Asian Journal of Water, Environment and Pollution, Vol. 5, No. 1, pp. 57-65.

Use of Back-propagation Artificial Neural Networks for Groundwater Level Simulation

Azhar K. Affandi*, Kunio Watanabe and Haryadi Tirtomihardjo[1]

Geosphere Research Institute of Saitama University
Shimo Okubo 255, Sakura-ku, Saitama 338-8570, Japan
[1]Center of Environmental Geology, Geological Agency
Department of Energy and Mineral Resources
Jalan Diponegoro No. 5, Bandung 40122, Indonesia
azhar_affandi@yahoo.com

Received August 31, 2006; revised and accepted September 10, 2007

Abstract: This article presents simulation of groundwater level fluctuation based on an artificial neural network modelling. The prediction used multi-layer back-propagation neural networks (BPANN). The case of study area was Jakarta, Indonesia, that has high population density and several purposes of groundwater resource usage. Input variables were using delay five-daily groundwater level fluctuation (GLF) of observation well interest to predict current GLF. The applicability of BPANN for GLF prediction was verified in three sets of input variables. The result showed that application of BPANN to simulate GLF gives satisfied prediction results.

Key words: Groundwater level fluctuation, forecast, artificial neural network, back-propagation, Jakarta, Indonesia.

Introduction

Ground water is the most important water resource, especially in developing countries such as Indonesia, Thailand, and the Philippines. The ground water is mainly used for domestic and municipal consumption, agricultural and industrial water supply. Over-exploitation of ground water has happened in many cities of Asia such as Bangkok, Manila and Jakarta. These examples show that long-term sustainable resource has been destroyed through the depletion of aquifers, which has caused salinization and land subsidence (McIntosh, 2003). The relation between land subsidence and groundwater exploitation is very close. Excess pumping of ground water causes severe land subsidence. It can break building and deform the water supply canal system, and increase flooding. Therefore, groundwater resources development should be performed under the controlled condition. Groundwater management can be done by forecasting groundwater level fluctuation to maintain the equilibrium of groundwater resource. The popular method that can predict groundwater level fluctuation is an artificial neural network.

An artificial neural network is recognized as a very promising tool for relating input to output data. Artificial Neural Network (ANN) is a mathematical methodology that describes relations between causes (input) and effects (output). ANN has attracted a rapidly growing interest in the past decade due to rapid theoretical advances relative to ANN. The most important properties, which favoured their popularity, are the capacity to express complex nonlinear behaviour and the ability to learn from numerical examples. The major advantage of an ANN is its ability to represent underlying nonlinear dynamic of the system modelled without any prior assumption regarding the process involved (Coulibaly et al., 2001).

ANN as black-box model can be developed based only on numerical examples of process input–output behaviour and little prior knowledge about the process required.

*Corresponding Author

Black-box modelling is a feasible approach, when processes are only poorly understood. Theoretical modelling would be too complex and time consuming. On the other hand, since such a model typically consists of hundreds of parameters, significant amounts of numerical data are required to properly train the network. The needed amount exponentially increases with the dimensionality of network input space, that is, the number of model inputs.

The applications of ANN in groundwater simulation are not so many compared to surface hydrology. Some such applications as literature reviews are as follows: Ranjithan et al. (1993) used ANN-based screening tool for identifying critical realizations with respect to a groundwater remediation design under condition of uncertainty in hydraulic conductivity parameter. Aziz and Wong (1992) used ANN for determining aquifer parameter values from normalized data drawdown obtained from pumping test. Balkhair (2002) used ANN to estimate aquifer parameter value namely transmissivity and storage coefficient from pumping test data for large diameter well as was used by. Rajanayaka et al. (2002) to estimate hydraulic conductivity and dispersion coefficient. Groundwater level simulation by means of ANN has been conducted by Coulibaly et al. (2001) using three types of functionally different ANN models; Daliakopoulos et al. (2005) comparing seven different types of network architectures and training algorithms; Nayak et al. (2006) forecasting GLF in an unconfined aquifer and Lallahem et al. (2004) estimating groundwater level in fracture media.

The purpose of this study is to simulate groundwater level fluctuation (GLF) by means of predicting GLF using artificial neural network, even with relatively short length of data for training and testing phase. The Back Propagation Artificial Neural Network (BPANN) algorithm was used for simulating groundwater level base on time lag groundwater level fluctuation.

Artificial Neural Network

Most discussions of ANN begin with a biological motivation. The ANN concept is originally developed to simulate the human brain. ANN is an information processing system that roughly replicates the behaviour of a human brain by emulating the operations and connectivity of biological neurons. An ANN is constructed of artificial neurons, the mathematical elements within the network.

Neural network systems are different from the conventional systems, such as analytical models or statistical model. ANN is a network consisting of arbitrary number of very simple elements called neurons. Each neuron is simple processing element that responds to the connected (weighted) inputs it received from other neurons. The activation function intensifies or weakens the signal that is transmitted to other neurons. The arrangement of the neuron is referred to as the network architecture. Various network architectures are available. One of them that are applicable in many fields is feed-forward neural network with error back propagation training algorithm or so called Back Propagation Artificial Neural Network (BPANN). In BPANN, the goal of learning (or training) is to minimize the error between the desired outputs (target) and the actual outputs of the network. The training is conducted by back propagation algorithm. In order to reduce the training time, some improvements are suggested by momentum and learning rate for setting the initial weights to reduce the convergence time.

The BPANN was introduced by Rumelhart et al. (1986), and a good description of the BPANN in groundwater problem can be found in Ranjithan et al. (1993), Govindaraju (2000), among others; however only the review is presented here below for completeness.

Feed-forward

A feed-forward neural network consists of at least three layers: an input layer, one or more hidden layers and an output layer (see Figure 1). The network is constructed in such a way that each layer is fully connected to the next layer. Every neuron in input layer will send its output to every neuron in the hidden layers, and every neuron in the hidden layers will send its output to every neuron in output layers. The number of neurons in the input layer is equal to the number of variable input data. The number of neurons in the hidden layer can be varied depending on the complexity of problem and the size of input.

From Figure 1 it can be explained that each neuron j receives incoming signals from every neuron i in the previous layer. Each incoming signal (y_i) associates with a weight (w_{ji}). The net input, x_j, to neuron j is a sum of the incoming signal times the weight as described in equation 1.

$$x_j = \sum_i y_i w_{ji} \tag{1}$$

Note that this includes an extra neuron we call bias neuron, which is assumed to have a value of 1 at all times. The weight on this extra neuron is called the bias as a threshold value.

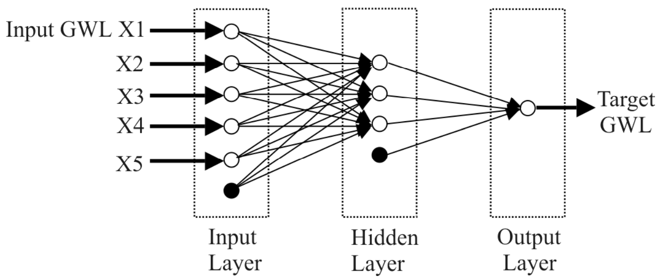

Figure 1: Topology of three-layer feed-forward Artificial Neural Network.

The output signal, y_j, which is a non-linear function is produced by transfer function of its net input. The most commonly used transfer or activation function is a sigmoid function. The sigmoid function is, in essence, a smooth version of a step function. It is zero for low input. At some point, it starts rising rapidly and then, at even higher levels of input, it saturates (see Figure 2). The characteristic of a sigmoid function is differentiable everywhere. The sigmoid function most frequently used for ANN is the logistic sigmoid function:

$$y_j = \frac{1}{1+e^{-x_j}} \tag{2}$$

The nonlinear nature of this sigmoid transfer function plays an important role in the performance of the ANN. Other functions can be used as long as they are continuous and possess a derivative at all points.

The output of ANN model from equation 2 is always between 0 and 1. Before training, the weight w_{ji} is initialized by random number. In this study, training of neural network is performed using back-propagation algorithms (Rumelhart et al., 1986).

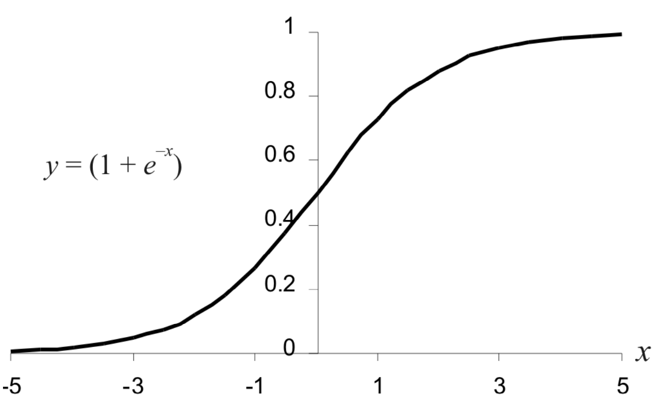

Figure 2: The sigmoid logistic activation function.

Back-propagation

The back-propagation algorithm is a gradient descent procedure used to minimize an objective function (error function) E. When the calculated outputs are carried out, next step is to calculate the difference or error between calculated outputs and desired (target) output. If the overall error value drops below some pre-determined threshold, then model is completed. If not, error back propagation, one of procedures used to adjust weights, is begun. It means that the error value is then propagated backwards through the network, and small changes are made to the weights in each layer. The sum-squared value is given by:

$$E = 0.5\sum_{p=1}^{np}\sum_{j}^{nj}\left(t_{pj} - y_{pj}\right)^2 \tag{3}$$

where p is an index of training pattern, j is an index of output neuron, t_{pj} is the target value of the j-th of the outputs for the pattern p, y_{pj} is the output of the j-th neuron of the actual output pattern produced by the input pattern p. The goal of the training process is to minimize this sum-squared error in overall training pattern. The weight update for the tth epoch between node i and j are generally given by:

$$w_{ij}(t) = w_{ij}(t-1) - \eta\frac{\partial E}{\partial w_{ij}} + \alpha \quad \Delta w_{ij}\left(t-1\right) \tag{4}$$

where $\partial E/\partial w_{ij}$ is a gradient of error with respect to weight, the η and α are called learning rate and momentum respectively.

Case Study

To investigate the ANN as a robust method for solving non-linear problem such as groundwater level, in this paper the BPANN simulates groundwater level fluctuation (GLF) of Jakarta, Indonesia (see Figure 3). Jakarta is located in the groundwater basin which is called the Jakarta groundwater basin. The bottom of the basin system is formed by impermeable Miocene sediment which also crop out at the southern boundary of the system. The basin fill consists of marine Pliocene and Quarternary sand and delta sediment up to 300 m thick. Thickness of sandy aquifer layer is about 1–5 m interconnected with a predominantly silty/clayey sequence and comprise only 20% of total sediment deposit. Fine sand and silt is very frequent component of aquifers (Tirtomiharja, 1996). The groundwater contribution to the actual supply is about 250 million m^3 $year^{-1}$ and is mainly abstracted from innumerable shallow wells (80%) and more than 3000 deep wells (20%). Between 1900 and 1950, groundwater abstraction was below 10 million m^3 $year^{-1}$ but since that time, mainly

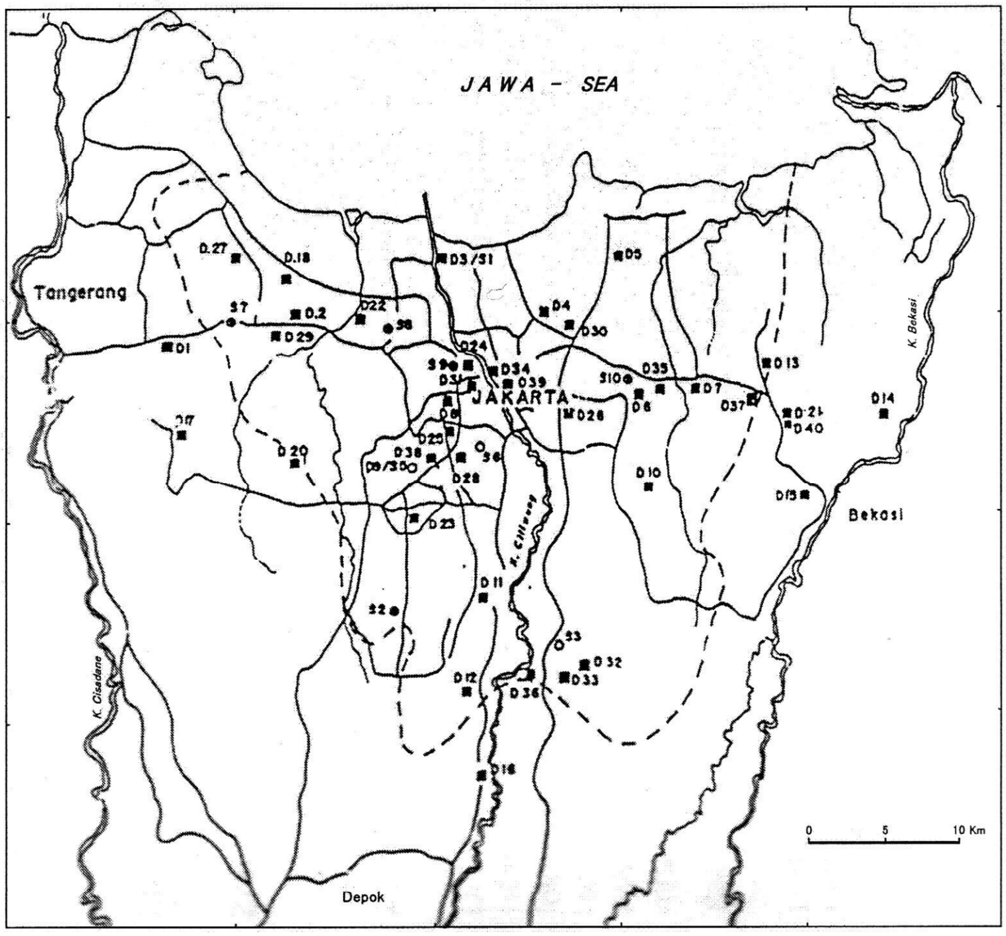

Figure 3: Map of observation wells in Jakarta, Indonesia.

after 1970, it has steadily increased in step with the growth in population and industrial development. In the year 1994, deep groundwater abstraction was estimated to be 53 million m^3 $year^{-1}$ which was about 50% higher than could be accounted for by registered wells of 33.8 million m^3 $year^{-1}$ (Soetrisno, 1999).

Representative data sets are necessary for training phase in which the ANN learns the model's input-output relations (Kamp and Savenije, 2006). The main reason for this is that ANN belongs to the class of data driven approaches (Maier and Dandy, 2000).

However, data is very limited due to the policy or budget of government. A five-day groundwater level fluctuation was simulated. The input variable was using five time lag of well interest (see Table 1). Length of training pattern is 365 or 5-year data (1989-1993) and 146 or 2-year data (1994-1995) as a testing pattern. The aim of using time lag is not only to overcome limitation of data but also to investigate the effect of time lag to the model. The optimal numbers of hidden neurons were determined by trial and error based on a suggestion that over-fitting does not occur if the number of training samples is at least 30 times the number of free parameters or free weights (Maier and Dandy, 2000; Amari et al., 1997).

Table 1: Input and output variables to the ANN models

Set	*Input variables*	*Output variable*
1	Water level at $(t-1)$, $(t-2)$, $(t-3)$, $(t-6)$, $(t-12)$	Water level at (t)
2	Water level at $(t - 2)$, $(t-3)$, $(t-6)$, $(t-12)$, $(t-18)$	Water level at (t)
3	Water level at $(t-3)$, $(t-6)$, $(t - 12)$, $(t-18)$, $(t-36)$	Water level at (t)

It is important to determine the appropriate network architecture in order to obtain satisfactory results. For this study, we tried two kinds of hidden nodes: 3 and 5 hidden nodes respectively for all types and models input. A number of trial and error methods were performed for back-propagation algorithms with learning rate of 0.1 and momentum of 0.5 and 0.9. The activation function used is the sigmoid logistic function. Target data is scaled into range of 0.1 to 0.9 values by normalizing with respect to minimum and maximum data before feeding into network. For convenience, input data is also scaled into value in range of 0.1–0.9.

Normally, the data set for ANN needs to be divided into three parts. The first part is for the training, the second part for validation and the third part for testing phase. However, the length of data is not so big, only two parts are considered in this study, namely training and testing phase. The only difference between a testing and a validation is that if the error of the validation increases the training stops. In this study, these two terms are used synonymously.

Result and Discussion

The BPANN was performed to forecast groundwater level fluctuation (GLF) for four observation wells in the area of study. After trial and error method applying to network model, we found that input variable set one and three hidden nodes were suitable for network model of GLF forecasting. It means that input variable with time lag $(t-1)$ to $(t-12)$ or until two months before data pattern has significant influence to forecast GLF at time (t).

The Efficiency Index (R-squared) and root mean square error (RMSE) were investigated as a major performance indicator to evaluate the relationship between calculated (output model) and desired (target). The R-squared for the identified models for five years training phase (1989-1993) and two years testing phase (1994-1995) for all three sets input variable are presented in Table 2. In general, the calculation results were relatively good. Note that observation wells of J3 and J4 consist testing sample outside range of training pattern. Typically, poor prediction occurs when the validation or testing phase contain values outside that used for training phase (Maier and Dandy, 2000). It means that the training and validation or testing phase are representative of the same population. To overcome the problems, in this study, the testing sample which contain outside value was added to training sample and hence trained with whole sample (Maier and Dandy, 2000). Therefore, the prediction result of wells J3 and J4 were satisfied. It is clear that the BPANN can achieve better performance for testing phase if the training process can learn or generalize the sample pattern in training phase itself and testing phase.

Table 2: The best performance of BPANN model in term of R-squared (efficiency index)

Input variables, sets	*Training phase*	*Testing phase*
	Well J1	
1	0.9467	0.9722
2	0.8416	0.9159
3	0.8281	0.8264
	Well J2	
1	0.9914	0.9609
2	0.9825	0.8682
3	0.9607	0.7853
	Well J3	
1	0.9652	0.9525
2	0.9563	0.9474
3	0.9396	0.9155
	Well J4	
1	0.9802	0.9625
2	0.9564	0.9140
3	0.9556	0.8795

Table 3 provides performance BPANN model in terms of RMSE. It can be seen that case one gives relatively better prediction result for all observation wells studied.

Table 3: The best performance of BPANN model in term of RMSE (m)

Input variables, sets	*Training phase*	*Testing phase*
	Well J1	
1	0.2383	0.1629
2	0.4109	0.2832
3	0.4281	0.4069
	Well J2	
1	0.2535	0.2788
2	0.3616	0.5119
3	0.5419	0.6534
	Well J3	
1	0.2409	0.2643
2	0.2699	0.2779
3	0.3174	0.3524
	Well J4	
1	0.2109	0.1765
2	0.3127	0.2674
3	0.3172	0.3216

The RMSE increase for cases two and three. It infers that input samples containing t-18 and t-36 or three months and six months before data have no significant influences to the model. For input set one, the maximum RMSE for testing phase is about 0.278 m of well J2. The RMSE of low value is corresponding to high value of R-squared.

A graphical view of BPANN model for testing phase of GLF is shown in Figure 4. The prediction result is close to observed GLF. The accuracy seems very high. For four GLF of input set one, the calculated model the RMSE lies between 0.16 and 0.28 m (Table 3). In general,

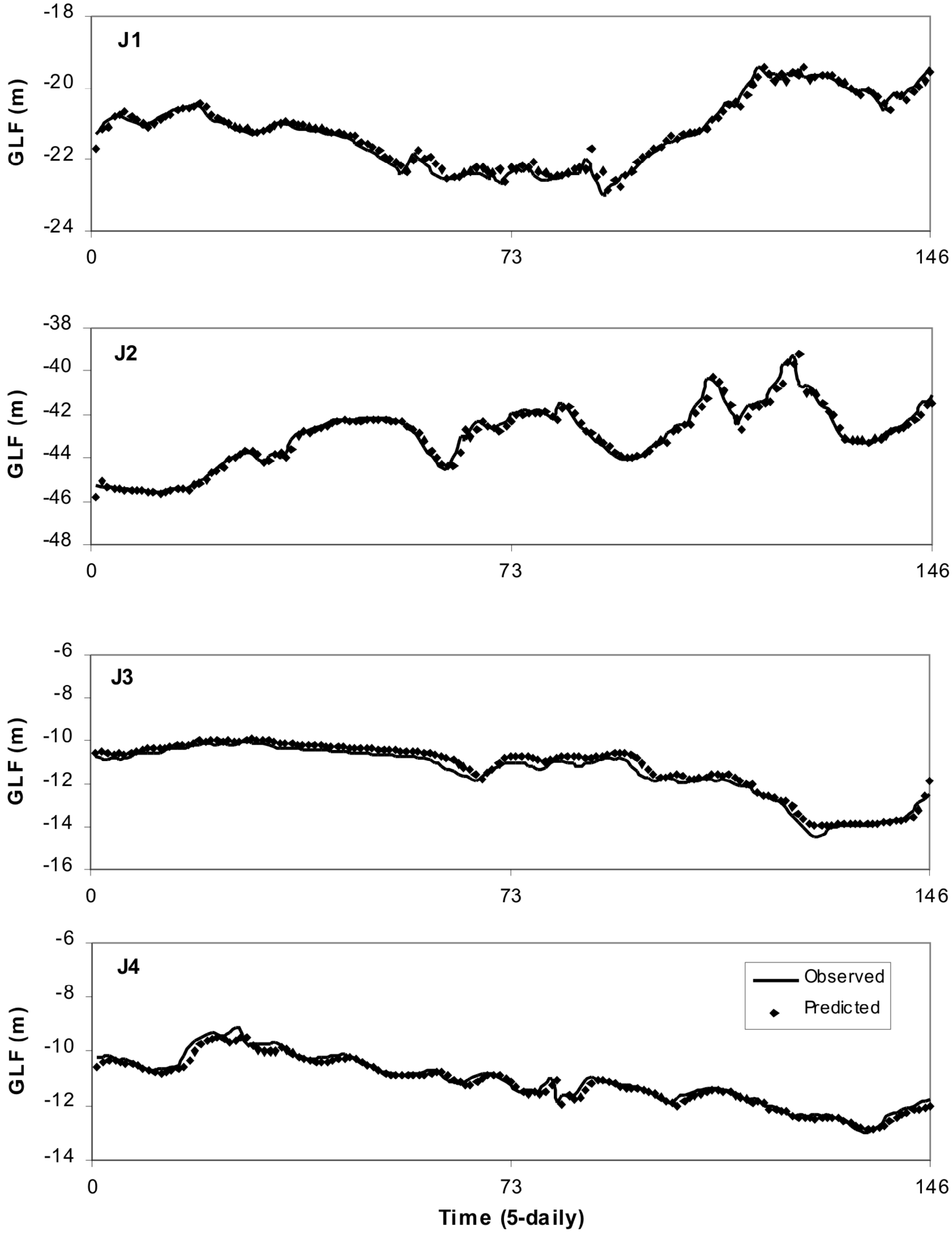

Figure 4: Result of prediction and observed groundwater level fluctuation of four observation wells for two years testing phase.

there are no extraordinary cases during testing phase because there is no significant difference between observation and prediction.

Figure 5 provides a scatter-plot of predicted against observed GLF for input set one of four observation wells for testing phase. The observed and predicted values were in close proximity to the line of best fit (diagonal line). The highest efficiency index (*R*-squared) for these scatter-plot had value 0.9722 of well J1 with RMSE of 0.1629 m. The results indicate that the BPANN was successful in learning the relationship between the input and output variables.

Figure 6 shows the prediction deviation from the observed groundwater level in testing phase. The result

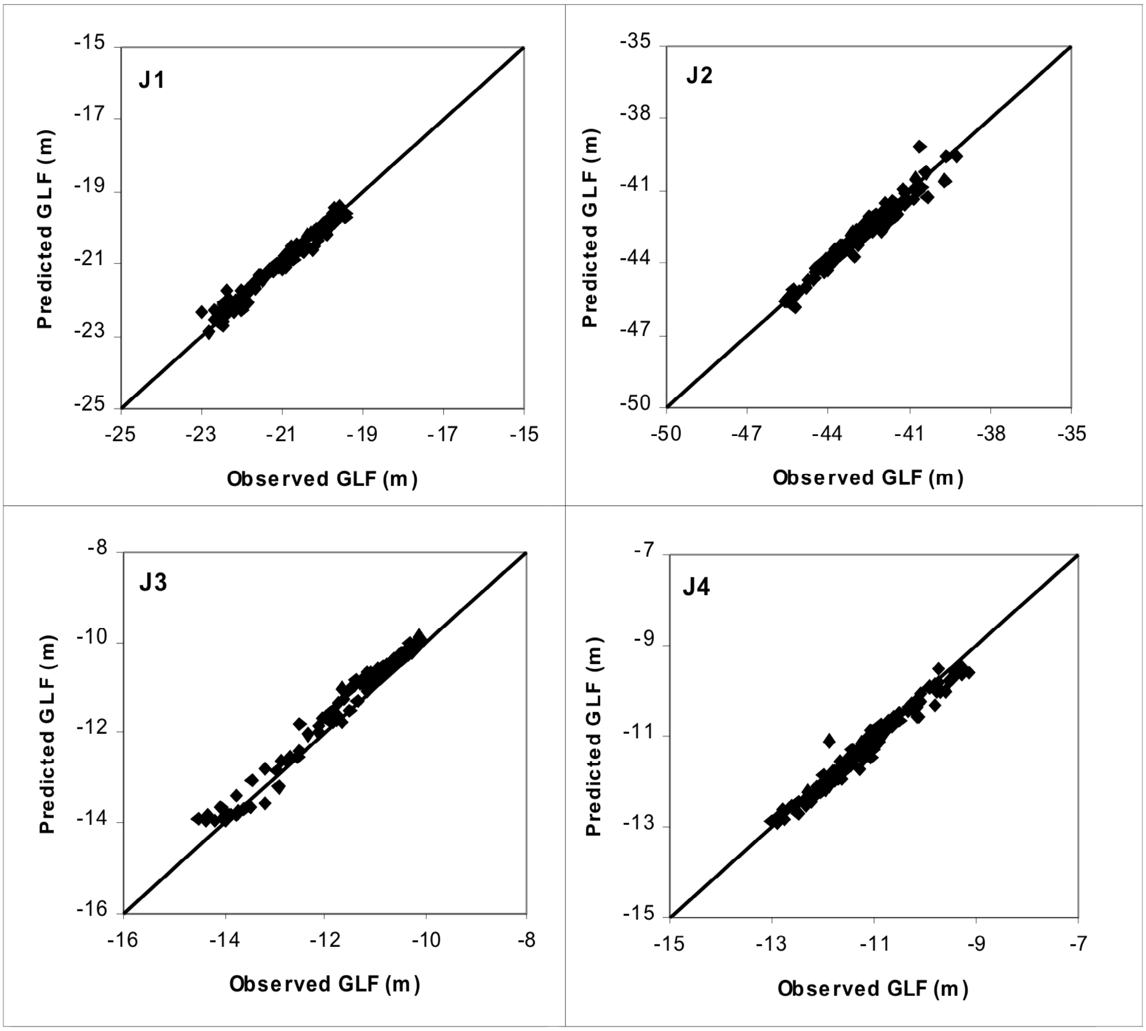

Figure 5: Scatter-plot of prediction and observed GLF for four observation wells in testing phase.

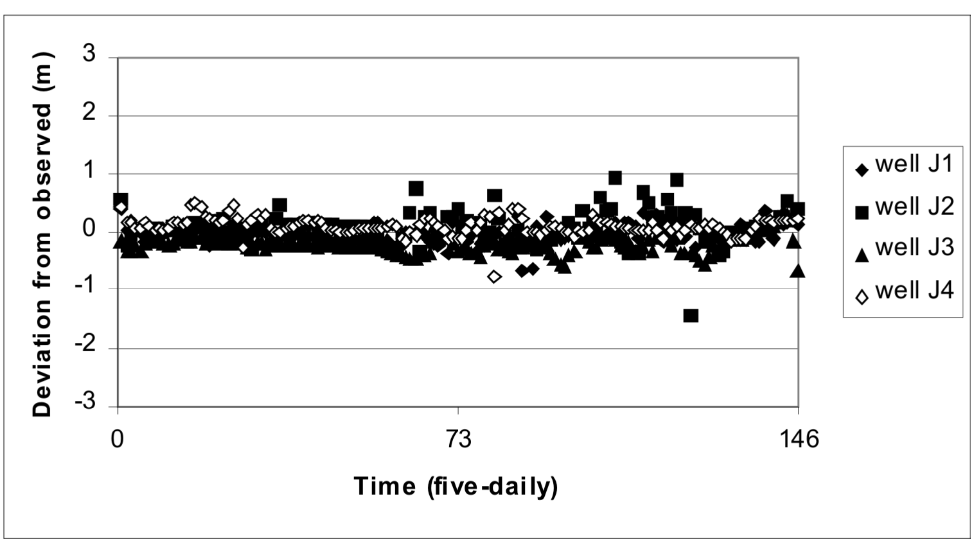

Figure 6: The prediction deviations from observed GLF for four observation wells using input set case one in testing phase.

of BPANN model for all observation wells was satisfied. The deviations generally were in between ±1 m, except one point of well J2 is about 1.45 m. The RMSE was in the range of 0.18 to 0.28 m. Positive deviation values indicate that the model over-prediction the groundwater level, whereas negative deviation values denote under-prediction (Coulibaly et al., 2001). Wells J1 and J3 were under-prediction with 94% negative deviation of well J3, whereas wells J2 and J4 were over-prediction with 79% positive deviation of well J4.

The results obviously indicate that time lags as an input can be used to forecast groundwater level fluctuation. This advantage is very useful for prediction with very limited observation data, especially for region or country, which only has several observation wells because of limited budget or policy such as Indonesia. In addition, the simulation result can be used for real time prediction and for management purposes. For real time prediction, we can use the model to manage groundwater fluctuation or to identify the optimal pumping rate. Furthermore, the groundwater equilibrium system can be well maintained.

Conclusion

This paper deals with the prediction of groundwater level fluctuation at several points using time lag as input set of artificial neural network. The applicability of BPANN for the prediction of groundwater level fluctuation was evaluated in three types of time lag as input variable. Under these conditions, the BPANN formalism has an advantage in being able to learn and generalize from examples without knowledge of rules and the successful prediction depends on the availability of good quality data recorded.

The result from BPANN model in general indicates that ANN is an effective tool for five-daily groundwater level prediction. From the simulation results it can be inferred that the BPANN model can be used as a tool for groundwater level fluctuation modelling where even the available data are relatively short. For more satisfied result, especially for shallow groundwater, the model can be improved by considering some information such as water withdrawal, stream discharge and rainfall. Moreover, to explore the BPANN method, different algorithm such as a Levenberg-Marquardt can be applied.

Acknowledgements

This research is supported by Technological and Professional Skills Development Sector Project (TPSDP), Directorate General of Higher Education, Ministry of National Education, the Republic of Indonesia. The authors are grateful to Mining Department, Government of Jakarta Province for providing the valuable data and giving us permission for publication. We would like to express sincere thanks to Mr. M. Harris Pindratno for valuable discussion during the early phases of this work.

References

Amari, S., Murata, N., Müller, K.R., Finke, M. and H. Yang (1997). Asymptotic statistical theory of overtraining and cross-validation. *IEEE Transactions on Neural Network*, **8(5):** 985-996.

Aziz, A.R.A. and K.F.V. Wong (1992). Neural network approach to the determination of aquifer parameter. *Ground Water*, **30(2):** 164-166.

Balkhair, K.S. (2002). Aquifer parameters determination for large diameter wells using neural network approach. *J. Hydrol.*, **265:** 118-128.

Coulibaly, P., Anctil, F., Aravena, R. and B. Bobée (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, **37(4):** 885-896.

Daliakopoulos, L.N., Coulibaly, P. and I.K. Tsanis (2005). Groundwater level forecasting using artificial neural networks. *J. Hydrol.*, **309:** 229-240.

Govindaraju, R.S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. By the ASCE task Committee on application of artificial neural networks in hydrology. *J. Hydrol. Eng*. **5:** 115-123.

Kamp, R.G. and H.H.G. Savenije (2006). Optimising training data for ANNs with genetic algorithms. *Hydrol. Earth Syst. Sci. Discuss.*, **3:** 285-297.

Lallahem, S., Mania, J., Hani, A. and Y. Najjar (2005). On the use of neural networks to evaluate groundwater levels in fractured media. *J. of Hydrol.*, **307:** 92-111.

Maier, H.R. and G.C. Dandy (2000). Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environmental Modelling & Software*, **15:** 101-124.

McIntosh A.C. (2003). Asian water supplies: Reaching the urban poor. A guide and sourcebook on urban water supplies in Asia for governments, utilities, consultants, development agencies, and nongovernment organizations. Asian Development Bank and International Water Association, 137 p.

Nayak, P.C., Satyaji Rao, Y.R. and K.P. Sudheer (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, **20:** 77-90.

Rajanayaka, C., Samarasinghe, S. and D. Kulasiri (2002). Solving the Inverse Problem in Stochastic Groundwater Modelling with Artificial Neural Networks. Proc. iEMSs 2002 sessions (part two), 154-159. Available at: *http://www.iemss.org/iemss2002/proceedings/pdf/volume%20due/385_rajanayaka.pdf*

Ranjithan, S., Eheart, J.W. and J.H. Garrett (1993). Neural network-based screening for groundwater reclamation under uncertainty. Water Resources Research, **29(3):** 563-574.

Rumelhart, D., Hinton, G.E. and R. Williams (1986). Learning representations by back-propagation errors. *Nature*, **323:** 533-536.

Soetrisno, S. (1999). Groundwater management problem: A Jakarta and Bandung (Indonesia) comparative city case studies. Available at: *http://www.geocities.com/Eureka/Gold/1577/index.html.*

Tirtomiharjo, H. (1996). Urban groundwater database. Directorate of Environmental Geology of Indonesia. Available at: http:*//www.scar.utoronto.ca/~gwater/IAHCGUA/UGD/jakarta.html.*