52. IWK

Internationales Wissenschaftliches Kolloquium
International Scientific Colloquium

PROCCEDINGS | 10 - 13 September 2007

FACULTY OF COMPUTER SCIENCE AND AUTOMATION

COMPUTER SCIENCE MEETS AUTOMATION

VOLUME II

Session 6 - Environmental Systems: Management and Optimisation
Session 7 - New Methods and Technologies for Medicine and Biology
Session 8 - Embedded System Design and Application
Session 9 - Image Processing, Image Analysis and Computer Vision
Session 10 - Mobile Communications
Session 11 - Education in Computer Science and Automation

TECHNISCHE UNIVERSITÄT
ILMENAU

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbiografie; detaillierte bibliografische Daten sind im Internet über **http://dnb.ddb.de** abrufbar.

ISBN 978-3-939473-17-6

Impressum

Herausgeber:	Der Rektor der Technischen Universität Ilmenau Univ.-Prof. Dr. rer. nat. habil. Peter Scharff
Redaktion:	Referat Marketing und Studentische Angelegenheiten Kongressorganisation Andrea Schneider Tel.: +49 3677 69-2520 Fax: +49 3677 69-1743 e-mail: kongressorganisation@tu-ilmenau.de
Redaktionsschluss:	Juli 2007
Verlag:	Technische Universität Ilmenau/Universitätsbibliothek Universitätsverlag Ilmenau Postfach 10 05 65 98684 Ilmenau www.tu-ilmenau.de/universitaetsverlag
Herstellung und Auslieferung:	Verlagshaus Monsenstein und Vannerdat OHG Am Hawerkamp 31 48155 Münster www.mv-verlag.de
Layout Cover:	www.cey-x.de
Bezugsmöglichkeiten:	Universitätsbibliothek der TU Ilmenau Tel.: +49 3677 69-4615 Fax: +49 3677 69-4602

Preface

Dear Participants,

Confronted with the ever-increasing complexity of technical processes and the growing demands on their efficiency, security and flexibility, the scientific world needs to establish new methods of engineering design and new methods of systems operation. The factors likely to affect the design of the smart systems of the future will doubtless include the following:

- As computational costs decrease, it will be possible to apply more complex algorithms, even in real time. These algorithms will take into account system nonlinearities or provide online optimisation of the system's performance.
- New fields of application will be addressed. Interest is now being expressed, beyond that in "classical" technical systems and processes, in environmental systems or medical and bioengineering applications.
- The boundaries between software and hardware design are being eroded. New design methods will include co-design of software and hardware and even of sensor and actuator components.
- Automation will not only replace human operators but will assist, support and supervise humans so that their work is safe and even more effective.
- Networked systems or swarms will be crucial, requiring improvement of the communication within them and study of how their behaviour can be made globally consistent.
- The issues of security and safety, not only during the operation of systems but also in the course of their design, will continue to increase in importance.

The title "Computer Science meets Automation", borne by the 52nd International Scientific Colloquium (IWK) at the Technische Universität Ilmenau, Germany, expresses the desire of scientists and engineers to rise to these challenges, cooperating closely on innovative methods in the two disciplines of computer science and automation.

The IWK has a long tradition going back as far as 1953. In the years before 1989, a major function of the colloquium was to bring together scientists from both sides of the Iron Curtain. Naturally, bonds were also deepened between the countries from the East. Today, the objective of the colloquium is still to bring researchers together. They come from the eastern and western member states of the European Union, and, indeed, from all over the world. All who wish to share their ideas on the points where "Computer Science meets Automation" are addressed by this colloquium at the Technische Universität Ilmenau.

All the University's Faculties have joined forces to ensure that nothing is left out. Control engineering, information science, cybernetics, communication technology and systems engineering – for all of these and their applications (ranging from biological systems to heavy engineering), the issues are being covered.

Together with all the organizers I should like to thank you for your contributions to the conference, ensuring, as they do, a most interesting colloquium programme of an interdisciplinary nature.
I am looking forward to an inspiring colloquium. It promises to be a fine platform for you to present your research, to address new concepts and to meet colleagues in Ilmenau.

Professor Peter Scharff
Rector, TU Ilmenau

Professor Christoph Ament
Head of Organisation

Table of Contents

CONTENTS

Page

6 Environmental Systems: Management and Optimisation

10 Mobile Communications

11 Education in Computer Science and Automation

D. Karimanzira/ M. Jacobi

Modelling yearly residential water demand using neural networks

ABSTRACT

Demand management plays an increasingly important role in dealing with water scarcity in Beijing. It is important to understand the level and pattern of water use in various sectors across the regions for any measures being put into effect. The objective of this study is to enhance the understanding of the factors that influence water demand by examining closely the water use in the domestic sector and to develop a predictor for the yearly domestic water demand. Neural network techniques are investigated in an application to identify and subsequent offline prediction of a process "residential water demand" exhibiting nonlinearities (coupled determinants) and typical disturbances. The design and development of neural network process model from measured data is described, and practical aspects of the identification procedure are discussed. Results demonstrate that the developed neural network representation of the process behaviour is sufficiently accurate to be used independently from the process, emulating the process response from only process input information. Accurate long range predictions from the neural network model are mainly due to the use of a novel spread encoding technique for representing data in the network. Implementation of a predictive strategy incorporating the identified neural network model is described. Results illustrate the improvements in prediction performance that can be achieved when compared to regression models.

INTRODUCTION

As water supply fails to meet the demand in many areas, careful analysis of decisions on the allocation of water is of great significance. The past policy responses to water scarcity are mainly supply oriented and aim at fostering the development and exploitation of new sources and expansion of the network infrastructure to guarantee the water supply. In recent year water policies have increasingly addressed demand management, which means development of water conservation and management programs to influence water demand. Demand driven measures include adoption of water-saving technologies and appliances, awareness raising and economic instruments such as price and tax. Population growth and urbanization and overall expansion in economic activities are the major factors underlying the increase in water consumption.

This work is part of the Chinese - German joint project "Towards Water-Scarcity Megalopolis' Sustainable Water Management System"[6]. This project takes the challenge of water shortage, the outstanding conflict between water supply and demand. It aims at a decision support system (DSS) for the sustainable development of economics and community in Beijing. An essential requirement for such a DSS is a simulation model of the water resources/supply system. The simulation model comprises the water supply, optimization and the water demand systems. On focus in this paper is part of the water demand system, namely the domestic water demand. Compared to the agricultural and the industrial water demand is the domestic water demand very difficult to model, because it is determined by several subjective factors.

This paper proceeds as follows. Section 2 provides a brief introduction to modelling with neural networks. Section 3 describes the issues related to water demand estimation. Section 4 provides the results of the performance of the developed prediction system in comparison to regression techniques applied in several literature, followed by a short summary and conclusions.

NONLINEAR MODELLING WITH NEURAL NETWORKS

In this work it has been chosen to restrict the attention to the so called multilayer perceptron neural networks (MLP) for modelling nonlinear processes. However, the particular choice of nonlinear model description is not vital for the predictor design. Other types of generic nonlinear model structures might be used instead.

Many different types of MLP based model structures can be considered when identifying nonlinear processes. However, it is typically assumed that the process can be described by the general model

$$y(t) = \hat{y}(t \mid t-1, \theta) + e(t) = g(\varphi(t), \theta) + e(t) \tag{1}$$

where $\varphi(t)$ is the regression vector composed of the past information, g is the function realized by the MLP network, θ is the model parameters (the weights), and e(t) is additive noise which is assumed to be white and independent of the past information. By inserting g, which is assumed to be a two-layer MLP network with tangent hyperbolic activation functions in the hidden units and a linear output unit, the predictor takes the form:

$$\hat{y}(t \mid \theta) = \sum_{j=1}^{q} W_j \tanh\left(\sum_{i=1}^{n_\varphi} w_{ji} \varphi(t) + w_j \right) + W_0 \tag{2}$$

where the components of θ, W_j and w_{ji}, specifies hidden-to-output layer and input-to-hidden layer weights, respectively. To reduce the degrees of freedom and because it is advantageous in many predictor system designs, it is common to consider model structures that are natural extensions of well-known linear model structures like *ARX* and *OE* in a similar regression vector is considered. The weights are estimated from a set of corresponding input-output pairs z^N={[u(t),y(t)]:t=1,…,N} acquired in a practical experiment with the residential water demand process.

DETERMINANTS OF THE RESIDENTIAL WATER DEMAND SYSTEM

Possible influencing factors for the domestic water demand for the Beijing region were selected and are listed in Figure 2 for the period from 1996-2005 [5]. They are factors from the weather, population and economy, which are thought to be obviously linked to the domestic water demand. The main objective of this study is to find out, which factor has the greatest influence and which ones are negligible in the models, and therefore simplify them. Correlation, regression and significance analysis described in the previous section were performed for the domestic water demand with respect to all input variables including the previous domestic water demand W_{prev}.

All estimated cross-correlation coefficients r_{xy} which were significant at the 5% level according to the three step procedure described above are summarized in Table 2 (bold) and Figure 3. Surprisingly, the magnitude of the total cross correlation coefficient of the domestic water demand and the number of employment is quite large (-0.777). A decreasing domestic water demand by increasing employment can not be explained logically. The partial correlation coefficient (a measure for the dependence of two variables after switching of the linear influences of other variables) between employment and the domestic water demand confirms this.

After switching of the linear influences of other variables the remaining partial correlation coefficient is only -0.065, which practically shows no linear dependence between the two variables. Also unexpected is the minimal correlation of the domestic water demand and the population (correlation coefficient of -0.016). Normally, one would think that the domestic water demand increases with increasing population growth. The correlation coefficients of the Beijing panel data were tested for significance and only the previous water demand, GDP per capita, employment, time and number of households (H) were significant at the 5% level (see Table 1). Therefore, it is recommended to include these variables as inputs in reduced forecasting models.

Regression results also show that several combinations of these variables are possible to obtain a reliable model. Most of the coefficients of the explanatory variables have expected signs. The positive value of temperature suggests domestic consumers use more water when the weather is relatively warm. Precipitation contributes negatively to water consumption, meaning that households tend to use less water when there is enough rainfall. Family size and water price (not shown) are not significant at any level, which may be due to the fact that both variables vary little with time.

	T	P	GDP	H	Tmp	prec	E	W	gdpc
T	1.000	0.934	0.991	0.993	0.047	-0.546	0.406	0.183	-0.294
P	0.934	1.000	0.928	0.947	-0.199	-0.413	0.548	-0.016	-0.537
GDP	0.991	0.928	1.000	0.996	0.079	-0.489	0.501	0.078	-0.373
H	0.993	0.947	0.996	1.000	0.002	-0.486	0.499	0.676	-0.384
Tmp	0.047	-0.199	0.079	0.002	1.000	-0.243	-0.133	0.332	0.382
prec	-0.546	-0.413	-0.489	-0.486	-0.243	1.000	0.069	-0.500	-0.199
E	0.406	0.548	0.501	0.499	-0.133	0.069	1.000	-0.777	-0.899
W	0.183	-0.016	0.078	0.676	0.332	-0.500	-0.777	1.000	0.824
gdpc Eva	-0.294 —	-0.537 —	-0.373 —	-0.384 —	0.382 —	-0.199 —	-0.899 —	0.824 -0.726	1.000 —

Table 1: Correlation coefficients of the variables

Important was also to find the robustness and the generality of the influencing parameters. Therefore, the correlation coefficient for the previous water demand, price, employment, time and number of households to domestic water demand were calculated for other regions of different nature where data could be obtained. Data for Germany and Canada was present and the calculated correlation coefficients were almost similar to that of Beijing, *T, E* and *H* had the highest correlation with the domestic water demand, which suggests that models for forecasting domestic water demand that include these variables are quite reliable.

To avoid multicollinearity in the models, the population density *P* is excluded to be an explanatory variable as it is highly correlated with *GDP* and *H*.

MODEL DEVELOPMENT AND TEACHING

NN MODEL FOR THE RESIDENTIAL WATER DEMAND PREDICTION

It was found that the yearly water demand in the previous year, the employment rate, the number of households and the *gdpc* were more significant variables affecting the yearly water demand in any given year [3]. Therefore, the neural network model consisted of four neurons in the input layer representing water demand in the previous year, the employment rate, the number of households and the *gdpc*. The whole data set was divided into two parts, namely, “training set” and “testing set”. The training set was used to train the neural network model, whereas the testing set was used to test the performance of the neural network model in terms of some statistical parameters explained later.

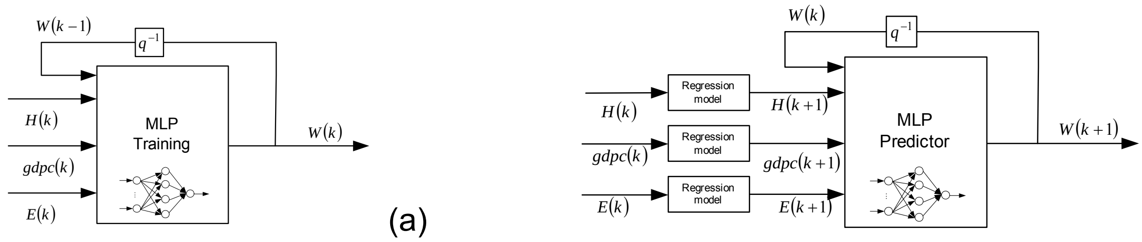

Figure 1: Structure of the neural networks (a) Training and (b) Prediction

REGRESSION MODEL FOR THE RESIDENTIAL WATER DEMAND PREDICTION

A linear multiple regression model was also developed using the yearly water demand in the previous year, the employment rate, the number of households and the *gdpc* as the explanatory variables. The structure of the model can be expressed as follows:

$$W = \alpha + \beta H + \gamma E + \delta W_{prev} + \mu gdpc \tag{3}$$

where α, β, γ, δ and μ are the regression coefficients to be determined, *H* is the number of households, *E* is the employment rate, W_{prev} is the previous yearly water demand and *gdpc* is the gross domestic product per capita**.** The linear multiple regression model was calibrated to determine the values of the regression coefficients using the same data set used for training the neural network model, and was tested using the same data set used to test the neural network model. Due to the fact that there could be some effects that are correlated to some explanatory variables the OLS is biased and inconsistent. Therefore, the parameter estimation was done using feasible generalized least squares analysis. The parameters in Table 2 were estimated for the regression model. Their standard deviation and coefficient of variation in % were also calculated.

Parameter	**Best Estimate**	**Standard Deviation**	**Coefficient of variation (%)**
α	12.799	0.7511	586.817
β	-8,9471E-4	0.0019	-2.108.406
γ	-6,8e74E-4	9,7115E-4	-142.036
δ	-0.2142	0.486	-2.269.412
μ	0.0038	0.0016	41.849

Table 2: Estimated parameters of the regression model

RESULTS

MODEL PERFORMANCE

Figure 2 and Table 3 shows the qualitative and quantitative results of the regression/neural network model, respectively. The sum of the squares of the residuals, the multiple correlation coefficients for I/O data and the linear correlation coefficient of measured output and the calculated output show very good results.

	Model results	
	Regression	NN
Sum of squares	0.0011	2.90e-007
Correlation: x - y data		
Multiple Correlation Coefficient	0.9879	
Correlation: y(experimental) - y(calculated)		
Linear Correlation Coefficient Linear Correlation Coefficient Probability	0.9879 2.2028E-6	

Table 3: Regression model results

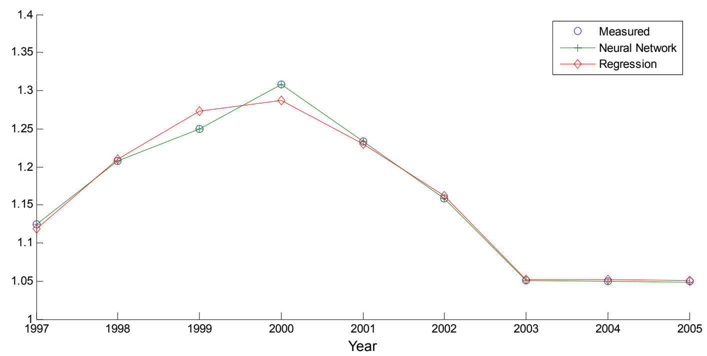

Figure 2: Results of the two models (Water Demand (100mil m^3))

To compare the two models, the performance of each was quantified using two different statistical measures. The first statistical measure was "average absolute relative error" (*f*). The average absolute relative error may be calculated using the following equations:

$$f = (1/N)\cdot abs(sum(RE_t)) \qquad (4)$$

$$RE_t = \{(DO_t - DF_t)/DO_t\}*100 \qquad (5)$$

where RE**,** is the relative error in the forecasting during year t, DO_t and DF_t, are the observed and forecasted water demand during year t. The second statistical measure was the "threshold statistic", which quantifies consistency in the forecasting errors from a particular model. The threshold static is calculated for a particular level of relative error (say, x %) in forecasting. A threshold statistic may be calculated using the following equation:

$$TS = \{n/N\}*100 \qquad (6)$$

where TS, is the threshold statistic for a level of x %, n is the number of data points forecasted having relative error in forecasting less than x %, and N is the total number of data points forecasted. In this study, the threshold statistics were calculated for levels of 1%, 2%, 5%.

Level of threshold *x(%)*	Average Absolute Relative Error *f*		TS (%)	
	Regression model	NN-Model	Regression model	NN-Model
	0.0498	0.0095		
1	-	-	100	71.43
2	-	-	100	100
5	-	-	100	100

Table 4: Statistical measures for the models

DISCUSSION OF RESULTS

The results in terms of various statistical measures obtained from all the models developed in this study are presented in Table 4. Observed and forecasted yearly water demands from the neural network and regression models for the testing period are shown in Figure 2. It is clear from the Table 1 that the neural network model performed the best in terms of and all levels of threshold statistics. The lowest average absolute relative error achieved by the best model was 0.0095%. In 71.43% of the forecasted water demands, the absolute relative error in forecasting was less than 0.3541%. All forecasted errors from the best model were less than 1.88 %.

CONCLUSIONS

In the study presented here, a neural network and a regression model have been developed to forecast yearly water demand for Beijing, China. In this study several variables have been tested for their influence on the domestic water demand. It has been shown that to predict domestic water reliably at least the *gdpc*, the previous water demand, employment rate, the time and the number of households must be included. The estimation can be improved by using panel data covering a longer time period or more disaggregated sub-regional level analyses. It would also be useful to extend the study with more adequate data especially regarding time series water prices for the domestic sector. Well-designed household surveys would provide richer information and greater insights into the factors influencing domestic water demand. It was observed that the water demand in Beijing is a function of past water demand, the employment rate, the *gdpc* and the number of households. Based upon the results obtained in this study, the following conclusions can be made: the results obtained in this study are highly promising demonstrating the superiority of artificial neural networks over the conventional techniques such as regression analysis.

References

[1] Rosegrant, M., Ringler, C., Msangai, S., Cline, S. *IMPACT-WATER: Model Description*. Int. Food Policy Research Inst, Washington, D.C., 2005.

[2] *Beijing Municipal Bureau of Statistics. Beijing Statistical Information net*, Mai 2005. http://www.bjstats.gov.cn/english/index.html.

[3] Rauschenbach T. *Water Resources in Beijing and the Demand Analysis*, Juli 2005.

[4] Brockwell, P.J. and Davis R. A. Time Series: Theory and Methods. Springer Verlag, New York. 1991

Authors:
Divas Karimanzira
Marco Jacobi
TU-Ilmenau, Gustav Kirchhoff 1, P.O.B. 1005065
98684, Ilmenau
Phone: 03677691421
Fax: 03677691434
E-mail:divas.karimanzira@tu-ilmenau.de