Biosystems Engineering (2007) **96** (1), 135–138
doi:10.1016/j.biosystemseng.2006.09.003
SW—Soil and Water

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER

RESEARCH NOTE

Reservoir Level Forecasting using Neural Networks: Lake Naivasha

S. Ondimu; H. Murase

Bio-instrumentation, Control and Systems (BICS) Engineering Laboratory, College of Life and Environmental Sciences, Osaka Prefecture University, 1-1, Gakueni- cho, Sakai, Japan; email of corresponding author: ondimu@bics.envi.osakafu-u.ac.jp

(Received 5 September 2005; accepted in revised form 21 September 2006; published online 13 November 2006)

Six feature groups comprising of water levels, rainfall, evaporation rate, discharges for rivers Malewa and Gilgil and one pair of time harmonics were used to develop neural network models to forecast water levels for Lake Naivasha in Kenya. Six elements were used from each feature group. Some feature groups were compressed using the Karhunen–Loeve Transform (KLT) to reduce their dimensions.

The neural network models developed were able to forecast effectively the reservoir levels for the lake for four consecutive months after a given month and given data for six consecutive months prior to the month. It was found that the more the number of feature groups used, the higher the ability of neural networks to forecast accurately the reservoir levels. Data compression generally reduced the size and computation time of the models.

1. Introduction

A neural network (NN) provides an attractive modelling technology for reservoir level forecasting and planning because historical data can be used to a reasonable level of certainty (Melesse, 2005; Sepaskhah & Akbari, 2004; Yu *et al.*, 2004). This paper reports the use of NN to forecast water levels in Lake Naivasha, which is located 100 km northwest of Nairobi, Kenya. The geographic location is 00° 40′ to 00° 53′ S latitude and 36° 30′E longitude. Its altitude is approximately 1890 m above sea level. The main recharge of the lake is rain and inflow from rivers Malewa and Gilgil. It has no known surface outlet and its main discharge is by evaporation and underground seepage.

The lake is of great economic importance (Perera, 2002; Lopez, 2002; Gorroxategi, 2001). However, its current annual abstraction is about six times higher than a safe yield determined by Becht and Harper (2002). Over-exploitation of the lake is visible in many environmental adverse effects (Amani, 2004; Hickley *et al.*, 2004) and has attracted attention from many media houses (News24, 2003; Al Jazeera, 2003; Nyambura, 2003a, 2003b; Ngware, 2002; Black, 2004; All Africa, 2003; FAO, 2002; The Star, 2003).

Many water models have been studied in relation to Lake Naivasha (Amani, 2004; Perera, 2002; Lopez, 2002), but none has focussed on forecasting its water levels.

The purpose of this study was to use NN to forecast the monthly water level of Lake Naivasha. The main objective was to study the effect of data compression and number of feature groups on reservoir water level forecasting using NN.

2. Method and theory

2.1. *Data features and data sets*

Historical data for the years 1960–1997 was used as input sets to develop forecasting NN models. The data was grouped into six feature groups: FG1 (past monthly water levels in m above sea level); FG2 (rainfall in mm per month); FG3 (evaporation in mm per month); FG4 (inflow from River Malewa in $m^3 s^{-1}$); FG5 (inflow from River Gilgil in $m^3 s^{-1}$) and FG6 (simple time harmonics

Notation

A	transformation matrix	n	month number
$E(k)$	total mean square error for the kth output unit	t_{pk}	kth element of the pth desired output vector
i	counter for 1, 2, …, k	x	part of the training vector to be compressed
j	number of pairs of harmonics	y_{pk}	the kth output for pth training pattern.
k	number of feature groups present	z	input vector after applying Karhunen–Loeve Transform (KLT)
N	number of elements in the input or output vector		
N_i	number of elements in the ith feature group		
N_v	dimension of input pattern		

of month of the year). Six elements from each feature group were joined to the input vector elements by the relation

$$N = \sum_{i=1}^{k} N_i \tag{1}$$

where: N is the number of elements in a data set vector; k is the number of feature groups; N_i number of elements per feature group.

Feature group elements were obtained for 6 months prior to the given month. One simple time harmonics was used for the month in question, giving time features $\sin 2\pi nj/12$ and $\cos 2\pi nj/12$, where: j denotes the number pairs of harmonics; and n is the month number of the year.

Therefore, the maximum dimension of the input vector was 32. However, this was fixed at 20 to avoid the curse of dimensionality due to a limited number of input data classes. To ensure this, the Karhunen–Loeve Transform (KLT) was applied [Eqn (2)] to feature groups FG2, FG3, FG4 and FG5.

$$z = A{\cdot}x \tag{2}$$

where: z is the compressed input vector; x is the part of the vector to be compressed; and A is a transformation matrix.

2.2. *Neural network models*

Six different multilayer Perceptron (MLP) NN models:—NN1, NN2, NN3, NN4, NN5 and NN6 were developed. A three-layered architecture with 10 neurons in the hidden layer and four neurons in the output layer was adapted for all the models after trial and error. The number of neurons in the input layer varied with the number of elements in the input vector for each model. The NN models were trained using output weight optimisation–hidden weight optimisation (OWO–HWO) training algorithm (Chen *et al.*, 1999; Yu & Manry, 2002). Training was completed in 500 iterations for each of the models.

Data from 1960 to 1990 was used for training, data from 1991 to 1995 was used for testing, and data from 1996 to 1997 was used for validating the models. Forecasting was done for four consecutive months following a given month. The desired outputs were the observed lake levels corresponding to the 4 months. Inputs were feature group elements for 6 months prior to the month in question and one simple time harmonics for the month.

Inputs to model NN1 were from feature groups FG1 and FG6; model NN2, from feature groups FG1, FG2, FG3 and FG6; model NN3, same as in model NN2, but with feature groups FG2 and FG3 compressed to two elements each; model NN4, same as in model NN2 but with feature groups FG2 and FG3 compressed to three elements each; model NN5, from all feature groups but with feature groups FG2, FG3, FG4, and FG5 compressed to two elements each and model NN6 same as in model NN5 but with feature groups FG2, FG3, FG4, and FG5 compressed to three elements each.

The overall performance of each NN model was evaluated using a total mean square error (TMSE) given by

$$E(k) = \frac{1}{N_v}\sum_{p=1}^{N_v}[t_{pk} - y_{tk}]^2 \tag{3}$$

where: $E(k)$is the TMSE of the kth output unit; N_v is the dimension of the input pattern; t_{pk} is kth element of pth desired output vector; and y_{pk} is the kth output of the pth training pattern.

3. Results and discussion

Table 1 shows the training and testing TMSE of the six models. The training TMSE was highest in Model NN1 and lowest in model NN6. This implies that the more the number of feature groups in the training set the

Table 1
Training and testing mean square errors of neural network models developed in the study

Model	*Model architecture (input:hidden:output)*	*Training total mean square error for each output in percentage (after 500 iterations)*	*Testing mean square error for each output in percentage*
NN1	8:10:4	0·43	0·12
NN2	20:10:4	0·35	0·39
NN3	12:10:4	0·27	1·68
NN4	14:10:4	0·22	5·79
NN5	16:10:4	0·20	3·1
NN6	20:10:4	0·12	13·32

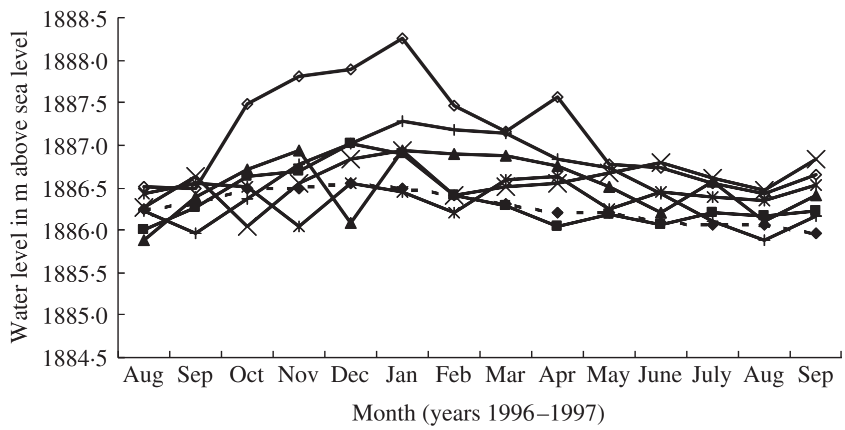

Fig. 1. Observed and forecast water levels at Lake Naivasha for 14 months (1996–1997): ···◆···, observed; —■—, model NN1; —▲—, model NN2; —✕—, model NN3; —✳—, model NN4; —◇—, model NN5; —+—, model NN6

greater the ability of the NN to forecast the reservoir level.

The TMSE for each output was more than that obtained during training for all the models. The error values for models—NN3, NN4, NN5 and NN6 were more than 1%.

Figure 1 represents the observed and forecast water levels for the Lake for 14 months in the years 1996 and 1997. Values forecast by model NN5 are the highest. Model NN1 does well in forecasting the reservoir levels, but its forecast values exhibit some oscillation. This is unexpected in light of the results in Table 1. Values forecast by model NN6 follow the pattern of the observed values closely and are more reliable.

Models with more feature groups were expected to exhibit better forecasting ability. Data compression may have introduced undesirable qualities into their input data by lowering their forecasting ability. Another possible reason for this observation could be that some feature groups were redundant. Use of river discharges in $m^3 s^{-1}$ is another possible source of error since the other feature groups were in monthly units.

Figure 2 shows the TMSE for each output for water levels forecast by the six models for 14 months (1996-1997). Models NN3 and NN5 have higher error values than models NN4 and NN6. Models NN3 and NN5 had elements in their input feature groups compressed to two, while the same elements were compressed to three in models NN4 and NN6. Thus, too much compression affects the forecasting ability of the NN models.

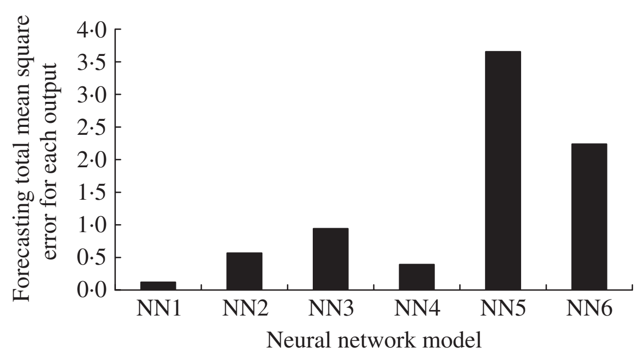

Fig. 2. Forecasting total mean square error per output for six neural network models developed in the study for years 1996–1997

4. Conclusion

The neural networks (NN) models developed in this study were able to forecast the water levels of Lake Naivasha for four consecutive months beginning after a given month and given data for six consecutive months prior to that month. Thus, NN provide an effective and timely method for forecasting water levels in the lake. This can help in water-use formulation and scheduling for domestic, municipal and agricultural uses. Timely forecasting can also help in disaster monitoring, response and control in areas prone to floods. For power generation, effective and timely reservoir level forecasting can help in predicting power loads and management of power generation for efficiency and optimisation.

The number of feature groups and the number of elements in each feature group used as inputs greatly influence the ability of NN to forecast reservoir levels accurately.

Data compression using the Karhunen–Loeve Transform (KLT) provides an optimal technique of reducing the size of input vectors and thus reduces the size of NN used. This greatly increases the forecasting ability of the NN models. However, data compression introduces undesirable qualities into the data that affects the forecasting ability of the NN. In addition, over-compression of data undermines the efficiency of NN in forecasting reservoir levels.

References

Al Jazeera (2003). Lake life dying as flowers bloom, 2003, Al Jazeera, Doha, Qatar. Available at: 〈http://English.aljzeera.net/〉 in the News, Accessed on: 1st January 2005

All Africa (2003). Flowers vs fish. Allafrica.com, Washington, DC, USA. Available at: 〈http://www.earthisland.org/eijournal〉, Accessed on: 1st January 2005

Amani A (2004). Modelling water resource management in Lake Naivasha. MS Thesis, International Institute for Geo-Information Science and Observation, Enschede, the Netherlands.

Becht R; Harper D M (2002). Towards an understanding of human impact upon the hydrology of Lake Naivasha, Kenya. Hydrobiologia, **488**(1–3), 1–11

Black R (2004). Cultivating an ethical flower. BBC News, London, UK. Available at: 〈http://news.bbc.co.uk〉, Accessed on: 1st January 2005

Chen H H; Manry M T; Chandrasekaran H (1999). A neural network training algorithm utilizing multiple sets of linear equations. Neurocomputing, **25**(1–3), 55–72

FAO (2002). A thorn on every rose for Kenya's flower industry. Food and Agriclture Organisation, Rome, Italy. Available at: 〈http://archives.foodsafetynnetwork.ca/〉, Accessed on: 1st January 2005

Gorroxategi G I (2001). Peak runoff estimation of the ungauged area in Naivasha basin using the slope area method and empirical formulas. Water Resources and Environmental Management. International Institute for Aerospace Survey and Earth Sciences, Enschede, The Netherlands.

Hickley P; Muchiri M; Boar R; Briton R; Adams C (2004). Habitat degradation and subsequent fishery collapse in Lakes Naivasha and Baringo, Kenya. Available at: 〈http://www.le.ac.uk/biology/〉, Accessed on: 5th August 2005

Lopez S M (2002). Papyrus conservation around Lake Naivasha: development of alternative management schemes, Kenya. MS Thesis, Enschede, The Netherlands: International Institute for Geo-Information Science and Observation

News24 (2003). Flowers threaten the lake. News24, Cape town, SA. Available at: 〈http://www.news24.com/〉, Accessed on: 1st January 2005

Melesse A M (2005). Multi-temporal scale flood prediction using Artificial Neural Network: A case study for Devils Lake and Red River of the North Basins. ASAE Paper No. 05-2115, St. Joseph, MI, USA

Ngware C (2002). Save Lake Naivasha waters. Daily Nation, Nairobi, Kenya. Available at: 〈http://www.afae.org/〉, Accessed on: 1st January 2005

Nyambura H (2003a). Cut-throat flowers. Reuters News, Nairobi, kenya. Available at: 〈http://www.animalfreedom.org/english/〉, Accessed on: 1st January 2005

Nyambura H (2003b). Kenya Fishers at odds with Flower Farms. Reuters news, Nairobi, kenya. Available at: 〈http://www.flmnh.ufl.edu/fish/innews/〉, Accessed on: 1st January 2005

Perera, L (2002). Development of information system to support hydroenvironmental–economic modeling on Lake Naivasha area. MS Thesis, International Institute for Geo-Information Science and Observation, Enschede, The Netherlands

Sepaskhah A R; Akbari D (2005). Deficit irrigation planning under variable seasonal rainfall. Biosystems Engineering, **92**(1), 97–106

The Star (2003). Lake's Problems stem from the flowers. The star, Cape Town, SA. Available at: 〈http://www.thestar.co.za/〉 in the News, Accessed on: 1st January 2005

Yu C; Manry M T (2002) A modified hidden weight optimization algorithm for feed-forward neural networks. Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers, Vol. 1, pp 1034–1038, IEEE press, Pacific Grove, California, USA

Yu C; Northcott W J; McIsaac G F (2004). Development of an artificial neural network for hydrologic and water quality modelling of agricultural water sheds. Transactions of the ASAE, **47**(1), 285–290