Sedentary	Active or moderately active	Vigorous	Total	Kappa	p
	n (%)	n (%)	n (%)	n (%)		
Steps/Day						
Sedentary & low active	136 (50.0)	69 (25.4)	6 (2.2)	211 (77.6)	0.296	0.00
Somewhat Active	12 (4.4)	25 (9.2)	3 (1.1)	40 (14.7)		
Active	2 (0.7)	6 (2.2)	13 (4.8)	21 (7.7)		
Total	150 (55.1)	100 (36.8)	22 (8.1)	272 (100.0)		
IPAQ						
Low	29 (10.7)	25 (9.2)	6 (2.2)	60 (22.0)		
Moderate	90 (33.1)	52 (19.1)	11 (4.0)	153 (56.3)	0.460	0.24
High	31 (11.4)	23 (8.5)	5(1.8)	59 (21.7)		
Total	150 (55.1)	100 (36.8)	22 (8.1)	272 (100.0)		
	IPAQ					
	Low	Moderate	High	Total		
Steps/Day[3]						
Sedentary & Low active	49 (18)	120 (44.1)	42 (15.4)	211 (77.6)		
Somewhat Active	5 (1.8)	25 (9.2)	10 (3.7)	40 (14.7)	0.037	0.18
Active	6 (2.2)	8 (2.9)	7 (2.6)	21 (7.7)		
Total	60 (22.0)	153 (56.3)	59 (21.7)	272 (100.0)		

The percentage of agreement and Kappa statistics were further calculated and compared across two factors: indices of obesity and socio-demographic factors (see Table 5). Indices of obesity include BMI, waist circumference, and percentage of body fat, whereas socio-demographic factors were sex, age, and educational level. The agreement between physical activity as determined by steps/day vs. the accelerometer was significant across all indicators of obesity and socio-demographic factors, with the percentage of agreement ranging from 76.2% to 43%. The percentage of agreement was higher among the overweight/obese (vs. those with a normal weight), those with abdominal obesity (vs. those with a normal waist circumference), those with a higher percent body weight (vs. those with a lower percent body weight), females vs. males, and younger respondents (vs. older respondents). The agreement between physical activity levels as determined by the accelerometer vs. IPAQ and IPAQ vs. steps/day was statistically insignificant across all the indicators of obesity and socio-demographic characteristics.

TABLE 5

Kappa statistics, p value, and percentage of agreement between physical activity levels as measured using an accelerometer vs. the IPAQ, IPAQ vs. steps/day, and an accelerometer vs. steps/day and selected obesity indices and socio-demographic characteristics.

Characteristics	Accelerometer vs. Steps/day {Kappa (95% CI); p value; % of agreement}	Accelerometer vs. IPAQ {Kappa (95% CI); p value; % of agreement}	IPAQ vs. Steps/day {Kappa (95% CI); p value; % of agreement}
Sex			
Male	K=0.326 (0.451, 0.201); p=0.00; 55.0%	K=-0.080 (0.014, -0.174); p=0.101; 25.8%	K=0.012 (0.088, -0.064); p=0.756; 25.8%
Female	K=0.216 (0.361, 0.071); p=0.001; 66.9%	K=-0.007 (0.115, -0.129); p=0.909; 38.8%	K=0.041 (0.108, -0.026); p0.272; 34.7%
Education			
Lower	K=0.338 (0.477, 0.199);p=0.00; 63.3%	K=0.022 (0.136, -0.092); p=0.699; 37.5%	K=0.036 (0.124, -0.052); p=0.419; 26.4%
Upper	K=0.239 (0.372, 0.106); p=0.000; 64.6%	K=-0.110 (-0.006, -0.214); p=0.037; 26.4%	K=0.030 (0.095, -0.035);p=0.359; 29.2%
Age			
≤40 years	K=0.315 (0.462, 0.168); p=0.00; 70.6%	K=-0.005 (0.089, -0.099); p=0.916; 31.4%	K=0.021 (0.094, -0.052); p=0.555; 28.8%
>40 years	K=0.274 (0.394, 0.154); p=0.00 ; 55.7%	K=-0.097 (0.021, -0.215); p=0.127; 31.9%	K=0.057 (0.139, -0.025); p=0.175; 31.1%
Physical Activity Level	K =0.296 (0.392, 0.200), p=0.000; 72.8%	K=-0.46 (-0.384, -0.536); p=0.238; 31.6%	K=0.037 (0.090, -0.016), p=0.175; 29,8%
BMI			
Normal weight	K=0.197 (0.301, 0.093); p=0.00; 46.8%	K=-0.09 (0.124, -0.142); p=0.195; 34.2%	K=0.021 (0.109, -0.067); p=0.618; 38.7%
Overweight/Obese	K=0.417 (0.568, 0.266); p=0.00; 74.3%	K=0.001 (0.085, -0.083);p= 0.974; 30.4%	K=0.045 (0.116, -0.026); p=0.237; 29.7%
Waist Circumference			
Acceptable	K=0.204 (0.318, 0.090); p=0.00; 49.2%	K=-0.036 (0.091, -0.163); p=0.584; 36.4%	K=0.007 (0.097, -0.083); p=0.814; 28.8%
At risk	K=0.404 (0.555, 0.253); p=0.00; 75.3%	K=-0.037 (0.047, -0.121); p=0.393; 27.9%	K=0.059 (0.124, -0.006); p=0.093; 30.5%
Percent Body Weight			
Acceptable	K=0.178 (0.290, 0.066);p=0.00; 43%	K=-0.116 (0.019, -0.251); p=0.11; 32%	K=0.004 (0.102, -0.094); p=0.925; 29%
High	K=0.413 (0.554, 0.272); p=0.00; 76.2%	K=0.014 (0.090, -0.062); p=0.734; 31.4%	K=0.055 (0.118,-0.008); p=0.104; 30.2%

Correlation tests were also performed (see Table 6). The results indicated that the accelerometer-determined physical activity was significantly associated with that determined by steps/day (r=0.354, p=0.00) but not significantly associated with that determined by IPAQ (p=0.531). However, the physical activity level as determined by IPAQ was correlated significantly with that determined by steps/day (r=0.131, p=0.00). The relationship between accelerometer-determined physical activity level and indices of obesity such as waist circumference, hip circumference, percentage of body fat, and BMI was stronger than the correlation observed between the physical activity level as determined by IPAQ and steps/day with the same indicators. In terms of indices of obesity and blood pressure, the physical activity level determined by IPAQ was significantly correlated with waist circumference, waist-hip ratio, and percentage of body fat. Physical activity level, as determined by steps/day, was significantly correlated with percentage of body fat, but physical activity level determined by the accelerometer was significantly associated with almost all indicators of obesity and blood pressure.

TABLE 6
Correlation matrices of relationship between IPAQ, accelerometer determined physical activity, steps/day, indices of obesity, blood pressure and selected indicators of socio-demographic characteristics

	IPAQ	Steps/day	Accelerometer
IPAQ	1*	0.131*	
Steps/Day	0.131*	1*	0.354*
Accelerometer		0.354*	1*
Age			0.213*
Income		-0.139*	
Education		-0.180*	
BMI			-0.580*
WHR	0.163*		
Waist Circumference	0.170*		-0.417*
Hip Circumference			-0.528*
Percent body fat	-0.145*	-0.147*	-0.537*
Systolic blood pressure			
Diastolic blood pressure			-0.185*

*p<0.05

DISCUSSION

The agreement between the subjective methods of physical activity assessment such as IPAQ and objective methods (use of accelerometers, pedometers, and double labelled water) has been reported in many studies (Craig *et al*., 2003; Bull *et al*., 2009; Boon *et al*., 2010; Maddison *et al*., 2007). However, the majority of these studies were carried out among the Western populations. Some studies were carried out among the Asian populations such as in Hong Kong (Lee *et al*., 2011), Singapore (Nang *et al*., 2011) and Vietnam (Lachat *et al*., 2008), but not among Malaysians.

According to the World Bank Group (2011), the gross national income (GNI) for Malaysia was USD7769, while the adult literacy rate (percentage of literate people aged 15 and above) was 92%. In this study, all of the respondents worked in government agencies, almost half of them received a monthly income of RM1500-2500 (or USD6000-10000 per year), and 53% had received at least 18 years of formal education. The prevalence of overweight (37.5%), obesity (21.7%), and abdominal obesity (56.5%) was high in this study. According to the National Health and Morbidity Survey III, the prevalence of overweight and obesity among adults 18 years and above in Malaysia was 29.1% and 14.01%, respectively (Institute for Public Health, 2008).

Previous studies reported that among a sample of respondents in Universiti Putra Malaysia, the prevalence of them who were overweight was 31.9% for males and 26.5% for females, while the total prevalence of obesity was 16.1% (Siew *et al.*, 2010). Meanwhile, a study among a group of security guards and their spouses at the University of Malaya, Kuala Lumpur, also showed a high prevalence (64%) of overweight and obese respondents (Moy & Atiya, 2003).

Based on IPAQ classification, 22% of the respondents were engaged in low level of physical activity. This prevalence of low physical activity is apparently lower than the national prevalence, which is 43.7% in 2006 (Institute for Public Health, 2008). The difference might be due to differences in the study populations, since our study focused on workers aged 18-65 years in a specific state, whereas in the National Health and Morbidity Survey III, the figure reported covered the entire population of Malaysia aged 18 and above. However, the prevalence of low levels of physical activity in the national study was lower compared to the data gathered using the accelerometer (55.1%) and steps/day (77.6%). These might be due to the differences in the instrument used for assessing the respondents' level of physical activity.

In Malaysia, the prevalence of physical inactivity, as measured using a questionnaire based on occupation, was higher among the unemployed group (60.8%), housewives (54.5%), craft and clerical workers (47.3%), senior officials and managers (46.3%), and professionals (46.2%) (Institute of Public Health, 2008). Among a sample of working women in Petaling Jaya, 28.8% of them were found to have low physical activity levels (Siti Affira *et al*., 2011). A low prevalence of adequate exercises (13.8%) was also shown amongst a group of security guards and their spouses at the University of Malaya, Kuala Lumpur (Moy & Atiya, 2003).

The prevalence of physical inactivity by state in Malaysia was higher (more than 50%) in Selangor and Kuala Lumpur, but lowest in Pahang (31.4%) and Terengganu (32.3%), (Institute of Public Health, 2008). On the other hand, among the respondents in Greece, Pitsavos *et al.* (2005) reported physically active people as compared to sedentary among those in higher occupation skills and were more likely to live in rural areas ($p < 0.05$). Another study among adults aged 45 to 68 years in France revealed that subjects aged ≥60 years and women with higher education levels or living in rural areas as more likely to be meeting the recommended physical activity levels (Bertrais *et al.*, 2004).

Accelerometers and pedometers have been used in many studies to validate the physical activity assessment techniques using questionnaires such as in the study by Craig *et al*. (2003) where accelerometer was used to access the validity of IPAQ and the work of Bull *et al*. (2009) where a pedometer was utilized to validate GPAQ. Both the objective techniques used in these studies were compared against the questionnaires data to measure the physical activity

(pedometer vs. GPAQ pooled $\rho = 0.31$, $p < 0.01$; IPAQ vs. accelerometer pooled $\rho = 0.33$, 95% CI 0.26–0.39). In our study, the prevalence of moderate physical activity varied according to the assessment techniques used. The prevalence of moderate physical activity was lowest if it was assessed based on steps/day (14.7%) as compared to the physical activity assessments using the accelerometer (36.8%) and IPAQ (56.3%).

Based on some previous studies, the correlation between physical activity level assessments using IPAQ and accelerometer varied across populations. The correlations between these two techniques ranged between -0.12 and 0.57 in a study by Craig *et al*. (2003), indicating variability in the validity of these instruments across populations. The study had a pool correlation of 0.30. The researchers concluded that IPAQ is a valid tool for assessing physical activity levels in surveillance studies. However, in a study by Johnson-Kozlow *et al*. (2006), IPAQ overestimated physical activity levels by up to 247% (as compared to physical activity levels determined using the accelerometer).

In our study, the percentage of agreement between physical activity levels determined by using the accelerometer vs. IPAQ and IPAQ vs. steps/day was around 30%. However, this agreement was not significant based on Kappa statistics. This result supports the conclusions of a recent study by Grimm *et al*. (2011) who found that the overall percentage of agreement was 44.8%. In contrast, a study by Lachart *et al*. (2008) showed that the agreement between physical activity levels, as determined by IPAQ and the accelerometer, was not significant among rural respondents (percentage of agreement = 32.9%; kappa = 0.00; $p = 0.51$) but it was significant among urban respondents (percentage of agreement = 47.9%; kappa = 0.22; $p < 0.0015$), which might be due to the differences in the respondents' socio-demographic characteristics. These results showed that physical activity levels could not be estimated accurately using this questionnaire for all populations.

As reported by Ainsworth *et al*. (2006), the disagreement between physical activity assessments using different questionnaire methods might be due to the differences in understanding the questionnaire among different demographic groups. The differences might also be due to people's tendency to over-report when a self-report questionnaire was used (Rzewnicki, Auweele, & Bourdeaudhuij, 2003). Furthermore, the classification of physical activity levels as sedentary, moderate, or vigorous is rather subjective. This classification may be interpreted differently by the participants based on certain factors such as age, educational level, and other environmental factors. For example, vigorous activity may be interpreted differently by younger groups as compared to the elderly and also people living in rural areas who are used to working and using a lot of energy such as those in the agriculture sector, who may interpret vigorous activity differently than urban executives.

Our study focused on the workers at government agencies in one Malaysian state. Therefore, in terms of socio-demographic characteristics such as educational and income level, our study can not be generalized to the whole Malaysian population. Differences in scoring protocols may also contribute to the differences in the physical activity levels observed in this study. For example, steps/day is an assessment of physical activity levels using walking only. In contrast, IPAQ and accelerometer take into account other activities as well. Nonetheless, accelerometer is often used as a point of comparison for self-report questionnaires such as in the studies of Craig *et al*. (2003), Bull *et al.* (2009), and Grimm *et al.* (2011).

As stated in other studies, accelerometer contains several inherent limitations that contribute to the underestimation of true physical activity levels. For example, accelerometer cannot measure upper body activities such as walking with a load and lifting, which may contribute to inaccuracy in its assessment of level of physical activity (Welk, 2002). Additionally, accelerometers cannot measure activities such as swimming because accelerometers are not waterproof. Further, accelerometers will not accurately measure activities that lack significant acceleration of the hip (e.g. cycling). On the other hand, as many researchers have shown, studies tend to overestimate physical activity levels when questionnaires are used in the assessment (Lee *et al*., 2011).

The results of this study are in line with those of a prospective birth cohort study in Pelotas, Brazil, which showed that the accelerometer-determined physical activity was longitudinally and inversely associated with diastolic blood pressure. Self-reported physical activity was not related to blood pressure and both methods of physical activity assessments were unrelated to systolic blood pressure (Hallal *et al*., 2011). However, as reported by Hedayati, Elsayed and Reiley (2011) in their review paper, there are still inconsistent results on the effects of exercises on blood pressure.

Our findings showed that IPAQ, steps per day, and accelerometer were significantly and negatively correlated with the percentage of body fat. Other studies have also reported significant associations between physical activity measures by questionnaire and accelerometer (questionnaire = -1.93, $p < 0.001$; accelerometer = -1.06, $p = .001$) with the percentage of body fat (Hearst *et al*., 2012). However, the results of the prospective study between objectively measured physical activity and fat mass suggest no association between these two variables (Wilks *et al*., 2011).

The correlation between energy expenditure as measured by IPAQ and body weight indices in this study was significant for waist circumference and percentage of body fat. However, the correlations were found to be stronger between the indices of obesity and accelerometer-determined physical activity. More parameters were significantly correlated to accelerometer-determined physical activity than to assessments using IPAQ or steps/day. The relationship between body weight indices and physical activity levels is still controversial (Cook & Schoeller, 2011). The results of this study showed that accelerometers were better correlated to physical activity levels than the questionnaire data.

Meanwhile, the relationship between energy expenditure, as measured by IPAQ and accelerometer-determined physical activity, was not significant in this study. The relationship between physical activity level, as determined by steps/day and IPAQ, was weak but significant. Other studies comparing objective method vs. IPAQ produced mixed results. Lee *et al*. (2011), Lachat *et al*. (2008), and Nang *et al*. (2011) showed a weak relationship [correlation (r) of 0.10-0.37] between IPAQ and accelerometer-determined physical activity but other studies (Hagstromer *et al*., 2006; Craig *et al*., 2003) showed stronger associations [correlation (r) of 0.55-0.67].

One must interpret the results of this study carefully. The measurement using IPAQ was done before using the accelerometer. Therefore, the two assessment methods were not used simultaneously to measure the same period and thus, the same components of physical activity. The measurement using the accelerometer was done based on 2-week days and one-

weekend day, whereas the IPAQ measurement used a 7-day recall technique. Trost *et al*. (2005) suggested that a 3- to 5-day period is adequate for assessing adults' physical activity using an accelerometer. A 7-day recall technique used with IPAQ is common for measuring habitual physical activity using questionnaires. Both of the above techniques are usually used to describe habitual physical activity levels for a population within a stated timeframe. Other researchers such as Lee *et al*. (2011) also confirmed the validity of using IPAQ and accelerometer during non-concurrent periods to determine physical activity levels.

The agreement of the methods used for physical activity assessments among the participants in Malaysia should be studied further. Any result would require a careful interpretation. Additionally, such studies should use concurrent assessments of physical activities using the same instruments utilized in this study and a larger number of respondents from a wider variety of socioeconomic backgrounds.

CONCLUSION

Based on the physical activity levels obtained using the accelerometer and steps/day, the majority of the respondents in our study were sedentary. However, the prevalence of sedentary lifestyles was lower when activity levels were assessed using IPAQ. The agreement between physical activity level as determined by the accelerometer vs. IPAQ and IPAQ vs. steps/day was low and not significant.

ACKNOWLEDGEMENTS

This study was funded by the Ministry of Science, Technology and Innovation, Malaysia (Grant No: 5450422) and the National Institute of Health and Nutrition Japan.

REFERENCES

Bertrais, S., Preziosi, P., Mennen, L., Galan, P., Hercberg, S., & Oppert, J.M. (2004). Sociodemographic and Geographic Correlates of Meeting Current Recommendations for Physical Activity in Middle-Aged French Adults: the Supplémentation en Vitamines et Minéraux Antioxydants (SUVIMAX) Study. *American Journal of Public Health, 94*, 1560–1566.

Boon, R. M., Hamlin, M. J., Steel, G. D., & Ross, J. J. (2010). Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with Accelerometry. *British Journal of Sports and Medicine, 44,* 741–746.

Bull, F. C., Maslin, T. S., & Armstrong, T. (2009). Global Physical Activity Questionnaire (GPAQ): Nine Country Reliability and Validity Study. *Journal of Physical Activity and Health, 6,* 790-804.

Chobanian, A. V., Bakris, G. L., Black, H. R., Cushman, W. C., Green, L. A., Izzo Jr, J. L., Jones, D. W., Materson, B. J., Oparil, S., Wright Jr, J. T., & Roccella, E. J. (2003). Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension 42,* 1206-1252.

Cook, C. M., & Schoeller, D. A. (2011). Physical activity and weight control: conflicting findings. *Current Opinion in Clinical Nutrition and Metabolic Care, 14*, 419–424.

Craig, C. L., Marshall, A. L., Sjöström M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja P. (2003). International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381-1395.

FAO/WHO/UNU. (2004). Human energy requirements: Energy Requirement of Adults. *Report of a Joint FAO/WHO/UNU Expert Consultation*. Retrieved November 1, 2011 from http://www.fao.org/docrep/007/y5686e/y5686e07.htm.

Farah Wahida, Z., Mohd Nasir, M. T., & Hazizi, A. S. (2011). Physical Activity, Eating Behaviour and Body Image Perception among Young Adolescents in Kuantan, Pahang, Malaysia. *Malaysian Journal of Nutrition, 17*(3), 325-336.

Grimm, E. K., Swartz, A. M., Hart, T., Miller, N. E., & Strath, S. J. (2011). Comparison of the IPAQ-Short Form and Accelerometry Predictions of Physical Activity in Older Adults. *Journal of Aging and Physical Activity,* 1-16.

Hagstromer, M., Oja, P., & Sjostro, M. (2006). The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutrition, 9*(6), 755–762.

Hallal, P. C., Dumith, S. C., Reichert, F. F., Menezes, A. M., Araújo, C. L., Wells, J. C., Ekelund, U., & Victora, C. G. (2011). Cross-sectional and longitudinal associations between physical activity and blood pressure in adolescence: birth cohort study. *Journal of Physical Activity and Health, 8*(4), 468-474.

Hearst, M. O., Sirard, J. R., Lytle, L., Dengel, D. R., & Berrigan, D. (2012). Comparison of 3 measures of physical activity and associations with blood pressure, HDL, and body composition in a sample of adolescents. *Journal of Physical Activity and Health,* 9(1), 78-85.

Hedayati, S. S., Elsayed, E. F., & Reilly, R. F. (2011). Non-pharmacological aspects of blood pressure management: what are the data? *Kidney International, 79*(10), 1061–1070. doi:10.1038/ki.2011.46.

Heng, K. S., & Hazizi, A. S. (2010). Validation of Kenz Lifecorder e-Steps Accelerometer in Assessing Steps Count and Energy Expenditure. Abstract of Paper Presented at 25th Scientific Conference and Annual General Meeting of the Nutrition Society Malaysia, 25-26 March 2010, Kuala Lumpur.

Institute for Public Health (2008). *The Third National Health and Morbidity Survey (NHMS III) 2006*. Ministry of Health Malaysia.

International Obesity Task Force/World Health Organization/ International Association for Study of Obesity. (IOTF/WHO/IASO) (2000). *The Asia Pacific Perspective: Redefining Obesity and its Treatment*. WHO, Hong Kong.

Johnson-Kozlow, M., Sallis, J. F., Gilpin, E. A., Rock, C. L., & Pierce, J. P. (2006). Comparative validation of the IPAQ and the 7-Day PAR among women diagnosed with breast cancer. *International Journal of Behavioral Nutrition and Physical Activity*, *3*(7).

Kumahara, H., Schtz, Y., Ayabe, M., Yoshioka, M., Yoshitake, Y., Shindo, M., Ishii, K., & Tanaka, H. (2004). The use of uniaxial accelerometry for the assessment of physical-activity-ralated energy expenditure: a validation study against whole-body indirect calorimetry. *British Journal of Nutrition*, *91*, 235-243.

Lachat, C. K., Verstraeten, R., Khanh, L. N. B., Hagströmer, M., Khan, N. C., Van, N. D. A., Dung N. Q., & Kolsteren, P. W. (2008). Validity of two physical activity questionnaires (IPAQ and PAQA) for Vietnamese adolescents in rural and urban areas *International Journal of Behavioral Nutrition and Physical Activity*, *5*, 37.

Lee, D. C., & Nieman R. D. (2003). *Nutritional assessment*. McGraw-Hill.

Lee, P. H., Yu, Y. Y., McDowell, I., Leung, G. M., Lam, T. H., & Stewart, S. M. (2011). Performance of the international physical activity questionnaire (short form) in subgroups of the Hong Kong chinese population. *International Journal of Behavioral Nutrition and Physical Activity*, *8*, 81.

Luke, A., Dugas, L. R., Durazo-Arvizu, R. A., Cao, G., & Cooper, R. S. (2011) Assessing Physical Activity and its Relationship to Cardiovascular Risk Factors: NHANES 2003-2006. *BMC Public Health, 11*, 387.

Maddison, R., Mhurchu, C. I., Jiang, Y., Hoorn, S. V., Rodgers,A., Lawes, C. M. M., & Rush, E. (2007). International Physical Activity Questionnaire (IPAQ) and New Zealand Physical Activity Questionnaire (NZPAQ): A doubly labelled water validation. *International Journal of Behavior Nutrition and Physical Activity*, *4*, 62.

Ministry of Health. (2006). *Malaysia NCD Surveillance-1 2005-2006 NCD Risk Factors in Malaysia*. Non-Communicable Disease Section, Disease Control Division, Ministry of Health Malaysia.

Moy, F. M., & Atiya, A. S. (2003). Lifestyle practices and prevalence of obesity in a community within a university campus. *Journal of University Malaya Medical Centre, 8*, 33–38.

Nang, E. E. K., Ngunjiri, S. A. G., Wu, Y., Salim, A., Tai, E. S., Lee, J., & Van Dam R. M. (2011). Validity of the international physical activity questionnaire and the Singapore prospective study program physical activity questionnaire in a multiethnic urban Asian population. *BMC Medical Research Methodology*, *11*, 141.

Park, J., Ishikawa-Takata, K., Tanaka, S., Hikihara, Y., Ohkawara, K., Watanabe, S., Miyachi, M., Morita, A., Aiba, N., & Tabata, I. (2011). Relation of body composition to daily physical activity in free-living Japanese adult women. *British Journal of Nutrition*, *106*, 1117–1127.

Pitsavos, C., Panagiotakos, D. B., Lentzas, Y., & Stefanadis, C. (2005). Epidemiology of leisure-time physical activity in socio-demographic, lifestyle and psychological characteristics of men and women in Greece: the ATTICA Study. *BMC Public Health, 5,* 37.

Rzewnicki, R., Auweele, Y. V., & Bourdeaudhuij, I. D. (2003). Addressing over-reporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutrition*, *6*(3), 299–305.

Siew Man, C., Kandiah, M., Chinna, K., Yoke Mun, C., & Hazizi, A. S. (2010). Prevalence of Obesity and Factors Associated with it in a Worksite Setting in Malaysia. *Journal of Community Health, 35*, 698–705.

Siti Affira, K., Mohd Nasir, M. T., Hazizi, A. S., & Kandiah, M. (2011). Socio-Demographic and Psychosocial Factors Associated with Physical Activity of Working Woman in Petaling Jaya, Malaysia. *Malaysian Journal of Nutrition, 17*(3), 315-324.

Soo, K. L., Wan Abdul Manan, W. M., & Wan Suriati, W. N. (2012). The Bahasa Melayu Version of the Global Physical Activity Questionnaire: Reliability and Validity Study in Malaysia. *Asia Pacific Journal of Public Health, 10*. doi:10.1177/1010539511433462.

The U.S. Department of Health and Human Services (2008). *2008 Physical Activity Guidelines for Americans*. The U.S. Department of Health and Human Services.

The World Bank Group. (2011). *The World at a Glance: Key development indicators from the World Bank.* Retrieved November 1, 2011 from http://data.worldbank.org.

Thompson, D. L., Rakow, J., & Perdue, S. M. (2004). Relationship between Accumulated Walking and Body Composition in Middle-Aged Women. *Medicine and Science in Sports and Exercise*, 911-914.

Troiano, R. P., Berrigan, D., Dodd, K., Masse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*, 181-188.

Trost, S. G., Mciver, K. L., & Pate, R. R. (2005). Conducting Accelerometer-based Activity Assessments in Field-based Research. *Medicine and Science in Sports and Exercise*, *37*(11), S531-S543.

Tudor-Locke, C., & Bassett, D. R. Jr. (2004). How Many Steps/day Are Enough?. *Sports Medicine*, *34*, 1-8.

Welk, G. J. (2002). *Physical Activity Assessments for Health-Related Research.* Human Kinetics.

Wilks, D. C., Sharp, S. J., Ekelund, U., Thompson, S. G., Mander, A. P., Turner, R. M., Jebb, S. A., & Lindroos, A. K. (2011).Objectively measured physical activity and fat mass in children: a bias-adjusted meta-analysis of prospective studies. *PloS One, 23*, 6, 2, e17205.

World Health Organization. (1998). Obesity: Preventing and managing the global epidemic. *WHO Technical Series Report No. 894*. World Health Organization: Geneva.

World Health Organization. (2008). *WHO STEPwise approach to surveillance (STEPS)*. World Health Organization: Geneva.

World Health Organization. (2010). *Global Recommendations on Physical Activity for Health*. World Health Organization.

Yokoyama, Y., Kawamura, T., Tamakoshi, A., Noda, A., Hirai, M., Saito, H., & Ohno, Y. (2002). Comparison of Accelerometry and Oxymetry for Measuring Daily Physical Activity. *Circulation Journal*, *66*(8), 751-754 .

Pertanika J. Sci. & Technol. 22 (2): 419 - 432 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Hypothesis Tests of Goodness-of-Fit for Fréchet Distribution

Abidin, N. Z.*, Adam, M. B. and Midi, H.

Institute for Mathematical Research, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

Extreme Value Theory (EVT) is a statistical field whose main focus is to investigate extreme phenomena. In EVT, Fréchet distribution is one of the extreme value distributions and it is used to model extreme events. The degree of fit between the model and the observed values was measured by Goodness-of-fit (GOF) test. Several types of GOF tests were also compared. The tests involved were Anderson-Darling (AD), Cramer-von Mises (CVM), Zhang Anderson Darling (ZAD), Zhang Cramer von-Mises (ZCVM) and L_n. The values of parameters μ, σ and ξ were estimated by Maximum Likelihood. The critical values were developed by Monte-Carlo simulation. In power study, the reliability of critical values was determined. Besides, it is of interest to identify which GOF test is superior to the other tests for Fréchet distribution. Thus, the comparisons of rejection rates were observed at different significance levels, as well as different sample sizes, based on several alternative distributions. Overall, given by Maximum Likelihood Estimation of Fréchet distribution, the ZAD and ZCVM tests are the most powerful tests for smaller sample size (ZAD for significance levels 0.05 and 0.1, ZCVM for significance level 0.01) as compared to AD, which is more powerful for larger sample size.

Keywords: Critical values Fréchet distribution, goodness-of-fit, rejection rate

INTRODUCTION

Extreme Value Theory (EVT) is a statistical discipline with the main focus is to estimate the probability of the occurrence of extreme phenomenon. The extreme event occurs either at maximum or minimum level (Coles, 2001). For instance, the likelihood of the flood during certain period of time is predicted by observing the volume of water at maximum level. Likewise, the study on the occurrence of draught was done by taking into account the minimum volume of rainfall (Castillo *et al*., 2005). These maximum and minimum values were collected and modelled based on the statistical extreme models.

There are several statistical extreme models in EVT that are broadly used for the extrapolation. The EVT comprises Gumbel,

Article history:
Received: 11 January 2012
Accepted: 28 August 2012

Email addresses:
Abidin, N. Z.(nahdiya@upm.edu.my), Adam, M. B. (bakri@upm.edu.my), Midi, H. (habshah@upm.edu.my)
*Corresponding Author

ISSN: 0128-7702

Weibull and Fréchet's models. A combination of the three extreme models by Jenkinson (1955) is known as Generalized Extreme Value (GEV). The combination was introduced for the purpose of simplifying the modelling steps. In other words, instead of testing several assumptions on which extreme model is more likely to fit the data, GEV allows for the determination of the most suitable model based on a single value of shape parameter, ξ, which represents the types of tail behaviour (Coles, 2001). Each kind of extreme models has certain interval value of ξ explained in Theorem 1.1. For the maximum level of extreme, the extreme behaviour is denoted as $M_n = \max(X_1, ..., X_n)$, where M_n is the maximum value of the observation X over n time.

Theorem 1.1: *Let $X_1, ..., X_n$ be an independent random variables with the distribution F, and let asymptotic argument be $M_n = \max(X_1, ..., X_n)$. As the constants $a_n > 0$ and b_n exist, denote*

$$\lim_{n \to \infty} \Pr\left(\frac{M_n - b_n}{a_n} \leq x \right) \to F(x) \qquad [1]$$

If the non-degenerate function F exists, then GEV is:

$$F(x) = \exp\left\{ -\left[1 + \xi\left(\frac{x - \mu}{\sigma} \right) \right]_+^{-\frac{1}{\xi}} \right\} \qquad [2]$$

where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ are the location, scale and shape parameters respectively. The GEV distribution should comply with $\left[1 + \xi\left(\frac{x - \mu}{\sigma} \right) \right] > 0$. The distribution of Gumbel, Fréchet, and Weibull corresponds to $\xi = 0$, $\xi > 0$, and $\xi < 0$, respectively.

Given that the main concern of EVT is to model the extreme behaviour, it is crucial that the extreme model is able to reflect the real event. This is because the choice of the extreme model affects the outcomes of decision making and problem solving (Castillo *et al*., 2005). Therefore, a careful model validation is necessary. The validation test is called Goodness-of-fit (GOF) test. GOF test is a statistical test used to measure the fit of the selected statistical model against the observed values (Kinnison, 1989). When the model fits the observed distribution, the model can be utilized to interpret the behaviour of the distribution as well as to predict the outcomes. Thus, the GOF test plays a significant role to ensure the selected model is able to precisely reflect the population of the observed values (Shabri & Jemain, 2008). The GOF test that is commonly used is the hypothesis test using empirical distribution function (Zempléni, 2004). Several classical GOF tests are Anderson Darling (AD), Cramer-von Mises (CVM) and Kolmogrov Smirnov (KS). The performances of the GOF tests vary, depending on the types of the distribution and the methods of parameter estimation. Hence, in order to select the most appropriate GOF test, studies have been done to examine the power of the GOF tests for certain statistical distributions.

For extreme circumstances, there have been many studies conducted on the performance of the GOF tests for Gumbel and Weibull distributions (Kimber, 1985; Coles, 1989; Kinnison, 1989; Lockhart & Stephens, 1994; Liao & Shimokawa, 1999; Kotz & Nadarajah, 2000; Shabri

& Jemain, 2009; Laio, 2004; Zempleni, 2004). Other than that, Zhang (2002) and Zhang and Wu (2005) modified the classical AD, CVM and KS tests. These modified tests are named after Zhang who is also known as Zhang Anderson Darling (ZAD), Zhang Cramer-von Mises (ZCVM) and Zhang Kolmogorov Smirnov (ZKS). These later tests are more powerful than the classical tests, except for ZKS. However, the Zhang tests were evaluated for Normal distribution. Therefore, it is of interest to test the power of the Zhang test for the GEV distribution, particularly for Fréchet distribution. This is because the assessment of the GOF test for Fréchet distribution has yet to receive extensive attention.

Although only a few studies on the GOF test for Fréchet distribution have ever been discussed (Koning & Peng, 2008; Abd-Elfattah *et al*., 2010), this distribution plays a major role in the modelling of extreme events. Fréchet has been used to model heavy tailed distribution such as the option pricing in the extreme financial losses (Markose & Alentorn, 2011), hydrology and internet traffic (Koning & Peng, 2008). Moreover, many practical problems have the limit of maxima values that converge to Fréchet distribution (Castillo *et al.*, 2005). Thus, the identification of the best GOF test for Fréchet is important to facilitate the practitioners to have more precise evaluation on the degree of fit between the model and the observed values so that more reliable prediction on the extreme events related to Fréchet can be achieved. Therefore, the purpose of this study is to identify the most powerful GOF test coupled with parameter estimate of Maximum Likelihood for Fréchet distribution.

METHODOLOGY

The hypotheses for the GOF test are shown below. The proposition that the hypothetical statistical model fits the observed distribution is equivalent to fail to reject H_0. Otherwise, the rejection of H_0 implies the model does not fit the observed values. In this study, the hypothetical distribution is the Fréchet distribution.

H_0: The hypothetical model fits the observed distribution or $F_h(x) = F(x)$
H_1: The hypothetical model does not fit the observed distribution or $F_h(x) \neq F(x)$

Critical Values of the GOF Tests

Critical values are the baseline of whether to reject or fail to reject H_0. If H_0 is to be rejected, the statistics value produced by the GOF test should be able to exceed the critical value. The development of critical values was made by means of Monte Carlo simulation:

Step (A): Generate the random variables for sample of size 15. Random variables were generated from the inverse function of GEV with parameters μ = 100, σ = 10 and ξ = 0.1, 0.2, 0.3, 0.4, 0.5. The values of ξ are within the values exhibited by Shin *et al* (2012), Shabri, and Jemain (2008), and Ahmad (1988). Moreover, standard values for μ and σ are 0 and 1, respectively (Shin *et al*., 2012; Shabri & Jemain, 2008; Ahmad, 1988). However, the value of σ = 10 was selected for simulation because the dispersion is wider than the standard value. Hence, it is easier to reach the convergence of parameter estimation using Maximum

Likelihood Estimation, especially for small sample size. Furthermore, to determine whether the critical values of $\mu = 100$ and $\sigma = 10$ are similar to the standard values, the rejection rate for Fréchet with $\mu = 0$, $\sigma = 1$ was compared. The inverse function of GEV is:

$$F(x)^{-1} = \mu - \frac{\sigma}{\xi}(1 - \log U)^{-\xi} \quad [3]$$

where U is hypothetical distribution function. The value for U is $U_{i,n} = \frac{i - 0.5}{n}$. These random variables were arranged in the ascending order.

Step (B): The parameters were estimated. The values of parameters μ, σ and ξ were estimated by Maximum Likelihood Estimation. The loglikelihood of Fréchet distribution is

$$l(\mu, \sigma, \xi) = -\sum_{i=1}^{n}\left[1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} - \left(\frac{1}{\xi} + 1\right)\sum_{i=1}^{n}\log\left[1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right] - n\log(\sigma) \quad [4]$$

The maximum likelihood estimate is obtained by maximizing the loglikelihood expression by way of partial derivatives method. These differentiations were carried out with respect to each parameter.

Step (C): Based on the arranged random variables and the estimated parameters, the cumulative distribution function, *F(x),* for Fréchet was determined.

Step (D): The values of *F(x)* were substituted into the expressions of the GOF tests. The expression of the GOF tests involved is shown below. The values produced by the expression of the GOF tests are called the statistics values:

Anderson Darling test (AD):

$$AD = \sum_{i=1}^{n}\frac{2i - 1}{n}\{\log[F(x_i)] + \log[1 - F(x_{n+1-i})]\} - n \quad [5]$$

Cramer von-Mises test (CVM):

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_i) - \frac{2i - 1}{2n}\right]^2 \quad [6]$$

Zhang Anderson Darling test (ZAD):

$$ZAD = \sum_{i=1}^{n}\left\{\frac{\log[F(x_i)]}{n - i + 0.5} + \frac{\log[1 - F(x_i)]}{i - 0.5}\right\} \quad [7]$$

Zhang Cramer von-Mises test (ZCVM):

$$ZCVM = \sum_{i=1}^{n} \left\{ \log \left[\frac{\frac{1}{F(x_i)} - 1}{\frac{n - 0.5}{i - 0.75}} - 1 \right] \right\}^2 \quad [8]$$

L_n test:

$$L_n = \frac{1}{\sqrt{n}} \frac{\max\left[\frac{i}{n} - F(x_i), F(x_i) - \frac{i-1}{n}\right]}{\sqrt{F(x_i)\left[1 - F(x_i)\right]}} \quad [9]$$

Step (E): The steps from A to D were iterated for 10,000 times. The 10,000 iterations yield 10,000 statistics values. Those 10,000 statistics values were arranged in the ascending order. The order of the statistics value at the percentiles of 99, 95, and 90 implies the critical value at the significance level of 0.01, 0.05 and 0.10, respectively.

The association between different values of ξ and the critical values was performed by means of average. In other words, the critical values obtained from $\xi = 0.1, 0.2, 0.3, 0.4, 0.5$ were averaged for each sample size and significance level. Thus, for $0.1 < \xi < 0.5$, there is a single critical value at a particular sample size, as well as significance level. The steps from A to E were done for a sample of size n=20, 30, 40, 50 and 100.

Power of the GOF Tests

Power is the statistics value of $F(x)$ that exceeds the critical value of $F_h(x)$. Similarly, it is also known as the probability of rejecting H_0. Low rejection rate implies low probability of rejecting H_0. Hence, the rejection rate is normally applied to check the probability of failing to reject H_0. When $F_h(x) = F(x)$ is true, the rejection rate approximates the respective significance level (Laio, 2004; Shin *et al.,* 2012). The ability of any GOF test to get these approximate values signifies that the respective critical values obtained are reliable. The assessment is done by comparing the hypothetical distribution which is Fréchet with other observed values from Fréchet distributions, which are: Fréchet with true parameters values (μ = 100, σ = 10 and ξ = 0.5) and Fréchet with different parameter values (μ = 0, σ = 1 and ξ = 0.3).

In contrast, high rejection rate implies a high probability of rejecting H_0. When $F_h(x) \neq F(x)$, the rejection rate is higher than the respective significance level (Laio, 2004; Shin *et al.,* 2012). The higher the rejection rate, the more powerful the GOF test is. Therefore, rejection rate is also used to evaluate the degree of power of the GOF test. This evaluation is done by comparing the Fréchet model with the alternative statistical distributions. Given by the alternative distributions, the power test was conducted to evaluate the ability of the GOF tests to exceed the critical values of Fréchet and subsequently reject H_0.

Meanwhile, the Monte Carlo method is used to assess the degree of rejection rate of the GOF tests using different sample size and at different significance levels. The steps of calculating the rejection rate are similar to the steps done for critical values mentioned previously. For step A, the random variables of Fréchet with true parameter values $\mu = 100$, $\sigma = 10$ and $\xi = 0.5$ were simulated. In addition, the random variables of Fréchet for different parameter values $\mu = 0$, $\sigma = 1$ and $\xi = 0.3$ were also generated. After that, the random variables of the following alternative distributions were simulated: Gamma distribution; Gamma~(3,1), Weibull distribution; weibull~(0,1,0.5), Lognormal distribution; Lognorm~(0, 1, 0.5), Normal distribution; N~(0,1), Generalized Logistic distribution; Glog~(0, 1, -0.5), and Exponential distribution; Exp~(1).

Next, the steps from B to D were followed orderly. At step E, the simulations were done for 10,000 iterations. The power values of GOF tests were determined by averaging the 10,000 statistics values exceeding the critical values of Fréchet.

RESULTS

Table 1 represents the critical values of the GOF tests for Fréchet distribution with respect to the sample of sizes, n= 15, 20, 30, 40 50 and 100. These critical values of the GOF tests for the Fréchet distribution are illustrated in Fig.1. In Fig.1, the critical values of AD is on the top left, ZAD is on the top right, CVM is on the middle left, ZCVM is on the middle right and L_n is on the bottom left. Based on Table 1 and Fig.1, the critical values of AD are within the interval of 0.4 to 0.9. On the other hand, the critical values of ZAD are around 3.4 as the sample size increases. The critical values for CVM ranged from 0.05 to 0.15. For ZCVM and L_n, the critical values increase along with the increment of sample size, starting from 4 to 16 and 2 to 2.8, respectively.

The reliability of the critical values was observed from the rejection rates for Fréchet ($\mu = 100$, $\sigma = 10$ and $\xi = 0.5$) and Fréchet ($\mu = 0$, $\sigma = 1$ and $\xi = 0.3$), as presented in Table 2. Table 2 shows that the rejection rates of the GOF tests for both the distributions are close to the respective significance level. The rejection rates of the GOF tests at the significance level of 0.01 are approximately 0.01. The same trend goes to the rejection rates at 0.05 and 0.1 significance levels. Moreover, Fig.2 (top left and top right) portrays the rejection rates at significance level of 0.05. The rejection rates of all the GOF tests are within 0.04 and 0.06.

In addition, Table 2 exhibits the rejection rates of the GOF tests for the alternative distributions. The values highlighted in bol represent the highest rejection rate at each sample size, significance level and alternative distribution. The results show that at 0.01 significance level, ZCVM is the most powerful test for small sample sizes, which are n=15 and 20. For n=30, ZAD outperforms the other GOF tests. For larger sample size (n=40, 50 and 100), AD test is superior to other GOF tests. On the other hand, the results of the rejection rates at 0.05 and 0.1 significance levels are similar. For n=15, 20, 30 and 40, the most powerful test is ZAD. For n=50 and 100, ShabriJemain AD produces the greatest rejection rates than other competitors. The results for the rejection rates at significance level of 0.05 are depicted by Figures 2 and 3.

TABLE 1
Critical values of the GOF tests for Fréchet distribution

Test	AD			ZAD		
sig.lvl / n	0.10	0.05	0.01	0.10	0.05	0.01
15	0.457	0.546	0.800	3.355	3.380	3.470
20	0.467	0.534	0.716	3.350	3.371	3.418
30	0.476	0.554	0.771	3.345	3.363	3.402
40	0.479	0.560	0.765	3.340	3.353	3.387
50	0.481	0.577	0.769	3.336	3.348	3.374
100	0.496	0.583	0.782	3.323	3.330	3.350
Test	CVM			ZCVM		
sig.lvl / n	0.10	0.05	0.01	0.10	0.05	0.01
15	0.076	0.092	0.133	4.648	5.455	7.888
20	0.077	0.091	0.125	5.231	6.076	7.905
30	0.078	0.095	0.132	6.275	7.356	9.579
40	0.079	0.093	0.133	6.943	8.206	10.745
50	0.079	0.096	0.130	7.598	8.755	11.780
100	0.082	0.098	0.133	9.600	11.176	15.057
Test	L_n					
sig.lvl / n	0.10	0.05	0.01			
15	2.037	2.216	2.671			
20	2.088	2.265	2.641			
30	2.155	2.345	2.728			
40	2.202	2.411	2.803			
50	2.220	2.417	2.804			
100	2.296	2.482	2.871			

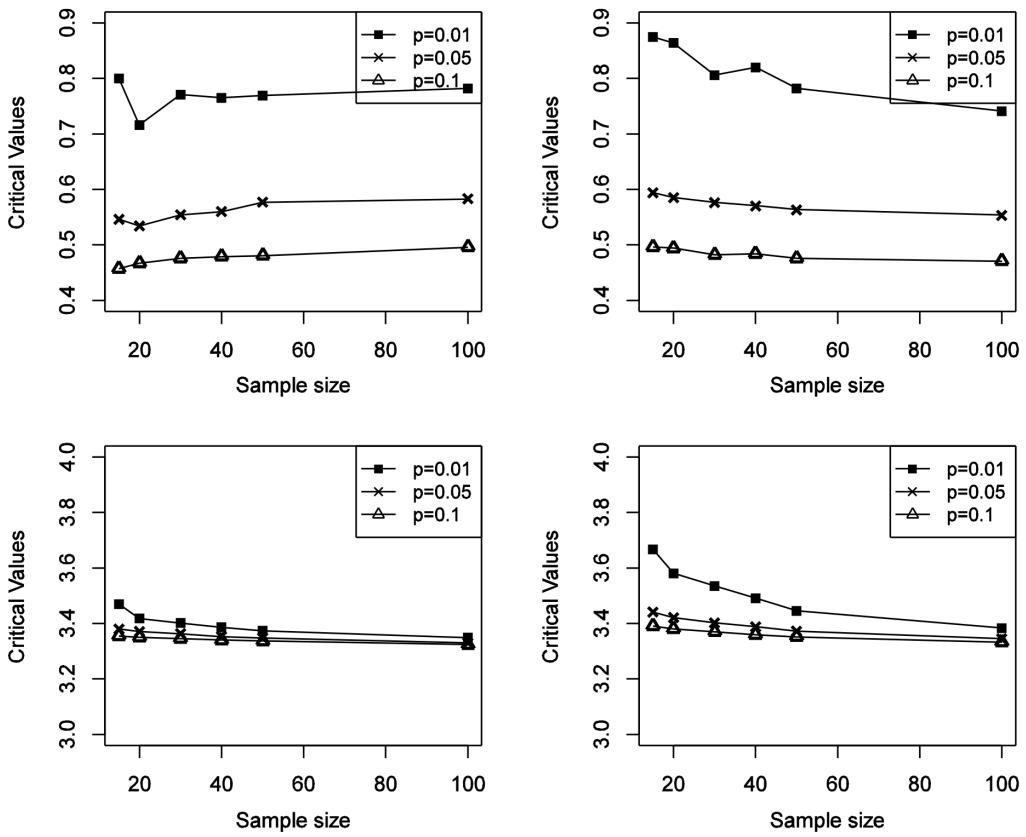


Fig.1: Critical values of the GOF tests for Fréchet distribution with respect to the sample of size, n= 15, 20, 30, 40 50 and 100. The critical values of AD is on the top left, ZAD is on the top right, CVM is on the middle left, ZCVM is on the middle right and L_n is on the bottom left.

TABLE 2
Rejection rate for GOF tests based on the selected distributions other than Fréchet

Distribution	Test	AD			ZAD			CVM		
	sig. lvl / n	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
GEV type-II (μ = 100, σ = 10 and ξ = 0.5)	15	0.100	0.054	0.012	0.093	0.056	0.012	0.104	0.054	0.013
	20	0.103	0.051	0.010	0.113	0.050	0.018	0.097	0.050	0.011
	30	0.086	0.052	0.009	0.103	0.049	0.010	0.095	0.046	0.012
	40	0.101	0.050	0.010	0.101	0.047	0.011	0.107	0.055	0.008
	50	0.103	0.050	0.014	0.101	0.048	0.012	0.097	0.054	0.015
	100	0.093	0.049	0.014	0.091	0.046	0.005	0.091	0.051	0.013

TABLE 2 (continue)

GEV type-II ($\mu = 0$, $\sigma = 1$ and $\xi = 0.3$)	15	0.109	0.055	0.011	0.118	0.059	0.014	0.113	0.057	0.011
	20	0.099	0.054	0.016	0.110	0.058	0.016	0.097	0.055	0.011
	30	0.097	0.051	0.009	0.099	0.052	0.019	0.090	0.044	0.008
	40	0.108	0.045	0.010	0.118	0.052	0.012	0.113	0.049	0.010
	50	0.106	0.045	0.012	0.111	0.054	0.015	0.115	0.054	0.013
	100	0.092	0.053	0.012	0.114	0.057	0.010	0.088	0.054	0.012
Gamma	15	0.811	0.695	0.460	**0.872**	**0.770**	0.500	0.752	0.625	0.438
	20	0.783	0.681	0.547	**0.834**	**0.766**	**0.600**	0.708	0.586	0.472
	30	0.786	0.674	0.472	**0.828**	**0.689**	**0.486**	0.734	0.605	0.421
	40	0.761	0.663	**0.511**	**0.778**	**0.675**	0.472	0.682	0.608	0.468
	50	0.769	**0.687**	**0.476**	0.750	0.661	0.457	0.715	0.620	0.418
	100	**0.743**	**0.674**	**0.497**	0.674	0.586	0.362	0.694	0.607	0.464
Weibull	15	0.793	0.717	0.470	**0.837**	**0.801**	0.506	0.733	0.651	0.437
	20	0.786	0.714	0.534	**0.845**	**0.758**	0.554	0.714	0.640	0.474
	30	0.781	0.684	0.494	**0.817**	**0.713**	0.482	0.698	0.601	0.455
	40	**0.776**	**0.686**	**0.485**	0.766	0.658	0.483	0.731	0.613	0.417
	50	**0.769**	**0.685**	**0.486**	0.761	0.667	0.452	0.695	0.616	0.440
	100	**0.723**	**0.656**	**0.509**	0.674	0.580	0.348	0.649	0.589	0.437
Generalized Logistic	15	0.818	0.692	0.467	**0.870**	**0.769**	0.514	0.746	0.624	0.434
	20	0.773	0.733	0.556	**0.831**	**0.788**	0.582	0.700	0.654	0.494
	30	0.767	0.669	0.487	**0.797**	**0.689**	**0.502**	0.691	0.591	0.458
	40	0.756	0.687	**0.497**	**0.804**	**0.697**	0.461	0.696	0.631	0.437
	50	0.775	**0.669**	**0.506**	**0.776**	0.663	0.476	0.709	0.612	0.443
	100	**0.758**	**0.681**	**0.532**	0.682	0.588	0.350	0.664	0.601	0.473
Exponential	15	0.801	0.718	0.477	**0.865**	**0.773**	0.503	0.729	0.655	0.453
	20	0.767	0.728	0.554	**0.835**	**0.796**	**0.587**	0.722	0.643	0.490
	30	0.776	0.701	0.487	**0.796**	**0.711**	0.503	0.690	0.635	0.442
	40	**0.785**	0.677	**0.505**	0.775	**0.682**	0.469	0.733	0.612	0.440
	50	0.789	0.672	0.504	0.769	0.656	0.465	0.709	0.610	0.458
	100	**0.749**	**0.675**	**0.499**	0.685	0.583	0.361	0.681	0.624	0.442
Lognormal	15	0.769	0.697	0.476	**0.833**	**0.768**	0.520	0.687	0.627	0.445
	20	0.796	0.720	0.563	**0.846**	**0.753**	0.587	0.705	0.651	0.478
	30	0.793	0.680	0.517	**0.832**	**0.712**	**0.529**	0.735	0.593	0.458
	40	0.785	0.672	**0.500**	**0.790**	**0.693**	0.482	0.728	0.611	0.434
	50	**0.786**	**0.700**	**0.504**	0.778	0.659	0.480	0.711	0.636	0.459
	100	**0.771**	**0.647**	**0.489**	0.702	0.587	0.364	0.701	0.585	0.443
Normal	15	0.793	0.696	0.466	**0.845**	**0.771**	0.495	0.727	0.639	0.434
	20	0.783	0.719	0.535	**0.848**	**0.747**	0.605	0.713	0.656	0.492
	30	0.750	0.708	0.511	**0.806**	**0.719**	**0.540**	0.678	0.629	0.473
	40	0.738	0.681	**0.500**	**0.790**	**0.704**	0.492	0.669	0.619	0.438
	50	**0.785**	**0.651**	**0.500**	0.775	0.626	0.483	0.716	0.605	0.455
	100	**0.775**	**0.641**	**0.490**	0.718	0.581	0.353	0.703	0.586	0.444

TABLE 2 (continue)

Distribution	Test: sig. lvl / n	ZCVM			L_n		
		0.10	0.05	0.01	0.10	0.05	0.01
GEV type-II ($\mu = 100$, $\sigma = 10$ and $\xi = 0.5$)	15	0.099	0.054	0.013	0.106	0.051	0.015
	20	0.099	0.049	0.013	0.107	0.043	0.013
	30	0.096	0.057	0.013	0.097	0.049	0.010
	40	0.110	0.043	0.010	0.099	0.048	0.014
	50	0.099	0.056	0.008	0.095	0.047	0.016
	100	0.100	0.046	0.008	0.094	0.057	0.007
GEV type-II ($\mu = 0$, $\sigma = 1$ and $\xi = 0.3$)	15	0.118	0.049	0.012	0.113	0.057	0.014
	20	0.116	0.058	0.015	0.091	0.058	0.018
	30	0.114	0.049	0.017	0.093	0.055	0.014
	40	0.119	0.055	0.006	0.109	0.048	0.011
	50	0.117	0.049	0.017	0.105	0.052	0.012
	100	0.107	0.057	0.008	0.104	0.052	0.015
Gamma	15	0.854	0.759	**0.532**	0.650	0.558	0.371
	20	0.801	0.736	**0.600**	0.634	0.531	0.421
	30	0.795	0.675	0.480	0.658	0.535	0.359
	40	0.762	0.648	0.472	0.611	0.503	0.347
	50	0.728	0.664	0.423	0.625	0.547	0.341
	100	0.664	0.566	0.353	0.624	0.542	0.373
Weibull	15	0.816	0.776	**0.537**	0.656	0.569	0.347
	20	0.819	0.752	**0.570**	0.658	0.555	0.382
	30	0.785	0.690	**0.499**	0.628	0.528	0.371
	40	0.756	0.626	0.464	0.653	0.492	0.323
	50	0.750	0.659	0.419	0.618	0.531	0.339
	100	0.649	0.562	0.334	0.572	0.499	0.345
Generalized Logistic	15	0.859	0.763	**0.527**	0.684	0.550	0.363
	20	0.817	0.758	**0.583**	0.638	0.564	0.390
	30	0.765	0.660	0.495	0.640	0.511	0.358
	40	0.785	0.664	0.454	0.610	0.520	0.344
	50	0.754	0.665	0.454	0.631	0.532	0.355
	100	0.661	0.564	0.349	0.603	0.503	0.374
Exponential	15	0.840	0.756	**0.525**	0.645	0.567	0.359
	20	0.808	0.780	**0.587**	0.641	0.573	0.394
	30	0.766	0.679	**0.504**	0.635	0.541	0.342
	40	0.765	0.664	0.477	0.646	0.509	0.366
	50	0.741	0.644	0.440	0.646	0.537	0.364
	100	0.677	0.564	0.360	0.615	0.528	0.364

TABLE 2 (continue)

Lognormal	15	0.813	0.758	**0.535**	0.634	0.538	0.358
	20	0.814	**0.753**	**0.589**	0.660	0.562	0.398
	30	0.776	0.683	0.527	0.660	0.533	0.405
	40	0.768	0.654	0.481	0.645	0.497	0.330
	50	0.748	0.651	0.441	0.627	0.547	0.350
	100	0.690	0.565	0.350	0.629	0.503	0.353
Normal	15	0.835	0.747	**0.520**	0.636	0.569	0.364
	20	0.821	0.736	**0.619**	0.644	0.581	0.382
	30	0.775	0.695	0.531	0.604	0.553	0.376
	40	0.769	0.672	0.479	0.595	0.492	0.338
	50	0.755	0.627	0.448	0.656	0.516	0.362
	100	0.693	0.551	0.348	0.631	0.518	0.364

* The value in **bold** denotes the highest rejection rate for each sample size, significance level and alternative distribution.

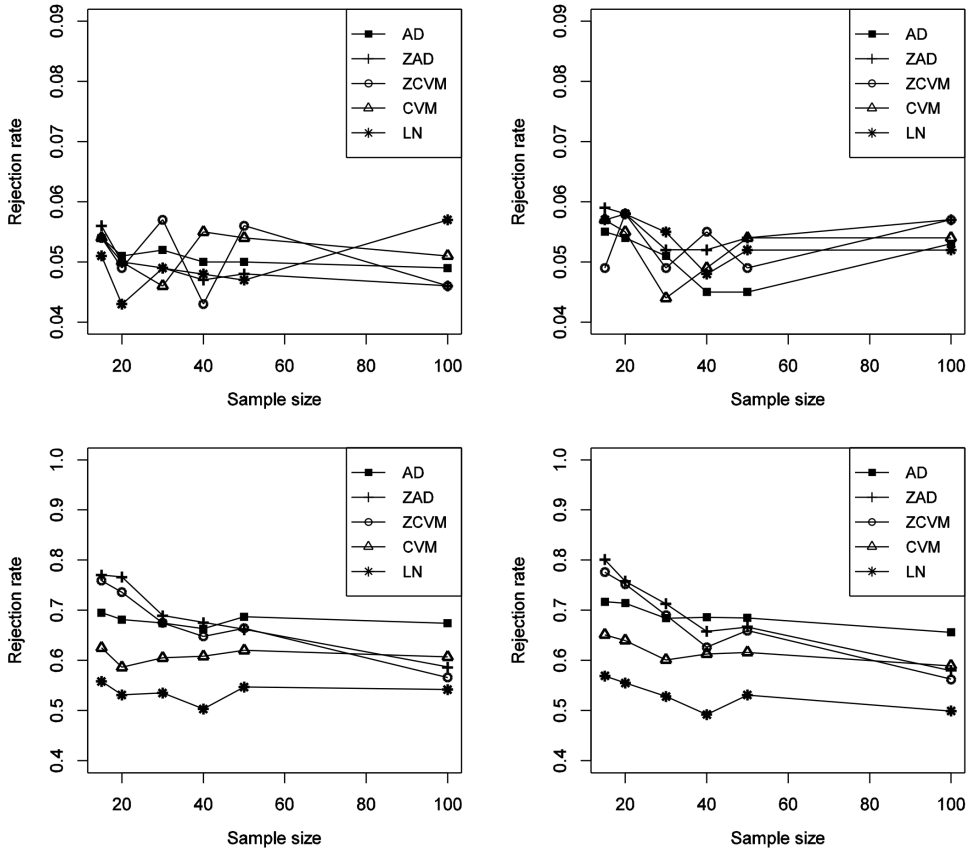


Fig.2: The rejection rate of the GOF tests for Fréchet distribution with respect to the sample of size, n= 15, 20, 30, 40, 50 and 100 at significance level 0.05. The rejection rate for Fréchet (μ = 100, σ = 10 and ξ = 0.5) is on the top left, Fréchet (μ = 0, σ = 1 and ξ = 0.3) is on the top right; Gamma is on the bottom left and Weibull on the bottom right.

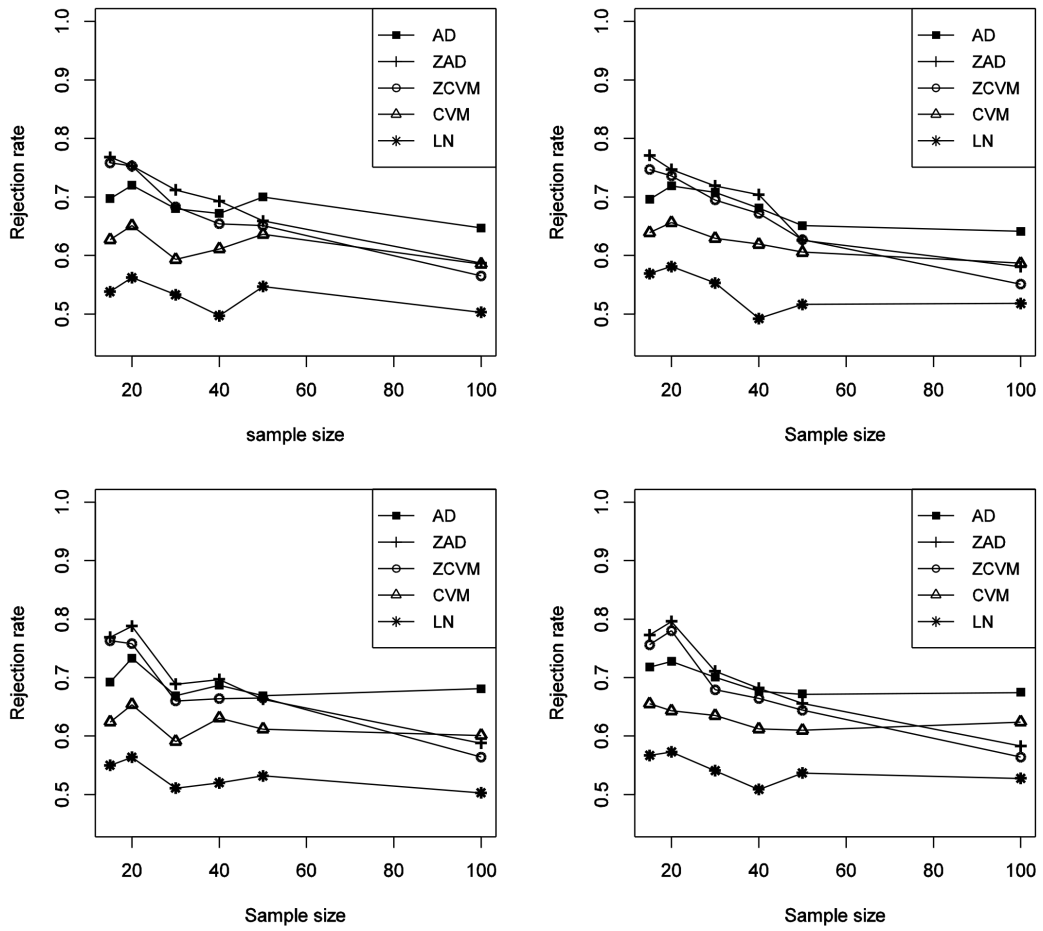


Fig.3: Rejection rate of the GOF tests for Fréchet distribution with respect to the sample of size n= 15, 20, 30, 40, 50 and 100 at significance level 0.05. The rejection rate for Lognormal is on the top left, Normal is on the top right; Generalized Logistics is on the bottom left and Exponential on the bottom right.

DISCUSSION

The assessment on the GOF tests begins with the development of critical values. The establishment of the critical values is crucial because they are the border points in deciding whether the selected statistical model fits the observed distribution or not. If the critical values are correct, the model will certainly match the observed distribution so the prediction of extreme event can be made effectively. Otherwise, the practitioners may deliver inaccurate information that will lead to the failure of prevention planning. Therefore, the validation of the critical values is crucial. For the purpose of validating the critical values, GOF for the true parameter values of the Fréchet distribution is evaluated first. Besides, it is important to determine whether the same result can be arrived at by using the same distribution with different parameters values. Despite having different parameters values, the GOF test should fail to reject H_0 because the distribution comes from the same parent distribution, which is the Fréchet distribution. The GOF test fails to reject H_0 if the rejection rates approximate the particular significance level.

For both Fréchet with true and different parameter values, the rejection rates of AD, ZAD, CVM, ZCVM and L_n tests are reliable because these rates are close to the respective significance level.

After the development and validation of critical values, the identification of the best GOF test is of interest. This is because the most powerful test yields the most accurate measurement on the degree of fit. In general, at significance level of 0.01, ZVCM is preferable for a small sample size (n=15 and 20), while the AD test performs the best for larger sample sizes (n=40, 50, 100). On the other hand, at 0.05 and 0.1 significance levels, ZAD is the most suitable test for GOF for small sample size (15 and 20) and sample of sizes n=30 and 40. For larger sample size (n=50 and 100), on the other hand, the AD critical values are the most powerful test.

CONCLUSION AND RECOMMENDATIONS

In this study, the performances of the GOF tests for Fréchet distribution were observed. Based on the parameter estimation by Maximum Likelihood, the ZAD and ZCVM tests are the most powerful ones for smaller sample sizes (ZAD for significance levels 0.05 and 0.1, ZCVM for significance level 0.01) as compared to AD which is more powerful for larger sample sizes. The GOF test for Fréchet can be assessed for different approaches of parameter estimations. In addition, the modification of the existing GOF can be extended to boost the power values. The sensitivity of the GOF tests can be evaluated in future studies.

REFERENCES

Abd-Elfattah, A. M., Fergany, H. A., & Omima, A. M. (2010). Goodness of fit tests for generalized frechet distribution. *Australian Journal of Basic and Applied Sciences*, *4*(2), 286-301.

Ahmad, M. I., Sinclairm C. D., & Spurrm B. D. (1988). Assessment of flood frequency models using empirical distribution function statistics. *Water Resour. Res.*, *24*(8), 1323–1328.

Castilo, E., Hadi, A. S., & Sarabia, J. M. (2005). *Extreme value and related models with applications in engineering and science.* New Jersey: John Wiley & Sons.

Coles, S. G. (1989). On goodness-of-fit tests for the two-parameter weibull distribution derived from the stabilized probability plot. *Biometrika, 76*(3), 593-598.

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* London: Springer.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorology Society, 87,* 145-158.

Kimber, A. C., (1985). Tests for the exponential, weibull and gumbel distributions based on the stabilized probability plot. *Biometrica, 72*(3), 661-3.

Kinnison, R. (1989). Correlation Coefficient goodness-of-fit test for the extreme-value distribution. *The American Statistician, 4*(2), 98-100.

Koning, A. J., & Peng, L. (2008). Goodness-of-fit tests for a heavy tailed distribution. *Journal of Statistical Planning and Inference, 138,* 3960-3981.

Kotz, S., & Nadarajah, S. (2000). *Extreme value distributions, theory and applications*. London: Imperial College Press.

Laio, F. (2004). Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research, 40,* 1–10.

Liao, M., & Shimokawa, T. (1999). A new goodness-of-fit test for type-1 extreme value and 2-parameter Weibull distributions with estimated parameters. *Journal of Statistical Computation and Simulation, 64*(1), 23–48.

Lockhart, R. A., & Spinelli, J. J. (1990). Comment on Correlation Coefficient goodness-of-fit Test for the extreme-value distribution. *The American Statistician, 44*, 259-260.

Lockhart, R. A., & Stephens, M. A. (1994). Estimation and tests of fit for the three-parameter weibull distribution. *Journal of the Royal Statistical Society. Series B (Methodological), 56*(3), 491-500.

Markose, S., & Alentorn, A. (in press). The generalized extreme value (GEV) distribution, implied tail index and option pricing. *Journal of Derivatives.*

Shabri, A., & Jemain, A.A. (2008). The Anderson-Darling test statistic of the generalized extreme value. *Matematika, 24*(1), 85-97.

Shabri, A., & Jemain, A. A. (2009). Goodness-of-fit test for extreme value type I distribution. *Matematika, 25*(1), 53–66.

Shin, H., Jung, Y., Jeong, C., & Heo, J. H. (2012). Assessment of modified Anderson–Darling test statistics for the generalized extreme value and generalized logistic distributions. *Stoch Environ Res Risk Assess, 26,* 105-114.

Zempléni, A. (2004). *Goodness-of-fit test in extreme value applications*. Unpublished manuscript, Eötvös Loránd University, Budapest, Hungary.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B, 64*(2), 281-294.

Zhang, J., & Wu, Y. (2005). Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis, 49,* 709-721.

Pertanika J. Sci. & Technol. 22 (2): 433 - 444 (2014)

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Distribution of Recent Ostracoda in Offshore Sediment of the South China Sea

Ramlan, O.* and Noraswana, N. F.

School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

ABSTRACT

A study on the distribution of Recent Ostracoda in offshore sediment was carried out around the South China Sea. A total of 30 sediment samples were taken from the sampling stations between latitude 1°48' and 7°25'N and longitude 102°09' and 105°16'E. From this study, 79 species of ostracods belonging to 16 families and 44 genera were identified. The dominant species was *Foveoleberis cypraeoides* with 937 individuals obtained. There were 13 to 43 species in total. Diversity Index, H(s), was in the range of 2.1 to 3.3, whereas the dominance values were between 4.4 and 14.7%. Several environmental parameters were measured including depth, temperature and salinity. The range values for each of these parameters are 13-72 m, 25.24-30.06°C and 27.74-34.91 ppt, respectively. The sediment texture in this study area can be categorized as sand, sandy mud, clayey mud, silty mud, silty clay, clayey sand, clayey silt and silty sand. The observations revealed that abundance and diversity of ostracod appeared to be principally controlled by depth. Two faunal assemblages were identified in terms of faunal composition, namely, shallow water (*Hemikrithe orientalis, Neomonoceratina iniqua, Stigmatocythere indica, Cytherelloidea leroyi* and *Neocytheretta snellii*) and deep water (*Paracypris* sp.*, Alataconcha pterogona, Bythocytheropteron alatum, Keijella paucipunctata* and *Actinocythereis scutigera*). A comparative analysis showed a high degree resemblance between the study area and south-eastern Malay Peninsula (the South China Sea).

Keywords: Abundance, diversity, Recent Ostracoda, depth, assemblages

Article history:
Received: 29 January 2010
Accepted: 18 January 2012

Email addresses:
Ramlan, O.(rbo@ukm.my),
Noraswana, N. F. (orangkita_05@yahoo.com)
*Corresponding Author

INTRODUCTION

Among crustaceans, ostracods are the most diverse groups of living crustaceans, whereas these are abundantly represented by fossil arthropods with about 33,000 species (Cohen *et al*., 1998). All of them are essentially aquatic, inhabiting both marine and non-marine environments, although some taxa are

ISSN: 0128-7702

adapted to semi-terrestrial life. The majority of ostacods are free living (benthonic or pelagic) but some are commensal on other crustaceans, echinoderms and even on sharks. The class of ostracoda is subdivided into two, namely, Myodocopa and Podocopa. The vast majority of ostracods encountered in Quaternary sediments are likely to be Podocopids (Horne *et al*., 2002). Ostracods are widely used in biostratigraphy, in determining palaeoenvironments and palaeoclimates and are indispensable as indicators of ancient shorelines and plate distributions. Recent ostracods from the South China Sea have been the subject of numerous investigations in the last decades (e.g., Gou *et al*., 1983; Zhao *et al.*, 1985; Gou, 1990; Whatley & Zhao, 1993; Zhao, 2005). However, information on ostracods from Malaysian waters, particularly in the South China Sea, is comparatively inadequate. An important work on the recent ostracoda of the south eastern Malay Peninsula has been contributed by Zhao and Whatley (1989), wherein ostracods from shallow waters (< 20 m) from the Sedili River and Jason Bay regions were described. Another study on the recent ostracoda of the Malacca Straits was done by Whatley and Zhao (1987 & 1988) who reported a total of 129 species, among which 22 species and 2 genera (*Bythocytheropteron* and *Alataconcha*) are described as new from 18 bottom samples taken from depths ranging from 20 to 100 m. The abundance and diversity of ostracods were found to be influenced by the types of substrate. Thus, this study will provide a more complete list of species and patterns of ostracods distribution in relation to different environmental parameters. The present paper is an attempt to study the distribution of ostracods in this area and to provide a comparative analysis of this group in this region.

MATERIALS AND METHODS

Thirty sediment samples were collected from thirty sampling stations around the South China Sea during June, 2008 (Fig.1). A Smith McIntyre grab sampler was used to collect the sediments over a surface of about 400 cm^2. Several environmental parameters such as temperature, salinity, depth, and sediment type were measured in the study area. The types of sediment were determined in the laboratory, as follows: grain coarser than 2 mm was considered as gravel, from 2 mm to 0.063 mm as sand, from 0.002 to 0.063 mm as silt and finer than 0.002 mm as clay. Using the information of grain distribution, all the samples were classified using a triangular diagram of Folk (1980) on the basis of percentages of sand, silt and clay. For ostracods, the samples were washed over three different size sieves, namely, 0.50 mm, 0.15 mm and 0.063 mm, and dried at 60°C in an oven. The specimens were picked, identified and counted. The data obtained were used to compute the species diversity (number of species in each sample), abundance (specimen number in each sample) and dominance (percentage of the most abundance species in each sample). The Shannon-Wiener's diversity indices, H(s), and correlation analysis were calculated using PAST (Palaeontological Statistics) software to elucidate the nature of the various ostracod communities and their relationships with environmental factors and types of sediment. The species were identified using Scanning Electron Microscopy (SEM) and Light Microscopy (LM). For identification purposes, a comparison was carried out on the morphology features between the collected specimens with the ostracods species that had been recorded by earlier researchers (Zhao *et al.,* 1985; Whatley & Zhao, 1987; 1988; Zhao & Whatley, 1989; Gou, 1990; Mostafawi, 1992; Dewi, 2000).

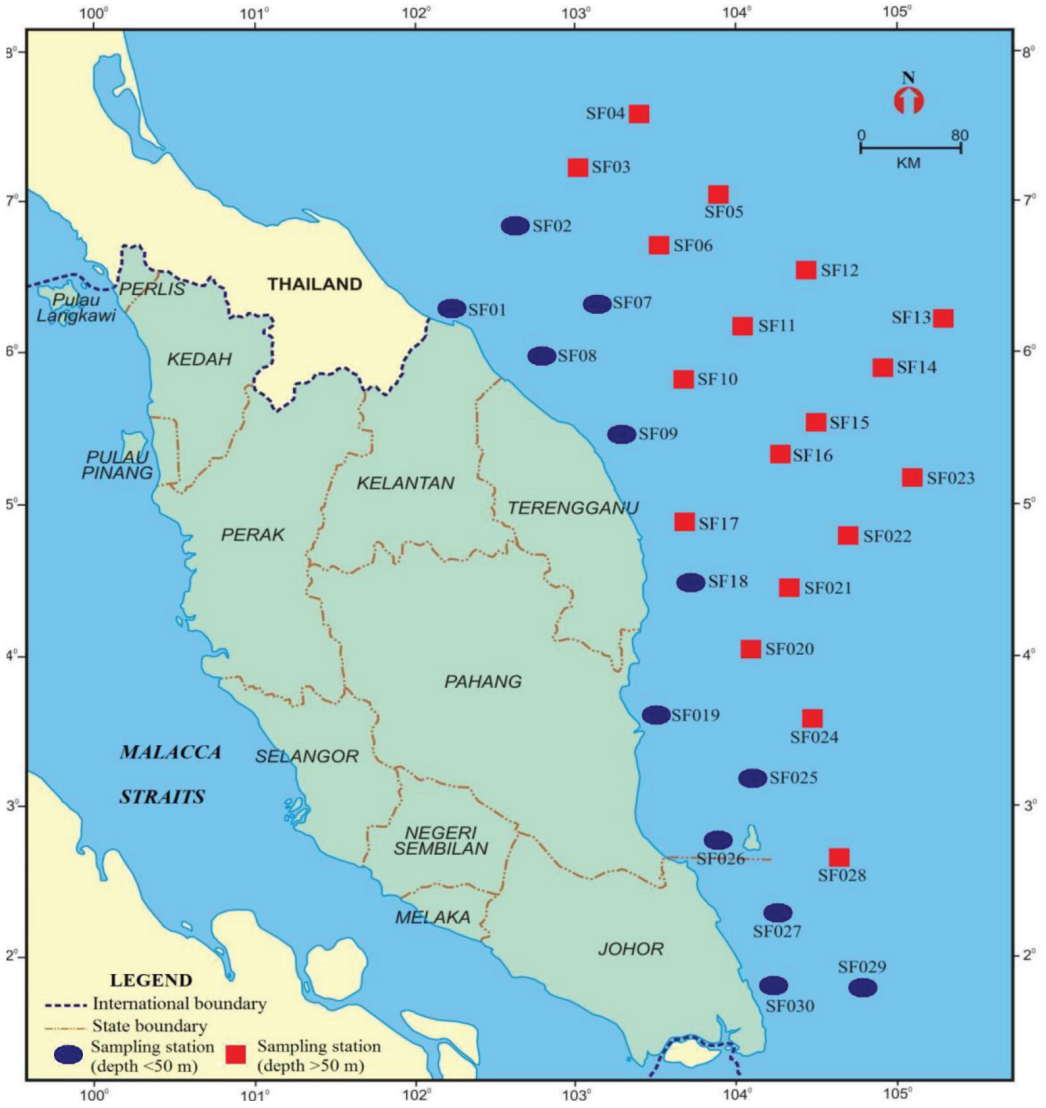


Fig.1: Location of the sampling stations in the South China Sea

RESULTS AND DISCUSSION

Environmental Factors

In marine environments, salinity, substrate types, temperature and depth mostly govern the distribution and diversity of ostracodas (Armstrong & Brasier, 2005). The environmental data and faunal composition are given in Table 1. Station SF11 showed the deepest depth (72 m) and station SF01 had the lowest depth (13 m) from all 30 sampling stations (Table 1). As for water temperature, the maximum value was 30.06°C at station SF01, and the minimum value was 25.24°C at station SF24 (Table 1). For salinity, station SF13 showed the maximum value, i.e. 34.91 ppt. The minimum value for salinity was 27.74 ppt at station SF24 (Table 1). The

TABLE 1
Environmental data and faunal composition of ostracoda at sampling stations

Station	Location		Sampling area	Depth (m)	Temperature (°C)	Salinity (ppt)	Type of sediment	No. of Species	Index of diversity
	Longitude (E)	Latitude (N)							
SF01	102°19'00.3"	06°13'59.0"	Kelantan	13	30.06	33.64	Clayey mud	25	2.6
SF02	102°47'03.2"	06°50'02.5"	Kelantan	47	27.06	34.80	Silty clay	29	2.7
SF03	103°05'03.6"	07°05'01.6"	Kelantan	50	27.00	34.78	Clayey mud	27	2.9
SF04	103°26'01.8"	07°25'59.4"	Kelantan	62	26.93	34.18	Clayey silt	38	3.2
SF05	103°56'03.2"	06°56'07.7"	Kelantan	54	26.58	28.90	Silty sand	33	3.0
SF06	103°35'03.7"	06°42'08.5"	Kelantan	52	26.55	34.75	Silty sand	36	3.0
SF07	103°01'00.1"	06°10'06.1"	Kelantan	45	26.79	34.77	Sandy mud	30	2.9
SF08	102°51'56.7"	05°52'06.6"	Kelantan	34	26.64	34.70	Clayey sand	30	2.7
SF09	103°21'58.6"	05°22'04.6"	Terengganu	47	26.19	34.60	Sandy mud	38	3.0
SF10	103°48'10.0"	05°48'59.9"	Terengganu	55	27.10	34.72	Clayey sand	31	2.8
SF11	104°09'07.8"	06°06'09.4"	Terengganu	72	26.98	34.71	Silty clay	42	3.3
SF12	104°22'06.9"	06°32'01.3"	Terengganu	56	27.34	34.71	Silty clay	29	2.8
SF13	105°16'54.9"	06°16'54.8"	Terengganu	55	29.54	34.91	Sand	27	2.9
SF14	104°58'08.5"	05°57'09.4"	Terengganu	56	29.37	34.64	Sand	29	2.9
SF15	104°29'00.5"	05°29'06.3"	Terengganu	60	28.87	34.86	Sand	29	2.9
SF16	104°12'37.1"	05 18'23.3"	Terengganu	60	29.02	34.89	Sand	21	2.6
SF17	103°42'59.5"	04°54'07.0"	Terengganu	55	27.57	34.68	Sand	37	3.0
SF18	103°49'51.7"	04 29'02.7"	Terengganu	45	28.51	34.47	Sand	31	3.0
SF19	103°41'04.5"	03°37'06.0"	Pahang	24	29.13	34.41	Sand	13	2.1
SF20	104°00'02.6"	03°55'05.1"	Pahang	51	26.29	34.67	Clayey sand	37	3.2
SF21	104°22'04.1"	04°22'05.8"	Pahang	66	25.32	34.76	Clayey silt	38	3.2
SF22	104°38'30.7"	04°44'30.3"	Pahang	63	25.34	34.81	Clayey silt	37	3.2
SF23	105°12'54.1"	05°08'04.1"	Pahang	67	26.41	31.24	Clayey mud	36	3.2

TABLE (Continue)

SF24	104°36'00.0"	03°31'59.6"	Pahang	62	25.24	27.74	Silty sand	37	3.1
SF25	104°09'01.5"	03°09'09.1"	Pahang	41	25.95	34.76	Sand	43	3.2
SF26	103°49'59.5"	02°56'08.8"	Pahang	20	27.72	34.48	Sand	36	3.0
SF27	104°16'59.1"	02°16'57.6"	Johor	30	28.82	34.28	Clayey sand	28	2.7
SF28	104°38'55.9"	02°39'08.5"	Johor	58	25.28	34.78	Silty sand	32	3.1
SF29	104°41'58.7"	02°00'30.1"	Johor	46	27.55	34.44	Sandy mud	30	3.0
SF30	104°15'01.2"	01°48'02.9"	Johor	14	29.20	34.10	Silty mud	24	2.5

nature of the substrate has a pronounced effect on the composition of ostracoda communities. The sediment-inhabiting species live either at the surface of the sediment or within the sediment. The sediment texture in this study area can be categorized as sand, sandy mud, clayey mud, silty mud, silty clay, clayey sand, clayey silt and silty sand (Table 1). Correlation analysis showed that only depth was positively and significantly correlated ($P<0.05$) with the abundance of ostracods, while other parameters (temperature, salinity and type of sediment) did not show any significant correlation.

Abundance and Diversity

A total of 79 species belonging to 44 genera were identified from about 11,148 specimens that had been picked from the sediment samples. Out of these, 77.3% of species belong to Cytheracea, 6.3% to the Cypridacea, 3.8% to Bairdiacea, 10.1% to Cytherellidae, and 2.5% to Polycopidae. All of the ostracods were podocopid and only one species was from the order of myodocopida. Most of the cytheraceans belong to Trachyleberididae (30 species, 38.0%), followed by Cytherellidae (8 species, 10.1%) and Hemicytheridae (6 species, 7.6%). Representatives of these families are typical of infralittoral marine environments around the world. The number of species ranged from 13 to 43 species. The highest species diversity was recorded at station SF25. The H(s) values were from 2.1 to 3.3, with the highest value at station SF11 and the lowest at station SF19 (Table 1). The dominance was from 4.4 to 14.7%. These illustrated the ostracod diversity in the study area. The dominant species was *Foveoleberis cypraeoides* (1137 specimens or 10.20%), followed by *Loxoconcha paiki* (1119 specimens or 10.03%). The selected species from the SEM micrograph are shown in Plate 1.

The abundance and diversity are most related to the depth factor. A correlation analysis showed that depth was positively and significantly correlated ($r=0.650$, $P<0.05$) with Index of Diversity, H(s). The highest value of H(s) lies at the deepest station (Table 1). Meanwhile, the values of H(s) seemed to increase slightly with water depth (Fig.2). A correlation analysis showed that depth was negatively and significantly correlated ($r=-0.726$, $P<0.05$) with dominance, D. The value of dominance decreased with increasing water depth (Fig.3).

Some published reports (Zhao & Whatley, 1989) suggested that depth was also a controlling factor in the abundance and diversity of recent podocopid ostracoda in the south-eastern of Malay Peninsula. More specifically, they found that the Index of Diversity, H(s), slightly increased with depth and the dominance decreased with water depth. The shelf or neritic assemblages occurred between 0 to 200 m depth, and they included many of the marginal marine forms whereas the densest populations were found in the marginal areas, with the highest diversities tended to occur in a shallow-shelf sea. The Ostracod species in marginal marine environments is markedly lower than in non-marine and fully marine ecosystems (Boomer & Eisenhauer, 2002). In high energy shallow waters, both diversity and density of the ostracods were lower than in deeper and more stable environments. The absence or the lowest number of species in shallower samples, closer to the coast, could be mainly due to the instability of the bottom sediment due to the wave action (Whatley *et al.*, 1995; Ramos *et al.*, 1999; Machado *et al.*, 2005).

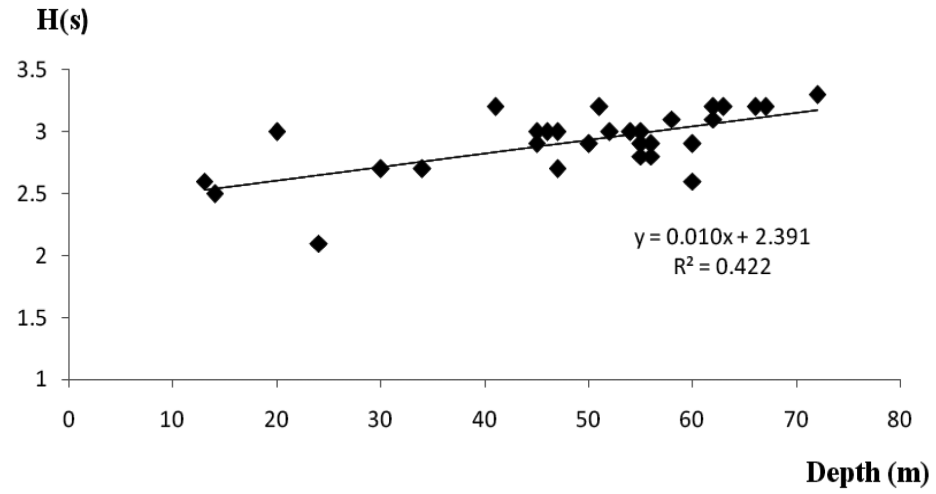


Fig.2: Variations in H(s) with water depth (m)

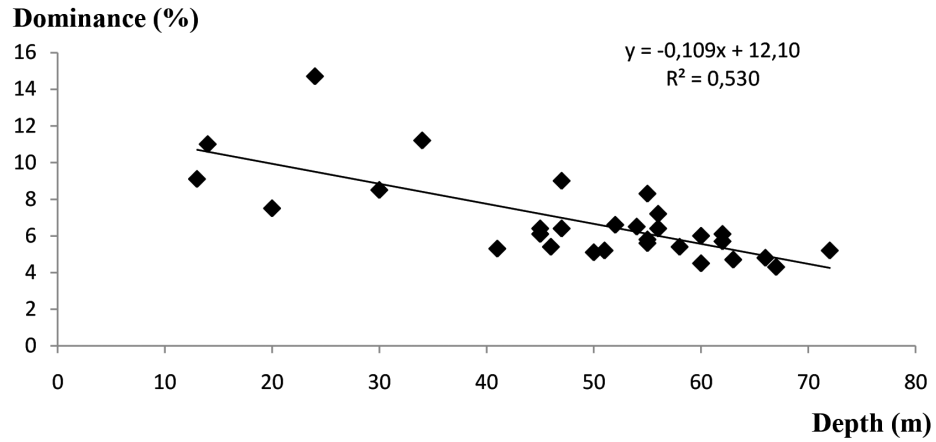


Fig.3: Variations in dominance (%) with water depth (m)

The presence of thick valves with eye spots, strong sculpture, amphidont hinges and conspicuously branched pore canals are the features common in extant shallow-water ostracods from coarse grained substrates such as *Neomonoceratina iniqua* and *Stigmatocythere indica* (Table 2). Deeper water neritic substrates, which also tend to be finer grained, support forms with smooth, thin, often translucent carapaces with relatively weak hinges and no eyes or eye spots (Armstrong & Brasier, 2005). The results also showed that the species *Paracypris* sp. and *Bythocytheropteron alatum* were found in deeper water (Table 2). The Platycopina such as *Cytherella semitalis, Cytherelloidea cingulata, Cytherelloidea leroyi and Cytherelloidea malaccaensis* become an important element of the fauna in the deeper parts of the Jason Bay (Zhao & Whatley, 1989).

Assemblages

In this paper, we use the following terms, 'abundant', 'common' and 'rare' to describe the incidence of one species in its sample. 'Abundant' indicates that one species has a percentage of more than 10% of the total specimens of the samples; 'common' indicates 5 to 10%, and 'rare' indicates less than 5%. Most of the species obtained in the study area are rare. Only two species are abundant (*Foveoleberis cypraeoides* and *Loxoconcha paiki*), while 14 other species are common in the study area (Table 2). Some of these species were found to be widespread throughout the entire study area (*Pistocythereis bradyi, Foveoleberis cypraeoides, Hemicytheridea reticulata, Phlyctenophora orientalis, Loxoconcha paiki* and *Keijella multisulcus*). It was observed that other species displayed a restricted distribution and dominated only in certain areas.

The observations made in the present study revealed two broad assemblages; shallow water fauna (< 50 m depth) comprising *Hemikrithe orientalis, Neomonoceratina iniqua, Stigmatocythere indica, Cytherelloidea leroyi* and *Neocytheretta snellii* (Table 2) and the deep water (> 50m depth) assemblage was distinguished by *Paracypris* sp.*, Alataconcha pterogona, Bythocytheropteron alatum, Keijella paucipunctata,* and *Actinocythereis scutigera* (Table 2). Earlier reports (Zhao & Whatley, 1989) on recent podocopid ostracoda from south-eastern Malay Peninsula found that the dominant species (*Hemicytheridea reticulata, Neomonoceratina delicata, Neomonoceratina iniqua* and *Lankacythere coralloides*) were ubiquitous and most abundant in all the open sea samples. The common species (*Actinocythereis scutigera, Phlyctenophora orientalis* and *Stigmatocythere roesmani*) occurred more frequently in deeper water (10 to 20 m) compared to *Keijella jankeiji* which was confined to shallow water (0 to 7 m).

TABLE 2
Distribution of the species in the South China Sea

SPECIES	STATION (SF)																													
	< 50 m												> 50 m																	
	1	2	7	8	9	18	19	25	26	27	29	30	3	4	5	6	10	11	12	13	14	15	16	17	20	21	22	23	24	28
Loxoconcha paiki	C	C	C	R	R	R	R	C	R	R	C	R	R	C	A	A	C	A	R	R	C	C	R	C	C	C	C	R	A	C
Foveoleberis cypraeoides	R	C	C	C	R	C	R	C	C	R	A	R	C	R	A	A	A	A	R	C	R	C	C	A	C	R	R	C	C	R
Pistocythereis bradyi	C	C	C	-	R	R	-	C	R	R	C	C	R	C	R	R	C	C	-	C	A	R	A	A	R	R	R	C	C	R
Phlyctenophora orientalis	R	C	R	-	R	C	-	C	R	C	C	R	C	R	-	C	C	A	R	C	A	C	C	C	R	R	C	R	C	R
Hemicytheridea reticulata	C	R	C	R	R	C	R	R	-	R	A	R	R	C	R	R	C	A	-	C	C	C	A	C	R	R	C	R	R	R
Keijella multisulcus	R	C	R	-	-	A	C	C	R	C	R	C	R	C	R	C	R	C	-	R	R	C	A	R	R	R	R	C	A	R
Hemikrithe orientalis	C	A	C	C	R	A	C	A	C	C	C	R	R	R	-	R	R	-	-	R	R	R	R	R	-	-	R	R	R	-
Neomonoceratina iniqua	R	C	C	R	C	A	C	A	C	-	C	C	-	R	R	R	R	R	R	R	R	-	-	R	R	R	R	-	R	R
Stigmatocythere indica	C	C	R	C	C	C	A	R	C	C	R	C	R	-	R	-	R	-	-	R	R	R	R	R	R	R	R	-	-	R
Cytherelloidea leroyi	C	A	R	R	C	A	C	C	C	R	C	C	R	R	R	R	-	-	-	R	R	-	R	R	-	-	R	R	R	-
Neocytheretta snellii	C	C	C	C	R	C	C	C	A	A	A	R	R	R	-	-	R	R	-	-	-	R	R	-	R	R	-	-	-	C
Actinocythereis scutigera	R	R	R	R	-	C	-	R	R	-	C	R	C	C	R	C	R	R	R	C	C	C	A	C	R	R	R	-	R	-
Keijella paucipunctata	C	C	C	R	R	-	-	R	R	-	-	-	R	C	C	C	R	C	R	C	C	C	A	R	-	-	R	C	C	-
Bythocytheropteron alatum	R	R	R	-	R	R	R	R	R	-	R	-	C	R	C	C	R	C	C	C	A	C	A	C	R	R	C	C	C	-
Alataconcha pterogon	-	-	R	-	-	-	-	R	-	-	-	-	R	C	C	R	R	R	R	C	C	R	R	C	-	R	C	C	C	-
Paracypris sp.	R	-	-	-	-	R	-	-	-	-	-	-	C	C	R	R	-	R	C	C	C	C	R	R	-	-	-	R	C	-

LEGEND
A – abundant (more than 10%) C – common (5 to 10%) R – rare (less than 5%)

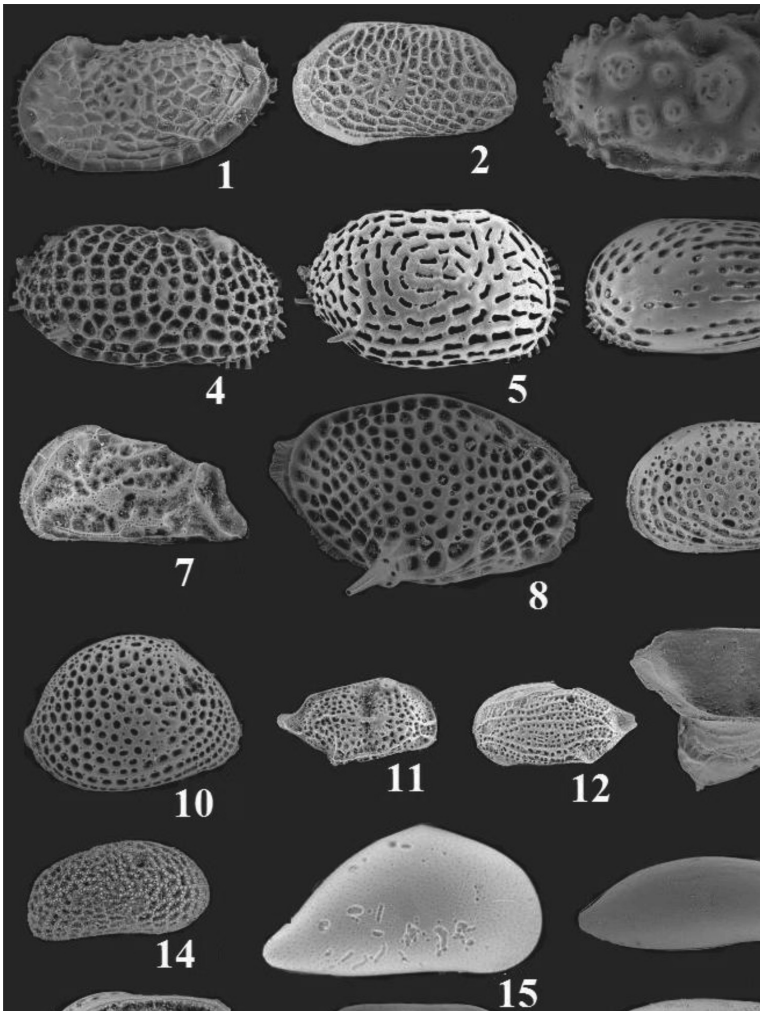


Fig.4: 1.) *Neocytheretta murilineata* (Zhao & Whatley, 1989); left valve, x400, sample station SF01; 2.) *Hemikrithe orientalis* (Van den Bold, 1950); left valve, x400, sample station SF03; 3.) *Actinocythereis scutigera* (Brady, 1868); right valve, x200, sample station SF04; 4.) *Pistocythreis bradyi* (Ishizaki, 1968); right valve, x300, sample station SF03; 5.) *Lankacythere euplectella* (Brady, 1869); right valve, x300, sample station SF03; 6.) *Keijella multisulcus* (Whatley & Zhao, 1988); left valve, x400, sample station SF04; 7.) *Caudites scopulicolus* (Hartmann 1981); left valve, x400, sample station SF01; 8.) *Alataconcha pterogona* (Zhao, 1985); right valve, x200, sample station SF03; 9.) *Loxoconcha paiki* (Whatley & Zhao, 1987); left valve, x400, sample station SF10;10.) *Foveoleberis cypraoeides* (Brady, 1868); right valve, x300, sample station SF10;11.) *Neomonoceratina macropora* (Kingma, 1948); right valve, x500, sample station SF01;12.) *Semicytherura contraria* (Zhao & Whatley, 1989); left valve, x500, sample station SF01;13.) *Bythocytheropteron alatum* (Whatley & Zhao, 1987); right valve, x400, sample station SF04;14.) *Corallicythere* sp., right valve, x400, sample station SF01; 15.) *Propontocypris* sp., right valve, x400, sample station SF10; 16.) *Paradoxostoma* sp., right valve, x300, sample station SF04; 17.) *Cytherelloidea cingulata* (Brady, 1869), left valve; x300, sample station SF10; 18.) *Cytherella semitalis* (Brady, 1868); right valve, x400, sample station SF07; Fig.19.) *Cytherella* sp., right valve, x500, sample station SF04.

Comparison

A comparative analysis of the data collected in the present study with the adjacent area showed a high degree of resemblance with south-eastern Malay Peninsula fauna, South China Sea (Zhao & Whatley, 1989). Out of the 79 species obtained, 51 were found to be common to both areas. Among these, a few species were found to be widespread (*Pistocythereis bradyi, Hemicytheridea reticulata* and *Phlyctenophora orientalis*). The above species were also found to be reported earlier from the Straits of Malacca (Whatley & Zhao, 1987; 1988), Sunda Shelf (Mostafawi, 1992), and Java Sea (Dewi, 2000).

The observations made revealed that a large number of species were common between the study area and south-eastern Malay Peninsula. However, the dominant species was different and this could probably be due to the difference in water depth in the study area. The dominant species in the south-eastern Malay Peninsula were *Hemicytheridea reticulata, Neomonoceratina bataviana, Neomonoceratina delicata* and *Lankacythere coralloides*. The abundant species included *Actinocythereis scutigera, Alocopocythere kendengensis, Keijella jankeiji, Keijella papuensis, Parakrithella pseudadonta, Phlyctenophora orientalis, Pistocythereis bradyi, Stigmatocythere roesmani* and *Tanella gracilis* (Zhao & Whatley, 1989). The present study reported six species (*Pterygocythereis* sp., *Hemicytherura* sp., *Cytherella* sp., *Caudites* sp., *Polycope orbulinaeformis* and *Trachyleberis* sp.) as new records in Malaysian waters.

CONCLUSION

The recognised ostracod fauna constituted by a total of 79 species belonging to 44 genera were identified from the study area. A shallower water (< 50 m) fauna is characterized by *Hemikrithe orientalis, Neomonoceratina iniqua, Stigmatocythere indica, Cytherelloidea leroyi* and *Neocytheretta snellii,* while a deeper water (>50 m) fauna is distinguished by *Paracypris* sp., *Alataconcha pterogona, Bythocytheropteron alatum, Keijella paucipunctata* and *Actinocythereis scutigera*. However, some of these fauna (*Pistocythereis bradyi, Feveoleberis cypraeoides, Hemicytheridea reticulata, Phlyctenophora orientalis, Loxoconcha paiki* and *Keijella multisulcus*) were widespread throughout the entire study area. The analysis and comparison showed that a high degree of faunal replication between the study area and south-eastern Malay Peninsula (South China Sea). A total of 51 species are common to the two areas, with relative abundance and distribution (*Pistocythereis bradyi, Hemicytheridea reticulata* and *Phlyctenophora orientalis*).

ACKNOWLEDGEMENTS

This research was funded by Universiti Kebangsaan Malaysia through FRGS grant (Code Project: UKM-ST-08-FRGS0241-2010). The authors wish to thank the SEM Unit, the Faculty of Science and Technology, UKM for making SEM photographs.

REFERENCES

Armstrong, H., & Brasier, M. D. (2005). *Microfossils* (2nd Edition). UK: Blackwell Publishing. p. 296.

Boomer, I., & Eisenhauer, G. (2002). Ostracod faunas as palaeoenvironmental indicators in marginal marine environments. *The Ostracoda: Applications in Quaternary Research Geophysical Monograph, 131*, 135-149.

Cohen, A. C., Martin, J. W., & Kornicker, L. S. (1998). Homology of Holocene ostracode biramous appendages with those of other crustaceans: the protopod, epipod, exopod and endopod. *Lethaia, 31*, 251-265.

Dewi, K. T. (2000). Distribution of ostracoda from South of Tanjung Selatan, South Kalimantan. *Marine Geological Institute Bulletin, 15*(1), 1-14.

Folk, R. L. (1980). *Petrology of sedimentary rocks*. Texas: Hemphill Publishing. p. 184.

Gou, Y., Zheng, S., & Huang, B. (1983). Pliocene ostracode fauna of Leizhou Peninsula and northern Hainan Island, Guangdong Province. *Paleontologica Sinica, 18*, 1-134.

Gou, Y. (1990). Recent ostracoda from Hainan Island, South China Sea. *Courier Forschungsintitut Senckenberg, 123*, 19-36.

Horne, D. J., Cohen, A., & Martens, K. (2002). Taxonomy, Morphology and Biology of Quaternary and living Ostracoda. *The Ostracoda: Applications in Quaternary Research Geophysical Monograph, 131*, 5-36.

Machado, C. P., Coimbra, J. C., & Carreno, A. L. (2005). The ecological and zoogeographical significance of the sub-Recent ostracoda off Cabo Frio, Rio de Janeiro State, Brazil. *Marine Micropaleontology, 55*, 235-253.

Mostafawi, N. (1992). Recent ostracods from the central Sunda Shelf, between the Malay Peninsula and Borneo. *Senckenbergiana Lethaea, 72*, 129-168.

Ramos, M. I. F., Coimbra, J. C., Whatley, R. C., & Moguilevsky, A. (1999). Taxonomy and ecology of the family Cytheruridae (Ostracoda) in recent sediments from the Northern Rio de Janeiro coast, Brazil. *Journal of Micropaleontology, 18*, 1-16.

Whatley, R. C., & Zhao, Q. (1993). The Krithe problem: a case history of the distribution of *Krithe* and *Parakrithe* (Crustacea, Ostracoda) in the South China Sea. *Palaeogeography, Palaeoclimatology, Palaeoecology, 103*, 281-297.

Whatley, R. C., & Zhao, Q. (1987). The recent ostracoda of Malacca Straits (Part I). *Revista Espanola de Micropaleontologia, 19*(3), 327-366.

Whatley, R. C., & Zhao, Q. (1988). The recent ostracoda of Malacca Straits (Part II). *Revista Espanola de Micropaleontologia, 20*(1), 5-37.

Whatley, R. C., Toy, N., Moguilevsky, A., & Coxill, D. (1995). Ostracoda from the South west Atlantic: Parte 1. The Falkland Island. *Revista Espanola de Micropaleontologia, 27*(1), 17-38.

Zhao, Q., Wang, P., & Zhang, Q. (1985). Ostracoda in bottom sediments of the South China Sea off Guangdong Province, China: their taxonomy and distribution. In P. Wang (Ed.). *Marine Micropaleontology of China* (pp. 196-217). Beijing: China Ocean Press.

Zhao, Q., & Whatley, R.C. (1989). Recent podocopid ostracoda of the Sedili River and Jason Bay, southeastern Malay Peninsula. *Micropaleontology, 35*(2), 168-187.

Zhao, Q. (2005). Late Cainozoic ostracod faunas and paleoenvironmental changes at ODP site 1148 South China Sea. *Marine Micropaleontology, 54*, 27-47.

Pertanika J. Sci. & Technol. 22 (2): 445 - 456 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Empirical Correlation of Refrigerant HC290/HC600a/ HFC407C Mixture in Adiabatic Capillary Tube Using Statistical Experimental Design

Shodiya, S.[1,2]*, Azhar, A. A.[1] and Darus, A. N.[1]

[1]*Automomotive Development Centre (ADC), Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*

[2]*Department of Mechanical Engineering, Faculty of Engineering, University of Maiduguri (UNIMAID), Maiduguri, Borno, Nigeria*

ABSTRACT

HCFCs, in addition to destroying the ozone layer, have been recognized as a contributing factor that increases global warming. It is widely used as working fluid in window air-conditioning system, where capillary tube serves as an expansion device. Literature reports have shown that no single refrigerant can solve the problem of ozone layer depletion and global warming. Refrigerant HC290/HC600a/HFC407C mixture, an eco-friendly refrigerant, has been recognized as an alternative to HCFC22. The objective of this study is to, for cost effectiveness, develop an empirical correlation to predict the refrigerant HC290/HC600a/HFC407C mixture mass flow rate using statistical experimental design approach. A review of relevant literature shows that refrigerant's mass flow rate depends on condensing temperature, degree of subcooling, inner diameter and length of capillary tube. The relationship between the mass flow rate and the four independent variables was established as an empirical mathematical correlation using central composite design (CCD), a response surface methodology (RSM). This empirical correlation was examined using analysis of variance (ANOVA) of 5% level of significance. The results of these analysis showed that the correlation fitted well with the experimental data yielding an average and standard deviation of 1.05% and 2.62%, respectively. The validity of the present correlation was further assessed by comparing it with published empirical correlation in literature and the result showed that the present correlation is consistent.

Keywords: Ozone layer, air-conditioning, capillary tube, refrigerant, empirical correlation

Article history:
Received: 2 February 2012
Accepted: 18 March 2013

Email addresses:
Shodiya, S. (sulaimonshodiya@yahoo.com), Azhar, A. A. (azhar@fkm.utm.my), Darus, A. N. (amer@fkm.utm.my)
*Corresponding Author

ISSN: 0128-7702

INTRODUCTION

To some extent, most people are aware of the consequences of ozone layer depletion and global warming, and thus, changes that address these issues are of utmost importance. Efforts to eliminate non-eco-friendly refrigerants such as chlorofluorocarbon (CFCs) and hydro chlorofluorocarbons (HCFCs) in refrigeration and air-conditioning industries are ongoing. In modern refrigeration and air-conditioning systems, capillary tube, a drawn copper tube with diameter ranging from 0.5 to 2 mm and length from 1.5 to 6 m, is often used as an expansion and metering device. This is due to its simplicity in construction, low cost, no moving part and requires small starting torque of the compressor. Capillary tube has a simple geometric configuration but refrigerant flow in it is complex. The performance of a refrigeration system depends on appropriate selection of length and inner diameter of the capillary tube for a given input conditions so that the desired mass flow rate can be achieved. Design of refrigeration system using a new refrigerant requires a fresh study of this alternative refrigerant flow in the capillary tube because of the differences in the thermophysical properties of the conventional and the new refrigerants.

The refrigerant behaviour through capillary tube has been studied extensively in both numerical and experimental approach. In the numerical approach, three different types of models are commonly used – two-phase homogenous, separated flow and drift flux models. In a homogenous model, the two-phase liquid/vapour mixture is simulated using mean refrigerant properties (Bansal & Rupasinghe, 1998; Kritsadathikarn *et al.*, 2002; Kumar *et al.*, 2009; Sami & Tribes, 1998; Shodiya *et al.*, 2011b; Wongwises & Chingulpitak, 2010; Wongwises & Pirompak, 2001; Wongwises & Suchatawut, 2003; Zhou & Zhang, 2006; Kim *et al.*, 2011; Shodiya *et al.*, 2012). Likewise, in the separated flow model, the slip that exists between the liquid and vapour is considered and a mixture variable called void fraction is introduced to the conservation equations (Wongwises *et al.*, 2000). In the drift flux model, the entire two-phase liquid-vapor mixture is considered when formulating the conservation equations (Liang & Wong, 2001).

Practically speaking, these numerical simulations that were used for analyzing and designing capillary tubes requires some computer programming skills by the users, as a result, it has not been widely used by the capillary tube engineering designers. In order to simplify these models, many researchers (Yilmaz & Unal, 1996; Ding *et al.*, 1999; Zhang & Ding, 2001; Hermes *et al.*, 2010) developed an approximate analytic model for the analysis and designing of the tube. Though, an appreciable success was recorded, but, their models resulted in an iterative calculation, as such, the analytic model is still not an explicit solution.

Unlike the numerical simulation and approximate analytic models, empirical correlations are simple to operate and more convenient for the capillary tube engineering designers. As such, several investigators carried out studies in order to develop empirical correlation that can be used to predict refrigerant mass flow rate in capillary tubes (Choi *et al.*, 2004; Choi *et al.*, 2003; Fiorelli *et al.*, 2002; Jabaraj *et al.*, 2006; Kim *et al.*, 2002; Park *et al.*, 2007; Shodiya *et al.*, 2011a; Vinš & Vacek, 2009; Zhou & Zhang, 2006). In developing these correlations, experimental data of various refrigerants (eco-friendly and non-eco-friendly) and a dimensionless parameter based on Buckingham pi theorem were used.

All the aforementioned empirical correlations do not use statistical experimental design to investigate the behaviour of refrigerant flow in capillary tube and develop their correlations. Statistical experimental design is economical because it requires a relatively small number of experiments and yet able to be analyzed by using statistical methods to yield results in valued and objective conclusions. This particular approach of experimental design is very important if meaningful conclusions are to be drawn from the data. A statistically designed approach is a systematic and scientific approach for planning and analyzing data when a problem that involves experimental errors is considered (Montgomery, 2005). Central composite design (CCD), which is a Response surface method (RSM), is one of the statistical designs of the experimental methods. It is a combination of the statistical and mathematical techniques used for planning and analyzing problems in which the output variables (responses) are influenced by input factors and the goal is to optimize the response(s). Using RSM, the interactions between the input factors and the response(s) can be established (Montgomery, 2005).

Research applying statistical experimental design to develop an empirical correlation to predict the refrigerant mass flow rate in capillary tubes is rather limited (see Bittle & Pate, 1996; Bittle *et al.*, 1995; Melo *et al.*, 2002). For example, Melo *et al*. (2002) performed an experiment using statistical factorial design of experiment on concentric diabatic capillary tube with R600a as working fluid. They proposed empirical correlations of refrigerant mass flow rate based on their experimental results. The empirical correlation is in good agreement with their experimental data.

However, the refrigerants used in those studies are mostly synthetic and contain chlorine and fluorine that cause depletion of ozone layer and lead to global warming. The major objective of this paper is to develop a new empirical correlation with new refrigerant mixture containing 20% HC (HC600a and HC290) and HFC407C to predict refrigerant mass flow rate through adiabatic capillary tube using statistical design of experiment. It should be noted that the main advantage of this design of experiment is to obtain maximum information with a minimum amount of experiment performed.

CORRELATION DEVELOPMENT

Data Source

As observed in some previous studies (see Li *et al.*, 1990; Lin *et al.*, 1991; Melo *et al.*, 1992), the refrigerant mass flow rate (m) in capillary tube depends on the length of capillary tube (L), inner diameter (D), condensing temperature (T_{cond}) and degree of subcooling (T_{sub}). Though the refrigerant properties (thermodynamic and transport) are also important in predicting the refrigerant mass flow rate, the effects of these properties have been to some extent taken care of by condensing temperature and degree of subcooling. As a result, these four factors are considered as independent variables and mass flow rate as a dependent variable. Design expert (version 7.1, Stat-Ease, Inc., Minneapolis, USA), an experimental design software, was used for this study. In order to reduce the experiment to be performed, which is the main advantage of this approach, statistical technique, central composite design (CCD), the most popular RSM design, was used to construct the test matrix. A four factor CCD rotatable option (four factors: 4! =24), with six replicate at the centre point, was employed in designing the experimental

test. A CCD rotatable option is applicable to design factors less than six. The centre points are usually repeated 4-6 times to get a good estimate of experimental error. In this study, six centre points were used to achieve an optimal performance. The total experimental data points that are now required to complete the experiment are 30 (24 from factors and 6 from centre points). In order to determine the numerical values of each design variable to be used to perform the experiment, a two-level alpha (low and high levels) was used. The lowest and highest values of each design parameter were input into the design expert software. Thereafter, the software generated the experiment to be performed. The ranges and level of factors used are shown in Table 1.

TABLE 1
Experimental range and level of central composite design

Variables	Range and levels				
	-2	-1	0	+1	+2
Condenser temperature, $T_{cond.}$ (°C)	37	42	44.5	47	52
Degree of subcooling, $T_{sub.}$(°C)	2	5	8	11	14
Length of capillary tube, L (m)	0.75	1.00	1.25	1.5	1.75
Diameter of capillary tube, D (mm)	1.12	1.19	1.27	1.33	1.40

RESULTS AND DISCUSSION

The design expert software (version 7.1, Stat-Ease, Inc., Minneapolis, USA) was used for the regression and graphical analysis of the data. Considering the values of each design parameter generated from the software, the responses were carefully selected from the experimental data of Jabaraj *et al*. (2006). The total experimental data points of Jabaraj *et al*. (2006) were 200. In selecting the experimental data points, some of these points were not directly available in the experimental data of Jabaraj *et al.* (2006). As a result, the empirical model developed by Jabaraj *et al.* (2006) from their experimental data was used to determine these points. The method of using the existing model to generate data to be used for developing new correlations has been used by many researchers (Bansal & Rupasinghe, 1996; Wongwises & Trisaksri, 2003; Shodiya *et al.*, 2011a; Sarker & Jeong, 2012).

The measured mass flow rates and their associated input parameters according to the experimental design are given in Table 2. The application of RSM with the design expert software yielded the following regression equation, which is an empirical relationship between the mass flow rate and the input variables given in Eq. 1. In developing this equation, the software used least square formulation to minimize the errors between the mass flow rate calculated by the regression equation and the measured data.

$$\begin{aligned} m = & -93.7361 + 2.4854T_{cond} + 0.8050T_{sub} - 10.4512L + 64.3006D \\ & - 0.005241T_{cond}T_{sub} + 0.03114T_{cond}L - 0.1925T_{cond}D \\ & + 0.03217T_{sub}L + 0.1329T_{sub}D + 2.9701LD - 0.02314T_{cond}^2 \\ & - 0.006352T_{sub}^2 + 0.1430L^2 - 16.2679D^2 \end{aligned} \quad (1)$$

TABLE 2
CCD response result for four input factors

Run	Condensing Temperature (°C) – Factor 1	Degree of Subcooling (°C) – Factor 2	Capillary tube Length (m) – Factor 3	Capillary tube Diameter (mm) – Factor 4	Mass flow rate (g/s) -Response
1	44.5	8	1.25	1.27	12.26
2	44.5	8	1.25	1.40	14.84
3	47.0	11	1.75	1.33	12.05
4	42.0	11	1.75	1.33	11.32
5	42.0	11	1.75	1.19	8.89
6	44.5	8	1.25	1.12	9.68
7	44.5	8	1.75	1.27	10.72
8	47.0	5	1.50	1.12	7.29
9	44.5	8	1.25	1.27	12.26
10	47.0	11	1.00	1.27	14.36
11	42.0	5	1.00	1.12	8.76
12	42.0	11	1.00	1.27	13.61
13	47.0	5	1.50	1.12	7.29
14	44.5	14	1.25	1.27	14.44
15	44.5	2	1.25	1.27	10.39
16	47.0	5	0.75	1.33	14.62
17	44.5	8	1.25	1.27	12.26
18	42.0	5	1.75	1.19	6.47
19	42.0	5	1.75	1.33	8.90
20	44.5	8	1.25	1.27	12.26
21	44.5	8	0.75	1.27	15.12
22	47.0	11	0.75	1.19	14.69
23	44.5	8	1.25	1.27	12.26
24	42.0	11	1.25	1.19	10.81
25	37.0	8	1.25	1.27	9.96
26	42.0	5	1.00	1.40	13.54
27	47.0	5	1.00	1.40	14.58
28	47.0	11	1.50	1.12	9.57
29	52.0	8	1.25	1.27	12.38
30	44.5	8	1.25	1.27	12.26

The fitting of the model can be checked using several criteria. The Analysis of Variance (ANOVA) for the empirical correlation model is summarized in Table 3. A model is significant at 95% confidence level if Fisher F-test has a probability value (Prob > F) less than 0.05. The F-test for lack of fit (LOF) describes the deviation of the actual points from the fitted surface, relative to pure error (Anderson & Whitcomb, 2005). Preferable is a large value of Prob > F for LOF which is greater than 0.05. For coefficient of determination (R^2), a higher value is preferred and a reasonable agreement with adjusted R^2 is crucial (Ghafari *et al.*, 2009). Adequate precision (AP) can be defined as a measure of experimental signal to noise ratio (Anderson & Whitcomb, 2005). AP that exceeds 4 usually indicates that the model will give

reasonable performance in prediction. PRESS is the prediction error sum of squares which is a measure of how well the model for the experiment is likely to predict the response in the new experiment. The standard deviation (SD), coefficient of variance (CV), and PRESS values are preferred to be small (Montgomery, 2005).

As shown in Table 3, the empirical correlation regression model (Eq.1) is a significant model since its Prob>F value is less than 0.05. In addition, AP, SD and PRESS for the regression model are satisfactory since the value of AP is more than 4. Similarly, SD, CV and also PRESS values are also small (see Table 3).

TABLE 3
ANOVA results for mass flow rate response parameter

Response	Model F Value	Prob >F	LOF Prob >F	R^2	Adjusted R^2	AP	SD	CV	PRESS
Mass flow rate	202.99	0.0001	0.038	0.9705	0.9654	69.92	0.0448	1.05	0.0528

Fig.1 compares the mass flow rate experimental data of refrigerant with the predicted data obtained from equation 1. The value of R^2 for the predicted mass flow rate model is 0.9705, showing that the model is adequate enough to explain most of the variability of the experimental data. The absolute deviation or the difference between the calculated mass flow rate and the measured one is about 0.24g/s.

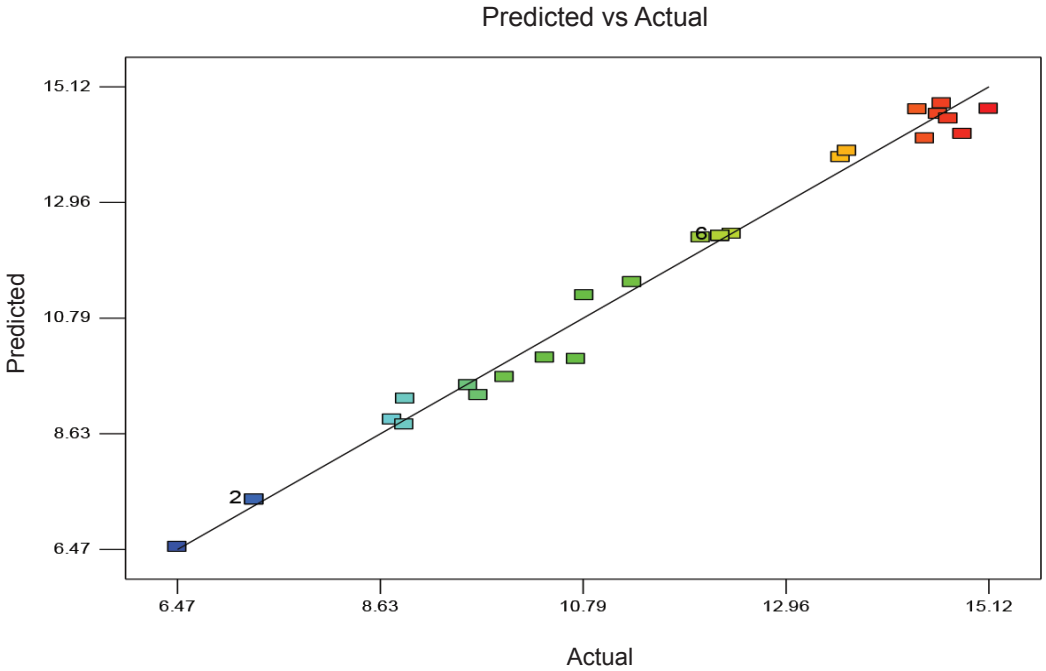


Fig.1. Predicted vs. actual values plot for refrigerant mass flow rate

To further validate Eq. (1), its mass flow rate predictions and the predictions from Jabaraj *et al*.'s (2006) empirical correlation was compared with the measured data as shown in Fig.2. The figure reveals that these two correlation predictions of mass flow rates are in good agreement with the measured data within a deviation of ±5%. About 93% of these predicted mass flow rates fall within this range. This assessment shows that the present correlation is consistent with the correlation of Jabaraj *et al*. (2006).

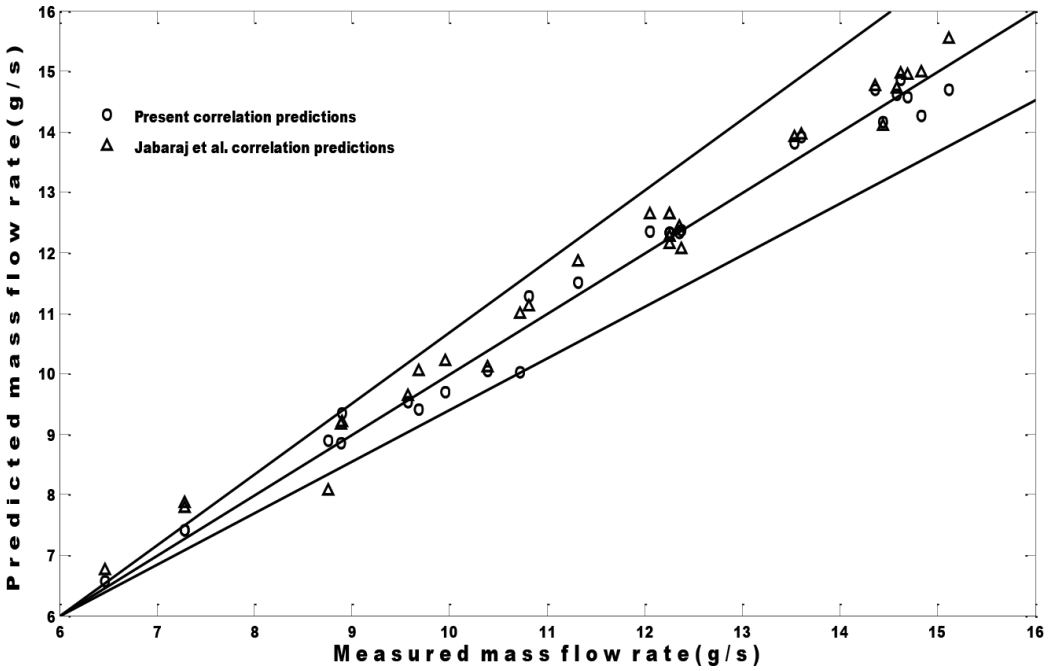


Fig.2. A comparison of the mass flow rate predictions with the measured data

The fitted model from equation (1) can be used to empirically determine the response function over the experimental region. The contour plot helps in assessing the effect of any two variables in combination on the response quality. A 3-D graphical surface plot illustrates the response value of the dependent variables.

It is a well-known fact in literature that the refrigerants mass flow rates increase with the increases in condensing temperature and degree of subcooling. The mass flow rate is strongly dependent on the length and inner diameter of the capillary tube. Fig.3 shows the effects of condensing temperature and degree of subcooling on mass flow rate. The optimum point is obtained on the coordinates of the curve lines (Fig.3). Fig.4 also shows the 3-D graphical surface plot of mass flow rate with the degree of subcooling and condensing temperature. In this figure, there is an increase in mass flow rate as the degree of subcooling increases. This is due to the fact that increased subcooling leads to increased liquid phase, and as a result, an increase in the mass flow rate. The liquid phase offers lesser resistance to the refrigerant flow compared to the vapour/ liquid phase in the capillary tube.

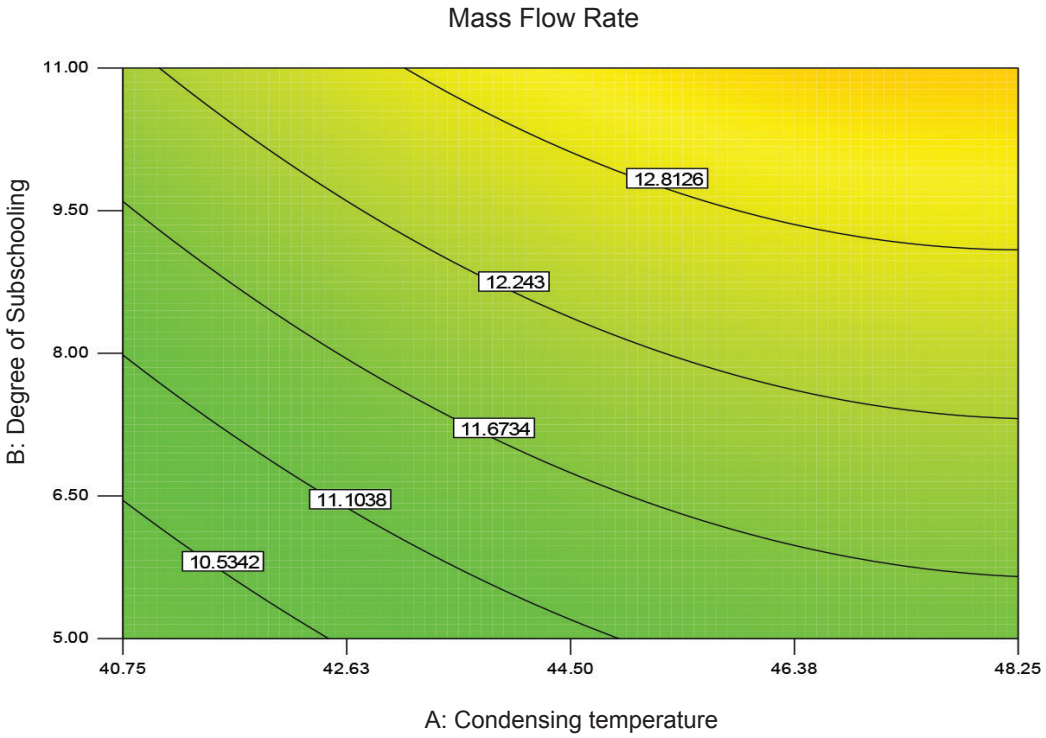


Fig.3: A contour surface plot of mass flow rate as a function of degree of subcooling and condensing temperature

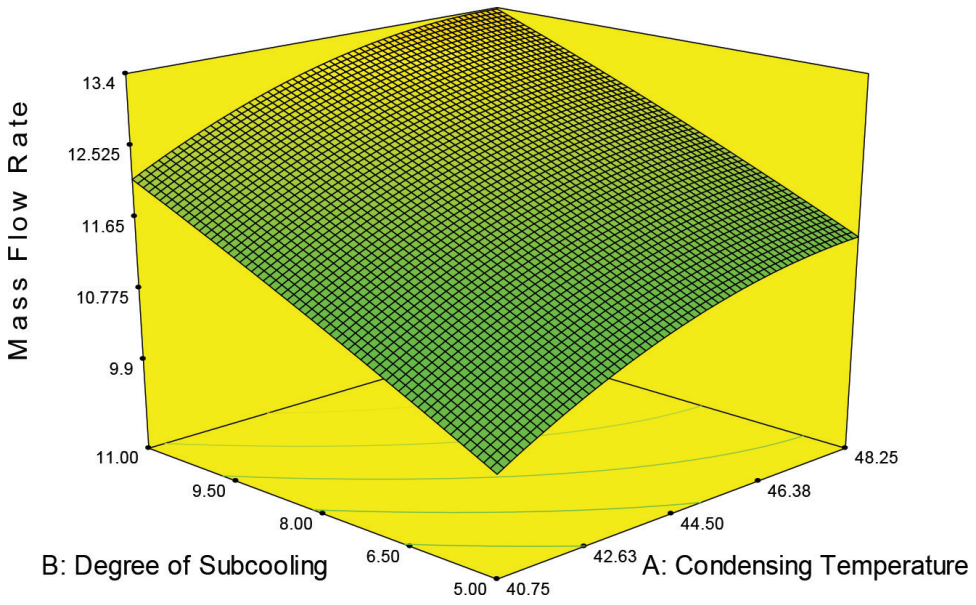


Fig.4. A 3-D graphical surface plot of mass flow rate versus degree of subcooling and condensing temperature

Fig.5 shows the 3-D graphical surface plot of mass flow rate for capillary tube length and diameter. It can be clearly seen in the figure that the mass flow rate increases with the increase in tube diameter and also increases with a decrease in tube length. This may be due to the fact that in a shorter capillary tube length, there is a smaller effect of wall frictional force on the refrigerant.

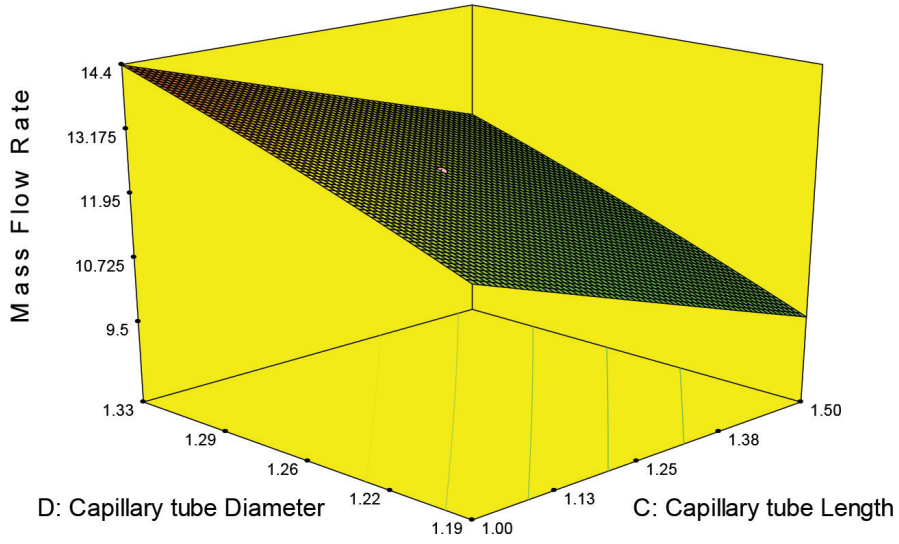


Fig.5. A 3-D graphical surface plot of mass flow rate versus capillary tube diameter and length

CONCLUSION

An experimental statistical design of response surface methodology has effectively been used in planning, analyzing and optimizing the four independent factors and assessing their effects on the response variable - refrigerant mass flow rate. The mass flow rate correlation of refrigerant HFC407C/HC600a/HC290, a new refrigerant mixture developed, showed that the mass flow rate strongly depends more on capillary tube length and diameter. In addition, the reliability and adequacy of this empirical correlation were also evaluated using the analysis of variance (ANOVA) and the results showed that this model had given a good estimation of measured refrigerant mass flow rate. The average deviation between the calculated mass flow rate and the measured was about – 1.05%. Based on the results of these comparisons, it can be concluded that, for cost effectiveness, statistical experimental design approach is an effective tool for developing refrigerant mass flow rate in capillary tubes. Though the empirical correlation developed in this study is in good agreement with the measured data and previous correlation, it should be applied for the specified refrigerant of this study.

ACKNOWLEDGEMENTS

The present study was financially supported by Research University Grant (RUG) programme: Tier 2 Cost Centre Code: QJ130000.7124.01J85, Universiti Teknologi, Malaysia. The financial support of Educational Trust Fund (ETF), Nigeria, is also acknowledged.

REFERENCES

Anderson, M. J., & Whitcomb, P. J. (2005). *RSM Simplified*. New York: Prductivity Press.

Bansal, P., & Rupasinghe, A. (1996). An empirical model for sizing capillary tubes. *International Journal of Refrigeration, 19,* 497-505.

Bansal, P. K., & Rupasinghe, A. S. (1998). An homogeneous model for adiabatic capillary tubes. *Applied Thermal Engineering, 18,* 207-219.

Bittle, R. R., & Pate, M. B. (1996). An experimental evaluation of capillary tube–suction line heat exchanger performance with refrigerant HFC-134a. *Proceeding of the ASME, Advanced Energy Systems Division, AES, 36,* 87–100.

Bittle, R. R., Stephenson, W. R., & Pate, M. B. (1995). An experimental evaluation of capillary tube-suction line heat exchanger performance with R-152a. *ASHRAE Transactions, 100*(Part 1), 124-135.

Choi, J., Kim, Y., & Chung, J. T. (2004). An empirical correlation and rating charts for the performance of adiabatic capillary tubes with alternative refrigerants. *Applied Thermal Engineering, 24,* 29–41.

Choi, J., Kim, Y., & Kim, H. Y. (2003). A generalized correlation for refrigerant mass flow rate through adiabatic capillary tubes. *Int. J. Refrig. , 26,* 881–888.

DING *et al.* (1999). An approximate analytic model for flow through capillary tubes. *Chin. Sci. Bull, 44,* 668–670.

Fiorelli, F. A. S., Huerta, A. A. S., & Silvares, O. M. (2002). Experimental analysis of refrigerant mixtures flow through adiabatic capillary tubes. *Experimental Thermal Fluid science, 26,* 499-512.

Ghafari, S., Abdul Azis, H., Isa, M. H., & Zinatizadeh, A. (2009). Application of response surface methodology (RSM) to opimize coagulation-flocculation treatment of leachate using aluminium chloride (PAC) and alum. *Journal of Hazardous Materials, 163,* 650-656.

Hermes, C. J., Melo, C., & Knabben, F. T. (2010). Algebraic solution of capillary tube flows, part I: Adiabatic capillary tube. *Applied Thermal Engineering, 30,* 229-457.

Jabaraj, D. B., Vettri Kathirvel, A., & Mohan Lal, D. (2006). Flow characteristics of HFC407C/HC600a/HC290 refrigerant mixture in adiabatic capillary tubes. *Applied Thermal Engineering, 26,* 1621–1628.

Kim, L., Son, K., Sarker, D., Jeing, J. H., & Lee, S. H. (2011). An assessment of models for predicting refrigerant characteristics in adiabatic and non-adiabatic capillary tube. *Heat Mass Transfer, 47,* 163-180.

Kim, S., Kim, M., & Ro, S. (2002). Experimental investigation of the performance of R22, R407C and R410A in several capillary tubes for air-conditioners. *International Journal of Refrigeration, 25,* 521-531.

Kritsadathikarn, P., Songnetichaovalit, T., Lokathada, N., & Wongwises, S. (2002). Pressure Distribution of Refrigerant Flow in an Adiabatic Capillary Tube. *ScienceAsia, 28,* 71-76.

Kumar, R., Mittal, M. K., & Akhilesh, G. (2009). Numerical analysis of adiabatic flow of refrigerant through a spiral capillary tube. *International Journal of Thermal Sciences, 48,* 1348–1354.

Li, R. Y., Lin, S., Chen, Z. Y., & Chen, Z. H. (1990). Metastable flow of R12 through capillary tubes. *lnternational Journal of Refrigeration, 13,* 181 -186.

Liang, S. M., & Wong, T. N. (2001). Numerical modeling of two-phase refrigerant flow through adiabatic capillary tubes. *Applied Thermal Engineering, 21,* 1035–1048.

Lin, S., Kwok, C. C. K., Li, R. Y., Chen, Z. H., & Chen, Z. Y. (1991). Local frictional pressure drop during vaporisation of R-12 through capillary tubes. *International Journal of Multiphase Flow, 17,* 95-102.

Melo, C., Ferreira, R. T. S., & Pereira, R. H. (1992). Modelling adiabatic capillary tubes: a critical analysis. *Proc lIR-Purdue Refrigeration Conf. West Lafayette, U.S.A.,* 113-123.

Melo, C., Torquato Vieira, L. A., & Pereira, R. H. (2002). Non-adiabatic capillary tube flow with isobutane. *Applied Thermal engineering, 22,* 1661-1672.

Montgomery, D. C. (2005). *Design and analysis of Experiments*. United States of America: John Wiley and Sons Inc.

Park, C., Lee, S., Kang, H., & Kim, Y. (2007). Experimentation and modeling of refrigerant flow through coiled capillary tubes. *International Journal of Refrigeration, 30,* 1168-1175.

Sami, S., & Tribes, C. (1998). Numerical prediction of capillary tube behaviour with pure and binary alternative refrigerants. *Applied Thermal Engineering, 18,* 491-502.

Sarker, D., & Jeong, J. H. (2012). Development of empirical correlations for non-adiabatic capillary tube based on mechanistic model. *International Journal of Refrigeration, 35,* 974 - 983.

Shodiya, S., Azhar, A. A., & Darus, A. N. (2012). Improved Refrigerant Characteristics Flow Predictions in Adiabatic Capillary Tube. *Research Journal of Applied Sciences, Engineering and Technology, 4,* 1922-1927.

Shodiya, S., Azhar, A. A., Henry, N., & Darus, A. N. (2011a). New empirical correlation for sizing adiabatic capillary tubes in refrigeration systems. *Proceedings of the 4th International Meeting of Advances in Thermofluids, Melaka, Malaysia, Oct. 3-4, pp: 787 - 801.*

Shodiya, S., Azhar, A. A., Henry, N., & Darus, A. N. (2011b). Numerical simulation of refrigerant flow in adiabatic capillary tubes including metastability phenomenon. *In Proc. 11th Asian International Conference on Fluid Machinery and 3rd Fluid Power Technology Exhibition Madras Chennai India*, 1-14.

Vins, V., & Vacek, V. (2009). Mass flow rate correlation for two-phase flow of R218 through a capillary tube. *Applied Thermal Engineering, 29,* 2816–2823.

Wongwises, S., Chan, P., Luesuwannatat, N., & Purattanark, T. (2000). Two-Phase separated flow model of refrigerant flowing through capillary tubes. *International Communications in Heat and Mass Transfer, 27,* 343–356.

Wongwises, S., & Chingulpitak, S. (2010). Two-phase flow model of refrigerants flowing through helically coiled capillary tubes. *Applied Thermal Engineering, 30,* 1927-1936.

Wongwises, S., & Pirompak, W. (2001). Flow characteristics of pure refrigerants and refrigerant mixtures in adiabatic capillary tubes. *Applied Thermal Engineering, 21,* 845-861.

Wongwises, S., & Suchatawut, M. (2003). A simulation for predicting the refrigerant flow characteristics including metastable region in adiabatic capillary tubes. *International Journal of Energy Research, 27,* 93-109.

Wongwises, S., & Trisaksri, V. (2003). Correlation for sizing adiabatic capillary tubes. *International Journal of Energy Research, 27,* 1145-1164.

Yilmaz, T., & Unal, S. (1996). General equation for the design of capillary tubes. *ASME Journal of Fluids Engineering, 118,* 150–154.

Zhang, C. L., & Ding, G. L. (2001). Modified general equation for the design of capillary tubes. *ASME Journal of Fluids Engineering, 123,* 914–919.

Zhou, G., & Zhang, Y. (2006). Numerical and experimental investigations on the performance of coiled adiabatic capillary tubes. *Applied Thermal Engineering, 26,* 1106-1114.

Pertanika J. Sci. & Technol. 22 (2): 457 - 470 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Identification of Hot Spots in Proteins Using Modified Gabor Wavelet Transform

D. K. Shakya[1]*, Rajiv Saxena[2] and S.N. Sharma[3]

[1]*Department of Biomedical Engineering, Samrat Ashok Technological Institute, Chandrasekhar Marg, Vidisha-464001 (Madhya Pradesh), India*

[2]*Jaypee University, Anoopshahr, 203390 (Uttar Pradesh), India*

[3]*Department of Electronics and Communication Engineering, Samrat Ashok Technological Institute, Chandrasekhar Marg, Vidisha 464001 (Madhya Pradesh), India*

ABSTRACT

Identification of hot spots is an important issue in proteomics. Identifying hot spots using Digital Signal Processing (DSP) based methods is quite useful in newly discovered proteins as these methods do not require the structural information of proteins. In this paper, Modified Gabor Wavelet Transform (MGWT) was used to predict hot spots from primary amino acid sequence of protein. Incorporation of MGWT into Resonant Recognition Model (RRM) improves the prediction of the hot spots. The proposed method only requires tuning of MGWT to the characteristic frequency of the proteins' functional group, which is determined using RRM. This DSP-based technique is illustrated using several protein examples and the results are compared with the other recently reported digital signal analysis methods, viz. digital filtering and S-Transform based approaches. Relative procedural simplicity of this method over S-transform based approach and better prediction performance than digital filtering method are the novel features.

Keywords: Electron-Ion-Interaction-Potential, Hot spot, protein, resonant recognition model, wavelet transform

Article history:
Received: 2 February 2012
Accepted: 18 March 2013

Email addresses:
D. K. Shakya (devendrashakya@rediffmail.com),
Rajiv Saxena (rsaxena2001@yahoo.com),
S.N. Sharma (sanjeev_n_sharma@rediffmail.com)
*Corresponding Author

INTRODUCTION

Proteins are polymers built up from amino acids (Alberts *et al*., 1998). Although numerous different amino acids are theoretically possible, only 20 of them are commonly found in proteins, and all proteins are made up of combinations of these molecules. The 20 amino acids are represented in a protein sequence as a string

ISSN: 0128-7702

of alphabetical symbols with typical lengths ranging from 100 to 10000 (Anastassiou, 2001). Proteins are the main conductors and workforce in any living process. They play a vital role in body functioning as catalysts accelerating chemical reactions, as carrier and storage molecules in muscle contractions, as antibodies imparting immunity and as receptors in the nervous system generating and transmitting nerve impulses. These cellular processes are largely governed by different types of interaction between proteins, and the function of a protein can be better understood considering its interactions (Uetz *et al.*, 2000). By means of its three dimensional (3-D) structure, protein expresses its biological function. 3-D shape allows the protein to interact with other molecules known as targets. These interactions are very selective in nature. Studies on protein interfaces have revealed that energies are not uniformly distributed. Instead, there are certain critical residues called hot spots comprising only a small fraction of interfaces, yet accounting for the majority of the binding energy. The broad recognition of the importance of characterizing protein interactions in a cell has rendered the development of experimental and computational techniques to detect and predict hot spots with an objective to produce new and more efficient drugs and other biotechnological products.

Experimentally, hot spot residues are identified via Alanaine Scanning Mutagenesis (ASM), as described by Bogan and Thorn (1998). If a residue has a significant drop in binding affinity ($\Delta\Delta G$) when mutated to alanine, it is labelled as a hot residue. Thorn and Bogan (2001) deposited hot spots from the ASM experiments in the ASEdb. ASM is expensive, time consuming and requires a lot of efforts. Hence, simpler and less expensive computational techniques are required by biologists for estimating hot spot locations. Wet lab experiments can then be selectively performed by using the estimates obtained, resulting in a considerable saving of laboratory resources. Hot spot related databases/web servers have been complied by Tuncbag *et al*. (2009). Ofran and Rost (2007a), and Ofran and Rost (2007b) stated that all databases/servers require protein structure for prediction of hot spots except Interaction Sites Identified from Sequence (ISIS). ISIS predicts the hot spots from the primary sequence only and uses the physicochemical features, evolutionary and structural features of the protein through neural network model to predict the hot spots. However, for a newly discovered protein molecule, the only information initially available is its amino-acid sequence. Hence, the Digital Signal Processing (DSP) based methods play important roles in the analysis of these sequences as they do not need any structural information or training for estimating hot spots, apart from the primary amino-acid sequence (Cosic, 1994; Cosic, 2001; Cosic *et al.,* 2002; Vaidyanathan, 2004; Ramchandran & Antoniou, 2008). All the reported DSP-based methods first extract characteristic frequency using Resonant Recognition Model (RRM) (Cosic, 1994) and then apply DSP algorithms. The first DSP-based reported method by Cosic (1994) alters the amplitude at the characteristic frequency and the positions of the amino acids mostly affected by the change of amplitude are defined as hot spots. However, changing a single DFT coefficient affects all the elements of the protein's numerical sequence, making this particular method not reliable. Ramchandran *et al.* (2004) improved the performance of this method using short-time discrete Fourier transform (STDFT) and this improvement was

attained by employing digital filters (Ramchandran & Antoniou, 2008). In a recently reported work, Sahu and Panda (2011) used S-transform to predict hot spots with better accuracy than digital filtering. S-transform approach is relatively complex as it requires multiplication of the S-transform with the consensus spectrum in each time instant, followed by band limited filtering in time-frequency domain to select the characteristic frequency. The band limited filter is to be activated during the specific regions in the time-frequency plane. Application of the different wavelets' functions for their possible uses in the identification of active sites in the proteins has also been reported (see Cosic, 2001; Cosic *et al.,* 2002; Rao & Swamy, 2008). These wavelet based approaches successfully identified the areas of high energy regions that embrace the active sites but the exact identification of hot spot residues is missing in these works.

In this work, Modified Gabor Wavelet Transform (MGWT) reported by Mena-Chalco *et al.* (2008) was used to identify hot spots by tuning it to the characteristic frequency of the proteins' functional group. The prediction accuracy of this method has been compared with the digital filtering method introduced by Ramchandran and Antoniou (2008) and S-transform approach suggested by Sahu and Panda (2011).

The rest of the paper is organized as follows. In Section, II RRM is discussed. Section III describes MGWT and its application to identify hot spots in combination with RRM. The potentiality of the proposed method was assessed using a set of 10 proteins from different functional family selected from the standard databases. The protein sequences and the evaluation criteria used for the experimental study are discussed in Section IV. The experimental results of the proposed method are presented in Section V. A comparative study with S-transform and digital filtering approach is also elaborated in this section. Finally, the research is concluded in Section VI.

RESONANT RECOGNITION MODEL (RRM)

RRM (Cosic, 1994) is a physicomathematical approach to gain insights into selective protein interactions relevant to their biological function. This model explains selectivity of these interactions in terms of the resonant energy transfer between interacting molecules. RRM shows that certain periodicities within the distribution of energies of delocalized electrons along a protein molecule are critical for proteins biological function, i.e., the interaction with its target. RRM interprets information from protein sequences using signal analysis methods. It comprises of two stages: the first step involves the transformation of the amino acid sequence into a numerical sequence by assigning to each amino acid its Electron-Ion Interaction Potential (EIIP) value (Veljkovic *et al.,* 1985). EIIP of an amino acid is a physical property denoting the average energy of the valence electrons in the amino acid, and it is known to correlate well with a protein's biological properties (Lazovic, 1996). Thus, the resulting numerical series represents the distribution of the free electrons' energies along the proteins. The EIIP values for the 20 amino acids are listed in Table 1.

TABLE 1
EIIP Values for 20 Amino-acids

S.No.	Amino-Acid Name	EIIP Values	S.No.	Amino-Acid Name	EIIP Values
1	Alanine	0.0373	11	Methionine	0.0823
2	Cysteine	0.0829	12	Asparagine	0.0036
3	Aspartic acid	0.1263	13	Proline	0.0198
4	Glutamic acid	0.0058	14	Glutamine	0.0761
5	Phenylalanine	0.0946	15	Arginine	0.0959
6	Glycine	0.0050	16	Serine	0.0829
7	Histidine	0.0242	17	Threonine	0.0941
8	Isoleucine	0.0000	18	Valine	0.0057
9	Lysine	0.0371	19	Tryptophan	0.0548
10	Leucine	0.0000	20	Tyrosine	0.0516

In the next step, numerical series are analyzed by transforming them into frequency domain using discrete Fourier transform (DFT). The common frequency components for a group of protein sequences are determined by computing the cross-spectral function $S(e^{j\omega})$.

$$S(e^{j\omega}) = \left| X_1(e^{j\omega}) X_2(e^{j\omega}) \_\_\_\_\_ X_m \right| \quad (1)$$

where $X_1, X_2, \_\_\_\_\_ X_m$ in equation (1) are DFTs corresponding to M proteins. The consensus spectrum obtained from this product has a distinct peak at a certain frequency. This frequency is termed as characteristic frequency. A sufficient number of protein sequences are used to get a distinct peak in the consensus spectrum for a clear identification of the characteristic frequency. For a successful protein target interaction, both the protein and the target signals must share the same characteristic frequency but they must have opposite phase. The matching resembles resonance and so this model of the protein-target recognition has been termed as the resonant recognition model. After determining the characteristic frequency for a particular protein function, the hot-spot locations in a protein or target molecule can be marked by identifying the regions in the numerical sequence where this frequency is dominant. For locating these regions, DSP algorithms can now be employed as hot spot detection is now translated into a time-frequency analysis problem.

PREDICTION OF HOT SPOTS USING MGWT

Gabor wavelet was modified (Mena-Chalco *et al.*, 2008) for analyzing a signal in specific frequency and multiple scales. This has been achieved by varying the Gaussian standard deviation of the analyzing function, while its complex exponential frequency has been kept constant. The following relationship describes the analyzing function corresponding to MGWT:

$$\varphi_{MGWT}(t,n,a) = e^{\frac{-(t-n)^2}{2a^2}} e^{j\omega_0(t-n)} \quad (2)$$

In (2), *n* is the position index of the sequence to be analyzed, *a* is the scaling parameter, and ω_0 is the base frequency to which the MGWT is tuned. MGWT of a signal *u*(*t*) is given as:

$$U(n,a) = \int u(t) e^{\frac{-(t-n)^2}{2a^2}} e^{j\omega_0(t-n)} dt \tag{3}$$

The exponential term in (3) contains only a single frequency ω_0. This makes MGWT capable to capture frequency component ω_0 using different scales that are present at different locations along the position index. The spectrum of the sequence is obtained by computing the squared complex modulus of the MGWT coefficients as:

$$M(n,a) = \left| U(n,a) \right|^2 \tag{4}$$

This spectrum is then projected onto the position axis in order to detect the locations of the presence of a specific frequency w_0 which corresponds to the local maxima regions of the projection. For a sequence of length N, this projection spectra is obtained by summing up the MGWT coefficients for all the scales. Equation (5) describes this computation.

$$MGWT(n) = \sum_a M(n,a), \qquad n = 0, \_\_\_ N-1 \tag{5}$$

In this work, MGWT was used to analyze protein sequences for identifying the hot spot locations. Multi resolution analysis of the protein sequences to capture specific periodicity locations corresponding to the protein's characteristic frequency was carried out. This was achieved by fixing ω_0 at the characteristics frequency of the protein's functional group and then computing MGWT for different scales. Therefore, the MGWT approach combines the features of wavelet (Pirogova *et al.*, 2002) and digital filtering method (Ramchandran & Antoniou, 2008), resulting in an improved performance with less computational complexity. Following are the steps involved in identifying hot spots by using this method:

(a) Convert protein character sequences of the functional group of interest into numerical sequences by using the EIIP values listed in Table 1.

(b) Using (1), plot the consensus spectrum and determine the characteristic frequency.

(c) Tune MGWT by making ω_0 equal to the characteristic frequency.

(d) Using (3), compute the MGWT of the protein sequence of interest at different scales.

(e) Determine the projection spectrum of the sequence using (4) and (5).

(f) Plot the projection spectra and identify the hot spots by locating the energy peaks based on a suitable peak-to-average ratio (Ramchandran & Antoniou, 2008).

PROTEIN SEQUENCES AND PERFORMANCE EVALUATION

In a recent publication, where DSP based identification of hot spots was reported by Sahu and Panda (2011), 10 proteins belonging to different functional families were selected from the standard databases for the experimental study. With an objective to compare the results of the proposed method with S-transform and digital filtering approaches, the same set of protein

sequences (Sahu & Panda, 2011; Ramchandran & Antoniou, 2008) was used in this work. The sequence length, characteristic frequency and PDB ID of these protein sequences are listed in Table 2. The amino acid sequences for these proteins were obtained from the freely available protein data bank (PDB) [http://www.rcsb.org/] and Swiss-Prot [http://us.expasy.org/sprot/]. A benchmark to compare the hot spots identified by the proposed approach had also been generated by combining the results of ASEdb by Thorn and Bogan (2001), and Robetta interface alanine scanning (Robetta-Ala) reported by Kortemme *et al.* (2004) and Kortemme and Baker (2002). In ASEdb, an interface residue is considered as a hot spot if its corresponding $\Delta\Delta G$ is equal to or higher than 2.0 kcal/mol. As for Robetta-Ala, the interface residues with $\Delta\Delta G$ more than 1.0 kcal/mol are taken as hot spots.

TABLE 2
Protein Sequences Investigated and their Characteristics Frequency

S.No.	Organism	Protein Name	PDB ID	Swiss-port ID	Sequence Length	Characteristics frequency
1	Human	basic fibroblast growth factor (bFGF)	4fgf	P09038	146	0.904
2	Human	Growth hormone (hGH)	3hhr	P01241	190	0.270
3	Human	Growth hormone binding Protein (hGHbp)	3hhr	P10912	205	0.270
4	Human	Interleukin (IL4)	1rcb	P05112	129	0.587
5	Human	Human alpha hemoglobin	1vwt	P69905	141	0.023
6	Bacteria	Barnase	1brs	P00648	110	0.321
7	Bacteria	Barstar	1brs	P11540	89	0.321
8	Bacteria	Tryptophin RNA-binding attenuator protein(TRAP)	1wap	P19466	75	0.247
9	*E.Coli*	Colicin-E9 immunity protein (IM9)	1bxi	P13479	86	0.190
10	*C.fumi*	Endoglucanse C	1ulo	P14090	152	0.093

Sahu and Panda (2011) identified the hot spots in the protein sequences by comparing the energy in the regions that contributed to characteristic frequency with a reference energy level. Evaluation criteria of peak to average ratio, proposed by Ramchandran *et al.* (2008), had been followed by Sahu and Panda (2011). In the current work, the author also used the same evaluation criteria with the purpose of establishing a comparative study of the results with the earlier reported DSP approaches. The average value of the projection spectra was computed and used as a reference level for indicating the hot spots in protein sequence. In

order to control the resolution of this method, the ratio of the peaks of the projection spectra to the average value was set as threshold (t_p). The efficiency of the method in identifying the hot spots can be varied by increasing or decreasing the threshold value.

EXPERIMENTAL RESULTS AND COMPARISON STUDY

The MGWT has been applied to protein sequences with transform calculated for different scales *a* and frequency ω_0 (characteristic frequency of the protein). The results in this work were obtained by using 40 analyzing functions corresponding to 40 scale values exponentially separated between 0.2 and 0.7. The lengths of these functions have been truncated to 15 sequence points. As a sample plot, the spectrogram for human basic fibroblast growth factor (ω_0 = 0.904) is shown in Fig.1(a). In Fig.1(b), projection of the spectrum values onto the position axis is plotted. The peaks at certain interface residues shown in Fig.1(b) correspond to the hot spot locations identified by MGWT. These locations were detected by taking 90% of the average energy as threshold to locate the hot spots. Sahu and Panda (2011) also used the same threshold value in their work for the comparative study.

The hot spot prediction performance of MGWT for the proteins listed in Table 2 was compared with the results of the S-transform technique, digital filtering technique and the alanine scan computed from both ASEdb and Robitta-Ala (Sahu & Panda, 2011). The comparative detection results for these methods are given in Table 3. To assess the prediction performance using these results, the following measures (Baldi *et al.*, 2000) were used in this paper:

(i) **Accuracy (A)** – Accuracy is the ratio of the number of correctly predicted residues to the number of all the predicted residues, formulated as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

where TP, FP, TN, and FN stand for the number of true positives (correctly predicted hot spot residues), number of false positives (non-hot spot residues incorrectly predicted as hot spots), number of true negatives (correctly predicted non-hot spot residues) and number of false negatives (hot spot residues incorrectly predicted as non-hot spots), respectively.

(ii) **Recall (R)** – Recall or sensitivity is the proportion of the number of correctly classified hot spot residues to the number of all hot spot residues.

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

(iii) **Specificity (S)** – Specificity is the proportion of the number of correctly predicted non-hot spot residues to the number of all the non-hot spot residues.

$$Specificity = \frac{TN}{TN + FP} \tag{8}$$

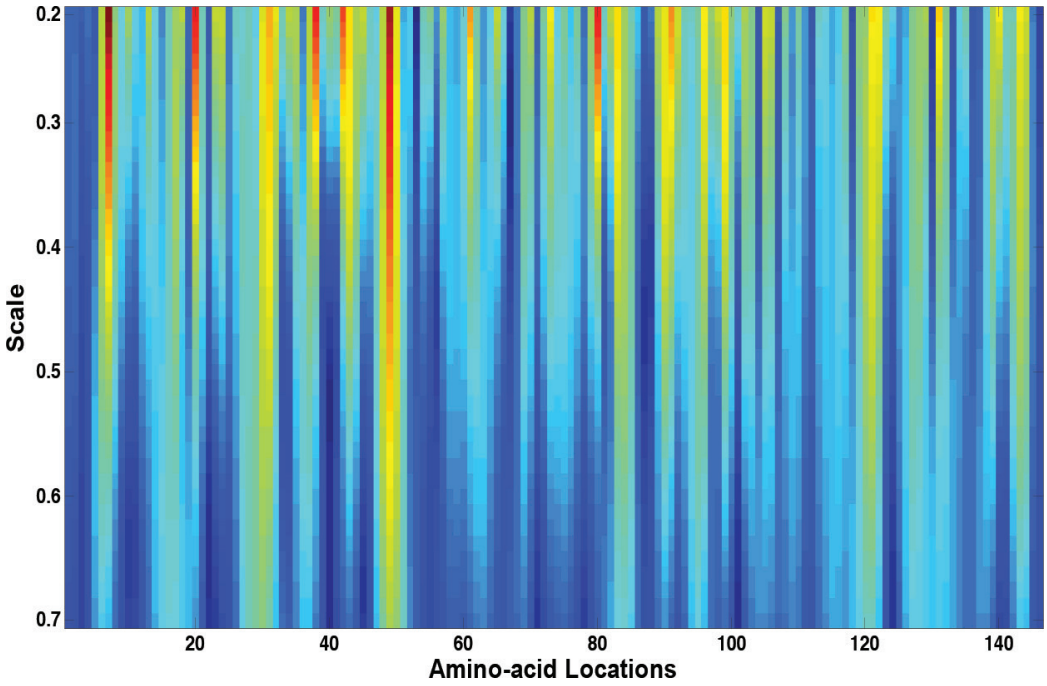


Fig.1(a): Spectrogram for human basic fibroblast growth factor

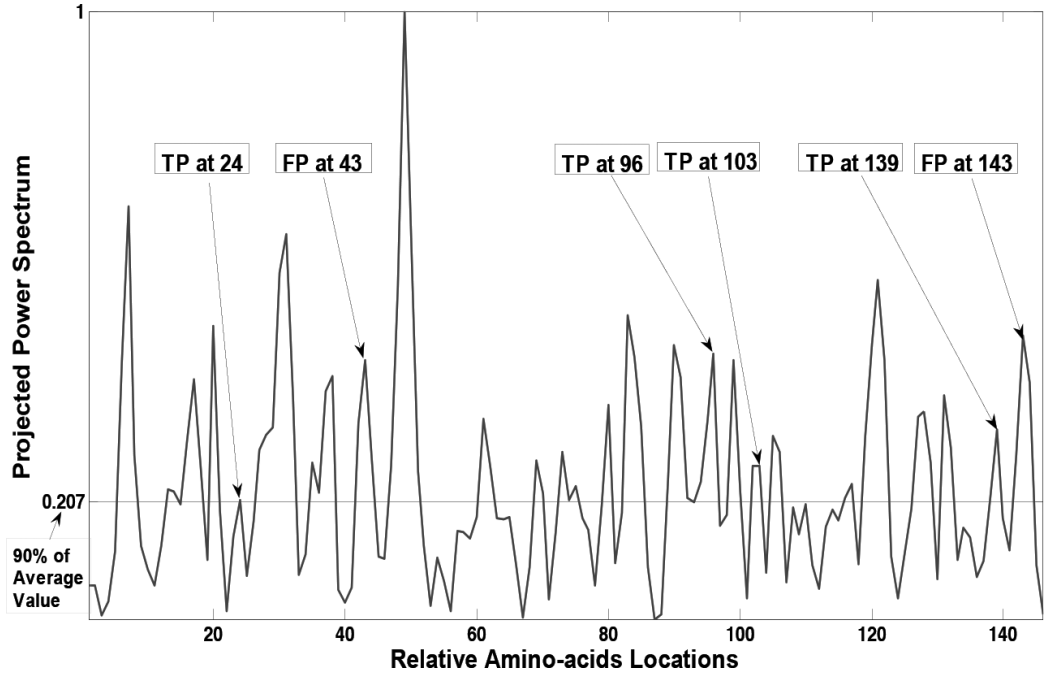


Fig.1(b): Projection spectra for human basic fibroblast growth factor

(iv) **Precision (P)** – Precision is the ratio of number of correctly classified hot spot residues to the number of all residues classified as hot spots.

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

(v) **F-measure (F)** – By using F-measure, we checked the balance between precision and recall, which is formulated as follows:

$$F - Measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{10}$$

(vi) **Matthews Correlation Coefficient (MCC)** – While there is no perfect way of describing true and false positives and negatives by a single number, the Matthews correlation coefficient is generally regarded as being one of the best such measures (Baldi *et al.*, 2000). The MCC can be calculated directly using this formula:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{11}$$

By using the detected hot spots and interface residues listed in Table 3, the average values of the six performance evaluation measures for the ten proteins have been obtained and are shown in Table 4. It is evident from the results in Table 4 that all the performance measures for MGWT are superior to the digital filtering technique. The values of TP and FN for MGWT are less than the corresponding values for S-transform, resulting in comparatively small value of Recall. However, the number of false positives identified by S-transform is greater than those identified by MGWT. Time-frequency filtering operation has been cited (Sahu & Panda, 2011) as the reason behind the large number of false positives. Therefore, the Specificity and Precision values for MGWT are better than the S-transform method. Meanwhile, the performance of S-transform is slightly better than MGWT in terms of F-measure and MCC. However, the Accuracy for S-transform and MGWT methods is the same.

TABLE 3
Comparative Results of the Detected Hot-Spots Using Experimental and Different Computational Methods

Protein Name	Interface Residues	Actual Hot Spots (ASEdb)			Detected Hot-Spots		
		ASEdb (2kcl/mol)	*Robetta Alanine (1kcl/mol)*	*ASEdb+ Robetta Alanine*	Digital Filter	S-Transform	MGWT
basic fibroblast growth factor (bFGF)	22, 24, 26, 44, 46, 96, 97, 101, 103, 107, 109, 110, 111, 113, 114, 140, 142	24, 96, 103, 140	----	24, 96, 103, 140	24, 26	24, 96, 103, 140	24, 96, 103, 139
Growth hormone (hGH)	2, 3, 4, 8, 9, 12, 15, 16, 18, 19, 21, 22, 25, 26, 29, 42, 45, 46, 48, 51, 56, 62, 63, 64, 65, 68, 164, 167, 168, 171, 172, 174, 175, 176, 178, 179, 182, 183, 186	172, 175, 178, 176	18, 25, 42, 45, 46, 64, 168, 171, 175, 179	18, 25, 42, 45, 46, 64, 168, 171, 172, 175, 178, 179	26, 41, 45, 64, 168, 171, 175, 178, 179	18, 25, 42, 47, 65, 168, 172, 175, 178, 180	17, 41, 45, 47, 63, 65, 167, 170, 172, 176, 179
growth hormone binding protein (hGHbp)	43, 44, 72, 76, 77, 80, 98, 102, 103, 104, 105, 108, 120, 121, 122, 124, 126, 127, 164, 165, 166, 167, 169	43, 104, 105, 165, 169	43, 76, 104, 127, 169	43, 76, 104, 105, 127, 165, 169	43, 105, 127, 164, 165, 169	43, 103, 105, 127, 165, 170	43, 126, 164
interleukin (IL-4)	5, 6, 8, 9, 11, 13, 15, 16, 19, 77, 78, 81, 82, 84, 85, 88, 89, 91	9, 88	----	9, 88	9, 88	9, 88	87
human alpha haemoglobin	5, 18, 22, 36, 43, 59, 76, 109, 111, 131	18, 22, 36, 43, 59	----	18, 22, 36, 43, 59	37, 60	22, 36, 60	37, 42
Barstar	29, 35, 39, 42, 76, 80	29, 35, 39	29, 35, 39, 42, 76	29, 35, 39, 42, 76	35, 38, 42	29, 35, 38, 42	36, 40, 43, 76
Barnase	27, 58, 59, 60, 73, 87, 102	27, 58, 59, 73, 87, 102	27, 59, 60, 83, 87, 102	27, 58, 59, 60, 73, 87, 102	27, 59, 73, 87, 102	27, 58, 60, 3, 86,102	27, 57, 60, 73, 87, 102

TABLE 3 (continue)

TRAP	20, 22, 28, 30, 32, 37, 42, 50, 51, 56, 58, 60	37, 40, 56, 58	56, 58	37, 40, 56, 58	37, 40, 56	40, 56, 59	36, 40, 59
colicin-E9 immunity protein (IM9)	23, 24, 27, 28, 29, 30, 33, 34, 37, 38, 41, 48, 49, 50, 51, 55, 56	33, 34, 41, 50, 51, 55	30, 33, 38, 50, 55	30, 33, 34, 38, 41, 50, 51, 55	34, 41, 50, 51, 55	33, 41, 50, 51, 55	30, 39, 41, 49, 52, 55
Endoglucanase C	19, 50, 75, 80, 81, 84, 87, 90, 124, 128	19, 50, 84	----	19, 50, 84	50	50, 84	No Detection

TABLE 4
Comparative Performance of Different Computational Methods

Performance Evaluation Measures in Percentage	Digital Filtering	S-Transform	MGWT
Accuracy	60	67	67
Recall	67	79	70
Specificity	56	60	65
Precision	46	52	53
F-Measure	54.5	62.7	60.3
MCC	21.65	37.29	33.49

Another dimension in which a comparison can be established is the computational load of the three methods. The computational complexity of S-transform is relatively more as it requires the following additional processing:

S-transform spectrum is to be multiplied with the consensus spectrum in each time instant to suppress the noisy frequencies and to boost up energy at the characteristic frequency.

(i) After multiplication by the consensus spectrum distinct energy concentrated areas in the time-frequency plane where the characteristic frequency is dominated are obtained. In order to separate the frequency of interest, a band limited time-frequency filter is to be designed and activated during the specific regions in the time-frequency plane.

In this context, digital filtering (Ramchandran & Antoniou, 2008) and the proposed method are at par, as multiplication by consensus spectrum and time-frequency filtering is not needed in these methods. Tuning of MGWT to the protein characteristic frequency is analogous to anti-notch filtering of the proteomic signal, with anti-notch frequency set to the characteristic frequency. Because of this filtering, multiplication by consensus spectrum in each time instant to boost the characteristic frequency component is not required. Multi-resolution analysis then identifies the hot spot locations and thus eliminates the need of filtering in time-frequency domain to locate the regions where characteristic frequency is dominant.

CONCLUSION

This paper proposes a simple and efficient method for the identification of hot spot residues in proteins using MGWT. The proposed method predicts hot spots from the amino acid sequence only and it does not requires structural information of the protein or any kind of prior training using other protein features. Hence, this method can be quite useful in estimating hot spot residues prior to performing wet lab experiments in a newly identified protein, for which the only information initially available is its amino acid sequence. The performance of this method has been evaluated on protein sequences selected from the standard protein data bases using sufficient number of prediction performance measures. A comparative study with the recently reported signal analysis methods, based on digital filtering and S-transform, was also carried out. Prediction accuracy of all the measures for the proposed method is observed to

be superior to digital filtering method. When compared with S-transform, MGWT gives a better performance in terms of specificity and precision. The accuracy of the two methods is the same. Better Recall, F-Measure, and MCC values have been obtained with S-transform but at the cost of increased computational load. The computational complexity of MGWT based hot spot detection is relatively less than the S-transform method as it does not require multiplication by consensus spectrum in each time instant and time-frequency filtering. Thus in this work, a novel DSP-based method for hot spots identification with satisfactory prediction performance and less computational load has been developed.

REFERENCES

Alberts, B., Bray, D., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (1998). *Essential Cell Biology*. New York: Garland Publishing.

Anastassiou, D. (2001). Genomic signal processing. *IEEE Signal Processing Magazine*, *18*(4), 8-20.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, F. A. C., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, *16*(5), 412–424.

Bogan, A., & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology*, *280*, 1-9.

Cosic, I. (1994). Macromolecular bioactivity: is it resonant interaction between macromolecules? – Theory and applications. *IEEE Transaction on Biomedical Engineering, 41*(12), 1101-1114.

Cosic, I. (2001). Analysis of HIV proteins using DSP techniques. *Proceedings of the 23rd Annual EMBS International Conference,* Istanbul, Turkey*,* 2886-2889.

Kortemme, T., & Baker, D. (2002). A simple physical model for binding energy hot spots in protein-protein complexes. *Proceeding of Natural Academic Science*, *U.S.A., 99*, 14116-14121.

Kortemme, T., Kim, D. E., & Baker, D. (2004). Computational alanine scanning of protein-protein interfaces. *Science STKE*, *219*, 12.

Lazovic, J. (1996). Selection of amino acid parameters for Fourier transform-based analysis of proteins. *Computer Application Bioscience, 12*(6), 553-562.

Mena-Chalco, J. P., Carrer, H., Zana, Y., & Cesar, JR. R. M. (2008). Identification of protein coding regions using the modified Gabor-wavelet transform. *IEEE/ACM Transaction on Computational Biology and Bioinformatics, 5*(2), 198-206.

Ofran, Y., & Rost, B. (2007a). Protein-Protein interaction hotspots carved into sequences. *PLOS Computational Biology, 3,* 1169-1176.

Ofran, Y., & Rost, B. (2007b). ISIS: Interaction sites identified from sequences. *Bioinformatics, 23,* 13-16.

Pirogova, E., Fang, Q., Akay, M., & Cosic, I. (2002). Investigation of the structural and functional relationships of oncogene proteins. *Proceeding of the IEEE, 90*(12), 1859-1867.

Ramchandran, P., Antoniou, A., & Vaidyanathan, P. P. (2004). Identification and location of hot spots in proteins using the short-time discrete Fourier transform. In *Proceeding 38th Asilomer Conference Signals, Systems, Computers*, Pacific Grove, CA*,* 1656-1660.

Ramchandran, P., & Antoniou, A. (2008). Identification of hot-spot locations in proteins using digital filters. *IEEE Journal of Selected Topics in Signal Processing*, *2*(3), 378-207.

Rao, K. D., & Swamy, M. N. S. (2008). Analysis of genomics and proteomics using DSP techniques, *IEEE Transaction on Circuits and System –I, 55*(1), 370-378.

Sahu, S. S., & Panda, G. (2011). Efficient localization of hot spots in proteins using a novel S-transform based filtering approach. *IEEE Transaction on Computational Biology and Bioinformatics, 8*(5), 1235-1246.

Thorn, K. S., & Bogan, A. (2001). ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics, 17*(3), 284-285.

Tuncbag, N., Kar, G., Keskin, O., Gursay, A., & Nussinov, R. (2009). A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. *Briefings in Bioinformatics*, *10*(3), 217-232.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., & Judson, R. S. (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature*, *403*, 623-627.

Vaidyanathan, P. P., & Yoon, B. J. (2004). The role of signal processing concepts in genomics and proteomics. *Journal of the Franklin Institute, 341*, 111-135.

Veljkovic, V., Cosic, I., Dimitrijevic, B., & Lalovic, D. (1985). Is it possible to analyze DNA and protein sequences by the methods of digital signal processing. *IEEE Transaction Biomedical Engineering*, *32*(5), 337-341.

Pertanika J. Sci. & Technol. 22 (2): 471 - 488 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Self-Care Behaviour among Type 2 Diabetes Patients

Siti Khuzaimah, A. S.[1]*, Aini, A.[2], Surindar Kaur, S. S.[2], Hayati Adilin, M. A. M.[3] and Padma, A. R.[4]

[1]Nursing Department, Faculty of Health Sciences, Universiti Teknologi MARA, Puncak Alam Campus, 42300 Puncak Alam, Selangor, Malaysia

[2]Department of Nursing Science, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia

[3]Nutrition and Dietetic Department, Faculty of Health Sciences, Universiti Teknologi MARA, Puncak Alam Campus, 42300 Puncak Alam, Selangor, Malaysia

[4]Occupational Therapy Department, Faculty of Health Sciences, Universiti Teknologi MARA, Puncak Alam Campus, 42300 Puncak Alam, Selangor, Malaysia

ABSTRACT

Self-care behaviour involves all activities type 2 diabetes patients engage in to care for their disease. In our local population, however, most patients do not manage their disease appropriately. This study aimed to determine the level of self-care behaviour and to examine the differences in self-care behaviour according to type 2 diabetes patients' demographic data and health condition at University Malaya Medical Centre. Sample of this study comprised 388 patients (respondents) and data were collected from December 2010 to February 2011 using self-administered questionnaires. Results showed that the level of self-care behaviour was moderately high (mean = 38.94, SD=11.93). There were significant differences between self-care behaviour and ethnicity [Wilk's Lambda = 0.92, $F(12, 1008) = 2.70$, $p < 0.05$], age group [Wilk's Lambda = 0.96, $F(4, 383) = 4.39$, $p < 0.05$], education level [Wilk's Lambda = 0.94, $F(12, 1008) = 1.85$, $p < 0.05$], type of treatment [Wilk's Lambda = 0.92, $F(12, 1008) = 2.84$, $p < 0.05$], health education [Wilk's Lambda = 0.97, $F(4, 383) = 3.33$, $p<0.05$] and smoking status [Wilk's Lambda = 0.96, $F(4, 383) = 4.53$, $p < 0.05$]. Respondents who are Indian, elderly, had lower level of education, on insulin treatment, had received health education on diabetes and not smoking had better self-care behaviour scores. It can be concluded that high risk type 2 diabetes patients should be taught individually so as to help them improve physical and psychological outcome.

Keywords: Self-care behaviour, type 2 diabetes

Article history:
Received: 2 April 2012
Accepted: 26 April 2013

Email addresses:
Siti Khuzaimah, A. S. (sitik123@yahoo.com),
Aini, A. (aini57@um.edu.my),
Surindar Kaur, S. S. (surindar@ummc.edu.my),
Hayati Adilin, M. A. M. (whofrance10@gmail.com),
Padma, A. R. (padma2106@yahoo.com)
*Corresponding Author

ISSN: 0128-7702

INTRODUCTION

Type 2 diabetes is a condition of relative insulin deficiency and the most common chronic condition affecting adults and the elderly (Feinglos & Bethel, 2008; Meiner & Lueckenotte, 2006). The incidences and its prevalence continue to rise due to aging and urbanization globally (Wild *et al*., 2004). In Malaysia, the Third National Health and Morbidity Survey (NHMS III) stated between 1996 and 2006, the prevalence of the disease among adults aged 18 and 30 years old and above rose to 11.6% and 14.9%, respectively (Letchuman *et al*., 2010). It is important to note that type 2 diabetes patients will have considerable functional impairment associated with their reduced health status (Sinclair *et al*., 2008). Patients with complications of type 2 diabetes are hospitalized 1.5 – 3 times more than those without the disease (Leonard *et al*., 2004). The major killer of patients suffering from the illness is macrovascular disease (Rizvi, 2007) such as renal impairment and co-morbidity (Prato *et al*., 2003). The burden of the disease could potentially overwhelm existing health care systems and may cause an escalation in health care cost. As such, one of the prevention strategies is to improve self-care behaviour in order to prevent and minimize serious and long-term complications. Adherence to self-care behaviour that includes healthy-eating, being physically active, monitoring blood glucose and foot care is an important factor in maintaining the disease process (Song *et al*., 2012; Feinglos & Bethel, 2008; Austin, 2005).

The findings of this study give an insight into and awareness among type 2 diabetes patients to understand the significance of self-care behaviour in the management of this disease. As a result, nurses who deal with type 2 diabetes patients will develop new techniques in health education, which can improve patients' compliance and confidence level in controlling the disease and increase their self-care behaviour. This will indirectly promote healthy living by improving glycaemic status and reducing any potential complications. Ultimately, it will improve diabetic patients' quality of life and reduce health care cost or hospital burden.

Self-care Behaviour

Diabetes is a self-managed disease as patients usually provide their own care (Feinglos & Bethel, 2008; Toobert *et al*., 2000). Self-care behaviour refers to active decisions and actions that an individual take to cope with a health problem or to improve his or her health or delay complications (Funnel & Haas, 1995; Weinger *et al*., 2005). It also encompasses an individual's learning from situations and experiences that have worked in the past. Diabetes self-care behaviour includes all the activities which the patients themselves engage in to care for their illness, promote health, improve physical, social, and emotional resources, as well as prevent the long-term and short-term complications from the disease (Bai *et al*., 2009; McCollum *et al*., 2005). It includes the ability, knowledge, skills and confidence to make daily decisions, as well as select and make behavioural changes and the ability to cope with the emotional aspects of their disease within the context of their lives (Barlow *et al*., 2002). There are a lot of advantages of maintaining good self-care behaviour such as lowering the cost required to get health care, increasing effective collaboration between patients and other health care team, and increasing patients' satisfaction and their perceptions of patients' health condition (The fourth Clinical Practice Guideline, CPG, 2009).

Effective management of type 2 diabetes requires a collaborative health care team approach. However, type 2 diabetes patients themselves are the most important individuals in the team because they are the ones who will do most of the disease management. Thus, having a self-care behaviour is essential for type 2 diabetes patients. Hence, the role of nurses is to assist patients to be as independent as possible in managing their health. In addition, self-care behaviour is seen as an empowerment through which gaining of the self-care skills, patients are able to participate more actively in nurturing their own health and in determining the good conditions that will ultimately influence their own health.

There are some factors that may affect the effectiveness of self-care behaviour among type 2 diabetes patients; these include age, gender, health state, developmental age, socio-cultural, health care variables, family system elements, and patterns of living arrangement (Fawcett, 2002; Munshi & Lipsitz, 2007; Johnston-Brooks *et al*., 2002; Wu *et al*., 2007, Wang & Shiu, 2004; Tan & Magarey, 2008; West & Goldberg, 2002; Huang & Hung, 2007; Lee *et al*., 2009; Bai *et al*., 2009; Ciechanowski *et al*., 2004). For example, determinants of non-compliance that may affect glycaemic control in patients includes older age, cost of therapy, complexity of prescribed medical therapies, poor family dynamics (Leichter, 2005), attitudes and health beliefs (Heisler *et al*., 2005). In addition, culture, ethnicity, socioeconomic status and psychosocial also play a big role in explaining type 2 diabetes patients' self-care and health outcomes (Munshi & Lipsitz, 2007; Weinger, 2007).

A few studies have been carried out to examine level of self-care behaviour and factors associated with them in the Malaysian setting. It is important to help diabetic patients to evaluate their daily self-care behaviours, identify possible barriers and understand why patients are unable to perform certain tasks and identify areas of self-care behaviour in which they need assistance. Thus, this study was done to examine the level of self-care behaviour and to identify the factors influencing self-care behaviour so that new nursing strategies can be implemented in health education for diabetic patients. The research questions put forward in this study are: (i) what is the level of self-care behaviour in type 2 diabetes patients?; (ii) Are there any differences in self-care behaviour according to demographic data (such as gender, ethnicity, age, educational level and monthly income) and health condition (such as HbAIc level, type of treatment, length of time of diabetes disease, health education and smoking status) among type 2 diabetes patients?

METHODS

Design, Sample and Setting

This is a cross-sectional study that was conducted from 1st December 2010 to 28th February 2011 at University Malaya Medical Centre (UMMC). The setting was in the diabetic clinic, as well as in the medical, nephrology, orthopedic, and surgery wards.

Non-probability convenience sampling with specific eligibility criteria was employed. This method is deemed feasible, particularly for a researcher with restricted time and resources (Polit & Beck, 2006). The patients who fulfilled the inclusion criteria were asked to participate in this study.

Sample size was calculated using the method by Raosoft (2004); a sample size of 377 respondents was considered as adequate ("rule of thumb"), with 5% margin of error, 95% confidence interval (CI), 50% response distribution, and an estimated population size of 20,000. Patients (the respondents) eligible for this study were those who had been diagnosed with type 2 diabetes with glycosylated hemoglobin (HbA1c) value within the last six months, aged 18 years old and above, speak and understand English and had no major complications (such as being legally blind, suffered severe strokes and were unconscious) that could interfere with self-care.

Research Instrument

The questionnaires comprising of three parts were prepared in the English language. Part A was intended to gather demographic data on the respondents; this part consisted of five items, namely, gender, ethnicity, age, educational level and monthly income. Part B was included to retrieved information on the respondents' health condition, and it also has five items (namely, HbAIc level, type of treatment, length of time of the diabetes disease, health education and smoking status).

Part C, the Summary of Diabetes Self-Care Activities (SDSCA) scale (Toobert *et al*., 2000) was adopted to measure about the frequency of self-care activities reported by patients with diabetes during the past seven days in relation to diet (4 items), exercises (2 items), blood glucose testing (2 items), and foot care (2 items). each respondent was required to circle the answer that best describes his or her self-care behaviour on diabetes management during the past seven days using the Likert scale (0 – 7). The mean number of items is based on the number of days of the week that the behaviour is carried out. The validity reports in this tool were high, with internal consistency of more than 0.50 and test-retest reliability from 0.55 to 0.64 (Glasgow *et al*., 1989; 1998).

Prior to this study, a pilot study was also conducted involving 10% of the target population. The purpose of the pilot study was to identify and investigate the feasibility of the suggested study and to detect any possible error in the data collection instrument such as ambiguous words and instructions, inadequate time and to confirm whether the variables defined by the operational definitions were actually measurable and observable (Brink, 2006). The Cronbach's alpha value for the 10 items in SDSCA scale were found to be moderately acceptable (α = 0.74). The diet subscale consisted of four items ($\alpha = 0.61$), the exercise subscale consisted of two items ($\alpha = 0.71$), the blood-glucose testing subscale consisted of two items ($\alpha = 0.78$) and the foot care subscale consisted of two items ($\alpha = 0.61$).

Ethical Considerations

The ethical approval (reference number 812.30) was granted on 22nd September 2010 by the Medical Ethics Committee, University Malaya Medical Centre.

The respondents' participation in this study was on voluntary basis and they also could opt not to be involved in the study if they eventually chose to. The document reviewed was conducted from patients' records to complete the questionnaire. The variables included most recent HbA1c level (within the last 6 months) and the types of treatment the patients received. It is important to note that full confidentiality and anonymity was maintained.

Data Collection Procedure

Data were collected and this was done in two phases. In Phase 1, the respondents were given as much information regarding the aims of the study and the ethical considerations. All the patients (respondents) taking part in this study were asked to complete the questionnaire.

In Phase 2, data on the latest (within 6 months) HbA1c level of each respondent and the type of treatment received were identified from a review of their medical records.

Data Analysis

The collected data were analyzed using SPSS version 16.0. A descriptive statistics was used to assess the frequency and percentage of the respondents' demographic data (Part A) and health conditions (Part B). Meanwhile, level of self-care behaviour was presented in terms of mean and standard deviation (SD). Multivariate analysis of variance (MANOVA) was used to identify the differences in the respondents' self-care behaviour based on their demographic information and health condition. MANOVA is the extension of ANOVA to more than one dependent variable and this procedure is used to test the significance of differences between the means of two or more groups on two or more dependent variables considered simultaneously (Polit & Beck, 2006). If the results were significant, LSD post-hoc test was performed to determine which group means differed significantly from the others. This helped specify the exact nature of the overall effect determined by the F test.

RESULTS

Respondents' Demographic Data

Data gathered from three hundred and eighty eight respondents data were analyzed (Table 1). Out of the total respondents, 57.5% were men, nearly half of them were Malays (46.9%), and 76% aged 64 years and below. Over half of the respondents (55.9%) had secondary level education. As for monthly income, most of the respondents were having an income below RM1000.

Health Condition

According to the Gribbles Pathology (2011) criteria, majority of the respondents in this study had poor glycaemic control (57.7%). About 43% indicated they are taking oral medication, 53.1% of them have been diagnosed with diabetes for more than ten years, and 65.2% have not received any health education about diabetes and self-management. Majority of the respondents do not smoke (89.4%). An overview of the respondents' health conditions is presented in Table 2.

TABLE 1
Description for demographic data of type 2 diabetes patients

Variables		Frequency	Percentage
Gender:	Male	223	57.5
	Female	165	42.5
Ethnicity:	Malay	182	46.9
	Chinese	70	18.0
	Indian	122	31.4
	Others	14	3.6
Age:	≤ 64 years (adult)	295	76.0
	≥ 65 years (elderly)	93	24.0
Educational level:	Never	8	2.1
	Primary	49	12.6
	Secondary	217	55.9
	Tertiary	114	29.4
Monthly income:	< RM1000	139	35.8
	RM1001-RM2000	100	25.8
	RM2001-RM3000	54	13.9
	RM3001-RM4000	28	7.2
	>RM4000	67	17.3

TABLE 2
Description for health condition of type 2 diabetes patients

Variables		Frequency	Percentage
HbA1c level:	Good control	60	15.5
	Satisfactory control	104	26.8
	Poor control	224	57.7
Type of treatment:	Diet control	19	4.9
	Oral medication	167	43.0
	Insulin	81	20.9
	Oral medication + insulin	121	31.2
Length of diabetes:	≤ 5 years	112	28.9
	6-10 years	70	18.0
	>10 years	206	53.1
Health education:	No	253	65.2
	Yes	135	34.8
Smoking status:	No	347	89.4
	Yes	41	10.6

Respondents' Level of Self-care Behaviour

Overall, the respondents' level of self-care behaviour was found to be moderately high (38.94; SD=11.93). For each scale, the highest score was for diet (mean = 4.85; SD=1.25), indicating that the respondents were certainly able to perform this task. It was followed by foot care (mean = 3.57; SD=2.54), exercises (mean = 3.20; SD=2.17), and blood-glucose testing (mean = 3.02; SD=2.48).

TABLE 3
Type 2 diabetes patients' level of self-care behaviour

Variable	Frequency	Minimum	Maximum	Mean	(SD)
Self-care behaviour	388	5.00	70.0	38.94	11.93
Subscale:					
Diet	388	0.00	7.00	4.85	1.25
Foot care	388	0.00	7.00	3.57	2.54
Exercise	388	0.00	7.00	3.19	2.17
Blood-glucose testing	388	0.00	7.00	3.02	2.48

Differences in Self-care Behaviour Based on Demographic Data and Health Condition

This study aimed to examine the differences in the respondents' self-care behaviour according to their demographic data and health condition. Self-care behaviour is the dependent variable which includes of diet, exercise, blood-glucose testing and foot care, whereas demographic data and health condition are the independent variables which consist of gender, ethnicity, age, educational level, monthly income, HbAIc level, type of treatment, length of time of the diabetes disease, health education and smoking status.

Normality of the dependent variable was assessed for each dependent variable and they were found to be normally distributed (Kolmogorov-Smirnov = 0.04; $p > 0.05$). Therefore, the mean score for each dependent variable (self-care behaviour) and for each independent variable (demographic data and health condition) sub-group, subjected to MANOVA, was considered as appropriate.

Ethnicity

The results from MANOVA in Table 4 indicated that there was a significant difference in the self-care behaviour in collective between ethnicity [Wilk's Lambda = 0.92, $F(12, 1008) = 2.70$, $p<0.05$]. Meanwhile, a follow-up univariate ANOVA revealed significant differences in terms of diet [$F(3, 384) = 6.24$, $p<0.01$)]. The difference was also found in the LSD *post hoc* test, whereby Indians had better self-care behaviour in relation to their diet as compared to other ethnic groups (see Table 4).

TABLE 4
Differences in the self-care behaviour according to type 2 diabetes patients' demographic data and health condition (n = 388)

Variable	*N* (%)	MANOVA Wilk's Lambda	ANOVA								LSD post hoc
			Diet		Exercise		Blood glucose testing		Foot care		
			M ± *SD*	*F*	*M* ± *SD*	*F*	*M* ± *SD*	F	M ± SD	F	
Gender											
Male	223 (57.5)	0.99,	4.81±1.28	0.38	3.29±2.14	1.11	2.98±2.44	0.16	3.56±2.58	0.01	
Female	165 (42.5)	F(4,383)	4.89±1.22		3.06±2.21		3.08±2.54		3.584±2.48		
		= 0.56									
Ethnicity		**		***							
Malay	182 (46.9)	0.92,	4.68±1.34	6.24	3.15±2.09	1.26	3.06±2.41	0.53	3.80±2.52	2.40	Indian >
Chinese	70 (18.0)	F(12,1008)	4.74±1.33		3.39±2.38		2.76±2.56		2.86±2.48		Chinese >
Indian	122 (31.4)	= 2.70	5.22±1.00		3.25±2.19		3.03±2.53		3.63±2.59		Malay > Others
Others	14 (3.6)		4.23±0.93		2.18±1.80		3.61±2.74		3.54±2.21		
Age (years)		**		**							
≤ 64	295 (76.0)	0.96,	4.74±1.28	8.92	3.28±2.15	2.05	3.12±2.47	2.18	3.66±2.53	1.59	
≥ 65	93 (24.0)	F(4,383)	5.18±1.09		2.91±2.24		2.69±2.51		3.28±2.55		
		= 4.39									
Education level		*		*							
Never	8 (2.1)	0.94,	5.16±1.49	3.13	3.44±2.60	0.45	2.88±2.46	0.54	3.19±2.87	2.51	Primary >
Primary	49 (12.6)	F(12,1008)	5.30±1.07		2.88±2.65		2.79±2.91		2.68±2.73		Secondary >
Secondary	217 (55.9)	= 1.85	4.72±1.34		3.26±2.22		2.95±2.47		3.76±2.58		Tertiary
Tertiary	114 (29.4)		4.88±1.09		3.17±1.81		3.25±2.32		3.61±2.28		

TABLE 4 (continue)

Monthly income:											
< RM1000	139 (35.8)	0.94,	4.83±1.28	1.67	3.39±2.36	1.31	3.04±2.54	1.34	3.23±2.63	1.56	
RM1001-RM2000	100(25.8)	F(16,1162)	5.03±1.21		3.35±2.14		2.73±2.52		3.47±2.78		
RM2001-RM3000	54 (13.9)	= 1.57	4.66±1.15		2.81±2.08		2.78±2.55		4.02±2.12		
RM3001-RM4000	28 (7.2)		4.44±1.74		2.66±1.68		3.05±2.46		3.73±2.39		
>RM4000	67 (17.3)		4.92±1.07		3.08±2.04		3.58±2.23		3.99±2.26		
HbA1c Level											
Good	60 (15.5)	0.96,	4.90±1.16	2.57	3.19±2.31	0.42	2.70±2.51	1.63	2.83±2.59	3.01	
Satisfactory	104 (26.8)	F(8,764)	5.06±1.34		3.35±2.13		2.79±2.60		3.70±2.61		
Poor	224 (57.7)	= 1.94	4.73±1.23		3.11±2.16		3.21±2.42		3.70±2.46		
Type of treatment		**						***			
Diet control	19(4.9)	0.92,	4.71±1.11	1.35	2.92±1.95	1.08	2.29±2.55	8.27	2.93±2.68	0.63	Insulin > oral
Oral	167(43.0)	F(12,1008)	4.91±1.32		3.25±2.11		2.44±2.37		3.73±2.50		medication >
Insulin	81(20.9)	= 2.84	4.99±1.29		3.48±2.44		3.93±2.54		3.48±2.60		Diet control
Oral + insulin	121(31.2)		5.68±1.15		2.96±2.10		3.32±2.38		3.50±2.53		
Length of time of diabetes disease											
≤ 5 years	112(28.9)	0.96,	4.80±1.45	0.18	3.30±2.23	1.71	2.65±2.50	2.52	3.21±2.58	4.46	
6-10 years	70(18.0)	F(8,764)	4.81±1.04		2.76±1.88		2.85±2.34		3.10±2.38		
>10 years	206(53.1)	= 1.89	4.88±1.21		3.27±2.22		3.28±2.50		3.93±2.52		
Health education		*								***	
No	253(65.2)	0.10,	4.82±1.27	0.27	3.13±2.20	0.47	2.93±2.54	0.92	3.23±2.56	13.26	
Yes	135(34.8)	F(4,383)	4.89±1.22		3.29±2.12		3.19±2.37		4.20±2.37		
		= 3.33									

TABLE 4 (continue)

Smoking status		**		**						
No	347(89.4)	0.96,	4.91±1.20	9.96	3.20±2.16	0.08	2.95±2.46	2.94	3.49±2.50	3.39
Yes	41(10.6)	F(4,383) = 4.53	4.27±1.56		3.10±2.30		3.65±2.64		4.26±2.78	

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Age Groups

As shown in Table 4, there was a significant difference of the self-care behaviour of the respondents aged 65 years and above (elderly) with those aged 64 and below (adults), with Wilk's Lambda = 0.96, $F(4, 383) = 4.39$, $p<0.05$). It was also revealed that the elderly had better self-care behaviour in relation to their diet as compared to adults or those in the other age category.

Level of Education

The results of MANOVA for the respondents' level of education and self-care behaviour showed significant difference (Wilk's Lambda = 0.94, $F(12, 1008) = 1.85$, $p<0.05$). In more specific, self-care behaviour was found to be significantly and collectively differed among the respondents of different levels of education. The follow-up univariate ANOVA revealed significant differences in terms of diet $F(3, 384) = 3.13$, $p<0.05$). The LSD *post hoc* test also indicated that the respondents with primary education had better self-care behaviour on diet as compared to those with secondary and tertiary education (see Table 4).

Type of Treatment

Referring to Table 4, the result of MANOVA on type of treatment and self-care behaviour was significant (Wilk's Lambda = 0.92, $F(12, 1008) = 2.84$, $p<0.05$). This showed that self-care behaviour differed significantly between types of treatment. Follow-up univariate ANOVA revealed significant differences on blood glucose testing $F(3, 384) = 8.27$, $p<0.01$).The LSD post hoc test as shown in Table 4 revealed that respondents on treatment with insulin had better self-care behavior which practiced more on blood glucose testing than respondents on diet control and on oral medication.

Health Education

The results in Table 4 revealed that collectively, there was a significant difference of self-care behaviour between those who received or did not receive health education; Wilk's Lambda = 0.97, $F(4, 383) = 3.33$, $p<0.05$. Univariate ANOVA test found that respondents who have received health education on diabetes had better self-care behaviour on foot care as compared to those who have not received health education on diabetes.

Smoking Status

It was found in Table 4 that there was a significant difference of self-care behaviour between smoking status (Wilk's Lambda = 0.96, $F(4, 383) = 4.53$, $p<0.05$). Thus, smoking status has influence on self-care behaviour. Respondents who were not smoking had better self-care behaviour on diet as compared to those who were smoking.

DISCUSSION

Level of Self-care Behaviour

Answers to the current research's objectives were obtained using the scale developed by Toobert *et al*. (2000) and the questions in the SDSCA scale measured the respondents' activities for the past seven days. The results showed that patients with type 2 diabetes have a moderately high level (based on the mean scores) of self-care behaviour. The findings in this study are quite similar to those by Ploypathrpinyo (2008) who also reported a moderate level of self-care behaviour in a study carried out on knowledge and self-care behaviour of type 2 diabetes patients among the population in Thailand. However, there is limited research measuring level of self-care behaviour as there is no cut-off point that has been determined to identify adherence versus non-adherence in the SDSCA scale (Eigennmann *et al*., 2009; Toobert, personal communication, October 6, 2010).

For each scale, the majority of patients were found having good level of self-care behaviour in terms of diet, which meant that they were following healthy eating plan and taking nutritious food (with less calories and from healthier sources of carbohydrates such as *cappati or tosei* with *Tandoori chicken*) almost every day. Similarly, Toobert *et al*. (2000) also found that the patients in their study typically reporting higher levels in terms of diet compared to other self-care tasks.

Foot care, which requires inspecting the feet thoroughly to check for abrasions, lesions, and early infections, may be thought of as a relatively solitary activity (Ciechanowski *et al*., 2004). On the other hand, the patients in this study were more likely to check and care further on their feet. Similarly, Ploypathrpinyo (2008) also found foot care was at a good level among all the aspects of patients' self-care behaviour. It might also due to the fact that some patients had received health education on diabetes and thus, they became more aware of its long-term effects.

In terms of exercise, the patients were found to be less likely to perform this activity. In fact, physical exercise was insufficient in the aspects of self-care behaviour (Ploypathrpinyo, 2008) because it often involved and depended on other important factors such as guidance from health care provider and exercising with a partner or in a group (Ciechanowski *et al*., 2004). Besides, barriers to doing exercise among patients with type 2 diabetes may due to working commitment (do not have enough time to do exercises) or due to other ailments such as obesity and arthritis (Austin, 2005).

Glucose monitoring is a relatively quick and straightforward procedure (Ciechanowski *et al*., 2004) but most of the patients were least concern with monitoring their own blood glucose. Most patients in this study were on oral medication and this was similar to NHMS III (2006) which reported the majority of the patients were on oral medication (Letchuman *et al*., 2010). Type 2 diabetes patients who are on oral medication typically found it unnecessary to check their blood-glucose level everyday (Edelman & Chae, 2009).

Differences in Self-care Behaviour Based on Respondents' Demographic Data and Health Condition

The results revealed the respondents of the Indian ethnic group aged 65 years and above (elderly), with low education level, on insulin treatment, and had received health education on diabetes and non–smoker, had better self-care behaviour scores as compared to adults of other ethnic groups with higher education level, not on insulin treatment, and had not received health education on diabetes and patients who were smoking.

Culture and ethnicity play a large role in explaining diabetic patients' self-care (Munshi & Lipsitz, 2007). This study found that Indian patients are on a better diet plan when compared to the respondents of other ethnicity. This might be due to the fact that the respondents of Indian ethnicity usually consume diets that are high in grains, vegetables and fruits. The dishes are typically served with basmati rice (whole-grain rice) and Indian breads made from wholemeal flour; all these are helpful in training them to plan for a good diet. Moreover, some Indian respondents are on vegetarian diet, while food prepared by those of other ethnic group is usually high in carbohydrates and calories which are based on rice and coconut milk such as *nasi lemak* and *nasi briyani*. Food that contains high fibre is good for health and diabetic patients should increase their intake of whole grain products and maintain a diet low in glycemic index (Qi *et al*., 2006).

As for age group, the elderly had better self-care behaviour and had better diet compared to younger adults. Similarly, it was found older patients have better self-care behaviour which includes having good diet (Wang & Shiu, 2004). In general, the elderly eat lesser amounts of calorie burning food and eat out less frequently compared to other adults (Harris & Blisard, 2002).

This study also found that the respondents with primary education had better self-care behaviour and scored better on diet than those with secondary and tertiary education. Likewise, Kim *et al*. (2004) stated that in developing countries, the respondents of this group seemed to maintain a healthy lifestyle. Individuals with higher level of education are more likely to socialize and gather with friends, become exposed to unhealthy lifestyle including taking diet that is high in fat and sugar (Kim *et al*., 2004).

Daily blood-glucose monitoring is important for all patients with diabetes (Austin, 2005). The study found that patients who are on insulin treatment have higher self-care behaviour in term of their blood glucose testing than those on a diet control and on oral medication. Self-monitoring of blood-glucose levels among diabetic patients with pharmacologically treatment was associated with better glycaemic control (Karter *et al*., 2001). However, self-monitoring of blood glucose in self-care was associated with high costs for diabetes non-insulin treatment (Simon *et al*., 2008). Patients who are not on insulin treatment need to consider the price of the equipment used such as syringes, needles, blood glucose testing machine, and test strips which are more expensive and may put a burden on them to purchase the equipment (Volman *et al*., 2008). However, not all type 2 diabetes patients need to check their blood sugar level every day, especially those who are on oral medication and diet control only.

The results revealed that collectively, patients who had received health education on diabetes were better in their self-care behaviour and more likely to perform foot care as compared to those who did not receive health education on diabetes at all. Other studies have also found that patients who had received health education had better self-care behaviour (Wu *et al*., 2007) and patients who had received education about diabetic foot care were more likely to examine their feet regularly (Lee *et al*., 2009).

Diabetic patients with poor self-care behaviour were more likely to smoke (Ciechanowski *et al*., 2004). There was also an association between smoking and poor diet, as previously reported by Maynard *et al*. (2005). Similarly, this study also found that the patients who are non-smokers have better self-care behaviour and are more likely to consume healthy diet as compared to those who are smokers.

CONCLUSION

In conclusion, special attention needs to be given to type 2 diabetes patients who are least likely to exercise regularly and monitor their blood glucose level. Other influential factors such as ethnicity, age group, level of education, type of treatment, health education and smoking status should also be highlighted. These type 2 diabetes patients are at high-risk and thus, special individualized nursing education initiative on self-care behaviour should be carried out regularly. Nurses should work collaboratively with other health care providers to encourage diabetic patients to be more aware of the risks they face and better manage the disease.

In this study, patients who are of other ethnic groups (except for Indians), adults with high level of education and smokers were less likely to consume healthy diet. Firstly, type 2 diabetes patients who smoke should be encouraged to attend smoking cessation programme or advised to stop smoking altogether. Besides, type 2 diabetes patients need to learn how food affects disease control and their overall health. Nurses as a diabetes educator should collaboratively work with dieticians during health education session on diabetes (e.g., setting the goals of health education strategy programme, preparing teaching materials such as pamphlets and booklet, etc.) to assist the patients and their family members to gain knowledge and be more aware of the effects of food on blood glucose, the sources of carbohydrates and fat, appropriate meal preparation, and resources to assist in making food choices. Type 2 diabetes patients should be taught to read labels, as well as to plan, prepare meals and measure food for portion and fat control, and carbohydrate counting (Austin, 2005). In addition, family members' involvement in this session should be encouraged as it would be more beneficial to the patients. Family members should know suitable food choices and the kind of diet to be taken by diabetic patients because this can help increase patients' self-esteem and awareness of their health condition.

This study found that patients who did not receive health education on type 2 diabetes were less likely to perform foot care. Hence, nurses need to help them, especially the newly diagnosed patients by educating them on the preventive strategies to reduce diabetic complications which include regular foot care hygiene and regular foot inspections. Patients also should be advised to visit their doctors according to the appointment given to them to closely monitor the progress of the disease.

This study found that type 2 diabetes patients were less likely to perform blood glucose testing especially those who are not on any insulin treatment. Patients often address barriers in blood glucose monitoring such as physical, financial, emotional, and cognitive issues (Austin, 2005). As health care professionals, nurses and medical doctors should explain to patients about the importance of blood glucose testing and maintaining it at a healthy/normal level. In addition, Type 2 diabetes patients need to be taught of necessary or relevant skills that include self-monitoring sign and symptoms of hypoglycemia or hyperglycemia situation and other things that they should do if they encounter those manifestations. Meanwhile, these professionals must know the financial burdens incurred by the treatment as these will ultimately have an influence on the patients' self-care behaviour, and for this reasons, they should be referred to psychologists for counselling sessions. For instance, relevant government and non-government agencies such as Social Security Organization (SOCSO), "Pusat Zakat" and Malaysian Diabetes Association may be useful for type 2 diabetes patients to get a personal assistance.

LIMITATIONS AND RECOMMENDATIONS

This study employed a cross-sectional survey and the use of this research design means that causality can not be established. Besides, non-probability convenience sampling approach may cause selection bias and the use of specific entry criteria of respondents into the study sample may not represent the whole population. In addition, the setting of this study was solely at UMMC and hence, the results could not be generalized to all type 2 diabetes patients in Malaysia. It is worth mentioning that some patients in this study felt uneasy to disclose their personal circumstances and health conditions.

A study employing qualitative and mix-method approaches needs to be carried out in the future to explore patients' natural feelings and behaviours. This would also include an intervention programme study design that needs to be implemented so as to achieve better effectiveness of the results. A longitudinal study may also yield better results as researchers can observe the trends and track factors associated that may have effects on the changes in diabetic patients' self-care behaviour.

ACKNOWLEDGEMENTS

Special thanks to the Department of Medicine, Department of Orthopaedic Surgery, Department of Surgery and Medical Ethics Committee, University Malaya Medical Centre, Department of Nursing Science and University Malaya Postgraduate Research Fund, University of Malaya, and Universiti Teknologi MARA.

REFERENCES

Austin, M. (2005). Importance of self-care behaviours in diabetes management. *Business briefing: US Endocrine Review*. Retrieved April 20, 2011, from http://www.touchbriefings.com/pdf/1479/austin_bookforweb.pdf.

Bai, Y. L., Chiou, C. P., & Chang, Y. Y. (2009). Self-care behaviour and related factors in older people with type 2 diabetes. *Journal of Clinical Nursing, 18*, 3308-3315.

Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: a review. *Patient Education and Counseling, 48*, 177-187.

Brink, P. (2006). *Basic steps in planning nursing research from question to proposal* (4th Edition). Boston: Jones and Bartlett Publishers.

Clinical Practice Guideline. (2009). *Management of type 2 diabetes mellitus* (4th Edition). Retrieved July 16, 2010, from http://www.moh.gov.my/cpgs/67.

Ciechanowski, P., Russo, J., Katon, W., Korff, M. V, Ludman, E., & Lin, E. (2004). Influence of patient attachment style on self-care and outcomes in diabetes. *Psychosomatic Medicine*, *66*(5), 720-728.

Edelman, S. V., & Chae, J. Y. (2009). Self-monitoring of blood glucose in non-insulin-using type 2 diabetes. *MedscapeCME Diabetes & Endocrinology*. Retrieved March 29, 2011, from http://www.medscape.org/viewarticle/710117.

Fawcett, J. (2002). *Overview of Nurse Theorist: Dorothea Orem's Self-Care Framework*. Retrieved March 29, 2011, from http://www4.desales.edu/~sey0/orem.html

Eigennmann, C. A., Colagiuri, R., Skinnert, T. C., & Trevena, L. (2009). Education and psychological aspects are current psychometric tools suitable for measuring outcomes of diabetes education? *Diabetic Medicine, 26*, 425-438.

Feinglos, M. N. & Bethel, M. A. (2008). *Type 2 Diabetes Mellitus; An Evidence-Based Approach to Practical Management.* Durham, NC: Humana Press.

Funnel, M. M. & Haas, L. B. (1995). National standards for diabetes self-management education program (technical review). *Diabetes Care*, *18*, 100-116.

Glasgow, R. E., Toobert, D. J, Riddle, M., Donnelly, J., Michell, D., & Calder, D. (1989). Diabetes-specific social learning variables and self-care behaviours among persons with type 2 diabetes. *Health psychology*, *8*(3), 285-303.

Glasgow, R. E., La Chance, P., Toobert, D. J., Brown, J., Hampson, S. E., & Riddle, M. C. (1998). Long-term effects and costs of brief behavioural dietary intervention delivered in the medical office. *Patient Education and Counseling*, *32*, 175-184.

Gribbles Pathology. (2011). *Clinical information*. Retrieved October 29, 2011, from http://www.gribbles.com.my/commontest.html.

Harris, J. M. & Blisard, N. (2002). Food Spending and the Elderly. *FoodReview*, *25*(2), 14-18.

Heisler, M., Piette, J. D., Spencer, M., Kieffer, E., & Vijan, S. (2005). The relationship between knowledge of recent HbA1c values and diabetes care understanding and self-management. *Diabetes Care*, *28*, 816-822.

Huang, M. C., & Hung, C. H. (2007). Quality of life and its predictors for middle-aged and elderly patients with type 2 diabetes mellitus. *Journal of Nursing Research*, *15*(3), 193-200.

Johnston-Brooks, C. H., Lewis, M. A., & Garg, S. (2002). Self-efficacy impacts self-care and HbA1c in young adults with type I diabetes. *Psychosomatic Medicine, 64*(1), 43–51.

Karter, A. J., Ackerson, L. M., Darbinian, J. A., D'Agostino, R. B., Ferrara, A., Liu, J., & Selby, J. V. (2001). Self-monitoring of blood glucose levels and glycemic control: The Northern California Kaiser Permanente Diabetes Registry. *The American Journal of Medicine, 111*(1), 1-9.

Kim, S., Symons, M., & Popkin, B. M. (2004). Contrasting socioeconomic profiles related to healthier lifestyle in China and the United States. *American Journal of Epidemiology*, *159*(2), 184-191.

Lee, H., Ahn, S., & Kim, Y. (2009). Self-care, self-efficacy and glycaemic control of Koreans with diabetes mellitus. *Asian Nursing Research*, *3*(3), 139-146.

Leichter, S. B. (2005). Making outpatient care of diabetes more efficient: analyzing noncompliance. *Clinical Diabetes, 23*(4), 187-190.

Leonard, J., Boseman, L., & Vinicor, F. (2004). Aging Americans and diabetes: A public health and clinical response. *Geriatrics*, *59*(4), 14-16.

Letchuman, G. R., Wan Nazaimoon, W. M., Wan Mohamad, W. B., Chandran, L. R., Tee, G. H., & Jamaiyah, H. (2010). Prevalence of diabetes in the Malaysian National Health Morbidity Survey III 2006. *Medical Journal Malaysia*, *65*(3), 173-179.

Maynard, M., Gunnell, D., Ness, A. R., Abraham, L., Bates, C. J., & Blane, D. (2005). What influences diet in early old age? Prospective and cross-sectional analyses of the Boyd Orr cohort. *European Journal of Public Health*, *16*(3), 315-323.

McCollum, M., Hansen, L. S., Lu, L., & Sullivan, P. W. (2005). Gender Differences in diabetes mellitus and effects on self-care activity. *Gender Medicine, 2*, 246-254.

Meiner, S. E., & Lueckenotte, A. G. (2006). *Gerontological Nursing (3rd Ed.).* St. Louis: Mosby.

Munshi, M. N., & Lipsitz, L. A. (2007). *Geriatric Diabetes.* New York: Informa Health Care.

Parchman, M. L., Pugh, J. A., Noel, P. H., & Larme, A. C. (2002). Continuity of care, self-management behaviours and glucose control in patients with type 2 diabetes. *Medical Care*, *40*(2), 137-144.

Prato, S. D., Heine, R. J., Keilson, L., Guitard, C., Shen, S. G., & Emmons, R. P. (2003). Treatment of patients over 64 years of age with type 2 diabetes. *Diabetes Care, 26*(7), 2075-2080.

Ploypathrpinyo, W. (2008). Knowledge and self-care behaviors of diabetes patients at NamNao Hospital, Petchabun Province. *Khon Kaen Hospital Medical Journal*, *32*(7), 87-92.

Polit, D. F., & Beck, C. T. (2006). *Essentials of nursing research: methods, appraisal and utilization (6th ed.).* Philadelphia: Lippincott, Williams, & Wilkins.

Raosoft. (2004). *Sample size calculator*. Retrieved July 15, 2010, from http://www.raosoft.com/samplesize.html

Rizvi, A. A. (2007). Management of diabetes in older adults. *The American Journal of the Medical Sciences, 333*(1), 35-47.

Qi, L., Van Dam, R. N., Liu, S., Franz, M., Mantzoros, C., & Hu, F. B. (2006). Whole-grain, bran, and cereal fiber intakes and markers of systemic inflammation in diabetic women. *Diabetes Care*, *29*(2), 207- 211.

Simon, J., Gray, A., Clarke, P., Wade, A., Neil, A., & Farmer, A. (2008). Cost effectiveness of self-monitoring of blood glucose in patients with non-insulin treated type 2 diabetes: economic evaluation of data from the DiGEM trial. *British Medical Journal, 336*, 1177-1180.

Sinclair, A. J. (2006). Special considerations in older adults with diabetes: meeting the challenge. *Diabetes Spectrum, 19*(4), 229-233.

Song, Y., Song H. J., Han, H. R., Park, S. Y., Nam, S., & Kim, M. T. (2012). Unmet Needs for Social Support and Effects on Diabetes Self-care Activities in Korean Americans with Type 2 diabetes. *The Diabetes Educator*.

Stys, A. N., & Kulkarni, K. (2007). Identification of self-care behaviors and adoption of lifestyle changes result in sustained glucose control and reduction of comorbidities in Type 2 diabetes. *Diabetes Spectrum, 20*(1), 55-58.

Tan, M. Y., & Magarey, J. (2008). Self-care practices of Malaysian adults with diabetes and sub-optimal glycaemic control. *Patient Education and Counselling*, *72*, 252-267.

Tonstad, S., Butler, T., Yan, R., & Fraser, G. E. (2009). Type of vegetarian diet, body weight and prevalence of type 2 diabetes. *Diabetes Care*, *32*(5), 791-796.

Toobert, D. J., Hampson, S. E., & Glasgow, R. E. (2000). The summary of diabetes self-care activities measure. *Diabetes Care*, *23*(7), 943-950.

Volman, B. Leufkens, B. Stolk, P. Laing, R. Reed, T., & Ewen, M. (September 2007 - March 2008). *Direct Costs and Availability of Diabetes Medicines in Low-income and Middle-Income Countries*. Retrieved February 26, 2013, from http://apps.who.int/medicinedocs/en/m/abstract/Js18387en/.

Wang, J. Q., & Shiu, T. Y. (2004). Diabetes self-efficacy and self-care behaviour of Chinese patients living in Shanghai. *Journal of Clinical Nursing*, *13*, 771-772.

Weinger, K., Butler, H. A., Welch, G. W., & La Greca, A. M. (2005). Measuring diabetes self-care: a psychometric analysis of the self-care inventory-revised with adults. *Diabetes Care, 28*(6), 1346-1352.

Weinger, K. (2007). Psychosocial issues and self-care. *American Journal Nursing, 107*(6), 34-38.

West, J. D., & Goldberg, K. L. (2002). Diabetes self-care knowledge among outpatients at a veterans affairs medical center. *American Journal of Health-System Pharmacy, 59*(9), 849-852.

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care, 27*(5), 1047-1053.

Wu, S. F., Courtney, M., Edwards, H., McDowell, J., Shortridge-Baggett, L. M., & Chang, J. R. (2007). Self-efficacy, outcome expectations and self-care behaviour in people with type 2 diabetes in Taiwan. *Journal Compilation*, 250-257.

Pertanika J. Sci. & Technol. 22 (2): 489 - 505 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Reduced τ-adic Naf (RTNAF) Representation for an Efficient Scalar Multiplication on Anomalous Binary Curves (ABC)

Faridah Yunos*, Kamel Ariffin Mohd Atan, Muhammad Rezal Kamel Ariffin and Mohamad Rushdan Md Said

Institute for Mathematical Research, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

Elliptic curve cryptosystems (ECC) provides better security for each bit key utilized compared to the RSA cryptosystem. For this reason, it is projected to have more practical usage than the RSA. In ECC, scalar multiplication (or point multiplication) is the dominant operation, namely, computing nP from a point P on an elliptic curve, where n is an integer defined as the point resulting from adding $P + P + ... + P$, n times. However, for practical uses, it is very important to improve the efficiency of the scalar multiplication. Solinas (1997) proposes that the τ-adic Non-Adjacent Form (τ-NAF) is one of the most efficient algorithms used to compute scalar multiplications on Anomalous Binary curves. In this paper, we give a new property (i.e., Theorem 1.2) of τ-NAF(n) representation for every length, l. This is useful for evaluating the maximum and minimum norms occurring among all length-l elements of $Z(\tau)$. We also propose a new cryptographic method by using randomization of a multiplier n to $\bar{n}$ an element of $Z(\tau)$. It is based on τ-NAF. We focused on estimating the length of RTNAF($\bar{n}$) expansion by using a new method.

Keywords: Anomalous Binary Curves (Koblitz Curves), Scalar Multiplication, τ -adic Non-Adjacent Form, Norm.

Article history:
Received: 9 April 2012
Accepted: 17 December 2012

Email addresses:
Faridah Yunos (faridahy@upm.edu.my),
Kamel Ariffin Mohd Atan (kamel@upm.edu.my),
Muhammad Rezal Kamel Ariffin (rezal@upm.edu.my),
Mohamad Rushdan Md Said (mrushdan@upm.edu.my)
*Corresponding Author

INTRODUCTION

ECC is a cryptosystem that utilizes shorter keys as compared to the RSA and it also provides the same level of security at the same time. Thus, Vanstone (2006) suggested that ECC as the preferred asymmetric cryptosystem compared to RSA. Another advantage that is conjectured is that the elliptic curve discrete

ISSN: 0128-7702

logarithm problem is believed to be harder than the integer factorization problem. Given the best known algorithms to factor integers and compute the elliptic curve logarithms, the recommended key size is 106 bits compared to 512 bits for RSA which is considered to be an equivalent strength based on 10^4 MIPS years needed to recover one key as reported by A Certicom White Paper (1998). The scalar multiplication is the main cryptographic operation in ECC which computes $Q = nP$, where a point P is multiplied by an integer n resulting in another point Q on the elliptic curve. The computational cost of (n times) is therefore expressed as the number of field operations (additions, multiplications, inversions). The discrete logarithm problem is the basis for the security of many cryptosystems including ECC. This means that given points P and Q in the group, it is computationally infeasible to obtain n (i.e., the discrete logarithm of Q to the base P), if n is sufficiently large. Point multiplication is achieved by two basic elliptic curve operations:

1. Point addition, adding two points J and K to obtain another point (i.e., $L = J + K$) and
2. Point doubling, adding a point J to itself to obtain another point L (i.e., $L = 2J$).

Fig.1 and Fig.2 give the geometrical and analytical explanations of the point addition and the point doubling, respectively (Hankerson *et al*., 2004; Coron, 1999):

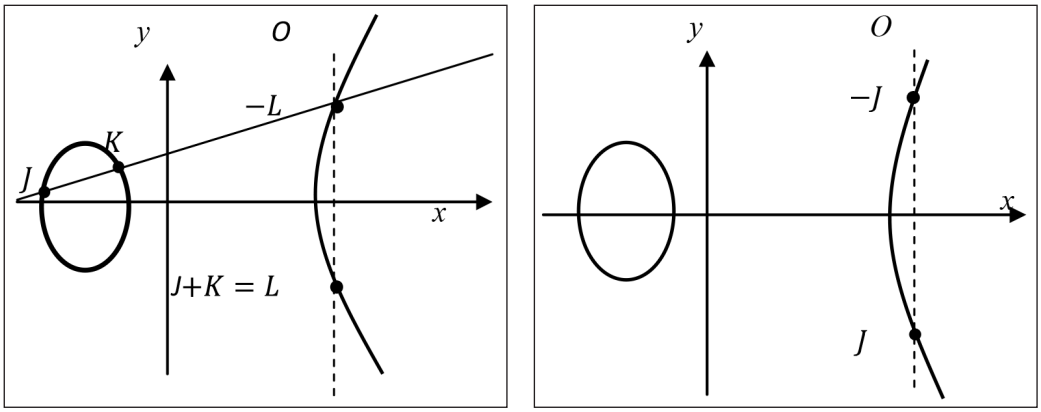


Fig.1: (Adding points in ECC): $J = (x_J, y_J)$ and $K = (x_K, y_K)$. Let $L = J + K$ where $L = (x_L, y_L)$, then $x_L = \lambda^2 - x_J - x_K$, $y_L = -y_J + \lambda(x_J - x_L)$, and $\lambda = \dfrac{y_J - y_K}{x_J - x_K}$ where λ is the slope of the line through J and K. If $K = -J$ (i.e. $K = (x_J, -y_J)$) then $J + K = O$. If $K = J$ then $J + K = 2J$ and point doubling equations are used. Also $J + K = K + J$.

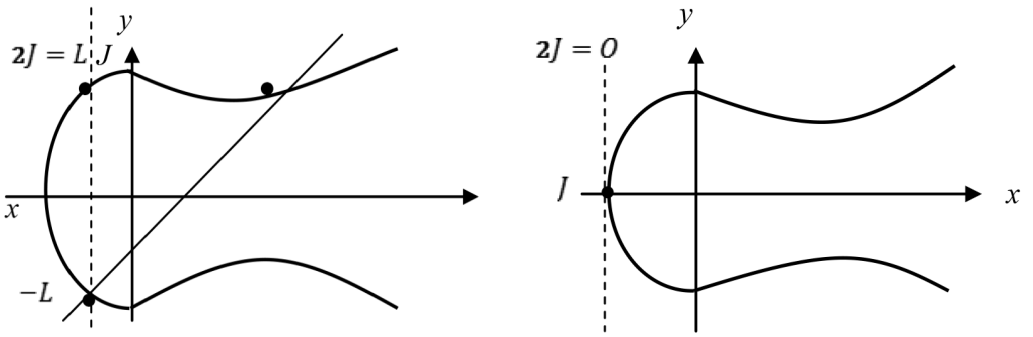


Fig.2: (Doubling points in ECC): Consider a point $J = (x_J, y_J)$ where $y_J \neq O$. Let $L = 2J$ where $L = (x_L, y_L)$. Then $x_L = \lambda^2 - 2x_J$, $y_L = -y_J + \lambda(x_J - x_L)$, and $\lambda = \frac{3x_J^2 + a}{2y_J}$ where λ is the tangent at point J and a is one of the parameters chosen with the elliptic curve.

Koblitz (1987) found that the *Koblitz curves* are a special type of curves for which the Frobenius endomorphism can be used for improving the performance of computing a scalar multiplication. The Koblitz curves are defined over F_2, as follows:

$$E_a : y^2 + xy = x^3 + ax^2 + 1$$

where $a \in \{0, 1\}$ as suggested by Koblitz (1992). The Frobenius map τ: $E_a(F_{2^m}) \rightarrow E_a(F_{2^m})$ for a point $P = (x, y)$ on $E_a(F_{2^m})$ is defined by

$$\tau(x, y) = (x^2, y^2), \qquad \tau(O) = O$$

where O is the point at infinity. It stands that $(\tau^2 + 2)P = t\tau(P)$ for all $P \in E_a(F_{2^m})$, where the trace, $t = (-1)^{1-a}$. Thus, it follows that the Frobenius map can be considered as a multiplication with complex number $\tau = \frac{t + \sqrt{-7}}{2}$ as stated earlier (Solinas, 2000).

Solinas produced an efficient procedure (i.e., Algorithm 3 in Solinas, 2000) for performing elliptic scalar multiplication and it requires about $\frac{m}{3}$ additions and no doubles. It gives at least 50% faster than any earlier version for operating on Koblitz curve, as shown in Table 1.

TABLE 1
Comparison of elliptic scalar multiplication

Method	Length of Expansion	Average Density	Average number of Elliptic Operations
Balanced expansion by Koblitz (1992)	$2m$	$\frac{3}{8}$	$3\frac{m}{4}$
Meier and Stafflebach (1993).	m	$\frac{1}{2}$	$\frac{m}{2}$
τ-NAF by Solinas (1997)	m	$\frac{1}{3}$	$\frac{m}{3}$

The computation of an average number of elliptic operations is dependent on the average density of non-zero coefficients among τ-NAF(n) representation (i.e., the number of non-zero coefficients divided by the length of expansion). Solinas (2000) estimated that the length of τ-NAF(n) is bounded by $\log_2 N(n) - 0.54626826939 < l < \log_2 N(n) + 3.5155941234$ when $l > 30$. This decision is achieved by obtaining the maximum and minimum norms occurring among all length-15 elements of Z(τ) by the direct evaluation method given by Solinas (2000, p. 213) although he does not give any detail on this method. The problem is how to get the maximum and the minimum norms of the element in Z(τ) that are more than 15 in length. In Section 1, a new property (i.e., Theorem 1.2) of τ-NAF(n) representation is given for every length, l. This is useful for evaluating the maximum and the minimum norms occurring among all the length-l elements of Z(τ) by using Equation [1.3]. In Section 2, it was observed that the property on estimation of length τ-NAF($\bar{n}$) with $l > 30$ (see Theorem 2.7). A new technique called the Reduced τ-adic Non-adjacent of Koblitz was also proposed for point multiplication for the Koblitz curves which focused on the estimation length of RTNAF.

τ- ADIC NON-ADJACENT FORM

This section begins with the meaning of a few definitions that are used in this study:

Definition 1.1

Let τ-NAF($\bar{n}$) = $\sum_{i=0}^{l-1} c_i \tau^i$ denote τ-adic Non-Adjacent Form for an integer an element of Z(τ) where l is the length of an expansion of τ-NAF($\bar{n}$), $c_i \in \{-1,0,1\}$ and $c_i c_{i+1} = 0$.

Definition 1.2

Let N: $Z(\tau) \rightarrow Z$ denote the norm function. If $x + y\tau$ is an element of Z(τ) then the norm is $x^2 + txy + 2y^2$.

It is well known that an integer of Z(τ) can be converted to τ-NAF form through Algorithm 1 in Solinas (2000). His algorithm is follows:

Algorithm 1.1 (τ-adic NAF)

Input : integers g_0, g_1
Output : τ-NAF
Computation:
Set $h_0 \leftarrow g_0$, $h_1 \leftarrow g_1$
Set $S \leftarrow \langle\ \rangle$
While $h_0 \neq 0$ or $h_1 \neq 0$
If h_0 odd
then
set $u \leftarrow 2 - (h_0 - 2h_1 \bmod 4)$
set $h_0 \leftarrow h_0 - u$

else

set $u \leftarrow 0$

Prepend u to S

Set $(h_0, h_1) \leftarrow \left(h_1 + \frac{th_0}{2}, -\frac{h_0}{2}\right)$

EndWhile

Output S

A question arises on the converse of the process. In this study, $\sum_{i=0}^{l-1} c_i \tau^i$ is transformed into an integer form with the length l in $Z(\tau)$. Thus, it is easy to get the norm of $\sum_{i=0}^{l-1} c_i \tau^i$. Furthermore, using this transformation, the maximum and minimum norms can be determined. First, the following theorem that gives an expansion of τ^i found in $\sum_{i=0}^{l-1} c_i \tau^i$ is presented.

Theorem 1.1

If $a_0 = 0$, $b_0 = 1$, $a_i = a_{i-1} + b_{i-1}$ and $b_i = -2a_{i-1}$ then

$$\tau^i = b_i t^i + a_i t^{i+1} \tau \qquad [1.1]$$

for $i > 0$.

Proof. We will give a proof by induction. If $i = 1$ then

$$\begin{aligned} \tau^1 &= b_1 t + a_1 t_2 \tau \\ &= -2a_0 t + (a_0 + b_0)\tau \\ &= \tau. \end{aligned}$$

So the equality [1.1] is verified for $i = 1$. Now, for $i = 2$,

$$\begin{aligned} \tau^2 &= b_2 t^2 + a_2 t^3 \\ &= -2a_1 t^2 + (a_1 + b_1)\, t^3 \tau \\ &= -2(a_0 + b_0)t^2 + (a_0 + b_0 - 2a_0)\, t^3 \tau \\ &= -2t^2 + t^3 \tau \\ &= t^2 + t\tau - 2 \\ &= t\tau - 2. \end{aligned}$$

So the equality [1.1] is verified for $i = 2$.
Assume that $\tau^k = b_k t^k + a_k t^{k+1} \tau$ is true up to $i = k$ where $k \geq 1$. Let us compute τ^{k+1}.

$$\begin{aligned} \tau^{k+1} &= \tau^k \bullet \tau \\ &= (b_k t^k + a_k t^{k+1} \tau)\, \tau \\ &= b_k t^k \tau + a_k t^{k+1}\, (t\tau - 2) \\ &= (b_k t^k + a_k t^{k+2})\, \tau - 2a_k t^{k+1}. \end{aligned}$$

Since $t^k = t^{k+2}t^{-2} = t^{k+2}$, we get

$$\begin{aligned}\tau^{k+1} &= (b_k t^{k+2} + a_k t^{k+2})\,\tau - 2a_k t^{k+1} \\ &= a_{k+1}t^{k+2}\tau + b_{k+1}t^{k+1}.\end{aligned}$$

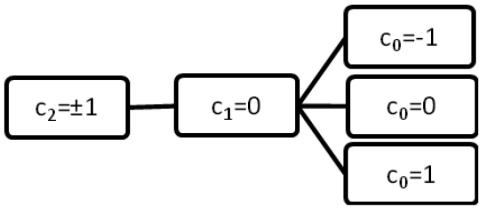
Therefore the relation [1.1] is true for all $i > 0$.

Now, let us start by making short analysis on τ-NAF that have length-3 as suggested in the following table.

TABLE 2
Combinations of c_0, c_1 and c_2 and the norm of $c_0 + c_1\tau + c_2\tau^2$

c_2	c_1	c_0	t	$r = c_0 - 2c_2$	$s = c_1 + c_2 t$	$N(r + s\tau)$
–1	0	–1	–1	1	1	2
–1	0	1	–1	3	1	8
–1	0	0	–1	2	1	4
–1	0	–1	1	1	–1	2
–1	0	1	1	3	–1	8
–1	0	0	1	2	–1	4
1	0	–1	–1	–3	–1	8
1	0	1	–1	–1	–1	2
1	0	0	–1	–2	–1	4
1	0	–1	1	–3	1	8
1	0	1	1	–1	1	2
1	0	0	1	–2	1	4

Table 2 shows all the combinations of c_0, c_1 and c_2 and the norm of $r + s\tau$ where $r + s\tau = c_0 + c_1\tau + c_2\tau^2$. From that table, we see that the maximum norm is 8 and the minimum norm is 2. There exist 12 combinations of c_0, c_1, c_2 and t to determine the maximum and minimum norm of $c_0 + c_1\tau + c_2\tau^2$. That combinations built based on the following tree diagram.



As such, there are 6 ways to arrange c_0, c_1 and c_2, while 2 ways are considered as the total number of t. Thus, there will be 12 combinations in total, as shown in Table 2. By using a similar method, all the outcomes of c_i can be obtained. Furthermore, the norms can also be obtained for every combination. Therefore, $\sum_{i=0}^{l-1} c_i \tau^i$ and the norm may be rewritten as follows:

Theorem 1.2

If $a_0 = 0$, $b_0 = 1$, $a_i = a_{i-1} + b_{i-1}$ and $b_i = -2a_{i-1}$ for $i > 0$ then

$$\sum_{i=0}^{l-1} c_i \tau^i = \sum_{i=0}^{l-1} c_i \left(b_i t^i + a_i t^{i+1} \tau\right) \qquad [1.2]$$

for $l > 0$.

Proof.

We will give a proof by induction. If $l = 1$ then $c_0 = c_0(b_0 t^0 + a_0 t^1 \tau) = c_0$.
We assume that if $l = k$ then $\sum_{i=0}^{k-1} c_i \tau^i = \sum_{i=0}^{k-1} c_i \left(b_i t^i + a_i t^{i+1} \tau\right)$ is true.
Now, if $l = k + 1$ then

$$\begin{aligned}
\sum_{i=0}^{k} c_i \tau^i &= c_0 + c_1 \tau + c_2 \tau^2 + \ldots + c_{k-1} \tau^{k-1} + c_k \tau^k \\
&= c_k \tau^k + \sum_{i=0}^{k-1} c_i \left(b_i t^i + a_i t^{i+1} \tau\right) \\
&= c_k \left(b_k t^k + a_k t^{k+1} \tau\right) + \sum_{i=0}^{k-1} c_i \left(b_i t^i + a_i t^{i+1} \tau\right) \\
&= \sum_{i=0}^{k} c_i \left(b_i t^i + a_i t^{i+1} \tau\right) \\
&= \sum_{i=0}^{(k+1)-1} c_i \left(b_i t^i + a_i t^{i+1} \tau\right)
\end{aligned}$$

Thus, [1.2] is true for $l = k + 1$ therefore it is true for all $l > 0$.

By using Definition 1.2 and Theorem 1.2, we obtain the norm of $\sum_{i=0}^{l-1} c_i \tau^i$ as follows.

$$\begin{aligned}
N\left(\sum_{i=0}^{l-1} c_i \tau^i\right) = &\left(\sum_{i=0}^{l-1} c_i b_i t^i\right)^2 + t\left(\sum_{i=0}^{l-1} c_i b_i t^i\right)\left(\sum_{i=0}^{l-1} c_i a_i t^{i+1}\right) \\
&+ 2\left(\sum_{i=0}^{l-1} c_i a_i t^{i+1}\right)^2
\end{aligned} \qquad [1.3]$$

Table 3 shows the maximum and the minimum norm of τ-NAF($\bar{n}$) occurring among all length-l elements of $Z(\tau)$ where $l = \{1,2,...,15\}$. The formula [1.3] for the norm obtained above can be made as a basis to estimate the length of τ-NAF representation, and the length of RTNAF expansion in this study. It can improve the technique to acquire the maximum and minimum norms of $\bar{n}$ by direct evaluation of all the length-15 element of $Z(\tau)$ mentioned by Solinas (2000, p. 213).

REDUCED τ-ADIC NAF

In this section, the elliptic scalar multiplication is developed on the Koblitz curve analogue of the binary method known as Reduced τ-adic Non-adjacent Form. Letting $\tau : (x,y) \rightarrow (x^2, y^2)$ the frobenius endomorphism where $\tau \in C$ will be an algebraic integer with $|\tau| > 1$. Routine

TABLE 3
The maximum and the minimum norms of τ-NAF($\bar{n}$) with l = {1,2,...,15}.

τ-NAF($\bar{n}$)	l	Maximum Norm	Minimum Norm
c_0	1	1	1
$c_0 + c_1\tau$	2	2	2
$(c_0 - 2c_2) + (c_1 + c_2t)\tau$	3	8	2
$(c_0 - 2c_2 - 2c_3t) + (c_1 + c_2t - c_3)\tau$	4	16	4
$(c_0 - 2c_2 - 2c_3t + 2c_4) + (c_1 + c_2t - c_3 - 3c_4t)\tau$	5	37	7
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t) + (c_1 + c_2t - c_3 - 3c_4t - c_5)\tau$	6	81	9
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6) + (c_1 + c_2t - c_3 - 3c_4t - c_5 - 5c_6t)\tau$	7	162	18
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t)$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7)\tau$	8	352	28
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8)$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t)\tau$	9	704	56
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t)$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9)\tau$	10	1421	112
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t + 34c_{10})$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9 - 11c_{10}t)\tau$	11	2921	197
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t + 34c_{10} + 22c_{11}t)$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9 - 11c_{10}t + 23c_{11})\tau$	12	5842	394
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t$ $+ 34c_{10} + 22c_{11}t - 46c_{12})$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9$ $- 11c_{10}t + 23c_{11} + 45c_{12}t)\tau$	13	11816	764
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t$ $+ 34c_{10} + 22c_{11}t - 46c_{12} - 90c_{13}t)$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9$ $- 11c_{10}t + 23c_{11} + 45c_{12}t - c_{13})\tau$	14	23662	1498
$(c_0 - 2c_2 - 2c_3t + 2c_4 + 6c_5t + 2c_6 - 10c_7t - 14c_8 + 6c_9t$ $+ 34c_{10} + 22c_{11}t - 46c_{12} - 90c_{13}t + 2c_{14})$ $+ (c_1 + c_2t - c_3 - 3c_4t - c_5 + 5c_6t + 7c_7 - 3c_8t - 17c_9$ $- 11c_{10}t + 23c_{11} + 45c_{12}t - c_{13} - 91c_{14}t)\tau$	15	47524	2996

72 in Solinas (2000) will be reused for division in $Z(\tau)$ and Algorithm 1.1 to transform an integer form in $Z(\tau)$ to τ-adic NAF expansion that was tested for its effectiveness for a long time. The division process is presented as follows:

Algorithm 2.1. (Division in $Z(\tau)$)

Input : the dividend $\gamma = u_0 + u_1\tau$ and divisor $\delta = v_0 + v_1\tau$
Output : the quotient $w = w_0 + w_1\tau$ and the remainder $z = z_0 + z_1\tau$
Computation:

$k \leftarrow u_0v_0 + tu_0v_1 + 2u_1v_1$
$l \leftarrow u_1v_0 - u_0v_1$
$N(\delta) \leftarrow v_0^2 + tv_0v_1 + 2v_1^2$

$$\lambda_0 \leftarrow \frac{k}{N(\delta)}$$

$$\lambda_1 \leftarrow \frac{l}{N(\delta)}$$

$(w_0w_1) \leftarrow$ Round (λ_0, λ_1) (Use Routine 60 in Solinas (2000) for rounding off λ_0 and λ_1)
$z_0 \leftarrow u_0 - v_0w_0 + 2v_1w_0$
$z_1 \leftarrow u_1 - v_1w_0 - v_0w_1 - tv_1w_1$
Output w_0, w_1, z_0, z_1

Our algorithm proceeds as follows:

Algorithm 2.2

(1) Consider $n \in Z$ and choose a random $\rho \in Z(\tau)$ such that $N(\rho) \leq N$ for some positive integer N;
(2) Compute the multiplier $\bar{n} \leftarrow n \mod\left(\rho\frac{\tau^m - 1}{\tau - 1}\right)$ by using Algorithm 2.1;
(3) Evaluate RTNAF of $\bar{n}$ as $\bar{n} \leftarrow n\sum_i \tau^i$ where $n_in_{i+1} = 0$ by using Algorithm 1.1;
(4) Compute Q as $Q \leftarrow \sum_i n_i\tau^i P$;
(5) Return Q.

Let $\rho = \rho_0 + \rho_1\tau$ in the element of $Z(\tau)$ and $\bar{n} \leftarrow n \mod\left(\rho\frac{\tau^m - 1}{\tau - 1}\right)$. We can find an integer ρ_0 and ρ_1 such that $N(\rho) \geq \frac{7N(n)}{4N\left(\frac{\tau^{m-1}}{\tau - 1}\right)}$ by using the following Lucas Sequence.

$$U_0 = 0,\ U_1 = 1 \qquad [2.1]$$

$$U_i = tU_{i-1} - 2U_{i-2} \quad \text{for} \quad i \geq 2 \qquad [2.2]$$

$$\tau^i = U_i\tau - 2U_{i-1} \qquad [2.3]$$

See Example 2.1 (in Appendix) in order to find ρ_0 and ρ_1.

Before multiplying the element ρ with $\frac{\tau^m - 1}{\tau - 1}$, $\frac{\tau^m - 1}{\tau - 1}$ has to be converted to $r_m + s_m\tau$ an element of $Z(\tau)$. The following theorem gives the formula of integers r_m and s_m.

Theorem 2.1

Let $\frac{\tau^m - 1}{\tau - 1} = r_m + s_m\tau$ an element of $Z(\tau)$ then $r_m = -2\sum_{i=1}^{m-2} U_i + 1$ and $s_m = \sum_{i=1}^{m-1} U_i$.

Proof.

$$\begin{aligned} r_m + s_m\tau &= \tau^{m-1} + \tau^{m-2} + \tau^{m-3} + \ldots + \tau^2 + \tau + 1 \\ &= (U_{m-1}\tau - 2U_{m-2}) + (U_{m-2}\tau - 2U_{m-3}) + (U_{m-3}\tau - 2U_{m-4}) \\ &\quad + \ldots + (U_2\tau - 2U_1) + (U_1\tau - 2U_0) + 1 \\ &= (U_{m-1} + U_{m-2} + U_{m-3} + \ldots + U_2 + U_1 + 1)\tau - 2(U_{m-2} + U_{m-3} \\ &\quad + U_{m-4} + \ldots + U_1 + U_0) + 1 \end{aligned}$$

Hence,

$$r_m = -2(U_{m-2} + U_{m-3} + U_{m-4} + \ldots + U_1) + 1 \qquad [2.4]$$

$$\equiv 1 \bmod 2 \qquad [2.5]$$

and

$$s_m = U_{m-1} + U_{m-2} + U_{m-3} + \ldots + U_2 + U_1 \qquad [2.6]$$

Now, using the formula s_m and r_m from Theorem 2.1, the following algorithm is constructed.

Algorithm 2.3: Conversion from $\rho\frac{\tau^m - 1}{\tau - 1}$ to $r + s\tau \in Z(\tau)$ by Lucas Sequence.

Input : Prime m, integer $t = (-1)^{1-a}$ for $a = 0$ or $a = 1$, nonzero elements ρ_0 and ρ_1 such that $\rho_0 + \rho_1\ \tau \in Z(\tau)$.
Output : $r + s\tau \in Z(\tau)$

(1) $U_0 \leftarrow 0,\ U_1 \leftarrow 1,\ s_0 \leftarrow 0,\ s_1 \leftarrow 1.$

(2) For i from 2 to m do $U_i \leftarrow tU_{i-1} - 2U_{i-2}$

(3) $s_m \leftarrow 1 + \sum_{i=2}^{m-1} U_i$

(4) $r_m \leftarrow -1 - 2\sum_{i=2}^{m-2} U_i$

(5) $s \leftarrow \rho_0 s_m + \rho_1 r_m + \rho_1 s_m t$

(6) $r_m \leftarrow \rho_0 r_m - 2\rho_1 s_m$

(7) Return (r, s)

The following theorem shows that the s_m in Step (3) Algorithm 2.3 is in mod 2 for $m \geq 4$.

Theorem 2.2

If $s_m = U_{m-1} + U_{m-2} + U_{m-3} + ... + U_2 + U_1$ then
$s_m \equiv$ for $t^{m-2} + t^{m-3} + t^{m-4} + ... + t^2 + t + 1 \bmod 2$ $m \geq 4$.

Proof.

$$\begin{aligned}
s_m &= U_{m-1} + U_{m-2} + U_{m-3} + ... + U_3 + U_2 + U_1 \\
&= (U_{m-2}t - 2U_{m-3}) + (U_{m-3}t - 2U_{m-4}) + (U_{m-4}t - 2U_{m-5}) + ... + (U_2t - 2U_1) + (U_1t - 2U_0) + U_1 \\
&= t(U_{m-2} + U_{m-3} + U_{m-4} + ... + U_2 + U_1) + 1 - 2(U_{m-3} + U_{m-4} + U_{m-5} + ... + U_1) \\
&= t((U_{m-3}t - 2U_{m-4}) + (U_{m-4}t - 2U_{m-5}) + (U_{m-5}t - 2U_{m-6}) + ... + U_1t + U_1) + 1 - 2(U_{m-3} + \\
&\quad U_{m-4} + U_{m-5} + ... + U_1) \\
&= t^2(U_{m-3} + U_{m-4} + U_{m-5} + ... + U_2 + U_1) + t + 1 - 2(U_{m-3} + 2U_{m-4} + 2U_{m-5} + ... + 2U_1) \\
&= t^2((U_{m-4}t - 2U_{m-5}) + (U_{m-5}t - 2U_{m-6}) + (U_{m-6}t - 2U_{m-7}) + ... + U_1t + U_1) + 1 - 2(U_{m-3} + \\
&\quad 2U_{m-4} + 2U_{m-5} + ... + 2U_1) \\
&= t^3(U_{m-4} + U_{m-5} + U_{m-6} + ... + U_1) + t^2 + t + 1 - 2(U_{m-3} + 2U_{m-4} + 3U_{m-5} + ... + 3U_1) \\
&= t^{m-2}U_{m-(m-1)} + t^{m-3} + t^{m-4} + ... + t^2 + t + 1 - 2(U_{m-3} + 2U_{m-4} + 3U_{m-5} + ... + (m-3)U_1).
\end{aligned}$$

Therefore, $s_m \equiv t^{m-2} + t^{m-3} + t^{m-4} + ... + t^2 + t + 1 \bmod 2$.

We can simplify Step (4) in Algorithm 2.3 by letting $r_m = U_m$ when $m \geq 3$ and $t = 1$.

Theorem 2.3

If $r_m = -2(U_{m-2} + U_{m-3} + U_{m-4} + ... + U_1) + 1$ and U_m is defined as in equation [2.2] then $r_m = U_m$ for $m \geq 3$ and $t = 1$.

Proof. We prove by using mathematical induction. Take $m = 3$, we have

$$r_3 = -2U_1 + 1 = -1 \text{ and } U_3 = U_2 - 2U_1 = (U_1 - 2U_0) - 2U_1 = -1$$

Therefore, for $m = 3$, $r_3 = U_3$. So the result is true for $m = 3$.
Now, our assumption asserts that $r_k = U_k$ is true for $m = k$.
Lastly, we prove that $r_{k+1} = U_{k+1}$ is true for $m = k + 1$.
We have,

$$\begin{aligned}
r_{k+1} &= -2(U_{k-1} + U_{k-2} + U_{k-3} + ... + U_1) + 1 \\
&= -2U_{k-1} - 2\,(U_{k-2} + U_{k-3} + ... + U_1) + 1 \\
&= -2U_{k+1} + r_k \\
&= -2U_{k-1} + U_k \qquad \text{using the inductive hypothesis} \\
&= U_{k+1}.
\end{aligned}$$

Hence, the result is true for all $m \geq 3$.

We give two properties of ρ as follows.

Theorem 2.4

If $\rho = \rho_0 + \rho_1\tau \in Z(\tau)$ and ρ_0 is even then f is even such that $\rho \mid (f + e\tau)$.

Proof.
Suppose $\rho_0 = 2k$ where $k \in Z$.

$$\begin{aligned}\rho(c + d\tau) &= (2k + \rho_1\tau)(c + d\tau)\\ &= 2ck + (2dk + \rho_1 c)\tau + \rho_1 d\tau^2\\ &= 2ck + (2dk + \rho_1 c)\tau + \rho_1 d(t\tau - 2)\\ &= 2(ck - \rho_1 d) + (2dk + \rho_1 c + t\rho_1 d)\tau.\end{aligned}$$

Suppose $f = 2\,(ck - \rho_1 d)$ and $e = 2dk + \rho_1 c + t\rho_1 d$. We get f is even where $ck - \rho_1 d \in Z$.

Theorem 2.5

If $\rho = \rho_0 + \rho_1\tau \in Z(\tau)$ and ρ_0 and ρ_1 are even then f and e are even such that $\rho \mid (f + e\tau)$.

Proof. Suppose $\rho_0 = 2k_1$ and $\rho_1 = 2k_2$ where $k_1, k_2 \in Z$.

$$\begin{aligned}\rho(c + d\tau) &= (2k_1 + 2k_2\tau)(c + d\tau)\\ &= 2k_1 c + 2(dk_1 + k_2 c)\tau + 2k_2 d\tau^2\\ &= 2k_1 c + 2(dk_1 + k_2 c)\tau + 2k_2 d(t\tau - 2)\\ &= 2(ck_1 - 2k_2 d) + 2(dk_1 + k_2 c + tk_2 d)\tau.\end{aligned}$$

Let $f = 2\,(ck_1 - 2k_2 d)$ and $e = 2(dk_1 + k_2 c + tk_2 d)$. Hence, we prove that f and e are even where $ck_1 - 2k_2 d,\ dk_1 + k_2 c + tk_2 d \in Z$.

If we change f and e by r and s respectively, the norm of $r + s\tau$ *is* obviously an even as well. It is helpful to observe the number of points in modulo $r + s\tau$ and the average number of non-zero coefficients in RTNAF.

Before presenting the estimation of length-l of τ-NAF($\bar{n}$), the bounds of norm of $\bar{n}$ is given as been shown in Theorem 2 of Solinas (2000). The theorem is rewritten as follows:

Theorem 2.6

Suppose that $N_{\min}(d)$ denote the minimal norm and $N_{\max}(d)$ denote the maximal norm occurring among all length-d elements of $Z(\tau)$. Let $l > 2d$, and let $\bar{n}$ be a length-l element of $Z(\tau)$. Then

$$\left(\sqrt{N_{\min}(d)} - \frac{\sqrt{N_{\max}(d)}}{2^{\frac{d}{2}} - 1}\right)^2 \cdot 2^{l-d} < N(\bar{n}) < \frac{N_{\max}(d)}{\left(2^{\frac{d}{2}} - 1\right)^2} \cdot 2^l.$$

Combining formula [1.3] and Theorem 2.6, the main result of this section can be obtained.

Theorem 2.7

The length l of τ-NAF($\bar{n}$) is bounded by

$$\log_2 N(\bar{n}) - 0.54626826939 < l < \log_2 N(\bar{n}) + 3.5155941234$$

when $l > 30$.

Proof.

We choose $d = 15$, there are 43692 combinations of c_i and t to determine the maximum and minimum norm of $\sum_{i=0}^{14} c_i \tau^i$. The norm of all the combinations is evaluated by using formula [1.3]. That is,

$$N\sum_{i=0}^{14} c_i \tau^i = \left(\sum_{i=0}^{14} c_i b_i t^i\right)^2 + t\left(\sum_{i=0}^{14} c_i b_i t^i\right)\left(\sum_{i=0}^{14} c_i a_i t^{i+1}\right) + 2\left(\sum_{i=0}^{14} c_i a_i t^{i+1}\right)^2$$

thus, we get $N_{\max}(15) = 47324$ and $N_{\min}(15) = 2996$.

Now, using Theorem 2.6, the following can be obtained:

$$\left(\sqrt{2996} - \frac{\sqrt{47324}}{2^{\frac{15}{2}} - 1}\right)^2 \cdot 2^{l-15} < N(\bar{n}) < \frac{47324}{\left(2^{\frac{15}{2}} - 1\right)^2} \cdot 2^l$$

$$0.087438100867 \cdot 2^l < N(\bar{n}) < 1.4603035291 \cdot 2^l$$

$$\log_2 N(\bar{n}) - 0.54626826939 < l < \log_2 N(\bar{n}) + 3.5155941234 \qquad [2.7]$$

when $l > 30$.

Although the limitation of the length of τ-NAF through the above theorem is similar to Solinas's study (Solinas, 2000), the approach used by him to get the maximum and the minimum norm was different from the one used in the present study. The result of this length limitation will be implemented on estimating the length of RTNAF expansion of an integer in $Z(\tau)$ in the following discussion.

Theorem 2.8

The length $\bar{l}$ of RTNAF($\bar{n}$) satisfies $\bar{l} \leq \log_2 N(\rho) + m + a$ for $\bar{l} > 30$.

Proof.

We have $N\left(\frac{\tau^m - 1}{\tau - 1}\right) = 2^{m-2+a} + O(2^{\frac{m}{2}})$ as given on page 224 of Solinas' (2000). Then, using $N(\bar{n}) \leq \frac{4}{7} N(\rho) N\left(\frac{\tau^m - 1}{\tau - 1}\right)$ the following is obtained:

$$N(\bar{n}) \leq N(\rho)\left[\frac{2^{m+a}}{7} + O(2^{\frac{m}{2}})\right]$$

$$\log_2 N(\bar{n}) \leq \log_2 N(\rho) + \log_2\left[\frac{2^{m+a}}{7} + O(2^{\frac{m}{2}})\right]$$

$$< \log_2 N(\rho) + m + a - \log_2 7 + \log_2 O(2^{\frac{m}{2}}) \qquad [2.8]$$

From equations [2.7] and [2.8], we get,

$$\bar{l} < \log_2 N(\rho) + m + a + 0.7082392013 \qquad [2.9]$$

Since $\bar{l}$ is an integer, it follows that $\bar{l} < \log_2 N(\rho) + m + a$ for $\bar{l} > 30$.

Example 2.2 (see the Appendix) is an illustration from the above theorem.

Since the hamming weight (i.e. the number of non-zero coefficients) of a scalar representation is the product of its length and density (i.e., the average of hamming weight), our bound of will help to estimate the hamming weight of scalar based on RTNAF.

Now, we give an illustration of Algorithm 2.2. Let us choose a = 1 and m = 5 to get a random ρ = 56000 + 50000τ such that $N(\rho)$ = 10936000000. Convert a divisor $\rho\frac{\tau^m - 1}{\tau - 1} = (56000 + 50000\tau)\frac{\tau^5 - 1}{\tau - 1}$ to –262000 + 144000τ by using Algorithm 2.3 and consider n = 60000000001 as a dividend. Now, compute the multiplier $\bar{n} \leftarrow n \bmod \rho\frac{\tau^m - 1}{\tau - 1}$ by using Algorithm 2.1, we obtain that the quotient is –58855 + 130678τ and the remainder is –151999 – 6000τ. Evaluate RTNAF of $\bar{n}$ as $\bar{n} \leftarrow n\sum_i \tau^i$ where $n_i n_{i+1} = 0$ by using Algorithm 1.1, we have, RTNAF(–151999 – 6000τ) =< 1,0,0,0,0,–1,0,0,0,0,–1,0,–1,0,0,1,0, –1,0,0,1,0,–1,0,0,0,1,0,–1,0,1,0,0,–1,0,0,–1> with the length is 37. Finally, we can compute the scalar multiplication by implementing any efficient algorithm such as Algorithm 3.66 in Hankerson *et al*. (2004).

CONCLUSION

As a conclusion, the new property of τ-NAF(n) representation for every length, l, as in Theorem 1.2, is useful for evaluating the maximum and the minimum norms occurring among all the length-l elements of $Z(\tau)$. An estimation of the length of RTNAF($\bar{n}$) expansion has also been presented. Retrieval from Theorem 2.8 is important to get the average of the non-zero coefficients in the RTNAF expansion that becomes the subject of our future discussion. Therefore, we can observe the effectiveness of Algorithm 2.2 in scalar multiplication as compared to the ordinary τ-NAF.

REFERENCES

A Certicom White Paper (1998). The Elliptic Curve Cryptosystems for Smart Cards. *The Seventh in a Series of ECC White Papers*. Certicom Corp, San Matco, California. http://www.certicom.com.

Coron, J. S. (1999). Resistence Against Differential Power Analysis for ECC. In C. Koc and C. Paar (Eds.), *Cryptographic Hardware and Embeddded Systems (CHES'*99*)*, *Lecturer Notes in Computer Science* 1717 (pp. 292-302). Springer-Verlag.

Gordon, D. M. (1998). A Survey of Fast Exponentiation Methods. *Journal of Algorithms 27* (*Article no AL970913*) 129-146.

Hankerson, D., Menezes, A., & Vanstone, S. (2004). Guide to Elliptic Curve Cryptography. Springer-Verlag.

Koblitz, N. (1987). Elliptic Curve Cryptosystem. *Mathematics Computation, 48*(177), 203-209.

Koblitz, N. (1992). CM Curves with Good Cryptographic Properties. *Advance in Cryptology, Proc. Crypto'91, Lecture Notes in Computer Science* 576 (pp. 279-287). Springer-Verlag.

Meier, W., & Stafflebach, O. (1993). Efficient Multiplication on Certain Non-supersingular Elliptic Curves. *Advance in Cryptology, Proc. Crypto'*92*, Lecture Notes in Computer Science, 740* (p. 333-344). Springer-Verlag.

Miller, V. S. (1986). Use of Elliptic Curve in Cryptography. In H.C. Williams (Ed.), *Advance in Cryptology, Proc.Crypto'*85, *Lecture Notes in Computer Science* 218 (pp. 417-426). Springer-Verlag.

Solinas, J. A. (1997). An Improved Algorithm for Arithmetic on a Family of Elliptic Curves. In B. Kaliski (Ed.), *Advance in Cryptology, Proc. CRYPTO'97*, *Lecture Notes in Computer Science* 1294 (pp. 357-371). Springer-Verlag.

Solinas, J. A. (2000). Efficient Arithmetic on Koblitz Curves. *Design, Codes, and Cryptography*, *9*, 195-249. Netherlands: Kluwer Academic Publishers, Boston.

Vanstone, S. (2006). ECC Holds Key to Next Generation Cryptography. Retrieved from http://www.design-reuse.com/articles/7409/ecc-hold-key-to-next-gencryptography.html.

APPENDIX

Example 2.1

Find ρ_0 and ρ_1 for $\bar{n} \equiv 10 \bmod \left(\rho \dfrac{\tau^3 - 1}{\tau - 1}\right)$ and $t = 1$.

Solution.

Let $r_3 + s_3\tau = \dfrac{\tau^3 - 1}{\tau - 1}$. Now,

$$\begin{aligned} r_3 + s_3\tau &= \tau^2 + \tau + 1 \\ &= (U_2\tau - 2U_1) + \tau + 1 \\ &= (tU_1 - 2U_0)\tau - 2U_1 + \tau + 1 \\ &= 2\tau - 1. \end{aligned}$$

Thus, we get $r_3 = -1$ and $s_3 = 2$.

Therefore, $N\left(\dfrac{\tau^3 - 1}{\tau - 1}\right) = N(-1 + 2\tau) = (-1)^2 + 1 \cdot (-1) \cdot 2 + 2 \cdot 2^2 = 7$
and $N(10) = 100$.

Now, consider $N(\rho) \geq \dfrac{7N(10)}{4N(r_3 + s_3\tau)}$ to get the value ρ_0 and ρ_1. The inequality admits $\rho_0^2 + \rho_0\rho_1 + 2\rho_1^2 \geq 25$ a solution if the discriminant $\rho_1^2 - 4(2\rho_1^2 - 25) = 100 - 7\rho_1^2$ is a positive integer, which means that $|\rho_1| \leq \dfrac{2(25)}{\sqrt{7}}$.

For a fixed positive integer $\rho_1 \leq \dfrac{2(25)}{\sqrt{7}}$, the solutions of the inequality $\rho_0^2 + \rho_0\rho_1 + 2\rho_1^2 \geq 25$ are the integers $\rho_0 \geq \dfrac{-\rho_1 + \sqrt{100 - 7\rho_1^2}}{2}$ or $\rho_0 \leq \dfrac{-\rho_1 - \sqrt{100 - 7\rho_1^2}}{2}$.

Let us choose the fixed $\rho_1 = 3$ such that $\rho_1 \leq \dfrac{2(25)}{\sqrt{7}}$. Then, by the inequality $\rho_0 \geq \dfrac{-\rho_1 + \sqrt{100 - 7\rho_1^2}}{2}$ we get $\rho_0 \geq \dfrac{-3 + \sqrt{17}}{2}$. We can take $\rho_0 = 2$. Lastly, we find that $\rho = 2 + 3\tau$ and $N(\rho) = 28$ where clearly $N(\rho) \geq 25$.

Example 2.2

Give an estimation of $\bar{l}$ when $\rho = -6000 + 6000\tau$, $m = 5$ and $a = 1$..

Solution.

Since $a = 1$ then $t = -1$, we have

$$N(-6000 + 6000\tau) = (-6000)^2 + (-1)(-6000)(6000) + 2(6000)^2$$
$$= 144000000.$$

By equation [2.9].

$$\bar{l} < \log_2 144000000 + 5 + 1 + 0.7082392013$$
$$\bar{l} < 33.8097327721$$

Since the length $\bar{l}$ is an integer, we estimate that $30 < \bar{l} \leq 33$.

Pertanika J. Sci. & Technol. 22 (2): 507 - 527 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Relevance of Integrated Geophysical Methods for Site Characterization in Construction Industry – A Case of Apa in Badagry, Lagos State, Nigeria

K. S. Ishola[1,2]*, L. Adeoti[1], F. Sawyerr[1] and K. A. N Adiat[2,3]

[1]*Department of Geosciences, University of Lagos, Akoka, Lagos, Nigeria*

[2]*School of Physics, Geophysics Program, Universiti Sains Malaysia, 11800 Penang, Malaysia*

[3]*Applied Geophysics Department, Federal University of Technology, Akure, Ondo, Nigeria*

ABSTRACT

Detailed geophysical investigations have been carried out using integrated geophysical methods with a view to characterising the subsurface lithologic features that might indicate suitable places for structural developments. An overview of the subsurface resistivity distribution has been achieved employing 8Vertical Electrical Soundings with the Schlumberger array and 4 2D resistivity imaging using Wenner array. In order to constrain the results of the electrical resistivity methods, we carried out a ground magnetic survey along E-W direction using the Proton precession magnetometer at 1m sampling interval. Analysis of well logs data available and VES results showed 4 to 5 geoelectric layers corresponding to sand, clayey sand, clay, silty sand and sandy clay. The 2D resistivity imaging sections showed relative decrease of apparent resistivity with depth implying a geological transition from sand with high resistivity value of about 508Ωm to clay with low resistivity value 16Ωm at depths of 0-20m and 25-50m respectively. The magnetic profiles showed that the study area was characterised by short wavelengths and amplitudes ranging from –3800 to 700 nT. The highs and lows of the magnetic responses occasioned by lithological variations and structural features were magnetically resolved. In view of the identified subsurface structures, the suggested depth to the competent layer is about 20m for low to medium structures while above 50m would be suitable for heavy or massive engineering structures. The use of integrated geophysical methods for the delineation, identification and imaging of the subsurface geological structures which could provide clues to the nature and type of foundation suitable for the development of the study area has been successfully achieved.

Keywords: Electrical resistivity, Wenner array, sand, magnetic anomaly, constrain, foundation design.

Article history:
Received: 4 June 2012
Accepted: 25 February 2013

Email addresses:
K. S. Ishola (saidisho@yahoo.co.uk),
L. Adeoti (lukuade@yahoo.com),
F. Sawyerr (fsawyyer12@yahoo.com),
K. A. N Adiat (adiatnafiu@yahoo.co.uk)
*Corresponding Author

ISSN: 0128-7702

INTRODUCTION

Lagos is a cosmopolitan city inhabited by millions of people from various parts of the world. At the economic level, it is Nigeria's economic capital as well as her financial and commercial nerve centre. The position of Lagos as a regional financial hub is universally acknowledged. Despite the relocation of Nigeria's capital from Lagos to Abuja, Lagos still holds so much potential that nowhere in Nigeria can be better suited (Peterside, 2007).

Recently, Lagos joined the Megacity group of cities like Tokyo, Bombay, New York and Los Angeles with a population estimation of about 18 million people. With such a dense population, Lagos can be said to have become a teeming tangle of humanity and enterprise (Dekolo & Oduwaye, 2011).Globally, megacities due to their huge population, have attendant issues and problems of inadequate socio-economic infrastructure necessary to sustain their population. For decades, the city of Lagos has grappled with problems such as gross shortage of housing, insufficient road networks and acute potable water supply.

In an effort to address this socio-economic malice, the state government through public and private partnership has embarked on various developmental programmes in the city with a view to drawing people away from the congested centre and have them shift to the suburbs. Badagry, one of the suburbs that is about 45km from the city, is expected to witness a huge influx of people not only because of its strategic location but also due to the fact that it isa trading outlet for Trans-Saharan Trade Routes and an inlet for the Trans-Atlantic Trade (Mabogunje, 1971).

Following this trend, and the fact that the number of people living in Lagos, Nigeria's fastest growing city, is expected to reach the 24 million mark in 2015 (Dekolo & Oduwaye, 2011), there is a need for adequate town planning projects in the suburbs of Lagos and Badagry, in particular, which should be focused on delivery of residential accommodation in addition to other social infrastructures.

Against this background, we embarked on the use of integrated geophysical methods involving electrical resistivity and magnetic methods to define the stratigraphy of the subsurface structural features and identified suitable place for housing developments in the study area. Although geotechnical methods may be used, the high cost of using such methods and the distortion of the subsurface compositions which might result compared to the outcome of using the geophysical methods employed in this study favour the latter.

SITE AND GEOLOGY

The study area was about 15km from the Badagry roundabout on the Seme Border route. It is located between longitudes2° 52'50" to 2°53' 40" East of the Greenwich meridian and latitudes 6°.24'30" to 6°25' 20"North of the Equator. It can be accessed via the Lagos-Badagry expressway, leading to Badagry Town. It is bounded on the westby Porto Novo and Weme; on the north by Ilogbo, Ipokia; on the south by the Osa lagoon and the Atlantic ocean and on the east by the Awori settlements of Ojo and Lagos (Akran,2001). The stratigraphy of the Dahomey Basin has been extensively discussed with various works, and several classification schemes have been proposed by notable authors (Jones & Hockey, 1964; Omatosola & Adegoke,1981; Coker *et al*., 1983; Billman 1992; Elueze & Nton, 2004).

The stratigraphy sequence is classified under 5major formations in terms of their geological formation age that include the Littoral and Lagoon deposits, Coastal Plain sand, the Ilaro formation, the Ewekoro formation and Abeokuta overlying the crystalline basement complex; their ages range from Recent to Cretaceous. With the exception of Ilaro, the other 4of these formations constitute aquifers in the Dahomey Basin from which the geological section of Lagos was drawn. Lagos state is underlain by a geological sequence composed of Sands, clayey sands, Clay/Shale, Sandy clay typical of Alluvium and Coastal plain sands (Fatoba *et al*., 2004).

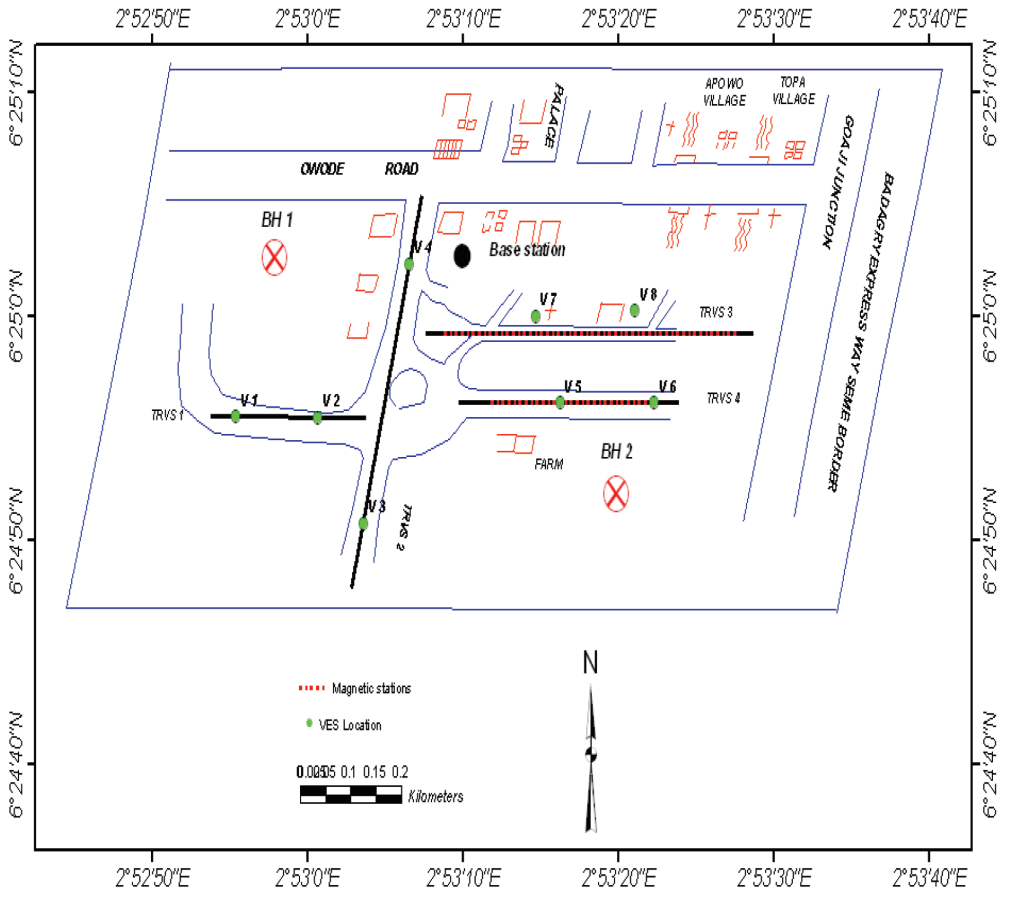


Fig.1: Location Map of the Study Area Showing the Traverses, VES Points and Borehole Locations

DATA ACQUISITION

Electrical Resistivity

In electrical resistivity method, electric current was injected into the ground through a pair of current electrodes. A second pair of electrodes was then used to measure the voltage resulting from the flow of current in the current electrode into the ground. Measurements were taken using Vertical Electrical Sounding and 2-D Electrical Resistivity Imaging techniques. Four

traverses were established in the field of study (Figure 1). ABEM SAS 1000 Terrameter was used for the data acquisition. A total number of 8vertical electrical sounding (VES) positions were occupied using the Schlumberger electrode array. Along traverses 1,2 and 3we occupied VES points 1 and 2, VES 3 and 4, VES 7 and 8 respectively while VES 5 and 6werecarried out along traverse 4 (Fig.1). The current electrode spread (AB) varied from 2m to a maximum of 600m.

Current (I) was injected into the ground through two current electrodes and the resulting potential difference (V) was measured through another pair of electrodes called the potential electrodes. The obtained I and V were used to determine the apparent resistivity. The r.m.s error for iterations of resistivity data are shown in Figures 2a and2b. For the 2D survey, 4electrodes were used at electrode separation (a) of 10m. The Wenner array configuration was utilized for all the traverses. The midpoint of the 4electrodes was mapped i.e. measurements of the apparent resistivity values were taken at the mid-point of the four electrodes. Measurements were repeated as the array was moved along this profile while the electrodes were maintained at fixed separation until the end of the traverse (say, traverse 1). On completion of the data acquisition for a = 10 m, the four electrodes were moved to the starting point of the same traverse, and measurements were taken with a=20m, 30m and 40m. This process was repeated for the other 3traverses.This technique assisted in obtaining both the lateral and vertical changes in the subsurface formations.

The high resolution for imaging the subsurface and relative depth of penetration favoured the choice of Wenner electrode configuration (Loke, 2004). The apparent resistivity was obtained by using the equation:

$$\rho_a = 2\pi aR \tag{1}$$

where is the apparent resistivity measured in Ωm and a is the electrode separation and the measured resistance of the field is denoted by R measured in ohm (Telford *et al.*, 1990).

Magnetic Method

Since many geophysical interpretations may fit the observed data and a given interpretation may not be unique, it is always useful to use other methods in the same area to constrain the interpretation. Thus, we deployed the magnetic method as a reconnaissance tool to constrain the electrical resistivity method in order to give a clue to the number of the VES point to be sounded in the study area. Two magnetic traverses trending approximately in the E-W (Figure 1) directions were established. Magnetic survey involved taking the measurement of the total component of the earth's magnetic field using a G-856 Proton precession magnetometer. A base station, which is about 20m away from traverse 3,was established with readings taken before commencement of data capturing and immediately after the traverses had been occupied to allow for diurnal corrections. A total of 85 magnetic stations were covered with 3readings taken at each station and averaged. The operation and principles of the magnetometer are explained by Breiner (1973) and Telford *et al*. (1990).

DATA PROCESSING AND INTERPRETATION

Vertical Electrical Sounding

The VES curves shown in Figures 2a and2b were obtained by plotting the apparent resistivity against half electrode spacing (AB/2) and were interpreted by the partial curve matching method and computer iteration techniques. Theoretical derivations and practical tests (Barker 1989, Pozdnyakova *et al*., 1996, Banton *et al*., 1997) have shown that the approximate penetration depth can be considered as $^1/_6$ of the current electrode spread (AB) for the arrays of Schlumberger and Wenner types used on wide ranges of soils and grounds. The forward modelling computes the true resistivity of the models. This is accomplished by the programme RESIST using the linear filter theory (Ghosh,1971a,b) while the inverse modelling tries to obtain a theoretical model whose apparent resistivity curve matches a set of field data to a good approximation. The inverse modelling is carried out applying Marquardt's algorithm (Marquart, 1963) to an initial model which is modified repeatedly until it matches with the field curve.

2 D ElectricalResistivity Imaging

In order to convert the apparent resistivity data sets of the Wenner array into 2D resistivity images, the data sets were inverted using the DIPPRO inversion software. The programme algorithm calculates the forward responses using a finite element method. After the inversion process, the results were displayed as measured apparent resistivity pseudo-sections at the top and the resulting inverted true resistivity 2D section as the bottom panel. The result of the interpretation of the data acquired is presented as pseudo- sections shown in Figures 5 and 6. It uses the numerical approach to optimise the initial multilayer model constructed directly from the observed apparent resistivity values.

Magnetic Survey

The ground raw magnetic survey measurements were subjected to relevant corrections. Since the area is characterized by smooth terrain and no elevation differences were noticeable between the stations, terrain and elevation corrections were not applied. The international geomagnetism reference field (IGRF) value of 32000nT for the study area was subtracted from the readings of the survey stations. The data acquired from the field werefiltered using drift correction equations given as:

$$B_{drift} = B_e - B_s \tag{2}$$

$$\Delta T = T_e - T_s \tag{3}$$

where B_e and B_s and are the initial magnetic response and final magnetic response readings respectively, T_e and T_s are the initial and final time measurements and is the time difference.

The filtered data were plotted as magnetic profiles using the Microsoft Excel spreadsheet package. The plots were interpreted by visual inspection and done so quantitatively using Peter's slope index given by:

$$S = 1.6h \quad (4)$$

where S is the distance between the points of inflexion and h is the depth to the anomalous structure (Telford *et al*., 1990).

RESULTS AND DISCUSSION

Geoelectrical Sections

The 1-D geoelectric sections resulting from the interpretation of the resistivity data given in Figures 3 and 4 comprise VES 1- 8. These geo-electric sections revealed 4to 5geo-electric layers. The first layer, the topsoil, is characterised by resistivity values ranging from 334.7Ωm to 5167.0Ωm with associated thicknesses varying from 0.5m to 1.1m. The second substratum in VES 1, 3, 4 and 5 denotes silty sand with thicknesses between 0.3m and 4.3m, and resistivity values from 1232.5Ωm to 2577.6Ωm.The corresponding layer in VES 2, 6, 7 and 8 typifies sand with thicknesses ranging from 1.5m to 4.3m, and resistivity between 369.2Ωm to 912Ωm.The third horizon in all the VES is typical of sand with resistivity ranging from 350.7Ωm to 582Ωm except for VES 2,3 and 4, which is presumptuous of clayey sand. The associated thicknesses are between 4.9m and 14.0m.

On account of the resistivity and thickness, this layer is promising for citing engineering structures. The fourth layer in VES 1,2 and 3 is symptomatic of sandy clay with resistivity values of 132.7Ωm, 176.6Ωmand 99.4 Ωmand layer thicknesses of 49.1m,16.6m and 27.4m respectively. It can serve as a good layer for foundations. The corresponding layer in VES 6 and 8 is typical ofclay, with resistivity in the range of 50.3Ωm to 56.8Ωm, and thicknesses ranging from 12.6m to 15.7m.This same layer in VES 4, 5 and 7 denotes clay, with resistivity values ranging from 36.8Ωm to 40.5Ωm and thicknesses between 11.6m and 43.5m.

The fifth layer in all the VESs except VES 5 is characterised by high resistivity values especially VESs 3,4,7 and 8 with resistivity from 3633.0 Ωm to 6339.2 Ωm representative of gravelly sand. The resistivity value of the substratum of VES 5 is 807 Ωm indicative of sand while VESs 1,2 and 6 correspond to silty sand with resistivity of 1039.1 Ωm to 1231.5 Ωm. The current penetration terminated at this layer due to the limitation in the length of the traverses covered and consequently, the thickness of this layer could not be determined. However, for a deeper investigation of the subsurface, the length of the traverses should be increased.

By visual inspection, the second layer under VES points 1and 3 together with the first layer beneath VES 4 correlates well with the first layer of the borehole log (BH1) shown in Figure3. Also, on comparison of the geoelectric section B-B' with the borehole lithologic log (BH2) in Figure 1, it can be seen that the second layer for VES 6,7 and 8 has a similar soil type(sand) with the borehole at a depth of 12m. In addition, the fourth layer underneath VES 5, which matches with the third layer for VES 6, 7 and 8,has a similar composition with the second layer of the borehole.

2D Resistivity Imaging

A more accurate geological model has been established by means of a 2D resistivity imaging section (Figures 5 and 6). The resistivity cross sections resulting from the inversion of the geoelectrical data exhibit significant variations in resistivity values with depths. The inverted resistivity images made it possible to obtain information on the variations of resistivity to a depth of about 50m.The upper image is the pseudo-section data, while the lower one is the inverted image of the raw data. The models obtained for all the traverses show similar stratagraphical units corresponding to sand, sandy clay and clay deposits.

Traverse 1

The traverse is oriented in the NNE-SSW direction (Figure 1) and the lateral extent is 300m. The field data pseudo-section and 2D resistivity inverted section delineate 3distinct geoelectrical layers which are the topsoil sand, sandy clay and clay (Figure 5a). There is gradation in geological formation from sand (topmost) to clay (bottom) via sandy clay formation. The topsoil is characterised by high resistivity value 508 Ωm , average thickness of about 15m.It is representative of undulating sand mixed with stones at shallow depths used during filling of the area. The next layer has an average resistivity value 160 Ωm with an average thickness of about 35m; it is made up of sandy clay formation.

Towards the end of the traverse between electrodes 18 and 27, the third layer made up of subsoil associated with low resistivity value of about 16 Ωm and average thickness of 25m. It typifies clay. The presence of a clay subsoil indicates the incompetence of this layer to support a massive engineering structure. In addition, some structural geological features are noticed in the inverted resistivity section -- synclinal of the sandy clay formation and fracture between electrodes 8and 12. These could result in structural instability and foundation failure for any development erected on this part of the traverse. However, foundation structures would require special techniques (piling) extending down to the stable soil strata or bedrock to avoid foundation failure and structural damage in the proposed area.

Traverse 2

The traverse trends in the NE–SW direction. The lateral extent is about 280m. The pseudo-section and 2D resistivity imaging section (Figure 5b) are closely similar to traverse 1. The first part of the section (0-10m) shows a zone of relatively high resistivity value of about 465Ωm It is composed of unconsolidated sand mixed with stones used for filling at a shallow depth of about 15m. Its closeness to the surface shows that it can only support light to medium foundational structures. Overlaying the clay layer is the middle part of the profile with resistivity values of 195 Ωm to 113 Ωm that correspond to sandy clay formation. Due to its good engineering property and zone of deposition, it can support low to medium engineering structures. Beneath it is a low resistivity layer (resistivity 94 Ωm, thickness 30m) common to clay. The presence of a clay subsoil and a fracture noticed between the 9^{th} and 12^{th} electrodes makes this layer unsuitable for heavy engineering structures. Alternatively, a well-designed engineering foundation with proper piling would need to be taken into consideration for a good bearing layer.

Traverse 3

It runs from E-W direction and is about 240m long. It reveals (Figure6a) 3distinct layers with the topsoil made up of undulating sand mixed with stones that spans almost the length of the traverse. This layer is characterised by relatively high resistivity values of 398 Ωm with thickness of about 18m. The observed decrease in resistivity at a depth below the topsoil could be ascribed to change in subsurface geological formations. The second layer which typifies sandy clay with average resistivity of 135 Ωm and thickness of 15m could have good engineering properties for building construction. The third layer is similar to a formation beneath the afore-mentioned traverses. It is composed of clay with an anticline shape having average resistivity of about 29 Ωm and thickness of 25m. The low resistivity value of this formation and its structural form explain why it might be incompetent to support a massive engineering structure unless special foundation designs are taken into account prior to development.

Traverse 4

It is oriented in the E-W direction. The topsoil is made up of undulating sand mixed with stones at shallow depths (Figure 6b).The resistivity values ranges from 205-309 Ωm to a depth of about 18m. A basin-like structure which depresses throughout the section is observed between electrodes 5 and 10 possibly due to a fracture in the subsurface. The subsoil in this zone is sandy clay with average resistivity of 135 Ωm from a depth of 18m to the extent of investigation. The third layer which has a resistivity of 54 Ωm is about 27m thick. It is composed of a segmented clayey formation. It might be due to a fracture observed between the 5^{th} and 12^{th}electrodes in the inverted section. As mentioned earlier, the development of this part of the site would also require special foundation designs.

Magnetic Survey

The magnetic data were interpreted using the total magnetic intensity profiles Figure 7a-d to identify possible anomalies. From visual inspection, the profiles show a number of marked anomalies with amplitudes ranging from -3800nT to 700 nT. Generally, in all the profiles the magnetic anomalies are associated with short wavelengths occasioned by near surface materials (Reynolds, 2011; Grant &West, 1965) suggestive of intercalations of sand and clay and pebbles or stones) which characterized the study area. The magnetic anomalies are located between 40-100m, 150-180m, 200-250m and 300-320m along the traverses. The high amplitude magnetic peaks noticed at 60-100m and 210-230m could be ascribed to a transition between geological formations (sand and clay due to a difference in their magnetic susceptibilities)and the presence of faults/fractures arising from the depositional nature of the study area. The low amplitude magnetic troughs identified between 40 m and 60m, 150 m and200m (Figure 7) may be due to undulating sand mixed with stones near the surface as observed in the 2-D resistivity sections.

INTEGRATION

An integrated interpretation of the acquired geophysical data shows that 3major resistivity zones were delineated from the resistivity sections. The presence of these zones is confirmed by variations observed in amplitude and size of the magnetic profiles. The depression observed in the2D sections might be due to a fault arising during the depositional period of the sediments or lithologic units at different depths in the subsurface. The variations in the amplitude of the magnetic profiles leading to peaks and troughs give indications of the presence of the alternating sequence of sand and clay formations in the study area.

A combination of electrical, magnetic and available well logs provides useful information of the subsurface geology of the area that is relevant to understanding and delineating the subsurface conditions for development of any category of engineering structure. The depth and extent of the magnetic responses are employed in constraining the depth obtained from the resistivity method, with both methods in good agreement.

Traverse 1

A combination of the 1-D geoelectrical section and the 2-D resistivity image shows the possibility that the VES 1 centre point was carried out between electrode positions 10 and 11while the start-off point for VES 2 was performed between electrodes 21 and 22. This is justified by the resistivity and depths obtained within this zone which correlate with that obtained from the geoelectric section for VES 1.It is evident that the third layer at a depth of about 2.2m with resistivity of 582.7Ωm matches that obtained in the 2-D resistivity imaging.

Traverse 2

In relating the geoelectric section and the 2-D resistivity structure, it is possible that VES 3 and VES 4 were carried out at the electrode positions 14 and between 9 and 10 respectively. This is evident from the resistivity and depths obtained for the layers beneath these locations agreeing with the results of the geoelectric sections. For example, at an approximate depth of 20m, clay with resistivity of 94Ωm, agrees with the fourth layer of VES 3.

Traverse 3

Using the 1Dgeoelectric section, the 2-D resistivity geoelectrical section, and the magnetic anomaly profiles, the depths to the competent layers estimated for the magnetic profiles match the sand layers obtained from the geoelectric section and the 2-D resistivity imaging. Furthermore, it can be inferred that the centre point for VES 7 was between electrodes 10 and 11.This is because all but the last layer are consistent with the resistivity of the layers shown in the 2-D resistivity section. For example, the 2-D resistivity section agrees with the fourth layer of VES 7at a depth of about 25m and with a resistivity of about 33Ωm. Also, the centre point for VES 8 which was conducted about 50m from VES 7 is between electrodes 16 and 17.

Traverse 4

On analysis of the geoelectric sections, the 2-D resistivity structure and the magnetic anomaly profiles, we noticed that the various depths to the anomaly estimated from the magnetic profiles presented in Table 2 were consistent especially with the sand layers obtained for both the geoelectric section and the 2-D resistivity structure. It then follows that VES 5 could have been performed between electrode points 8 and 9 andVES 6 between electrodes positions 14 and 15.

CONCLUSION AND RECOMMENDATIONS

This paper has discussed the use of integrated geophysical methods in the delineation and identification of geological structures which can guide the nature and type of foundations for the development of the study area. The qualitative interpretation of the 2D resistivity sections showed that the area is characterised mainly by sand, sandy clay and clay materials with the sand and sandy clay formations at shallow depths. Also, structural features mainly of fractures and synclinal were observed. This illustrates one of the advantages of 2-D resistivity imaging over the 1D geoelectrical section since none of the VES points could identify the fracture region on the 2-D imaging sections.

The undulating sand at shallow depths, the clayey materials as well as the fracture identified shows the incompetence of the subsurface in supporting massive or heavy engineering structures that could be developed in the study area. Worthy of note is the clayey materials present in the third layer of all the resistivity imaging sections. Its nature makes it highly incompetent for massive engineering structures except for proper foundation designs. Erecting structures could lead to differential settlement as clay has little or no integrity to support massive engineering structures because of its low shear strength. Since the clay layer extends to a substantial depth below the surface, a situation such as fluctuation in the water level could cause the clay to become oversaturated, leading to uplift of the ground's surface.

The presence of sandy materials near the surface suggests the suitability of this stratum (soil) for shallow foundations. The sandy clay formation observed in the second layer of the inverted 2-D resistivity sections reaching to an average depth of 20m offers good engineering properties such that the formation can support low to medium engineering structures. Deep foundations may not be feasible or convenient in the survey area owing to the presence of clayey materials in all the sections. It is strongly recommended that reinforcement and concrete packing should be done in parts of the study area where there is need for massive engineering structures. This is imperative in preventing tilting of buildings and cracks surfacing on the walls and floor as that would besigns of differential settlement and subsequent collapse of such structures in the future. Also, foundation footing should not be placed on clayey material. Foundations of large engineering structures around the study area should be sited on competent material via pilings. This study has demonstrated the usefulness of both the electrical resistivity and magnetic methods in characterizing subsurface soil for citing engineering structures.

TABLE1
Summary of VES Model Resistivity Values and Their Corresponding Thicknesses

Ves Point	Layer 1 (Topsoil)		Layer 2 (Clayey/Sand)		Layer 3 (Sand)		Layer 4 (Sand/Sandy Clay)		Layer 5 (Clayey and/Sand)		Curve Type
	Resistivity (Ωm)	Thickness (m)	Resistivity (Ωm)	Thickness (m)	Resistivity (Ωm)	Thickness (m)	Resistivity (Ωm)	Thickness (m)	Resistivity (Ωm)	Thickness (m)	
VES 1	5167.0	0.9	2171.5	1.4	582.7	7.3	132.7	49.1	1119.2	-	QQH
VES 2	2809.1	0.9	912.0	3.2	205.0	7.3	176.6	16.6	1039.1	-	QQH
VES 3	2314.5	0.9	1596.1	1.2	296.9	14.0	99.4	27.4	6339.2	-	QQH
VES 4	334.7	0.5	2577.6	0.3	200.6	9.3	40.5	43.5	4702.2	-	KQH
VES 5	1897.0	1.1	1232.5	1.9	375.8	8.0	36.8	22.7	807.5	-	QQH
VES 6	1581.2	1.1	844.0	2.5	423.5	6.4	56.8	15.7	12331.5	-	QQH
VES 7	1054.6	0.8	426.9	4.3	350.7	4.9	40.5	11.6	3633.0	-	QQH
VES 8	1003.0	1.0	369.2	1.9	459.1	5.4	50.3	12.6	4001.0	-	QQH

TABLE 2
Depth, Extent and Nature of Magnetic Anomaly

S/N	Traverse (m)	Depth (m)	Anomaly Range(m)	Nature of Anomaly
1	3(10)	3.15	100 - 120	Negative
		3.15	250 -270	Positive
		4.5	260 - 280	Negative
		4.5	310 - 330	Positive
		6.25	30 - 50, 150 - 200	Negative
		12.5	60 - 90, 220 - 250	Negative
2	3(20)	4.5	250 - 270	Positive
		5.1	130 - 180	Positive
		5.2	100 - 120	Negative
		9.5	220 - 240	Negative
		15.5	30 - 50, 280 - 300	Negative
3	4(10)	5.3	50-70	Positive
		21.5	100 - 120	Negative
		25.2	130 - 150	Negative
		9.7	180 - 200	Positive
		19.8	210 - 230	Negative
		5.1	240 - 250	Positive
4	4(20)	13	60 - 100	Negative
		9.2	110 - 140	Negative
		9.7	150 - 170	Negative
		17	200 - 240	Negative

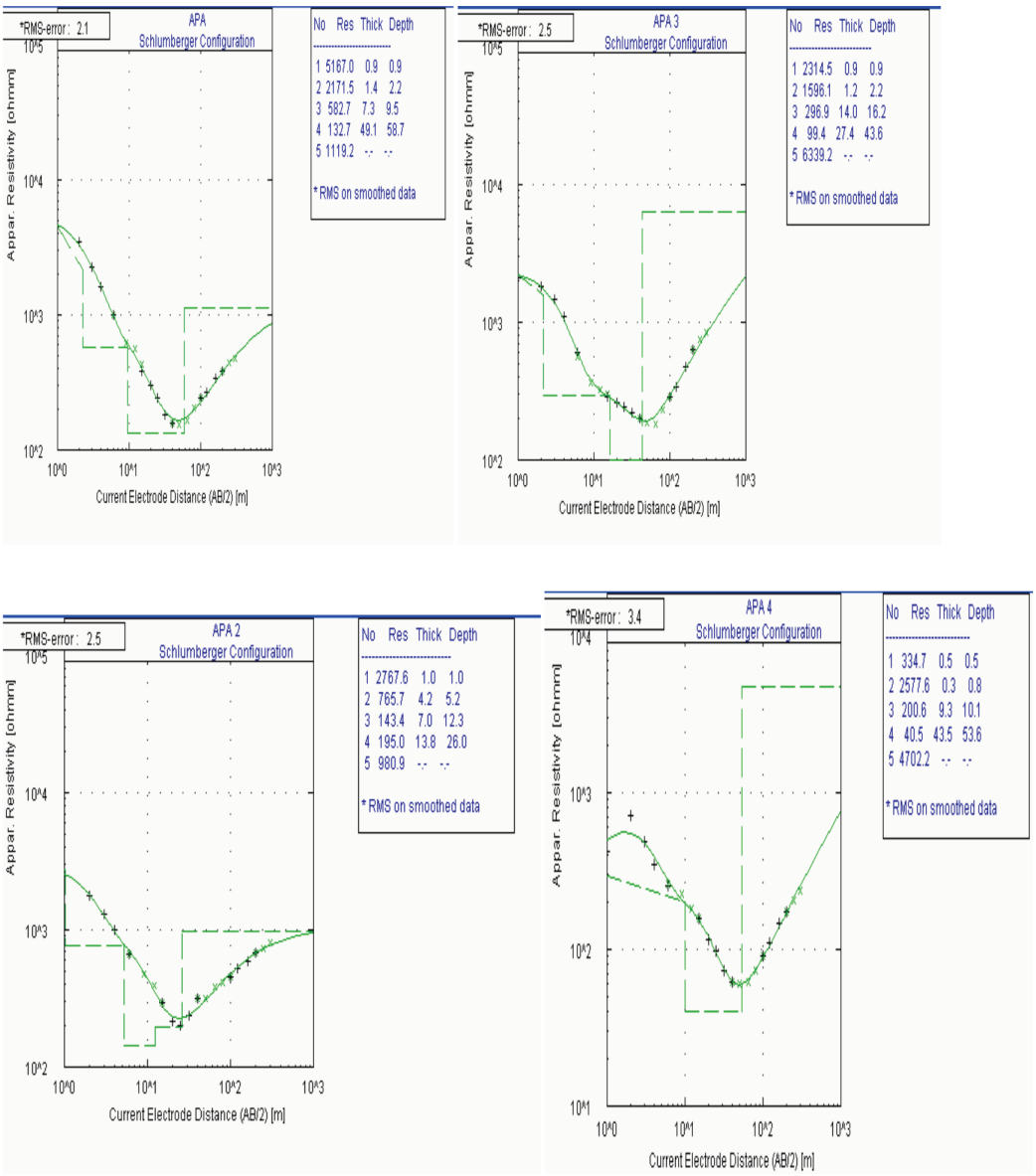


Fig.2a: Observed, Computed VES Points (1-4) and Layered Inversion Model

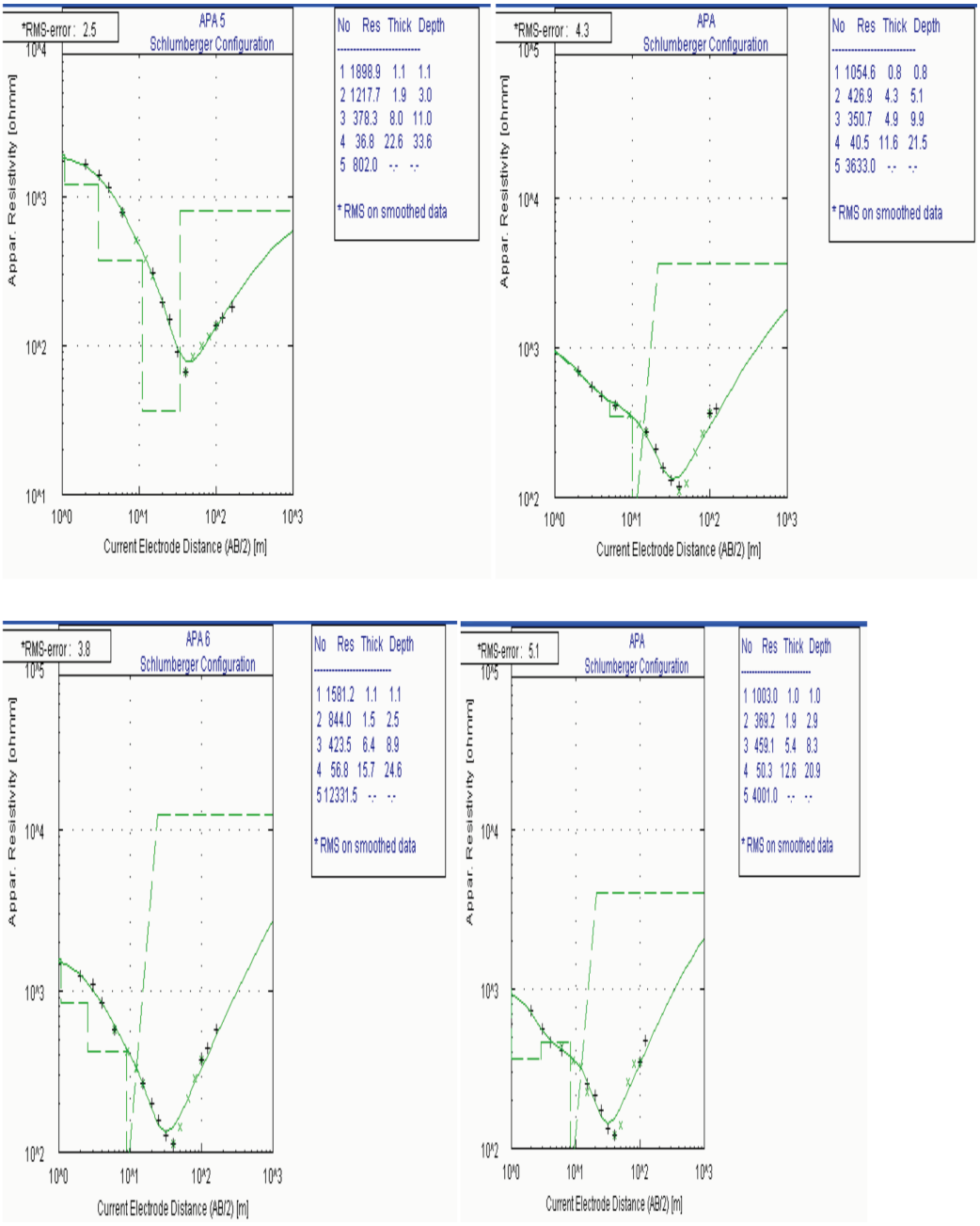


Fig.2b: Observed,Computed VES Points(5-8) and Layered Inversion Model

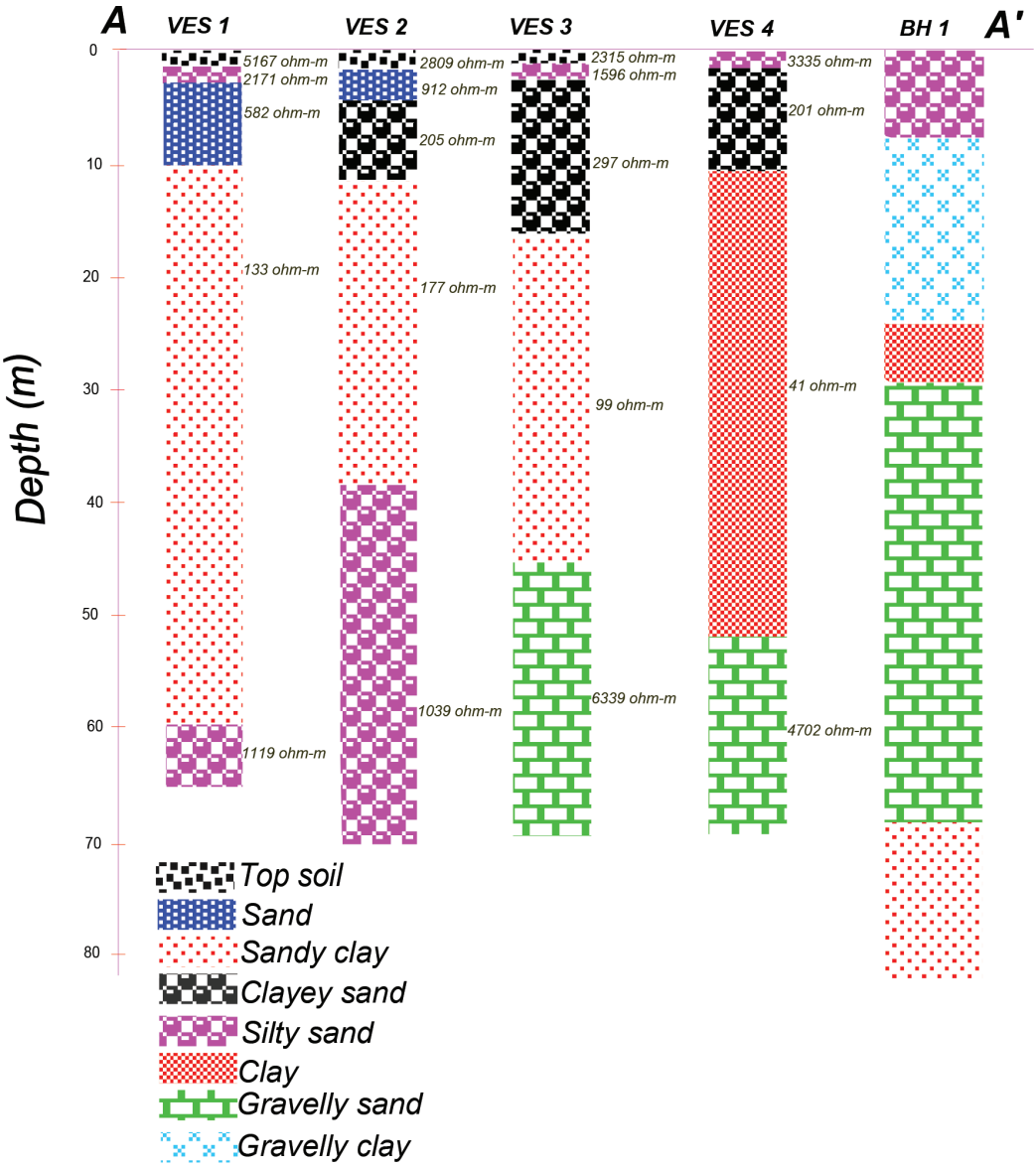


Fig.3: Geoelectric Section Along AA'

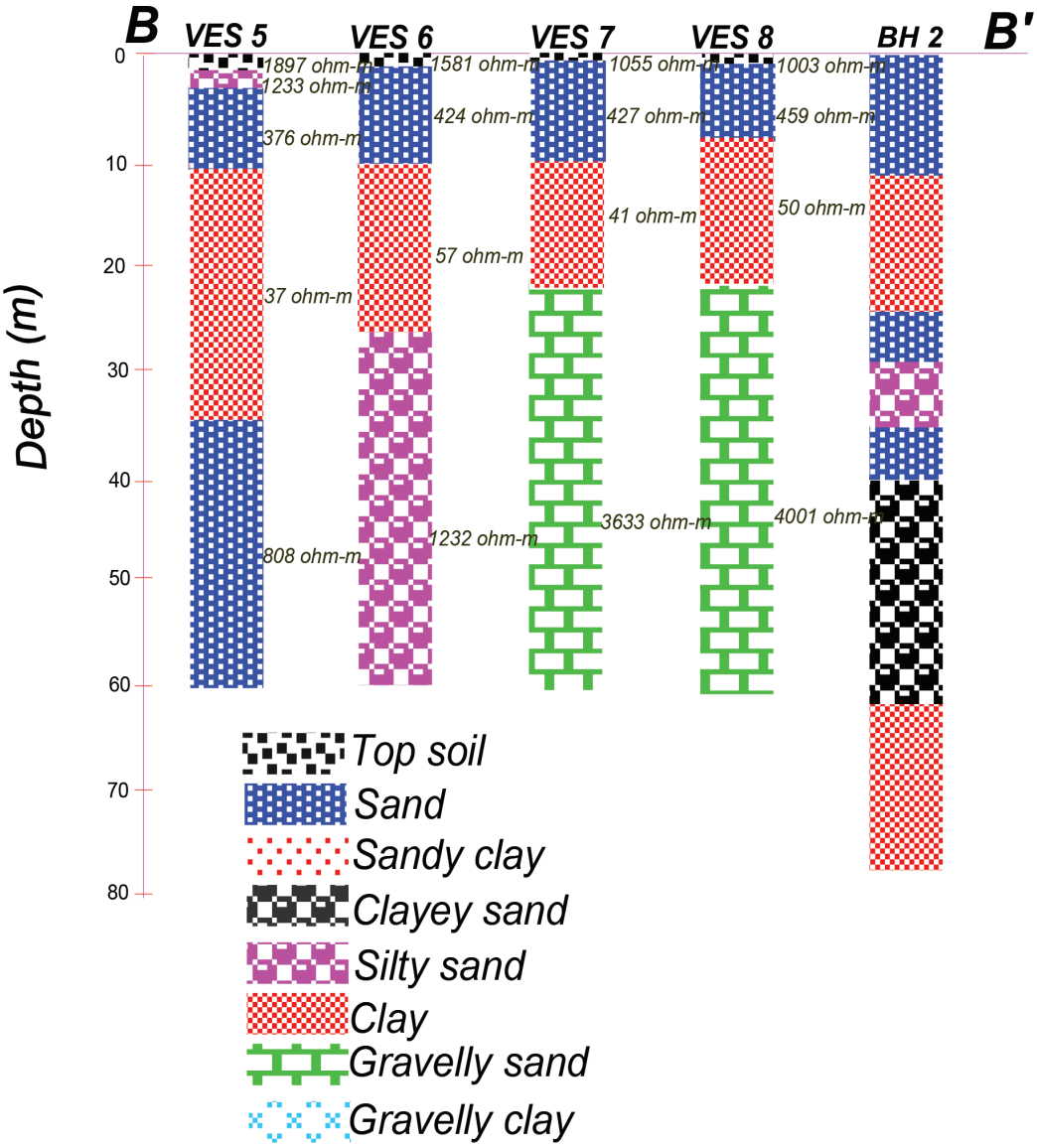
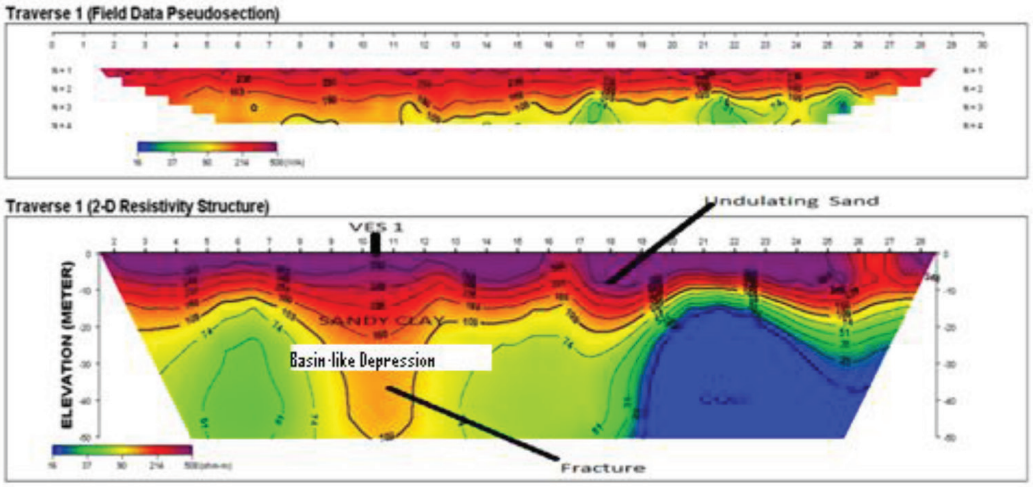


Fig.4: Geoelectric Section Along BB'

(a)



(b)

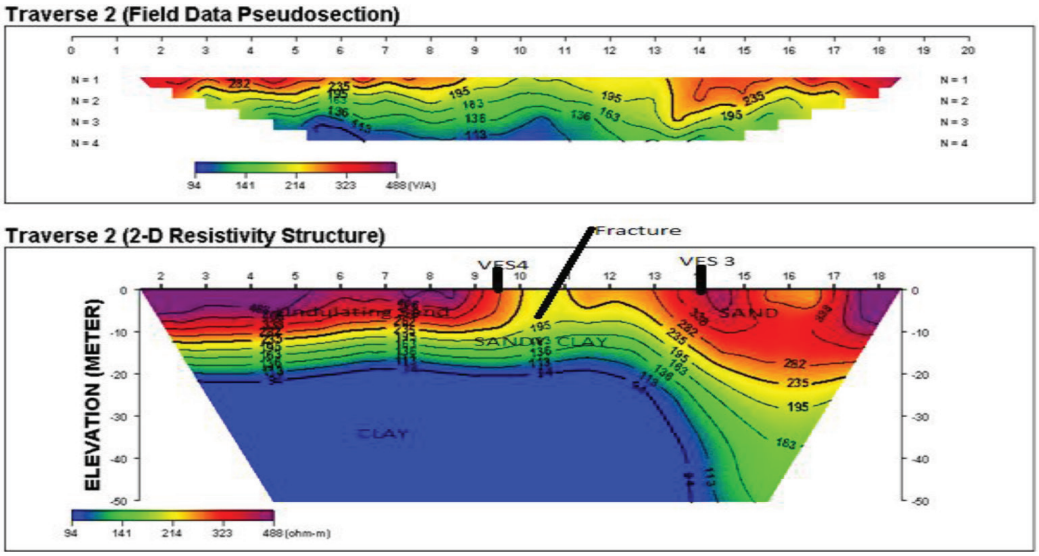
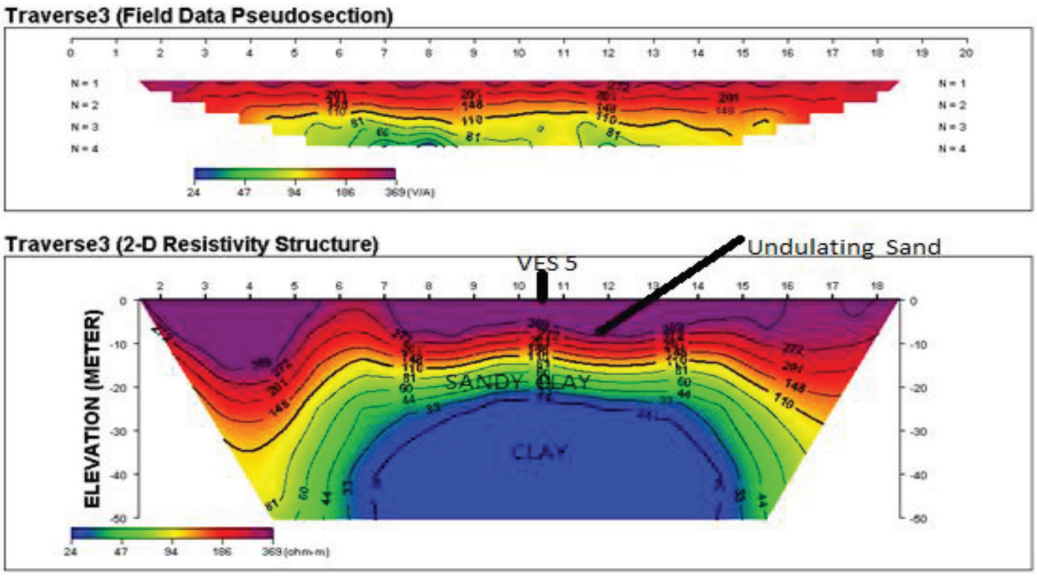


Fig.5: (a)Measured Apparent Resistivity Pseudosection(top) and Inverse Model Resistivity Section (bottom)Along Traverse 1 (b) Measured Apparent Resistivity Pseudosection (top) and Inverse Model Rresistivity (bottom) Along traverse 2.

(a)



(b)

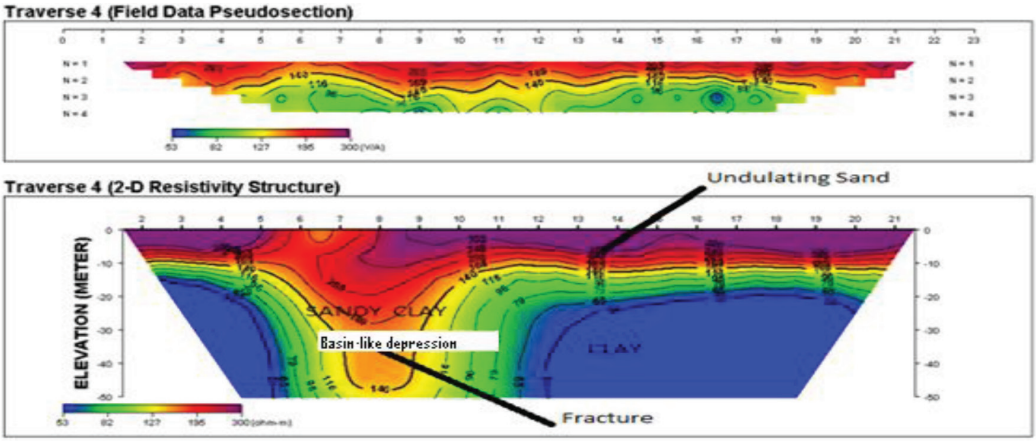


Fig.6: (a) Measured Apparent Resistivity Pseudosection (top) and Inverse Model Resistivity Section (bottom) Along Traverse 3 (b) Measured Apparent Resistivity Pseudosection (top) and Inverse Model Resistivity (bottom) Along Traverse 4.

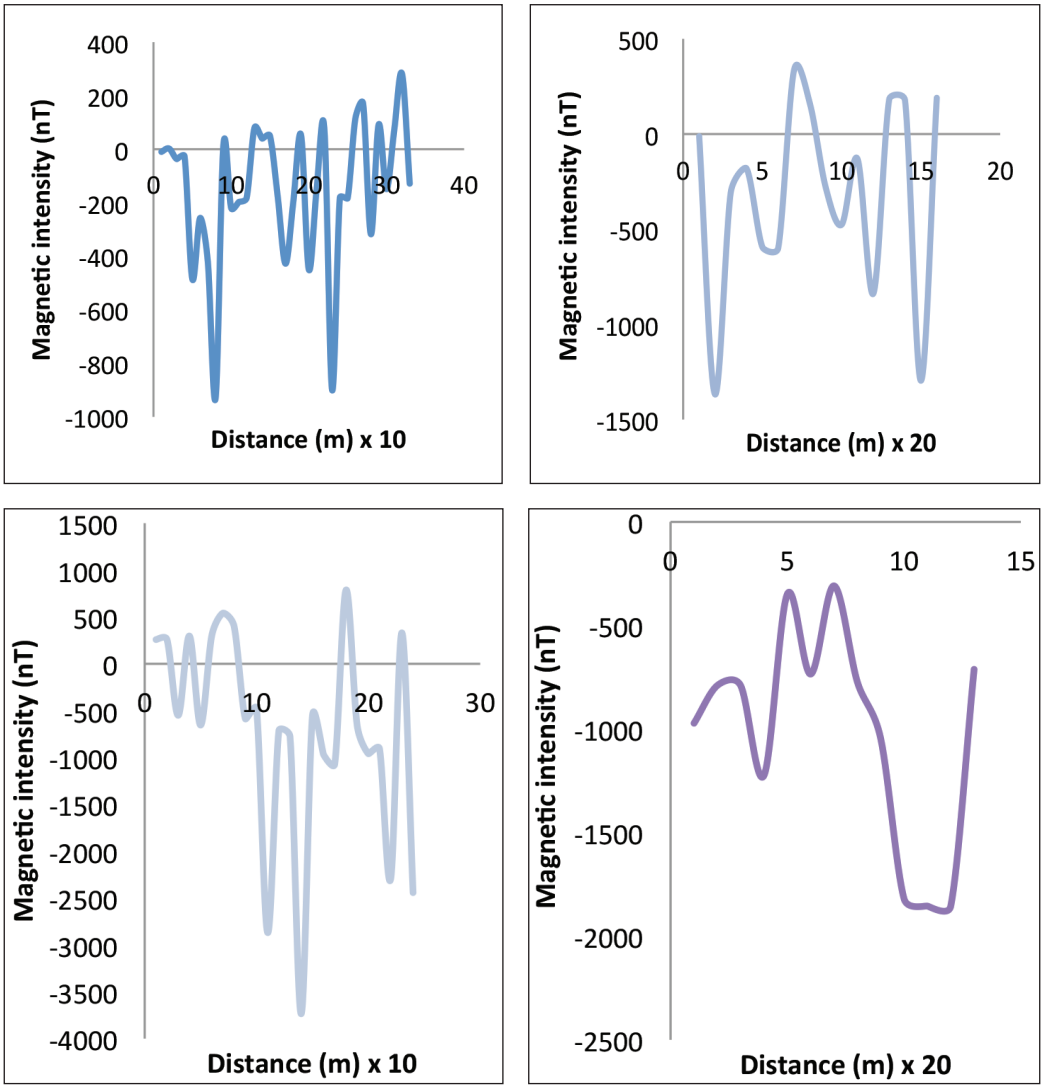


Fig.7: Magnetic Anomaly ProfilesAlong (a) Traverse 3 at 10m Spacing (b)Traverse 3 at 20m Spacing (c)Traverse 4 at 10m Spacing (d) Traverse 4 at 20m Spacing

REFERENCES

Adegoke, O. S., Jeje, L. K., Durotoye, B., Adeleye, D. R., & Ebukanson, E. E. (1980). The geomorphology and aspects of sedimentology of coastal region of Western Nigeria. *J. Min. Geol.* 17-19.

Akran, V. S. (2001). Lagos Historical Review. *European Journal of Scientific Research, 1*, 33-44.

Ayolabi, E. A., Folorunso, A. F., Eleyinmi, A. F., & Anyah, E. O. (2009). Applications of 1D and 2D Electrical Resistivity methods to Map Aquifers in a complex Geologic Terrain of Foursquare Camp, Ajebo, Southwestern, Nigeria. *The pacific Journal of Sci. and Tech., 10*(2). 657 - 666.

Breiner, S. (1973). *Application manual for portable magnetometers.* Geometrics: California.

Cardarelli, E., Marrone, C., & Orlando, L. (2003). Evaluation of Tunnel Stability Using Integrated Geophysical Method. *Journal of Applied Geophysics, 52*(2-3), 93-102.

Coker, S. J., Ejedawe, J. E, & Oshiorienua. (1983). Hydrocabon source potentials of Cretaceous rocks of Okitipupauplift,Nigeria. *Journal of Mining and Geology, 22*, 163- 169.

Dabas, M., Tabbagh, A., & Tabbag, J. (1994). 3D Inversion in subsurface electrical surveying - 1.*Theory. Geophys. J. Int*, *119*, 975 - 990.

Dekolo, S. O., & Oduwaye, A. O. (2011). Managing the Lagos megacity and its geospatial imperative. *International Archives of the Photogrammetry, Remote Sensing and Spatial information Science, XXXVIII* (4), C21

Egwuonwu, G. N., & Osazuwa, I. B. (2011). Geophysical and Geotechnical Invetsigation of the origin of Structural Instabilities shown on some Low Rise Buildings in Zaira, North-Western Nigeria. *The pacific Journal of Science and Technology, 12*(2), 534-546

Elueze, A. A., & Nton, M. E. (2004).Organic geochemical appraisal of limestone and shales in part of eastern Dahomey Basin, Southwestern Nigeria. *Journal of Mining and Geology, 40*(1), 29 - 40.

Fatoba, J. O., & Olorunfemi, M. O. (2004). Subsuface sequence delineation and saline water Mapping of Lagos State,Southwetern, Nigeria. *Global journal of geological sciences,* 111-123.

Grant, F. S., & West, G. F. (1965). *Interpretation theory in applied geophysics*. New York

(McGraw). p.584.

Ghosh , D. (1971a). Application of linear filter interpretation of geo-electrical resistivity theory to the direct sounding measurements. *Geophysical Prospecting, 19*, 192-217.

Ghosh, D. (1971b). Inverse filter coefficients for the computation of apparent resistivity standard curves for horizontally layered earth. *Geophysical Prospecting, 19,* 769-217.

Jones, H. A., & Hockey, R. D. (1964). The geology of part of south western Nigeria. *Geol. Survey. Nigeria*, *Bull*, *31*, 1-101.

Kayode, J. S., & Adelusi, A. O. (2010). Ground magnetic data interpretation of Ijebu-Jesa Area, Southwestern Nigeria, Using total component magnetic field. *Research Journal of Applied Sciences, Engineering and Technology*, *2*(8), 703-708.

Loke, M. H. (2000). Electrical Imaging surveys for Environmental and Engineering studies: A practical Guide to 2-D and 3-D surveys. *Heritage Geophysics.*

Lucet, N., & Mavko, G. (1991). Images of rock properties estimated from a crosswell seismic velocity tomogram. *Technical Program and Abstracts of Papers*. *Eur. Assoc. Of Explor. Geophys*, *53*, 520-521

Mabogunje, A. L.(1971). *Urbanization of Nigeria*. New York, African Publishing Corportation. pp.48-60.

MacDonald, A. M., Davies, J., & Peart, R. J. (2001). Geophysical methods for locating Groundwater in low permeability sedimentary rocks: examples from southeast Nigeria. *Journal of African Earth Sciences, 32*(1), 115- 131.

Marquardt, D. O. (1963). An algorithm for least-squares estimation of nonlinearparameters: *Journal of the Society for Industrial and Applied Mathematics, 11*, 431-441.

Mondal, N. C., Rao, V. A., Singh, V. S., & Sarwade, D. V. (2008). Delineation of Concealed Lineaments Using Electrical Resistivity Imaging in Granitic Terrain. *Current Science*, *94*(8), 1023-1030

Omatsola, M. E., & Adegoke, O. S. (1981). Tectonic evolution and cretaceous stratigraphy of Dahomey Basin. *Journ. Min. Geol., 18*, 130-137.

Oyedele, K. F., Oladele, S., & Okoh, C. (2011). Geophysical and Geotechnical Assessment of Geo-engineeing properties of soil in Magodo Brook Estate, Lagos, Nigeria. *International Jounal of Geophysics, 2*(1), 31-37.

Peterside, C. S. (2007). Bond as alternative Funding Instrument for Housing and infrastructure development in Nigeria (2007). *NESG Economic Indicators, 12*(4), 24-26.

Reyment, R. A. (1965). *Aspects of the Geology of Nigeria: The Stratigraphy of the cretaceous and Cenozoic Deposits.* Ibadan University Press. pp. 1-133.

Reynolds, J. M.(2011). *An introduction to Applied and Environmental Geophysics. 2nd Edition.* ISBN 978-0-471-48535-3.

Sharma, P. V. (1997). *Geophysical Methods in Geology.* PTR Prentice-Hall, Inc., A Simon and Schuster Company, Engle Cliffs, New Jersey, United States.

Telford, W. M., Geldart, L. P., & Sheriff, R. E. (1990). Applied *Geophysics (2nd Edition)*. Cambridge University Press, UK. p.770.

Tsourlos, P. I., Szymanski, J. E., & Tsokas, G. N.(1998). A smoothness constrained algorithm for the fast 2- D inversion of DC resistivity and Induced Polarization data. *Journal of the Balkan Geophysical Society, 1*(1), 3 - 13.

Zohdy, A. A. R., Eaton, G. P., & Mabey, D. R. (1980). *Application of Surface Geophysics to Ground-Water Investigations. Techniques of Water Resources Investigations of the United States Geological Survey.* United States Government Printing Office, Washington.

Pertanika J. Sci. & Technol. 22 (2): 529 - 540 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Antimicrobial Activities of Three Different Seed Extracts of *Lansium* Varieties

H. Alimon[1], A. Abdullah Sani[1], S. S. Syed Abdul Azziz[2]*, N. Daud[3], N. Mohd Arriffin[2] and Y. Mhd Bakri[2]

[1]*Department of Biology, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia*

[2]*Department of Chemistry, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia*

ABSTRACT

Lansium domesticum Corr. is a fruit tree of the Meliaceae family, which is commonly found in South-East Asia with a wide range of varieties. This study investigated three varieties of *L. domesticum*; *Duku*, *Langsat* and *Dokong* for the phytochemical screening and antimicrobial activity. Seeds from the matured fruits were extracted using hexane, methanol and water. The crude extracts were screened for antimicrobial activities toward three bacteria, namely, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus*. The findings showed that *Langsat* seed extracts contained more groups of compounds compared with the other two varieties, and its methanol extract demonstrated the highest inhibition zones against the three bacteria. The crude methanol extract of *Duku* seeds showed inhibition zones only towards *Bacillus subtilis* at a high concentration (1.0 mgL^{-1}), whilst the seed extracts of *Dokong* showed no inhibition zones towards any of the tested bacteria.

Keywords: *Lansium domesticum*, antimicrobial activities, inhibition zones

Article history:
Received: 2 July 2012
Accepted: 26 April 2013

Email addresses:
H. Alimon (hasimah@fsmt.upsi.edu.my),
S. S. Syed Abdul Azziz (saripah@fsmt.upsi.edu.my),
Y. Mhd Bakri (yuhanis@fsmt.upsi.edu.my)
*Corresponding Author

INTRODUCTION

Lansium domesticum Corr. or lanzones, a species of Meliaceae family, are evergreen tropical trees that grow upright, tall, slim, and can reach up to 30 metres in height. *L. domesticum* is a species that vary widely, and leads to a classification of several varieties as different species by some taxonomists. Difficulties in classifying the *L. domesticum* varieties have led to the usage of high-

ISSN: 0128-7702

technology methods to determine the differences between them. Beng *et al*. (2000) classified the *L. domesticum* varieties of Peninsular Malaysia into four main groups (*Dokong*, *Duku*, *Langsat* and *Duku-langsat*) based on random amplified polymorphic DNA (RAPD) analysis, while the researchers from our neighbouring countries, Thailand and Indonesia, classified these varieties into three main groups (*Dokong*, *Duku* and *Langsat*), which are clearly differentiated through several techniques mainly RAPD, RAPD-based PCR and flow-cytometry (Te-chato *et al*., 2005; Sri Yulita, 2011). Our study focused on three varieties of *L. domesticum*, namely, *Duku*, *Langsat* and *Dokong*, due to their availability during the period of study.

The fruit from all the three varieties emerge from the flower inflorescences that grow and hang from large branches or trunk. The fruit can be elliptical, oval, or round, measuring 2 to 7 cm (0.79 to 2.8 in) by 1.5 to 5 cm (0.59 to 2.0 in) in size. The thickness of the fruit's skin varies. The skin of *Langsat* is thin and has latex, but there is little latex in the thicker skin of *Duku* and very little latex or none in the thick skin of *Dokong* (Yunus, 1997). The fruit usually contain 1 to 3 seeds, flat, and bitter tasting, whereas the seeds are covered with a thick, clear-white aril that has sour flavour and sweet taste (Sapii *et al*., 2000). The sweet juicy flesh contains monosaccharides such as sucrose, fructose and glucose. Among the three varieties, *Dokong* is sweeter, with very little seeds or no seed and relatively small size (Te-chato *et al*., 2005).

In addition to food, some parts of the tree are traditionally used as a remedy for treating parasites in the intestine and diarrhoea. The dried fruit with burnt skin are known to be used as mosquito repellent and treat malarias and scorpion stings. Meanwhile, the seed powder of *Langsat* is used to reduce fever (Morton, 1987). These traditional practices have initiated numerous research in identifying the chemical constituents of the *Lansium* seeds. A study on the *Lansium* plant has reported the presence of high amount of alkaloid content in barks (Lense, 2011) but only trace amount in their seeds (Morton, 1987). Saewann *et al.* (2006) successfully isolated five tetranortriterpenoids and 11 known triterpenoids.

A recent report shows that the methanol extract of the dried seeds and bark of *Lansium domesticum* cv. kokossan yielded two tetranortriterpenoids, kokosanolide A and C, and an onoceranoid-type triterpenoid, kokosanolide B, along with two onoceranoid-type triterpenoids; 8,14-*seco*gammacera-7,14-diene-3,21-dione, and a mixture of 8,14-*seco*gammacera-7,14(27)-diene-3,21-dione and 8,14-*seco*gammacera-7,14-diene-3,21-dione (1.5:0.5) (Mayanti *et al*., 2011). Several compounds isolated from *Lansium* seeds have been identified to exhibit anti-malarial activity against *Plasmodium falciparum* (Saewann *et al*., 2006), anti-feedant activity against the fourth instar larvae of *Epilachna vigintioctopunctata* (Mayanti *et al*., 2011), and anti-bacterial activity against Gram-positive bacterium (Dong *et al*., 2011).

Generally, *Lansium* has been reported in very few studies focusing on its varieties. In this study, the potential of the seed crude extracts of the three closely related varieties of *Lansium* (*Duku*, *Langsat*, and *Dokong*) in inhibiting the growth of the three selected bacteria, *Pseudomonas aeruginosa*, *Staphyllococcus aureus* and *Bacillus subtilis,* was investigated. These bacteria are common in Malaysia and in most places in the world. For example, *Staphylococcus aureus* is a facultative anaerobic Gram-positive bacterium that is usually found

to coexist with human in normal skin flora and in nasal passages (Foster, 2004). However, it can cause a wide range of infections from minor ones such as pimples and impetigo to severe infections such as the methicillin-resistant *S. aureus* (MRSA) that are becoming a serious problem in many places in the world as these virus clones are resistant to many antibiotics (Foster, 2004; Grundmann *et al*., 2012). *S. aureus* produces enterotoxin, which is responsible for staphylococcal food poisoning that cause rapid onset, nausea, violent vomiting, and may cause chronic diarrhoea (Argudin *et al*., 2010). In fact, scientists are still working on searching for the right vaccine or special drug to prevent or to cure this problem. *Pseudomonas aeroginosa*, a Gram-negative bacterium, is capable of living in both aerobic and anaerobic conditions. It is an important pathogen of plants and animals, which can infect damaged tissues or those with reduced immunity, and may become fatal if colonized in vital organs such as kidney, liver and urinary tracts (Stojek *et al*., 2008). As this bacterium has high prevalence of antibiotic resistant strains, a new therapeutic agent is required to overcome this problem. Another bacterium chosen in this study, *Bacillus subtilis,* is a well-known Gram-positive model bacterium for laboratory studies. Many *B. subtilis* strains have been genetically manipulated, selected, improved, and well studied (Buxton & Ward, 1980; Tam *et al*., 2006). On this basis, the researchers hoped that the results from the current work would complement excellent findings previously reported for the *Lansium* species.

MATERIALS AND METHODS

Preparation of the Seed Extracts

The matured seeds of *Duku*, *Langsat*, and *Dokong* were collected from several areas in Kelantan, Malaysia, from July to October 2010. The plant samples were identified by the plant systematic experts. The seeds were separated from the pulps and dried in an oven at 40-45°C for 5 days. From 1 kg of fruits, approximately 30 g, 25 g, and 20 g of the dried seeds of *Langsat*, *Duku* and *Dokong* were produced, respectively. The dried seeds were ground to coarse powder before extracted with hexane, methanol and water in cold condition. The various extracts were filtered and evaporated using rotavap to give crude extracts. The crude extracts produced from hexane, methanol and water of *Langsat's* dried seeds were 1.44 g, 2.75 g, and 1.00 g, respectively, *Duku* (1.20 g, 1.69 g and 1.00 g) and *Dokong* (1.12 g, 11.27 g and 1.00 g). All the crude extracts were stored in -4 ^{0}C until to be used for antimicrobial activity.

Preparation of the Stock Solutions of Plant Extracts

Stock solutions of the plant extracts were prepared by dissolving 200 mg of each plant extract in 1 mL of sterilized distilled water. The mixture was vortexed to ensure that the extracts were homogeneous. The working stock solutions were protected from light by covering the bottle with aluminium foil. The seed extracts of stock solutions were used in disc diffusion test for the antimicrobial activity and for the phytochemical screening.

Phytochemical Screening of Seed Extracts

Seed extracts were subjected to phytochemical screening to identify the chemical constituents such as alkaloids, flavonoids, saponins and tannins using standard procedures described below.

Alkaloids

The test for alkaloids was carried out by basifying 20 g of crushed seeds in 10 % of ammonia solution and soaking with dichloromethane, heated and filtered. The dichloromethane extract was re-extracted with 5 % HCl and the aqueous layer was tested with the Mayer's reagent (Bruneton, 1999). A positive test for alkaloids was indicated by the production of a turbid solution or a yellowish creamy precipitate colour.

Flavonoids

The presence of flavonoids was determined by adding a few drops of sodium hydroxide into the extracts. An intense colour was produced and it later became colourless with an addition of a few drops of dilute acid (Kumar *et al*., 2009).

Saponins

The extract was subjected to Froth test to identify saponin (Onwukaeme *et al*., 2007). A small quantity of the seed extract was boiled in 20 ml of distilled water in a water bath and filtered. The mixture was then filtered and 10 ml of the filtrate was added into 5 ml of distilled water in a test tube. The test tube was stoppered and shaken vigorously for a stable persistent froth. The frothing was mixed with 3 drops of olive oil and shaken vigorously, and the mixture was then observed in the formation of emulsion.

Tannins

The tests for tannins were carried out by subjecting the seed extracts in 1 ml of 10% of potassium hydroxide and the formation of dirt precipitation showed the existence of tannin (Kumar *et al*., 2009).

Culture Media

The bacterial cultures used in this study were *Staphylococcus aureus* (ATCC 1026), *Bacillus subtilis (B. spizizenii*) (ATCC 6633) and *Pseudomonas aeruginosa* (ATCC 10145). The bacteria were purchased from Choice Care Sdn. Bhd. The test organisms were purified and maintained on slant agar kept at 4°C until further use.

Preparation of the Media

Nutrient agar was used as the media for bacterial enumeration. Nutrient broth was also used for the generation of exponential culture of each organism.

Nutrient Agar

15 g nutrient agar was suspended with 1 litre of distilled water in a Duran bottle. The media was dissolved by fast cooking in the microwave and autoclaved for about 15 min at 121°C. The sterilized melted agar was poured into sterile Petri dishes in the laminar flow for the agar plate preparation.

Nutrient Broth

13 g nutrient broth was suspended with 1 litre of distilled water. The media dissolved by fast cooking in the microwave. 5 ml of melted broth were pipetted into test tube and sterilized by autoclaving for about 15 min at 121°C.

Screening for Antibacterial Activity

The antibacterial activity of the extracts on the microorganisms was determined by using the disc diffusion methods.

Disk Diffusion Method

The disk diffusion method was used to measure the rate of inhibition in growth of bacteria by different concentrations of the plant extract on paper disc. The volumes of the extract used were 20 µl, 30 µl, 40 µl, and 50 µl from the 200 mg/ml of each stock solution. The discs were allowed to dry in the laminar flow before they were placed on top of the agar. The 24-hour broth culture of each test bacterium species was aseptically introduced and spread on the surface of sterile nutrient agar using sterile cotton swab. The sterile paper discs (6 mm) impregnated with extract were placed on the cultured plates and sterile forceps were used to gently press down each disc to ensure complete contact with agar surface. A disc with distilled water alone served as negative control. The plates were incubated at 37^0C for 48 hours. All the experiments were performed in four replicates and each experiment was reproduced a minimum of three times and all these procedures were carried out aseptically.

Determination of Antibacterial Properties

Reading of the inhibition zones was done at 48 hour-intervals. The antibacterial activity was interpreted from the size of the diameter of zone inhibition measured to the nearest mm as it was observed from the clear zone surrounding the disc. The zone of inhibition is measured from the edge of the disc to the edge of the growth. It is measured on the undersurface of the plate without opening the lid.

RESULTS AND DISCUSSION

Firstly, the mature seeds of *Duku*, *Langsat* and *Dokong* were screened to determine their phytochemical components. In some previous studies, phytochemical compounds such as tannins, saponins, flavonoids, steroids, and glucose-lowering were found in many plant seeds (Anago *et al*., 2011). The phytochemical screening in this study revealed the presence of

tannins and flavonoids in the seeds from all the *Lansium* varieties (see Table 1). Alkaloid, a substance that is always related to bitterness taste, was only found in the *Langsat* seed extract, but not in the other two varieties. Several findings have shown a high amount of alkaloid in the bark extract (Lense, 2011) and fruit peels (Solidum, 2012) of *Lansium domesticum*, which exhibit its medicinal potential. *Langsat* is also the only variety studied that contains saponin, another bitter taste secondary metabolite. The presence of the compounds screened in the *Langsat* seed extracts shows its potential in anti-microbial activity as these compounds have been identified to exhibit this particular effect (Cushnie & Lamb, 2005).

TABLE 1
Phytochemical screening of the seed extracts from *Lansium domesticum* varieties

Lansium domesticum variety	Constituents			
	Alkaloid	Saponin	Tannin	Flavonoid
Langsat	+	+	++	+
Duku	-	-	+	+
Dokong	-	-	+	+

++ = present in high amount; + = present; - = absent

In order to determine the potential of the *Duku*, *Langsat* and *Dokong* seed extracts as antimicrobial agents, seed extraction was done using hexane, methanol and water. All the crude extracts were tested against *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Bacillus subtilis*, and observed after 48 hours. The results are shown in Tables 2, 3 and 4, respectively.

Inhibition zones could be clearly seen in the hexane and methanol extracts of *Langsat*, but not in the water extract (Table 2). The hexane extract of *Langsat* seeds inhibited the growth of two of the tested bacteria (*S. aureus* and *B. subtilis*), while the methanol extract of Langsat was shown to inhibit all the tested bacteria. The minimal inhibition concentration (MIC) of the hexane and methanol extracts towards *S. aureus* and *B. subtilis* was at 0.25 g/ml and the MIC for methanol extract towards *P. aeroginosa* was at 0.50 g/ml. The increase in the inhibition zones was observed with the increasing concentration of the hexane and methanol seed extracts. The biggest inhibition zones occurred in the hexane extract against *S. aureus* and in the methanol extract against *B. subtilis*. In both solvents, increase in the inhibition zones was observed when the concentration of the extracts increased with the highest inhibition zones shown against *S. aureus* for the hexane extract and *B. subtilis* for the methanol extract.

TABLE 2
The inhibition zones of the *Langsat* seed extracts after 48 hours of incubation

Solvent	Bacteria	Concentration of plant extracts (g/ml)				
		0.0625	0.125	0.25	0.5	1.0
		Inhibition zones (cm)				
Hexane	PA	-	-	-	-	-
	SA	-	-	0.67 ± 0.12	0.77 ± 0.09	1.07 ± 0.09
	BS	-	-	0.27 ± 0.15	0.23 ± 0.12	0.70 ± 0.06
Methanol	PA	-	-	-	0.33 ± 0.09	0.73 ± 0.09
	SA	-	-	0.23 ± 0.15	0.80 ± 0.10	0.97 ± 0.09
	BS	-	-	0.83 ± 0.07	0.90 ± 0.06	1.13 ± 0.12
Water	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-

PA = *Pseudomonas aeruginosa* SA = *Staphylococcuss aureus* BS = *Bacillus subtilis*

An inhibition zone was observed in the methanol seed extract of *Duku* towards *B. subtilis* at the highest concentration (1.0 g/ml). Nonetheless, no inhibition against other bacteria was observed from either the methanol extracts or other extracts (Table 3).

TABLE 3
The inhibition zones of the *Duku* seed extracts after 48 hours of incubation

Solvent	Bacteria	Concentration of plant extracts (g/ml)				
		0.0625	0.125	0.25	0.5	1.0
		Inhibition zones (cm)				
Hexane	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-
Methanol	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	0.67 ± 0.03
Water	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-

PA = *Pseudomonas aeruginosa* SA = *Staphylococcuss aureus* BS = *Bacillus subtilis*

On the other hand, none of the *Dokong* seed extracts showed any inhibition zones against any bacterium (Table 4). The finding revealed that using any of the solvent in this study, the *Dokong* seed extract is an ineffective inhibitor of all the three bacteria.

TABLE 4
The inhibition zones of the *Dokong* seed extracts after 48 hours of incubation

Solvent	Bacteria	Concentration of plant extracts (g.mL-1)				
		0.0625	0.125	0.25	0.5	1.0
		Inhibition zones (cm)				
Hexane	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-
Methanol	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-
Water	PA	-	-	-	-	-
	SA	-	-	-	-	-
	BS	-	-	-	-	-

PA = *Pseudomonas aeruginosa* SA = *Staphylococcuss aureus* BS = *Bacillus subtilis*

This study employed the disc diffusion method to determine the antimicrobial activity of the *Lansium* seed extracts against the three selected bacteria. In some previous reports of antimicrobial activity studies, the disc diffusion methods have been shown to yield similar results with an excellent agreement using broth dilution or microdilution methods and other tests (Serrano *et al.*, 2004; Milici *et al*., 2007; Kumar *et al.*, 2010). Therefore, the disc diffusion method was applied in this study as it is cheaper, reliable and easy to perform (Wiegand & Hilpert, 2008).

When the effectiveness of the solvents was compared, methanol was shown to have exhibited a better performance in extracting the antimicrobial compounds from the *Lansium* seed extracts compared to hexane, as the methanol extract showed better inhibition zones and inhibited more bacteria. Meanwhile, alcoholic solvents such as ethanol and methanol showed better performance in extracting flavonoids and phenolic compounds in many medicinal plants (Rispail *et al*., 2005; Sultana *et al.,* 2009; Tomsone *et al*., 2012). Methanol is an efficient solvent to degrade the cell walls and penetrate the cellular membranes, causing the intercellular compounds to be released from the cells (Tiwari *et al*., 2011). On the other hand, hexane is efficient in extracting oils and other non-polar compounds from seeds and plant parts (Gidwani *et al.,* 2010; Singh *et al*., 2012). As oil is difficult to dissolve in water or agar medium, this might have affected the results of the antibacterial activity using disc diffusion method, which was employed in the current study. However, several studies have shown that the hexane extracts of the seeds and other plant parts are capable of inhibiting the growth of *B. subtilis* and *S. aureus* using either disc diffusion or agar dilution methods (Gidwani *et al*., 2010; Ahmed-Hassan *et al*., 2011; Singh *et al*., 2012). In the present study using hexane as a solvent, the *Langsat* seed extracts demonstrated the ability to inhibit two of the bacteria studied. Nevertheless, no positive results were found from any water extracts of the *Lansium* varieties. In traditional medicine, water is always used to extract different

parts of plant (Abdalah, 2011; Borhade, 2012). However, higher water content in an extract increases the concomitant extractions of other compounds and reduces the selectivity of the extracted compounds (Tomsone *et al*., 2012). This reason may result in lower extraction rate of antimicrobial compounds in the water extracts of *Lansium* seeds.

The antimicrobial activity results show that the *Langsat* seed extract is the most effective inhibitor against the bacteria tested compared with *Duku* and *Dokong*. Some previous studies have reported the effectiveness of *Langsat* seed and organ extracts as antimicrobial agents. For examples, the fruit skin, leaf and seed methanol extracts of *L. domesticum* were found to be effective in inhibiting the growth of *Plasmodium falciparum,* demonstrating the potential as antimalarial agent (Yap & Yap, 2003; Saewann *et al*., 2006). Mayanti *et al.* (2011) reported that the methanol extract of *L. domesticum* showed a strong antifeedant activity against the fourth instar larvae of *Epilachna vigintioctopunctata.* The potential of the *Lansium* ethanol extract as antimicrobe had also been studied by Dong *et al.* (2011). The researchers reported a class of onoceranoid-type triterpenoids, found in the plant twigs (Dong *et al*., 2011), exhibited a moderate antibacterial activity against Gram-positive bacteria. The finding is in agreement with the results of the present study which showed that the *Langsat* and *Duku* seed extracts exhibited antibacterial activities towards Gram-positive bacteria, namely, *S. aureus* and *B. subtilis*. Nevertheless, has been reported on the potential of the *Lansium* seed methanol extract against the Gram-negative bacteria. The findings of this study showed that the *Langsat* seed extract showed a great potential in inhibiting the growth of harmful bacteria such as *P. aeruginosa*.

Mohamed *et al.* (1994) studied *Duku* fruit skin extracts and found their activities against *Candida lypolytica* but none was reported on the seed extracts. In the present study, the methanol extract of *Duku* seeds was found to exhibit the antibacterial activity towards *B. subtilis*.

CONCLUSION

Three varieties of *Lansium domesticum* (*Langsat*, *Duku* and *Dokong*) have been shown to have interesting patterns in their antimicrobial activities towards *Staphylococcuss aureus*, *Bacillus subtilis* and *Pseudomonas aeruginosa*. Even though these varieties shared similar morphological characteristics, the findings revealed that *the Langsat* seed extract exhibits a good potential in inhibiting the growth of two Gram-positive bacteria (*S. aureus* and *B. subtilis*) whilst moderately inhibiting the growth of Gram-negative bacterium, *P. aeruginosa.* Meanwhile, the methanol seed extract of *Duku* showed a moderate inhibition only towards *B. subtilis*. On the other hand, the seed extracts of *Dokong* did not show any potential antimicrobial inhibition. Therefore, the results of this study provide a basis for further phytochemical and antimicrobial studies on the seed extracts of the *Lansium* varieties.

ACKNOWLEDGEMENTS

Special gratitude is given to the Ministry of Science and Technology, Malaysia (MOSTI), for the FRGS grant (vote 03-03-09-09) to carry out this research.

REFERENCES

Abdalah, M. E. (2011). The study of antibacterial activity of fenugreek (*Trigonellafoenum graecum*) seeds extract. *Iraq Journal of Market Research and Consumer Protection*, *3*(6), 145-155. ISSN: 20713894.

Ahmed-Hassan, L. E., Sirat, M., Sakina, H., Ahmed Yagi, M., Koko, W. S., & Abdelwahab, S. I. (2011). *In vitro* antimicrobial activities of chloroformic, hexane and ethanolic extracts of *Citrullus lanatus* var. *citroides* (Wild melon). *Journal of Medicinal Plants Research*, *5*(8), 1338-1344.

Anago, E., Lagnika, L., Gbenou, J., Loko, F., Moudachirou, M., & Sanni, A. (2011). Antibacterial activity and phytochemical study of six medicinal plants used in Benin. *Pakistan Journal of Biological Sciences, 14*, 449-455.

Argudin, M. A., Mendoza, M. C., & Rodicio, M. R. (2010). Food poisioning and *Staphylococcus aureus* enterotoxins. *Toxin*, *2*(7), 1751-1773.

Beng, K. H., Clyde, M. M., Wickneswari, R., & Mohd Noor, N. (2000). Genetic relatedness among *Lansium domesticum* accessions using RAPD markers. *Annals of Botany*, 86, 299-307.

Borhade, S. (2012) Antibacterial activity, phytochemical analysis of water extracts *Zyzyum cumini* and analytical study by HPLC. *Asian Journal of Experimental Biological Sciences*, *3*(2), 320-324.

Buxton, R. S., & Ward, J. B. (1980) Heat-sensitive Lysis Mutants of *Bacillus subtilis* 168 Blocked at Three Different Stages of Peptidoglycan Synthesis. *Microbiology*, *120*(2), 283-293. doi: 10.1099/00221287-120-2-283.

Bruneton, J. (1999). *Pharmacognosie: Phytochimie Plantes Medicinales* (3rd Edition). TEC and DOC, France, ISBN-10: 2-7430-0315-4, pp: 794-796.

Cushnie, T. P. T., & Lamb, A. J. (2005). Antimicrobial activity of flavonoids. *International Journal of Antimicrobial Agents*, *26*, 343–356.

Dong, S. H., Zhang, C. R., Dong, L., Wu, Y., & Yue, J. M. (2011). Onoceranoid-Type Triterpenoids from *Lansium domesticum*. *Journal of Natural Product*, *74*(5), 1042–104.

Foster, T. J. (2004). The *Staphylococcus aureus* "superbug". *The Journal of Clinical Investigation, 114*(12), 1693–1696. doi:10.1172/JCI23825.

Gidwani, B., Alaspure, D., & Duragkar, N. J. (2010). Anti-inflammatory and antimicrobial activity of hexane extract of seed of *Psoralea corylifolia* linn. *International Journal of Pharma Research and Development*, *2*(10). Retrieved from www.ijprd.com129. ISSN 0974 – 9446.

Grundmann, H., Aanensen, D. M., Wijngaard, C. C., Spratt, B. G., Harmsen, D., & Friedrich, A. W. (2012). Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: A Molecular-Epidemiological Analysis. *PLOSS Medicine*, Oct 22-28.

Kumar, A., Ilavarasan, R., Jayachandran, T., Decaraman, M., Aravindhan, P., Padmanabhan, N., & Krishnan, M. R. V. (2009). Phytochemicals Investigation on a Tropical Plant, *Syzygium cumini* from Kattuppalayam, Erode District, Tamil Nadu, South India. *Pakistan Journal of Nutrition*, *8*(1), 83-85.

Kumar, R., Sandeep Kumar Shrivastava, S. K., & Chakraborti, A. (2010). Comparison of broth dilution and disc diffusion method for the antifungal susceptibility testing of *Aspergillus flavus. American Journal of Biomedical Sciences*, *2*(3), 202-208; doi: 10.5099/aj100300202 nwpii.com/ajbms. ISSN: 1937-9080.

Lense, O. (2011). Biological screening of selected tropical medicinal plant species utilised by local people, Manokwari, West Papua Province. *Bioscience*, *3*(3), 145-150.

Mayanti, T., Tjokronegoro, R., Supratman, U., Mukhtar, M. R., Awang, K., & A. Hadi, A. H. (2011). Antifeedant Triterpenoids from the Seeds and Bark of *Lansium domesticum* cv Kokossan (Meliaceae). *Molecules*, *16*, 2785-2795. doi: 10.3390/molecules16042785.

Milici, M. E., Maida, C. M., Spreghini, E., Ravazzolo, B., Oliveri, S., Scalise, G., & Barchiesi, F. (2007). Comparison between disk diffusion and microdilution methods for determining susceptibility of clinical fungal Isolates to Caspofungin. *Journal of Clinical Microbiology*, *45*(11), 3529–3533. PMCID: PMC2168513.

Mohamed, S., Hassan, Z., & Abd Hamid, N. Z. (1994). *Antimicrobial activity of some tropical fruit wastes (Guava, Starfruit, Banana, Papaya, Passionfruit, Langsat, Duku, Rambutan and Rambai)*. Retrieved from http://psasir.upm.edu.my/3194/1/.

Morton, J. (1987). *Lansium domesticum Corr.* In *Fruits of warm climates*. Retrieved from http://www.hort.purdue.edu/newcrop/morton/langsat.html. Miami, FL. pp. 201–203.

Onwukaeme, D. N. O., Ikuegbvweha, T. B., & Asonnye, C. C. (2007). Evaluation of phytochemical constituents, antibacterial activities and effect of exudate of *Pycanthus angolensis* Weld Warb (Myristicaceae) on corneal ulcers in rabbits. *Tropical Journal of Pharmaceutical Research, 6*(2), 725-730.

Rispail, N., Morris, P., & Webb, K. J. (2005) in A.J. Marquez. *Lotus japonicas* handbook. *Phenolic Compounds: Extraction and Analysis*. Springer, 349-355.

Saewann, N., Sutherland, J. D., & Chantrapromma, K. (2006) Antimalarial tetranortriterpenoids from the seeds of *Lansium domesticum* Corr. *Phytochemistry*, *67*, 2288–2293.

Sapii, A. T., Yunus, N., Muda, M., & Tham, S. L. (2000). Postharvest quality changes in Dokong (*Lansium domesticum* Corr.) harvested at different stages of ripeness. Quality assurance in agricultural produce, ACIAR Proceedings 100. pp.201-205.

Serrano, M. C., Ramírez, M., Morilla, D., Valverde, A., Chávez1, M., Espinel-Ingroff, A., Claro, R., Fernández, A., Almeida, C., & Martín-Mazuelos, E. (2004). A comparative study of the disc diffusion method with the broth microdilution and Etest methods for voriconazole susceptibility testing of *Aspergillus* spp. *Journal of Antimicrobial Chemotherapy*, *53*, 739–742. DOI: 10.1093/jac/dkh172.

Singh, R., Dar, S. A., & Sharma, P. (2012). Antibacterial activity and toxicological evaluation of semi purified hexane extract of *Urtica dioica* leaves. *Research Journal of Medicine Plant*, *6*(2), 123-135.

Solidum, J. N. (2012) Potential nutritional and medicinal sources from fruit peels in Manila, Philippines. *International Journal of Bioscience, Biochemistry and Bioinformatics, 2*(4), 270-274.

Sri Yulita, K. (2011). Genetic variations of *Lansium domesticum* Corr. accessions from Java, Sumatra and Ceram based on Random Amplified Polymorphic DNA fingerprints *Biodiversitas*, *12*(3), 125-130. ISSN: 1412-033X (printed edition), ISSN: 2085-4722 (electronic).

Stojek, N. M., Szymanska, J., & Dutkiewicz, J. (2008). Gram-negative bacteria in water distribution systems of hospitals. *Annals of Agricultural and Environmental Medicine*, *15*, 135-142.

Sultana, B., Anwar, F., & Ashraf, M. (2009). Effect of extraction solvent/technique on the antioxidant activity of selected medicinal plant extracts. *Molecules, 14*, 2167-2180. doi: 10.3390/molecules14062167.

Tam, N. K. M., Uyen, N. Q., Hong, H. A., Duc, L. H., Hoa, T. T., Serra, C. R., Henrigues, A. O., & Cuttinmg, S. M. (2006). The Intestinal Life Cycle of *Bacillus subtilis* and Close Relatives. *Journal of Bacteriology, 188*(7), 2692–2700. doi: 10.1128/JB.188.7.2692-2700.2006PMCID:PMC1428398

Te-chato, S., Lim, M., & Masahiro, M. S. (2005). Comparison of cultivar identification methods of longkong, langsat and duku: *Lansium* spp. *Journal of Science and Technology*, *27*(3), 465-472.

Tiwari, A. K., Swapna, M., Ayesha, S. B., Zehra, A., Agawane, S. B. & Madhusudana, K. (2011). Identification of proglycemic and antihyperglycemic activity in antioxidant rich fraction of some common food grains. *International Food Research Journal, 18*(3), 915-923.

Tomsone, L., Kruma, Z., & Galoburda, R. (2012). Comparison of different solvents and extraction methods for isolation of phenolic compounds from Horseradish roots (*Armoracia rusticana*). *World Academy of Science, Engineering and Technology*, *64*, 903-908.

Wiegand, I., & Hilpert, K. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Hancock RENat Protoc.*, *3*(2), 163-75. doi: 10.1038/nprot.2007.521.

Yapp, D. T. T., & Yap, S. Y. (2003). *Lansium domesticum*: skin and leaf extracts of this fruit tree interrupt the lifecycle of *Plasmodium falciparum*, and are active towards a chloroquine-resistant strain of the parasite (T9) in vitro. *Journal of Ethnopharmacology*, *85*(1), 145–150.

Yunus, N. (1997). Flowering and fruiting of dokong (*Lansium domesticum* Corr.). In S. Vejaysegaran, M. Pauziah, M.S. Mohamed, and S. Ahmad Tarmizi (Eds.), *Proceedings of An International Conference on Tropical Fruits*. 23–26 July 1996, Kuala Lumpur. Serdang, Malaysian Agricultural Research and Development Institute (MARDI), 3, 281–286.

Pertanika J. Sci. & Technol. 22 (2): 541 - 552 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Study on the Performances of Multivariate Exponentially Weighted Moving Average (MEWMA) and Multivariate Synthetic Charts

Ellappan, S.[1] and Khoo Michael, B. C.[2]*

[1]*School of General and Foundation Studies, AIMST University, 08100 Bedong, Kedah Darul Aman, Malaysia*

[2]*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia*

ABSTRACT

A multivariate control chart is a common tool used for monitoring and controlling a process whose quality is determined by several related variables. The objective of this study is to compare the performances of the multivariate exponentially weighted moving average (MEWMA) and the multivariate synthetic T^2 control charts, for the case of a multivariate normally distributed process. A comparative study is made based on the average run length (ARL) performances of the control charts, using the simulation method, in order to identify the chart having the best performance in monitoring the process mean vector. The performances of the two charts, for different sample sizes and correlation coefficients, are presented in this paper. It was found that the MEWMA chart outperformed synthetic T^2 chart for small shifts but the latter prevailed for moderate shifts. Both charts performed equally well for larger shifts. In addition, the performances of both MEWMA and synthetic T^2 charts were found to be influenced by sample size and correlation coefficient. The two charts' performances improved as the sample size and correlation coefficient increased for small and moderate shifts, but the charts' performances did not depend on sample size and correlation coefficient when the shift was large.

Keywords: Average run length (ARL), MEWMA chart, multivariate synthetic chart, out-of-control, in-control

Article history:
Received: 15 August 2012
Accepted: 8 November 2012

Email addresses:
Ellappan, S. (sugunes@gmail.com),
Khoo, Michael B. C. (mkbc@usm.my)
*Corresponding Author

INTRODUCTION

Statistical process control (SPC) is used to describe a set of problem solving tools that have been used to measure and analyze variations in processes. SPC can be implemented by common key monitoring and investigation tool, which is the control chart.

ISSN: 0128-7702

It is a statistical tool used to detect excessive process variability due to specific assignable causes that can be corrected (Shewhart, 1986).

Univariate control charts involving the measurement of a single variable include the popular Shewhart $\overline{X}-R$ and $\overline{X}-S$ charts, as well as the moving average and moving range charts. When more than one variable is collected, the relationship between several different variables need to be shown, resulting in multivariate charts such as Hotelling's T^2, multivariate exponentially weighted moving average (MEWMA), and multivariate cumulative sum (MCUSUM) charts (Montgomery, 2009).

Ghute and Shirke (2008) and Lee and Khoo (2006) suggested two approaches on multivariate synthetic and MEWMA charts. Lee and Khoo (2006) also proposed an optimal statistical design of the MEWMA chart based on the average run length (ARL) and median run length (MRL) criteria. The optimal design of the MEWMA chart was made simple by using plots of the optimal smoothing constant r and plots of the optimal control limit, H. The constructed plots give approximation of the optimal r and its corresponding H for the given in-control ARL and MRL. Meanwhile, Ghute and Shirke (2008) proposed a multivariate synthetic control chart, which is a combination of the Hotelling's T^2 chart and the conforming run length (CRL) chart for monitoring the mean vector of a multivariate normally distributed process.

Some of the recent works on the multivariate synthetic charts were done by Khoo *et al*. (2009), Khoo *et al*. (2011), and Aparisi and de Luna (2009), while those on the MEWMA charts were carried out by Niaki and Ershadi (2012), Alkahtani and Schaffer (2012), Xie *et al*. (2011), Alipour and Noorossana (2010), and Khoo and Quah (2004).

The objective of this study was to compare the ARL performances of the optimal MEWMA chart and the optimal multivariate synthetic T^2 chart. In the optimal design of the MEWMA chart, the method proposed by Lee and Khoo (2006) is used. Note that the optimal multivariate synthetic T^2 chart was proposed by Ghute and Shirke (2008). ARL, which is the expected value of the run length, is used to measure the performance of the control charts. When comparisons were made, the charts were first designed to have a common in-control ARL and then the out-of-control ARLs for a given shift in the process were compared. The common in-control ARL values of the charts were set arbitrarily by the investigator. The chart with a better performance is the chart with a smaller out-of-control ARL.

In the next section, the MEWMA and the multivariate synthetic T^2 charts are reviewed. This is followed by a comparison between the performances of MEWMA and the multivariate synthetic T^2 charts for different sample sizes and correlation coefficients through a simulation study. Finally, conclusions and some suggestions for further research are presented.

MULTIVARIATE CONTROL CHARTS

The MEWMA Chart

The multivariate Shewhart type control chart, like the Hotelling's T^2 chart, uses the information from the current sampling and is insensitive to small and moderate shifts in the mean vector. The MEWMA control chart is one of the multivariate control charts developed to solve this problem (Bersimis *et al*., 2007).

The MEWMA chart proposed by Lowry *et al*. (1992) is a logical extension of the univariate exponentially weighted moving average (EWMA) chart. The MEWMA statistics are written as follows:

$$Z_i = r\overline{X}_i + (1-r)Z_{i-1}, \quad i = 1, 2, \ldots, \qquad [1]$$

where Z = 0 and $0 < r \leq 1$. Note that Z_0 is an initial vector and r is a smoothing constant. The multivariate observations $\overline{X}_1, \overline{X}_2, \ldots$, are assumed to be independently and identically distributed multivariate normal random vectors, each with p quality characteristic. It is assumed that μ_0 is the in-control process mean vector of zeros and Σ is the covariance matrix. The multivariate control charts for the mean vector are designed to detect shifts over time from this in-control vector. The quantity plotted on the MEWMA chart is:

$$T_i^2 = Z_i' \sum\nolimits_{Z_i}^{-1} Z_i, \qquad i = 1, 2, \ldots, \qquad [2]$$

where $\sum_{Z_i} = \{r[1-(1-r)^{2i}]/[(2-r)n]\} \sum$ is the covariance matrix of Z_i. The MEWMA chart gives an out-of-control signal as soon as $T_i^2 > H$, where H is the control limit which is a constant chosen to give a desired in-control ARL.

Lowry *et al*. (1992) have shown that the ARL performance of the MEWMA chart depends on the mean vector μ and the covariance matrix Σ only through the value of the non-centrality parameter δ, where,

$$\delta = \left(\mu' \sum\nolimits^{-1} \mu\right)^{1/2} \qquad [3]$$

The Optimal MEWMA Control Chart

A MEWMA chart is optimal in detecting a shift if the MEWMA chart has the smallest out-of-control ARL among all MEWMA charts with the same in-control ARL for a particular shift. The optimal design specifies the optimal selection of the chart parameters, namely, the smoothing constant, r and its corresponding control limit, H for a MEWMA chart. Prabhu and Runger (1997) developed a design strategy for the MEWMA chart to obtain the optimal values of the chart's parameters. Lee and Khoo (2006) extended the optimal design approach of Prabhu and Runger (1997) for the MEWMA chart by recommending the four steps method in order to select the optimal chart parameters. In designing the optimal MEWMA chart, Lee and Khoo (2006) used the Markov chain method to select the control chart's parameters based on the desired size of a shift that must be detected quickly and the desired in-control ARL.

The Multivariate Synthetic Control Chart

The recent approaches in improving the performances of the univariate control charts include combining the Shewhart $\overline{X}$ control chart with a conforming run length (CRL) chart, leading to a synthetic control chart (Wu & Spedding, 2000). Ghute and Shirke (2008) extended the univariate synthetic chart for multivariate processes by developing the synthetic T^2 chart for monitoring the mean vector of a multivariate normally distributed process. This synthetic T^2 chart is a combination of the Hotelling's T^2 chart and the CRL chart.

Hotelling's T^2/S Sub-chart

The Hotelling's T^2/S sub-chart used in this paper is based on the Hotelling's T^2 chart that was used to construct the synthetic T^2 chart proposed by Ghute and Shirke (2008). The Hotelling's T^2 statistic is a generalization of the Student's t statistic used in the multivariate hypothesis testing. Consider the vectors representing the measurements of the p process characteristics, $X_1, X_2, ..., X_n$, as a random sample from the p – variate normal distribution. The mean vector is μ and the covariance matrix is Σ. μ_0 is the desired mean vector and Σ_0 is the desired covariance matrix. The intent is to detect the shifts in μ.

The Hotelling's T^2 charting statistic used in monitoring the p quality characteristics is given by:

$$T^2 = n(\overline{X} - \mu_0)' \sum_0^{-1} \overline{X} - \mu_0, \quad [4]$$

where n is the sample size, $\overline{X}$ is the ($p \times 1$) vector of sample means, and $\sum_0^{-1}$ is the inverse of the ($p \times p$) covariance matrix.

The upper control limit is used to determine if a process is in-control or out-of-control based on the location of the plotted points taken from the sample statistics computed from independent subgroups of observations. The upper control limit takes the form:

$$UCL = \chi^2_{p,\alpha} \quad [5]$$

where $\chi^2_{p,\alpha}$ is upper 100α percentage point of the chi-square distribution with p degrees of freedom and α is the probability of the Type I error for the T^2 control chart.

When the process is in-control, the plotted points lie below *UCL*, indicating that the T^2 statistic is distributed as a chi-square variable with p degrees of freedom. When the process is out-of-control, a plotted point lies above *UCL*, showing that the T^2 statistic is distributed as a non-central chi-square variable with p degrees of freedom and a non-centrality parameter, λ^2, given as:

$$\lambda^2 = n(\mu - \mu_0)' \sum_0^{-1} (\mu - \mu_0) = nd^2 \quad [6]$$

The Mahalanobis distance, $d = \sqrt{(\mu - \mu_0)' \sum_0^{-1} (\mu - \mu_0)}$, is used to measure the change in the process mean vector.

The ARL for the Hotelling's T^2 chart can be calculated as:

$$ARL = \frac{1}{P} \quad [7]$$

where $P = \Pr(T^2 > UCL \mid d)$. The on-target value of P can be determined as:

$$P(0) = \Pr(T^2 > UCL \mid d = 0) = 1 - F_p(UCL) \quad [8]$$

where $F_p(x) = \Pr(\chi^2_p \leq x)$ is the cumulative distribution function of a central chi-square distribution with p degrees of freedom. The off-target value of P can be determined from:

$$P(d) = \Pr(T^2 > UCL \mid d \neq 0) = 1 - F_{p,\lambda^2}(UCL) \quad [9]$$

where $F_{p,\lambda^2}(x) = \Pr(\chi_p^2 \leq x)$ is the cumulative distribution function of a non-central chi-square distribution with p degrees of freedom and non-centrality parameter, λ^2. Equations [4]-[9] are obtained from Ghute and Shirke (2008).

Conforming Run Length, CRL/S Sub-chart

The CRL chart was proposed by Bourke (1991) and it was originally developed for attribute quality control to detect shifts in the fraction non-conforming, θ. The CRL/*S* sub-chart employed in this paper is based on the CRL chart used to construct the synthetic T^2 chart proposed by Ghute and Shirke (2008).

Consider CRL which follows a geometric distribution with parameter, θ. Its expected value μ_{CRL} and the cumulative distribution function F_θ (*CRL*) can be calculated as follows:

$$\mu_{CRL} = \frac{1}{\theta} \quad [10]$$

$$F_\theta(CRL) = 1 - (1-\theta)^{CRL},\ CRL = 1, 2, \ldots \quad [11]$$

Equation [11] shows that the distribution of *CRL* changes with θ. The expected value of CRL decreases as θ increases and vice versa.

The CRL chart detects an increase in θ when $CRL \leq L$, where L is the lower limit required for the chart. The average number of the CRL samples required to detect an increase in θ is the average run length of the *CRL* chart, denoted as ARL_{CRL} (Wu & Spedding, 2000) and can be calculated as:

$$ARL_{CRL} = \frac{1}{F_\theta(L)} = \frac{1}{1-(1-\theta)^L} \quad [12]$$

The average number of the inspected units (ANS_{CRL}) required to signal a fraction non-conforming shift can be calculated as the product of Equations (10) and (12) shown below:

$$ANS_{CRL} = \mu_{CRL} \times ARL_{CRL}$$

$$= \frac{1}{\theta} \times \frac{1}{1-(1-\theta)^L} \quad [13]$$

Equations [10] – [13] were obtained from Ghute and Shirke (2008).

The Optimal Multivariate Synthetic Control Chart

Ghute and Shirke (2008) developed a synthetic T^2 chart based on ARL. Let $ARL_s(d)$ denote the average number of the T^2 samples required for a synthetic T^2 chart to signal a shift of magnitude d in the mean vector. The $ARL_s(d)$ can be calculated from Equations [9] and [13] and is given as follows:

$$ARL_s(d) = \frac{1}{P(d)} \times \frac{1}{1-(1-P(d))^L} \quad [14]$$

where L is the lower control limit of the CRL/S sub-chart and $P(d)$ is the detecting power. The $ARL_s(0)$ should be large in order for the false alarm rate to be kept low. The $ARL_s(d)$ should be small so that the detection of the process can be made quickly. Suitable values of L and the upper control limit of the T^2/S sub-chart, UCL, must be obtained in order to design a synthetic T^2 chart. Ghute and Shirke (2008) solved the optimization problem and designed the synthetic control chart by minimizing the following out-of-control ARL:

$$ARL_s(d^*) = \frac{1}{P(d^*)} \times \frac{1}{(1-P(d^*))^L} \quad [15]$$

where d^* is the optimal shift, and a quick detection is needed.

RESULTS AND DISCUSSION

This paper involves a comparative study on the performances of MEWMA and the multivariate synthetic T^2 control charts based on the case of a multivariate normally distributed process. The Statistical Analysis System (SAS, version 9.1) software was used to conduct all the simulations in this paper.

A brief explanation on how the simulation was conducted is as follows: Firstly, samples of size $n \in \{5,10\}$ having bivariate normal observations (the number of variables, $p = 2$ is considered) are generated. Then, Equations [1] and [2] are used to compute the MEWMA statistics, while Equation [4] is used to compute the synthetic T^2 sub-chart's statistic. The number of samples required for each of the two charts to signal an out-of-control signal is recorded as the run length of the chart. The processes of generating samples with bivariate normal observations and recording the run length at each trial are repeated for 10000 trials for each of the charts. Finally, the average of the run lengths, based on the 10000 trials for each of the charts, is computed as ARL of the charts.

The out-of-control ARL performances of MEWMA and synthetic T^2 charts were compared for various sizes of shifts regarded as important, where the in-control ARL is set as 370. The ARL profiles for the optimal MEWMA and optimal synthetic T^2 control charts were also developed based on the sample sizes, $n \in \{5,10\}$, the number of variables, $p = 2$, design shifts, $d^* \in \{0.5,1.0\}$, and correlation coefficients, $\rho \in \{0.2,0.5,0.8\}$. Note that due to cost considerations, in practice, a large sample size is rarely used in process monitoring. For this reason, only small and moderate sample sizes, i.e. $n \in \{5,10\}$, were considered in this study. Thus, it should be pointed out that considering other sample sizes would give similar results as the MEWMA and synthetic T^2 charts' performances only depend on the Mahalanobis distance, d. The comparative study on the ARL performances of the control charts will identify the better chart in monitoring the process mean vector. The performances of the charts for different sample sizes and correlation coefficients are presented.

The ARL Profiles of the MEWMA Control Chart

Table 1 shows the combinations of optimal parameters of the MEWMA chart derived using the four-step method by Lee and Khoo (2006). This four-step method is as follows:

Step 1 : Specify the smallest acceptable in-control ARL.

Step 2 : Decide on the smallest shift, *d,* that must be detected quickly. For this size of shift, *d*, determine the smoothing constant, *r,* that gives the in-control ARL in Step 1.

Step 3 : Based on the optimal *r* obtained in Step 2, determine the control limit, *H*.

Step 4 : Perform a sensitivity analysis by comparing the out-of-control ARL of the optimal pair (*r*, *H*) to other choices of *r* and *H* that produce the same in-control ARL. Then, select the pair (*r*, *H*) with the most desirable performance, in terms of the out-of-control ARL.

The smallest acceptable in-control ARL of 370, for the design shifts, $d^* \in \{0.5,1.0\}$, is used to derive the smoothing constant, *r*, which is then used to derive the control limit *H*.

TABLE 1
The combination of d^*, *r*, and *H* for $p = 2$ with in-control ARL of 370

Design shifts d^*	Smoothing constant *r*	Control limit *H*
0.5	0.045	8.7
1.0	0.13	10.45

The values from Table 1 are then used to compute the ARL values. The ARL profiles for the optimal MEWMA chart when $d^* \in \{0.5,1.0\}$, $n \in \{5,10\}$ and for the correlation coefficients, $\rho \in \{0.2,0.5,0.8\}$ are shown in Tables 3 – 6.

The ARL Profiles of the Multivariate Synthetic Control Chart

Table 2 shows the combinations of the optimal parameters of the synthetic T^2 control chart using the six steps procedure proposed by Ghute and Shirke (2008). The set of (*L*,*UCL*) which generated the smallest $ARL_s(d^*)$ is used as the optimal design parameters for the synthetic T^2 control chart. The optimal values for *L* and *UCL* are shown for the sample size, $n \in \{5,10\}$ and the design shifts, $d^* \in \{0.5,1.0\}$. These optimal values of (*L*,*UCL*) are then used to compute the ARL values for the correlation coefficients, $\rho \in \{0.2,0.5,0.8\}$. The ARL profiles for the optimal synthetic T^2 control chart when $d^* \in \{0.5,1.0\}$ are provided in Tables 3 – 6.

TABLE 2
The optimal design parameters of the synthetic T^2 control chart

Sample size, n	Design shifts, d^*	Lower control limit of CRL sub-chart, L	Upper control limit of T^2 sub-chart, UCL
5	0.5	15	8.52408
	1.0	5	7.47532
10	0.5	8	7.92678
	1.0	2	6.58792

A Comparison of the MEWMA and Multivariate Synthetic Control Charts

A control chart is optimal in detecting a shift if the chart has the smallest out-of-control ARL among all the charts of its kind having the same in-control ARL for a particular shift of interest. From Tables 3– 6, it can be seen that the optimal MEWMA control chart performs better than the optimal synthetic T^2 control chart when the shift, d, is small. For $d < 0.6$, the MEWMA control chart has a better out-of-control ARL performance in detecting the process shift. For moderate shifts, $0.8 \leq d \leq 3.0$, the synthetic T^2 control chart detects an out-of-control signal quicker than the MEWMA control chart. For larger shifts, $d > 3.0$, both the control charts have the same performances in detecting the out-of-control signals.

In the case of the performances of the MEWMA and the synthetic T^2 control charts for different sample sizes, both charts display similar characteristics. Both the control charts perform better in detecting out-of-control signals for sample size $n = 10$ compared to $n = 5$. This is true for both moderate and small shifts, $d < 2.0$. For bigger shifts ($d > 2.0$), the performances of both the charts are independent of the sample size.

There are some similarities and variations in the performances of the charts for different correlation coefficients. Generally, it can be seen from Tables 3 – 6 that as ρ increases, the out-of-control ARL values decrease for both charts. This means that the charts perform better in detecting the out-of-control signals when ρ increases. This is evident especially in the case when $d < 2.0$ for both charts. Similarly, this is also true for $2.0 \leq d \leq 3.0$ for the MEWMA control chart. As for the synthetic T^2 control chart, its performance is independent of the correlation coefficient for moderate shifts, whereas for bigger shifts, the performances of both the charts are independent of the correlation coefficients.

TABLE 3
The ARL profiles for optimal MEWMA and optimal synthetic T^2 control charts when $d^* = 0.5$ and $n = 5$

D	OPTIMAL MEWMA			OPTIMAL SYNTHETIC T^2 CHART		
	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$
0	371.6	371.6	371.6	370.5	370.5	370.5
0.2	36.8	30.7	18.7	177.9	149.5	72.8
0.4	14.6	12.6	8.2	44.6	30.9	10.1
0.6	9.1	7.9	5.4	13.0	9.0	3.3
0.8	6.6	5.8	4.1	5.5	4.0	1.7
1.0	5.3	4.7	3.3	3.1	2.3	1.2
1.5	3.6	3.2	2.2	1.3	1.2	1.0
2.0	2.8	2.4	2.0	1.0	1.0	1.0
2.5	2.2	2.0	1.7	1.0	1.0	1.0
3.0	2.0	2.0	1.1	1.0	1.0	1.0
3.5	2.0	1.8	1.0	1.0	1.0	1.0
4.0	1.8	1.3	1.0	1.0	1.0	1.0
4.5	1.4	1.0	1.0	1.0	1.0	1.0
5.0	1.1	1.0	1.0	1.0	1.0	1.0
5.5	1.0	1.0	1.0	1.0	1.0	1.0
6.0	1.0	1.0	1.0	1.0	1.0	1.0

TABLE 4
The ARL profiles for optimal MEWMA and optimal synthetic T^2 control charts when $d^* = 1.0$ and $n = 5$

D	OPTIMAL MEWMA			OPTIMAL SYNTHETIC T^2 CHART		
	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$
0	368.8	368.8	368.8	368.5	368.5	368.5
0.2	45.6	36.2	18.6	187.2	160.8	81.7
0.4	13.2	10.9	6.5	49.4	35.0	10.6
0.6	7.3	6.2	4.0	14.2	9.2	2.9
0.8	5.0	4.4	2.9	5.3	3.6	1.5
1.0	3.9	3.4	2.4	2.7	2.0	1.1
1.5	2.6	2.3	1.8	1.3	1.1	1.0
2.0	2.0	1.9	1.2	1.0	1.0	1.0
2.5	1.8	1.5	1.0	1.0	1.0	1.0
3.0	1.4	1.1	1.0	1.0	1.0	1.0
3.5	1.1	1.0	1.0	1.0	1.0	1.0
4.0	1.0	1.0	1.0	1.0	1.0	1.0
4.5	1.0	1.0	1.0	1.0	1.0	1.0
5.0	1.0	1.0	1.0	1.0	1.0	1.0
5.5	1.0	1.0	1.0	1.0	1.0	1.0
6.0	1.0	1.0	1.0	1.0	1.0	1.0

TABLE 5
The ARL profiles for optimal MEWMA and optimal synthetic control charts T^2 when $d^* = 0.5$ and $n = 10$

d	OPTIMAL MEWMA			OPTIMAL SYNTHETIC T^2 CHART		
	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$
0	372.3	372.3	372.3	370.6	370.6	370.6
0.2	22.6	19.3	12.1	107.4	82.5	31.4
0.4	9.7	8.4	5.7	16.1	10.9	3.4
0.6	6.2	5.5	3.8	4.4	3.1	1.5
0.8	4.7	4.1	3.0	2.1	1.7	1.1
1.0	3.8	3.4	2.4	1.4	1.2	1.0
1.5	2.6	2.3	2.0	1.0	1.0	1.0
2.0	2.0	2.0	1.3	1.0	1.0	1.0
2.5	2.0	1.8	1.0	1.0	1.0	1.0
3.0	1.6	1.1	1.0	1.0	1.0	1.0
3.5	1.1	1.0	1.0	1.0	1.0	1.0
4.0	1.0	1.0	1.0	1.0	1.0	1.0
4.5	1.0	1.0	1.0	1.0	1.0	1.0
5.0	1.0	1.0	1.0	1.0	1.0	1.0
5.5	1.0	1.0	1.0	1.0	1.0	1.0
6.0	1.0	1.0	1.0	1.0	1.0	1.0

TABLE 6
The ARL profiles for optimal MEWMA and optimal synthetic T^2 control charts when $d^* = 1.0$ and $n = 10$

d	OPTIMAL MEWMA			OPTIMAL SYNTHETIC T^2 CHART		
	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$
0	363.8	363.8	363.8	373.2	373.2	373.2
0.2	23.9	19.2	10.5	123.0	97.9	38.7
0.4	7.8	6.7	1.2	20.7	13.6	3.9
0.6	4.7	4.1	2.8	5.1	3.4	1.4
0.8	3.4	3.0	2.1	2.1	1.6	1.0
1.0	2.7	2.4	1.9	1.3	1.1	1.0
1.5	2.0	1.8	1.1	1.0	1.0	1.0
2.0	1.5	1.2	1.0	1.0	1.0	1.0
2.5	1.1	1.0	1.0	1.0	1.0	1.0
3.0	1.0	1.0	1.0	1.0	1.0	1.0
3.5	1.0	1.0	1.0	1.0	1.0	1.0
4.0	1.0	1.0	1.0	1.0	1.0	1.0
4.5	1.0	1.0	1.0	1.0	1.0	1.0
5.0	1.0	1.0	1.0	1.0	1.0	1.0
5.5	1.0	1.0	1.0	1.0	1.0	1.0
6.0	1.0	1.0	1.0	1.0	1.0	1.0

CONCLUSION

The comparisons made show that the MEWMA control chart performs better than the synthetic T^2 control chart when the shifts are small. However, the synthetic T^2 control chart performs better for moderate shifts. Nonetheless, both the charts perform equally well for larger shifts. The sample sizes and the correlation coefficients have been found to influence the detection of the out-of-control signals. Both charts perform better for larger sample sizes for both small and moderate shifts. However, the performances of both charts are not dependent on the sample sizes for larger shifts. When the correlation coefficient increases, the performances of both charts also improve, especially for small shifts. The performances of both the charts are independent of the correlation coefficient for larger shifts.

There are many areas on the multivariate control charts that are worthy of further research. The construction of a multivariate control chart is usually based on the multivariate normality assumption. Hence, the designs of the multivariate charts for skewed or heavy tailed populations, i.e. when the normality assumption is not fulfilled, are potential topics for further research. In addition, future works can also be conducted to compare the MRL performances of the MEWMA and synthetic T^2 charts to give more meaningful interpretations of the in-control and out-of-control performances of the charts.

REFERENCES

Alkahtani, S., & Schaffer, J. (2012). A double multivariate exponentially weighted moving average (dMEWMA) control chart for a process location monitoring. *Communications in Statistics-Simulation and Computation*, *41*, 238-252.

Aparisi, F., & de Luna, M. A. (2009). The design and performance of the multivariate synthetic T^2 control chart. *Communications in Statistics-Theory and Methods*, 38, 173-192.

Bersimis, S., Psarakis, S., & Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International*, *23*, 517-543.

Bourke, P. D. (1991). Detecting a shift in fraction nonconforming run-length control charts with 100% inspection. *Journal of Quality Technology*, *23*, 225-238.

Ghute, V. B., & Shirke, D. T. (2008). A multivariate synthetic control chart for monitoring process mean vector. *Communications in Statistics-Theory and Methods*, *37*, 2136-2148.

Lee, M. H., & Khoo, M. B. C. (2006). Optimal statistical design of a multivariate EWMA chart based on ARL and MRL. *Communications in Statistics-Simulation and Computation*, *35*, 831-847.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, *34*, 46-53.

Khoo, M. B. C., Abdu A. M. A., & Wu, Z. (2009). A multivariate synthetic control chart for monitoring the process mean vector of skewed populations using weighted standard deviations. *Communications in Statistics-Simulation and Computation*, *38*, 1493-1518.

Khoo, M. B. C., & Quah, S. H. (2004). Pengiraan persentil taburan panjang larian bagi carta kawalan purata bergerak berpemberat eksponen multivariat. *Pertanika Journal of Science & Technology*, *12*, 33-43.

Khoo, M. B. C., Wong, V. H., Wu, Z., & Castagliola, P. (2011). Optimal designs of the multivariate synthetic chart for monitoring the process mean vector based on median run length. *Quality and Reliability Engineering International, 27,* 981-997.

Montgomery, D. C. (2009). *Statistical quality control: a modern introduction* (6th Edition). Asia: John Wiley & Sons.

Niaki, S. T. A., & Ershadi, M. J. (2012). A hybrid ant colony, Markov chain, and experimental design approach for statistically constrained economic design of MEWMA control charts. *Expert Systems with Applications*, *39*, 3265-3275.

Prabhu, S. S., & Runger, G. C. (1997). Designing a multivariate EWMA control chart. *Journal of Quality Technology*, *29*, 8-15.

Shewhart, W. A. (1986). *Statistical method from the viewpoint of quality control*, Mineola, NY: Dover publications.

Wu, Z., & Spedding, T. A. (2000). A synthetic control chart for detecting small shifts in the process mean. *Journal of Quality Technology*, *32,* 32-38.

Xie, Y., Xie, M., & Goh, T. N. (2011). Two MEWMA charts for Gumbel's bivariate exponential distribution. *Journal of Quality Technology*, *43*, 50-65.

Pertanika J. Sci. & Technol. 22 (2): 553 - 566 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Experimental and Modelling Study of Selected Heavy Metals Removal from Aqueous Solution Using *Scylla serrata* as Biosorbent

Aris A. Z.[1]*, Ismail F. A.[1], Ng, H. Y.[1] and Praveena, S. M.[2]

[1]*Environmental Forensics Research Centre, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

[2]*Environmental and Occupational Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

This study was conducted using crab shells as a biosorbent to remove Cu and Cd with different initial concentrations of 1, 5, 10, 15, and 20 mg/L in a biosorption treatment process. Crab shells were selected as biosorbents due to their abundance in the environment and ready availability as waste products from the market place. This study aimed to determine the ability of *Scylla Serrata* shells to remove Cu and Cd in an aqueous solution, as well as to provide a comparison of the removal rate between the two metals. The data were incorporated into hydrochemical software, PHREEQC, to investigate the chemical speciation distribution of each heavy metal. The shells of *S. serrata* were found to have a significant (p<0.05) ability to remove Cu and Cd following the treatment. After six hours of treatment, the crab shells had removed 60 to 80% of both metals. However, the highest removal percentage was achieved for Cu with up to 94.7% removal rate in 5 mg/L initial Cu concentration, while 85.1% of Cd was removed in 1 mg/L initial solution, respectively. It can be concluded that the shells of *S. serrata* could remove Cu and Cd better with significant results (p<0.05) in 1 and 5 mg/L initial concentrations, respectively.

Keywords: Biosorption, Cd, Cu, crab shell, PHREEQC

Article history:
Received: 1 Oktober 2012
Accepted: 7 January 2013

Email addresses:
Aris A. Z. (zaharin@upm.edu.my),
Ismail F. A. (farhah.amalya.ismail@gmail.com),
Ng, H. Y. (huiying1803@gmail.com),
Praveena, S. M. (smpraveena@gmail.com)
*Corresponding Author

INTRODUCTION

Biosorbents can be divided into living or dead materials, which are capable of carrying biosorption mechanisms (Shareef, 2009; Yadanaparthi *et al.,* 2009). The benefits of using living biosorbents are that they are self-renewing, and therefore, the active transport may take higher metal concentrations from

ISSN: 0128-7702

the media of aqueous solution. Some examples of these biosorbent activities include bacteria (Adeogun *et al.*, 2012; Castillo-Zacarías *et al.*, 2011), fungi (Yazdani *et al*., 2009), algae (Leborans & Novillo, 1996), and living plants (Park *et al.,* 2008). These biosorbents are only different in their uptake mechanisms. In contrast, dead biosorbents are chosen because of their ability to avoid the toxicity problems towards the sample and the environment. They do not cause secondary pollution or have any side effect prior to the treatment. In addition, defunct biosorbents often require shorter time, are easier to apply, and most importantly, provide a non-destructive recovery of bound metals, which allow regeneration of the biosorbent materials. Thus, the majority of biosorption studies preferred using defunct biosorbents, which offer more advantages compared to living organisms (Schiewer & Volesky, 2000). Biosorption of heavy metals is an alternative technology in wastewater treatment, instead of using additional chemical and high cost technology. Natural biological materials from different kinds of waste have proven the same efficiency rate with less harm to the environment and human health (Demirbas, 2008).

Crab shell is one of the most applied and efficient natural products used as a biosorbent in treating metal-polluted water (Evan *et al*., 2008; Kim, 2003) and toxic materials (Alkan *et al*., 2004). Generally, crab shells consist of calcium carbonate and protein (29.19%), ash (40.60%), lipid (1.35%) and chitin (26.65%) on a dry weight basis (Tokura & Nishi, 1995). Calcium carbonate and chitin in the crab shells are found to be the most effective materials in removing heavy metals, especially copper and lead. This is because calcium carbonate forms strong copper-carbonate and lead-carbonate bonds when reacting with copper and lead, while chitin acts as an adsorbent for precipitation in the presence of those heavy metals (Lee *et al*., 1998). Vijayaraghavan *et al*. (2006) utilized crab shells for the biosorption of copper (II) and cobalt (II) in their study. Based on their study, crab shells have been proven to be the best adsorbent with the highest efficient removal of heavy metal. The biosorbent, which is around 0.767 mm in size with a dosage of 5g/L and an initial pH solution of 6, adsorbed heavy metals up to 90% from solution, i.e. 30% higher than other adsorbents. In addition, the biosorption capacity depends on the pH solution, for which the maximum copper (II) and cobalt (II) uptakes were 243.9 mg/g and 322.6 mg/g, respectively, and the results were observed at pH6. Another study by An *et al*. (2001) revealed that crab shells had the highest removal efficiency among other adsorbents such as zeolite, powdered activated carbon and granular activated carbon. This makes crab shells the most effective and applicable adsorbents for metal removing treatment, which combined with their abundance in the environment as waste and its cost effectiveness, making them a reasonable adsorbent in larger treatments such as industrial treatment plants.

Heavy metals exist naturally in the environment in different orders such as sedimentary, magmatic, metamorphic rocks, weathering and soil structure. Anthropogenic activities, mainly from the human activities also contribute to their existence, which usually originate from industrial and agricultural sources (Bradl, 2005). Copper and its compounds, which are very complex and involved in many metabolic processes in living organisms, are ubiquitously distributed in the environment and consist of the redox potential of Cu(I) or Cu(II) (Scheinberg, 1991). Copper is generally known as a non-toxic element in lower concentrations, but it may have an adverse effect on health (Bisone *et al.*, 2012) such as nausea (Ali & Aboul-Enein, 2006) when in higher concentrations.

Cadmium, which is a volatile element, is not essential for plants, animals or human beings. High doses of cadmium can have toxic effects on plants, animals and human beings. The target of cadmium concentration in the human organs is the kidney. In more specific, cadmium is considered toxic when its concentration reaches15 mg/L or higher (Stoeppler, 1991).

The shell of *S. serrata*, a waste material disposed of by the fishing industry, could be a viable and cost effective adsorbent for the removal of heavy metals. It is cheap and abundantly available in the environment, making it the most suitable material to remove copper and cadmium from an aqueous solution. The present study was conducted to determine the *S. serrata* shells' ability to remove copper (Cu) and cadmium (Cd) from an aqueous solution based on the removal rate of the heavy metal. The efficiency of removal was also calculated to determine the most suitable application for the removal of heavy metal using *S. serrata* shells as the biosorbent. Lastly, the chemical speciation of each heavy metal was modelled using hydrochemical software known as PHREEQC.

MATERIALS AND METHODS

Sorbent Collection and Preparation

The sorbent used throughout this study was crab (*S. serrata*) shells obtained from the local wet market. The crab *S. serrata* belongs to the Phylum Arthropode, Class Malacostrace and Family Portunidae. *S. serrata* (mud crab) is easily and abundantly found in the estuaries or mangrove areas in Malaysia and other developing countries in Asia such as Indonesia, Thailand, Vietnam and Sri Lanka (Overton á *et al*., 1997). The smallest *S. serrata* is found to have a shell width of 11.0 cm, and it can reach up to 24.0 cm in size (Jirapunpipat *et al.,* 2008; Overton á *et al*., 1997), depending on the surrounding conditions. For this study, mature *S. serrata* (in the range size of 8.6-9.2 cm) (Ikhwanuddin *et al.,* 2011) of similar shapes were selected and transferred to the laboratory for further analysis. The shells were steamed at 120˚C to separate the crabmeat and then thoroughly rinsed in tap water, followed by deionized water (Milli Q; 18 MΩ c/m). Subsequently, the shells were dried in the oven overnight at 30^{o}C to make sure they were totally dry. There was no special treatment applied at this step of the experiment. After they had been totally dried and no weight changes were observed, the shells of *S. serrata* were pulverized (Juang *et al.*, 1999) into fragment in the sizes between 2.00-2.50 mm and stored at room temperature (Evan *et al.*, 2008) until further analysis.

Sorption Experiment

All the sorption experiments were carried out in batch experiments with different initial concentrations of Cu and Cd solutions (1, 5, 10, 15 and 20 mg/L), representing the low range (<100 mg/L) of heavy metal concentrations (Liu *et al.*, 2009). All the solutions used in this experiment were analytical grade (Merck, India) reagent cadmium standard solution and copper standard solution (for Flame Atomic Absorption Spectrometry). The batch experiments were done in triplicate by using an Erlenmeyer flask containing 100ml solution for each metal. The laboratory apparatus used in this experiment were acid-washed by soaking them overnight in 5% (v/v) nitric acid (HNO_3) before rinsing thoroughly with distilled water. This procedure was

done to ensure that any available contaminants, ions and trace elements of cleaning reagent were removed before the experiments (APHA, 2005). With 100ml heavy metal in each flask, the pH of the solutions was adjusted to pH6 (±0.1) by adding a significant amount of 1.0M NaOH or 1.0M HCl. This pH value was chosen as the optimal pH solution because the uptake of heavy metals could be increased with the increase in pH, in which the maximum pH needed for biosorption ranged from 4.5 to 6.0 (±0.1) (Vijayaraghavan *et al*., 2006). Each flask was added 0.1g of crushed crab shells before sealing with a piece of parafilm to avoid any external contamination. Immediately afterwards, all the flasks were placed in an incubator rotary shaker and the temperature was set at 30°C, with a shaking speed of 150rpm. Samples of the heavy metal solution (3ml) were drawn out at every 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, and 1.75 h to obtain a precise breakthrough curve, which was followed by an interval time of 2 h up to 6 h of contact time. Different flasks indicate a different interval time. The use of the incubator rotary shaker was to ensure the crushed crab shells could mix well with the heavy metal solutions in the flasks. The samples were immediately centrifuged for 10 minutes in a centrifuge at 3000 rpm. After 10 minutes, the supernatant sample was pipetted out into another centrifuge tube and preserved until heavy metal analysis. Preservation was undertaken according to APHA (2005), in which the samples were kept in the refrigerator at a temperature of 4°C. The selected heavy metals were analyzed using Flame Atomic Absorption Spectrometry (FAAS) (Perkin Elmer AAS 3300). The accuracy of the FAAS performance was assessed based on the external standards, which had been prepared by diluting the specific stock standards of Cd and Cu (Merck, Germany) into a series of concentrations with the same mixture used for the sample pre-treatment. The recovery rates calculated for Cd and Cu ranged from 100.07% to 101.88% (±0.5%), depending on the elements analyzed (Table 1).

TABLE 1
Determination of the recovery percentage of metals

Element	Concentration (mg/)	Recovery (%)	Working Curve (r)	Detection Limit Ranges
Cd	1	100.29	1.0000	1.0-10.0
	5	101.14	1.0000	
	10	100.76	1.0000	
	15	100.27	0.9998	
	20	101.38	0.9999	
Cu	1	100.07	1.0000	1.0-10.0
	5	100.15	1.0000	
	10	101.88	0.9997	
	15	101.03	1.0000	
	20	101.67	0.9999	

Modelling Using PHREEQC

Data from the removal experiments were calculated using the hydrochemical model PHREEQC. PHREEQC is a useful tool for simulating a variety of reactions and processes in natural waters

or laboratory experiments (Parkhurst & Appelo, 2005). In this study, PHREEQC was applied to obtain the chemical speciation of Cu and Cd during the experiments (Parkhurst & Appelo, 2005). PHREEQC performs speciation calculations based on the law of mass action, which states that for a reaction of the generalized type:

$$aA + bB \leftrightarrow cC + dD \quad \text{(Eq. 1)}$$

Whereas, the distribution at thermodynamic equilibrium of the species at the left and right side of the reaction is given by:

$$K = \frac{[C]^c \cdot [D]^d}{[A]^a \cdot [B]^b} \quad \text{(Eq. 2)}$$

Where, K = equilibrium constant and the bracketed quantities denote activities or effective concentrations. In more specific, a, b, c and d = the number of moles of the reactants A, B, and the end products C, and D, respectively, for the given reaction of the mass-action law. In particular, the term K is defined as the solubility product of dissolution/ precipitation of minerals.

Data Analysis

The raw data obtained from the experiments were analyzed for the statistical analysis using the Statistical Package for Social Science (SPSS version 17) software. The statistical analysis was done for the descriptive statistics, one-way ANOVA and correlation analyses based on the heavy metals and time consumption of the treatment. The variances of the treated heavy metal solutions can be described using this statistical analysis in terms of variability.

The removal rate of heavy metals, R (%) in each interval time was calculated based on the following equation:

$$R = \left[\frac{C_\beta - C_\alpha}{t}\right] \quad \text{(Eq. 3)}$$

where C_β indicates the concentration before that interval time (mg/L) and C_α indicates the concentration at the time interval of contact time (mg/L). This calculation was performed as a percentage to show the removal amount of heavy metal in respective time. The higher it is, the better the performance of heavy metal removal will be within the time.

In order to obtain the efficiency of heavy metal removal, E (%) was calculated based on the percentage of efficient removal after the treatment, as shown in the equation below:

$$E = \left[\frac{C_o C_f}{C_o} \times 100\right] \quad \text{(Eq. 4)}$$

C_o refers to the initial heavy metal concentrations (mg/L) and the C_f refers to the final heavy metal concentrations (mg/L) after the treatment.

RESULTS AND DISCUSSION

Sorption Experiment

The results for the Cu removal are shown in Figure 1, with a rapid increase of cadmium removal at the initial stages of the 2-h treatment. Specifically, the Cu removal within 2 h of adsorption was between 38.9 and 90.23% (1-20 mg/L) before it slowed down after a few hours, depending on the initial Cu concentration used. From the graph of the remaining concentration of Cu, the treatment using 1 mg/L Cu at the beginning of experiments clearly resulted in the best Cu removal of 0.416 ± 0.010 mg/L within 0.25 h or 15 minutes of the treatment. In comparison, other concentrations only achieved between 10.2% and 30.9% of Cu removal, while Cu in 1 mg/L had already achieved 58.4% of Cu removal. However, in the middle of the treatment (3h), the removal trend shifted to the 5 mg/L initial concentration, which had previously been dominated by 1 mg/L. This trend continued until the end of the six hour treatment, whereby 5 mg/L gave the highest Cu removal of 94.7% (Table 2). Finally, the remaining concentration of Cu in the solution was 0.266 ± 0.024 mg/L, with a total concentration of 4.774 ± 0.024 mg/L being successfully removed by using the crushed crab shells.

In the 10 mg/L Cu initial concentration, only 6.35 ± 0.072 mg/L of Cu metal was removed after four hours of treatment using the crushed crab shells. However, by the end of the experiment (6 hours), the remaining concentration was slightly increased to 3.93 ± 0.067 mg/L (see Figure 1). This indicates that the desorption mechanism of Cu occurred using the crushed crab shells. As for the 15 and 20 mg/L Cu treatments, 10.61 ± 0.053 and 16.38 ± 0.059 mg/L were removed, respectively, by the use of the crab shells by the end of the experiment (6 hours).

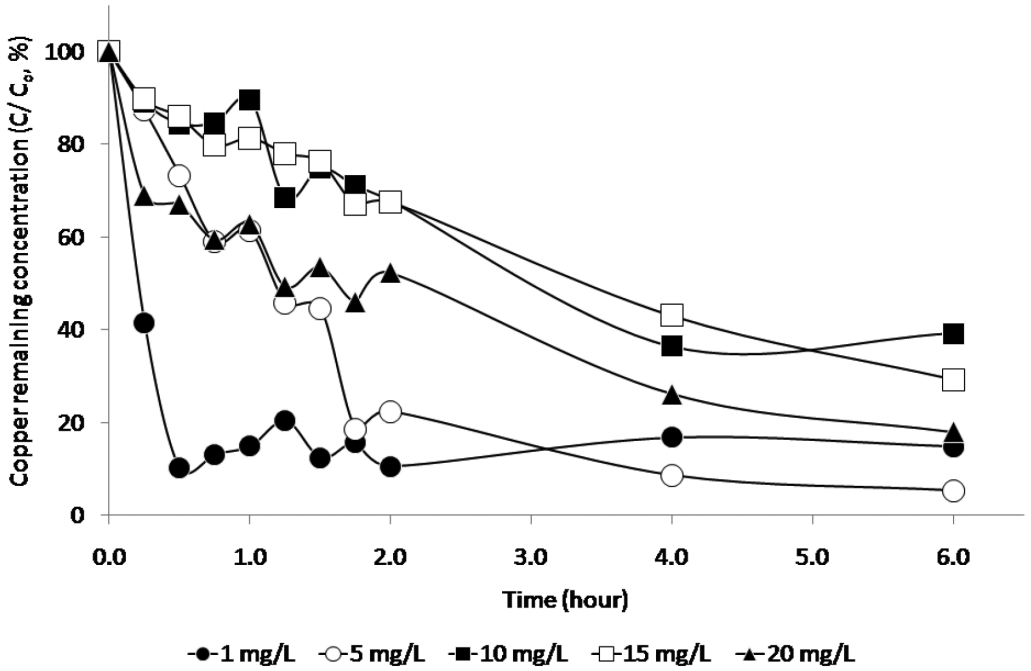


Fig.1. Copper remaining concentration (C/ C_o, %) after six hours of treatment using crab shells

Both the Cu solutions, in the concentrations of 1 mg/L and 5 mg/L, were found to be in the equilibrium state after 0.5 and 2 hours, respectively. For the Cu solutions in the concentrations of 10, 15 and 20 mg/L, however, no equilibrium state was observed, even after the experiments had been carried out for 6 hours. Similarly, it is suggested that the time for the attainment of equilibrium increased with the concentration of metals. These findings are parallel with the study conducted by Vijayaraghavan *et al*. (2006), who found a significant change in the initial Cu concentration ranging from 50 to 200 mg/L. Eventually, the metal amount being adsorbed was increased from 75.4 to 197.7 mg/L. Therefore, the equilibrium state for the higher concentrations of Cu (10, 15 and 20 mg/L) requires a longer time to be achieved.

The maximum sorption capacity was recorded for Cu removal with an initial concentration of 5 mg/L, with 94.7% of Cu being removed after 6 h of treatment. The 6-h treatment might seem inadequate for a complete (100%) removal, however, it showed that it could still achieve >90% removal within that time. Having shown its ability to remove heavy metals, the next factor to consider is the time required. The faster it can remove heavy metals, the greater chance it has of being an efficient sorbent.

The results of the Cd removal experiments are shown in Figure 2. In the Cd removal treatment, the variations of Cd concentration were also obtained. The initial cadmium removal is rapid throughout the treatment but it became slow after a few hours, depending on the initial Cd concentration being used. The slow removal rate is likely to be due to the saturation of Cd concentration into the biosorbent. Since it attached more and more Cd to its surface, it became saturated and therefore, the remaining Cd can only be adsorbed in minimal amounts. Similarly, Evans *et al*. (2002) found that the longer time for Cd uptake resulted in a slower Cd uptake rate in respect of crab shells as a biosorbent.

In the first 2 hours, all the treatments showed that they had achieved the equilibrium state, while the removal rate after that time interval showed a similar trend. When identical trend was repeated for Cu removal, the 1 mg/L initial metal concentration always gave the best metal removal in the first 15 minutes of the treatment. This phenomenon continued until the end of the treatment through which 1 mg/L initial Cd concentration achieved the highest Cd removal rate of 85.1% or a total of 0.851 ± 0.004 mg/L of Cd was removed by the end of experiment.

From the experiment, 5 mg/L of Cd concentration showed the highest desorption (0.69 ± 0.011 mg/L) from 0.25 h to 1.25 h of the treatment. This was repeated for the next 15 minutes but in a lower amount (0.22 ± 0.103 mg/L). Higher initial Cd concentrations (10, 15 and 20 mg/L) were only able to remove Cd in amounts ranging from 49.3 to 64.7%, while the lowest percentage was detected for the 20 mg/L initial concentration. This might be due to the saturation of the solution, which required a longer time to gradually remove Cd (Evans *et al*., 2002).

In this study, the ability of all the initial Cd concentrations and their removal rates after the treatment using the crushed crab shells to remove Cd were assessed. Figure 2 shows the ability of these initial Cd concentrations since the decreasing amount of Cd showed a downward trend. To support this result, the results obtained from the statistical analysis using the one-way ANOVA showed a significant difference ($p< 0.05$) after the use of crushed crab shells. The efficiency of the crushed crab shells to remove Cd in aqueous solution is shown in Table 2.

TABLE 2
Results of Cu and Cd removal efficiency from different initial concentrations (mg/L)

Initial Heavy Metal Concentrations (%)	Removal Efficiency (%)	
	Cu	Cd
1	85.1	85.1
5	94.7	80.2
10	60.7	61.7
15	70.7	64.7
20	81.9	49.3

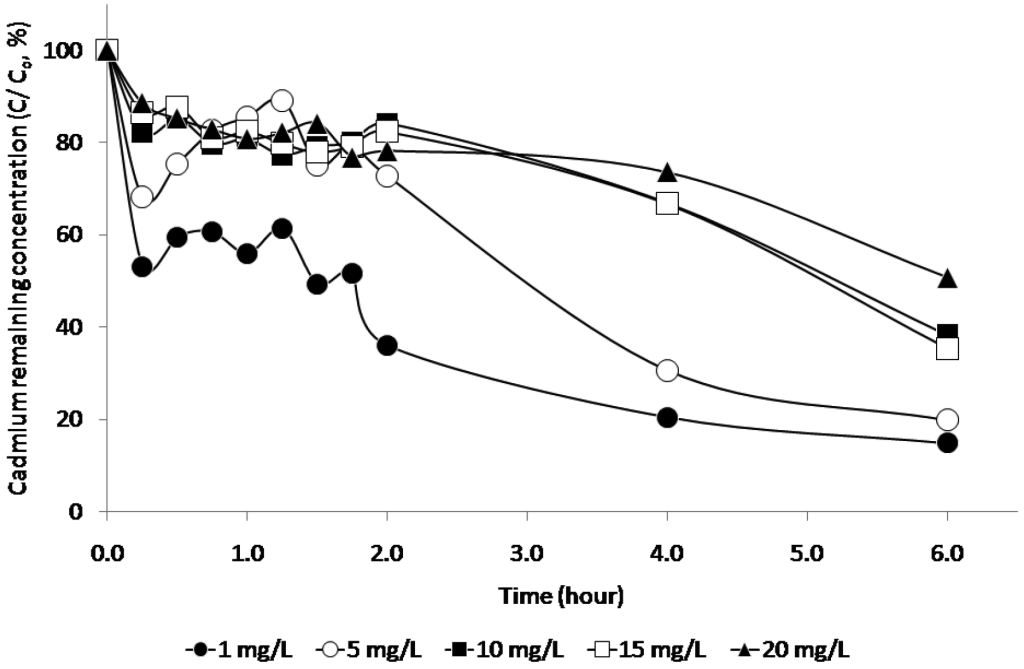


Fig.2. Cadmium remaining concentration (C/ C_o, %) after six hours of treatment using crab shells

Additionally, this study also showed that with increasing initial concentration, the removal efficiency by the crushed crab shells also decreased (see Figure 3). A strong correlation (r= -0.950, p< 0.05) between the initial concentration solution and the removal efficiency for Cd is shown in Figure 3. In regards to the variation in the initial Cd concentration, the more concentrated the solution, the lower the tendency of the Cd metal to bind to the surface of the crushed crab shells. This might be due to the equilibrium state of the surface environment in which the mechanism happened (Ismail & Aris, 2012; Reddad *et al*., 2002). Therefore, it is suggested that future experiments use the lowest initial concentration (1 mg/L) of Cd because of its cost effectiveness and ability to produce the most efficient result. Thus, it is suggested that the trend for total Cd adsorbed by the crushed crab shells follow a decreasing trend of 1 mg/L > 5 mg/L > 15 mg/L > 10 mg/L > 20 mg/L (r= -0.950, p< 0.05).

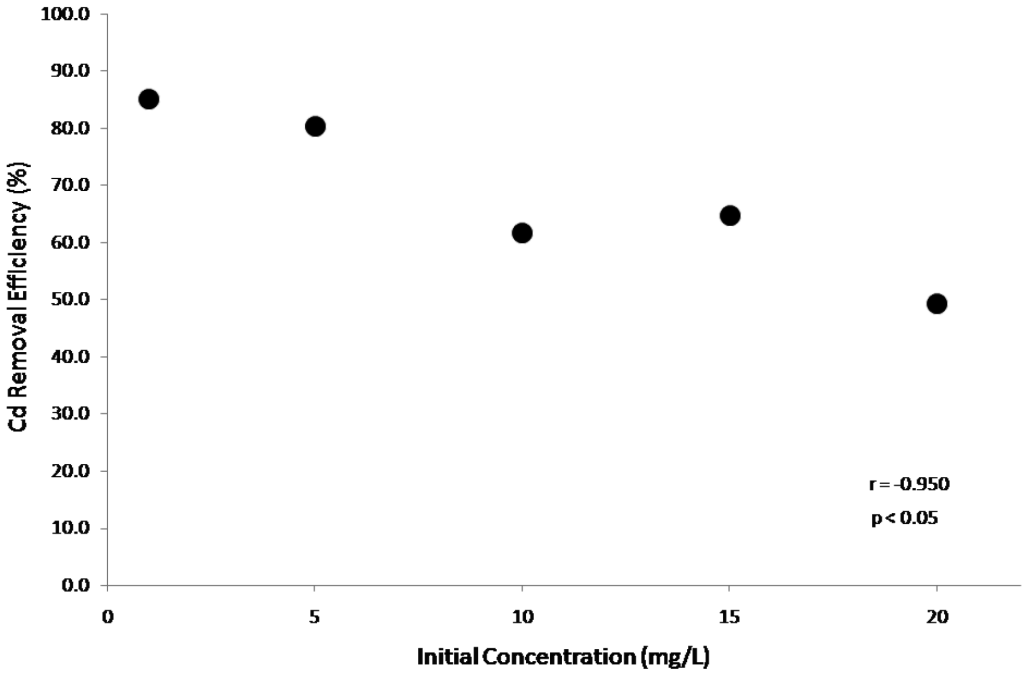


Fig.3. The relationship between total Cd removed with the initial concentration for the treatment

However, the removal experiments for Cu showed a weak correlation (r= - 0.363, p< 0.05) between the removal efficiency and the initial metal concentrations (Figure 4). Even though the highest amount of Cu was removed at the end of the experiments using 20 mg/L Cu (16.38 ± 0.159 mg/L out of 20 mg/L), it was not as efficient as using 1 mg/L (0.851 ± 1.021 mg/L out of 1 mg/L) of Cu in the initial stage. Based on the experiments, it is suggested to choose 5 mg/L of initial Cu concentration in the future when attempting to treat Cu from industrial wastewater effluent. This is particularly due to the effective and efficient removal rate of 94.7% attained when using the suggested initial Cu concentration (Table 1). Despite the concentration amounts being used, another crucial aim is to remove as much Cu as possible in a short time period. This has been proven by the statistical analysis, which showed no significant difference (p = 0.391) between the 1 and 5 mg/L solutions.

Comparing the removal efficiency rates of Cu and Cd, Figure 5 illustrates that the fastest removal rate achieved for Cu was within two hours of treatment, with almost 92.4% of Cu being removed, and only 0.076 ± 0.002 mg/L Cu remained in the solution. The removal rate of Cu was more rapid than Cd (p<0.05). An experiment using 1 mg/L initial metal concentration was chosen as a comparison since this experiment yielded the same removal efficiency of 85.1% by the end of the experiment. As a more abundant element in the environment compared to Cd (Ellingsen *et al*., 2007; Nordberg *et al*., 2007), Cu seems to be adsorbed faster. In the Cu experiments, however, the negative amounts indicated that desorption most likely occurred because the surface area of the crab shells had reached their capacity for adsorption.

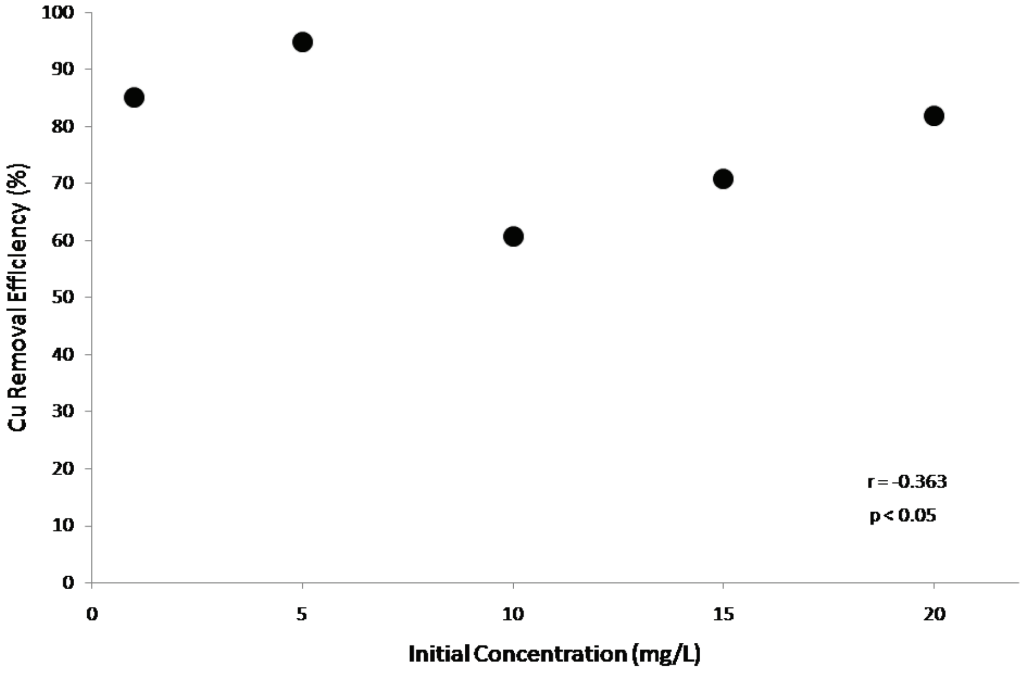


Fig.4. The relationship between total Cu removed and the initial concentration for the treatment

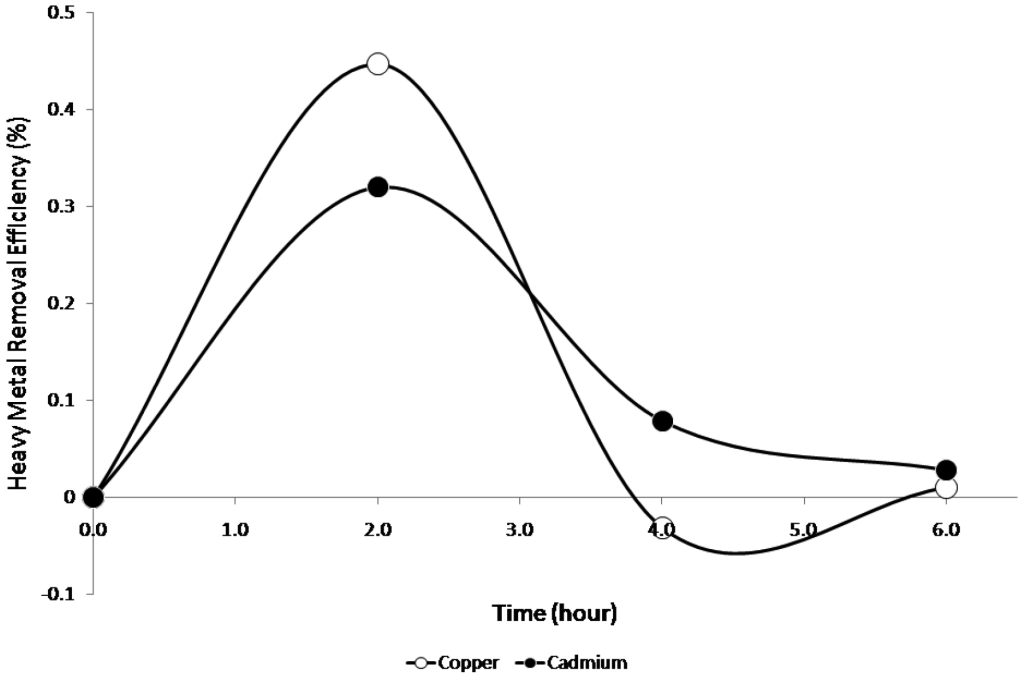


Fig.5. Variation of the heavy metals removal efficiency rates from beginning until the end of the treatment using the initial 1 mg/L metals solution

PHREEQC Modelling

The data obtained were also modelled using hydrochemical software with PHREEQC (Parkhurst & Appelo, 2005). Basically, PHREEQC was used to identify the dominant chemical speciation of heavy metals during the experiment, based on the results obtained. This software determined the most abundant species during the treatment, and came out with the amount of species prediction. However, to obtain the exact amount of speciation, further experiments must be done. In this experiment, PHREEQC was used to calculate the dominant metal species in the solution during the experiment. The Cu experiments (see Figure 6) suggest that divalent Cu (Cu^{2+}) seems to be the most abundant species in the aqueous solution. In fact, this metal species is the most stable form that exists in normal atmospheric and pressure conditions (Appelo & Postma, 2005). A similar observation was encountered for the Cd experiment (Figure 7), in which the divalent Cd (Cd^{2+}) also tended to be the most abundant in that solution. Following that, cadmium hydroxide ($Cd(OH)^{+}$) was the second most abundant, with approximately 44.01% of $Cd(OH)^{+}$.

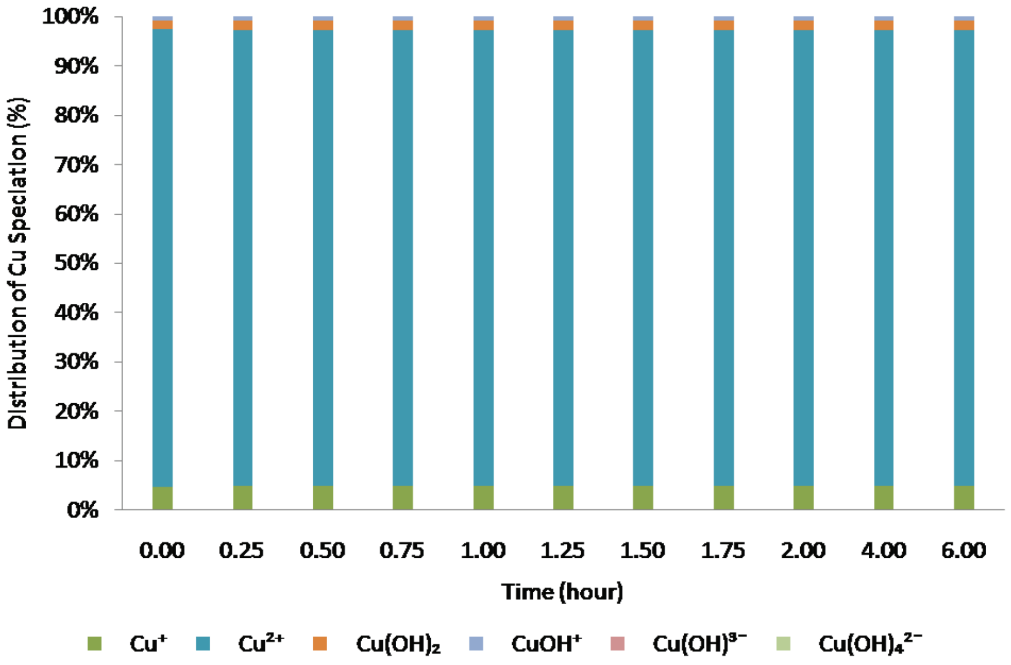


Fig.6. Distribution of Cu speciation throughout the treatment of six hours using initial mg/L Cu

CONCLUSION

In summary, the crushed shells of *S. serrata* have a high sorption capacity for Cu and Cd. Since this study mainly focused on the ability of crushed crab shells to remove Cu and Cd, it was proven that this biosorbent could remove up to 94.7 and 85.1% Cu and Cd, respectively. From the calculated removal efficiencies, *S. serrata* was found to be a more effective adsorbent in removing Cu than Cd for both 5 and 1 mg/L. The application of *S. serrata* in low amounts

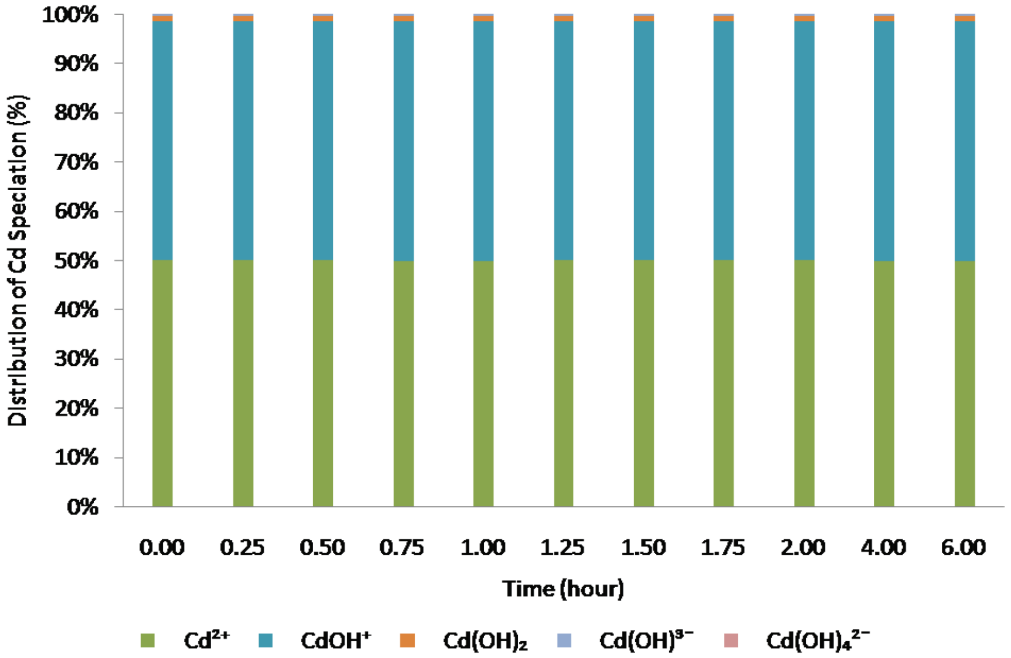


Fig.7. Distribution of Cd speciation throughout the treatment of six hours using initial 1 mg/L Cd

that eventually achieved high removal of heavy metals for Cu and Cd suggests that *S. serrata* could be a very effective and efficient biosorbent. With the constant pH of 6 and the temperature maintained at 25°C, this experiment perfectly suited the wastewater effluent characteristics from mining or metal refining industries. Coupled with its low cost, easy availability (since it is a waste product from the fishing industry), as well as its abundance in the environment, crab shells are undeniably suitable adsorbents for heavy metals from industrial wastewater.

REFERENCES

Adeogun, A. L., Kareem, S. O., Durosanya, J. B., & Balogun E. S. (2012). Kinetics and equilibrium parameters of biosorption and bioaccumulation of lead ions from aqueous solutions by *Trichoderma Longibrachiatum*. *Journal of Microbiology, Biotechnology and Food Sciences, 1*(5), 1221-1234.

Ali, I., & Aboul-Enein, H. Y. (2006). *Instrumental Methods in Metal Ion Speciation*. London: Taylor & Francis, pp. 376.

Alkan, M., Demirbaş, **Ö.,** & Doğan, M. (2004). Removal of Acid Yellow 49 from Aqueous Solution by Adsorption.*Fresenius Environmental Bulletin, 1311*(a), 1112-1121.

An, H. K., Park, B. Y., & Kim D. S. (2001). Crab shell for the removal of heavy metals from aqueous solution. *Water Research, 35*(15), 3551-3556.

APHA. (2005). *Standard Methods for the Examination of Water and Wastewater* (21st Edn.). Washinton: American Water Works Association, Water Environment Federation, pp. 1368.

Appelo C. A. J., & Postma D. (2005). *Geochemistry, Groundwater and Pollution* (2nd Edn,). Balkema, Rotterdam, pp. 649.

Bisone, S., Blais, J. F., Drogui, P., & Mercier, G. (2012). Toxic metal removal from polluted soil by acid extraction. *Water Air Soil Pollution, 223*, 3739-3755.

Bradl, H. B. (2005). *Heavy Metal in the Environment: Origin, Interaction and Remediation*. London: Elsevier Academic Press, pp. 268.

Castillo-Zacarías, C. J., Suárez-Herrera, M. A., Garza-González, M. T., Sánchez-González, M. N., & López-Chuken, U. J. (2011). Biosorption of metals by phenol-resistant bacteria isolated from contaminated industrial effluents. *Journal of Microbiology Research, 5*(18), 2627-2631.

Demirbas, A. (2008). Heavy metal adsorption onto agro-based waste materials: A review. *Hazardous Materials, 157*(2-3), 220-229.

Ellingsen, D. G., Hom, N., & Aaseth J. (2007). Copper. In *Handbook of Toxicology of Metals* (p. 975). New York: VCH Publishers.

Evans, J. R., Davids, W. G., MacRae, J. D., & Amirbahman, A. (2002). Kinetics of Cadmium uptake by chitosan-based crab shells.*Water Research, 36*(13), 3219-3226.

Ikhwanuddin, M., Azmie G., Juariah H. M., Zakaria M. Z., & Ambak M. A. (2011). Biological information and population features of mud crab, genus *Scylla* from mangrove areas of Sarawak, Malaysia. *Fisheries Research*, *108*(2-3), 299-306.

Ismail, F. A., & Aris, A. Z. (2013). Experimental determination of Cd2+ adsorption mechanism on low-cost biological waste. *Frontiers of Environmental Science and Engineering.* DOI 10.1007/s11783-013-0488-1.

Jirapunpipat, K., Aungtonya, C., & Watanabe, S. (2008). Morphological study and application of multivariate analysis for the mud crab genus *Scylla* Klong Ngao Mangrove, Ranong Province, Thailand. *Phuket mar. Boil. Cent. Res. Bull., 69*, 7-24.

Juang, R. S., Wu, F. C., & Tseng, R. L. (1999). Adsorption removal of Copper(II) using chitosan from simulated rinse solutions containing chelating agents. *Wat. Res., 33*(10), 2403-2409.

Kim, D. S. (2003). The removal by crab shell of mixed heavy metal ions in aqueous solution. *Bioresource Technology*, *87*(3), 355-357**.**

Leborans, G. F., & Novillo, A. (1996). Toxicity and bioaccumulation of Cadmium in *Olisthodiscus Luteus* (Raphidophyceae). *Wat. Res., 30*(1), 57-62.

Lee, M. Y., Shin, H. J., Lee, S. H., Park, J. M., & Yang J. W. (1998). Removal of Lead in a Fixed-Bed Column Packed with Active Carbon and Crab Shell. *Separation Science and Technology, 33*(7), 1043-1056.

Nordberg, G. F., Nogawa K., Nordberg, M., & Friberg, A. L. T (2007). Cadmium. In *Handbook of Toxicology of Metals* (p. 975). New York: VCH Publishers.

Overton, J. L., Macintosh, D. J., & Thorpe, R. S. (1997). Multivariete analysis of the mud crab Scylla serrata (Brachyura: Portunidae) from four locations in Southeast Asia. *Marine Biology*, *128*, 55-62.

Park, D. P., Lim, S. R., Yun, Y. S., & Park, J. M. (2008). Deveopment of a new Cr(VI)-biosorbent from agricultural biowaste. *Bioresource Technology*, *99*, 8810-8818.

Parkhurst, D. L., & Appelo, C. A. J. (2005). User's Guide to PHREEQC Version 2-A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations. *U.S. Geological Survey, Water Resource Investigation Report.* pp. 99-4259.

Reddad, Z., Gerente, C., Andres, Y., & Cloires, P. L. (2002). Adsorption of several metal ions onto a low-cost biosorbent: kinetic and equilibrium studies. *Environmental Science Technology, 36*(9), 2067-2073.

Scheinberg, H. (1991). Copper. In *Metals and Their Compounds in the Environment, Occurrence, Analysis, and Biological Relevance* (pp. 891-904). New York: VCH Publishers.

Schiewer, S., & Volesky, B. (2000). Biosorption process for Heavy metal removal. In *Environmental Microbe-Metal Interactions* (pp. 329-362). Washington, D.C.: ASM Press, Washington.

Shareef, K. M. (2009). Sorbents for contaminants uptake from aqueous solutions. Part I: Heavy metals. *World Journal of Agricultural Sciences, 5*(S), 819-831.

Stoeppler, M. (1991). Cadmium. In *Metals and Their Compounds in the Environment, Occurrence, Analysis, and Biological Relevance* (pp. 803-835). New York: VCH Publishers.

Tokura, S., & Nishi, N. (1995). Specification and Characterizationof Chitin and Chitosan. In *Chitin and Chitosan the Versatile Environmentally Friendly Modern Materials* (pp. 67-86). Bangi: Penerbit Universiti Kebangsaan Malaysia.

Vijayaraghavan, K., Palanivelu, K., & Velan, M. (2006). Biosorption of copper (II) and cobalt (II) from aqueous solutions by crab shell particles. *Bioresource Technology, 97*(12), 1411-1499.

Yadanaparthi, S. K. R., Graybill, D., & Wandruszka, R. V. (2009). Review: Adsorbents for the removal of arsenic, cadmium, and lead from contaminated waters. *Journal of Hazardous Materials*, *171*, 1-15.

Yazdani, M., Yap, C. K., Abdullah, F., & Tan, S. G. (2009). Trichoderma atroviride as a bioremediator of Cu pollution: An in vitro study. *Toxicological and Environmental Chemistry*, *91*(7), 1305-1314.

Pertanika J. Sci. & Technol. 22 (2): 567 - 575 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Bayesian Network Classification of Gastrointestinal Bleeding

Nazziwa Aisha[1]*, Mohd Bakri Adam[2], Shamarina Shohaimi[1] and Aida Mustapha[3]

[1]Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

[2]Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

[3]Department of Computer Science, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

The source of gastrointestinal bleeding (GIB) remains uncertain in patients presenting without hematemesis. This paper aims at studying the accuracy, specificity and sensitivity of the Naive Bayesian Classifier (NBC) in identifying the source of GIB in the absence of hematemesis. Data of 325 patients admitted via the emergency department (ED) for GIB without hematemesis and who underwent confirmatory testing were analysed. Six attributes related to demography and their presenting signs were chosen. NBC was used to calculate the conditional probability of an individual being assigned to Upper Gastrointestinal bleeding (UGIB) or Lower Gastrointestinal bleeding (LGIB). High classification accuracy (87.3 %), specificity (0.85) and sensitivity (0.88) were achieved. NBC is a useful tool to support the identification of the source of gastrointestinal bleeding in patients without hematemesis.

Keywords: Bayesian network classifiers, Emergency department, hematemesis, upper gastrointestinal bleeding, Naive Bayes classifier, lower gastrointestinal bleeding, data mining

Article history:
Received: 31 December 2012
Accepted: 27 May 2013

Email addresses:
Nazziwa Aisha (aishanazziwa@yahoo.ca),
Mohd Bakri Adam (bakri@upm.edu.my),
Shamarina Shohaimi (sshohaimi@yahoo.com),
Aida Mustapha (aida_m@upm.edu.my)
*Corresponding Author

INTRODUCTION

Acute gastrointestinal bleeding (GIB) is considered a potentially life threatening condition that requires prompt assessment and aggressive medical management. Older patients, especially those over age 50 years, make up a rising proportion of patients with GIB, and mortality rates in this group have remained relatively high (Vreeburg *et al*., 1997; Crooks *et al*., 2011). Further reductions in mortality will require the introduction of

ISSN: 0128-7702

novel methods to help with identification of the cohort requiring aggressive resuscitation and endoscopic intervention to prevent complications and death from ongoing bleeding (Dhahab & Barkun, 2012). Delays in intervention often result from failure to adequately recognise the source and severity of the bleeding. GIB is described by the anatomical area that is bleeding and is classified as either upper gastrointestinal bleeding (UGIB) or lower gastrointestinal bleeding (LGIB). The anatomic landmark is the ligament of Treitz which extends from the small intestine at the duodenojeunal junction. When the bleeding is proximal to the ligament, it is classified as UGIB and if it is distal to the ligament, it is classified as LGIB.

In the emergency department (ED), when patients show signs of hematemesis i.e. vomiting of blood, (Society and Committee, 2002), it is clear that the source of bleeding is from the upper gastrointestinal tract. However, when there is no hematemesis the source of bleeding is unclear. The source of bleed determines management of the bleeding. It will determine the type of physician to be assigned and the timing (Barkun *et al*., 2010). The primary diagnostic tool of choice in patients with acute upper gastrointestinal bleeding is esophagogastroduodenoscopy (EGD) (Kovacs *et al*., 2002; Manning-Dimmitt *et al*., 2005) and for lower gastrointestinal bleeding it is colonoscopy (Jensen *et al*., 1997). These tools aid diagnosis, allow treatments that can stop bleeding to be delivered and yield information that help in prediction of outcomes (Albeldawi *et al*., 2010).

Although diagnosis of GIB is best done by a gastroenterologist, it is not always feasible because of resource, time and cost constraints (Gralnek & Dulai, 2004; Quirk *et al*., 1997). Without the gastroenterologist, physicians are left to determine the source of bleeding using symptoms and demographic data only. Of late, the use of nasogastric aspiration (NGA) has been advocated to localise bleeding (Witting *et al*., 2004; Anderson 7 Witting, 2010) but it is a painful procedure and it does not work well for patients without hematemesis. The predictors identified in clinical and epidemiological studies to predict UGIB in patients without hematemesis are that patients are aged below 50; the colour of the material passed through the rectum is black and the ratio of blood urea nitrogen to creatinine is 30 and above (Witting *et al*., 2006), while the factors for LGIB are hemodynamic instability (SBP=90mmHg, Heart rate > 100/min), a Hemoglobin level of 6g/dl and initial hematocrit of 35 % (Velayos *et al*., 2004; Parkes *et al*., 1993). Medications, especially those of NSAID, increase the risk of GIB, leading to hospital admission (Lanas *et al*., 2006). The incidence of LGIB is higher in men than in women, and patients with prior episodes of UGIB are more likely to bleed from the same lesion. Tarone *et al*. (2004) state that 60 % of patients with a history of UGIB bleed from the same lesion.

To assist the emergency department physician in diagnosing the patients more efficiently and effectively, mathematical models must be developed to identify the source of GIB. Classification models have the ability to identify the source of GIB, which is needed for intervention and to allow optimisation of care and healthcare resource allocation amongst patients with acute GIB (Chu *et al*., 2008). In this study, we use a naive Bayesian classifier (NBC) to predict the source of GIB. The graphical nature of NBC makes it easy for physicians to understand and use (Mittal and Kassim, 2007; Jensen, 1996). The NBC can be used to predict the source of bleeding using clinical and laboratory information available within a few hours of patient presentation.

BAYESIAN NETWORK CLASSIFIERS

Classification is a basic task in data analysis that constructs a function from labelled training data. The training data has both input and output objects. A supervised learning algorithm is used to construct a function called a classifier. When output is discrete we call it a classification algorithm and if continuous, a regression function. Bayesian network classifiers (BNC) are a group of generative classifiers that have performed well in many classification tasks. The NBC is a BNC that has a predictive performance which is competitive with state-of-the-art classifiers like C4.5 (Quinlan, 1993). This classifier has previously been used in other medical studies (Kazmierska & Malicki, 2008; Wei *et al*., 2011; Al-Aidaroos *et al*., 2012). The NB assumes conditional independence among the variables or attributes and learns conditional probabilities of each attribute, A_i given the class label, C. NBC predicts a new data point as the class with the highest posterior probability during classification by applying the Bayesian rule to compute the probability of $C_i (1 \leq i \leq k)$ given the particular attributes, as shown in the equation $\arg\max_{C_i} P(C_i) \prod_{j=1}^{n} P(A_j / C_i)$.

When performing a classification, the NB partitions the data sets according to their class label into sub datasets. Then for each sub data set labelled C_i a maximum likelihood (ML) estimator $P(A_j = a_{jk} | C_i)$ can be given by $\frac{n_k + mp}{n + m}$ where n = the number of training examples for which $C = C_i$, n_k = number of examples for which $C = C_i$ and $A = A_j$, p = a priori estimate for $P(A_j = a_{jk} | C_i)$ and m = the equivalent sample size.

METHODS

Data were collected from a retrospective cohort study of patients admitted through the ED for GI tract bleeding from unknown sources and followed until hospital discharge. This study has been described in detail previously (Witting *et al*., 2006). The study used logistic regression analysis to identify clinical variables that independently predict an UGIB source. A total of 325 patients were admitted through the emergency department for GIB and were followed until hospital discharge. Eligible patients were 17 years or older, had heavy bleeding, as indicated by bloody or hemoccult positive black stools, or hemoccult positive dark stools if NGA was performed in the ED, were admitted in hospital through the ED for a principal diagnosis of GI tract bleeding and had confirmatory diagnostic testing within 3 days after admission. The exclusion criteria were: hematemesis, ostomy, an obvious anorectal source, such as hemorrhoids and admission for GI tract bleeding within the previous month. The institutional review board at each participating hospital approved the protocol.

Assessment

The following information was recorded: sex, age, history of UGIB or LGIB, history of upper or lower GI cancer (yes/no), alcohol use (yes/no), tobacco use (yes/no), epigastric pain or tenderness, use of medication [prophylactic aspirin, non-steroidal anti-inflammatory drugs (NSAIDs), cirrhosis, steroids, warfarin iron, or bismuth] within the previous 2 weeks, colour of

blood in stools (black vs red), stool consistency (clots, tarry, diarrhea and not specified), initial resting and orthostatic vital signs and signs of cirrhosis. The colour and consistency of stools were based on the physician or patient's description. The diagnosed source of bleeding in a patient was obtained from the hospital discharge summary based on confirmatory testing such as EGD, colonoscopy, nuclear medicine scan, arteriography or surgery. In patients with unspecified gastrointestinal tract hemorrhage, classification was made based on a gastroenterologist's statement as to whether a finding indicated the site of hemorrhage.

Continuous variables were discretised e.g. age [50 and above (AboveEq50), below 50], Blood Urea Nitrogen to creatine ratio (BUN/CR) into [BUN/CR above or equal to 30 (AboveEq30), BUN/CR below 30 (below30)], Hematocrit into [above or equal to 30 (AboveEq30) and below 30 (below30)]. The discrete variables were consistence of stools (N=Not specified, T=Tarry, C=Clots, D=diarrhea), History of GIB (None, Lower=LGIB, Upper=UGIB) and colour of stools (red, black).

Methodology

Predictors of GIB as seen in the literature were collected from the data. The model was trained to predict the source of bleeding and a 10-fold cross validation was carried out to assess the accuracy of the classifier (Kohavi, 1995). In the 10-fold cross-validation, the original data were randomly partitioned into 10 sub-samples. Of the 10 sub-samples, a single sub-sample was retained as the validation data for testing the model, and the remaining 9 sub-samples were used as training data. The cross validation process is then repeated 10 times, with each of the 10 sub-samples used only once as the validation data. The results from the 10-fold cross validation were then averaged to produce a single estimation.

During training, all information in the training set including the source of bleeding was provided to the NBC. In the testing phase, each patient datum, apart from source of bleeding, was entered into the trained NBC to infer the probability of the disease. The predicted source of bleeding probability and the known source of bleeding were then analysed using prediction accuracy and area under ROC curve (Hanley *et al*., 1982). To evaluate the classifier, a standard approach to estimating the accuracy of NB was the prediction accuracy. Sensitivity analysis was done to identify the factors that influenced the source of bleeding. All calculations in this study were made with Naive Bayesian Classifier implemented in WEKA environment (Hall *et al*., 2009). WEKA is a collection of machine learning algorithms for solving real-world data mining problems. It is written in Java and runs on almost any platform and enables testing databases applying different artificial intelligence systems.

RESULT

A high classification accuracy of 87.3 % was achieved i.e. the results showed that out of a total of 325 cases, 284 patients had been accurately classified and 41 wrongly classified. Sensitivity of the system was 0.88 and specificity, 0.85. Fig.1 shows the initial probabilities in the naive Bayesian model used for predicting the source of GIB and their states e.g. the probability of

LGIB is 60.9 % and UGIB is 39.1 %. The probability of LGIB was higher than UGIB because in patients who show no signs of hematemesis, the most likely source of bleeding is LGIB. Since the factors in a naive Bayesian model are independent of each other, we can always include a number of factors even without feature selection. We included consistence and hematocrit to investigate the probabilities of the sources of bleeding, when only these factors were known. In general terms, the results indicate that source of bleeding is sensitive to a number of variables. The results of the sensitivity analysis showed two relatively high levels of mutual information for two variables: colour of stools at 33.2 % and history of gastrointestinal bleeding (HGIB) at 13.7 %, as shown by sensitivity analysis results captured in Table 1. Other factors are BUN/CR and hematocrit. The sensitivity analysis identified which data had a significant impact on the result, such that concentration could be given to finding accurate data for those items. The degree of sensitivity of one node (variable) to the class variable (source of bleeding) was shown by the mutual information (i.e. entropy reduction) while for continuous nodes it was shown by the variance reduction. All our variables were discretised; hence, we used the mutual information. The higher the mutual information, the greater the degree of sensitivity.

Application of the Model

With the naive Bayesian model, one is able to answer questions such as, "What is the probability of source of GIB given that age of a patient is below 50 years?" We can also predict the source of bleeding given any known symptoms. Because of maximum *a posteriori* (MAP) estimate in the NBC, the source with a higher posterior probability is the predicted source. If the information we know about the patient is that he passes black stools, the probability that he has UGIB rises to 76 % from the initial 39 % while the probability for LGIB falls to 23 %, and if he passes red stools, the probability is 87 % from the initial 60.9 %. This means that black stools is indicative of UGIB and red stools of LGIB. Figure 2 shows the posterior probability given only one piece of evidence that the patient has BUN/CR that is greater or equal to 30.

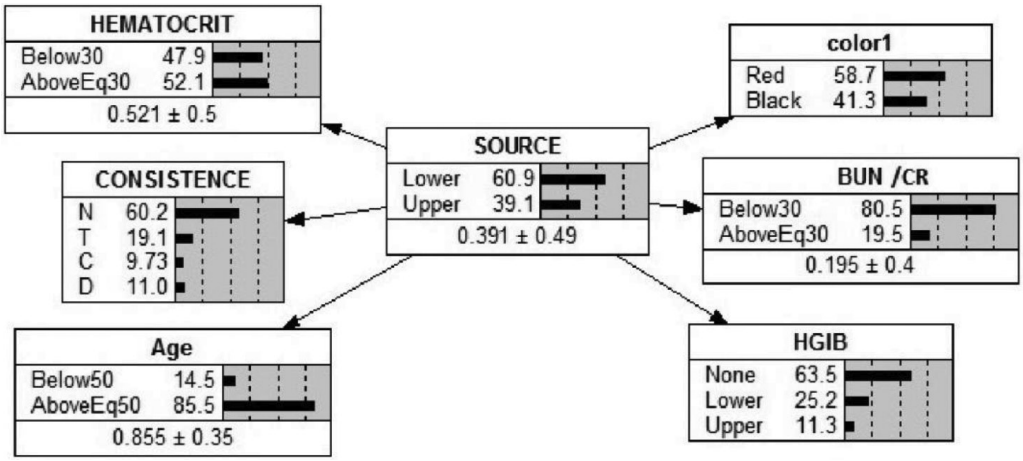
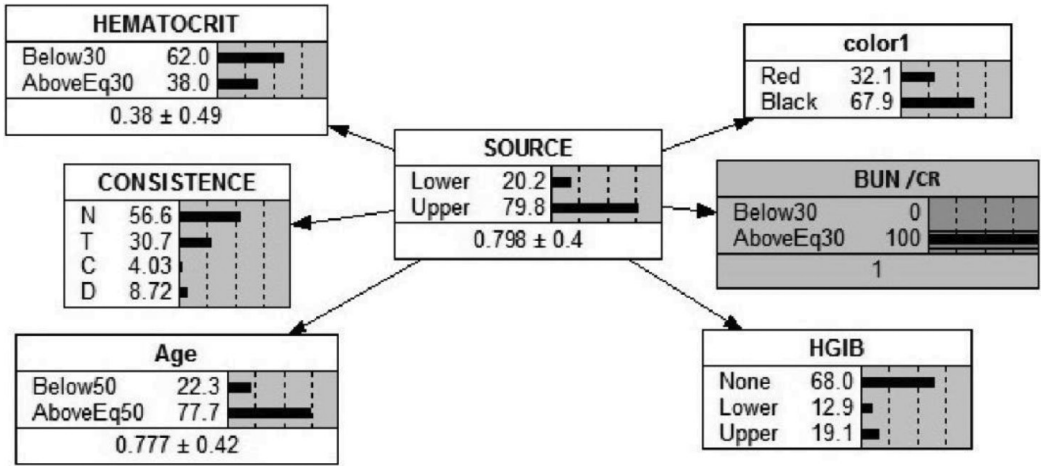


Fig.1: Naive Bayesian Classifier of Gastrointestinal Bleeding with Initial Probabilities

TABLE 1
Sensitivity of Source of Gastrointestinal Bleeding to Findings at Other Nodes

Node	Variance Reduction	Percent	Mutual Info	Percent	Variance of Beliefs
Source	0.23810	100	0.9653	100	0.238080
Colour1	0.09953	41.8	0.3207	33.2	0.099529
BUN/CR	0.04031	16.9	0.1226	12.7	0.040311
HGIB	0.03966	16.7	0.1322	13.7	0.039655
Consistency	0.03771	15.8	0.1240	12.9	0.037713
Hematocrit	0.02703	11.4	0.0835	8.65	0.027025
Age	0.01688	7.09	0.0502	5.20	0.016876



The figures in the boxes are the posterior probabilities when there is evidence that a patient has BUN/CR ≥ 30 (shown by the 100 % bar mark in the node).

Fig.2: Posterior Probability in Naive Bayesian Classifier Model When BUN/CR Is Greater or Equal to 30

We compare the original model without evidence as seen in Figure 1 with some evidence as seen in Figure 2. With evidence that BUN/CR ≥ 30, the probability that a patient has UGIB rises from 39.1 % in Figure 1 to 79.8 % in Fig.2, indicating that the patient has UGIB.

DISCUSSION

We aimed at constructing a diagnostic model that could combine simplicity of use and high performance accuracy to allow easy application. One of the most important possible limitations of NBC use is the assumption of independence of attributes. Although in practice this assumption is not quite true, in medical applications, the NBC has been shown to be effective and gives relatively good classification accuracy in comparison with other, more elaborate learning methods. Referring to the NBC, Kononenko (1993) states that, "Physicians found such explanations (using conditional probabilities) as natural and similar to their classification. They also summed up evidence for / against a diagnosis."

The NBC was, however, previously shown to be robust to obvious violations of this independence assumption (Domingo, 1997) and it yielded accurate classification models even when there were clear conditional dependencies. According to Rish (2001) the accuracy of the naive Bayesian model for zero-Bayes-risk problems is not directly correlated with the degree of feature dependencies measured as the class conditional mutual information between the features. Instead, a better predictor of naive Bayesian accuracy is the amount of information about the class that is lost because of the independence assumption.

In contrast with more complex models like the general Bayesian network, NBC allows for easy incorporation of additional attributes. The capability to add attributes in an easy way is a great benefit when working in fast changing fields like medicine, where new predictors may be identified. Adding new attributes or data unknown to the model may result in momentary deterioration of classification accuracy. Results improve again when the number of sample cases including new attributes increases. The NBC is also resistant to missing data i.e. with a few known attributes, one can still predict using the model e.g. if a patient is unconscious and HGIB and demographic factors like age cannot be obtained from him, a physician can still use any available data like hematocrit level to determine source of GIB. In clinical practice missing data is a common problem and a challenge to many research projects especially during classification (Quinlan, 1989). Classification results could offer valuable suggestions for the source of gastrointestinal bleeding in such difficult cases e.g. absence of hematemesis. Considering these conditions, NBC seems to be useful and deserves further study.

CONCLUSION

The trained naive Bayesian classifier to identify the source of gastrointestinal bleeding revealed an accuracy of 87.3 %, a specificity of 0.85 and a sensitivity of 0.88. These values show the classifier's potential and credibility in supporting physicians to identify the source of bleeding in patients with GIB. The results achieved are very encouraging and they support further development of NBC as a valuable tool for supporting everyday clinical decisions.

REFERENCES

Al-Aidaroos, K. M., Bakar, A. A., & Othman, .Z. (2012). Medical data classification with Naïve Bayes approach. *Information Technology Journal*, *11*(9), 1166-1174.

AL Dhahab, H., & Barkun, A. (2012). The Acute Management of Nonvariceal Upper Gastrointestinal Bleeding. *Ulcers*, 1-8.

Albeldawi, M., Qadeer, M. A., & Vargo, J. J. (2010). Managing acute upper GI bleeding, preventing recurrences. *Cleveland Clinic journal of medicine*, *77*(2), 131-142.

Anderson, R. S., & Witting, M. D. (2010). Nasogastric aspiration: a useful tool in some patients with gastrointestinal bleeding. *Annals of emergency medicine*, *55*(4), 364-365.

Barkun, A. N., Bardou, M., Kuipers, E. J., Sung, J., Hunt, R. H., Martel, M., & Sinclair, P. (2010). International consensus recommendations on the management of patients with nonvariceal upper gastrointestinal bleeding. *Annals of internal medicine*, *152*(2), 101-113.

Chu, A., Ahn, H., Halwan, B., Kalmin, B., Artifon, E. L. A., Barkun, A., Lagoudakis, M. G., & Kumar, A. (2008). A decision support system to facilitate management of patients with acute gastrointestinal bleeding. *Artificial intelligence in medicine*, *42*(3), 1-41.

Crooks, C., Card, T., & West, J. (2011). Reductions in 28-day mortality following hos-pital admission for upper gastrointestinal hemorrhage. *Gastroenterology*, *141*(1), 62-70.

Domingos, P. (1997). On the Optimality of the Simple Bayesian Classi er under Zero-One Loss. *Machine learning*, *29*(2), 103-130.

Gralnek, I. M., & Dulai, G. S. (2004). Incremental value of upper endoscopy for triage of patients with acute non-variceal upper-GI hemorrhage. *Gastrointestinal endoscopy*, *60*(1), 9-14.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10-18.

Hanley, A., Mcneil, J., & Ph, D. (1982). Under a Receiver Characteristic. *Radiology*, *143*, 29-36.

Jensen, D. M., Machicado, G. A. (1997). Colonoscopy for diagnosis and treatment of severe lower gastrointestinal bleeding. Routine outcomes and cost analysis. *Gastrointestinal endoscopy clinics of North America*, *7*(3), 477.

Jensen, F.V., (1996). *An introduction to Bayesian Networks*. University College London press London.

Kazmierska, J., & Malicki, J. (2008). Application of the Naive Bayesian Classi er to optimize treatment decisions. *Radiotherapy and Oncology*, *86*(2), 211-216.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International joint Conference on artificial intelligence*, *14*, 1137-1145.

Kononenko, I. (1993). Inductive and Bayesian Learning in Medical Diagnosis. *Applied Artificial Intelligence, 7*, 317-337.

Kovacs, T. O., Jensen, D. M. (2002). Recent advances in the endoscopic diagnosis and therapy of upper gastrointestinal, small intestinal, and colonic bleeding. *The Medical clinics of North America*, *86*(6), 1315-1319.

Kuncheva, L. I., Hoare, & Zoe, S. J. (2008). Error-dependency relationships for the Naive Bayes classifier with binary features. *Pattern Analysis and Machine Intelligence, IEEE Transactions, 30*(4), 735-740.

Lanas, A., Garc a-Rodr guez, L. A., Arroyo, M. T., Gomollon, F., Feu, F., Gonzalez- Perez, A., Zapata, E., Bastida, G., Rodrigo, L., & Santolaria, S. (2006). Risk of upper gastrointestinal ulcer bleeding associated with selective cyclo-oxygenase-2 Inhibitors, traditional non-aspirin non-steroidal anti-inflammatory drugs, aspirin and combinations. *Gut*, *55*(12), 1731-1738.

Manning-Dimmitt, L. L., Dimmitt, S. G., & Wilson, G. R. (2005). Diagnosis of gastrointestinal bleeding in adults. *American family physician*, *71*(7), 1339-1346.

Mittal, A., & Kassim, A. (2007). *Bayesian network technologies: applications and graphical models*. IGI publishing.

Palmer, K. R. (2002). Non-variceal upper gastrointestinal haemorrhage: guidelines. *Gut*, *51*(suppl 4), iv1-iv6.

Parkes, B. M., Obeid, F. N., Sorensen, V. J., Horst, H. M., Fath, J. J. (1993). The management of massive lower gastrointestinal bleeding. *The American surgeon*, *59*(10), 676-678.

Quinlan, J. R. (1989). Unknown Attribute Values In Induction. In *Proceedings of the sixth international workshop on Machine learning*, pp.164-168.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*, volume 240 of *Morgan Kaufmann series in Machine Learning*. Morgan Kaufmann.

Quirk, D. M., Barry, M. J., Aserko, B., & Podolsky, D. K. (1997). Physician specialty and variations in the cost of treating patients with acute upper gastrointestinal bleeding. *Gastroenterology*, *113*(5), 1443-1448.

Rish, I. (2001). An empirical study of the naive Bayes classifier. *In IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pp. 41-46.

Tarone, R. E., Blot, W. J., & McLaughlin, J. K. (2004). Nonselective nonaspirin non-steroidal anti-inflammatory drugs and gastrointestinal bleeding: relative and absolute risk estimates from recent epidemiologic studies. *American journal of therapeutics*, *11*(1), 17-25.

Velayos, F. S., Williamson, A., Sousa, K. H., Lung, E., Bostrom, A., Weber, E. J., Ostro , J. W., & Terdiman, J. P. (2004). Early predictors of severe lower gastrointestinal bleeding and adverse outcomes: a prospective study. *Clinical Gastroenterology and Hepatology*, *2*(6), 485-490.

Vreeburg, E. M., Snel, P., De Bruijne, J. W., Bartelsman, J. F., Rauws, E. A., & Tytgat, G. N. (1997). Acute upper gastrointestinal bleeding in the Amsterdam area: incidence, diagnosis, and clinical outcome. *The American journal of gastroenterology*, *92*(2), 236-243.

Wei, W., Visweswaran, S., & Cooper, G. F. (2011). The application of naive Bayes model averaging to predict Alzheimer's disease from genome-wide data. *Journal of the American Medical Informatics Association*, *18*(4), 370-375.

Witting, M. D., Magder, L., Heins, A. E., Mattu, A., Granja, C. A., & Baumgarten, M. (2004). Usefulness and validity of diagnostic nasogastric aspiration in patients without hematemesis. *Annals of emergency medicine*, *43*(4), 525-532.

Witting, M. D., Magder, L., Heins, A. E., Mattu, A., Granja, C., & Baumgarten, M. (2006). ED predictors of upper gastrointestinal tract bleeding in patients without hematemesis. *The American journal of emergency medicine*, *24*(3), 280-285.

Pertanika J. Sci. & Technol. 22 (2): 577 - 585 (2014)

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Removal of Toluene from Aqueous Solutions Using Oil Palm Shell Based Activated Carbon: Equilibrium and Kinetics Study

Kwong, W. Z.[1], Tan, I. A. W.[2], Rosli, N. A.[1] and Lim, L. L. P.[1]*

[1]*Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia*

[2]*Department of Chemical Engineering and Energy Sustainability, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia*

ABSTRACT

This study is an attempt to investigate the adsorption of petroleum hydrocarbon (toluene) from aqueous solutions using granular activated carbon (GAC) synthesized from oil palm shell (OPS) (referred as OPS-based GAC). This study involved a series of batch experiments to determine the adsorption equilibrium and kinetics. The batch experiments were conducted by shaking 200 mL toluene solution containing 0.4 g GAC (initial concentrations of 5, 15, 25 and 30 mg/L) at 180 rpm at 30°C. The OPS-based GAC achieved more than 80% toluene removal in all the experiments. The adsorption capacity of the OPS-based GAC estimated using Freundlich isotherm was 6.039 mg/g $(L/mg)^{1/n}$. The adsorption kinetic study showed that the adsorption of toluene was of chemisorption as the experimental data fitted better to the pseudo-second-order kinetic model than the pseudo-first-order kinetic model.

Keywords: Adsorption, equilibrium study, granular activated carbon, kinetics study

INTRODUCTION

Groundwater pollution caused by the petroleum industry usually occurs due to the leakage from gasoline storage tanks and pipelines, which is considered as point source pollution (Haest *et al.*, 2010). Groundwater contaminated by petroleum hydrocarbons becomes undrinkable due to the unpleasant taste and odour produced by petroleum hydrocarbons such as methyl tert butyl ether (MTBE) at a concentration as low as 20 µg/L (Khan *et al.*, 2011; Lim & Lynch, 2011). Some petroleum hydrocarbons such as benzene are toxic and carcinogenic (Yu *et al.*, 2011; Caprino & Togna, 1998). In addition, the monoaromic hydrocarbons

Article history:
Received: 21 February 2013
Accepted: 24 April 2013

Email addresses:
Kwong, W. Z. (go_go_kwong@hotmail.com),
Tan, I. A. W. (awitan@feng.unimas.my),
Rosli, N. A. (rnazalina@feng.unimas.my),
Lim, L. L. P. (llpleonard@feng.unimas.my)
*Corresponding Author

ISSN: 0128-7702

benzene, toluene, ethylbenzene and xylene (BTEX) can pose a serious threat to human health if consumed (Hameed *et al.*, 2004). Meanwhile, long-term exposure of toluene to human can lead to irreversible and progressive harm to the central nervous system, kidney and liver (ATSDR, 2000). Therefore, there is a need to remediate the groundwater sources contaminated by petroleum hydrocarbons, especially where groundwater is the main drinking water source (Reddy, 2008).

Various groundwater remediation technologies have been developed to remove petroleum hydrocarbons from groundwater such as pump-and-treat, air sparging, and soil vapour extraction. However, the efficiency of these technologies, which typically involve injection or withdrawal operation, is limited by several factors such as soil heterogeneity, strata and hydraulic conductivity (Reddy, 2008). Furthermore, some of these technologies need to be combined with other technologies such as air sparging and soil vapour extraction in order to achieve the desired clean up level, which leads to the increase of the remediation cost. A groundwater remediation approach, which is not affected by soil heterogeneity, is the permeable reactive barrier (PRB). PRB is a trench with subsurface emplacements of reactive materials designed to inhibit migration of a contaminant plume (Craig *et al.*, 2006). Adsorption process is commonly used in PRB to remove contaminants in groundwater (Yu *et al.*, 2011).

Adsorption is a physical process involving the attraction of adsorbate molecules to the surface of the adsorbent. This process has been proven to be effective in removing petroleum hydrocarbons from water (Oehm *et al.*, 2007). Granular activated carbon (GAC) synthesized from coconut had been reported to achieve up to 67.05 mg/g MTBE removal, with an initial concentration of 300 mg/L (Khan *et al.*, 2011) and the adsorption capacity of toluene using adsorbent synthesized from discarded rubber tyre showed 7.7535 mg/g at 30°C using the Freundlich isotherm (Hameed *et al.*, 2004).

Adsorption, a passive process, is suitable for removing contaminants from groundwater due to the slow groundwater flow (33 m/year) (Mackay *et al.*, 1986), which implies the slow migration of contaminants and long residence time for contaminated groundwater to flow through PRB. Unlike bioremediation and chemical remediation, the adsorption process does not produce harmful by-products as it mainly immobilizes contaminant molecules from being transported by groundwater. There is a potential for GAC to be synthesized from oil palm shell (OPS) (referred as OPS-based GAC) for groundwater remediation due to the use of cheaper raw material and the low operational cost using an adsorption process. In Malaysia, the production of OPS-based GAC can also reduce significant amount of oil palm wastes to be land filled (Chin, 2011; Pilo, 2011).

Therefore, OPS-based GAC was used to remove toluene from aqueous solutions in a batch experiment in this study. In particular, this study investigated the adsorption performance of toluene onto the OPS-based GAC, encompassing the adsorption equilibrium and kinetic studies.

MATERIALS AND METHOD

Adsorbent

Commercially available OPS-based GAC was obtained from Bravo Green Sdn. Bhd. located in Kota Samarahan, Sarawak. The size of the OPS-based GAC typically ranges from 1.68 to

3.36 mm (US mesh size of 6 × 12). Meanwhile, the BET surface area and pore volume of the OPS-based GAC ranged from 1050 to 1200 m^2/g and 0.48 cm^3/g, respectively.

Adsorbate

The adsorbate used in this study was toluene (purity: ≥99.8%, Fisher Chemical, 108-88-3). The initial concentrations of toluene used in this study were 5, 15, 25 and 30 mg/L. The highest toluene concentration was chosen based on the concentration used by Lim and Lynch (2011). The synthesized groundwater was prepared by spiking a predetermined amount of toluene into distilled water.

Batch Experiment

Batch experiment was conducted to study the effects of initial concentration on the adsorption capacity of the OPS-based GAC. The experiments were conducted by adding 0.4 g OPS-based GAC into 200 mL toluene solution of different initial concentrations. The initial concentrations of toluene used were 5, 15, 25 and 30 mg/L. The conical flasks were agitated in an incubated benchtop orbital shaker (Forma) at 180 rpm at 30 ± 1°C for 24 hours to obtain a well-mixed solution and to increase the contact frequency between the adsorbent and the adsorbate.

Sample Analysis

The solutions were sampled and analyzed in a 4 mL cuvette using a UV-visible spectrophotometer (Shimadzu UV Mini 1240) at room temperature of 22 ± 2°C. Methanol was spiked into the solution to emulsify the petroleum hydrocarbons in water prior to analysis. The methanol is an organic solvent which disperses the petroleum hydrocarbons into little droplets so that it is easily detected using the spectrophotometer.

The amount of adsorption at equilibrium, q_e (mg/g), was calculated using Equation (1):

$$q_e = \frac{(C_o - C_e)V}{W} \quad (1)$$

where C_o and C_e (mg/L) are the liquid-phase concentrations of toluene at initial and equilibrium, respectively. V is the volume of the solution (L) and W is the mass of adsorbent used (g).

RESULTS AND DISCUSSION

Effect of the Initial Concentration of Toluene on Adsorption Equilibrium

Figure 1 shows the effects of toluene initial concentrations on the adsorption of toluene by the OPS-based GAC at 30°C. The OPS-based GAC achieved more than 80% toluene removal in all the batch experiments. The adsorption curves obtained in all the experiments exhibited an increasing trend leading to plateau, indicating that the toluene molecules were adsorbed into the OPS-based GAC until the adsorption equilibrium was achieved. The curve reached a plateau at about 180 minutes after the experiment started for all the initial concentrations. At

the equilibrium point, the adsorption of toluene achieved dynamic equilibrium with the OPS-based GAC. The equilibrium concentration (C_e) increased from 0.186 to 3.871 mg/L with the increase of the initial concentration from 5 to 30 mg/L. The amount of toluene adsorbed at the equilibrium time reflected the maximum adsorption capacity of the OPS-based GAC under the specified operating conditions. In this study, the adsorption capacity at equilibrium (q_e) increased from 1.193 to 12.171 mg/g with the increase of the initial concentration from 5 to 30 mg/L. This was because the mass transfer driving force would become larger when the initial concentration increased, resulting in a higher adsorption of toluene.

The contact time required by the adsorption process to achieve equilibrium was circa 180 mins for all the initial concentrations (see Fig.1). This trend was due to the fact that pores were abundantly available for adsorption in the initial stage, and after some time, these pores became fully occupied. In addition, the faster adsorption rate at the beginning of the adsorption process was due to the adsorption of toluene by the surface of the OPS-based GAC. After some time, the surface reached saturation and the toluene was adsorbed into the interior surface of the OPS-based GAC by pore diffusion (Tan *et al.*, 2008). This phenomenon could be explained by the adsorption process in which the toluene molecules encountered the boundary layer first, diffused from the boundary layer onto the surface of activated carbon and finally diffused into the porous structure of the activated carbon (Faust & Aly, 1983).

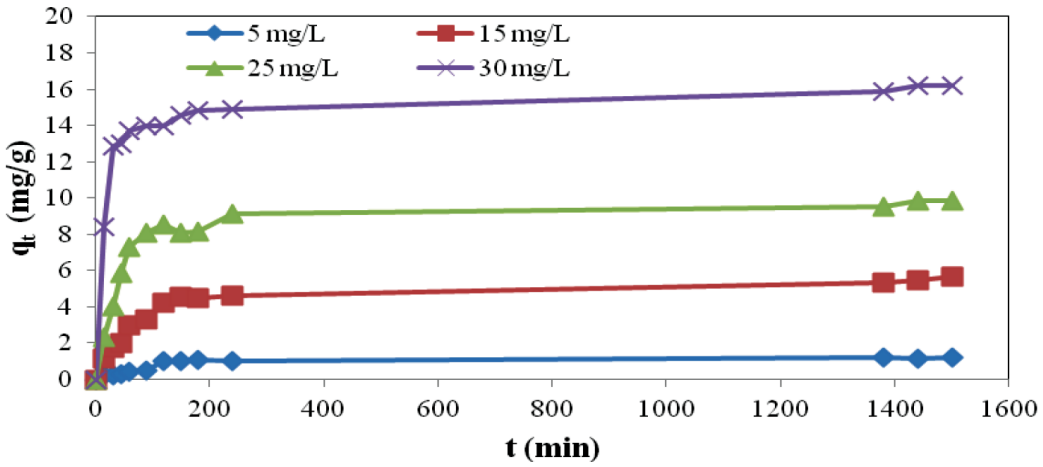


Fig.1: Adsorption of toluene using OPS-based GAC at 30°C at various initial concentrations of 5, 15, 25 and 30 mg/L

Adsorption Kinetics Study

The adsorption kinetic study was used to determine the adsorption rate of toluene into activated carbon and the equilibrium time. The two rate equations, namely, pseudo first-order [Equation (2)] (Lagergren, 1898) and pseudo-second-order [Equation (3)] were used to evaluate the adsorption kinetics and the results were compared with the adsorption data (Ho & McKay, 1999).

$$\log (q_e - q_t = \log q_e - (k1/2.303)t \quad (2)$$

$$t/q_t = 1/(k_2\, q_e^2) + (1/q_e)t \quad (3)$$

where q_e (mg/g) and q_t (mg/g) are the amounts of toluene adsorbed into the activated carbon per unit mass of activated carbon when the adsorption equilibrium is achieved and at time, t (min), respectively. Meanwhile, k_1 (min^{-1}) and k_2 (g/mg/min) are the rate constants of the pseudo-first-order and pseudo-second-order adsorption, respectively.

Table 1 summarizes the adsorption coefficients obtained using the pseudo-first-order and pseudo-second-order rate equations. The R^2 values, obtained from the pseudo-first-order and pseudo-second-order kinetics plots, were used to determine the adsorption behaviour (physisorption or chemisorption) of toluene molecules into the OPS-based GAC. The R^2 values were higher for the pseudo-second-order model than that of the pseudo-first-order model for all concentrations of toluene, indicating that the OPS-based GAC adsorption of toluene was of the pseudo-second-order adsorption. This also implied that the adsorption of toluene molecules into the OPS-based GAC was of chemisorption rather than physisorption. Chemical adsorption involves the adherent of toluene to the surface of adsorbent with covalent bonding through sharing or exchanging of electron between adsorbent and adsorbates (Ho & McKay, 1999).

TABLE 1
Pseudo-first-order model, pseudo-second-order model, and correlation coefficients for adsorption of toluene on OPS-based GAC at 30°C

Initial Conc. (mg/L)	q_e exp. (mg/g)	% Removal	Pseudo-first order parameter			Pseudo-second order parameter		
			q_e calc. (mg/g)	K_1 (min^{-1})	R^2	q_e calc. (mg/g)	k_2 (g/mg/min)	R^2
5	1.193	92.760	0.603	0.002	0.739	1.239	1.05E-02	0.993
15	5.636	97.400	2.812	0.002	0.773	5.714	3.36E-03	0.998
25	9.850	83.830	3.611	0.001	0.819	10.000	3.51E-03	0.999
30	16.171	90.920	3.579	0.002	0.586	16.393	4.61E-03	1.000

Adsorption Isotherms

It is important to determine the most appropriate correlation for the equilibrium curve using adsorption isotherms in order to optimize the design of adsorption system. The adsorption isotherms are used to indicate how the adsorption molecules are distributed between liquid phase and solid phase when the adsorption process reaches an equilibrium state. In this study, two well-known adsorption isotherms, namely, the Langmuir (Langmuir, 1916) and Freundlich (Freundlich, 1906) adsorption models were used and compared by judging the correlation coefficient, R^2, as a criterion (Hussain *et al*., 2006; Nuithitikul *et al*., 2010). Equations (4) and (5) represent the linear forms of Langmuir and Freundlich models, respectively:

$$C_e/q_e = 1/(Q_0\, K_L) + (1/Q_0)\, C_e \quad (4)$$

Q_0 (mg/g) and K_L (L/mg) are Langmuir constants in relevant to maximum monolayer adsorption capacity onto a surface with no transmigration of adsorbates onto the surface (Webi & Chakravort, 1974).

$$\text{Log } q_e = \log K_F + (1 / n) \log C_e \quad (5)$$

K_F (mg/g(L/mg)$^{1/n}$) and n are Freundlich constants representing the adsorption capacity on a heterogeneous surfaces and how favourable the adsorption process is, assuming that the stronger binding sites are occupied first and that the binding strength decreases with the increasing degree of site occupation, respectively (Freundlich, 1906; Nuithitikul *et al.*, 2010). The slope of 1/n is used to measure the adsorption intensity or surface heterogeneity ranging between 0 and 1, with the surface becoming more heterogeneous as the value becomes closer to zero (Haghseresht & Lu, 1998). Another essential characteristic of the Langmuir isotherm is a dimensionless equilibrium parameter (R_L) (Nuithitikul *et al.*, 2010; Webi & Chakravort, 1974):

$$R_L = 1 / (1 + K_L C_0) \quad (6)$$

where K_L is the Langmuir constant and C_0 is the initial contaminant concentration (mg/L). The value of R_L indicates whether the isotherm is unfavourable ($R_L > 1$), linear ($R_L = 1$), favourable ($0 < R_L < 1$) or irreversible ($R_L = 0$) (Nuithitikul *et al.*, 2010).

For the Langmuir isotherm, the correlation coefficient was found to be 0.704, while the value of R_L was 0.053 at the temperature of 30°C. This indicated that the Langmuir isotherm was favourable for the adsorption of toluene on GAC under the conditions used in this study. Meanwhile, Langmuir constants K_L and Q_0 were calculated using Equation (4) and their values are shown in Table 2. The maximum monolayer adsorption capacity obtained from the Langmuir isotherm was 16.129 mg/g. The adsorption capacity of the OPS-based GAC was higher than the adsorbent synthesized from montmorillonite (6.71 mg/g) (Nourmoradi *et al*., 2012).

TABLE 2
Langmuir and Freundlich isotherm model constants and correlation coefficient for adsorption of toluene onto GAC at 30°C

Langmuir Isotherm				Freundlich Isotherm		
Q_0 (mg/g)	K_L (L/mg)	R^2	R_L	1/n	K_F (mg/g(L/mg)$^{1/n}$)	R^2
16.129	0.756	0.704	0.053	0.618	6.039	0.743

For the Freundlich isotherm, the correlation coefficient was 0.743. The slope of 1/n obtained was 0.618, indicating that the adsorption of the toluene onto the GAC surface was homogeneous since it was closer to one. The correlation coefficient, R^2, for the Freundlich isotherm model was higher than that of the Langmuir model, which showed that the Freundlich model was more favourable for the adsorption of toluene onto OPS-based GAC under the specified condition. The adsorption capacity of the OPS-based GAC obtained using the Freundlich model was 6.039 mg/g(L/mg)$^{1/n}$. This result was comparable with that of discarded

rubber tyre-based adsorbent, with the K_F value of 7.7535 mg/g(L/mg)$^{1/n}$ (Hameed *et al.*, 2004). The adsorption capacity of the GAC for toluene was compared with other adsorbents reported in the literature as shown in Table 3.

TABLE 3
A comparison of the adsorption capacity of toluene on various adsorbents

Adsorbents	Adsorption capacity	References
OPS-based GAC[a]	16.129 mg/g	This work
OPS-based GAC[b]	6.039 mg/g(L/mg)$^{1/n}$	This work
Montmorillonite modified with nonionic surfactant[a]	6.71 mg/g	Nourmoradi *et al.* (2012)
Discarded rubber tyre-based adsorbent[b]	7.7535 mg/g(L/mg)$^{1/n}$	Hameed *et al*. (2004)
Bentonite[c]	66 mg/g	Amari *et al*. (2010)
Manganese-embedded PAN-based activated carbon nanofibers[c]	400 mg/g	Oh *et al*. (2008)

[a]Langmuir isotherm model
[b]Freundlich isotherm model
[c]Experimental data

CONCLUSION

This study showed that OPS-based GAC was successfully used in removing toluene from aqueous solutions. In more specific, all the adsorption achieved more than 80% toluene removal. The equilibrium data were best described by Freundlich isotherm with adsorption capacity of 6.039 mg/g(l/mg)$^{1/n}$. The adsorption capacity obtained from the Langmuir isotherm was 16.129 mg/g, while the adsorption kinetics was found to follow closely the pseudo-second-order kinetic model. This study also highlighted the potential of the activated carbon for groundwater remediation which would be used for the removal of petroleum hydrocarbons in its intended application in future study.

ACKNOWLEDGEMENTS

We would like to thank Universiti Malaysia Sarawak and Osaka Gas for funding this research through the Small Grant Scheme (grant no. 02(S84)/820/2011(18)) and Osaka Gas Foundation in Cultural Exchange (OGFICE) Research Grant Scheme, respectively, as well as Bravo Green Sdn. Bhd. for providing the OPS-based GAC.

REFERENCES

Amari, A., Chlendi, M., Gannouni, A., & Bellagi A. (2010). Optimised activation of bentonite for toluene adsorption. *Applied Clay Science*, *47*(3-4), 457-461.

ATSDR. (2000). *Public Health Statement: Toluene.* Atlanta: Agency for Toxic Substances and Disease Registry Division of Toxicology and Environmental Medicine.

Caprino, L., & Togna, G.I. (1998). Potential health effects of gasoline and its constituents: A review of current literature (1990-1997) on toxicological data. *Environmental Health Perspectives*, *106*(3), 115-125.

Chin, J. (2011). *Sarawak's share in Malaysia oil palm.* The Borneo Post, Kuching.

Craig, J. R., Rabideau, A. J., & Suribhatla, R. (2006). Analytical expressions for the hydraulic design of continuous permeable reactive barriers. *Advances in Water Resources*, *29*(1), 99-111.

Faust, D. S., & Aly, M. O. (1983). *Chemistry ofWastewater Treatment.* Butterworthworts: Boston.

Freundlich, H. M. F. (1906). Over the adsorption in solution. *Journal of Physical Chemistry*, 385-470.

Haest, P. J., Lookman, R., Keer, I. V., Patyn, J., Bronders, J., Joris, M., Bellon, J., & De Smedt, F. (2010). Containment of groundwater pollution (methyl tertiary butyl ether and benzene) to protect a drinking-water production site in Belgium. *Hydrogeology Journal*, *18*(8), 1917-1925.

Haghseresht, F., & Lu, G.Q. (1998). Adsorption characteristics of phenolic compounds onto coal-reject-derived adsorbents. *Energy Fuels*, *12*(6), 1100-1107.

Hameed, B. H., Mohamed, A. R., & Chong, H. Y. (2004). Adsorption of toluene using low cost adsorbent. *Jurnal Teknologi*, *40*(F), 17-26.

Ho, Y. S., & McKay, G. (1999). Pseudo-second order model for sorption processes. *Process of Biochemistry*, *34*(5), 451-465.

Khan, M. A., Lee, S.-H., Kang, S., Paeng, K.-J., Lee, G., Oh, S.-E., & Jeon, B.-H. (2011). Adsorption studies for the removal of methyl tert-butyl ether on various commercially available GACs from an aqueous medium. *Separation Science and Technology, 46*(7), 1121–1130.

Lagergren, S. (1898). Zur Theorie der Sogenannten Adsorption Geloster Stoffe. *Veternskapsakad Handlingar*, *24*(4), 1-39.

Langmuir, I. (1916). The constitution and fundamental properties of solids and liquids. Part I: Solids. *Journal of the American Chemical Society*, *38*(11), 2221-2295.

Lim, L. L. P., & Lynch, R. J. (2011). Feasibility study of a photocatalytic reactor for in situ groundwater remediation of organic compounds. *Journal of Hazardous Materials*, *194*, 100-108.

Mackay, D. M., Freyberg, D. L., Roberts, P. V., & Cherry, J. A. (1986). A natural gradient experiment on solute transport in a sand aquifer: Approach and overview of plume movement. *Water Resources Research, 22*(13), 2017-2029.

Nourmoradi, H., Nikaeen, M., & Khiadani (Hajian), M. (2012). Removal of benzene, toluene, ethylbenzene and xylene (BTEX) from aqueous solutions by montmorillonite modified with nonionic surfactant: Equilibrium, kinetic and thermodynamic study. *Chemical Engineering Journal, 191,* 341-348.

Nuithitikul, K., Srikhun, S., & Hirunpraditkoon, S. (2010). Kinetics and equilibrium adsorption of basic green 4 dye on activated carbon derived from durian peel: Effects of pyrolysis and post-treatment conditions. *Journal of the Taiwan Institute of Chemical Engineers*, *41*(5), 591–598.

Oehm, C., Stefan, C., Werner, P., & Fischer, A. (2007). Adsorption and abiotic degradation of methyl tert-butyl ether (MTBE). In D. Barcelo (Ed.), *The Handbook of Environmental Chemistry* (pP. 191-212). New York: Springer-Verlag Berlin Heidelberg.

Oh, G.-Y., Ju, Y.-W., Jung, H.-R., & Lee, W.-J. (2008). Preparation of the novel manganese-embedded PAN-based activated carbon nanofibers by electrospinning and their toluene adsorption. *Journal of Analytical and Applied Pyrolysis*, *81*(2), 211-217.

Pilo, W. (2011). *Sarawak Needs More Palm Oil Mills to Stem Loss.* The Borneo Post, Kuching.

Reddy, K. R. (2008). Physical and chemical groundwater remediation technologies. In C. J. G. Darnault, *Overexploitation and Contamination of Shared Groundwater Resources* (pp. 257-274). Chicago: Springer Science + Business Media BV.

Tan, I. A. W., Ahmad, A. L., & Hameed, B. H. (2008). Enhancement of basic dye adsorption uptake from aqueous solutions using chemically modified oil palm shell activated carbon. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *318* (1-3), 88-96.

Webi, T. W., & Chakravort, R. K. (1974). Pore and solid diffusion models for fixed-bed adsorbers. *AIChE Journal*, *20*(2), 228-238.

Yu, F., Ma, J., & Wu, Y. (2011). Adsorption of toluene, ethylbenzene and xylene isomers on multi-walled carbon nanotubes oxidized by different concentration of NaOCl. *Frontiers of Environmental Science and Engineering in China*, *6*(3), 320-329.

Pertanika J. Sci. & Technol. 22 (2): 587 - 592 (2014)

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Delay in Diagnosis of Lung Cancer: A Case Report

Ching, S. M.[1]*, Chia, Y. C.[2] and Cheong, A. T.[1]

[1]*Department of Family Medicine, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

[2]*Department of Primary Care Medicine, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia*

ABSTRACT

This case report highlights delay in the diagnosis of adenoma carcinoma of the lung in a female patient who has never smoked. It took three months to reach the diagnosis of stage IV lung carcinoma despite the presence of symptoms and an abnormal chest radiograph finding from the beginning. The clinical characteristics and predictors of missed opportunities for an early diagnosis of lung cancer are discussed. In this case, patient and doctor factors contributed to the delay in diagnosis. Thus, early suspicions of lung cancer in a woman with the presence of respiratory symptoms despite being a non-smoker are important in primary care setting.

Keywords: Adenocarcinoma, lung cancer, female, delay diagnosis, non-smoker

INTRODUCTION

Carcinoma of the lung is the second commonest cancer among men and the sixth most common cancer among women in Peninsular Malaysia (Ministry of Health, 2008). Lung cancer is notoriously known to be at an advanced stage by the time it is diagnosed and this unfortunately means a poorer prognosis (Peake, 2008). Cough and dyspnoea are the common symptoms of lung cancer, which also are common complaints in the primary care setting (Hamilton, Peters, Round, & Sharp, 2005). However, when these symptoms are accompanied by haemoptysis, alarm bells should have been sounded. Haemoptysis may not be present in the early stages, and this may cause delay in diagnosis of lung cancer. Thus, this raises the concern, i.e. whether this could increase the risk of missing or delaying the diagnosis of lung cancer in patients who have cough and dyspnoea in the primary care setting (Ellis & Vandermeer, 2011; Bjerager, Palshof, Dahl, & Olesen, 2006).

Article history:
Received:
Accepted:

Email addresses:
Ching, S. M. (sm_ching@upm.edu.my),
Chia, Y. C. (chiayc@um.edu.my),
Cheong, A. T. (cheaitheng@upm.edu.my)
*Corresponding Author

ISSN: 0128-7702

The study by Myrdal *et al.* which reported that shorter delay in diagnosis might not confer to the better prognosis on a survival rate, which could be due to the fact that the majority of the patients in that study were already in the advanced stage (Myrdal G *et al*., 2004). For a patient with a large but potentially operable tumour, however, delay may be detrimental (O'Rourke & Edwards, 2000). The following case study provides an example that a greater sense of awareness is needed by physicians when dealing with a female patient who is a non-smoker but who could potentially be harbouring a lung cancer.

CASE STUDY

A 64-year-old para 4, post-menopausal housewife with underlying hypertension on diet control, complained of cough, shortness of breath, haemoptysis and occasional chest pain for the past 10 weeks. Otherwise, she was active in terms of her activities of daily living. She went for a chest radiograph (CXR) at a general practitioner's clinic and was treated as community-acquired pneumonia. She was reassured that there was no need to be worried as a course of antibiotic would be prescribed for her condition. She was neither given a follow-up or a referral. Her symptoms did not improve even completing the course of antibiotic. The family members started to worry about her after three months of unresolved symptoms and decided to come to the hospital for a second opinion instead of going back to the previous doctor.

She was not a primary or secondary smoker. She denied having any symptoms of fevers, night sweat and did not have any other known risk factors for developing lung cancer. There was no family history of malignancy as well.

Upon examination, she appeared thin but there was no jaundice. Dilated vein was noted on the anterior chest wall and finger clubbing was also present. However, there was no cervical lymphadenopathy or facial swelling noted. Her respiratory rate was 18 breaths per minute, blood pressure of 156/ 80 mmHg, pulse rate of 80 beats per minute with a BMI of 20.1 kg/m^2.

The chest expansion and air entry were reduced together with stony dullness noted on the lower zone of the right lung. There was no crepitation or rhonchi heard. On top of that, there was a palpable liver, which is firm in texture and measuring about 2cm below the right costal margin. Breast and spine examinations were also normal.

Investigations

The patient went for investigations which consisted of full blood count test, sputum for acid fast bacilli, sputum for culture and sensitivity test, chest radiograph and a Mantoux test. The full blood count test showed that the patient had mild anaemia with normal white cell count and ESR (Table 1). The mild anaemia could be due to the underlying chronic disease or malignancy in this case. The tuberculosis workout was negative in view of the negative findings for the acid fast bacilli and Mantoux test (Table 2).

However, her CXR which was done in the clinic showed an ill-defined opacity and consolidation at the right middle lobe, with an ill-defined mass obscuring right heart border. There was a large right pleural effusion with an elevated right horizontal fissure seen (Figure 1). Subsequently Computed tomography (CT) scans (Figure 2) and bronchoscopic guided biopsy

of the chest were done two weeks later. There was an irregular heterogeneously enhancing mass with necrotic component at the lower lobe of the right lung medial segment, measuring about 4.5 x 4.3x 7cm. It showed a metastatic nodular lesion in the apical segment of the right lung's lower lobe. The left lung was clear and right hilar lymph node was enlarged by 2.1 cm. The right moderate pleural effusion was also present and two hypodense lesions were noted in the liver as well. Histopathology examination revealed a moderately differentiated adenocarcinoma (stage 4 N2 M1b) of the lung.

Patient's Progress

Despite being treated with targeted therapy (Gefitinib), the patient developed brain metastasis. Her condition deteriorated and she succumbed to the disease 18 months after the diagnosis.

DISCUSSION

A delay in any diagnosis of cancer is associated with substantial disability to patients (Phillips Jr., Bartholomew, & Dovey, 2004) and more recently, malpractice claims from physicians or hospitals (Singh, Sethi, & Raber, 2007).

Basically, delays could be due to three factors. Firstly, the delay could be due to patient's factors whereby the patient delays her/his decision in consulting a doctor when the symptoms first present. Then, the patient may have poor adherence to the advice or appointment, or she/he may ask for a second opinion from different health care centres, which can subsequently contribute to late diagnosis (Corner, Hopkinson, & Roffe, 2006). Secondly, physician's factors such as misinterpreting symptoms and treating the disease as some other diseases which share similar presentations cause a delay in referral. Thirdly, "system" factors where there is a long waiting period for appointments, imaging or diagnostic tests (Devbhandari, Bittar, & Quennell, 2007) often cause further delay. Besides all the above factors, there is more delay for treatment to be initiated upon diagnosis by the specialist. In other words, it may take around 3 to 7 months from the time of presentation before the treatment is started (Eija-Riitta, Sällinen, Hiekkanen, & Liippo, 2005).

The delay in the interval between the first symptom and patient's presentation to the clinic is complex (Muers, Holmes, & Littlewood, 1999). In this patient, she took 10 weeks before she first consulted a doctor. This is longer compared to another study, which reported that the mean interval between the first symptom and patient's visit to the family doctor was 21 days (Koyi, Hillerdal, & Brandén, 2002). The most common reason for the delay in the first consultation was the patient's perception that the present complaint was not serious (Ellis & Vandermeer, 2011).

Meanwhile, the delay between the first consultation to the primary-care doctor and a confirmed diagnosis in the secondary care are important issue highlighted in this patient. The primary-care doctor is often very busy and can easily miss the early signs and symptoms of lung cancer, especially if the suspicion is not kept at the back of the mind (Bjerager *et al*., 2006). This may cause the patient to go for doctor shopping, although this is not the case here. On top of that in the Malaysia present, neither has any established cancer control programme,

nor any policy for early detection of any cancer.

This may be quite challenging for the doctor who first saw her to reach the correct diagnosis. This is because her presenting symptoms could be due to the underlying infections like pneumonia or pulmonary tuberculosis, which is relatively more prevalent in Asian countries like Malaysia. To further complicate matters, lung cancer also commonly presents initially with chest infection (Singh *et al*., 2010), 2010). However, test for sputum acid-fast bacilli (AFB) should be performed by the doctor who initially saw her in view of the high prevalence of pulmonary tuberculosis in the local setting. In a case where the sputum AFB is negative in the presence of an abnormal CXR finding, the patient should be referred for an early CT scan or seen early by the chest physician for further evaluation. Despite haemoptysis being one of the cardinal and pertinent features of lung cancer (Hamilton, 2005), this symptom did not manage to alarm the doctor attending to this patient as she was neither given a follow-up or a referral following the first consultation.

The delay in diagnosis of this patient could be avoided if the initial doctor had monitored the patient's progress by giving a follow-up or a safety netting where she needed to come back for re-evaluation if the condition did not improve. Studies have shown that the duration of delay in diagnosing lung cancer could be shortened to 19 days if the predictors of missed opportunities for early diagnosis of lung cancer can be recognised (Singh *et al*., 2010). The predictors identified from this study are: recurrent bronchitis (odds ratio [OR], 3.31; 95%CI, 1.20 to 9.10), failure in following up an abnormal chest X-ray (OR, 2.07; 95% CI, 1.04 to 4.13) and completion of first needle biopsy (OR, 3.02; 95% CI, 1.76 to 5.18).

As for the screening of asymptomatic lung cancer, studies have shown that using a low dose-computered tomography demonstrated 20% reduction of lung cancer compared with screening with a chest radiograph (Jett, 2012). However, with limited resources and possibility of high false positive (24.2%), the authors are still uncertain of its role as a tool for screening for lung cancer in an asymptomatic patient.

CONCLUSION

In conclusion, a higher index of suspicion of lung cancer is needed in a symptomatic non smoker, female patients. The clinical features may mimic other lung conditions. Thus, close follow-up for the response of a patient to treatment is of utmost importance, and if there is no response then referral is warranted.

TABLE 1
Full blood count of the patient

Full blood count	Result	Reference Range
Haemoglobin	11.8	(12.0-15.0)g/L
White blood count	4.1	(4-10) × 10^9/L
Platelet	336	(150-400) × 10^9/L
ESR	34	mm/hr

TABLE 2
Tuberculosis workout

Tuberculosis workout	Result	Reference Range
Sputum culture and sensitivity	Normal upper respiratory tract flora	
Sputum acid fast bacilli × 3	Negative	
Mantoux test	No induration	(<10 mm)

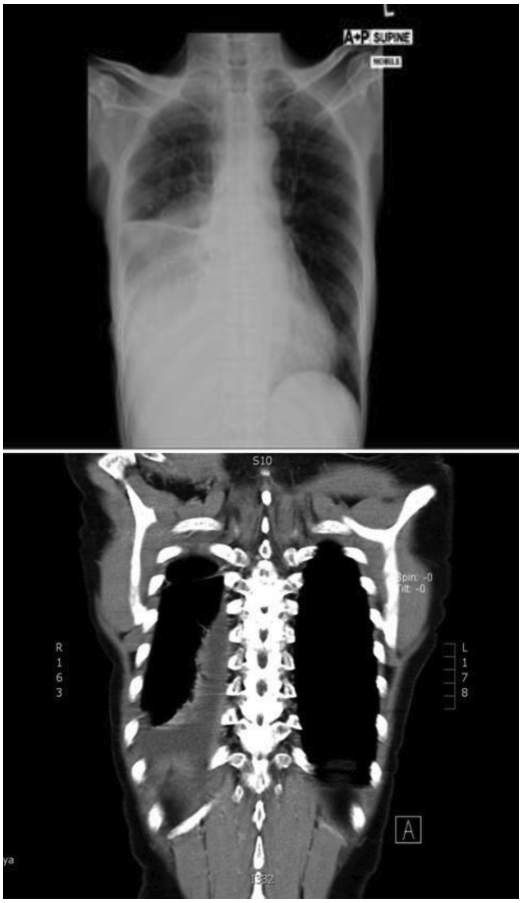


Fig.1 and Fig.2: Chest radiography (left) and CT scan of the patient shows a right pleural effusion (right)

REFERENCES

Bjerager, M., Palshof, T., Dahl, R., & Olesen, F. (2006). Delay in diagnosis of lung cancer in general practice. *Br. J. Gen. Pract., 56*(532), 863-868.

Corner, J., Hopkinson, J., & Roffe L. (2006). Experience of health changes and reasons for delay in seeking care: A UK study of the months prior to the diagnosis of lung cancer. *Soc. Sci. Med., 62*, 1381-1391.

Devbhandari, M.P., Bittar, M.N., & Quennell, P. (2007). Are we achieving the current waiting time targets in lung cancer treatment? Result of a prospective study from a large United Kingdom teaching hospital. *J. Thorac. Oncol., 2*, 590-592.

Eija-Riitta, S., Sällinen, S., Hiekkanen, H., & Liippo, K. (2005). Delays in the Diagnosis and Treatment of Lung Cancer. *CHEST, 128*, 2282-2288.

Ellis, P. M. & Vandermeer, R. (2011). Delays in the diagnosis of lung cancer. *Journal of Thoracic Disease, 3*(3), 183-188.

Hamilton, W., Peters, T.J., Round, A., & Sharp, D. (2005). What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax, 60*, 1059-1065.

Jett, J. (2012). Screening for lung cancer: who should be screened? *Arch. Pathol. Lab. Med., 136*(12), 1511-1514.

Koyi, H., Hillerdal, G., & Brandén, E. (2002). Patient's and doctors' delays in the diagnosis of chest tumors. *Lung cancer, 35*(1), 53-57.

Ministry of Health. (2008). *The Third Report of the National Cancer Registry, Malaysia.* Retrieved December 16, 2012, from http://www.radiologymalaysia.org/Archive/NCR/NCR2003-2005Bk.pdf.

Muers, M. F., Holmes, W. F., & Littlewood, C. (1999). The challenge of improving the delivery of lung cancer care. *Thorax, 54*(6), 540-543.

Myrdal, G., Lambe, M., Hillerdal, G., Lamberg, K., Agustsson, T., & Ståhle, E. (2004). Effect of delays on prognosis in patients with non-small cell lung cancer. *Thorax, 59*(1), 45-49.

O'Rourke, N., & Edwards, R. (2000). Lung cancer treatment waiting times and tumour growth. *Clin. Oncol. (R. Coll. Radiol.), 12*(3), 141-144.

Peake, M. D. (2008). Lung cancer and its management. *Medicine, 36*(3), 162-167.

Phillips, R. L. Jr, Bartholomew, L. A., & Dovey, S. M. (2004). Learning from malpractice claims about negligent, adverse events in primary care in the United States. *Qual. Saf. Health Care, 13*, 121-126.

Singh, H., Hirani, K., Kadiyala, H., Rudomiotov, O., Davis, T., & Khan, M. M. (2010). Shortening the diagnostic and treatment delay times might be possible with if the general practitioner can have a higher suspicion of this disease especially among those ex-smoker and non-smokers. *Journal of Clinical Oncology, 28*(20), 3307-3315.

Singh, H., Sethi, S., & Raber, M. (2007). Errors in cancer diagnosis: Current understanding and future directions. *J. Clin. Oncol., 25*, 5009-5018.

Pertanika J. Sci. & Technol. 22 (2): 593 - 601 (2014)

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Intelligent Monitoring Interfaces for Coal Fired Power Plant Boiler Trips: A Review

Nistah, N. N. M.[1]*, Motalebi, F.[1], Samyudia, Y.[1] and Alnaimi, F. B. I.[2]

[1]*School of Engineering, Curtin University Sarawak, CDT 250, 98009 Miri, Sarawak, Malaysia*

[2]*Faculty of Engineering, Asia Pacific University of Technology and Innovation (APU), Kuala Lumpur, Malaysia*

ABSTRACT

A major source of contemporary power is a Coal-fired Power Plant. These power plants have the capacity to continuously supply electricity to almost 500,000 residential and business units. An essential component of a Coal-fired Power plant is automation. A feature of this automation is an Intelligent System developed for the Power Plant. These Intelligent Systems have different configurations and design. This research studies the various Intelligent Monitoring Interfaces developed for Coal-fired Power Plant Trips, their advantages, disadvantages and proposes a new Intelligent Monitoring Interface that would alleviate the disadvantages of the existing systems. Current systems that use Neural Network models are investigated. The improved Intelligent Monitoring Interface as proposed in this paper is a modification of the existing monitoring system for the Coal-fired Power Plant Boiler Trips. It is expected to improve the overall system by implementing remote accessibility and interactability between the plant operator and the control system interface. The interface will also assist the operator by providing guidelines to troubleshoot the identified trips and the remote server application will allow data collected to be viewed anytime, anywhere.

Keywords: Component, Intelligent Systems, Neural Network, Coal-fired Power Plant Trips

Article history:
Received: 10 September 2013
Accepted: 18 January 2014

Email addresses:
Nistah, N. N. M. (nong.nurnie@curtin.edu.my),
Motalebi, F. (foad.m@curtin.edu.my),
Samyudia, Y. (yudi.samyudia@curtin.edu.my),
Alnaimi, F. B. I. (firas@apu.edu.my)
*Corresponding Author

INTRODUCTION

Electricity is an essential utility for a household; thus, the existence of power plants that could process our natural resources such as coal or gas into energy is very important. The main interest of the present research is to report the current development of various intelligent monitoring interface system of

ISSN: 0128-7702

thermal power plant focusing mainly on the fault detection and diagnosis for boilers. This research suggests alternate improvements to the existing monitoring interface systems. It does this by suggesting a novel interactive design using a high level programming language.

Several authors have reported on a variety of applications such as prediction, process control, and condition monitoring, diagnosing and evaluating boiler behaviour. For instance, Fast and Palmé (2010) simulated the Combined Heat and Power (CHP) plant components with an Artificial Neural Network (ANN) as the data modelling tool which is then integrated on a Power Generation Information Manager (PGIM) server. In their paper, the ANN model is continuously fed with actual data to predict the outcome of the power plant overall performance in a Graphical User Interface (GUI).

Another work includes the development of neural network models, which are trained using data collected from a power plant for simulation of a monitoring system. These developed models are reported to have a close to real-time response time which could be applied for an online monitoring system. One of the models in review implements the feed forward back propagation neural network modelling, which was trained using real plant data and reported to be quick in predicting the plant overall performances quite accurately (De *et al*., 2006). In another research, the feasibility of a coal-based or CHP power plant using neural network modelling was examined. De *et al*. (2006) developed two models which uses data from the plant for the training phase. One of the models includes the coal mass flow rate as its input parameter, while the other includes the speed of the conveyor belt as one of its input parameter. The outcome of both models were found to accurately predict real data performances even for data not included in the training, proving its usability for both online or offline real-time monitoring (Smrekar *et al*., 2009).

Another study shows the advantages of using the design of neural network modelling for a biomass boiler monitoring system. The research combines the traditional methods of monitoring using equation based modelling with the NN model. The result of the new structure proves that the new model was able to provide a close prediction with actual data reading. Thus, the conclusion of their findings was that NN is a better monitoring tool than an equation based model (Romeo & Gareta, 2005). In yet another study of a modelling method, a NN approach was used to analyse and predict a boiler's efficiency and also to generate data that could be used for efficiency enhancements. A database was created, which is combined with a soft computing assistance to identify and assess the factors that influence the critical evaluation of its current practices (Kljajicet *et al*., 2012).

This paper aim to identify the modification which best enhanced the coal-fired thermal power plant boiler trip mechanism. The enhancement can rectify the shortfall of the current system by providing an alternate solution which acts as an advisory guideline to the operator for the trips encountered while a physical examination of the equipment is conducted. The proposed project will implement artificial neural network models for the fault and diagnosis monitoring for the boiler, where Java programming language will be used to design the remotely accessible interactive interface that accurately presents the physical reality.

METHODOLOGY

Description of a Power Plant Boiler

The two basic components of the power plant are the furnace boiler and steam turbine generator. As illustrated in Fig.1 (*Thermal Boiler*, n.a.), the boiler structure is equipped with a furnace in which the pulverized coals are fed through to the load burners which ignites it, producing the maximum heat possible. Highly purified water is pumped through the water pipes to produce high pressured steam and the formed hot gases or steam is directed into the smokestack to be released to the air after desulphurization. In order to provide continuous supplies of electricity, the steam is cooled and condensed back into water which is then circulated back into the boiler to repeat the whole process.

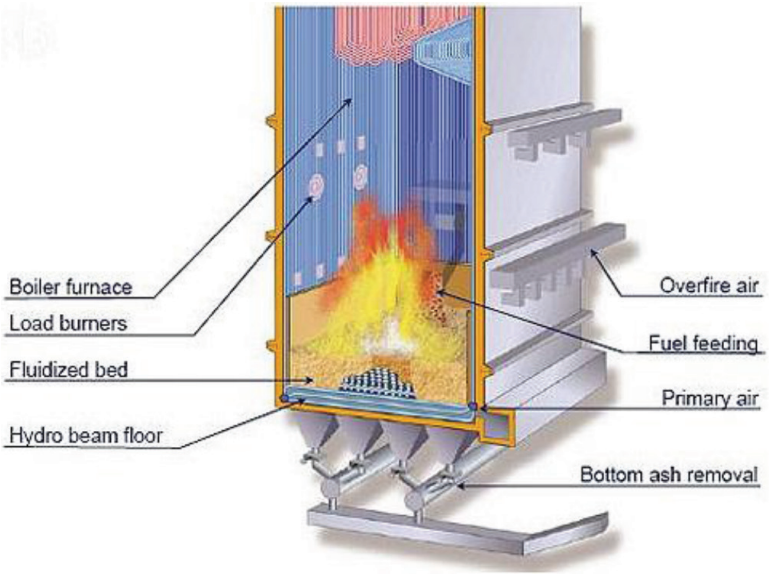


Fig.1: Components of a boiler for thermal power plant

The boilers of a power plant here in Malaysia are designed with a sub-critical pressure, single reheat and controlled circulation. Each boiler is fired with bituminous ranked pulverized coal to produce steam for the continuous generation of 700 MW of energy. The combustion circuit consists of a single furnace, with direct tangential firing and balanced draught. To comply with the Malaysian environmental requirements, the boiler has been design with a low nitro oxide (NO_x) combustion burner system which includes the over fire air (OFA) ports. An Electro-Static Precipitator (ESP) is applied to remove the dust in the flue gas at the boiler outlet and a Flue Gas Desulphurization (FGD) plant scrubs the flue gas to control the sulphur dioxide (SO_2) emission level at the stack. Apart from the boilers, the other supporting instruments of the power plant include three boiler circulating pumps, two forced draft fans, two steam air preheaters, a soot blowing equipment, two ESPs, a single coal milling plant consisting of seven vertical bowl mills and a FGD. A schematic diagram of the steam boiler is shown in Fig.2 (Alnaimi & Al-Kayiem, 2010).

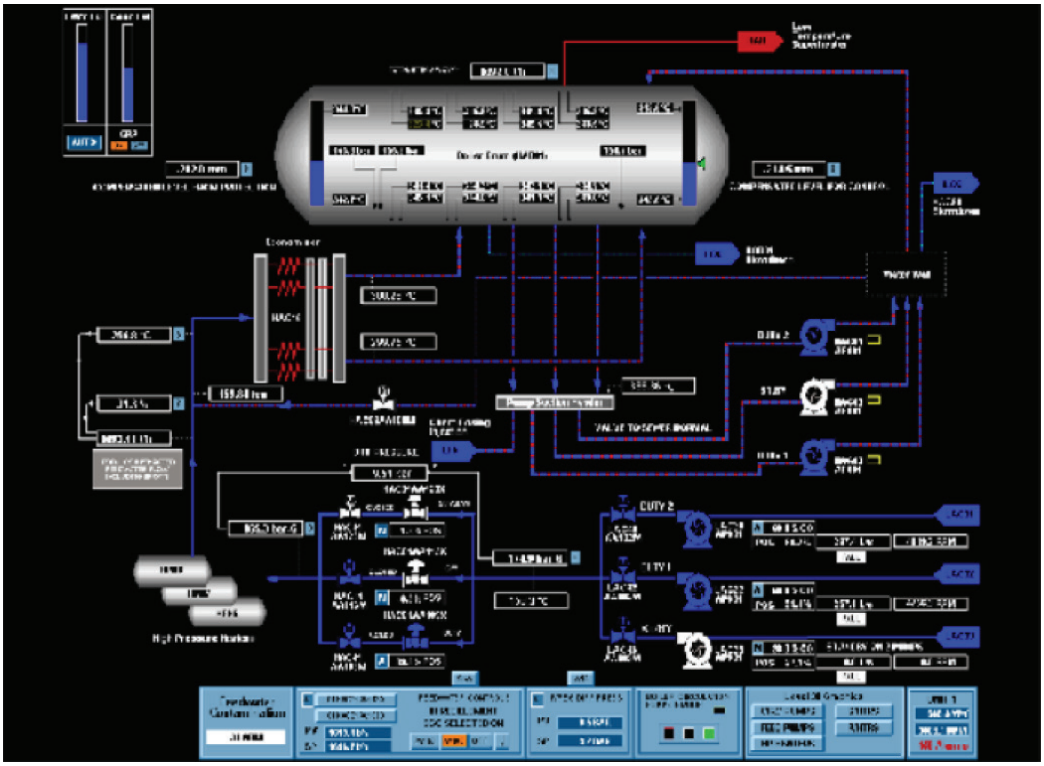


Fig.2: A schematic diagram of the steam boiler displayed on a monitoring & control room of the power plant

Monitoring System Development Phases

The improvement of a systems' output which closely resembles the desired goal can be achieved using an Artificial Neural Network. It is a group of interconnected artificial neurons, which can be trained through systematic procedures by adjusting the network parameters. The performance can be continuously improved using a performances feedback loop which implements a cost function as part of its equation. Information of these neurons is processed in parallel. Haykin (1999) describes ANN as a non-linear statistical data modelling tool that is adaptable and has the ability to learn from experience. Instead of being built from a specific set of parameter values, the parameters are automatically set from external data in its system (Principe *et al*., 2000).

Phase I: Neural Network modelling

In order to develop the NN model, there is no specific structure that can be referenced. It is determined by the complexity of the desired outcome. Hence, the appropriate number of neurons and hidden layers can only be decided through trial and error procedures. However, it should be noted that fewer number of neurons is preferable, as long as the accuracy of the prediction is not compromised (Fast & Palmé 2010). For example, Alnaimi and Al-Kayiem (2010) developed their NN model based on the lowest value of the Root Mean Squared Error

(RMSE) achievable using a single hidden layer with only one neuron and Broyden Fletcher-Goldfarb-Shanno (BFGS) Quasi-Newton training algorithm.

Phase II: Real data acquisition using neural network

It is common practice for operational data in a power plant to be recorded and stored as historical data in the system's database for future references. These readings can be beneficial for a power plant simulation programme development using ANN model.

It has been proven in a number of research that ANN has the capabilities to establish corresponding communication between the points of input and output domain in order to interpret the behaviour of energy conversion plants. The reason for ANN models is simply because it can be trained with the latest data to assess and predict the degradation of plant equipment, and hence,, providing easy, fast responding offline and online applications (Smrekar *et al*., 2009).

Data acquisition is an important step because the results will determine the success execution of the following phases. Here, real-time data were captured from the thermal power station control room, pre-processed and pre-randomized for training, validation and testing.

INTERFACES FOR BOILER TRIPS MONITORING SYSTEM

An extension to the model is the interface development for both client side data representation and control and maintenance operations.

On-Site Server Interface Application

The most common implementation of a monitoring system is through an on-site server which allows authenticated user to have access to the data using their local workstation connected to a Virtual Private Network (VPN).

Fast and Palme (2010) developed an on-site server monitoring system interface using MS EXCEL which was linked to a PGIM server. The operational data are continuously fed into the server which is integrated with an ANN model to generate prediction of the CHP plant performance. When any abnormal readings are generated, alerts and warning will be triggered and sent out to the operators through the main view of the GUI. The advantage of this GUI structure is that all the data is made available to any workstation connected to the server. This allows any authorized user to conduct individual analysis of any chosen parameter in any time intervals whenever they are required.

Remote Server Interface Application

A remote server monitoring system allows data collected on site to be sent to a remote server and be accessible through any standard web browsers (e.g., Mozilla, Internet Explorer, or Google Chrome). Anyone with an authentic username and password can view the data without installing client software. Furthermore, plant staff, analyst and equipment and maintenance operators and technicians can simultaneously look at the data to collaborate on remedial actions (Baxter & De Jesus, 2006).

Rather than delivering machine data in-house only (on-site control rooms), the Internet offers information to be accessible from anywhere, anytime. One such system is implemented on a Nucor mill, Arkansas (Rainey & Henderson, n.d.). They benefit from the system by allowing the monitoring and reading of the machine condition to be carried out without the physical presence of plant personnel on the site. Any changes on the machine condition such as its load, speed and operation will be monitored from a secure workstation or a control room using the local area network connection.

Improved Interface System

An alternate system would be developed using a high level programming language. In this case the alternate system should also be accessible from the internet. So Java programming language would be a good choice as we can create Java applets and also because it is platform independent. In addition, Java offers good possibilities for programming numerical calculations and creating visualization suitable for simulations (Bistak, 2009). These simulations can run online, thus giving the operator the chance to react with it. A diagram in Fig.3 (*Remote Monitoring System*, 2006) illustrates the plant boiler operator's experience with the improved interface.

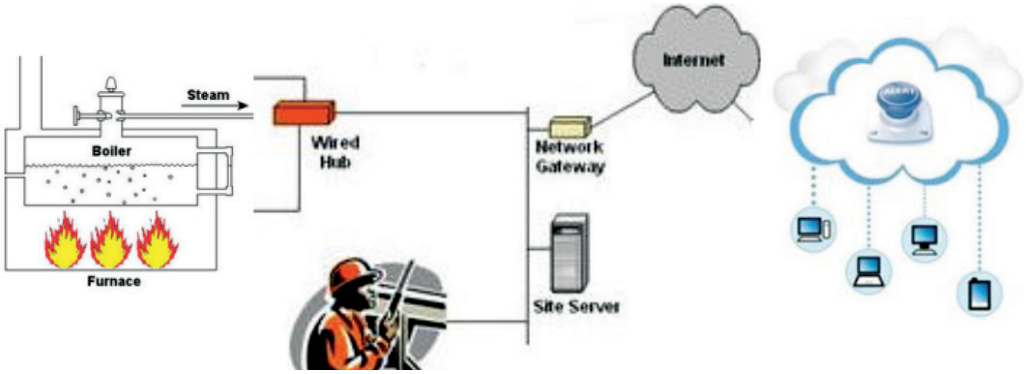


Fig.3: An improved interface monitoring system

The concept of Object Oriented Programming (OOP) became popular in the 90s when the internet hit the public domain. One of the most sought after programming languages for embedded system development is JAVA. Its distinctive features are unique and flexible allowing the developed system to be executed on any operating system platform. These features include robustness, extensive compiling and run time checking, distributed and secure, portability, high performance, dynamic and multithread (Gosling, 2006).

However, with all its high recognition of advantages, there are limitation to the real time performance and feedback. To overcome these issues, RTSJ has been specifically developed to allow JAVA to be used for real-time applications. This is achieved by using a scheduler to dispatch the execution of the logic periodically (Alhussian *et al*., 2012). With this continuous development, RTSJ can be considered as a starting point for JAVA to be widely used as a programmable real-time system such as the monitoring system of a thermal power plant boiler.

In addition, Ventura *et al*. (2002) stated in their paper that industrial automation is becoming increasingly decentralized, relying on distributed embedded devices to acquire and pre-process data and run increasingly sophisticated application software programs for control and self-diagnosis. Increasingly, these systems are employed in safety critical environment, where human life and health depends on the proper functioning of the computer systems. Hence, the safety and accuracy of these systems are of utmost importance. Therefore, they proposed the implementation of High Integrated Distributed Object Oriented Real-time Systems (HIDOORS) as an improvement to Java's shortcomings.

With automated, remote monitoring, health metrics are collected at whatever frequency is needed – without the expense or hazard associated with manual round. This paradigm is based on the concept that a plant will become more efficient if it is smarter about how equipment is monitored (Baxter & De Jesus, 2006).

Over the past few years, the internet technology has developed rapidly to ensure usability and accessibility is available for use in any industry. One of these advancements is the wireless technology. It has enabled plants to wirelessly collect data and make it available via the web browsers. Thus, a machine analyst can be at an airport, home or in a motel with a wireless internet connection and is able to view the machine condition and send correspondence to operations on whether the equipment will make it to the next outage.

Baxter and De Jesus (2006) highlighted the 10 reasons to embrace remote monitoring to ensure that a plant facility can operate more efficiently:

1. Spend more time analyzing data and less time collecting it
2. Collect data from previously inaccessible machine
3. Increase safety
4. Automated data collection
5. Collect additional metrics
6. Consistent data collection
7. Increase collection frequency for problematic machines
8. View plant-wide data
9. Setting of accurate alarms
10. Ability to monitor supervisory panels mounted on vital machines

CONCLUSION

Many approaches to the development of an intelligent monitoring systems used Matlab environment. This is due to its usefulness in designing an ANN model which can be trained to predict the expected performances of a plant's facility. The data collected from these monitoring systems are then recorded and stored on an on-site server for analysis and maintenance. On the other hand, a remote server interface application utilizes the internet to allow monitoring of plant performance to be accessible anytime, anywhere via the internet. Such systems are commonly implementing the Visual Basic and C++ programming language to develop its interface. The

existing monitoring systems can be improved by implementing the proposed alternate interface with a high level programming language such as Java. In this case, the alternate system should also be accessible from the internet. So Java programming language would be a good choice as we can create Java applets and also because it is platform independent. By combining the intelligent feature of trainable networks using ANN and Java's portability and secured features, a plant with a remotely accessible, interactive and intelligent monitoring system will become more efficient.

ACKNOWLEDGEMENT

The authors would like to acknowledge the GMG Power Station for the overall support and assistance in this paper.

REFERENCES

Alhussian, H. N., Zakaria, F. A., Hussin, & H. T. Bahbouh. (2012). A Review of The Current Status of Java Programming on Embedded Real-Time Systems. *Computer and Information Science (ICCIS). 2012 International Conference on Computer and Information Science* (p. 836 – 842). Kuala Lumpur, Malaysia.

Alnaimi, F. B. I., & Al-Kayiem, H. H. (2010). Multidimensional Minimization Training Algorithms for Steam Boiler Drum Level Trip using Artificial Intelligent (AI) Monitoring System. *Intelligent and Advanced Systems. 2010 International Conference on Intelligent and Advanced System (ICIAS)* (pp. 1 – 6). Kuala Lumpur, Malaysia.

Baxter, N., & De Jesus, H. (2006). *Remote Machine Monitoring: A Developing Industry*. Retrieved from http://azimadli.com/training-resources/technical-papers/

Bistak, P. (2009). Matlab and Java based virtual and remote laboratories for control engineering. *17th Mediterranean Conference on Control & Automation* (pp. 1439 – 1444). Thessaloniki, Greece.

De, S., Kaiadi, M., Fast, M., & Assadi, M. (2007). Development of an Artificial Neural Network Model for the Steam Process of a Coal Biomass Co-fired Combined Heat and Power (CHP) plant in Sweden. *Energy, 32*, 2099 – 2109.

Fast, M., & Palmé, T. (2010). Application of Artificial Neural Networks to the Condition Monitoring and Diagnosis of a Combined Heat and Power Plant. *Energy, 35,* 1114 – 1120.

Gosling, J. *et al.* (2006). *The Java programming language* (4th revised ed.). Addison-Wesley.

Haykin, S. (1999). *Neural networks: a comprehensive foundation* (2nd Ed.). New Jersey: Prentice Hall.

Kljajić, M., Gvozdenac, D., & Vukmirović, S. (2012). Use of Neural Networks for modelling and predicting boiler's operating performance. *Energy*, 1 - 8.

Principe J. C., Euliano, N. R., & Lefebvre, W. C. (2000). *Neural and adaptive systems*. New York: John Wiley & Sons Inc.

Rainey, D. E., & Henderson, J. (n.d.). *How a Nucor mill benefit from condition monitoring.* Retrieved from http://www.reliableplant.com/Articles/Print/27552

Remote Monitoring System [Image]. (2006). Retrieved from http://www.reliableplant.com/Read/27552/Nucor-benefits-condition-monitoring

Romeo, L. M., & Gareta, R. (2006). Neural Network for Evaluating Boiler Behaviour. *Applied Thermal Engineering, 26*, 1530 – 1536.

Smrekar, J., Assadi, M., Fast, M., Kuŝtrin, I., & De, S. (2009). Development of artificial neural network model for a coal-fired boiler using real plant data. *Energy, 34,* 144 – 152.

Thermal boiler [Image]. (n.d.). Retrieved from http://coalfiredpowerplants.blogspot.com/p/thermal-boilers.html

Ventura, J., Siebert, F., Walter, A., & Hunt, J. (2002). HIDOORS – A high integrated deterministic Java environment. *Proceedings of the 7th International Workshop on Object-Oriented Real-Time Dependable Systems (WORDS 2002)* (p. 113 – 118). San Diego, California, U.S.A.

Pertanika J. Sci. & Technol. 22 (2): 603 - 611 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Storm Runoff Pollution from a Residential Catchment in Miri, Sarawak

Ho, C. L. I.[1]* and Choo, B. Q.[2]

[1]*School of Engineering and Science, Curtin University Sarawak Campus, Miri, Sarawak, Malaysia*

[2]*Frasers Centrepoint Limited Management Services Pte Ltd., Singapore*

ABSTRACT

In this paper, stormwater runoff from a residential catchment located in Miri, Sarawak, was characterized to determine the pollutant concentrations and loading. The observed average event mean concentrations were 116 mg/L for TSS, 115 mg/L for COD, 1.5 mg/L for NH_3-N, and 0.23 mg/L for Pb. Based on Interim National Water Quality Standards (INWQS) for Malaysia, the average event mean concentration, EMC value for TSS exceeded class II (50 mg/L), exceeded class V (>100 mg/L) for COD, and exceeded class III (0.9 mg/L) for NH_3-N. All four water quality parameters exhibited first flush characteristic but to varying magnitude which was influenced by the storm characteristics.

Keywords: Storm runoff, residential catchment, NPS pollutant

INTRODUCTION

Developing countries such as Malaysia are characterized by rapid urbanization. In an urbanized area, anthropogenic activities and land use highly influence storm runoff volume and quality. The urbanization process results in an increase in imperviousness of the surface area and it also changes natural land covers as well as the drainage network. This results in an increase in runoff volume and produces hydrographs which have higher and earlier peaks (Elliott & Trowsdale, 2007; Gupta & Saul, 1996). Urban storm runoff has also been identified as one of the leading causes of degradation in the quality of receiving waters (USEPA, 2002). Rainfall washes off pollutants from the air and surfaces, carried into the drainage systems, and eventually leading to water contamination and river pollution (McLeod *et al.*, 2006). In many cities in the tropics, urban stormwater runoff is a major contributor to river pollution

Article history:
Received: 17 October 2013
Accepted: 18 January 2014

Email addresses:
Ho, C. L. I. (carrie.ho@curtin.edu.my),
Choo, B. Q. (kevincbq@gmail.com)
*Corresponding Author

ISSN: 0128-7702

(Nazahiyah *et al.*, 2007; Novotny & Olem, 1994). However, studies on urban runoff are still very limited, particularly in Malaysia, and the research is non-existent in Miri, Sarawak.

Urban runoff pollution can be broadly categorised into point source (PS) and non-point source pollution (NPS). Major point source pollution includes effluent from a wastewater treatment plant and industrial discharges. Non-point source pollution is not readily identifiable and the pollutants are transported in a diffuse manner. Typical non-point source pollutants that can be found in urban storm runoff include organic materials, suspended solids and nutrients such as ammonia nitrogen and phosphorus and heavy metals (Tsihrintzis & Hamid, 1997).

In Malaysia, development of preventive strategies to mitigate the impacts of NPS pollution from urban runoff is still in its infancy. An Urban Stormwater Management Manual for Malaysia has been developed and some guidelines are presented for Best Management Practices (BMPs) for control of urban stormwater (DID, 2000). However, formulation of effective management strategies is hampered by the lack of data on the diffuse pollutant loads. In Miri, Sarawak, the contribution of pollutants from urban runoff into the receiving water has not been well understood and limited data are available on urban runoff pollutants.

Miri is a fast developing city, characterized by rapid urbanization in the last decade. Large tracts of land are largely converted into residential developments. Waterways located within the city are heavily polluted, with Miri River, which enters the South China Sea, being one of the most polluted rivers in Sarawak. However, the contribution of NPS pollution to the quality of the surface water is still largely unknown.

The amount and types of pollutants found in storm runoff vary according to the catchment's predominant land use and associated human activities. Rainfall characteristics such as intensity and size will also impact pollutant transport (Zoppou, 2001). The variability of source loadings for urban runoff necessitates the sampling of storm events and analysis of runoff pollutants to estimate pollutant loading to the receiving water courses. The objectives of this study, which was conducted in Miri, Sarawak, were to characterize the storm runoff pollution from a residential area and to analyze the first flush effect of the storm runoff pollution.

MATERIALS AND METHODS

Study Area

The study site is located in Miri, Sarawak. The climate is tropical, with an annual precipitation ranging between 2800 and 3200 mm. The residential catchment is 9.8 ha in size, with no significant slope in the terrain. The residential houses are single storey terraced houses, high in density, which results in low perviousness of about 13%. Impervious areas include paved roads, concrete pavements, and roofs.

As-built drainage drawings were used to delineate the catchment boundary, according to the direction of flow. The drainage channels are open, rectangular in shape and constructed using concrete, which is typical for Sarawak. The drainage system conveys both stormwater runoff and wastewater effluents from individual septic tanks from the houses. Any washoff from the surface, as a result of human activities such as from car wash and paved surface cleanings also discharges into the drainage channels. Runoff from the catchment under study outfalls into an

existing retention pond, located north of the study area. Fig.1 shows the catchment boundary and the location of the retention pond.

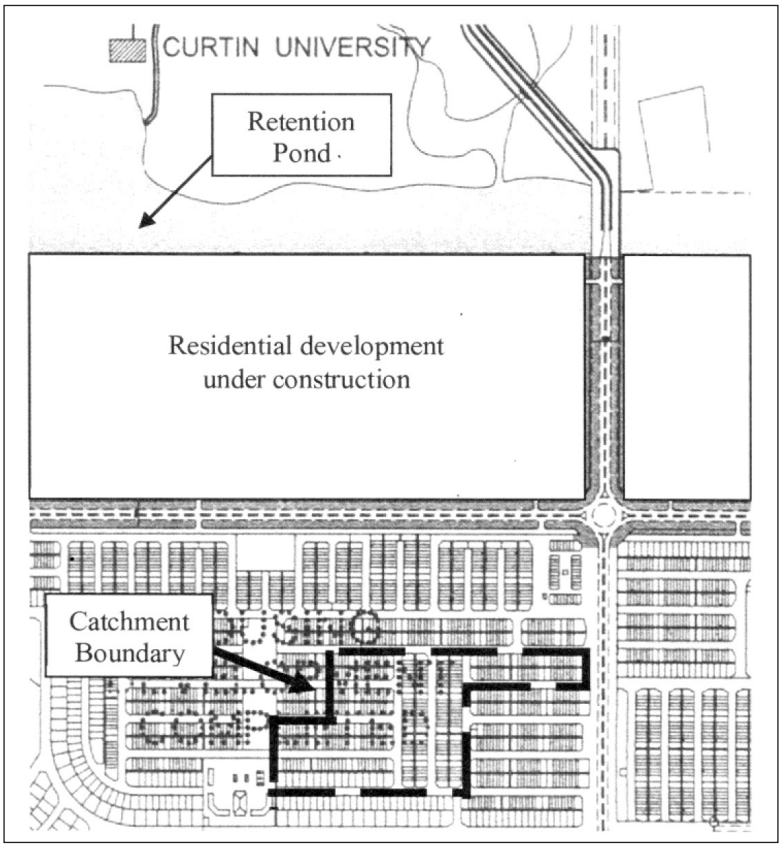


Fig.1: Study catchment in Miri, Sarawak

Rainfall and Runoff

Rainfall was monitored continuously using a tipping bucket rain gauge (HACH) with a data logger. Rainfall is recorded every two minutes. Runoff discharge and volume was determined using the velocity-area method. During sampling, a velocity meter was used to measure the velocity of the flow at the catchment outlet. The runoff discharge was calculated as the product of the flow velocities and the cross section area of the drainage channel at various water levels measured.

Sampling and Analysis

The samples were collected from 3 rainfall events from August to December 2011. The samples were collected at the outlet of the catchment. A total of 10 samples were collected for each of the rainfall events monitored at various intervals throughout the hydrograph. Descriptions of the events sampled are summarized in Table 1. Water quality parameters tested for in this study

include total suspended solids (TSS), chemical oxygen demand (COD), ammonia nitrogen (NH_3-N), and lead (Pb).

TABLE 1
Rainfall events monitored in this study

	1st Event	2nd Event	3rd Event
	28th Aug	20th Sept	28th Dec
Rainfall depth (mm)	8	103	2.8
Rainfall intensity (mm/h)	6.3	41	21
Max. rainfall intensity (mm/h)	30	145	38
Rainfall duration (h)	0.63	2.5	0.13
Antecedent dry period (h)	167	82	19

RESULTS AND DISCUSSION

Temporal Variation of Pollutants

A set of hydrograph and pollutograph for each of the water quality parameters tested can be produced based on the runoff monitored. Fig.2 and Fig.3 show the precipitation and the development of the flow and concentrations of TSS, COD and Pb during the 1st and 2nd storm events monitored in this study. Event 1 was a short but intense rainfall, characteristics of a tropical thunderstorm. Event 2 was a heavy storm, totalling 103 mm, with maximum rainfall intensity of 145 mm/h. Event 1 was high intensity at the start of the event, and event 2 was high intensity at the later period of the event.

In both events, pollutant concentration peaked before maximum flow occurred. Pollutant concentration rapidly decreased after the peak. It can be observed that event 2, which had a much higher rainfall intensity and duration than event 1, resulted in high TSS being sustained over the first half period of the event, whereas for event 1, TSS rapidly reduced from the initial high concentration. This finding indicates that the characteristic of the rainfall events greatly influences the pollutant generation pattern.

Pollutant Load Distribution

Concentration of pollutants found in storm runoff can be highly variable over the duration of the storm. Thus, an event mean concentration (EMC) is often used to evaluate the impacts of stormwater runoff on receiving waters. An EMC represents a flow-weighted average concentration of runoff computed as the total storm pollutant load divided by the total runoff volume, expressed by Equation 1.

$$EMC = \frac{\sum_{i=1}^{n} C_i Q_i}{\sum_{i=1}^{n} Q_i} \quad [1]$$

C (unit is mg/L) is the time-variable concentration of the pollutant and Q (unit is m^3/s) is instantaneous discharge.

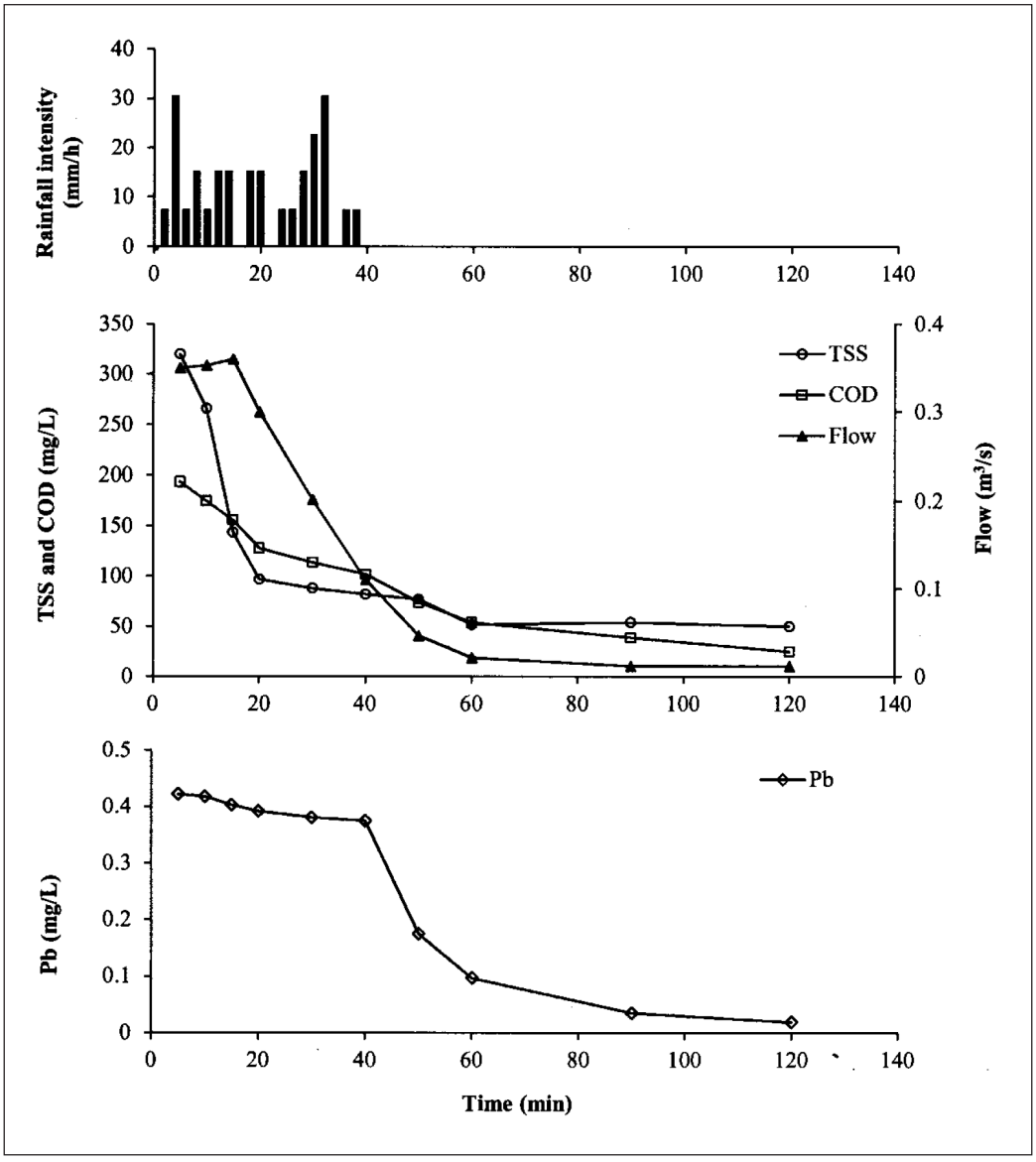


Fig.2: Flow and concentration of TSS, COD, and Pb for Event 1

Table 2 presents EMCs and loads from the three events monitored. Average EMC values were 116 mg/L for TSS, 115 mg/L for COD, 1.5 mg/L for NH_3-N, and 0.23 mg/L for Pb. It can be seen that the EMC values for the pollutants vary depending on the characteristics of the rainfall. Meanwhile, the EMC values for TSS and NH_3-N increased with the length of the antecedent dry period. Event 1 had the longest antecedent dry period of 7 days, followed by event 2 with 3.5 days, and event 3 with <1 day. The EMC values for the pollutant COD, however, showed correlation with the size of the rainfall, with event 2 having the lowest EMC value for COD, but the largest rainfall at 103 mm. The large rainfall event would have diluted

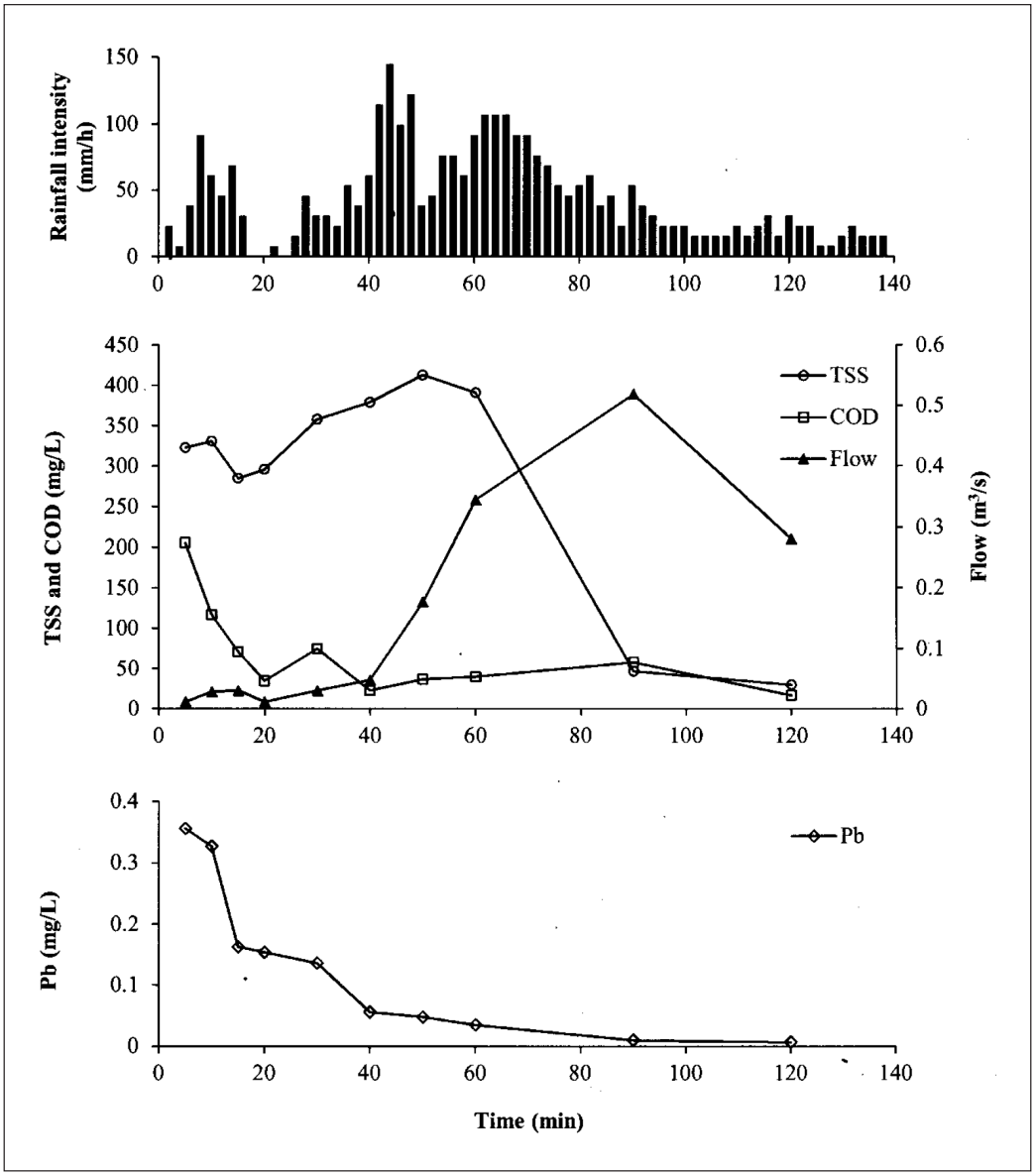


Fig.3: Flow and concentration of TSS, COD, and Pb for Event 2

the strength of the pollutant. The average EMC value for COD exceeded class V (>100 mg/L) of the Interim National Water Quality Standards for Malaysia (INWQS). The average EMC value for TSS exceeded class II (50 mg/L) and the average EMC value for NH_3-N exceeded class III (0.9 mg/L) of the Malaysian INWQS. This shows that the quality of the runoff from the study site is slightly polluted. However, it does result in significant contaminant loading to the receiving water.

TABLE 2
EMCs and total discharged loads for monitored runoff

Pollutants	1st Event		2nd Event		3rd Event	
	EMC (mg/L)	Load (kg)	EMC (mg/L)	Load (kg)	EMC (mg/L)	Load (kg)
TSS	158	107	115	209	76	3
COD	135	92	44	79	167	7
NH_3-N	2.17	1.5	0.89	1.6	-	-
Pb	0.36	0.25	0.02	0.04	0.31	0.01

In urban runoff, pollutant load delivery is typically not proportional to the amount of runoff volume. Pollutant load is often significantly higher during the early portion of the storm runoff and this characteristic is known as the first flush phenomenon. One widely used method to identify first flush is by plotting a dimensionless curve of cumulative discharge volume against cumulative discharge mass (Bedient *et al.*, 1980). A 45^o bisector line is also drawn and first flush is considered present when the initial slopes of the dimensionless curves are greater than the 45^o bisector.

From Fig.4, the plot of the distribution of pollutant load with volume showed that all the pollutants exhibit the first flush phenomenon, but to varying degrees. The dimensionless curves of all four water quality parameters studied were largely above the bisector line for all three events. This suggests that large fractions of the pollutant load were transported by the initial portion of the runoff. The difference between the curves and the bisector line can be used to indicate the magnitude of the first flush. TSS showed a strong first flush characteristic for all three storm events, whereas COD exhibited a weak first flush characteristic. NH_3N and Pb both exhibited a strong first flush characteristic for one storm event only.

The characteristic of the rainfall event also influences the magnitude of first flush, as it can be seen that for event 2 (20 September), TSS, NH_3-N, and Pb exhibited a very strong first flush characteristic. Event 2 had an initial high intensity rainfall which washed off the pollutants, followed closely by the maximum intensity which resulted in the peak hydrograph in the later period of the event, and thus resulted in the distinctive first flush for the pollutants tested.

All four water quality pollutants exhibited first flush for event 1 (28 August), which was a short, but high intensity rainfall. The magnitude of the first flush, however, is weak. As for event 3 (28 December), the pollutants exhibited weak or no first flush, as the rainfall event was small and of short duration, at only 2.8 mm over 10 minutes. This indicates that the presence and strength of first flush is dependent on the characteristic of the rainfall.

CONCLUSION

The NPS pollutant load and the distribution of the load with volume are determined in this study. The results derived from this study are useful to estimate the contribution of runoff pollutants from catchments with residential land use to the waterways in and around Miri City. As Miri is a rapidly expanding city with significant land developments planned for residential housing, the potential NPS loadings due to the developments and its potential impact on surface water can

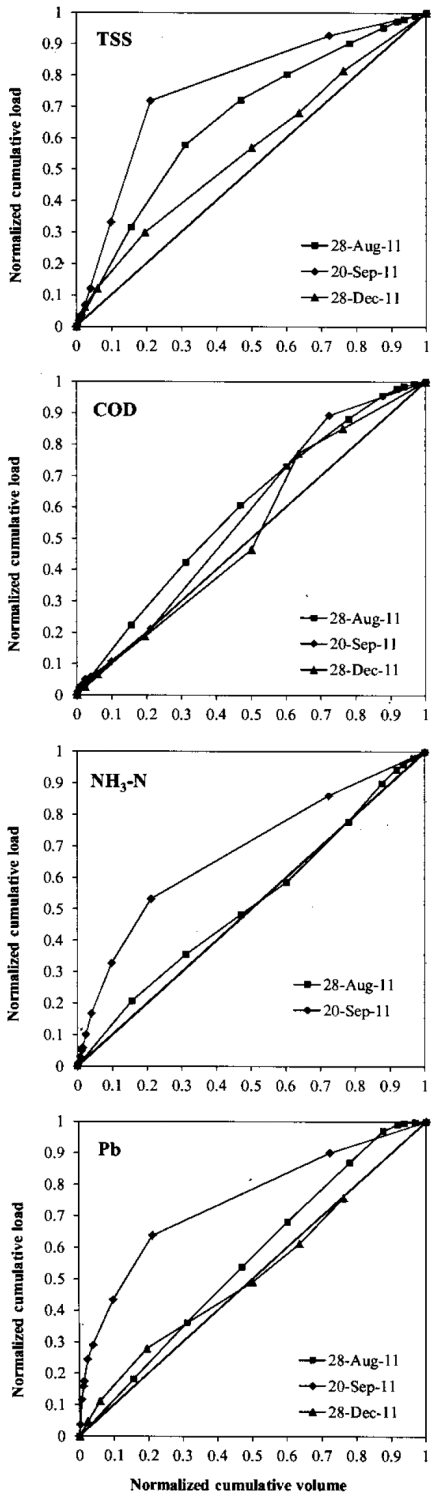


Fig.4: Normalized cumulative curves for TSS, COD, NH3-N, and Pb of the three storm events monitored

be estimated. In this study, observed average EMC values were 116 mg/L for TSS, 115 mg/L for COD, 1.5 mg/L for NH_3-N, and 0.23 mg/L for Pb. All the four water quality parameters exhibited first flush characteristic, but the magnitude of the first flush varied depending on the rainfall characteristic. In conclusion, the study of storm runoff and its associated pollutants is essential to understand the contribution of urban runoff pollutants to the degradation of surface water quality, and an important step in the design of management strategies to address the issue of deterioration of the water quality of rivers in Sarawak.

ACKNOWLEDGEMENTS

The authors would like to acknowledge United Consultant, Miri, for the information provided in carrying out this study at Desa Senadin, Miri.

REFERENCES

Bedient, P. B., Lambert, J. L., & Springer, N. K. (1980). Stormwater pollution load-runoff relationship. *Journal of Water Pollution Control Federation*, *52*(9), 23962404.

Department of Irrigation and Drainage Malaysia (DID). (2000). *Urban stormwater management manual for Malaysia*.

Elliott, A. H. & Trowsdale, S. A. (2007). A review of models for low impact urban stormwater drainage. *Environmental Modelling & Software*, *22*(3), 394405.

Gupta, K., & Saul, A. J. (1996). Specific relationships for the first flush load in combined sewer flows. *Journal of Water Research*, *30*(5), 12441252.

McLeod, S. M., Kells, J. A., & Putz, G. J. (2006). Urban runoff quality characterization and load estimation in Saskatoon, Canada. *Journal of Environmental Engineering*, *132*(11), 14701481.

Nazahiyah, R., Yusop, Z., & Abustan, I. (2007). Stormwater quality and pollution loading from an urban residential catchment in Johor, Malaysia. *Water Science and Technology*, *56*(7), 19.

Novotny, V., & Olem, H. (1994). *Water quality: prevention, identification, and management of diffuse pollution*. Van Nostrand Reinhold.

Tsihrintzis, V. A., & Hamid, R. (1997). Modeling and management of urban stormwater runoff quality: a review. *Water Resources Management*, *11*, 137164.

US EPA (2002). *2000 National Water Quality Inventory Report to Congress*. Office of Water, Washington, DC.

Zoppou, C. (2001). Review of urban storm water models. *Environmental Modelling & Software*, *16*, 195231.

Pertanika J. Sci. & Technol. 22 (2): 613 - 625 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Septage Treatment Using Pilot Vertical-flow Engineered Wetlands System

Jong, V. S. W. and Tang, F. E.*

Department of Civil and Construction Engineering, School of Engineering and Science, Curtin University Sarawak, Miri, Malaysia

ABSTRACT

This paper presents a two-staged, pilot-scale vertical flow engineered wetland-based septage treatment system (VFEWs), which was designed and constructed in Curtin University Sarawak Campus to determine the system efficiency in treatment of septage. The treatment system consists of storage tanks, vertical flow wetlands, and a network of influent and effluent distribution pipes. The first stage of the VFEWs treatment system consists of three vertical flow wetlands placed in parallel to provide pre-treatment to raw septage to reduce solids and organic matters mainly by physical filtration and sedimentation processes. The percolate from the first stage is then further treated in the second stage, with four vertical flow wetlands, each with variation in operational regime and substrate (filter) type. The influences of various system and application-related parameters such as substrate material, presence of plants and plant types, and septage feeding practices (solid loading rate (SLR), batch and intermittent loading, and frequency of daily feeding) on pollutant removal efficiency were studied. Results from the first stage wetlands indicate that the removal of total solids and organic matter (BOD and COD) from the raw septage is promising (> 80%) at both SLR of 100 kg TS/m^2.yr and 250 kg TS/m^2.yr, respectively. However, a higher SLR decreased the average NH_3-N removal efficiency. The findings on bed clogging assessment during the study period are also presented in this paper. Validation and expansion of these results are carried out with ongoing assessments on the system performance.

Keywords: Vertical-flow, engineered wetlands, septage, dewatering, substrate materials, plants, feeding regimes, removal efficiency;

Article history:
Received: 23 October 2010
Accepted: 18 January 2014

Email addresses:
Jong, V. S. W. (Valerie.jong@curtin.edu.my),
Tang, F. E. (Tang.fu.ee@curtin.edu.my)
*Corresponding Author

INTRODUCTION

Engineered wetlands (EWs) are rapidly emerging as a feasible method for treatment of wastewater due to their low investment costs and ease of operation and maintenance. In many sites, EWs are considered economically

ISSN: 0128-7702

and ecologically viable alternatives to conventional methods of wastewater treatment. Engineered wetlands have been used in many applications, ranging from the secondary treatment of domestic, agricultural and industrial wastewaters to the tertiary treatment and polishing of stormwater and wastewater treated by means of activated sludge plants. Factors affecting the performance of an engineered wetland system are generally dependent on a variety of operational factors relating to the system itself and the influent characteristic, as well as the way it is applied to the bed (Prochaska et al., 2007). System-related factors include substrate type, size and depth (Torrensa *et al*., 2009), maturity of bed and climate (Merlin *et al.*, 2002). Other factors could be type of vegetation, system configuration and the sizing of the bed. The application-related factors include the hydraulic loading rate (HLR), sludge loading rate (SLR), influent concentration and the influent feeding regime. From literature, most studies published over the past five decades reported on the performance of engineered wetlands in the temperate or subtropical regions. In the recent years, however, there is an increase in research interests on exploring the potential of using EWs to treat wastewater in the tropical regions. Their elevated temperatures were found to be beneficial in increasing the treatment efficiency of the system comparing to the non-tropical systems. EWs in tropical regions has been reported to show organic and nutrient removal rates at almost a factor of 10 higher than in temperate regions (Diemont, 2006).

Engineered wetlands have also been used to treat high-strength organic wastes such as septage, a by-product of on-site wastewater sanitation system. Engineered wetlands designed for the treatment of septage feature a combination of traditional sludge drying beds with natural wetlands, which have been productively used for solids dewatering and stabilization in small cities across Europe and Asia (Burgoon *et al*., 1997; Cooper *et al*., 1996; Kengne *et al*., 2009). Malaysia produces over six million cubic meters of raw sewage and septage annually and this results in over 100,000 tonnes of stabilized sludge each year (AECOM International Development & Sandec/EAWAG, 2010). Treatment and disposal of septage has become a problem due to its rich pollutants concentration and being highly heterogeneous. Besides, conventional septage treatment is non-existent in many small cities and suburban areas in Malaysia, and septage is managed by disposal to nearby water courses. This management method causes pollution and damage to the receiving water body.

Several researchers (see Koottatep *et al*., 2005; Nielsen, 2003) have suggested the use of vertical flow engineered wetlands (VFEWs) to treat septage as a cost-effective and technically feasible approach for septage dewatering, stabilisation and mineralisation. Vertical flow engineered wetlands (VFEWs) are flat beds comprise of graded gravels or aggregates topped with or without sand and planted with vegetation (macrophytes). VFEWs are fed either intermittently in batches or continuously with influent wastewater onto the bed surface. Wastewater will percolate down through the substrate and be collected by a drainage network at the bottom. There are relatively few studies focusing on septage treatment with EWs in tropical regions. Koottatep T. from the Asian Institute of Technology (AIT), Thailand had studied the treatment of faecal sludge or septage by means of engineered wetlands (Koottatep & Polprasert, 1997; Koottatep *et al*., 2001a; Koottatep *et al*., 2005; Panuvatvanich *et al.*, 2009).

From Koottatep's research, engineered wetlands have been found to be a promising and stable technology for septage treatment in tropical regions (Koottatep *et al*., 2001a). In one of his studies, 25 m^2 of pilot-scale wetland beds filled with 0.65 m sand-gravel substrate and planted with narrow-leaf cattails (*Typha augustifolia*) were used to treat septage collected from the Bangkok city, Thailand. Solid loading rate (SLR) of 80–500 kg TS/m^2.yr was applied every once to twice weekly onto the beds and an optimal SLR of 250 kg TS/m^2.yr were found to result in the highest total solid (TS), total chemical oxygen demand (TCOD) and total Kjeldahl nitrogen (TKN) removal. Further, a yard-scale experimental plant in Yaounde (Cameroon) was constructed and studied by Kengne to evaluate the treatment of VFEWs planted with *Echinochloa Pyramidalis* on faecal sludge dewatering ,beside experimenting on the effects of different SLRs on growth of the wetland macrophytes (Kengne *et al*., 2009). The study revealed that the system performed well for solid–liquid separation at loading rate of 100–200 kg TS/m2/yr, with an average dry matter content of biosolids ≥30% and pollutant removal efficiencies higher than 77%, 86%, 90%, 90% and 95% for ammonia (NH_4^+), total suspended solids (TSS), TS, TKN and COD, respectively.

In this paper, an engineered wetland-based pilot septage treatment system, which is built and still operating in Curtin University Sarawak Campus, Malaysia, is presented. Components of the treatment system are described along with the materials used and the septage application regimes practiced are explained. While the study is underway, the results presented in this paper focused only on the effects of variation of SLR on the first 8 weeks of plant operation of two of the pilot EW beds which act as the first stage of treatment in the system.

MATERIALS AND METHODS

Description of the Pilot-scale Treatment System

The project site is located beside the wastewater treatment plant in Curtin University Sarawak Campus in Miri, Sarawak, Malaysia. The system consists of pilot-scale vertical flow engineered wetlands (VFEWs) set in an open field that is exposed to sunlight and wind. A semi-transparent roof was constructed to shelter the system from rainfall to prevent stormwater from disturbing and affecting the experimental output due to dilution. The system is a 2-staged treatment comprises of vertical flow wetlands and storage tanks, as shown in Figure 1.

In the first stage, there are three VFEWs, each with a surface area of approximately 2.2 m^2 and a height of 1.3 m. Septage is stored in two elevated receiving tanks and is gravity-fed onto the first stage wetlands (wetland As) on a weekly basis through 3-inch PVC perforated pipes complete with stopcocks. The resulting filtrate is stored in the effluent collection tank before being piped onto the second stage wetlands (wetland Bs) for further treatment. Modified pumps and timers are used at the second stage to control the influent feeding frequency and volume of loading for each session. The macrophytes used in the study are *Phragmites Karka*, a common reed found in abundance in the local swamp areas. All the wetlands are equipped with vent pipes to prevent anaerobic environment in the substrate.

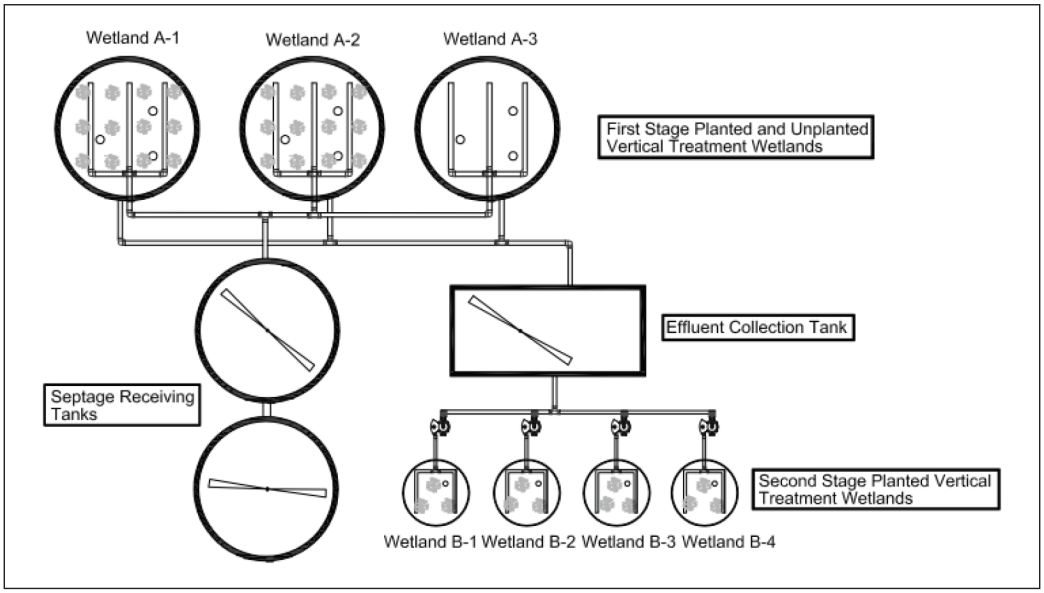


Fig.1. A schematic diagram of the pilot VFEWs system

Septage supply for the plant is obtained from a local environmental service provider and sourced from residential and communal septic tanks which receive only blackwater. Vacuum trucks loaded with raw septage visited the project site every 1-2 weeks to deliver the septage for the study. Coarse materials (>32mm) are removed from the septage by manual filtration using sieves to screen out the gross solids before storage. The septage is stored in two 400-gallon receiving tanks fitted with a mechanical mixer each to stir the septage to uniformity before being loaded onto the wetlands.

Crushed limestone is used as substrate in the vertical filter beds. The aggregate is a common construction material which is easily available locally for road construction. The total depth of the substrate in wetland As is 800mm, with 500mm freeboard for sludge accumulation. From bottom to top, the substrate filter consists of a 200mm layer of coarse aggregates (30–50mm diameter), a 300mm layer of medium aggregates (10-30mm diameter), and a 300mm layer of fine aggregates (3-8mm diameter). Wetland Bs are filled with limestone aggregates of 50mm thickness as the drainage layer (20-30mm diameter), overlaid by a 300mm thick intermediate stratum of palm kernel shell (PKS), topped with 300mm of pea gravel (3 mm) and finally covered by a layer of sand.

System Operational Regime

Wetland A1 – A3 are fed once in 7 days with nominal septage loads after the 6 months commissioning period. The septage is gravitationally loaded onto the beds for primary treatment before the resulting percolate is directed into an effluent collection tank for storage prior to further treatment by the subsequent wetlands (wetland B1 - 4). In the first stage wetlands, septage is applied in batches, with the wetlands receiving the influent in the range

of 0.10 - 0.35 m3 (with respect to the designed SLR) in one go and within approximately 15 minutes. The outflow of the wetlands is controlled by stopcocks and water taps at the bottom of the basins. Table 1 indicates the operating conditions (Solid Loading Rate, SLR) at the first stage wetlands. Table 2 indicates the operating conditions at wetlands B1 – B4. The wetlands are loaded with percolate from the first stage, with the varying HLR as shown. Under the experimental regime, the wetlands are operated with intermittent loading (4 and 8 times of daily feedings) or with batch loading (up to 3 days flood : 3 days rest period). Wetlands are planted with either *Phragmites Karka* or *Costus Woodsonii* and contain substrate that is with or without PKS. At the time of writing, the experiments are still underway and only the partial results from operating the first stage wetlands are being presented in this paper.

TABLE 1
Operating conditions of the pilot VFEWs system (1st stage of the treatment)

Wetland	Description
A-1	100 kg TS/m^2. Yr (Planted)
A-2	250 kg TS/m^2. Yr (Planted)
A-3	250 kg TS/m^2. Yr (Unplanted)
All wetlands are filled with graded aggregates and vegetated wetlands planted with *phragmites karka*	

TABLE 2
Operating conditions of the pilot VFEWs system (2nd stage of the treatment)

Wetland	B1-B4	
Feeding and Draining Pattern	Intermittent Loading (Free drainage) No of daily feedings: 4, 8	Batch Loading (Cyclic effluent ponding and draining) Flood and Rest (Days): 1:1, 2:2, 3:3
Substrate	Sand and aggregates only; Sand and aggregates with PKS	
Macrophytes	Reeds: Phragmites Karka; Ornamentals: Costus Woodsonii	
HLR (cm/d)	8.75; 17.5	

PKS= Palm kernel shell
HLR = Hydraulic Loading Rate

Monitoring System Performance

The influent and effluent of each wetland at every stage were collected and analyzed weekly for the following water quality parameters: total solids (TS), total suspended solids (TSS), total volatile solids (TVS), total volatiles suspended solids (TVSS), chemical oxygen demand (COD), biochemical oxygen demand (BOD), total nitrogen (TN), nitrite-nitrogen (NO_2-N), nitrate-nitrogen (NO_3-N), and ammonia-nitrogen (NH_3-N) with HACH DR 2800 spectrophotometer based on USEPA approved standard procedure for wastewater analyses. Removal efficiencies are obtained by calculating the influent pollutant load reduction from the effluent at each

treatment stage. *In situ* tests such as pH, temperature, dissolved oxygen (DO), oxidation-reduction potential (ORP) and electric conductivity (EC) were also carried out weekly on the influent before each feeding and effluents immediately after the collection.

A general assessment on clogging phenomena was carried out based on the visual observation as preliminary evaluation. The liquid fraction of the septage applied on the wetlands can infiltrate within a few minutes for well-drained substrate, to more than a week if the substrate is clogged. An observation was carried out on how noticeable the presence of standing water was on the beds days after septage application. A maximum of 3 days period for the draining of the liquid fraction was considered for well-drained substrate, 4 to 6 days for beds that were slightly clogged and more than 7 days for clogged beds.

For analysis of the collected results, one-way ANOVA tests were conducted to assess the effects of solid loading rate (SLR) on the pilot system's pollutant removal efficiencies. Post Hoc Multiple Comparisons test were also performed using Tukey's family error rate when necessary. The level of significance was set at $P<0.05$. The statistical software used was SPSS statistics 17.0 for Windows.

RESULTS AND DISCUSSIONS

Miri's Septage Characterisation

Septage is the combination of sludge, scum and liquid pumped from septic tanks. It is greatly heterogeneous and its concentration and organic matters content depend highly on users' habits, climate, septic tank size, and emptying frequency. Table 3 presents the characteristic of the raw domestic septage delivered to our VFEWs pilot plant, whereas Table 4 shows the characteristic of septage from different countries retrieved from the literature.

Our septage exhibited on average anaerobic or reducing conditions (<-200 mV) with high solids content and high organic and nutrient concentrations. Capillary suction time (CST) is the fundamental measure of the filterability and ease of removing moisture from sludge. Measured CST values for Miri's septage indicate a higher dewaterability in comparison to the septage from Andancette, France (Vincent *et al*., 2011), as shown in Table 4. Septage particle size, SS concentration and biochemical compositions (fats, protein and polysaccharides) are among factors that could influence the CST measurements (Jin *et al*., 2004; Vincent *et al*., 2011). Miri's septage has on average BOD:COD ratio of 0.09, which indicates low biodegradability due to decomposition of most degradable fractions after long storage time in the septic tanks. The characteristic of the septage from Miri is similar to that of the septage from Bangkok, as reported by Koottatep (see Table 4).

Plants and Wetlands Acclimatization

Preliminary planting trials were carried out to decide between two common wetland macrophytes, common reeds (*Phragmites Karka)* and cattail (*Typha Latifolia)* to be planted in the pilot system. Both the species were transplanted from the nearby river banks and monsoon drains into small pots by stem cutting for *Phragmites* and rhizome cutting for *Typha* on July 2011. The plants were planted in 37.5 mm aggregates and ponded with tap water for the first

TABLE 3
Physicochemical characteristics of Miri's septage

Parameter	N	Range	Average	Std
CST (s)	6	85.5-158.9	107.9	27.03
COD (mg/l)	14	12,400-54,870	33,442.19	13,963.70
BOD_5 (mg/l)	14	672-8,740	3,315.54	2,544.98
NH3-N (mg/l)	14	68-695	353.91	190.25
TKN (mg/l)	14	280-1,657	988	482.93
NO_3-N (mg/l)	14	5.39-35.7	14.53	9.18
TN (mg/l)	14	279-1,660	988.79	474.63
TP (mg/l)	6	275.9-2754	1,081.82	861.37
Temperature (◦C)	14	27.5-30.1	28.89	0.85
EC (ms/cm)	14	1.04-2.36	1.53	0.44
pH	14	5.93-7.69	6.82	0.51
DO (mg/l)	14	0.06-0.30	0.13	0.07
ORP (mV)	9	-100 to -546	-231.63	151
TS (mg/L)	14	13,962-57,600	32,588.50	14,954.44
TVS (mg/l)	14	8,054-59,318	21,888.21	12,556.11
TSS (mg/l)	14	5,200-50,500	23,875.38	12,121.11

N= No. of samples

1 month, after which the pots were drained and the plants were consistently watered with tap water on daily basis. After 3 months of growing period, it was observed that *Phragmites* had a faster growing rate than *Typha* in terms of number of emergent new shoots and plant height. Besides, *Phragmites* require less water to survive well, in terms of having minimal signs of wilting during the unponded period compared to *Typha.* Thus, *phragmites* was selected as the wetland macrophyte in our VFEWs system.

TABLE 4
Physicochemical characteristics of Septage from various regions

Parameter	N	Thailand*	N	Ghana**	N	Cameroon#	N	France##
CST (s)	-	N/A	-	N/A	-	N/A	-	360±142
COD (mg/l)	30	N/A	-	8,400	42	31,100	-	42,000±13
BOD_5 (mg/l)	30	2225±395	-	3,700	-	N/A	-	N/A
NH_3-N (mg/l)	30	320±70.89	-	500	42	600	-	287±76
TKN (mg/l)	30	N/A	-	N/A	42	1,100	-	1,423±435
NO_3-N (mg/l)	30	4.81±1.65	-	N/A	-	N/A	-	N/A
TN (mg/l)	30	950±99.18	-	N/A	-	N/A	-	N/A
TP (mg/l)	30	N/A	-	N/A	-	N/A	-	517±438
Temperature (°C)	30	28.67±1.5	-	N/A	-	N/A		N/A
EC (ms/cm)	30	N/A	-	17.27	44	2.79	-	N/A
pH	30	7.48±0.5	-	7.7	44	7.5	-	N/A
DO (mg/l)	30	N/A	-	N/A	-	N/A	-	N/A
ORP (mV)	30	-291±30	-	N/A	41	-54.2	-	N/A
TS (mg/L)	30	22,420±7702.6	-	11,800.00	44	3.7 (% DM)	-	30,000±10.6
TVS (mg/l)	30	N/A	-	6,726.00	43	64.4 (% DM)	-	21,300±2,100
TSS (mg/l)	30	19,500±7,250	-	N/A	-	N/A	-	23,0300±8,600

N/A= Not Available
N = number of samples
* Characteristics of septage from Bangkok, Thailand (Koottatep *et al.*, 2005)
** Characteristics of septage from Kumasi, Ghana (Cofie *et al.*, 2006)
# Characteristics of septage from Yaoundé, Cameroon (Kengne *et al.*, 2008)
Characteristics of septage from Andancette, France (Vincent *et al.,* 2011)

An engineered wetland is essentially an ecological system that requires time to establish itself before commencement of operation in which the full load will be introduced onto the substrate. There are varying guidelines in relation to the length of acclimatization, and these range from two to six months in the tropical climate (Ahmed *et al*., 2008; Trang *et al*., 2010). In our system, a six-month period was employed where the wetlands were progressively fed with diluted sludge every 3 days. This is to allow for plants and microbial community establishment besides preventing plants nutrient shock.

System Performance with Different SLRs

System Percolate

The results obtained from the preliminary experiments performed with the system are shown in Table 5. Removal of total solids and organic matter with the reduction in COD and BOD from the raw septage was the highest amongst the main parameters tested. Based on these preliminary results, it appears that variation in SLR has no significant effect on the removal efficiencies of these parameters at the 95% confidence level. This is because physical sedimentation and filtration were the major mechanisms for the removal of organic matters and solids from the system influent. Koottatep claimed that at a short hydraulic retention time (HRT), COD removal generally depends on filtration capacity of the wetland units rather than biological degradation of the organic matter (Koottatep *et al*., 2001b).

TABLE 5
Experimental results showing pollutant concentration and removal efficiency of wetland A1 and A2 with different SLR

Sample	Wetland (SLR in kg TS/m^2.yr)	Parameter*, mg/L					
		COD	*BOD*	*NH_3-N*	*NO_3-N*	*TN*	*TS*
Influent (Raw Septage)		31,957.71	3,592.50	427.99	13.26	1,209.20	24,573.60
Effluent (Resulting Filtrate)	A1 (100)	890.42 (96.87)	285.93 (91.92)	80.83 (79.24)	20.45	212.08 (80.12)	2,681.88 (89.97)
	A2 (250)	1,634.38 (94.5)	392.84 (87.86)	92.45 (76.79)	10.43	393.04 (76.01)	3,090.88 (88.81)
	F-value	1.966	2.064	4.767	-	2.932	0.305
	P**	0.183	0.173	0.047	-	0.109	0.589

* Percentage of removal as shown in parentheses were based on an average of 8 sets of data obtained from the 8 weeks of experimental run
** Significant at $P < 0.05$

Based on the statistical analysis of the data collected over a period of 8 weeks which reflected the system performance for the first 2 months, NH_3-N removal efficiency was found to vary according to different SLRs. At a loading rate of 100 kg TS/m^2.yr, ammonia removal was 79.24% and is significantly higher than that of the wetland fed with SLR of 250 kg TS/

m^2.yr which yielded a removal of 76.79% with $P<0.05$. In comparison with assimilation of organic matter by heterotrophic bacteria, autotrophic nitrification is a relatively more sensitive process which requires environment with sufficient DO, suitable pH and temperature, and the presence of the specific bacteria colonies. As shown in Table 6, effluent collected from A1 had higher DO concentration than that of effluent from A2. At SLR of 100 kg TS/m^2.yr, the wetland was relatively more aerated with thinner sludge residuals retained on top of the bed that dried out more rapidly and thus promoting oxygen diffusion into the substrata and increase the septage infiltration time.

TABLE 6
In situ test readings on effluent for wetland A1 and A2

	DO (mg/l)	Temperature (°C)	pH	Electric Conductivity (mS/cm)
A1 (SLR 100)	3.24	28.2	7.12	2.33
A2 (SLR 250)	2.40	28.3	7.25	1.95

Both wetlands produced effluents with increased DO from an average of 0.13 mg/l in the raw septage to 3.24 and 2.40 mg/l for A1 and A2, respectively. This indicates that our vertical-flow beds with weekly batch feeding performs aerobically, which is important for an efficient aerobic degradation of the oxygen consuming substances delivered with influent septage. The ventilation tubes placed in the substrate could also be a factor that helps support bottom ventilation to allow for oxygen diffusion through the substrate layers that is colonized by plant roots and nitrifying biomass, which could increase nitrification rate. Other authors have also highlighted the importance of oxygen concentration in EWs (see Cooper, 2005; Green *et al*. 2008; Noorvee *et al*., 2005). Good oxygen conditions favor nitrification and limit denitrification. Accordingly, total nitrogen removal in both the wetlands with SLR 100 and 250 did not differ significantly. From the experimental results, it was evident that SLR can affect ammonia removal and nitrate concentration in the system, with lower SLR achieving higher ammonia removal efficiency in a relatively more aerobic environment in the substrate.

Bed Clogging Assessment

Fig.2 depicts the solid retained on the bed surface 3 days after septage feeding onto the wetland and Fig.3 pictures the dried septage layer on the 8th day after sludge application, both with SLR of 250 kg TS/m^2.yr. Although wetland A2 was loaded with 250 kg TS/m^2.yr of SLR, which has 2.5 times higher the volume of septage applied per each batch, there was no bed clogging observed as evidenced from the draining efficiency of the septage during the course of operation of 8 weeks. Continuous growth of *phragmites* roots and rhizomes in the septage layers and the filter media, besides movement of plant stems helped to create airway into the beds and thus prevented media clogging. The beds remained aerobic with the applied load and feeding regime, as evidenced from the resulting effluent DO concentartion (Table 6). The preservation of aerobic conditions in the wetland substrate is the vital requisite for clogging prevention,

besides effective nitrification (Platzer & Mauch, 1997). Quantification and qualification of the dried sludge layers are still afoot and the output is not presented in this paper.

Fig.2. Septage residual retained on top of the wetland A2 3 days after feeding with SLR 250 kg/m^2.yr

Fig.3. Dried septage residual retained on top of the wetland A2 after 7 days (SLR 250 kg/m^2.yr)

CONCLUSION

A specially-designed, two-staged, engineered wetland-based treatment system for septage was constructed, tested, and presented in this paper. Based on the results obtained from the preliminary experiment conducted for 8 weeks, the first stage of the septage treatment by the pilot VFEWs system was efficient, with ≥ 80% of removal for COD, BOD, NH3-N, and TS at solid loading of 100 kg TS/m^2.yr. It was shown that even with higher SLR at 250 kg TS/

m^2.yr, the wetland could still achieve similar removal with no statistically different treatment efficiency for most of the parameters tested except for NH_3-N. The overall performance of the wetlands is encouraging and prolonged and extended studies are still on-going to assess on how both various system and application-related parameters could affect the treatment efficiency of this system.

REFERENCES

AECOM International Development, & Sandec/EAWAG. (2010). A Rapid Assessment of Septage Management in Asia: Policies and Practices in India, Indonesia, Malaysia, the Philippines, Sri Lanka, Thailand, and Vietnam *Project title: Environmental Cooperation-Asia (ECO-Asia)*: USAID.

Ahmed, S., Popov, V., & Trevedi, R. C. (2008). *Constructed wetland as tertiary treatment for municipal wastewater.* Paper presented at the Proceedings of Institution of Civil Engineers: Waste and Resource Management.

Burgoon, P. S., Frank Kirkbride, K., Henderson, M. E., & Landon, E. (1997). Reed beds for biosolids drying in the arid northwestern United States. *Water Science and Technology, 35*, 287-292.

Cofie, O. O., Agbottah, S., Strauss, M., Esseku, H., Montangero, a., Awuah, E., & Kone, D. (2006). Solid-liquid separation of faecal sludge using drying beds in Ghana: implications for nutrient recycling in urban agriculture. *Water Research, 40*(1), 75-82.

Cooper, P. (2005). The performance of vertical flow constructed wetland systems with special reference to the significance of oxygen transfer and hydraulic loading rates. *Water Science and Technology, 51*(9), 81-90.

Cooper, P. F., Job, G. D., Green, M. B., & Shutes, R. B. E. (1996). *Reed Beds and Constructed Wetlands for Wastewater Treatment*. Swindon, UK: WRc Publications.

Diemont, S. A. (2006). Mosquito larvae density and pollutant removal in tropical wetland treatment systems in Honduras. *Environment International, 32*(3), 332-341.

Green, M., Friedler, E., & Safrai, I. Enhancing nitrification in vertical flow constructed wetland utilizing a passive air pump. *Water Research, 32*(12), 3513-3520.

Jin, B., Wilén, B. M., & Lant, P. (2004). Impacts of morphological, physical and chemical properties of sludge flocs on dewaterability of activated sludge. *Chemical Engineering Journal, 98*(1-2), 115-126.

Kengne, I. M., Amougou, A., Soh, E. K., Tsama, V., Ngoutane, M. M., Dodane, P.-H., & Koné, D. (2008). Effects of faecal sludge application on growth characteristics and chemical composition of Echinochloa pyramidalis (Lam.) Hitch. and Chase and Cyperus papyrus L. *Ecological Engineering, 34*, 233-242.

Kengne, I. M., Dodane, P.-H., Akoa, A., & Koné, D. (2009). Vertical-flow constructed wetlands as sustainable sanitation approach for faecal sludge dewatering in developing countries. *Desalination, 248*(1-3), 291-297.

Koottatep, T., & Polprasert, C. (1997). Role of plant uptake on nitrogen removal in constructed wetlands located in the tropics. *Water Science and Technology, 36*(12), 1-8.

Koottatep, T., Polprasert, C., Oanh, N. T. K., Heinss, U., Montangero, A., & Strauss, M. (2001a). Potentials of vertical-flow constructed wetlands for septage treatment in tropical regions. *Advances in Water and Wastewater Treatment Technology*, 315-323.

Koottatep, T., Polprasert, C., Oanh, N. T. K., Heinss, U., Montangero, A., & Strauss, M. (2001b).

Septage dewatering in vertical-flow constructed wetlands located in the tropics. *Water Science and Technology, 44*(2-3), 181-188.

Koottatep, T., Surinkul, N., Polprasert, C., Kamal, A. S. M., Kone´, D., Montangero, A., M. Strauss, M. (2005). Treatment of septage in constructed wetlands in tropical climate: lessons learnt from seven years of operation. *Water Science Technology, 51*, 119-126.

Merlin, G., Pajean, J. L., & Lissolo, T. (2002). Performance of constructed wetlands for municipal wastewater treatment in rural mountainous area. *Hydrobiologia, 469*, 87-98.

Nielsen, S. (2003). Sludge drying reed beds. *Water Science Technology, 48*, 101-109.

Noorvee, a., Põldvere, E., & Mander, U. (2005). The effect of a vertical flow filter bed on a hybrid constructed wetland system. *Water Science and Technology, 51*(9), 137-144.

Panuvatvanich, A., Koottatep, T., & Kone, D. (2009). Influence of sand layer depth and percolate impounding regime on nitrogen transformation in vertical-flow constructed wetlands treating faecal sludge. *Water Research, 43*(10), 2623-2630.

Platzer, C., & Mauch, K. (1997). Soil clogging in vertical flow reed beds mechanisms, parameters, consequences and.....solutions? *Water Science and Technology, 35*(5), 175-181.

Prochaska, C. A., Zouboulis, A. I., & Eskridge, K. M. (2007). Performance of pilot-scale vertical-flow constructed wetlands, as affected by season, substrate, hydraulic load and frequency of application of simulated urban sewage. *Ecological Engineering, 31*(1), 57-66.

Torrensa, A., Molleb, P., Boutinb, C., & Salgota, M. (2009). Impact of design and operation variables on the performance of vertical-flow constructed wetlands and intermittent sand filters treating pond effluent. *Water Research, 43*(7), 1851-1858.

Trang, N. T. D., Konnerup, D., Schierup, H., Chiem, N. H., Tuan, L. A., & Brix, H. (2010). Kinetics of pollutant removal from domestic wastewater in a tropical horizontal subsurface flow constructed wetland system: Effects of hydraulic loading rate. *Ecological Engineering, 36*(4), 527-535.

Vincent, J., Molle, P., Wisniewski, C., & Liénard, A. (2011). Sludge drying reed beds for septage treatment: Towards design and operation recommendations. *Bioresource Technology, 102*(17), 8327-8330.

Pertanika J. Sci. & Technol. 22 (2): 627 - 636 (2014)

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Improvement of Engineering Properties of Peat with Palm Oil Clinker

M. E. Rahman[1]*, M. Leblouba[1] and V. Pakrashi[2]

[1]*School of Engineering and Science, Curtin University Sarawak, CDT 250 98009 Miri, Sarawak, Malaysia.*

[2]*Department of Civil Engineering, University College Cork, Cork, Ireland*

ABSTRACT

The aim of this study is to investigate the effects of Palm Oil Clinker (POC) added as a stabilizer for improving the strength of peat. Cement and POC are added into peat up to 50% of the maximum dry unit weight. Treated peat achieved higher dry unit weight, almost 2.5 times as compared to untreated peat. Unconfined compressive strength (UCS) of treated peat is also investigated for soaked and unsoaked conditions. The results show that curing time improved the unconfined compressive strength of treated sample and increased by a factor of 20 and 11 for unsoaked and soaked conditions after 28 days of curing, respectively. The treated samples added with POC can be related to an increase in unconfined compressive strength for long time curing.

Keywords: Peat, Palm Oil Clinker (POC), Unconfined Compressive Strength (UCS)

INTRODUCTION

The understanding of soil characteristics in terms of bearing capacity is crucial for engineers to ensure construction on all types of ground surface can prove sufficient bearing resistance to the construction load. According to Huat *et al*. (2009), 8% of total Malaysia soil consists of organic soils. A high organic content soil, otherwise known as peat, is of particular concern due to its low durability and high compressibility. Organic soils usually have at least 20% organic contents in the soil. The organic matter in the soil represents an accumulation of partially decomposed vegetation or plants under suitable conditions. High water content and poor ground drainage contribute to the higher rate of dry matter accumulation compared to rate of decay. Hence, ground surface begin to be overlain by organic matters. At the extreme, some soil can contain up to 75% and more of organic contents in the soil (Huat, 2004). Replacing

Article history:
Received: 28 September 2013
Accepted: 18 January 2014

Email addresses:
M. E. Rahman (merahman@curtin.edu.my),
M. Leblouba (mlablouba@gmail.com),
V. Pakrashi (v.pakrashi@ucc.ie)
*Corresponding Author

ISSN: 0128-7702

peat with good quality soil is still a common practice even though this effort may lead to uneconomical design (Said & Taib, 2009).

Some key problems related to peat soil are low shear strength, low bearing capacity and high compressibility leading to excessive settlements (Huat *et al.,* 2009). High water content in peat contributes to its low density and low resistance towards bearing loads when compared to other soil types. Additionally, the shear strength of the peat soil is exceptionally low due to loss of cohesiveness between soil particles in the presence of excessive water as lubricant. These deficiencies related to of strength capacity pose a challenge for construction to be built on peat soil.

Improvement of engineering properties of peat can be investigated by noting the increase in unconfined compressive strength (UCS) of stabilized peat as compared to an untreated sample. Hashim and Islam (2008) found that Unconfined Compressive Strength (UCS) of pure peat ranges between 4.7kPa to 6.9kPa.

Peat stabilization can be performed by adding additives such as silica fume, cement, fly ash and lime. These are some common soil conditioners used for the research in recent studies. Kalantari *et al* (2009) carried out a peat stabilization study using silica fume as additives and found that the untreated peat achieved up to 28kPa of unconfined compressive strength. Said and Taib (2009) carried out a study on peat stabilization with carbide lime and concluded that stabilized peat achieved an unconfined compressive strength of 115kPa. Kolay *et al*. (2011) found that 20% of cement added into peat could achieve up to 115kPa of unconfined compressive strength after 28 days of curing time. For the same 28 days of curing, treated peat added with lime and fly ash achieved 70kPa and 60kPa of unconfined compressive strength, respectively. Kalantari *et al.* (2009) reported an increase of unconfined compressive strength of treated peat with 25% cement and 5% silica fume up to 320kPa after 28 days of curing time.

According to Said and Taib (2009), a longer curing is associated with a higher unconfined compressive strength gained. Similarly, Kolay and Pui (2010) also concluded that curing time do significantly improve the strength of the peat soil. Kolay *et al*. (2011) reported that curing period increased the unconfined compressive strength of treated sample with fly ash and lime from 70kPa up to 90kPa based on 28 days of curing time. The increase of unconfined compressive strength corresponding to the curing period is caused by the hydration process between cement and water. Calcium hydroxide produced through the hydration process could promote the bonding between the particles of peat by acting as a paste.

In this study, classification of peat was carried out for geotechnical testing such as natural moisture and organic content, specific gravity, particle size distribution and Atterberg limits. In addition, unconfined compression strength of peat stabilized with POC is determined and the effect is investigated as compared to untreated peat soil.

MATERIALS

The main materials used for this study are peat soil, cement and palm oil clinker (POC). Cement used for laboratory testing is Ordinary Portland cement. Peat soils were transported from Matang, Kuching, Sarawak, while POC was collected from Lambir, Sarawak.

Peat

According to Huat *et al.* (2009), 8% of total Malaysia soil consists of organic soils. Peat soils used for this study are collected from Matang area. The area is near the river in Kuching, Sarawak and is swampy. The peat samples collected are at a depth between 500mm to 1000mm. All obvious large foreign objects such as wood bark and man-made objects were removed on the spot. Peat soils are collected in bags and were subsequently taken out and dried under the sun.

Palm Oil Clinker

Palm Oil Clinker (POC) collected from the furnace at Lambir Palm Oil mill were brought to CMS cement Sdn. Bhd. for grinding. POC was ground to fines passing 35 micron and stored inside containers.

POC is the residue waste following combustion of fibres used for steam turbine in palm oil mill. Palm oil mill is abundant nationwide in Malaysia, the second highest palm oil exporter in the world and POC production is significant. Generally POC is used for land-fills. Land-filling of POC has a negative effect on environment and occupies large land space. Hence, POC is often regarded as an industrial waste material with little commercial value. However, POC has also been reported to contain up to 81.8% of silica oxide (Rafidah & Chan, 2009). Presence of silica oxide is important for the longer term strength gain in stabilized peat soil (Kalantari *et al.*, 2009). Hence, silica oxide–rich POC can promote secondary pozzolanic effects of stabilized peat soil. With a longer time of curing, stabilized peat soil rich in calcium hydroxide will react with silica oxide supplied by POC which will result in further strength gain. The chemical composition of Palm coil clinker is presented in Table 1.

TABLE 1
Palm Oil Clinker Properties (Rafidah & Chan, 2009)

Element	Symbol	Concentration (%)
Sillicca Dioxide	SiO_2	81.8
Feric Oxide	Fe_2O_3	5.18
Potassium	K_2O	4.66
Aluminium Oxide	Al_2O_3	3.5
Calcium Oxide	CaO	2.3
Magnesium Oxide	MgO	1.24
Phosphorus Dioxide	P_2O_5	0.76
Titanium Dioxide	TiO_2	0.17
Sodium Oxide	Na_2O	0.14

METHODOLOGY

Before peat soils were stabilized, classification of the peat soil was carried out. Geotechnical tests of peat were performed to determine the moisture and organic content, specific gravity, particle size distribution and Atterberg limits.

Peat Soil Classification

Moisture and organic contents were determined for 11 packs of peat soil. The range of the peat moisture and organic content is shown in Table 2. The specific gravity test was performed in accordance with ASTM Method A 854. Particle size analysis was carried out in accordance to ASTM D 422-63 by performing hydrometer and sieving test. The test procedures were carried out according to Cheng and Evett (2003). The particle size distribution is illustrated in Fig.1. Atterberg limits are determined for both liquid and plastic limit of the peat soil in accordance to BS1377 (Huat, 2004). Table 2 shows the list of the geotechnical properties of the peat soil.

TABLE 2
Geotechnical Properties of Peat

Geotechnical Properties	Value
Natural moisture content	306% – 450%
Organic content	80.0 % – 93.1%
Specific gravity	1.39
Liquid limit	203
Plastic limit	143
Plasticity index	60
Liquidity index	-1.69

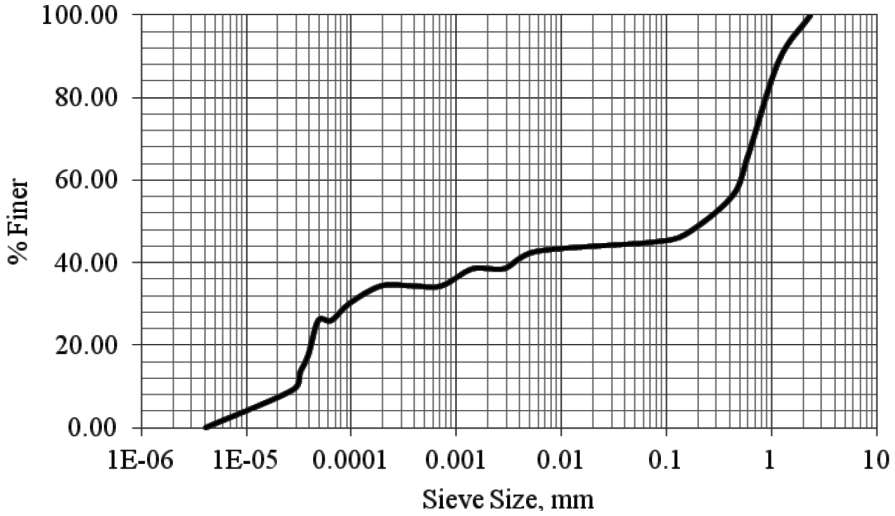


Fig.1: Particle Size Distribution

Compaction Test

Compaction test is performed in accordance with BS1377 for 5 types of binder ratio and untreated peat, as shown in Table 3 (British Standards, 1990). By using an automated compactor machine, a 2.5kg hammer was placed 300mm above the surface of the mould and it was

ensured that all soil particles pass through a 20mm sieve. Stabilized peat soil was filled into one third of a 1L mould and blown 27 times for each layer. Once the peat sample is finished compacted, the mass of the peat is weighted and the dry unit weight (ρ_d) is computed using the following equation:

$$\rho_d = \frac{1000\rho g}{w + 100} \tag{1}$$

where ρ is the bulk unit weight of the peat sample, g is the gravitational acceleration and w is the water content of the peat sample. The air void lines for 0%, 5% and 10% of air voids are computed as (2).

$$\rho_d = \frac{\left(1 + \frac{V_a}{100}\right)g}{\frac{1}{\rho_s} + \frac{w}{100\rho_w}} \tag{2}$$

where V_a is the air void ratio, ρ_s and ρ_w are the soil and water density, respectively. The maximum dry unit weight is the peak point of the line on the compaction curve and the corresponding moisture content is the optimum moisture content for the peat sample.

TABLE 3
Binder Ratio Summary

Binder	Peat Soil (%)	Palm Oil Clinker (%)	Cement (%)	OMC (%)
1	100	0	0	83.1
2	50	50	0	39.9
3	50	35	15	49.8
4	50	25	25	29.1
5	50	15	35	22.5
6	50	0	50	38.5

Unconfined Compression Strength Test

Peat soils are passed through 2mm sieve and compact using Humboldt Compactor Apparatus. The mould used for preparing the peat sample is 38mm in diameter and 76mm in height. Water was added at 5mL per 100g of peat sample to produce moisture content about 27% of water content. The peat sample was weighed to ensure the weight is between 1620g to 1700g. This ensured that the unit weight of the peat samples fall close to the optimum dry unit weight range. The peat samples were extruded using the Humboldt extruder and tested using unconfined compression strength (UCS) apparatus. The load shall be applied at the rate of about 2 revolutions per second to achieve the 0.5% to 2% per minute axial strain rate, as stated in Cheng and Evett. (2003) The area, A_0 of the contacting surface of the peat sample to the platen of the UCS machine (3) and the axial strain, ϵ were computed with respect to the change in height of sample as

$$A_0 = \frac{\pi D_0^2}{4} \tag{3}$$

$$\varepsilon = \frac{\Delta H}{H_0} \tag{4}$$

where After obtaining the loading, N value form the proving ring dial of the UCS machine, the pressure of upon the peat sample is computed based on the reduced area, A, following:

$$A = \frac{A_0}{1-\varepsilon} \tag{5}$$

$$P = \frac{N}{1000A} \tag{6}$$

where A is the corrected area, A_0 is the original area, N is the axial load and p is the axial stress.

Curing Time

The treated peat samples added with 15% POC and 35% cement, 25% POC and 25% cement, 35% POC and 15% cement were air-cured for 7 days and 28days, respectively. According to Kalantari *et al*. (2009), air-curing period allowed the cement to complete almost all its reactions after being cured for 90 days, and the water content of the peat was also reduced, resulting in an increase in strength.

Furthermore, both soaked and un-soaked stabilized peat samples were prepared for cured sample. Consequently, the unconfined compressive strength gained after curing could be obtained for both conditions. Soaked conditions, however, would portray a more realistic value as the ground conditions for construction. The ratio of the maximum dry unit weight achieved by treated peat sample to untreated sample was about 2.5.

RESULTS AND DISCUSION

Moisture Content – Dry Unit Weight Relationship

The compaction curve for all the treated and untreated peat samples are illustrated in Fig.2. The untreated peat (100/0/0) showed a lower optimum dry density at 5.4kN/m^3. Meanwhile, the stabilized peat sample containing 50% peat, 15% POC and 35% cement shows that the optimum dry density achieved is 12.6kN/m^3. This indicates that mixing of more cement than POC produces higher compatibility, resulting in lesser air voids. Furthermore, peat sample containing 50% peat, 35% cement and 15% POC can achieve close to maximum dry unit weight as compared to binder with 50% peat and 50% cement at 9% lower water content, as shown in Fig.3 below. Hence, it can be observed that POC can reduce cost by substituting cement as filler for achieving compatibility similar to peat stabilized using cement.

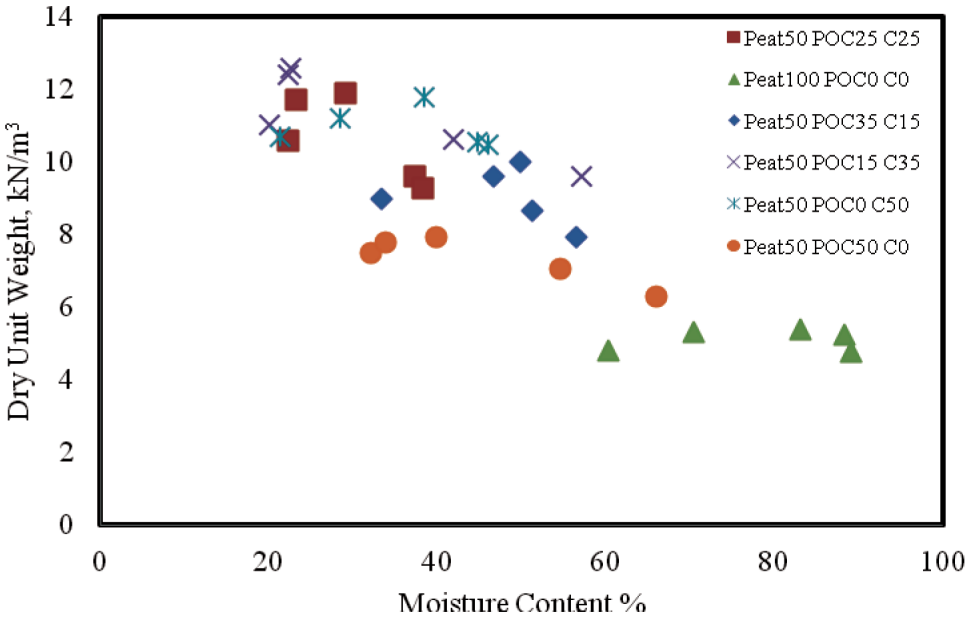


Fig.2: Compaction curves

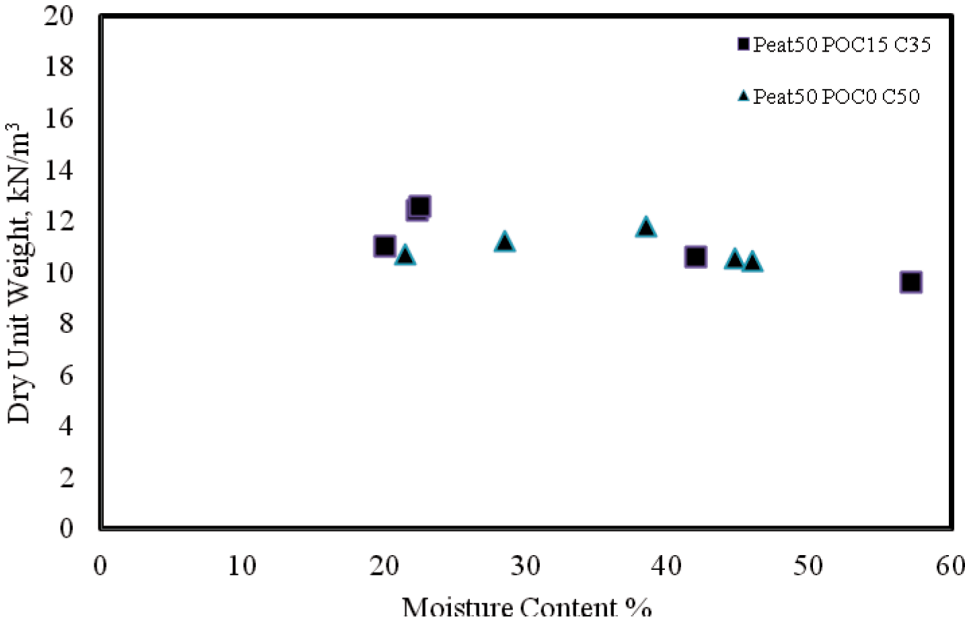


Fig.3: Compaction Curve for Peat Sample Containing 50% Cement and Peat Sample Containing 35% Cement and 15% POC

Unconfined Compressive Strength

Unconfined Compression Strength (USC) test was carried out for cohesive soils in order to obtain an approximate value of un-drained shear strength (Bardet, 1997). A summary of UCS for the treated and untreated peat samples is illustrated in Fig.4. The treated peat samples were cured for 7 days and 28 days respectively for both the soaked and un-soaked conditions so as to simulate realistic condition of construction, where the soils are often in soaked condition. From Fig.4, it can be seen that the untreated peat portrays a much lower unconfined compressive strength at 18kPa on average. The stabilized peat sample, containing 50% peat, 25% POC and 25% cement respectively, showed higher unconfined compressive strength at 109kPa after 7 days of curing time. Meanwhile, the corresponding soaked sample showed a lower unconfined compressive strength at 67kPa. For both un-soaked and soaked peat samples of 50% peat, 35% POC, 15% cement, the unconfined compressive strengths were 101kPa and 82kPa, respectively, after 7 days of curing time. In contrast, for both the un-soaked and soaked binder of 50% peat, 15% POC, 35% cement, the unconfined compressive strengths were 128kPa and 88kPa respectively after 7 days of curing.

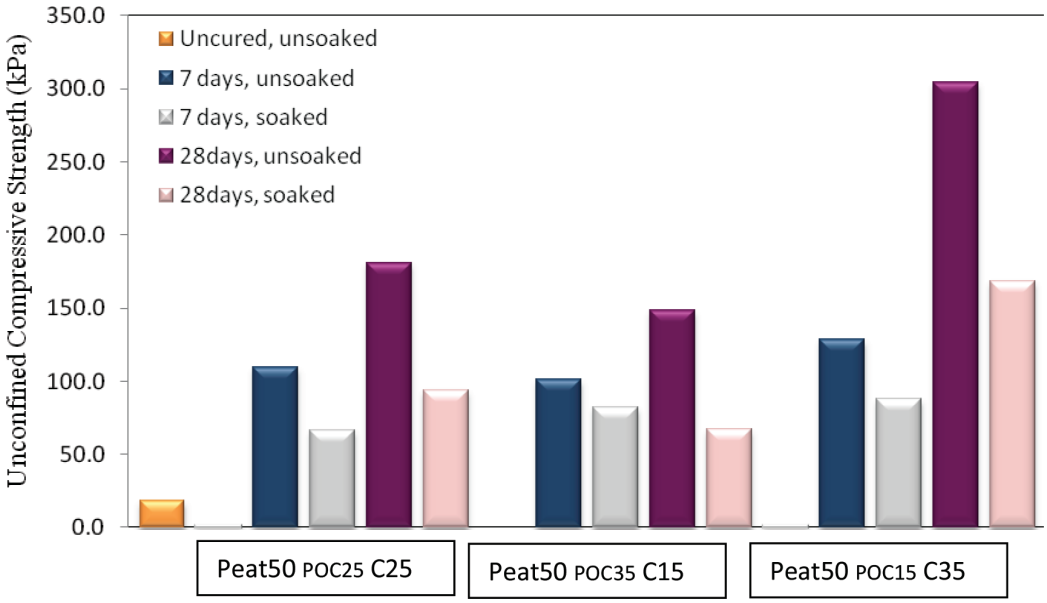


Fig.4: Unconfined Compressive Strength for All Peat Sample

At 28 days, stabilized peat sample containing 50% peat, 25% POC and 25% cement shows an increment of unconfined compressive strength at 181kPa and 94kPa for the un-soaked and soaked samples, respectively. This pattern is evident for un-soaked binder of 50% peat, 35% POC, 15% cement, showing a compressive strength of 149kPa. When compared to 7 days of curing, the 28days cured un-soaked and soaked samples with binder of 50% peat, 15% POC, and 35% cement showed an increase in unconfined compressive strength up to 304kPa and 169kPa, respectively.

The stabilized peat samples showed strength gain after 7 days of curing. The pozzolanic effects of cement promoted the strength gain in the stabilized peat samples. On the other hand, peat added with POC is results in a higher strength gain for a longer curing time. A possible reason behind this is that the high content of silica in POC that reacts with calcium hydroxide, which is produced through the hydration process between water and cement. During the first few days of curing, calcium hydroxide production is rapid, since presence of water is sufficient for hydration process to take place. After 7 days, the water content decreased, causing the strength gain in stabilized peat to slow down. Therefore, secondary pozzolanic reaction takes places generating secondary cement hydration with silica oxides contained in POC. Therefore, the peat samples having higher amount of POC result in a higher strength gain for longer curing times. The curing effect can be observed in Fig.5 for the stabilized peat sample for 7 days and 28days, respectively.

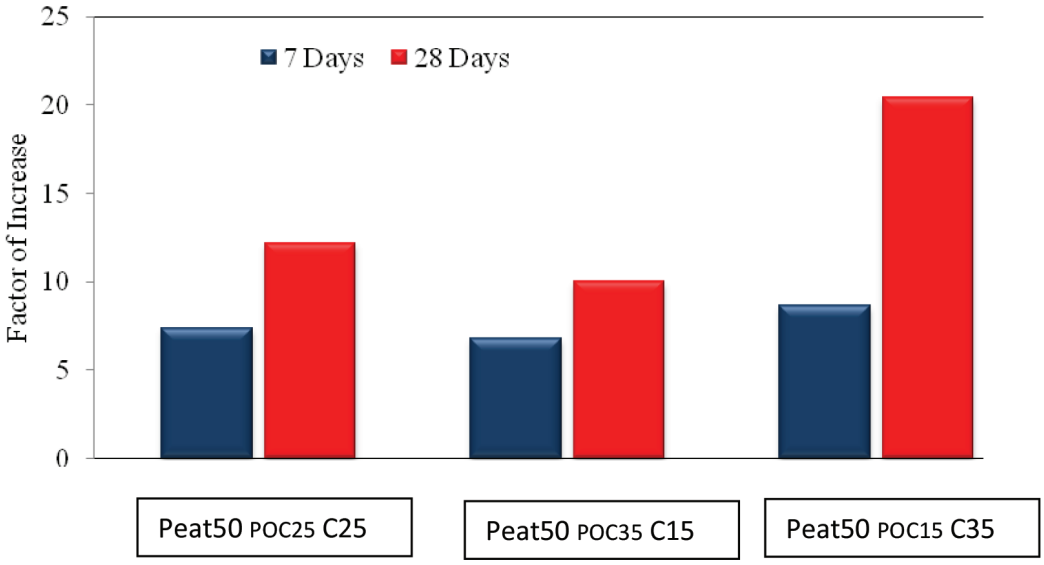


Fig.5: Factor of Increase in Unconfined Compressive Strength to Curing Period

CONCLUSION

Palm Oil Clinker (POC), an industrial waste, was found to be suitable for use in stabilizing peat soil in this paper through a number of experiments. Peat, comprising of around 8% of the soil of Malaysia, is often regarded as problematic due to its low strength related to high water retention and presence of organic content. However, POC is abundant around Malaysia and can be employed to strengthen such soils.

From the UCS test, the stabilized peat with ratio of 50% peat, 15% POC and 35% cement showed the highest optimum dry density at12.6kN/m3. Furthermore, 28 days cured the un-soaked peat sample with ratio of 50% peat, 15% POC and 35% cement showed the highest unconfined compressive strength at 304kPa. Hence, the bearing strength of the peat soil is increased through stabilization.

Furthermore, the results showed that POC produced higher strength gain at longer curing duration. Curing duration has effect on the strength gain of stabilized peat. In addition, addition of POC as peat soil stabilizer portrayed increase in soil optimum dry density and unconfined compressive strength by promoting secondary cement hydration.

Hence, POC is suitable to be utilized as soil stabilizer for peat soil as POC does contribute to the strength gain of treated peat soil. Peat soil that is stabilized by using POC portrays a higher strength gain for long-term curing.

ACKNOWLEDGMENTS

The first Author is grateful to the R&D Department, Curtin University Sarawak, for financial support to conduct this research work under CSRF scheme (Approval No: 2011-14-AG-ZZ).

REFERENCES

Bardet, J. P. (1997). *Experimental Soil Mechanics*. Pearson Education, New Jersey.

British Standards (1990). *British Standard Method of Testing for Soil Properties for Engineering Purposes* (BS1377-1990).

Cheng, L., & Evett, J. B. (2003). *Soil Properties Testing, Measurement, and Evaluation* (5th Edn.). University of North Carolina, Charlotte.

Hashim, R., & Islam, S. (2008). Engineering Properties of Peat Soils in Peninsular, Malaysia. *Journal of Applied Sciences, 8*(22), 4215-4219.

Huat, B. B. K. (2004). Deformation and Shear Stength Characteristics of Some Tropical Peat and Organic Soils. *Pertanika Journal Sci. Technol., 14*(1&2), 61-74.

Huat, B. B. K., Afshin Asadi, & Sina Kazemian. (2009). Experimental Investigation on Geomechanical Properties of Tropical Organic Soils and Peat. *American Journal of Engineering and Applied Sciences, 2*(1), 184-188.

Kalantari, B., Prasad, A., & Huat, B. B. K. (2009). Stabilizing Peat Soil with Silica Fume. *Proceedings of Institution of Civil Engineering*, 164

Kolay, P. K., & Pui, M. P. (2010). Peat Stabilization using Gypsum and Fly Ash. *UNIMAS E-Journal of Civil Engineering*, *1*. 1-5.

Kolay, P. K., Aminur, M. R., Taib, S. N. L., & Mohd Zain, M. I. S. (2011). Stabilization of Tropical Peat Soil from Sarawak with Different Stabilizing Agents. *Geotech Geo1. Eng. 29*, 1135-1141.

Rafidah, B. R., & Chan C. M. (2009). Reusing Soft-Soils with Cement-Palm Oil Clinker (POC) Stabilization. *International Conference on Engineering and Education in the 21st Century,* 1-4.

Said, J. M., & Taib, S. N. L. (2009). Peat Stabilization with Carbide Lime. *UNIMAS E-Journal of Civil Engineering, 1*, 1-6.

Pertanika J. Sci. & Technol. 22 (2): 637 - 644 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Optimization and Analysis of Bioethanol Production from Cassava Starch Hydrolysis

Liew, E. W. T.

Department of Chemical Engineering, School of Engineering and Science, Curtin University, CDT 250 98009 Miri, Sarawak, Malaysia

ABSTRACT

Current ethanol production processes utilizing crops such as sugar cane and corn starch have been well established over the decade. Other crop such as cassava is a potential candidate in producing ethanol. However, thermal processes are required to hydrolyze starch for the production of fermentable sugars. The processes are energy intensive and could lead to undesirable by-products generation. In this work, the hydrolysis of cassava starch is studied following an experimental design as a statistical problem solving approach. Central composite design (CCD) is used in order to select the most important variables from the simultaneous study on the effect and influence of operating conditions of bioreactor utilized, namely, pH, temperature and substrate concentration, as well as to optimize the process of cassava starch hydrolysis. From the results obtained, it can be concluded that the cassava starch hydrolysis is enhanced by pH and temperature. Model validations show good agreement between experimental results and the predicted responses.

Keywords: Cassava, central composite design, hydrolysis, optimization, pH, temperature

INTRODUCTION

During the last few decades, there has been an increasing demand for alternative sources of fuels due to the excessive consumption of fossil fuels globally. These alternative sources such as ethanol may reside in the production of renewable energies. Currently, ethanol is being produced commercially by using starch crops such as cassava (Cardona & SánchezÓ, 2007). Cassava is a potential candidate to produce ethanol in large scale since it can be easily cultivated and has high carbohydrate content. On the other hand, cassava is able to yield 3-15 tons/hectare in an agricultural environment and even 20-40 tons/hectare in an extensive cultivation area (Daubresse &

Article history:
Received: 12 February 2013
Accepted: 18 January 2014

Email addresses:
Liew, E. W. T. (emily.liew@curtin.edu.my)
*Corresponding Author

ISSN: 0128-7702

Ntibashirwa, 1987). Due to its high drought tolerance and low demand for nutrients, it can produce acceptable amount of yield even under marginal environmental conditions (Cock, 1982; Stupak *et al*., 2006). It is suggested to utilize *Saccharomyces cerevisiae* as the cultivation microbe since this type of yeast is most commonly used for cell growth in fermentation. This is due to the fact that this type of yeast has an active glucose transport system, whereby it metabolizes glucose through the glycolytic pathway, a metabolic pathway to convert glucose to pyruvate and energy and subsequently to ethanol (Nath & Das, 2004).

Generally, very high ethanol performances in fermentation are affected by process conditions such as pH, temperature and substrate concentration (Aldiguier *et al*., 2004). These process conditions are crucial for optimizing the fermentation process so that high production could be attained with optimum settings of these process conditions. There are very few studies reported on the impact of temperature on the dynamic behaviour of *Saccharomyces cerevisiae* during fermentation processes (Torija, 2003). It is important to note that pH has a great impact on the microbial cell activities and can modify the chemical pathways of the biological reaction as well as the kinetics (Akin, 2008). The significance of these three combined process and operating conditions have yet to be studied in fermentation processes. Based on literature studies, most studies so far focused on combined conditions of pH and temperature.

Due to the diversity and importance of process conditions of an alcoholic fermentation process, it is vital to ensure that each condition are well operated in order to ensure good bioreactor operation and production rate of ethanol under optimum process conditions. Therefore, the objective of this study is to investigate the optimum conditions of the three conditions in achieving high ethanol productivity. It is of interest to develop a low energy requiring process to convert cassava starch to fermentable sugars in order to reduce the cost of bioethanol production.

MATERIALS AND METHODS

Materials and Instruments

The bioreactor used is in this study is the BIOSTAT A Plus 2L, MO-Assembly. Industrial Baker's yeast, i.e. *Saccharomyces cerevisiae*, is utilized as the inoculum culture. 1.5L of fermentation medium is prepared by adding 0.75L of solution medium and 0.75L of hydrolyzed cassava into the bioreactor tank.

Solution Medium

The solution medium is prepared by adding the following components: 1.5g yeast extract, 3.75g NH_4Cl, 4.37g Na_2HPO_4, 4.5g KH_2PO_4, 0.38g $MgSO_4$, 0.12g $CaCl_2$, 6.45g citric acid and 4.5g sodium citrate.

Starch Hydrolysis

150g of fresh cassava starch in powder form is added into a 0.75L of 0.1M sulphuric acid solution. Both are mixed evenly in a 1L beaker and sterilized at 121°C for 45 minutes to break

down the cassava starch into fermentable sugars. The hydrolyzed cassava starch is then cooled to room temperature.

Fermentation Medium

Both the solution medium (0.75L) and hydrolyzed starch (0.75L) is mixed evenly and sterilized again at 121°C for 45 minutes to avoid contamination of the fermentation medium. The fermentation medium cooled to room temperature after sterilization before fermentation starts.

Sampling and Analysis

Sampling is taken every 2 hours during the first 24 hours of the fermentation process. After 24 hours, sampling was taken in every 3 hours since it is observed that cell growth starts to decrease and plateau. Samples were analyzed straight away for the concentrations of glucose and ethanol in order to prevent contamination of the samples. Enzymatic test kits (R-Biopharm) and UV-VIS spectrophotometer were utilized to analyze the concentrations of glucose and ethanol.

Response Surface Methodology (RSM) Optimization

RSM is a statistical technique which is useful for modelling and analyzing problems in which a response of interest is influenced by several variables and to optimize the response (Aldiguier *et al*., 2004; Torija, 2003).

In this study, three independent variables are studied at three levels (-1, 0, +1) with eight (2^3) factorial points and three replicate central points. A central composite design (CCD) is employed to determine the effects of independent variables on the response, namely, glucose and ethanol concentrations, as well as factor interactions. Table 1 shows the input variables and levels employed.

TABLE 1
Input Variables and Their Levels Employed

Factor	Variable	Units	Low Level (-)	Middle Level (0)	High Level (+)
X1	pH		2	6	10
X2	Temperature	°C	25	32.5	40
X3	Substrate Concentration	g/L	30	40	50

CCD is one of the most commonly used response surface designs for fitting second-order models in fermentation studies (Akin, 2008). This design provides a solid foundation for the generation of a response surface map. The response pattern and synergy in the optimum region are investigated and to identify the optimum conditions for glucose and ethanol concentrations. The results of each CCD are analyzed using Design Expert® software version 8, from Statease, Inc., Minneapolis, USA. Their interactions and significance were evaluated by variance analysis (ANOVA) test. Three-dimensional surface plots are drawn to illustrate the effects of the

independent variables on the dependent variables, being described by a polynomial equation, fitted on experimental data. R^2 coefficient is used to evaluate the fit of the models.

RESULTS AND DISCUSSION

Optimization by RSM

The coded values of experimental variables in CCD and response values are shown in Table 2. Lack-of-fit tests were carried out for deriving the best correlation between independent variables and responses. It is indicated that the LFT is not significant which supports the fitness of the model. On the other hand, probability value (p-value < 0.05) indicated that the model is significant, as shown in Table 3.

TABLE 2
Central Composite Design (CCD) for Optimization and Values of Observed Responses

Run	Block	X1	X2	X3	Glucose Conc. (g/L)	Ethanol Conc. (g/L)
1	1	2	30	25	0.135	1.10
2	1	10	50	40	1.70	3.00
3	1	2	50	25	0.13	1.50
4	1	10	50	25	1.51	3.50
5	1	2	50	40	0.50	0.95
6	1	10	30	25	4.84	0.75
7	1	10	50	25	2.38	0.36
8	1	2	30	25	0.15	0.96
9	1	10	50	40	1.50	3.70
10	1	10	30	25	4.92	0.73
11	1	6	40	32.5	0.03	21.38
12	1	2	30	40	0.17	0.85
13	1	6	40	32.5	0.03	21.36
14	1	6	40	32.5	0.04	21.02
15	1	10	30	40	1.90	2.45
16	1	2	30	40	0.17	0.86
17	1	2	50	40	0.15	0.88
18	1	10	30	40	1.95	2.50
19	1	2	50	25	0.14	1.55
20	2	2	40	32.5	0.14	0.98
21	2	6	40	17.5	0.52	18.63
22	2	6	40	32.5	0.03	21.36
23	2	14	40	32.5	1.56	2.50
24	2	6	60	32.5	0.07	10.50
25	2	6	40	47.5	0.08	15.33
26	2	6	20	32.5	0.07	11.50
27	2	6	40	32.5	0.04	19.64
28	2	6	40	32.5	0.04	18.97

TABLE 3
Lack-of-Fit Test and Model Summary Statistics

Source	SS[a]	df[b]	MS[c]	*p*-value[d]	
Glucose concentration					
Model	34.42	6	5.74	0.0019	Significant
Lack-of-Fit	13.97	8	1.75	0.0503	not significant
Ethanol concentration					
Model	1767.61	6	294.60	0.0111	significant
Lack of Fit	967.52	8	120.94	0.0771	not significant

[a]SS, sum of squares; [b] df, degree of freedom; [c] MS, mean squares; [d] Statistically significant at 95% of confidence level.

The ANOVA results on the models are shown in Table 4. The *p*-values are less than 0.05, indicating that models and their terms are significant. In all cases, the insignificant model terms (*p*-value > 0.05) have been omitted to give a better fit. The fitness of the model is subsequently examined by the coefficient of determination R^2. The R^2 value for glucose concentration is 92.46% and for ethanol concentration, the R^2 value is 95.30%. Meanwhile, the adjusted R^2 value of glucose and ethanol concentrations are 90.56% and 93.40%, respectively. R^2 and adjusted R^2 values of regression model higher than 90% are considered to be high correlated (Bao, 2011). The models are denoted by Equations [1] and [2], as follows:

$$\text{Glucose concentration} = f(X_1.X_2.X_3) = -0.41 + 0.84X_1 + 0.09X_2 - 0.07X_3 - 0.02X_1X_2 \quad [1]$$

$$\text{Ethanol concentration} = f(X_1.X_2.X_3) = 19.37 - 5.72X_1 - 0.39X_2 - 0.11X_3 + 0.2X_1X_2 \quad [2]$$

TABLE 4
Analysis of Variance (ANOVA) Results

Source	df[a]	Glucose conc.		Ethanol conc.	
		SS[b]	*p*-value[c]	SS[b]	*p*-value[c]
X_1	1	11.15	0.0042	172.50	0.0015
X_2	1	1.25	0.0096	855.58	0.0055
X_3	1	12.18	0.0030	28.00	0.0034
X_1X_2	1	5.97	0.0282	565.12	0.0139
X_1X_3	1	2.78	0.1221	60.56	0.3882
X_2X_3	1	1.09	0.3234	85.86	0.3061

[a]df, degree of freedom; [b] SS, sum of squares; [c] Statistically significant at 95% of confidence level.

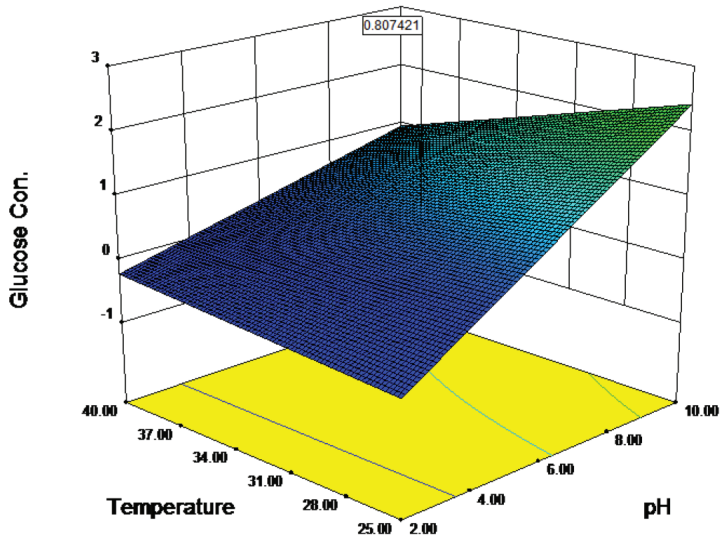


Fig. 1: Response Surface Plot for The Effect of pH and Temperature on Glucose Concentration. Substrate Concentration is Constant at Zero Level.

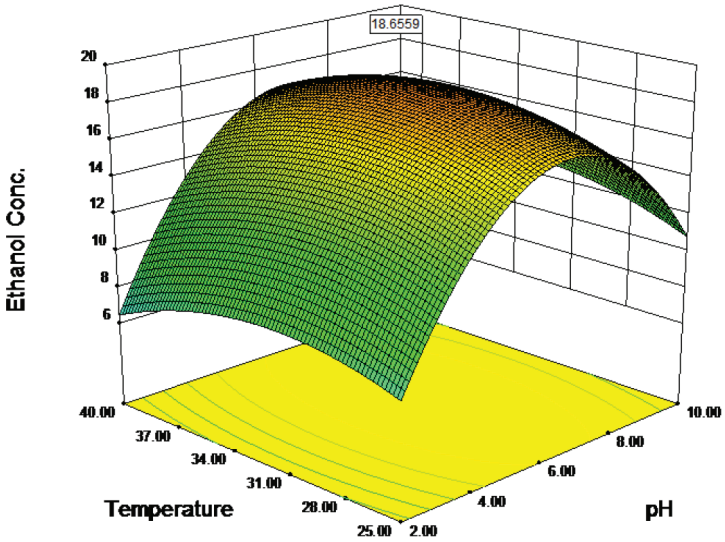


Fig. 2: Response Surface Plot for The Effect of pH and Temperature on Ethanol Concentration. Substrate Concentration is Constant at Zero Level.

It is observed that the linear terms for both pH and temperature have significant effects on glucose and ethanol concentrations. This shows that both pH and temperature highly affect the amount of glucose and ethanol concentrations produced during the fermentation process. At the same time, the interaction terms, i.e. X_1X_2 between pH and temperature show significant effect than the rest of the interaction terms since *p*-value is less than 0.05, as shown in

Table 4. Therefore, it is of interest to investigate the relationship between pH and temperature by visualizing the response surface plot. The polynomial equations to experimental data, i.e., Equations [1] and [2] can be described by the response surface plots as a function of pH and temperature, maintaining substrate concentration fixed at level zero, as shown in Fig.1 and Fig.2.

Based on the response surface plot, maximum ethanol concentration is obtained under high pH and high temperature conditions. The optimum values suggested to achieve high ethanol concentration is under the operation of pH 6.5, substrate concentration of 40.25g/L and temperature of 32.4°C.Under these conditions, it is predicted to achieve a maximum of 18.65g/L of ethanol and minimum of 0.81g/L of glucose.

Confirmation runs were conducted to validate the predicted ethanol concentration with respect to the optimum values suggested. The experimental response was 19.58g/L of ethanol and 0.85g/L of glucose. These values are in good agreement with the predicted values, considering a range of 95% confidence level. This shows the adaptation of the model to experimental data, confirming the validity of the models.

CONCLUSION

Several studies have evaluated the influence of pH and temperature in the production of ethanol. It is recommended in several studies that yeast generally grows well in pH range of 4 to 4.5, whereas for temperature, it is suggested to be operated in the range of 20 to 30°C (Stanbury *et al*., 2006). It is important to ensure that both pH and temperature are well conditioned so as to allow good yeast growth and achieve desirable amount of ethanol. In addition, different amounts of substrate do make a difference in achieving desirable amount of ethanol. Low substrate level will not be able to achieve maximum amount of ethanol. Substrate level which is too high will result in substrate inhibition. Therefore, it is vital to ensure that the substrate concentration used is optimum as well in order to ensure good production of ethanol.

From the results, it can be concluded that pH and temperature highly affect the amount of ethanol compared to substrate concentration. Both pH and temperature play important roles in ensuring optimum yeast growth since it is statistically proven that both have significant interactions within each other. The model shows good predictions between experimental results and predicted responses.

REFERENCES

Akin, H., Brandam, C., Meyer, X. M., & Strehaiano, P. (2008). A model for pH determination during alcoholic fermentation of a grape must by *Saccharomyces cerevisiae*. *Chemical Engineering and Processing: Process Intensification*, *47*, 1986-1993.

Aldiguier, A. S., Alfenore, S., Cameleyre, X., Goma, G., Uribelarrea, J. L., Guillouet, S. E., & Molina-Jouve, C. (2004). Synergistic temperature and ethanol effect on *Saccharomyces cerevisiae* dynamic behaviour in ethanol bio-fuel production. *Bioprocess Biosys Eng.*, *26*, 217-222.

Bao, Y., Chen, L., Wang, H., Yu, X., & Yan, Z. (2011). Multi-objective optimization of bioethanol production during cold enzyme starch hydrolysis in very high gravity cassava mash. *Bioresource Technology*, *102*, 8077-8084.

Cardona, C. A., & SánchezÓ, J. (2007). Fuel ethanol production: process design trends and integration opportunities. *Bioresource Technology*, *98,* 2415-2457.

Cock, J. H. (1982). Cassava - a basic energy-source in the tropics. *Science*, *218*, 755-762.

Daubresse, P., & Ntibashirwa, S. (1987). A process for protein enrichment of cassava by solid substrate fermentation in rural conditions. *Biotechnology and Bioengineering*, *29*, 962-968.

Nath, K., & Das, D. (2004). Improvement of fermentative hydrogen production: various approaches. *Appl Microbiol Biotechnol*, *65,* 520-529.

Stanbury, P. F., Whitaker, A., & Hall, S. J. (2006). *Principles of Fermentation Technology*. Butterworth Heinemann.

Stupak, M., Vanderschuren, H., Gruissem, W., & Zhang, P. (2006). Biotechnological approaches to cassava protein improvement. *Trends in Food Science & Technology*, *17,* 634-641.

Torija, M.J., Rozes, N., Poblet, M., Guillamon, J.M., & Mas, A. (2003). Effects of fermentation temperature on the strain population of *Saccharomyces cerevisiae*. *Int J Food Microbiol.*, *80*, 47-53.

Pertanika J. Sci. & Technol. 22 (2): 645 - 655 (2014)

PERTANIKA JOURNALS

SCIENCES & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Numerical Modelling of Molten Carbonate Fuel Cell: Effects of Gas Flow Direction in Anode and Cathode

Tay C. L.[1]* and Law M. C.[2]

[1]*Chemical Engineering Department, School of Engineering and Science, Curtin University Sarawak, Miri, Sarawak, Malaysia*

[2]*Mechanical Engineering Department, School of Engineering and Science, Curtin University Sarawak, Miri, Sarawak, Malaysia.*

ABSTRACT

The modelling of a three-dimensional (3-D) molten carbonate fuel cell (MCFC) was developed to study the effects of gas flow direction (co-flow and counter-flow) in anode and cathode on the generated power density by solving the mass and momentum conservation equations, electrochemical reaction and heat transfer. The simulation result of the co-flow temperature distribution was compared with the experimental data obtained from open literature. The molar fraction distribution of gases in the anode and cathode gas channels and temperature distribution across the cell were compared between two different flow directions. Furthermore, the performance of MCFC, which operates in the temperature range of 823 - 1023 K, was analysed by comparing the generated power density. The results showed that MCFC with co-flow attained higher power density compared to that of counter-flow at 873 K. However, at higher temperature of 1023 K, the generated power density was the same for both gas flow directions.

Keywords: Counter flow, Heat transfer, Molten Carbonate Fuel Cell, Numerical simulation, Three-dimensional model

INTRODUCTION

Nowadays, various types of molten carbonate fuel cell (MCFC) have been commercialized to generate electricity of around 200 kW to one megawatt for manufacturing factories, research centres, hospitals and commercial buildings (Ma *et al.,* 2009). The modelling of MCFC has become an important analysis method to better understand and optimize the operating conditions in order to improve its performance and lifespan.

According to Munoz *et al.* (2011), voltage calculation of MCFC could be categorised into three major approaches. The first approach was derived from the models developed for

Article history:
Received: 28 September 2013
Accepted: 18 January 2014

Email addresses:
Tay C. L. (taycl@curtin.edu.my),
Law M. C. (m.c.law@curtin.edu.my)
*Corresponding Author

ISSN: 0128-7702

solid oxide fuel cells (SOFC), which combined the experimental and theoretical approach; the second approach looked into the electrodes kinetics based on Butler-Volmer equation which was first developed by Yuh and Selman (1989), and the third approach focused on the voltage losses at cathode based on the Arrhenius-type equations (Baranak & Atakul, 2007).

He and Chen (1995) developed a 3-D, 5-cells stack model with a thickness of 14 mm for each cell and electrochemical area size of 1 m^2. The temperature and pressure distributions at the operating temperature of 873 K for three different gas flow types (co-flow, cross-flow and counter-flow) have been reported. Kim and co-workers (2010) developed a 3-D MCFC with corrugated gas channels and studied the current density distribution, ohmic resistance, and anode and cathode polarizations for three different gas flow directions at the operating temperature of 853 K. It was reported that the counter-flow model gave the best performance due to its smaller ohmic resistance and electrodes polarization.

The operating temperature of MCFC is in the range of 873 - 973 K and the maximum temperature gradient over the cells is recommended not to exceed 90 K in order to maintain the lifespan of the cells (Huppmann, 2009). It is believed that the high operating temperature may cause hot corrosion of current collector and separator, creep or compaction of electrodes and electrolyte matrix, as well as electrolyte loss that could reduce the lifespan of cells (Yuh & Farooque 2009). However, a low operating temperature may decrease the cell power density due to high ohmic resistance and electrodes polarization. The modelling of the MCFC is vital in achieving the balance between its performance and lifespan by optimising its operating condition and efficiency.

In the current study, a 3-D single cell MCFC model was developed using commercial software, COMSOL (version 4.1). This model solves a set of Navier-Stokes equations which includes multi-component gas species and Yuh and Selman's model (1984), which is used for the voltage-current relationship. The result of the temperature distribution was compared with the literature data. The effects of gas flow direction in the anode and cathode gas channels on the generated power density at different operating temperatures were also studied using the developed model.

MODELLING APPROACH

The modelling of MCFC was carried out based on several assumptions. Due to the symmetry of MCFC, only a pair of cathode and anode gas channels was simulated. The gas in the anode and cathode was assumed to be an ideal gas mixture, and the flow is laminar. The carbonate ions diffusion across the electrolyte was assumed to have a negligible effect on the performance of MCFC.

There are six main components in MCFC: current collectors, anode gas channel, anode electrode, electrolyte, cathode electrode and cathode gas channel (see Fig.1). The size of MCFC is 4.5 mm (W) × 7.0 mm (H) x 140 mm (D). The model contains 12800 elements, with 10 grids in the width direction for the fuel and oxidant channel. Details for the six main components are shown in Table 1.

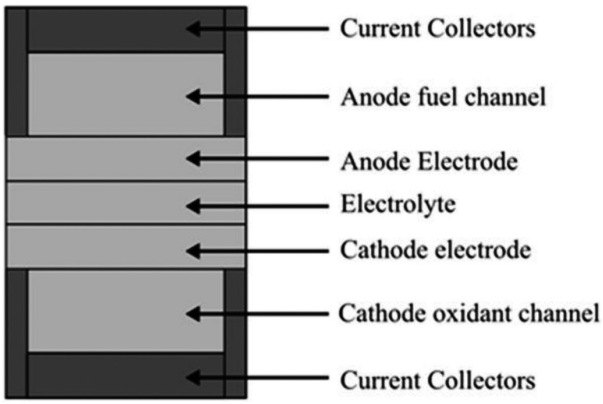


Fig.1. Front view of the MCFC 3-D model

TABLE 1
Specification of the components of MCFC cell (Yu *et al.,* 2008)

Components	Thickness (mm)	Porosity (%)	Mesh grids
Current collectors	0.8		3
Anode fuel channel	1.5		7
Anode electrode	0.8	0.5	5
Electrolyte	0.8		10
Cathode electrode	0.8	0.55	5
Cathode oxidant channel	1.5		7

The electrochemical reactions in MCFC's electrodes are shown as follows:

$$\text{Anode} \quad : H_2 + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e \qquad [1]$$

$$\text{Cathode} : \tfrac{1}{2}O_2 + CO_2 + 2e \rightarrow CO_3^{2-} \qquad [2]$$

$$\text{Overall} \quad : H_2 + \tfrac{1}{2}O_2 \rightarrow H_2O \qquad [3]$$

Based on the overall reaction, the electrons, carbonate ions are conserved within the overall reactions [3]. The carbonate ions, CO_3^{2-}, are consumed in the anode and generated in the cathode at the same reaction rate. The chemical reactions are presumed to occur in the electrode domains. The exterior surfaces of MCFC are adiabatic. In the counter-flow model, the flow direction in anode fuel channel remained the same but the flow direction in cathode oxidant channel is the opposite of that of fuel flow.

The oxidant and fuel in the channels are governed by the, single phase, laminar Navier-Stokes equations:

$$\nabla \cdot (\rho u) = Sp \qquad [4]$$

$$\rho(u \cdot \nabla)u = \nabla \cdot \left[-\rho I + \mu\left(\nabla u + (\nabla u)^T\right) - \frac{2}{3}\mu(\nabla \cdot u)I\right] \qquad [5]$$

where, ρ is the pressure (Pa), u is the velocity vector (m/s), I is the identity matrix, and T is the temperature (K). In the gas channels, the mass source term, S_p, is zero.

For the porous anode and cathode electrodes, the gas phase momentum is described by Darcy equation:

$$\frac{\rho}{\epsilon_p}\left((u\cdot\nabla)\frac{u}{\epsilon_p}\right)=\nabla\cdot\left[-\rho I+\frac{\mu}{\epsilon_p}(\nabla u+(\nabla u)^T)-\frac{2\mu}{3\epsilon_p}(u\cdot\nabla)I\right]-\frac{\mu u}{\kappa_{br}} \quad [6]$$

where, ϵ_p is the porosity, μ is dynamic viscosity (Pa·s) and κ_{br} is the permeability (m^2).

The rate of production/destruction of each species, S_j (mol/m^3s) in anode and cathode electrodes is:

$$S_j=v_j\frac{i}{2F\delta}MW_j \quad [7]$$

where v_j is stoichiometric constant for a gas species j, i is electrical current density (A/m^2), F is Faraday constant (96485.3 C/mol), MW_j is molecular weight of a gas species j (g/mol) and δ is the thickness (m).

The source terms in the anode and cathode electrodes are shown as follows (Kim *et al.,* 2010):

$$\text{Anode: } S_p\text{, anode} \quad = S_{CO_2}+S_{H_2}-S_{H_2} \quad [8]$$

$$\text{Cathode: } S_p\text{, cathode} = -(S_{CO_2}+S_{O_2}) \quad [9]$$

The transport of these concentrated gas species is described as:

$$\nabla\cdot J_j+\rho(u\cdot\nabla)\omega_j=S_j \quad [10]$$

where, ω_j is the mass fraction of a gas species of j, and J_j is defined by a mixture-average equation:

$$J_j=-\left(\rho D_j^m\nabla\omega_j D_j^m\frac{\nabla M_n}{M_n}\right) \quad [11]$$

where, the is the mixture-average diffusion coefficient and is defined as:

$$M_n=\left(\sum_j\omega_j/M_j\right)^{-1} \quad [12]$$

The heat transfer of MCFC is described with convection and conduction mechanisms as:

$$\rho C_p u\cdot\nabla T=\nabla(k_{eq}\nabla T)+Q \quad [13]$$

where, is specific heat capacity at constant pressure (J/(kg·K)), is the thermal conductivity (W/(m·K)) and Q is total heat source.

The total heat generated from MCFC was calculated by adding the electrical energy and the enthalpy change of the overall reaction [3] above (Koh *et al.,* 2002).

$$\text{Energy source}: \quad Q = \left[-r\Delta H + V_{cell}\, i\right]\frac{1}{\delta} \qquad [14]$$

where r is the reaction rate, (mol/m^{2}s), ΔH is molar enthalpy (J/mol), V_{cell} is electric voltage (V).

The molar enthalpy of formation of water was calculated as:

$$\Delta H = -\,(240506 + 7.3835T) \qquad [15]$$

The cell potential of the MCFC depends on the thermodynamic reversible potential, partial pressure of fuel gases and oxidant gases in the electrodes, and the irreversible losses of the cell due to the limits of the electrical current and the kinetic rate (Yuh & Selman, 1984).

$$V_{cell} = E_{rev} - i(\eta_\alpha + \eta_c + \eta_{ohm}) \qquad [16]$$

where E_{rev} is derived from the Nernst equation that considering the partial pressure of fuel gases in anode and oxidant gases in cathode (Kim *et al.,* 2010).

$$E_{rev} = E^0 + \frac{RT}{2F}\ln\frac{P^a_{H_2}}{P^a_{H_2O}\cdot P^a_{CO_2}} + \frac{RT}{2F}\ln\left(\sqrt{P^c_{O_2}P^c_{CO_2}}\right) \qquad [17]$$

The ideal standard potential E^0 for the cell reaction is calculated from Koh *et al.* (2002):

$$E^0 = 1.2723 - 2.7645 \times 10^{-4} \times T \qquad [18]$$

The ohmic loss (Ω m^2) is the electrical resistance caused by current flow (Kim *et al.,* 2010).

$$\eta_{ohm} = 0.5 \times 10^{-4}\exp\left[3016\left(\frac{1}{T} - \frac{1}{923}\right)\right] \qquad [19]$$

The activation polarization of anode and cathode (Yuh & Selman, 1984):

$$\eta_a = 2.27 \times 10^{-9}\exp\left[\frac{6435}{T}\right] \times P_{H_2}^{-0.42} \times P_{CO_2}^{-0.17} \times P_{H_2O}^{-0.1} \qquad [20]$$

$$\eta_c = 7.505 \times 10^{-10}\exp\left[\frac{9298}{T}\right] \times P_{O_2}^{-0.43} \times P_{CO_2}^{-0.09} \qquad [21]$$

The operating and boundary condition of MCFC are shown in Table 2.

TABLE 2
Operating and boundary conditions of MCFC (Yu *et al.,* 2008)

Parameter	Value
Operating pressure	101,325 Pa
Operating temperature	873 K
Current density	1400 A/m^2
Inlet mole fraction of fuel gases in the anode, H_2:CO_2:H_2O	0.40 : 0.40 : 0.20
Inlet mole fraction of oxidant gases in the cathode, O_2:CO_2:N_2	0.148 : 0.296 : 0.556
Anode inlet gas flow rate	40 litre/min
Cathode inlet gas flow rate	112.7 litre/min

RESULTS AND DISCUSSION

In Fig.2, it was observed that the temperature of the co-flow model increased from 873 K at the cathode inlet to the maximum temperature of 923 K at the cathode outlet. However, the temperature profile for the counter-flow is different compared to that of co-flow, where the temperature of the counter-flow model increases from 873K to 888K at the middle of the cell and decreases to 876K towards the end of the cell. This temperature profile could be due to the heat energy in cathode oxidant channel was regulated by the lower temperature from the anode inlet of MCFC. This happened to the anode side as well, whereby, the high temperature at the outlet of anode was regulated by the lower temperature at the cathode inlet of counter-flow model.

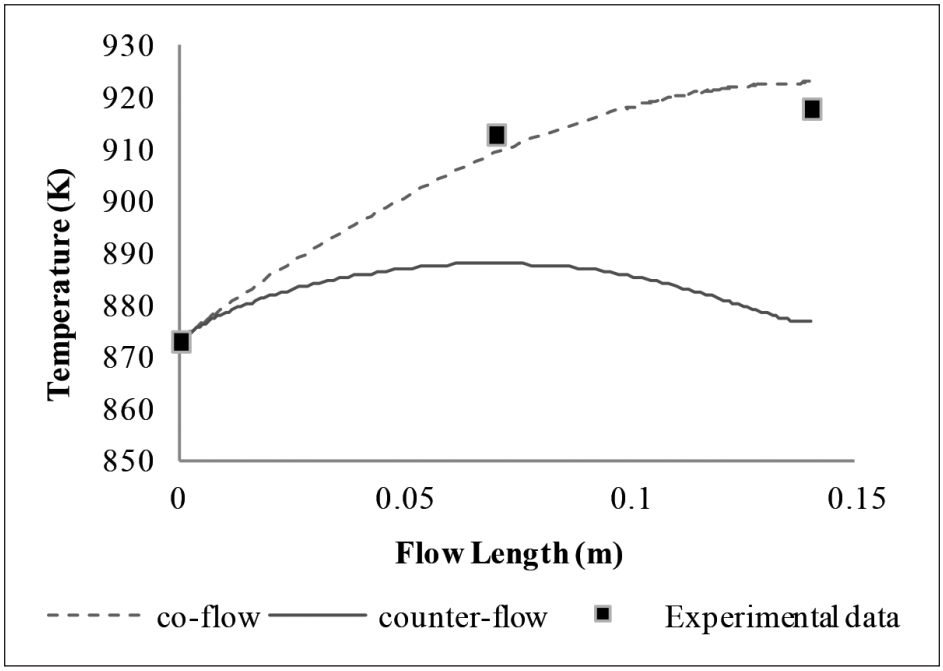


Fig.2. Temperature distributions of co-flow and counter-flow models operated at 873 K

The developed MCFC model has been validated with the literature data by comparing the temperature distribution across the cathode oxidant fuel channel (Yu *et al.,* 2008). By comparing the result with the published experimental result, it was found that there are 0.44% temperature difference at the middle of the cell and 0.54 % temperature difference at the end of the cell (see Fig.2). The temperature differences were in the range of ±5 K. The developed model was considered valid to further study the effects of the gas flow direction in the anode fuel channel and cathode oxidant channel.

Fig.3 and Fig.4 show the comparison for the distribution of hydrogen and water vapour in the anode fuel channel and the distribution of carbon dioxide and oxygen in the cathode oxidant fuel channel between two different gas flow models. From the figures, the reactant consumption and product formation rates for both models are the same from the inlet to the outlet of the cell. This result was supported with equation [7], whereby the rate of production/ destruction of each gas species in the anode and cathode was not temperature-dependent.

The performances of the MCFC cell for two different gas flow directions were analysed by comparing the generated power density from temperature of 823 K to 1023 K (Fig.5). It was observed that the MCFC cell with co-flow attained higher performance compared to the counter-flow model at a lower temperature. As the temperature increased from 823 K to 1023 K, the generated power density was also increased. However, the performances of the MCFC cell for these two flow configurations got closer at the temperature of 973K and became the same at 1023K.

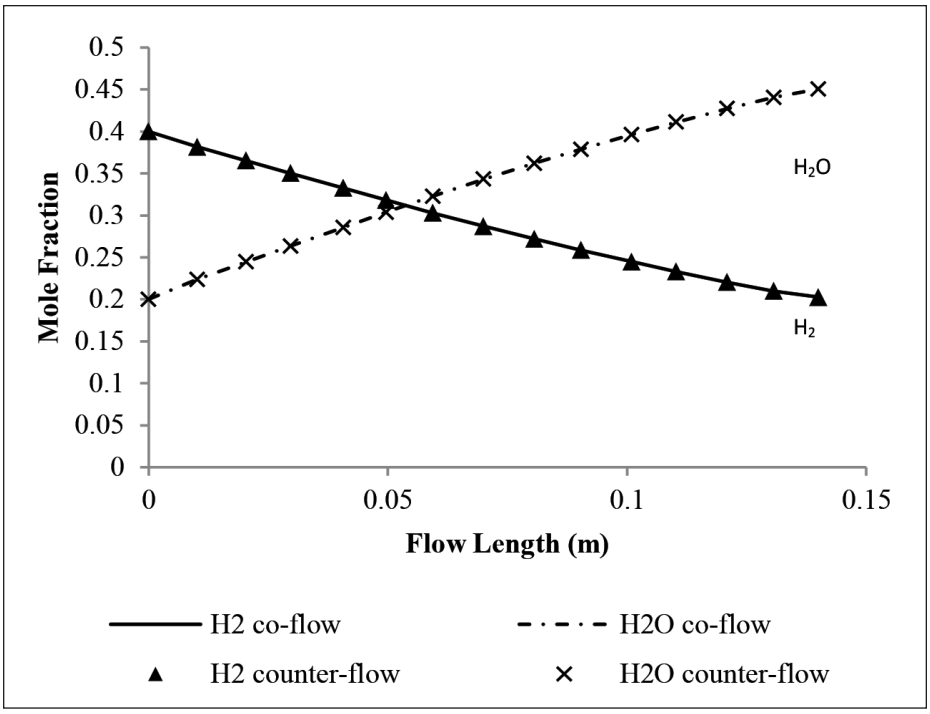


Fig.3. Distributions of hydrogen and water vapour in anode fuel channel of co-flow and counter flow models

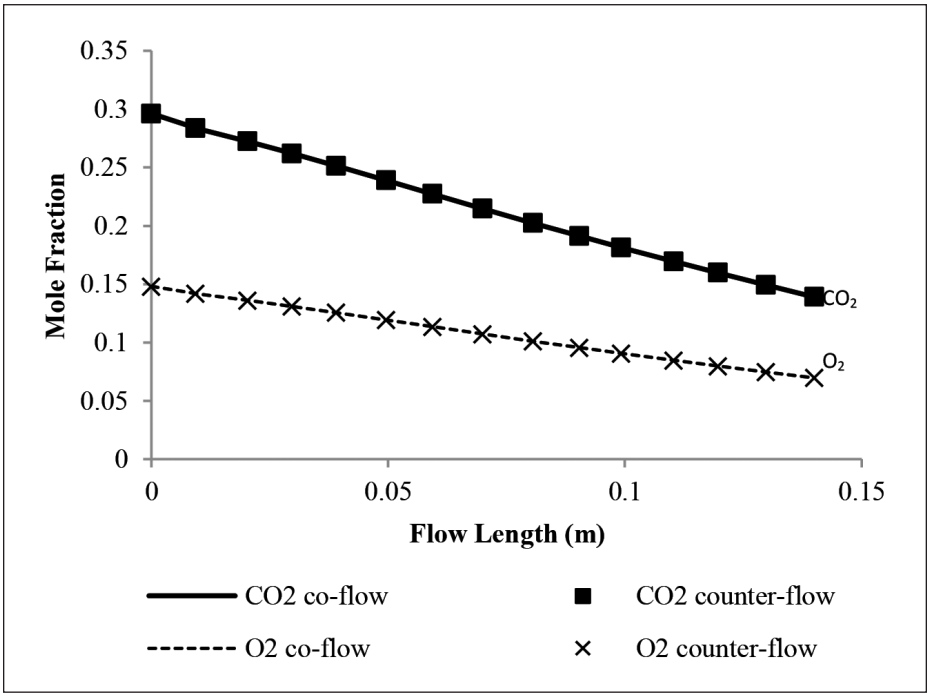


Fig.4. Distributions of oxygen and carbon dioxide in cathode oxidant channel of co-flow and counter-flow models

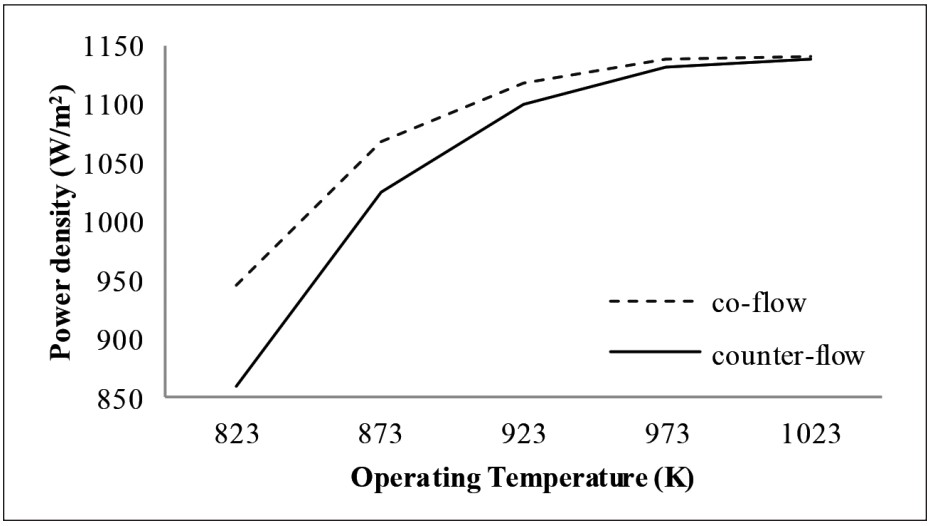


Fig.5. Generated power densities of co-flow and counter-flow models at various temperatures

The different performances of the co-flow and counter-flow models can be investigated through the overall cell voltage and total voltage losses due to the Nernst loss, ohmic loss, cathode and anode activation losses (Eqn. [19]-[21]). By comparing the voltage loss for the different gas flow models at 873 K, it was observed that cathode activation loss for the co-flow is 30% lower as compared that of counter-flow (Fig.6). Cathode and anode activation losses contributed to a high percentage of total voltage losses for the counter-flow model, which caused the lower overall cell voltage and lower power density, compared to the co-flow model. The higher cathode activation loss for the counter-flow model was caused by the lower temperature distribution across the inlet and outlet of the cathode oxidant channel as mentioned above. The relationship of the temperature and the cathode activation polarization was explained in equation [21]. At the temperature of 873 K, the co-flow model attained a higher overall cell voltage of 0.7596 V as compared to that of counter-flow model, which was 0.7331 V. Higher overall cell voltage contributed to higher generated power density at constant current output.

At the operating temperature of 973 K, the co-flow model produced a higher overall cell voltage of 0.8132 V as compared to counter-flow model, which was 0.8021 V. The differences of the overall cell voltage were small as compared to the different gas flow models that operated at the temperature of 873 K. The total voltage losses for both co-flow models operated at 873 K (Fig.6) and 973 K (Fig.7) were reduced as compared to counter-flow models, especially the for the cathode and anode activation losses. It is believed that the higher temperature distribution from the cathode inlet to the outlet of the co-flow models was the main reason for the dramatically reduction of the cathode activation loss (Fig.8).

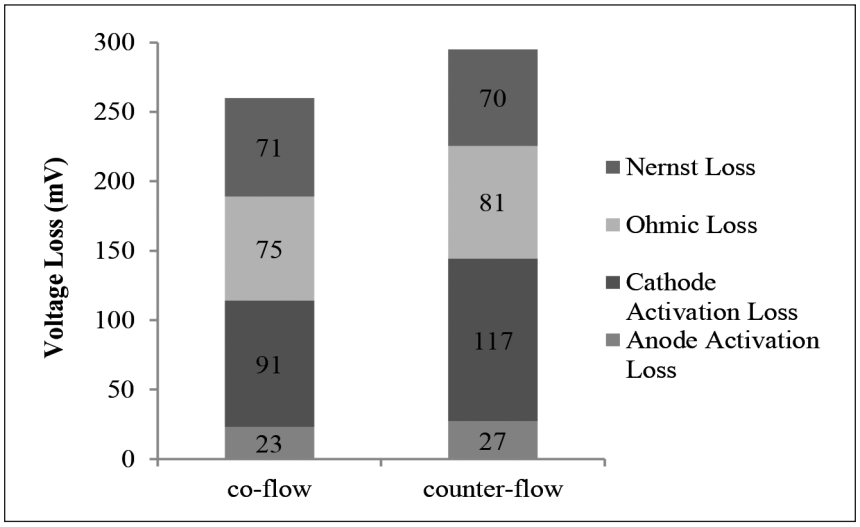


Fig.6. Voltage losses of co-flow and counter-flow models operated at 873 K and average current density of 1400 A/cm^2

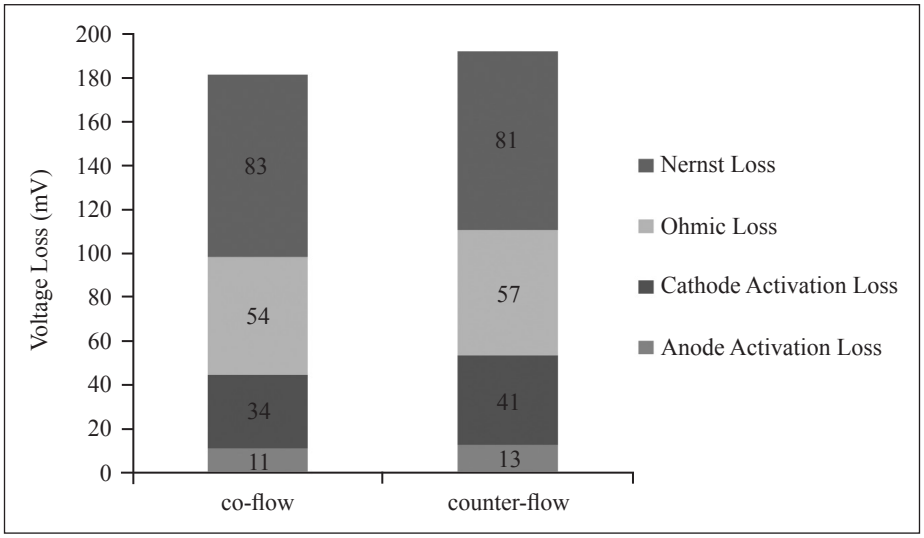


Fig.7. Voltage losses of co-flow and counter-flow models operated at 973 K and average current density of 1400 A/cm^2

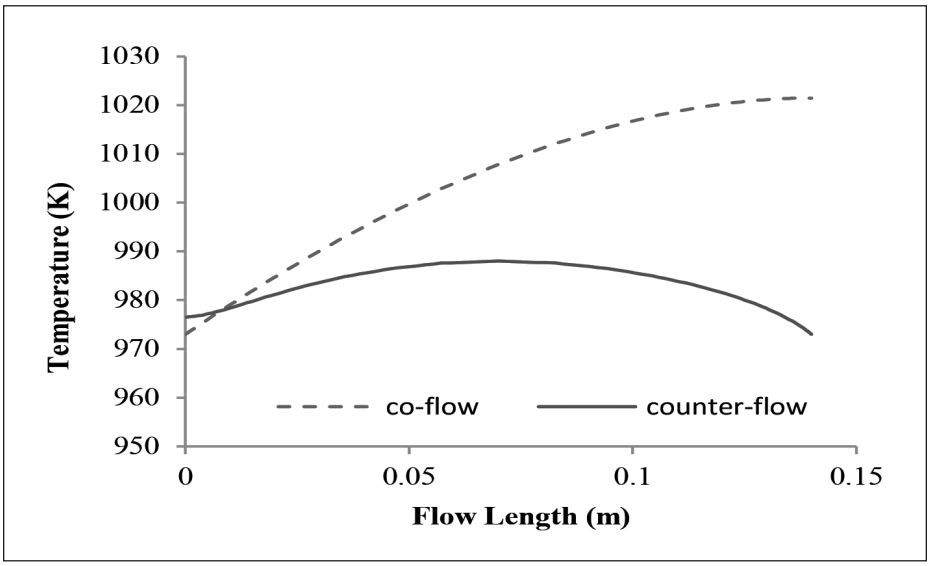


Fig.8. Temperature distributions of co-flow and counter-flow models operated at 973 K

CONCLUSION

The co-flow model was found to have a higher performance as compared to the counter-flow model because of the higher temperature distribution across the cathode inlet to the outlet. The lower temperature distribution of the counter-flow model was caused by the heat regulation from the lower temperature of both anode and cathode inlets. The higher temperature

distribution has the advantages in reducing the total voltage loss especially in the reduction of cathode and anode activation losses. By reducing the total voltage losses, the overall cell voltage and generated power density was increased. It is suggested that the co-flow model is suitable for the MCFC to be operated at lower temperatures of 823-923K to generate a higher power density. The counter-flow MCFC model is more appropriate for MCFC operated at temperature of 923-973K to generate high power density and at the same time maintaining the lower temperature profiles.

ACKNOWLEDGMENTS

The authors acknowledge the Sarawak Energy Berhad, Malaysia, for sponsoring the licenses of COMSOL 4.1 Multiphysics engineering simulation software and Curtin Sarawak, Malaysia, for the High Performance Computer facilities.

REFERENCES

Baranak, M., & Atakul, H. (2007). A basic model for analysis of molten carbonate fuel cell. *Journal of Power Sources*, *72*, 831 - 839.

He, W., & Chen, Q. (1995). Three-dimensional simulation of a molten carbonate fuel cell stack under transient conditions. *Journal of Power Sources*, *55,* 25 - 32.

Huppmann, G. (2009). Fuel cells – molten carbonate fuel cells | Systems. In J. Garche (Ed.), *Encyclopedia of Electrochemical Power Sources* (p. 479 - 496). Amsterdam: Elsevier.

Kim, Y. J., Chang, I. G., Lee, T. W., & Chung, M. K. (2010). Effects of relative gas flow direction in the anode and cathode on the performance characteristics of a Molten Carbonate Fuel Cell. *Fuel*, *89*, 1091 - 1028.

Koh, J. H., Seo, H. K., Yoo, Y. S., & Lim, H. C. (2002). Consideration of numerical simulation parameters and heat transfer models for a molten carbonate fuel cell stack. *Chemical Engineering Journal*, *87*, 367 - 379.

Ma, Z., Venkataraman, R., & Farooque, M. (2009). Fuel cells – molten carbonate fuel cells | Modeling. In J. Garche (Ed.), *Encyclopedia of Electrochemical Power Sources* (pp. 519 – 532). Amsterdam: Elsevier.

Munoz de Escalona, J. M., Sanchez, D., Chacartegui, R., & Sanchez, T. (2011). A step-by-step methodology to construct a model of performance of molten carbonate fuel cells with internal reforming. *International Journal of Hydrogen Energy*, *36,* 15739 - 15751.

Yu. L. J., Ren, G. P., & Jiang, X. M. (2008). Experimental and analytical investigation of molten carbonate fuel cell stack. *Energy Conversion & Management*, *49*, 873 - 879.

Yuh, C., & Farooque, M. (2009). Fuel cells – molten carbonate fuel cells | Materials and Life Considerations. In J. Garche (Ed.), *Encyclopedia of Electrochemical Power Sources* (pp. 497 506). Amsterdam: Elsevier.

Yuh, C. Y., & Selman, J. R. (1984). Polarization of the molten carbonate fuel cell anode and cathode. *Journal of The Electrochemical Society*, *131*, 2062 - 2068.

REFEREES FOR THE PERTANIKA
JOURNAL OF SCIENCE AND TECHNOLOGY

VOL. 22(2) JUL. 2014

The Editorial Board of the Journal of Science and Technology wishes to thank the following for acting as referees for manuscripts published in this issue of JST.

Abdul Ghapor Hussin
(UPNM, Malaysia)

Aileen Tan Shau-Hwai
(USM, Malaysia)

Ajau Danis
(UiTM, Malaysia)

Ani Shabri
(UTM, Malaysia)

Azni Zain Ahmed
(UiTM, Malaysia)

Bassim H. Hameed
(USM, Malaysia)

Bidisha Ghosh
(University of Dublin, Ireland)

Biswa Mohan Biswal
(USM, Malaysia)

Chai Siang Piao
(Monash Universiti, Malaysia)

Chandrashekher Rivonker
(Goa University, India)

Dzati Athiar
(USM, Malaysia)

Gwendoline Ee Cheng Lian
(UPM, Malaysia)

Harijono Djojodihardjo
(UPM, Malaysia)

Hejar Abdul Rahman
(UPM, Malaysia)

Heng Swee Huay
(MMU, Malaysia)

Ibrahim Mohamed
(UM, Malaysia)

Ismail Ibrahim Latif
(Diyala University, Iraq)

Jobrun Nandong
(Curtin University Sarawak, Malaysia)

Lee Yeong Yeh
(USM, Malaysia)

Lim Chin Wai
(UNITEN, Malaysia)

Lim Chy Ing
(Curtin University Sarawak, Malaysia)

Maria Justine
(UiTM, Malaysia)

Mariam Mohamad
(UiTM, Malaysia)

Melati Khairuddean
(USM, Malaysia)

Miroslav Radojevic
(The University of Nottingham, Malaysia Campus)

Mohd Nazrul Hakim Abdullah
(UPM, Malaysia)

Noor Atinah Ahmad
(USM, Malaysia)

Ong Hong Choon
(USM, Malaysia)

Perumal Kumar
(Curtin University Sarawak, Malaysia)

Ranjan Sarukkalige
(Curtin University Australia)

Satish Kumar Singh
(Jaypee University of Engineering and Technology, India)

Semeru Kasni
(Politeknik Negeri Bandung, Indonesia)

Shaharin Ibrahim
(UPM, Malaysia)

Shariful Islam
(Bangladesh University of Engineering and Technology, Bangladesh)

Sharipah Soaad Syed Yahaya
(UUM, Malaysia)

Sugandhi Rao
(Manipal University, India)

Sunil Kumar Maity
(Indian Institute of Technology, India)

T. Mohanraj
(SASTRA University, India)

Tang Fu Ee
(Curtin University Sarawak, Malaysia)

Tapas Mandal
(National Institute of Technology, India)

Thong Meow Keong
(UM, Malaysia)

Umar Hamzah
(UKM, Malaysia)

Valerie Jong Siaw Wee
(Curtin University Sarawak, Malaysia)

Wan Ainun Mior Othman
(UM, Malaysia)

Yap Chee Kong
(UPM, Malaysia)

Yap Yee Jiun
(UM, Malaysia)

Yeoh Kar Kheng
(USM, Malaysia)

Zaleha Kassim
(UMT, Malaysia)

UPM- Universiti Putra Malaysia
USM- Universiti Sains Malaysia
UM- Universiti Malaya
UKM- Universiti Kebangsaan Malaysia
UTM- Universiti Teknologi Malaysia
UiTM- Universiti Teknologi MARA Malaysia
UUM- Universiti Utara Malaysia
UPNM- Universiti Pertahanan Nasional Malaysia
UNITEN- Universiti Tenaga Nasional
UMT- Universiti Malaysia Terengganu

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at nayan@upm.my.

Any inclusion or exclusion of name(s) on this page does not commit the *Pertanika* Editorial Office, nor the UPM Press or the University to provide any liability for whatsoever reason.

Pertanika

Our goal is to bring high quality research to the widest possible audience

Journal of Science & Technology

INSTRUCTIONS TO AUTHORS

(Manuscript Preparation & Submission Guidelines)

Revised: June 2014

We aim for excellence, sustained by a responsible and professional approach to journal publishing.
We value and support our authors in the research community.

Please read the guidelines and follow these instructions carefully; doing so will ensure that the publication of your manuscript is as rapid and efficient as possible. The Editorial Board reserves the right to return manuscripts that are not prepared in accordance with these guidelines.

About the Journal

Pertanika is an international peer-reviewed journal devoted to the publication of original papers, and it serves as a forum for practical approaches to improving quality in issues pertaining to tropical agriculture and its related fields. *Pertanika* began publication in 1978 as Journal of Tropical Agricultural Science. In 1992, a decision was made to streamline *Pertanika* into three journals to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university. The revamped Journal of Science and Technology (JST) is now focusing on research in science and engineering, and its related fields. Other *Pertanika* series include Journal of Tropical Agricultural Science (JTAS); and Journal of Social Sciences and Humanities (JSSH).

JST is published in **English** and it is open to authors around the world regardless of the nationality. It is currently published two times a year i.e. in **January** and **July**.

Goal of *Pertanika*

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 12 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

Indexing of *Pertanika*

Pertanika is now over 33 years old; this accumulated knowledge has resulted in Pertanika JST being indexed in SCOPUS (Elsevier), EBSCO, Thomson (ISI) Web of Knowledge [CAB Abstracts], DOAJ, Google Scholar, ERA, ISC and MyAIS.

Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

We also have views on the future of our journals. The emergence of the online medium as the predominant vehicle for the 'consumption' and distribution of much academic research will be the ultimate instrument in the dissemination of the research news to our scientists and readers.

Aims and Scope

Pertanika Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: *bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management*, and related fields of study.

Editorial Statement

Pertanika is the official journal of Universiti Putra Malaysia. The abbreviation for *Pertanika* Journal of Science & Technology is *Pertanika J. Sci. Technol.*

Guidelines for Authors

Publication policies

Pertanika policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in Proceedings. Please refer to Pertanika's **Code of Ethics** for full details.

Editorial process

Authors are notified on receipt of a manuscript and upon the editorial decision regarding publication.

Manuscript review: Manuscripts deemed suitable for publication are sent to the Editorial Board members and/or other reviewers. We encourage authors to suggest the names of possible reviewers. Notification of the editorial decision is usually provided within to eight to ten weeks from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

Author approval: Authors are responsible for all statements in articles, including changes made by editors. The liaison author must be available for consultation with an editor of *The Journal* to answer questions during the editorial process and to approve the edited copy. Authors receive edited typescript (not galley proofs) for final approval. Changes **cannot** be made to the copy after the edited version has been approved.

Manuscript preparation

Pertanika accepts submission of mainly four types of manuscripts. Each manuscript is classified as **regular** or **original** articles, **short communications**, **reviews**, and proposals for **special issues**. Articles must be in **English** and they must be competently written and argued in clear and concise grammatical English. Acceptable English usage and syntax are expected. Do not use slang, jargon, or obscure abbreviations or phrasing. Metric measurement is preferred; equivalent English measurement may be included in parentheses. Always provide the complete form of an acronym/abbreviation the first time it is presented in the text. Contributors are strongly recommended to have the manuscript checked by a colleague with ample experience in writing English manuscripts or an English language editor.

Linguistically hopeless manuscripts will be rejected straightaway (e.g., when the language is so poor that one cannot be sure of what the authors really mean). This process, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

The instructions for authors must be followed. Manuscripts not adhering to the instructions will be returned for revision without review. Authors should prepare manuscripts according to the guidelines of *Pertanika*.

1. Regular article

Definition: Full-length original empirical investigations, consisting of introduction, materials and methods, results and discussion, conclusions. Original work must provide references and an explanation on research findings that contain new and significant findings.

Size: Should not exceed 5000 words or 8-10 printed pages (excluding the abstract, references, tables and/or figures). One printed page is roughly equivalent to 3 type-written pages.

2. Short communications

Definition: Significant new information to readers of the Journal in a short but complete form. It is suitable for the publication of technical advance, bioinformatics or insightful findings of plant and animal development and function.

Size: Should not exceed 2000 words or 4 printed pages, is intended for rapid publication. They are not intended for publishing preliminary results or to be a reduced version of Regular Papers or Rapid Papers.

3. Review article

Definition: Critical evaluation of materials about current research that had already been published by organizing, integrating, and evaluating previously published materials. Re-analyses as meta-analysis and systemic reviews are encouraged. Review articles should aim to provide systemic overviews, evaluations and interpretations of research in a given field.

Size: Should not exceed 4000 words or 7-8 printed pages.

4. Special issues

Definition: Usually papers from research presented at a conference, seminar, congress or a symposium.

Size: Should not exceed 5000 words or 8-10 printed pages.

5. Others

Definition: Brief reports, case studies, comments, Letters to the Editor, and replies on previously published articles may be considered.

Size: Should not exceed 2000 words or up to 4 printed pages.

With few exceptions, original manuscripts should not exceed the recommended length of 6 printed pages (about 18 typed pages, double-spaced and in 12-point font, tables and figures included). Printing is expensive, and, for the Journal, postage doubles when an issue exceeds 80 pages. You can understand then that there is little room for flexibility.

Long articles reduce the Journal's possibility to accept other high-quality contributions because of its 80-page restriction. We would like to publish as many good studies as possible, not only a few lengthy ones. (And, who reads overly long articles anyway?) Therefore, in our competition, short and concise manuscripts have a definite advantage.

Format

The paper should be formatted in one column format with at least 4cm margins and 1.5 line spacing throughout. Authors are advised to use Times New Roman 12-point font. Be especially careful when you are inserting special characters, as those inserted in different fonts may be replaced by different characters when converted to PDF files. It is well known that 'μ' will be replaced by other characters when fonts such as 'Symbol' or 'Mincho' are used.

A maximum of eight keywords should be indicated below the abstract to describe the contents of the manuscript. Leave a blank line between each paragraph and between each entry in the list of bibliographic references. Tables should preferably be placed in the same electronic file as the text. Authors should consult a recent issue of the Journal for table layout.

Every page of the manuscript, including the title page, references, tables, etc. should be numbered. However, no reference should be made to page numbers in the text; if necessary, one may refer to sections. Underline words that should be in italics, and do not underline any other words.

We recommend that authors prepare the text as a **Microsoft Word** file.

1. Manuscripts in general should be organised in the following order:

    - **Page 1: Running title.** (Not to exceed 60 characters, counting letters and spaces). This page should **only** contain the running title of your paper. The running title is an abbreviated title used as the running head on every page of the manuscript.

      In addition, the **Subject areas m**ost relevant to the study **must be indicated on this page.** Select the appropriate subject areas from the Scope of the Journals provided in the Manuscript Submission Guide.

      A **list** of n**umber of black and white / colour figures and tables should also be indicated on this page.** Figures submitted in color will be printed in colour. See "*5. Figures & Photographs"* for details.

    - **Page 2: Author(s) and Corresponding author information.** This page should contain the **full title** of your paper with name(s) of all the authors, institutions and corresponding author's name, institution and full address (Street address, telephone number (including extension), hand phone number, fax number and e-mail address) for editorial correspondence. The names of the authors **must** be abbreviated following the international naming convention. e.g. Salleh, A.B., Tan, S.G., or Sapuan, S.M.

      Authors' addresses. Multiple authors with different addresses must indicate their respective addresses separately by superscript numbers:

      George Swan[1] and Nayan Kanwal[2]
      [1]Department of Biology, Faculty of Science, Duke University, Durham, North Carolina, USA.
      [2]Office of the Deputy Vice Chancellor (R&I), Universiti Putra Malaysia, Serdang, Malaysia.

    - **Page 3**: This page should **repeat** the **full title** of your paper with only the **Abstract** (the abstract should be less than 250 words for a Regular Paper and up to 100 words for a Short Communication). **Keywords** must also be provided on this page (Not more than eight keywords in alphabetical order).

    - **Page 4 and subsequent pages**: This page should begin with the **Introduction** of your article and the rest of your paper should follow from page 5 onwards.

    Abbreviations. Define alphabetically, other than abbreviations that can be used without definition. Words or phrases that are abbreviated in the introduction and following text should be written out in full the first time that they appear in the text, with each abbreviated form in parenthesis. Include the common name or scientific name, or both, of animal and plant materials.

    Footnotes. Current addresses of authors if different from heading.

2. **Text.** Regular Papers should be prepared with the headings **Introduction**, **Materials and Methods**, **Results and Discussion, Conclusions** in this order. Short Communications should be prepared according to "*8. Short Communications."* below.

3. **Tables.** All tables should be prepared in a form consistent with recent issues of *Pertanika* and should be numbered consecutively with Arabic numerals. Explanatory material should be given in the table legends and footnotes. Each

table should be prepared on a separate page. (Note that when a manuscript is accepted for publication, tables must be submitted as data - .doc, .rtf, Excel or PowerPoint file- because tables submitted as image data cannot be edited for publication.)

4. **Equations and Formulae.** These must be set up clearly and should be typed triple spaced. Numbers identifying equations should be in square brackets and placed on the right margin of the text.

5. **Figures & Photographs.** Submit an original figure or photograph. Line drawings must be clear, with high black and white contrast. Each figure or photograph should be prepared on a separate sheet and numbered consecutively with Arabic numerals. Appropriate sized numbers, letters and symbols should be used, no smaller than 2 mm in size after reduction to single column width (85 mm), 1.5-column width (120 mm) or full 2-column width (175 mm).

6. Failure to comply with these specifications will require new figures and delay in publication. For electronic figures, create your figures using applications that are capable of preparing high resolution TIFF files acceptable for publication. In general, we require **300 dpi or higher resolution for coloured and half-tone artwork** and **1200 dpi or higher for line drawings**. For review, you may attach low-resolution figures, which are still clear enough for reviewing, to keep the file of the manuscript under 5 MB. Illustrations may be produced at extra cost in colour at the discretion of the Publisher; the author could be charged Malaysian Ringgit 50 for each colour page.

7. **References.** Literature citations in the text should be made by name(s) of author(s) and year. For references with more than two authors, the name of the first author followed by *'et al*.' should be used.

   Swan and Kanwal (2007) reported that …
   The results have been interpreted (Kanwal *et al.* 2009).

   - References should be listed in alphabetical order, by the authors' last names. For the same author, or for the same set of authors, references should be arranged chronologically. If there is more than one publication in the same year for the same author(s), the letters 'a', 'b', etc., should be added to the year.

   - When the authors are more than 11, list 5 authors and then et al.

   - Do not use indentations in typing References. Use one line of space to separate each reference. The name of the journal should be written in full. For example:

     - Jalaludin, S. (1997a). Metabolizable energy of some local feeding stuff. *Tumbuh, 1,* 21-24.
     - Jalaludin, S. (1997b). The use of different vegetable oil in chicken ration. *Malayan Agriculturist, 11,* 29-31.
     - Tan, S. G., Omar, M. Y., Mahani, K. W., Rahani, M., & Selvaraj, O. S. (1994). Biochemical genetic studies on wild populations of three species of green leafhoppers *Nephotettix* from Peninsular Malaysia. *Biochemical Genetics*, *32,* 415 - 422.

   - In case of citing an author(s) who has published more than one paper in the same year, the papers should be distinguished by addition of a small letter as shown above, e.g. Jalaludin (1997a); Jalaludin (1997b).

   - Unpublished data and personal communications should not be cited as literature citations, but given in the text in parentheses. 'In press' articles that have been accepted for publication may be cited in References. Include in the citation the journal in which the 'in press' article will appear and the publication date, if a date is available.

8. **Examples of other reference citations**:

   Monographs: Turner, H. N., & Yong, S. S. Y. (2006). *Quantitative Genetics in Sheep Breeding.* Ithaca: Cornell University Press.

   Chapter in Book: Kanwal, N. D. S. (1992). Role of plantation crops in Papua New Guinea economy. In Angela R. McLean (Ed.), *Introduction of livestock in the Enga province PNG* (p. 221-250). United Kingdom: Oxford Press.

   Proceedings: Kanwal, N. D. S. (2001). Assessing the visual impact of degraded land management with landscape design software. In Kanwal, N. D. S., & Lecoustre, P. (Eds.), *International forum for Urban Landscape Technologies* (p. 117-127). Lullier, Geneva, Switzerland: CIRAD Press.

9. **Short Communications** should include **Introduction, Materials and Methods, Results and Discussion, Conclusions** in this order. Headings should only be inserted for Materials and Methods. The abstract should be up to 100 words, as stated above. Short Communications must be 5 printed pages or less, including all references, figures and tables. References should be less than 30. A 5 page paper is usually approximately 3000 words plus four figures or tables (if each figure or table is less than 1/4 page).

   *Authors should state the total number of words (including the Abstract) in the cover letter. Manuscripts that do not fulfill these criteria will be rejected as Short Communications without review.

STYLE OF THE MANUSCRIPT

Manuscripts should follow the style of the latest version of the Publication Manual of the American Psychological Association (APA). The journal uses American or British spelling and authors may follow the latest edition of the Oxford Advanced Learner's Dictionary for British spellings.

SUBMISSION OF MANUSCRIPTS

All articles should be submitted electronically using the ScholarOne web-based system. ScholarOne, a Thomson Reuters product provides comprehensive workflow management systems for scholarly journals. For more information, go to our web page and click "**Online Submission**".

Alternatively, you may submit the electronic files (cover letter, manuscript, and the **Manuscript Submission Kit** comprising *Declaration* and *Referral* forms) via email directly to the Executive Editor. If the files are too large to email, mail a CD containing the files. The **Manuscript Submission Guide** and **Submission Kit** are available from the *Pertanika*'s home page at http://www.pertanika.upm.edu.my/ or from the Executive Editor's office upon request.

All articles submitted to the journal **must comply** with these instructions. Failure to do so will result in return of the manuscript and possible delay in publication.

Please do **not** submit manuscripts to the editor-in-chief or to any other office directly. All manuscripts must be **submitted through the executive editor's office** to be properly acknowledged and rapidly processed at the address below:

Dr. Nayan KANWAL
Chief Executive Editor
Pertanika Journals, UPM Press
Office of the Deputy Vice Chancellor (R&I)
IDEA Tower II, UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor
Malaysia

E-mail: nayan@upm.my; journal.officer@gmail.com; tel: + 603-8947 1622.
or visit our website at http://www.pertanika.upm.edu.my/ for further information.

Authors should retain copies of submitted manuscripts and correspondence, as materials can not be returned. Authors are required to inform the Executive Editor of any change of address which occurs whilst their papers are in the process of publication.

Cover letter

All submissions must be accompanied by a cover letter detailing what you are submitting. Papers are accepted for publication in the journal on the understanding that the article is original and the content has not been published or submitted for publication elsewhere. This must be stated in the cover letter.

The cover letter must also contain an acknowledgement that all authors have contributed significantly, and that all authors are in agreement with the content of the manuscript.

The cover letter of the paper should contain (i) the title; (ii) the full names of the authors; (iii) the addresses of the institutions at which the work was carried out together with (iv) the full postal and email address, plus facsimile and telephone numbers of the author to whom correspondence about the manuscript should be sent. The present address of any author, if different from that where the work was carried out, should be supplied in a footnote.

As articles are double-blind reviewed, material that might identify authorship of the paper should be placed on a cover sheet.

Peer review

Pertanika follows a **double-blind peer-review** process. Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

In the peer-review process, three referees independently evaluate the scientific quality of the submitted manuscripts. Authors are encouraged to indicate in the **Referral form** using the **Manuscript Submission Kit** the names of three potential reviewers, but the editors will make the final choice. The editors are not, however, bound by these suggestions.

Manuscripts should be written so that they are intelligible to the professional reader who is not a specialist in the particular field. They should be written in a clear, concise, direct style. Where contributions are judged as acceptable for publication on the basis of content, the Editor reserves the right to modify the typescripts to eliminate ambiguity and repetition and improve communication between author and reader. If extensive alterations are required, the manuscript will be returned to the author for revision.

The Journal's review process

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are seven steps to the editorial review process:

1. The executive editor and the editorial board examine the paper to determine whether it is appropriate for the journal and should be reviewed. If not appropriate, the manuscript is rejected outright and the author is informed.

2. The executive editor sends the article-identifying information having been removed, to three reviewers. Typically, one of these is from the Journal's editorial board. Others are specialists in the subject matter represented by the article. The executive editor asks them to complete the review in three weeks and encloses two forms: (a) referral form B and (b) reviewer's comment form along with reviewer's guidelines. Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the literature.

3. The executive editor, in consultation with the editor-in-chief, examines the reviews and decides whether to reject the manuscript, invite the author(s) to revise and resubmit the manuscript, or seek additional reviews. Final acceptance or rejection rests with the Editorial Board, who reserves the right to refuse any material for publication. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the author) are forwarded to the author. If a revision is indicated, the editor provides guidelines for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors submit a revised version of the paper to the executive editor along with specific information describing how they have answered' the concerns of the reviewers and the editor.

5. The executive editor sends the revised paper out for review. Typically, at least one of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the executive editor in consultation with the editorial board and the editor-in-chief examine their comments and decide whether the paper is ready to be published, needs another round of revisions, or should be rejected.

7. If the decision is to accept, the paper is sent to that Press and the article should appear in print in approximately three months. The Publisher ensures that the paper adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any queries by the Publisher. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, only essential changes are accepted. Finally, the article appears in the pages of the Journal and is posted on-line.

English language editing

Pertanika **emphasizes** on the linguistic accuracy of every manuscript published. Thus all authors are required to get their manuscripts edited by **professional English language editors**. Author(s) **must provide a certificate** confirming that their manuscripts have been adequately edited. A proof from a recognised editing service should be submitted together with the cover letter at the time of submitting a manuscript to Pertanika**. All costs will be borne by the author(s)**.

This step, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

Author material archive policy

Authors who require the return of any submitted material that is rejected for publication in the journal should indicate on the cover letter. If no indication is given, that author's material should be returned, the Editorial Office will dispose of all hardcopy and electronic material.

Copyright

Authors publishing the Journal will be asked to sign a declaration form. In signing the form, it is assumed that authors have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the form, and must sign the form or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed form has been received.

Lag time

A decision on acceptance or rejection of a manuscript is reached in 3 to 4 months (average 14 weeks). The elapsed time from submission to publication for the articles averages 5-6 months.

Hardcopies of the Journals and off prints

Under the Journal's open access initiative, authors can choose to download free material (via PDF link) from any of the journal issues from Pertanika's website. Under "Browse Journals" you will see a link entitled "Current Issues" or "Archives". Here you will get access to all back-issues from 1978 onwards.

The **corresponding author** for all articles will receive one complimentary hardcopy of the journal in which his/her articles is published. In addition, 20 off prints of the full text of their article will also be provided. Additional copies of the journals may be purchased by writing to the executive editor.

Why should you publish in *Pertanika*?

BENEFITS TO AUTHORS

PROFILE: Our journals are circulated in large numbers all over Malaysia, and beyond in Southeast Asia. Our circulation covers other overseas countries as well. We ensure that your work reaches the widest possible audience in print and online, through our wide publicity campaigns held frequently, and through our constantly developing electronic initiatives such as Web of Science Author Connect backed by Thomson Reuters.

QUALITY: Our journals' reputation for quality is unsurpassed ensuring that the originality, authority and accuracy of your work is fully recognised. Each manuscript submitted to Pertanika undergoes a rigid originality check. Our double-blind peer refereeing procedures are fair and open, and we aim to help authors develop and improve their scientific work. Pertanika is now over 35 years old; this accumulated knowledge has resulted in our journals being indexed in SCOPUS (Elsevier), Thomson (ISI) Web of Knowledge [BIOSIS & CAB Abstracts], EBSCO, DOAJ, Google Scholar, AGRICOLA, ERA, ISC, Citefactor, Rubriq and MyAIS.

AUTHOR SERVICES: We provide a rapid response service to all our authors, with dedicated support staff for each journal, and a point of contact throughout the refereeing and production processes. Our aim is to ensure that the production process is as smooth as possible, is borne out by the high number of authors who prefer to publish with us.

CODE OF ETHICS: Our Journal has adopted a Code of Ethics to ensure that its commitment to integrity is recognized and adhered to by contributors, editors and reviewers. It warns against plagiarism and self-plagiarism, and provides guidelines on authorship, copyright and submission, among others.

PRESS RELEASES: Landmark academic papers that are published in Pertanika journals are converted into press releases as a unique strategy for increasing visibility of the journal as well as to make major findings accessible to non-specialist readers. These press releases are then featured in the university's UK-based research portal, ResearchSEA, for the perusal of journalists all over the world.

LAG TIME: The elapsed time from submission to publication for the articles averages 4 to 5 months. A decision on acceptance of a manuscript is reached in 3 to 4 months (average 14 weeks).

An
Award-Winning
International-Malaysian
Journal
—MAY 2014

About the Journal

Pertanika is an international multidisciplinary peer-reviewed leading journal in Malaysia which began publication in 1978. The journal publishes in three different areas — Journal of Tropical Agricultural Science (JTAS); Journal of Science and Technology (JST); and Journal of Social Sciences and Humanities (JSSH).

JTAS is devoted to the publication of original papers that serves as a forum for practical approaches to improving quality in issues pertaining to **tropical agricultural research-** or related fields of study. It is published four times a year in **February**, **May**, **August** and **November**.

JST caters for **science and engineering research-** or related fields of study. It is published twice a year in **January** and **July**.

JSSH deals in **research or theories in social sciences and humanities research**. It aims to develop as a flagship journal with a focus on emerging issues pertaining to the social and behavioural sciences as well as the humanities, particularly in the Asia Pacific region. It is published four times a year in **March**, **June**, **September** and **December**.

PERTANIKA JOURNALS

Call for Papers 2014-15

now accepting submmissions...

Pertanika invites you to explore frontiers from all key areas of **agriculture**, **science and technology** to **social sciences and humanities**.

Original research and review articles are invited from scholars, scientists, professors, post-docs, and university students who are seeking publishing opportunities for their research papers through the Journal's three titles; JTAS, JST & JSSH. Preference is given to the work on leading and innovative research approaches.

Pertanika is a fast track peer-reviewed and open-access academic journal published by **Universiti Putra Malaysia**. To date, Pertanika Journals have been indexed by many important databases. Authors may contribute their scientific work by publishing in UPM's hallmark SCOPUS & ISI indexed journals.

Our journals are open access - international journals. Researchers worldwide will have full access to all the articles published online and be able to download them with **zero subscription fee**.

Pertanika uses online article submission, review and tracking system for quality and quick review processing backed by Thomson Reuter's ScholarOne™. Journals provide rapid publication of research articles through this system.

For details on the Guide to Online Submissions, visit
http://www.pertanika.upm.edu.my/guide_online_submission.php

Questions regarding submissions should only be directed to the **Chief Executive Editor**, Pertanika Journals.

Remember, ***Pertanika*** is the resource to support you in strengthening research and research management capacity.

Address your submissions to:
The Chief Executive Editor
Tel: +603 8947 1622
nayan@upm.my

Journal's profile: www.pertanika.upm.edu.my

Pertanika Journal of Science & Technology
Vol. 22 (2) Jul. 2014

Contents

Pertanika Editorial Office, Journal Division
Office of the Deputy Vice Chancellor (R&I),
1st Floor, IDEA Tower II,
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Malaysia

http://www.pertanika.upm.edu.my/
E-mail: executive_editor.pertanika@upm.my
Tel: +603 8947 1622/1620

PENERBIT UPM PRESS
http://penerbit.upm.edu.my
E-mail : penerbit@putra.upm.edu.my
Tel : +603 8946 8855/8854
Fax : +603 8941 6172

ISSN 0128-7680
07 >
9 770128 768083