UNIVERSITI PUTRA MALAYSIA

SOFT SENSOR MODELLING FOR OPTIMIZATION OF DISTRIBUTION CONTROL SYSTEM IN OIL REFINERIES BY APPLYING HYBRID DATA MINING TECHNIQUES

ALI HUSSEIN HUMOD AL JLIBAWI

FK 2022 24

SOFT SENSOR MODELLING FOR OPTIMIZATION OF DISTRIBUTION CONTROL SYSTEM IN OIL REFINERIES BY APPLYING HYBRID DATA MINING TECHNIQUES

By

ALI HUSSEIN HUMOD AL JLIBAWI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

May 2021

COPYRIGHT

DEDICATION

This study is sincerely dedicated to my beloved parents, who have been my source of inspiration and gave me strength, who continually provide their moral, spiritual, emotional, and financial support.

To my esteemed supervisors who were the river of knowledge.

To my brother, sisters, relatives, teachers, friends, and classmates who shared their words of advice and encouragement to finish this study. Tag Mint with the scent of roses.

And lastly, I dedicated this book to almighty Allah, thank you for the guidance, strength, power of the mind, protection, and skills, and for giving me a healthy life. All of these, I offer to you.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

SOFT SENSOR MODELLING FOR OPTIMIZATION OF DISTRIBUTION CONTROL SYSTEM IN OIL REFINERIES BY APPLYING HYBRID DATA MINING TECHNIQUES

By

ALI HUSSEIN HUMOD AL JLIBAWI

May 2021

Chairman : Associate Professor Ir. Mohammad Lutfi bin Othman. PhD
Faculty : Engineering

A data-driven soft sensor is a sensor that uses data from available online sensors (such as temperature, pressure, and flow rate) to forecast quality attributes that cannot be monitored naturally or can only be measured at a high cost, infrequently, or with long delays. Oil refineries use control systems, which are connected to PLCs or distributed control systems (DCS). The DCS system is the unit responsible for attaining and providing such data as daily reports for the process, to construct soft sensors utilizing past data from the laboratory observations/measurements and processes data. To determine the quality of crude oil, one can consider the prolonged in-depth laboratory-based tests or the rather expensive approach of online analysers. Implementing light naphtha product quality criterion measurement is surrounded with several essential concerns such as missing data, detecting outliers, selecting input variables and training, validating and maintaining the soft sensor which must be addressed and delt with beforehand. Hence, obtaining heavy-duty soft sensors for oil refineries remained a challenge which in return makes it difficult to improve the end product while simultaneously increasing production.

The adaptive neuro-fuzzy inference system, a hybrid soft computing technology combining a fuzzy logic system (FLS) and a neural network (NN), was used to develop a virtual sensor adaptive neural fuzzy inference system (ANFIS) in this research. Rough set theory (RST) and its discretization approach were used to minimize the fuzzy rule sets and redact characteristics of the decision table attributes. It was then used to create the soft sensor modelling for ANFIS, while

using the discretization method helped in converting continuous data into a comprehensible data mining format that can be used for data mining. This research is aimed at monitoring and controlling light naphtha production by examining the American petroleum institute gravity (API gravity) and Reid vapour pressure (RVP) variables in real time. It further aims at breaking the privacy barriers between the oil industries as well as soft sensor modelling for data source interpretation to predict API gravity and RVP in real time for crude oil unit's top splitter in the refinery. By comparing the prediction models to other machine learning techniques, with regard to the proposed prediction model, the root means square error improved to be 0.019 for RVP prediction model and 0.4137 for API prediction model and the determination of correlation yield satisfactory results with value of 0.96 for RVP prediction model and 0.99 for API prediction model , the model has been proven to be accurate, and the simulated soft sensor model has been employed as feedback for the cascade PID controller.

Whether for operator information, cascaded to base-layer process controller, or multivariable controllers, the proposed virtual sensor model can competently replace the actual online analyser. The objectives of this research were realized by the optimization of the controller of the splitter in the crude distillation unit of the AlDoura Oil Refinery's crude distillation unit. The ability to translate the expert's knowledge into the created model using the gaussian membership function, as demonstrated by the ANFIS model, results in excellent generalisation ability. It has been determined that the Al Doura oil refinery's real-time process was explored, and the data received from these two sources was used to expand the information provided by the data collected. Feedback measurement values from a cascade controller positioned at the top of the splitter in a rectifying section's crude distillation unit (CDU) are used to determine each response variable.

A steady-state control system was achieved through the incorporation of an embedded virtual sensor into the suggested adaptive soft sensor paradigm. For the oil refinery's quality control, a cascade ANFIS controller and a soft sensor model were used in the predictive control system's implementation to keep the distillate product's purity within the stipulated range. Overshoots and undershoots are eliminated in the proposed ANFIS-based cascade control compared to the conventional proportional-integral-derivative (PID)-based cascade control. The rise time and settling time are also significantly improved by 26.65 percent and 84.63 percent. Results from other machine learning techniques are also compared to those from the prediction and control models.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PEMODELAN PENDERIA LEMBUT YANG CEKAP DALAM KAWALAN RAMALAN MAJU BAGI MENINGKATKAN PRESTASI PROSES DALAM SISTEM PERKILANGAN MAJU DENGAN MENERAPKAN TEKNIK PERLOMBONGAN DATA PINTAR HIBRID

Oleh

ALI HUSSEIN HUMOD AL JLIBAWI

Mei 2021

Pengerusi : Profesor Madya Ir. Mohammad Lutfi bin Othman, PhD
Fakulti : Kejuruteraan

Penderia lembut dipacu data ialah penderia yang menggunakan data daripada penderia dalam talian yang tersedia (seperti suhu, tekanan dan kadar aliran) untuk meramalkan atribut kualiti yang tidak boleh dipantau secara semula jadi atau hanya boleh diukur pada kos yang tinggi, jarang atau dengan kelewatan yang lama. Penapisan minyak menggunakan sistem kawalan, yang disambungkan kepada PLC atau sistem kawalan teragih (DCS). Sistem DCS ialah unit yang bertanggungjawab untuk mencapai dan menyediakan data seperti laporan harian untuk proses, untuk membina penderia lembut menggunakan data lepas daripada data pemerhatian/ukuran makmal dan proses. Untuk menentukan kualiti minyak mentah, seseorang boleh mempertimbangkan ujian berasaskan makmal mendalam yang berpanjangan atau pendekatan penganalisis dalam talian yang agak mahal. Melaksanakan ukuran kriteria kualiti produk naphtha ringan dikelilingi dengan beberapa kebimbangan penting seperti kehilangan data, pengesanan outlier, memilih pembolehubah input dan latihan, mengesahkan dan mengekalkan sensor lembut yang mesti ditangani dan diselesaikan terlebih dahulu. Oleh itu, mendapatkan penderia lembut tugas berat untuk penapisan minyak kekal sebagai cabaran yang sebaliknya menyukarkan untuk menambah baik produk akhir sambil meningkatkan pengeluaran pada masa yang sama. Sistem inferens neuro-fuzzy adaptif, teknologi pengkomputeran lembut hibrid yang menggabungkan sistem logik kabur (FLS) dan rangkaian saraf (NN), telah digunakan untuk membangunkan sistem inferens kabur saraf adaptif sensor maya (ANFIS) dalam penyelidikan ini. Teori set kasar (RST) dan pendekatan pendiskretannya digunakan untuk meminimumkan set peraturan kabur dan menyunting ciri atribut jadual keputusan. Ia kemudiannya digunakan untuk mencipta pemodelan sensor lembut untuk ANFIS, sambil menggunakan kaedah pendiskretan membantu dalam menukar data berterusan kepada format perlombongan data yang boleh difahami yang boleh digunakan untuk perlombongan data. Penyelidikan ini bertujuan untuk memantau dan mengawal pengeluaran nafta ringan dengan memeriksa pembolehubah graviti institut petroleum Amerika (graviti API) dan tekanan wap Reid (RVP) dalam masa nyata. Ia seterusnya bertujuan untuk memecahkan halangan privasi antara industri minyak serta pemodelan sensor lembut untuk tafsiran sumber data untuk meramalkan graviti API dan RVP dalam masa nyata untuk pembahagi teratas unit minyak mentah dalam penapisan. Dengan membandingkan model ramalan dengan teknik pembelajaran mesin yang lain, berkenaan dengan model ramalan yang dicadangkan, ralat punca bermakna kuasa dua bertambah baik menjadi 0.019 untuk model ramalan RVP dan 0.4137 untuk model ramalan API dan penentuan hasil korelasi hasil yang memuaskan dengan nilai 0.96 untuk model ramalan RVP dan 0.99 untuk model ramalan API , model tersebut telah terbukti tepat, dan model penderia lembut simulasi telah digunakan sebagai maklum balas untuk pengawal PID lata. Sama ada untuk maklumat pengendali, dilantunkan kepada pengawal proses lapisan asas, atau pengawal berbilang pembolehubah, model penderia maya yang dicadangkan boleh menggantikan penganalisis dalam talian sebenar dengan cekap. Objektif penyelidikan ini direalisasikan melalui pengoptimuman pengawal pembahagi dalam unit penyulingan mentah unit penyulingan mentah AlDoura Oil Refinery. Keupayaan untuk menterjemah pengetahuan pakar ke dalam model yang dicipta menggunakan fungsi keahlian gaussian, seperti yang ditunjukkan oleh model ANFIS, menghasilkan keupayaan generalisasi yang sangat baik. Telah ditentukan bahawa proses masa nyata penapisan minyak Al Doura telah diterokai, dan data yang diterima daripada kedua-dua sumber ini digunakan untuk mengembangkan maklumat yang diberikan oleh data yang dikumpul. Nilai pengukuran maklum balas daripada pengawal lata yang diletakkan di bahagian atas pembahagi dalam unit penyulingan mentah (CDU) bahagian pembetulan digunakan untuk menentukan setiap pembolehubah tindak balas. Sistem kawalan keadaan mantap telah dicapai melalui penggabungan penderia maya terbenam ke dalam paradigma penderia lembut adaptif yang dicadangkan. Untuk kawalan kualiti kilang penapisan minyak, pengawal ANFIS lata dan model penderia lembut digunakan dalam pelaksanaan sistem kawalan ramalan untuk memastikan ketulenan produk sulingan dalam julat yang ditetapkan. Overshoot dan undershoot dihapuskan dalam kawalan lata berasaskan ANFIS yang dicadangkan berbanding kawalan lata berasaskan proportional-integral-derivative (PID) konvensional. Masa naik dan masa menetap juga meningkat dengan ketara sebanyak 26.65 peratus dan 84.63 peratus. Keputusan daripada teknik pembelajaran mesin lain juga dibandingkan dengan hasil daripada model ramalan dan kawalan.

ACKNOWLEDGEMENT

To the inspired leader in this world, my father. To Tigris River, my mother, that she fed me with kindness. My brother and sisters, my cousins, and best friends. To my supervisor, Assoc. Prof. Dr. Mohammed Lutfi bin Othman who supports me; without him, this thesis would not be achievable. To my supervisory committee, Prof. Ishak bin Iris with his valued comments and notes, and I would like to thank Dr. Samsul Bahari as well, and Dr. Muayed Sattar Sajitt for his encouragement and support. I would like to have special thanks to the Ministry of Oil in Iraq and AlDoura oil refinery for assisting and providing me with the data and information about the refinery processes. Many thanks for University of Wassit – Faculty of Engineering for their help and support. All thanks and gratitude to my friends and colleagues and everyone who had a great role in this achievement by sharing my joys and the challenges I faced.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Mohammad Lutfi bin Othman, PhD
Associate Professor Ir. Ts.
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Ishak bin Aris, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Samsul Bahari bin Mohd Noor, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Muayed Sattar Sajit Al-Husseiny, PhD
Associate Professor
College of Engineering
Universiti of Wassitt
(Member)

____
ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 14 April 2022

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ____

Name of Chairman of Supervisory Committee: Associate Professor, Ir. Ts Dr. Mohammad Lutfi bin Othman

Signature: ____

Name of Member of Supervisory Committee: Professor Dr. Ishak bin Aris

Signature: ____

Name of Member of Supervisory Committee: Associate Professor Dr. Samsul Bahari bin Mohd Noor

Signature: ____

Name of Member of Supervisory Committee: Associate Professor Dr. Muayed Sattar Sajit Al-Husseiny

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

A/D	Analogue to Digital Convertor
ANFIS	Adaptive Neuro-Fuzzy Inference System
API	American Petroleum Institute
ASTM	The American Society for Testing and Materials
BPSD	Barrels Per Stream Day
CDU	Crude Distillation Unit
CR	Crude Oil
CRD	Crude Oil Desalter
CV	Control Variable
D/A	Digital to Analogue Convertor
DCS	Decentralised Control System
DCS	Distributed Control System
DDMs	Data-Driven Models
DOF	Degree of Freedom
DW	Data Warehouse
EP	End Boiling point
FIS	Fuzzy Inferential System
FLC	Fuzzy Logic Controller
GUI	Graphical User Interface
HPC	High Performance Computing
IBP	Initial Boiling point
IOT	Internet of Thing

KDD	Knowledge Discovery Data Mining
LCN	Local Control Network
LTI	Linear Time Invariant Systems
NN	Neural Network
NP	Naphtha
Obs/sec	Observation/Second
OP	Operation Set Point
OPC	Open Platform Communications
P&ID	Pipe and Instrument Diagram
PCA	Principal Component Analysis
PFD	Process Fault Detection
PID	Proportional Integral Derivative
PIS	Plant Information System
PLC	Programmable Logic Controller
PMs	Process Managers
PV	Process Variable
RST	Rough Set Theory
RVP	Reid Vapour Pressure
SFD	Sensor Fault Detection
SP	Set Point
SVM	Support Vector Machine
VIF	Variance Inflation Factor

CHAPTER 1

INTRODUCTION

1.1 Overview

After providing a brief overview of the study's background, which focused on advanced control systems in oil refineries that utilized machine learning algorithms as the soft sensor concept in order to predict the quality of light naphtha in oil refining, this chapter presents the research problems and research objectives pertaining to the aforementioned issues. Following that, a thorough explanation of the scope of this study with respect to soft sensors of advanced control systems in refining processes is provided. A last section of the thesis outlines how it will be organized.

1.2 Study Background

In industrial environments, it is common to employ a number of different sensors. Online quality monitoring of compositional factors, on the other hand, can be accomplished with sensors that are less reliable and accurate in nature. The sensors' primary goal is to provide data for process monitoring and control, as well as plant process control. (Kadlec, Gabrys, & Strandt, 2009). The study's objectives are to improve data quality while also reducing redundancy. Methods such as data mining and machine learning were used to assist in the investigation of the data.It is not enough to rely solely on technology to maintain a competitive advantage. Knowledge of human brains, sustainable ingredients, procedures, and company experience is essential. As a result, the gathering and analysis of data is critical for these approaches. (Wang, 2007). Real time process sensors, such as temperature, flow rate, and pressure sensors, are used in industrial processes, and interpretation sensors are also used in some of these processes. It is common for a data historian to be associated with a process machine, as it collects and maintains historical data during the course of the process. There are some quality elements, in particular, that do not have online sensors due to the fact that the sensors are either too expensive or too unreliable. It is assumed that product quality attributes are operational and can be calculated in real time, together with other processing variables. (Devogelaere et al., 2002).

The introduction of new process control methodologies has a substantial impact on every aspect of process management and control. The advancement of control theory, the introduction of new adjusting techniques, current actuators, smart sensors, and the introduction of sophisticated processes all present new challenges. In order to demonstrate better productivity and conformity with rules, manufacturing plants are subjected to stringent controls regarding product quality and pollutant emission levels.

Small industries, such as water treatment plants, power stations, and irrigation systems, are controlled by SCADA/PLC systems. The SCADA system is often used in the management of industrial processes. In this research, the distributed control system (DCS) is discussed and demonstrated to aid the development of soft sensors for oil and gas refineries. (Morsi & El-Din, 2014).

Traditional and cutting-edge control technologies are employed in the refinery process control of petroleum. The conventional technique makes use of small microcontrollers with low input/output needs. Across extensive environmental area, PLCs, SCADAs, and DCSs are used to control activities involving a diverse set of process variables. (K & Shivappa, 2013).

DCS is used by the vast majority of refinery control systems. Processes are then subdivided and controlled locally by boxes or process managers (PMs) which further helps in the isolation of specific sections during the events of refinery malfunction. Boxes can handle information transfer between systems, sensors, and actuators, including all necessary conversions in A/D and D/A converters. A local control network (LCN) connects workstations with cutting-edge control, supervision, and optimization technologies. An interconnection exists between the distribution system and the plant information system (PIS) server, such that the latter can store historical data for refinery's performance monitoring and optimization. (L. Fortuna et al., 2002). The crude oil distillation tower is responsible of the petroleum separation process. Later, in pre-production units, the products are refined and blended with other industrial products such as oil and diesel. Distillation towers separate the elements, and optimizing the control of the distillation towers lifts the quality of the end product. (Macias-Hernandez et al., 2007).

Fractionation, conversion, processing, formulation and mixing, and other refining procedures, such as light-end recovery; sweat removal; solid waste; waste process water and wastewater treatment; cooling; storage, handling, and transportation of materials; water electrolysis; and acid treatment are all examples of refining procedures. Among the many modern approaches to treat petroleum, there are separation (distillation, solvent refinement), convection (carbon removal, hydrogen addition), reforming (catalytic processing, reforming of steam/hydrocarbon), rearrangement (isomerization), mixture (catalytic polymerization), alkylating (treatment, curing and mixing of gasoline and diesel; lubricants and waxes, asphalt), and oil preservation (preservation by oxidation) (wastewater treatment, disposal of solids, sulphur recovery). Distillation columns are physically equipped with several features that are either utilized to convey heat or to facilitate mass transmission. Vertical separation of the components is achieved through the use of trays or plates, and a boiler heats up the column from the bottom up, resulting in condensing the vapour from the splitter column.

A real-world oil project to build a light naphtha processing facility in order to boost the utility value of control systems served as the basis for this investigation's distillation splitter column computation. The splitter column has a nominal capacity of 24 000 kg/hr of light naphtha, which is approximately six percent of the entire product. When the operation of the system is 24 hours a day and 350 working days a year is taken into consideration, any proposed control sensing system must be well designed and durable enough to withstand such demands. The determination of the processing of light naphtha product is mainly dependent on the purity of the distillate, X_D in the range equal to 96% to 98% and the impurity, X_B similar to or less than 2%.

These two ranges for purity and impurity serve as a reference point for developing any sensor system. (Brno, 2010). Hence, this research will focus on the industrial processes associated with petroleum productions and developed the modelling of soft sensors. The investigation will begin by listing the details of virtual sensor construction, especially in an oil refinery as this industry contributes more effectively to the local revenue of Iraq and other oil countries.

An energy balanced (feeding rate, reflux rate, product withdrawal rate) structure control system for a splitter distillation column has been developed as the basis for the construction of a computational model and simulation. A petroleum project's feasibility assessment and design can benefit from simulation and analytic findings at the outset. (Minh & Abdul Rani, 2009) The integration of design, production, marketing and management in the chemical sector of all products and process development chains must be adapted to modelling and simulation (Balazs Balasko et al., 2014). Sensors are used in every industrial site to keep track of everything that is going on. Because the sensor output gives only objective information about the production processes, their reliability is critical. (Wang, 2010).

Refineries, chemical plants, cement plants, power plants, pulp and paper mills and other similar facilities all rely on soft sensors. SFD is a well-known feature of soft sensors that can be utilized in conjunction with physical sensors and industrial processes for the identification of process faults. Other characteristics of soft sensors include their ability to be implemented quickly using plant hardware and real-time data estimation. Mechanical, statistical, and artificial modelling techniques can be used to model soft sensors in industrial processes. The data-driven approach is used in order to establish correlations between primary and secondary variables in data sets (e.g., quantitative, artificial intelligence). Because of the ambiguity and complexity of industrial processes, mechanistic modelling has been therefore ignored. (Jianxu & Huihe, 2002). Soft sensors are not only affordable, but they may also be used in conjunction with real sensors, replace it when a flaw occurs or during the events when the sensor is sent for repair/maintenance.

The quality of the data collected determines the performance of soft sensors. Historical plant databases stores information that is frequently required for purposes other than modelling. Software-defined sensors are intended to give additional information for process control online. Soft sensors are intended to engage in active monitoring and supplement physical sensors for process management by providing additional information in real time. (Kadlec et al., 2009). Embedded computer sensors can be used to measure target's variables values of product quality. These smart sensors are designed using intelligent computer techniques, where objective criteria such as product quality are linked to online process variables (Bakhtadze et al., 2008).

Computer and networking technologies, as well as new data collection techniques employed in industrial organizations, have resulted in massive, automated databases containing information with regards to manufacturing processes, products, and machinery. Possible patterns in the parameters that control an industrial process or a product's quality can be established from such data. (Sadoyan et al., 2006).

The availability of consistent data that can be depended upon is quite limited. Miscalibration problems, measurement errors, and computer interface issues are all factors that contribute to inaccurate process variable samples taken by analysers and erroneous laboratory measurements. It is possible that outliers and measurement errors will lead the data to become distorted. Therefore, a number of considerations must be taken into account while developing and deploying real-time sensors. (Liu, Srinivasan, & Selvaguru, 2008).

All the operational procedures and process are dependent on the set values of these variables with difficulty being in how to determine the individual impact and the influence as a whole on each process. Such processes include backup device measurement, real-time plant control prediction, sensor validation and failure detection methods. Due to the complexity of industrial processes, it is difficult to compile historical data.

The sheer volume of data collected may make it difficult to discern the various elements that could influence plant operations. Dimensionality, noise, precision, redundant and erroneous values, selection flaws as well as recording techniques all contribute to the difficulty of data mining. Extensive knowledge of data-driven soft sensor modelling is therefore required, such as input variable selection, system order, operational range, time delay, nonlinearity, and sampling rates. These characteristics are critical to model designers. Soft sensors are frequently employed in inferential control systems in the industrial setting. In addition to measuring device backup, soft sensors are also used for supporting measuring device, real-time monitoring and control estimations, sensor validation, error detection and diagnostics as well as IF-then analyses. (Luigi Fortuna et al., 2007).

1.3 Research Hypothesis

Data mining, machine learning strategies, such as neuro-fuzzy based rough set theory (RST) as well as computing techniques and quantile discretisation will be adopted to design and develop soft sensor model aims to improve the performance of DCS in oil refineries throughout supporting physical sensors and calculate the value of quality metrics (API and RVP) of light naphtha which is continuously produced in oil refinery. Choosing the optimum methods for filtering and pre-processing the collected data from the process variables database in the control system of oil refinery will contribute to the reduction of the prediction model's complexity. Real-time prediction of the quality metrics of a product shall enhance the overall performance of control system and improve its stability.

1.4 Problem Statement

Oil refineries are sophisticated processes that can continue to operate 24 hours a day, and often throughout the year. The degradation of sensors and instruments may have an impact on the product quality. In addition to determining real time quality indicators, predictive models can assist physical sensors in maintaining their predictive ability. Computers, embedded systems, and machines must recognize and grasp the relationships between process variables in order to imitate human behaviour within the processing system.

In addition to changing data owing to variations in feedstock and alterations to process conditions and working environment, one of the disadvantages of soft sensors is the inability to adapt to fit different scenarios. It is noteworthy that proportional-integral-derivative (PID) controllers and cascade PID controllers are widely utilized in most manufacturing processes, including predictive control, adaptive control, and expert control, when compared to modern control systems in most manufacturing processes.

However, in the crude oil distillation industry, PID control techniques or enhanced PID controllers via nonlinear function block have shown to have limitations.

These performance constraints are caused by the differences in the chemical characteristics of the crude oil that is supplied into the unit. The inability of controllers such as the cascade PID controller to cope with the changes in the variables in the CDU is of a significant problem.

Soft sensor modelling, on the other hand, faces a number of hurdles, including the sophistication, nonlinearity, and data quality issues associated with industrial processes. It is critical to transmit data between industries in order to establish a robust and efficient soft sensor, as well as to break through privacy barriers to improve overall industrial support.

Soft sensing with shifting feedstock continues to be a concern for a crude oil distillation tower, especially as the control process factors alter over time. Despite the fact that process variables are straightforward to calculate, measured quality variables are complex and vary depending on the type of crude oil being processed.

The types of crude oil available from different vendors differ. Even crude oil from the same supplier can have varying amounts of hydrocarbons in it, according to the manufacturer. In addition, multiple refineries operate using a variety of crude oil sources blended in a variety of ratios to produce finished products. Oil refineries suffer from a lack of uniformity in their products, which, according to laboratory testing, has an impact on the finished goods.

A dearth of literature that investigates soft sensors in oil refineries and associated control systems, as well as a lack of data sharing within businesses, are additional issues. According to the findings of the literature research and conference papers, there are numerous limitations to the models employed in various sectors. The limitations can be summarized to: the improved model with limited data, the difficulty in maintaining some models of soft sensors or manual maintenance, a lack of data in some industries that results in the use of only models that require a small number of data, the complexity and nonlinearity of operating the models, as well as the privacy of data sharing with industry.

1.5 Research Objectives

Through the use of data mining and supervised machine learning approaches, this research attempt to envision the oil refinery control system, which will help to solve the challenges outlined above in greater detail. This is accomplished through mining of the obtained and reduce the affected variables in the production of light naphtha by utilizing RST after the discretization of continuous sensor readings. Neural networks can be used to identify new patterns of data in order to pre-process input in fuzzy systems and therefore improve fuzzy-system response to fresh input data.

Both neural networks and fuzzy inferential systems are used as sensors to forecast the Reid vapour pressure (RVP) and American Petroleum Institute (API) metric for light naphtha. The soft sensor design improves the control system for reflux cycling at the top of an oil refinery's splitter tower. The targeted activities to be achieved are as follows:

To enhance the quality of data and system controls by merging the decision tables obtained from soft computing model of different sources (of different parties) with the case studies' database.

To develop a soft sensor model that uses hybrid neuro-fuzzy-RST soft computing methodology and quantile discretisation method in an oil refinery control system for predicting and forecasting the quality of cuts (fractions) of light naphtha and adopting the model for processing data of refining light naphtha.

To validate the neuro-fuzzy-RST-based soft sensor model through comparison with regression and machine learning models as well as simulating the adaptive neuro-fuzzy inference system (ANFIS) soft sensor to support the physical sensor in a cascade PID controller of the flow and temperature of the reflux ratio to the head of the splitter in the oil refinery.

1.6 Research Scope

It is stated in the problem statement section that, in order for the soft sensor model's output to stay correct, it must be capable of adapting to changes in the regular data process. Several studies focused on a critical component of the petroleum industry: the crude oil refining process. Higher-octane hydrocarbons such as gasoline, light diesel, and lighter naphtha can be converted into lighter hydrocarbons such as gasoline, light diesel, and heavier naphtha by the use of a crude oil distillation machine. This unit is responsible for a significant portion of the refinery's overall economic efficiency. Using these sensors, API gravity and light naphtha's relative vacuum pressure could be monitored in real time.

Iraq refinery (Middle Refineries Co./Al-Doura refinery, Baghdad) units of two separate crude oil sources were used to acquire primary data in order to improve prediction accuracy, refinery specialists were approached to customize the factors. In order to anticipate crude oil quality at distillation towers and to improve the quality of soft sensor data based on the RST and discretion approaches, soft computing technique has been created according to a case study using the proposed model, the control system may be improved based on information gathered to help different businesses work together better. Demonstrates how well global smart data mining strategies work when they use the same soft sensor modelling for multiple refineries.

This means that the primary issues for sectors using soft sensors are the ability to use soft sensor modelling in harsh environments (such as oil refineries), the lack of high-quality raw data and data changes following sensor shifts, and feedstock degradability. There will be more cooperation between industries in the future as manufacturing moves toward Industry 4.0. This will allow their products to be more automated. Using algorithms for good and meaningful data approximation, machine learning then determines or predicts.

Simulation the oil refinery requires working with an operating system model on a computer, which evolves over time in order to better understand and refine the system. MATLAB and Simulink platforms could mimic the behaviour of their suggested soft sensor to adapt the light naphtha product using the C04 splitter top and imitate the model's effect to assist temperature and flow physical sensors for maintaining control system stability.

During the research phase, the study ran into some obstacles. Due to Iraq's reliance on petrochemical sectors, which affects the country's revenue, the amount of collected data was restricted, making data analysis and mining more difficult. Because of this, the quality of the collected data for API and RVP dependent variables may suffer. The laboratory test method is limited to four times per day rather than six times per day as indicated in the operation manual and standard of the cut product test. The proposed model's prediction of light naphtha response values is more reliable if the sensor readings are manually established in the manufacturing processes rather than employing advanced interface hardware systems to log sensor readings directly into soft sensor modelling in MATLAB Software.

1.7 Thesis Outline

In Chapter 1, the problem of predicting one of the first phases of crude oil refining with various algorithms and theories is illustrated. A new soft sensor based on the ANFIS, which uses various intelligent methods to increase performance modelling of the soft sensors, is recommended. The sensor is developed in MATLAB, a simulation system to evaluate the functioning of the algorithm. In addition, different environments have been designed to assess the strengths and benefits of the proposed model against existing approaches. Similar assessment measures have been used to help the evaluation.

Chapter 2 presents a detailed study on current works of soft sensors and machine learning techniques regarding the problem in predicting the quality of products in complex manufacturing system environments.

Chapter 3 describes the research methodology in detail. Different empirical and intelligent methods used in the study to reach the research objectives are specified.

Chapter 4 presents the results of the study. Detailed discussion about the proposed algorithm, performance analyses, and comparison outcomes are provided with supplemented charts, graphs, and tables.

Chapter 5 concludes the results of the study with additional graphs and discussion. Furthermore, the research contributions are outlined and recommendations for further studies in this area are given.

REFERENCES

Ahmed, A. A. M., & Shah, S. M. A. (2017). Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. *Journal of King Saud University - Engineering Sciences*, *29*(3), 237–243. https://doi.org/10.1016/j.jksues.2015.02.001

Al-Saraj, M. A. A., Rahman, A. M., & Al-Hassani, M. H. (2013). Enhancement of Iraqi Light Naphtha Octane Number Using Pt Supported HMOR Zeolite Catalyst. *Al-Khwarizmi Engineering Journal*, *9*(4), 1–11. Retrieved from http://kecbujournal.com/LionImages/News/1-11.pdf

Arnold, K., & Stewart, M. (2008). Crude Stabilization. In *Surface Production Operations* (pp. 457–481). Elsevier. https://doi.org/10.1016/b978-075067853-7.50011-9

Arya Azar, N., Ghordoyee Milan, S., & Kayhomayoon, Z. (2021). The prediction of longitudinal dispersion coefficient in natural streams using LS-SVM and ANFIS optimized by Harris hawk optimization algorithm. *Journal of Contaminant Hydrology*, *240*, 103781. https://doi.org/10.1016/j.jconhyd.2021.103781

Bakhtadze, N., Maximov, E. M., & Valiakhmetov, R. T. (2008). *Fuzzy Soft Sensors for Chemical and Oil Refining Processes*. *IFAC Proceedings Volumes* (Vol. 41). IFAC. https://doi.org/10.3182/20080706-5-kr-1001.00007

Balasko, B., Nemeth, S., Nagy, G., & Abonyi, J. (2008). Integrated process and control system model for product quality control-application. *Chemical Product and Process Modeling*, *3*(1), Article ID: 50. https://doi.org/10.2202/1934-2659.1213

Balasko, Balazs, Nemeth, S., & Abonyi, J. (2014). Integrated process and control system model for product quality control - A soft-sensor based application. *2009 European Control Conference, ECC 2009*, 3498–3502. https://doi.org/10.23919/ecc.2009.7074941

Bariani, M., Cucchiara, R., Mello, P., & Piccardi, M. (1997). Data mining for automated visual inspection. *Proc. of PADD97, London, UK*, 51–64. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.56.7366&rep=rep1&type=pdf

Bhaskar, M. K., Pal, N. S., & Yadav, V. K. (2018). A comparative performance analysis of automatic generation control of multi-area power system using PID, fuzzy and ANFIS controllers. *2018 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES 2018*, 132–137. https://doi.org/10.1109/ICPEICES.2018.8897477

Bidar, B., Khalilipour, M. M., Shahraki, F., & Sadeghi, J. (2018). A data-driven soft-sensor for monitoring ASTM-D86 of CDU side products using local instrumental variable (LIV) technique. *Journal of the Taiwan Institute of Chemical Engineers*, *84*, 49–59. https://doi.org/10.1016/j.jtice.2018.01.009

Bo, C., Li, J., & Sun, C. (2003). The application of neural network soft sensor technology to an advanced control system of distillation operation. *Neural Networks, 2003.*, 1054–1058.

Caldwell, J. A. (2015). MULTIMEDIA MINING RESEARCH – AN OVERVIEW. *International Journal of Computer Graphics & Animation*, *5*(1), 1–33. https://doi.org/10.1016/B978-0-323-35572-8.00030-0

Cao, J., Zhou, B. B., & Li, Y. (2010). Soft sensor for on-line estimation of gelatin concentration. In *2010 International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2010* (Vol. 3, pp. 406–409). https://doi.org/10.1109/ICMTMA.2010.136

Carlos A. Duchanoya, Marco A. Moreno-Armendáriza, Leopoldo Urbinaa, Carlos A. Cruz-Villarb, Hiram Calvoa, J. de J. R. a. (2005). A Fuzzy Adaptive Differential Evolution Algorithm. *Springer-Verlag 2004*, 448–462. https://doi.org/10.1007/s00500-004-0363-x

Chang, Y., Lv, Z., Wang, F., Mao, Z., & Wang, X. (2006). Soft sensing modeling based on stacked least square-support vector machine and its application. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, *1*, 4846–4850. https://doi.org/10.1109/WCICA.2006.1713305

Chen, B., Zhong, C., Yang, S., & Yan, M. (2007). The application of soft-sensor technology in measuring water boiling point. *Proceedings of the 26th Chinese Control Conference, CCC 2007*, 372–376. https://doi.org/10.1109/CHICC.2006.4347228

Chen, K., & Wang, J. (2008). The data mining technology based on CIMS and its application on automotive remanufacturing. *Proceedings - 1st International Workshop on Knowledge Discovery and Data Mining, WKDD*, 48–52. https://doi.org/10.1109/WKDD.2008.125

Chen, R. S., Wu, R. C., & Chang, C. C. (2005). Using data mining technology to design an intelligent CIM system for IC manufacturing. In *Proceedings - Sixth Int. Conf. on Softw. Eng., Artificial Intelligence, Netw. and Parallel/Distributed Computing and First ACIS Int. Workshop on Self-Assembling Wireless Netw., SNPD/SAWN 2005* (Vol. 2005, pp. 70–75). https://doi.org/10.1109/SNPD-SAWN.2005.78

Chen, R s., Tsai, Y. s., & Chang, C. c. (2006). Design and Implementation of an Intelligent Manufacturing Execution System for Semiconductor Manufacturing Industry. In *Industrial Electronics, 2006 IEEE International Symposium on* (Vol. 4, pp. 2948–2953). https://doi.org/10.1109/ISIE.2006.296085

Chen, Ruey-shun, Wu, R., & Chang, C. C. (2010). Using Data Mining Technology to Design an Intelligent control System in Manufacture Industry. *Proceedings of European Conference on Computer Science*, 272–276.

Chen, W.-C., Tseng, S.-S., & Wang, C.-Y. (2005). A Novel Manufacturing Defect Detection Method using Association Rule Mining Techniques. *Expert Systems with Applications*, *29*(4), 807–815. https://doi.org/10.1016/j.eswa.2005.06.004

Chien, C.-F., Wang, W.-C., & Cheng, J.-C. (2007). Data mining for yield enhancement in semiconductor manufacturing and an empirical study. *Expert Systems with Applications*, *33*(1), 192–198. https://doi.org/10.1016/j.eswa.2006.04.014

Choudhary, A. K., Harding, J. A., & Tiwari, M. K. (2009). Data mining in manufacturing: A review based on the kind of knowledge. *Journal of Intelligent Manufacturing*, *20*(5), 501–521. https://doi.org/10.1007/s10845-008-0145-x

Dan Braha and Armin Shmilovici. (2015). Data Mining for Improving a Cleaning Process in the Semiconductor Industry. *IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING*, *15*(1), 91–101.

Devogelaere, D., Rijckaert, M., Leon, O. G., & Lemus, G. C. (2002). Application of feedforward neural networks for soft sensors in the sugar industry. *Proceedings - Brazilian Symposium on Neural Networks, SBRN*, *2002-Janua*, 2–6. https://doi.org/10.1109/SBRN.2002.1181426

Dote, Y., Ovaska, S. J., & Member, S. (1996). Industrial Applications of Soft Computing : *PROCEEDINGS OF THE IEEE, 89*(9), 1–26.

Duraid Fadhil Ahmed, A. M. S. (2016). Neural Network Control for a Batch Distillation Column Duraid. *Tikrit Journal of Engineering Sciences*, *23*(1), 40–52.

Eastwood, M., & Kadlec, P. (2011). Interpretable, online soft-sensors for process control. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 581–587. https://doi.org/10.1109/ICDMW.2011.105

Elovici, Y., & Braha, D. (2003). A Decision-Theoretic Approach to Data Mining. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, *33*(1), 42–51. https://doi.org/10.1109/TSMCA.2003.812596

Fayyad, U. (1997). Data mining and knowledge discovery in databases: implications for scientific databases. In *Proceedings. Ninth International Conference on Scientific and Statistical Database Management (Cat. No.97TB100150)* (pp. 2–11). https://doi.org/10.1109/SSDM.1997.621141

Fernandez de Canete, J., Del Saz-Orozco, P., Baratti, R., Mulas, M., Ruano, A., & Garcia-Cerezo, A. (2016). Soft-sensing estimation of plant effluent concentrations in a biological wastewater treatment plant using an optimal neural network. *Expert Systems with Applications*, *63*, 8–19. https://doi.org/10.1016/j.eswa.2016.06.028

Flynn, D., Ritchie, J., & Cregan, M. (2015). DATA MINING TECHNIQUES APPLIED TO POWER PLANT PERFORMANCE MONITORING D. Flynn, J. Ritchie and M. Cregan. *IFAC Proceedings Volumes*, *38*(1), 369–374.

Fortuna, L., Graziani, S., & Xibilia, M. G. (2005). Soft sensors for product quality monitoring in debutanizer distillation columns. *Control Engineering Practice*, *13*(4), 499–508. https://doi.org/10.1016/j.conengprac.2004.04.013

Fortuna, L., Graziani, S., & Xibilia, M. G. (2009). Comparison of soft-sensor design methods for industrial plants using small data sets. *IEEE Transactions on Instrumentation and Measurement*, *58*(8), 2444–2451. https://doi.org/10.1109/TIM.2009.2016386

Fortuna, L., Rizzo, A., Sinatra, M., & Xibilia, M. G. (2002). Soft analysers for a sulfur recovery unit. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *15*(1), 289–294. https://doi.org/10.1016/S0967-0661(03)00079-0

Fortuna, L., Rizzo, A., Sinatra, M., & Xibilia, M. G. (2003). Soft analyzers for a sulfur recovery unit. *Control Engineering Practice*, *11*(12), 1491–1500. https://doi.org/10.1016/S0967-0661(03)00079-0

Fortuna, Luigi, Giannone, P., Graziani, S., & Xibilia, M. G. (2007). Virtual instruments based on stacked neural networks to improve product quality monitoring in a refinery. *IEEE Transactions on Instrumentation and Measurement*, *56*(1), 95–101. https://doi.org/10.1109/TIM.2006.887331

Fortuna, Luigi, Graziani, S., Rizzo, A., & Xibilia, M. G. (2007). *Soft Sensors for Monitoring and Control of Industrial Processes*. *Advances in Industrial Control*. https://doi.org/10.1007/978-1-84628-480-9

Fu, Y. S., Liu, F. Z., Zhang, W. Z., Zhang, Q., & Zhang, G. X. (2008). The fault diagnosis of power transformer based on rough set theory. *2008 China International Conference on Electricity Distribution, CICED 2008*, 1–5. https://doi.org/10.1109/CICED.2008.5211667

He, T. (2010). Data mining for the optimization in manufacturing process. *Proceedings - 2010 2nd International Workshop on Intelligent Systems and Applications, ISA 2010*, 5–8. https://doi.org/10.1109/IWISA.2010.5473459

Holes, L. (2010). Operation Manual. In *Creating the Built Environment* (pp. 166–191). https://doi.org/10.4324/9780203476154_chapter_9

Hou, Z., Lian, Z., Yao, Y., & Yuan, X. (2006). Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Conversion and Management*, *47*(15–16), 2479–2490. https://doi.org/10.1016/j.enconman.2005.11.010

Iliyas, S. A., Elshafei, M., Habib, M. A., & Adeniran, A. A. (2013). RBF neural network inferential sensor for process emission monitoring. *Control Engineering Practice*, *21*(7), 962–970. https://doi.org/10.1016/j.conengprac.2013.01.007

Jackson, B., Scargle, J. D., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., & Gwin, E. (2003). An Algorithm for Optimal Partitioning of Data on an Interval. *IEEE Signal Processing Letters*, *12*(2), 105–108.

Jalee, E. A., & Aparna, K. (2016). Neuro-fuzzy Soft Sensor Estimator for Benzene Toluene Distillation Column. *Procedia Technology*, *25*(Raerest), 92–99. https://doi.org/10.1016/j.protcy.2016.08.085

Jang, J. S. R. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man and Cybernetics*, *23*(3), 665–685. https://doi.org/10.1109/21.256541

Jianxu, L. (2002). Selecting secondary measurements for soft sensor modeling using rough sets theory. In *Proceedings of the 8 World Congress on Intelligent Control and Automation* (pp. 415–419).

Jianxu, L. J. L., & Huihe, S. H. S. (2002). Soft sensing modeling using neurofuzzy system based on rough set theory. In *Proceedings of the 2002 American Control Conference (IEEE Cat. No.CH37301)* (Vol. 1, pp. 0–5). https://doi.org/10.1109/ACC.2002.1024863

Juan, C., Yang, Y., & Yanlei, Q. (2011). Application of least squares support vector machine in soft sensor of traditional Chinese Medicine extraction. *Proceedings - 2011 7th International Conference on Natural Computation, ICNC 2011*, *2*, 747–751. https://doi.org/10.1109/ICNC.2011.6022223

K.Anitha, D. P. V., & Asst. (2013). Feature Selection By Rough – Quick Reduct Algorithm. *International Journal of Innovative Research in Science, Engineering and Technology*, *2*(8), 3989–3993.

K, R. K. S., & Shivappa, D. N. (2013). Adopting Distributed Control System To Control Desalting Process in Petroleum Refinery. *International Journal of Emerging Technology and Advanced Engineering*, *3*(10), 1–6.

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven Soft Sensors in the process industry. *Computers and Chemical Engineering*, *33*(4), 795–814. https://doi.org/10.1016/j.compchemeng.2008.12.012

Kalaivani.R, M.V.Suresh, N. S. (2017). A Study of Rough Sets Theory a nd It ' s Applicati on Over Various Fields. *Journal of Applied Science and Engineering Methodologies*, *3*(2), 447–455.

Kämpjärvi, P., Sourander, M., Komulainen, T., Vatanski, N., Nikus, M., & Jämsä-Jounela, S. L. (2008). Fault detection and isolation of an on-line analyzer for an ethylene cracking process. *Control Engineering Practice*, *16*(1), 1–13. https://doi.org/10.1016/j.conengprac.2007.03.007

Karim, A. R., Khanaqa, P., & Shukur, D. A. (2017). Kurdistan crude oils as feedstock for production of aromatics. *Arabian Journal of Chemistry*, *10*, S2601–S2607. https://doi.org/10.1016/j.arabjc.2013.09.038

Kinematic, A. N., & Planning, C. T. (2011). Data Mining of Printed-Circuit Board Defects. *Ieee Transactions on Robotics and Automation*, *17*(2), 1–8.

Köksal, G., Batmaz, I., & Testik, M. C. (2011). A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications*, *38*(10), 13448–13467. https://doi.org/10.1016/j.eswa.2011.04.063

Kotsiantis, S., & Kanellopoulos, D. (2006). Discretization Techniques : A recent survey. *GESTS International Transactions on Computer Science and Engineering*, *32*(1), 47–58.

Krupka, J., & Jirava, P. (2014). Rough-Fuzzy classifier modeling using data repository sets. *Procedia Computer Science*, *35*(C), 701–709. https://doi.org/10.1016/j.procs.2014.08.152

Kusiak, A. (2001). Rough set theory: a data mining tool for semiconductor manufacturing. *IEEE Transactions on Electronics Packaging Manufacturing*, *24*(1), 44–50. https://doi.org/10.1109/6104.924792

Lawal, S. A., & Zhang, J. (2007). Fault Monitoring and Fault Tolerant Control in Distillation Columns. *Automation and Computing (ICAC)*, 1–6.

Li, X., Yan, Q., & Yu, D. (2010). Application of neural network based on improved ant colony optimization in soft sensor modeling of polymer electrolyte membrane moisture. *ICCASM 2010 - 2010 International Conference on Computer Application and System Modeling, Proceedings*, *8*, 4–8. https://doi.org/10.1109/ICCASM.2010.5619086

Lin, B., Recke, B., Knudsen, J. K. H., & Jørgensen, S. B. (2007). A systematic approach for soft sensor development. *Computers and Chemical Engineering*, *31*(5–6), 419–425. https://doi.org/10.1016/j.compchemeng.2006.05.030

Liu, J., Srinivasan, R., & Selvaguru, P. N. (2008). Practical Challenges In Developing Data-Driven Soft Sensors For Quality Prediction. In *18th European Symposium on Computer Aided Process Engineering – ESCAPE18* (pp. 961–966).

Lu, C., Li, W., & Liu, W. (2009). Application of soft sensor based on LS-SVM on estimation of alumina powder flow. *2009 International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2009*, *1*, 281–284. https://doi.org/10.1109/ICMTMA.2009.236

Lu, Y.-Z., He, M., & Xu, C.-W. (1996). Fuzzy modeling and expert optimization control for industrial processes. *Control Systems Technology, IEEE Transactions On*, *5*(1), 2–12.

Luo, J., & Shao, H. (2002). Soft sensing modeling using neurofuzzy system based on rough set theory. *Proceedings of the American Control Conference*, *1*, 543–548. https://doi.org/10.1109/acc.2002.1024863

Luo, J. X., & Shao, H. H. (2006). Developing soft sensors using hybrid soft computing methodology: A neurofuzzy system based on rough set theory and genetic algorithms. *Soft Computing*, *10*(1), 54–60. https://doi.org/10.1007/s00500-005-0465-0

Macias-Hernandez, J., Angelov, P., & X., Z. (2007). Soft Sensor for Predciting Crude Oil Distillation Side Streams using Takagi Sugeno Evolving Fuzzy Models. *International Conference on Systems, Man, and Cybernetics*, 3305–3310.

Macias, J. J., Angelov, P., & Zhou, X. (2006). A method for predicting quality of the crude oil distillation. In *Proceedings of the 2006 International Symposium on Evolving Fuzzy Systems, EFS'06* (pp. 214–220). https://doi.org/10.1109/ISEFS.2006.251167

Mardani, A., Nilashi, M., Antucheviciene, J., Tavana, M., Bausys, R., & Ibrahim, O. (2017). Recent fuzzy generalisations of rough sets theory: A systematic review and methodological critique of the literature. *Complexity*, *2017*, 1–36. https://doi.org/10.1155/2017/1608147

Mendes, G., Aleme, H. G., & Barbeira, P. J. S. (2017). Reid vapor pressure prediction of automotive gasoline using distillation curves and multivariate calibration. *Fuel*, *187*, 167–172. https://doi.org/10.1016/j.fuel.2016.09.046

Minh, V. T., & Abdul Rani, A. M. (2009). Modeling and control of distillation column in a petroleum process. *Mathematical Problems in Engineering*, *2009*. https://doi.org/10.1155/2009/404702

Morsi, I., & El-Din, L. M. (2014). SCADA system for oil refinery control. *Measurement: Journal of the International Measurement Confederation*, *47*(1), 5–13. https://doi.org/10.1016/j.measurement.2013.08.032

Nair, A., & Daryapurkar, R. (2016). Analysis of fluidized catalytic cracking process using fuzzy logic system. *2016 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2016*, 1891–1895. https://doi.org/10.1109/ICACCI.2016.7732326

Nowicki, R. K. (2010). On classification with missing data using rough-neuro-fuzzy systems. *International Journal of Applied Mathematics and Computer Science*, *20*(1), 55–67. https://doi.org/10.2478/v10006-010-0004-8

Pani, A. K., Vadlamudi, V., Bhargavi, R. J., & Mohanta, H. K. (2011). Neural network soft sensor application in cement industry: Prediction of clinker quality parameters. *Proceedings of 2011 International Conference on Process Automation, Control and Computing, PACC 2011*, 0–5. https://doi.org/10.1109/PACC.2011.5979038

Park, S., & Han, C. (2000). A nonlinear soft sensor based on multivariate smoothing procedure for quality estimation in distillation columns. *Computers and Chemical Engineering*. https://doi.org/10.1016/S0098-1354(00)00343-4

Pawlak, Z. (1997). Rough set approach to knowledge-based decision support. *European Journal of Operational Research*, *99*(1), 48–57. https://doi.org/10.1016/S0377-2217(96)00382-7

Pawlak, Z. (1998). Rough set theory and its applications. *Journal of Telecommunications and Information Technology*, *29*(7), 7–10. https://doi.org/10.1080/019697298125470

Pawlak, Z. (2002). ROUGH SETS AND DATA ANALYSIS Zdzislaw Pawlak Institute. *Information Sciences*, *147*, 1–6.

Peña, M., Biscarri, F., Guerrero, J. I., Monedero, I., & León, C. (2016). Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach. *Expert Systems with Applications*, *56*, 242–255. https://doi.org/10.1016/j.eswa.2016.03.002

Pi, D., Zhang, F., Wang, N., & Qin, X. (2007). A data mining framework oriented CIM for cooperative manufacturing. In *2006 Chinese Control Conference Proceedings, CCC 2006* (pp. 1357–1362). https://doi.org/10.1109/CHICC.2006.280674

Qin, S. . J. (1997). *Neural Networks for Intelligent Sensors and Control - Practical Issues and Some Solutions. Neural Systems for Control*. https://doi.org/10.1016/B978-012526430-3/50009-X

Red, C. I., & Zhang, W. (2009). THE CONDITION ESTIMATE OF POWER TRANSFORMER BASED ON ROUGH SET THEORY. In *20th International Conference on Electricity Distribution* (pp. 8–11).

Rissino, S., & Lambert-Torres, G. (2009). Rough Set Theory — Fundamental Concepts, Principals, Data Extraction, and Applications. *Data Mining and Knowledge Discovery in Real Life Applications*, (February). https://doi.org/10.5772/6440

Rodríguez, M., & Gayoso, J. A. (2006). Degrees of freedom analysis for process control. *Computer Aided Chemical Engineering*, *21*(C). https://doi.org/10.1016/S1570-7946(06)80258-0

Rogina, A., Si, I., Mohler, I., & Bolf, N. (2011). Soft sensor for continuous product quality estimation (in crude distillation unit). *Chemical Engineering Research and Design*, *9*(June 2010), 2070–2077. https://doi.org/10.1016/j.cherd.2011.01.003

Sadoyan, H., Zakarian, A., & Mohanty, P. (2006). Data mining algorithm for manufacturing process control. *International Journal of Advanced Manufacturing Technology*, *28*(3–4), 342–350. https://doi.org/10.1007/s00170-004-2367-1

Serpas, M. R. (2012). Soft Sensors for Process Monitoring of Complex Processes, (August).

Sets, F. R., Zhou, P., Lu, S., & Chai, T. (2014). Data-Driven Soft-Sensor Modeling for Product Quality Estimation Using Case-Based Reasoning. *IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING*, *11*(4), 992–1003.

Sharifi, S.-S., Delirhasannia, R., Nourani, V., Sadraddini, A., & Ghorbani, A. (2013). Using Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS) for Modeling and Sensitivity Analysis of Effective Rainfall. *Recent Advances in Continuum Mechanics, Hydrology and Ecology*, (June), 133–139.

Sharmin, R., Sundararaj, U., Shah, S., Vande Griend, L., & Sun, Y. J. (2006). Inferential sensors for estimation of polymer quality parameters: Industrial application of a PLS-based soft sensor for a LDPE plant. *Chemical Engineering Science*, *61*(19), 6372–6384. https://doi.org/10.1016/j.ces.2006.05.046

Srivastava, N. (2014). Matlab Simulation of Temperature Control of Heat Exchanger using Different Controllers. *Automation, Control and Intelligent Systems*, *2*(1), 1. https://doi.org/10.11648/j.acis.20140201.11

Stankovic, J. A. (2008). When sensor and actuator networks cover the world. *ETRI Journal*, *30*(5), 627–633. https://doi.org/10.4218/etrij.08.1308.0099

Sugawara, E., & Nikaido, H. (2014). Properties of AdeABC and AdeIJK efflux systems of Acinetobacter baumannii compared with those of the AcrAB-TolC system of Escherichia coli. *Antimicrobial Agents and Chemotherapy*, *58*(12), 7250–7257. https://doi.org/10.1128/AAC.03728-14

Sujatha K.Pappa N.Kalaivani A.Pappa, N., & Kalaivani, A. (2011). Soft Sensor for Flue Gas Monitoring in Power Station Boilers. In *2011 International Conference on Process Automation, Control and Computing* (pp. 1–6). https://doi.org/10.1109/PACC.2011.5979031

Sulaiman, N. S., & Mohd Yusof, K. (2015). Artificial neural network-based model for quality estimation of refined palm oil. *ICCAS 2015 - 2015 15th International Conference on Control, Automation and Systems, Proceedings*, (Iccas), 1324–1328. https://doi.org/10.1109/ICCAS.2015.7364843

The Mathworks, I. (2018). Using MATLAB with Big Data from Sensors and IoT Devices, 1–15.

Tiwari, S., Babbar, R., & Kaur, G. (2018). Performance Evaluation of Two ANFIS Models for Predicting Water Quality Index of River Satluj (India). *Advances in Civil Engineering*, *2018*. https://doi.org/10.1155/2018/8971079

Tung-HSU Hou, Wang-Lin Liu, and L. L. (2003). Intelligent remote monitoring and diagnosis of manufacturing processes using an integrated approach of neural networks and rough sets. *Journal of Intelligent Manufacturing*, (1999), 239–253.

Variny, M., Furda, P., Kováč, N., & Mierka, O. (2019). Analysis of C3 fraction splitting system performance by mathematical modeling in MATLAB environment. *Acta Chimica Slovaca*, *12*(1), 127–135. https://doi.org/10.2478/acs-2019-0018

Wang, D., Liu, J., & Srinivasan, R. (2010). Data-driven soft sensor approach for quality prediction in a refining process. *IEEE Transactions on Industrial Informatics*, *6*(1), 11–17. https://doi.org/10.1109/TII.2009.2025124

Wang, K.-S. (2013). Towards zero-defect manufacturing (ZDM)—a data mining approach. *Advances in Manufacturing*, *1*(1), 62–74. https://doi.org/10.1007/s40436-013-0010-9

Wang, Keqin, Roucoules, L., Tong, S., Matta, N., & Eynard, B. (2007). Review on application of data mining in product design and manufacturing. *Proceedings - Fourth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2007*, *4*(Fskd), 613–618. https://doi.org/10.1109/FSKD.2007.482

Wang, Kesheng. (2006). Data mining in manufacturing: The nature and implications. In *IFIP International Federation for Information Processing* (Vol. 207, pp. 1–10). https://doi.org/10.1007/0-387-34403-9_1

Wang, Kesheng. (2007). Applying data mining to manufacturing: The nature and implications. *Journal of Intelligent Manufacturing*, *18*(4), 487–495. https://doi.org/10.1007/s10845-007-0053-5

Wang, P. (2010). Design and application of soft sensor object in water treatment plant. *2010 International Conference on Networking, Sensing and Control, ICNSC 2010*, 372–375. https://doi.org/10.1109/ICNSC.2010.5461526

Wang, S., & Cui, J. (2005). Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method. *Applied Energy*, *82*(3), 197–213. https://doi.org/10.1016/j.apenergy.2004.11.002

WANG Tao sheng, Z. H. Y. (2009). Fuzzy least squares support vector machine soft measurement model based on adaptive mutative scale chaos immune algorithm WANG. *Journal of Central South University of Technology ·*, *60*(2), 444–449. https://doi.org/10.1007/s11771

Wang, W. L., Chang, P., Huang, C. T., & Wang, B. S. (2009). ARFID-enabled with Data Mining Model for Exhibition Industry. In *6th International Conference on Service Systems and Service Management* (pp. 664–668).

Webb, G. I., Pazzani, M. J., & Billsus, D. (2001). Machine learning for user modeling. *User Modeling and User-Adapted Interaction*, *11*(1–2), 19–29. https://doi.org/10.1023/A:1011117102175

Xu, Y., & Liu, C. (2013). A rough margin-based one class support vector machine. *Neural Computing and Applications*, *22*(6), 1077–1084. https://doi.org/10.1007/s00521-012-0869-6

Yan, X., Tu, N., & Fu, H. (2009). Hybrid method for prediction of coal and gas outburst based on data fusion and soft sensor. *ICEMI 2009 - Proceedings of 9th International Conference on Electronic Measurement and Instruments*, 2997–21001. https://doi.org/10.1109/ICEMI.2009.5274397

Yang, H., Park, M., Cho, M., Song, M., & Kim, S. (2014). A system architecture for manufacturing process analysis based on big data and process mining techniques. *Big Data (Big Data), 2014 IEEE International Conference On*, 1024–1029. https://doi.org/10.1109/BigData.2014.7004336

Yang, S. H., Chen, B. H., & Wang, X. Z. (2000). Neural network based fault diagnosis using unmeasurable inputs. *Engineering Applications of Artificial Intelligence*, *13*(3), 345–356. https://doi.org/10.1016/S0952-1976(00)00005-1

Ying-shuan, F. Y. F., Fa-zhan, L. F. L., Wei-zheng, Z. W. Z., Qing, Z. Q. Z., & Gui-xin, Z. G. Z. (2008). Transformer Fault Diagnosis Utilizing Rough Set and Support Vector Machine. *2008 China International Conference on Electricity Distribution*, 2(August), 19–22. https://doi.org/10.1109/CICED.2008.5211667

Zhang, Yagang. (2009). Applications of Data Mining Theory in Electrical Engineering. *Engineering*, *01*(03), 211–215. https://doi.org/10.4236/eng.2009.13025

Zhang, Ying. (2010). An improved T-S fuzzy neural network and its application in soft sensing for FCCU. In *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)* (pp. 3712–3716). https://doi.org/10.1109/WCICA.2010.5553966

Zhao, L., & Chai, T. (2004). Adaptive moving window MPCA for online batch monitoring. *2004 5th Asian Control Conference*, *2*(August 2004), 1290-1295 Vol.2. https://doi.org/10.1109/ASCC.2004.185040

Zhou, C., Liu, Q., Huang, D., & Zhang, J. (2012). Inferential estimation of kerosene dry point in refineries with varying crudes. *Journal of Process Control*, *22*(6), 1122–1126. https://doi.org/10.1016/j.jprocont.2012.03.011