UNIVERSITI PUTRA MALAYSIA

RESERVOIR INFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORK AND ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM TECHNIQUES

SHAHRAM KARIMI GOOGHARI

RESERVOIR INFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORK AND ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM TECHNIQUES

SHAHRAM KARIMI GOOGHARI

DOCTOR OF PHILOSOPHY
UNIVERSITI PUTRA MALAYSIA

2007

RESERVOIR INFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORK AND ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM TECHNIQUES

By

SHAHRAM KARIMI GOOGHARI

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

December 2007

Dedicated to the author’s beloved Father, Mother and his dear family

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

RESERVOIR INFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORK AND ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM TECHNIQUES

By

SHAHRAM KARIMI GOOGHARI

December 2007

Chairman: Professor Ir. Lee Teang Shui, PhD

Faculty: Engineering

A feed forward Artificial Neural Network (ANN) and an Adaptive Neuro-Fuzzy Inferences System (ANFIS) reservoir inflow models were developed to investigate their potential in forecasting reservoir inflows. The site for the study is the Sembrong dam catchment which is located about 10km from Air Hitam town on the Air Hitam-Kluang road in the state of Johor, with an area of 130 square kilometers. The models consists of 9 inputs (previous last five-day reservoir inflow and last four-day average rainfall across the catchment) and are able to forecast the next day inflow into the reservoir. Average rainfall across the catchment was calculated by Theissen polygons. The 6 years daily data from 1995-1997 and 2002-2004 were used for training and validation of the models. Cross validation of training and validation data sets was also considered to obtain the best data set. Daily reservoir inflow was computed using a water balance equation. The reservoir inflow and rainfall data sets were examined for normal distribution and the best data transformation was used. Autocorrelation, partial autocorrelation and cross

correlation functions were used to find the best model inputs. The ANN models were trained and simulated using a written program in MATLAB environment (M-file) with raw and transformed data. The ANFIS models were built using the Fuzzy Toolbox of MATLAB. The Subtractive Clustering (SC) technique was employed to find the optimal number of rules. Different ANFIS structures were constructed by changing the SC parameters. All models were trained by the ANFIS editor of MATLAB with hybrid method. An M-file was written for calculating the different performance criteria of ANFIS models after simulating models during training, validation and testing. After selecting the best ANFIS structure, the response of the model to different types of membership functions was investigated.

The models were tested with the 10 months daily data of 2005. The best architecture of the ANN model was a 9-13-1 model which means a model with 9 inputs, 1 hidden layer with 13 neurons and 1 output. The model was trained based on the Leven-berg Marquardt algorithm with sigmoid activation functions. Simulation results for the independent testing data series showed that the model can perform well in simulating peak flows as well as base flows. The ANN model has been constructed for a strong non-linear input/output data. Comparisons of different ANN models for different data sets revealed that cross validation of data was effective in improving models performances. Data pre-processing to transform data to normal distribution before the training, results in better generalization and persistency of ANN models during testing.

The ANFIS models were built using the best data subset resulting from ANN modeling. The models were trained with normalized and non-normalized data. The selected ANFIS model was trained with normalized data with 6 Gaussian membership functions for each of 9 inputs and 6 rules. Comparisons of different performances of ANFIS models showed that data normalization can improve the model performances during training and testing. Simulation results for the independent test data series by the ANFIS model showed the ability of this model to forecast daily reservoir inflow in a tropical ungauged catchment. Sensitivity of the ANFIS model using different types of membership functions indicated that the best one is the Gaussian membership function.

The simulation results from the selected ANFIS and ANN models during training, validation and testing revealed the superiority of the ANN model. The selected ANFIS model gives lower values in most of the performance indices during training. For validation and testing, all performance indices of selected ANFIS model were inferior to those of the ANN model. The weakness of ANFIS model is shown in its inability to forecast individual peak flows. The sudden flow changes in these small tropical catchments resulting in these peak flows are common due to their small areal extent and to the intense localized phenomenon of tropical showers.

Keywords: ungauged tropical catchment, reservoir inflow forecasting model, artificial neural networks, neuro-fuzzy inference system, data transformation, data clustering.

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

PERAMALAN ALIRAN MASUK TAKUNGAN MENGGUNAKAN TEKNIK RANGKAIAN SARAF DAN TEKNIK SISTEM TAABIR KELAM SARAFSUAI

Oleh

SHAHRAM KARIMI GOOGHARI

Desember 2007

Pengerusi: Profesor Ir. Lee Teang Shui, PhD

Fakulti: Kejuruteraan

Model peramalan Rangkaian Saraf Buatan suapan ke depan (ANN) dan model peramalan Sistem Taabir Kelam Saraf Suai (ANFIS) bagi aliran masuk takungan telah dibangunkan untuk menyiasat pontensinya bagi meramal aliran masuk takungan. Tapak kajian ialah kawasan tadahan Empangan Sembrong, sebuah kawasan tadahan tropika 130 kilometer persegi, terletak 10km dari bandar Air Hitam di Jalan Air Hitam ke Kluang di dalam Negeri Johor. Model mengandungi sembilan masukan, aliran masuk lima-hari terakhir dan purata hujan empat-hari terakhir di kawasan tadahan dan boleh digunakan untuk meramalkan aliran masuk takungan pada hari selepasnya. Hujan purata di kawasan tadahan dikira dengan membentuk poligon Theissen. Data harian selama enam tahun daripada 1995-1997 dan 2002-2004 telah diguna untuk latihan dan pengesahan model.Pengesahan silang data diambilkira untuk menghasilkan set data terbaik. Aliran masuk takungan harian dikira dengan persamaan perseimbangan air. Data aliran masuk takungan dan hujan diperiksa untuk taburan normal dan penjelmaan data

terbaik telah diguna. Autosekaitan, autosekaitan separa dan fungsi sekaitan silang telah diguna untuk menghasilkan masukan model terbaik. Model ANN dilatih dan disimulasikan menggunakan program ditulis dalam sekitaran MATLAB dengan menggunakan data asas dan terjelma. Model ANFIS dibangunkan menggunakan Fuzzy Toolbox MATLAB. Teknik gugusan subtraktif telah digunakan untuk mendapatkan jumlah peraturan optimum. Struktur ANFIS yang berbeza telah dibina dengan menukar parameter gugusan subtraktif. Semua model telah dilatih menggunakan Penyunting ANFIS MATLAB dengan kaedah kacukan. Satu- fail-M ditulis bagi mengira kriteria berbeza perlakuan model ANFIS selepas model disimulasi pada masa latihan, pengesahan dan pengujian. Selepas struktur ANFIS terbaik dipilih, respon model kepada jenis fungsi keahlian yang berbeza telah disiasat.

Model tersebut telah diuji dengan data harian sepanjang sepuluh bulan dalam 2005. Seni bina terbaik model ANN ialah 9-13-1 yang bermakna 9 masukan, 1 lapisan tersembunyi dengan 13 neuron dan 1 keluaran. Model itu dilatih berdasarkan algoritma Leven-berg Marquardt dengan fungsi pengaktifan sigmoid. Keputusan simulasi untuk siri data ujian tak bersandar menunjukkan bahawa model boleh berprestasi baik bagi aliran kemuncak serta aliran asas. Model ANN dibentuk untuk data masukan/keluaran tak-linear yang padu. Perbandingan model ANN yang berbeza bagi set data berlainan, menunjukkan bahawa pengesahan silang data berkesan untuk memperbaiki prestasi model. Pra-proses data untuk menghasilkan data ke taburan normal sebelum latihan menghasilkan model ANN yang lebih baik ketika diuji.

Model ANFIS dibangunkan menggunakan subset data terbaik hasil daripada pemodelan ANN. Model diuji dengan data ternormal dan tak-ternormal. Model ANFIS terpilih diuji dengan data ternormal dengan 6 fungsi keahlian Gaussian bagi setiap 9 masukan dan 6 peraturan. Perbandingan prestasi berbeza model ANFIS menunjukkan bahawa penormalan data boleh memperbaiki prestasi model semasa latihan dan ujian. Keputusan simulasi bagi siri data ujian tak bersandar oleh model ANFIS menunjukkan kebolehan model itu untuk meramalkan aliran masuk takungan harian di dalam sebuah kawasan tadahan tropika yang tak bertolok. Kepekaan model ANFIS yang disemak berdasarkan jenis fungsi keahlian yang berbeza mununjukkan bahawa fungsi keahlian Gaussian adalah yang terbaik.

Keputusan simulasi daripada model ANFIS dan ANN terpilih semasa latihan, pengesahan dan ujian mendedahkan kelebihan model ANN. Model ANFIS yang terpilih itu menghasilkan nilai lebih rendah bagi kebanyakan indeks prestasi semasa latihan. Bagi pengesahan dan ujian, semua indeks prestasi model ANFIS pilihan adalah kurang baik daripada yang dihasilkan oleh model ANN. Kelemahan model ANFIS dalam meramalkan alir masuk kemuncak adalah nyata. Perubahan aliran mengejut di dalam kawasan tadahan tropika yang kecil adalah biasa oleh kerana keluasan yang kecil dan oleh kerana fenomena hujan tropika yang lebat dan setempat.

ACKNOWLEDGEMENTS

I thank GOD for all His blessings on me and thank Him for giving me courage and strength to finish my study.

I would like to express my sincere appreciation and gratitude to my supervisor, Professor Ir. Dr. Lee Teang Shui, for his guidance throughout the duration of this research. I would like also to take this opportunity to thank the supervisory committee members, Professor Ir. Dr. Mohd. Amin Mohd. Soom and Associate Professor Dr. Abdul Halim B. Ghazali, for their valuable assistance and support.

I wish to express my sincere appreciations to the staff of the Hydrology Division, the Department of Irrigation and Drainage, Jalan Ampang, especially to engineer Azmi bin. Md. Jafri and Puan Norhayati Md. Fadzil, for their assistance in providing the hydrological data; Puan Masilah Saim, technician in charge of the operations of Sembrong Dam of the Department of Irrigation and Drainage in Batu Pahat for providing operating policies on Sembrong Dam and Reservoir, and daily records of rainfall data, water supply data and reservoir water level and gate opening measurements.

I would thank the Ministry of Science, Research, and Technology (MSRT) of the Islamic Republic of Iran for the financial support given to me. Last but not least, I wish to convey my sincere thanks and love to my wife and my son for their sacrifices and patience throughout the duration of my study.

I certify that an Examination Committee met on 06 December 2007 to conduct the final examination of SHAHRAM KARIMI GOOGHARI on his Doctor of Philosophy thesis entitled “Reservoir Inflow Forecasting Using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System Techniques” in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Desa Ahmad, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Mohamed bin Daud, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Samsul Bahari Mohd Noor, PhD
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Van-Thanh-Van Nguyen, PhD
Professor
Department of Civil and Applied Mechanics,
Faculty of Engineering
McGill University
Canada
(External Examiner)

HASANAH MOHD. GHAZALI, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

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 are as follows:

Lee Teang Shui, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Mohd Amin Mohd Soom, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Abdul Halim Ghazali, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

AINI IDERIS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 21 February 2008

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at University Putra Malaysia or at any other institutions.

SHAHRAM KARIMI GOOGHARI

Date: 21/12/2007

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES