European Water 57: 361-366, 2017.

Artificial intelligence and regression analysis in predicting ground water levels in public administration

G.N. Kouziokas[1*], A. Chatzigeorgiou[2] and K. Perakis[1]
[1] *University of Thessaly, School of Engineering, Department of Planning and Regional Development, Volos, Greece*
[2] *University of Macedonia, Department of Applied Informatics, Thessaloniki, Greece*
[*] *e-mail: gekouzio@uth.gr*

Abstract: Water level prediction of ground water can be considered as a very important tool in water resources management. This research implements artificial neural network models in order to build the optimal forecasting models for predicting the water levels and regression analysis in order to evaluate the prediction accuracy. In this research, an Artificial Neural Network Perceptron is applied in order to construct forecasting models for predicting water levels of ground water. The developed models are then compared each other in order to find the optimal one according the best performance that will lead to the most accurate predictions. Several topologies were tested in order to discover the best forecasting model. The different predictive models were constructed by implementing different number of the nodes in the hidden layers, also by testing different number of the hidden layers. The results showed an increased prediction accuracy of the developed Artificial Neural Network Models. This research is aiming at providing the scientists and engineers with the optimal prediction models in order to be used for forecasting the water levels of ground water with increased accuracy compared to other prediction techniques and methods. The developed forecasting models can provide accurate predictions of water levels of ground water which can be very valuable for the public administrators and stakeholders in water resources management.

Key words: Artificial intelligence; Neural networks; Environmental management; Public Administration; Water level prediction

1. INTRODUCTION

Groundwater levels monitoring is of high importance in public administration and in environmental management and planning, since water is highly connected with the quality of human life and the health of the natural environment (Jacobs and Holway, 2004; Kemper, 2004).

Groundwater levels forecasting can play a significant role in public management and planning since the authorities will have the opportunity to take the appropriate measures in order to prevent the rapid decrease of the groundwater levels and to protect the public health (Davies et al. 2001; Howard et al., 2006; Schmoll, 2006).

Information technology and especially artificial intelligence have been applied in various fields of environmental management and public administration the last years (Kouziokas, 2016a, b, c, d; Metaxiotis et al., 2003). Artificial neural networks have been also used as a predictive method by some researchers in order to forecast groundwater levels (Adamowski and Chan, 2011; Daliakopoulos et al., 2005; Giustolisi and Simeone, 2006; Sreekanth et al., 2009).

Giustolisi and Simeone (2006) have studied groundwater level predictions by using artificial neural networks and a multi-objective strategy. The results have showed satisfactory prediction results by using a linear model in short time forecasting and a nonlinear model in distant time predictions.

Adamowski and Chan (2011) have used artificial neural networks and coupling discrete wavelet transforms for predicting groundwater levels. The developed prediction models have provided very accurate groundwater level forecasts in a monthly basis.

Sreekanth et al. (2009) have used artificial neural networks to develop prediction models to predict groundwater levels in the Maheshwaram watershed. The results have showed better prediction results than the conventional forecasting methods.

In this research, artificial intelligence and regression are used for forecasting groundwater levels. The levels of groundwater were chosen to be predicted, since groundwater resources are associated with public health and are considered of high importance. Artificial neural network models are developed in order to build the optimal forecasting model. The developed forecasting models can provide predictions of groundwater levels which can be used by the public administrators in water resources management and planning. The research methodology, the discussion and the results are described in the next sections.

2. THEORETICAL BACKGROUND

2.1 Artificial neural networks

Artificial Neural Networks (ANNs) were utilized in this research in order to construct and compare forecasting models regarding groundwater predictions. Artificial Neural Networks are computing systems that their construction simulates the neural structure of the human brain (Basheer and Hajmeer, 2000; Suykens et al., 2012). Artificial Neural Networks process the input data and the information traverses the neural network connections so as to produce the output values according to the input. Their advantage is that they can be used also at non-linear relationships between the input and the output (Almeida, 2002; Poznyak et al., 1999; Zhang et al., 1998).

2.2 Feed Forward Multilayer Perceptron

In this research, a Feedforward Multilayer Perceptron (FFMLP) was used in order to construct the artificial neural network prediction models. FFMLP is implemented in this study, since several studies have showed that Feed Forward Multilayer Perceptron is the most suitable regarding problems that have to deal with time series forecasting (Kouziokas et al., 2016a, 2016b; Tang and Fishwick, 1993; Zhang et al., 2001). In a Feed Forward Multilayer Neural Network, the neurons are connected only forward. The neural network layers have connections to the next layers, but they do not have any connections backwards (Blum and Li, 1991; Hornik, 1991; Hornik et al., 1989). Figure 1 shows the topology of a typical feed forward neural network, where n is the number of neurons in the input layer, k is the number of neurons in the hidden layer and o is the number of neurons in the output layer. I_1, I_2, …,In are the input values and O_1, O_2, .., On are the output (predicted) values.

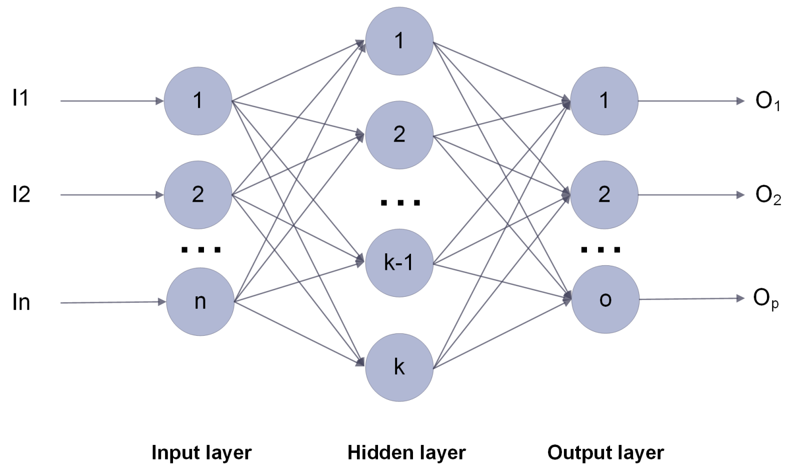

Figure 1. A typical feedforward neural network. n is the number of neurons in the input layer, k is the number of neurons in the hidden layer and o is the number of neurons in the output layer.

2.3 Levenberg Marquardt algorithm

The Levenberg Marquardt algorithm was selected as a training algorithm for the Feed Forward Multilayer Neural Network, since it produced better results than the other training algorithms.

The Levenberg Marquardt algorithm is considered as one of the fastest learning algorithms compared to other algorithms (Marquardt, 1963). In this research, the Levenberg Marquardt algorithm yield better outputs than the most common algorithms that were tested (Resilient Backpropagation, BFGS Quasi-Newton, Scaled Conjugate Gradient) by using different neural network topologies.

The Levenberg Marquardt algorithm combines the steepest descent algorithm and the Gauss-Newton algorithm. The Levenberg Marquardt algorithm can be used for solving non-linear least-squares problems (Liu, 2010; Lourakis, 2005).

3. RESEARCH METHODOLOGY

The research methodology that was followed, consists of five stages: data collection and preparation, neural network prediction models development, comparison of the developed neural network models so as to discover the optimum one according to the performance, testing the optimum developed neural network model in groundwater forecasting and finally assessing the forecasting results by using RMSE error and linear regression.

Firstly, groundwater and climate data were collected and prepared so as to feed the artificial neural network models. In the second stage, artificial neural network prediction models were developed by testing different topologies, learning algorithms and transfer functions in order to construct the optimal neural network forecasting model. In the next phase, a comparison of the constructed neural network models was performed, in order to find the optimum model according to the performance results. In the next stage, optimum developed neural network model was tested in groundwater forecasting. In the last stage, the forecasting results were evaluated by using Root Mean Squared Error (RMSE) and linear regression. Figure 2 shows an overview of the stages of the followed research methodology.

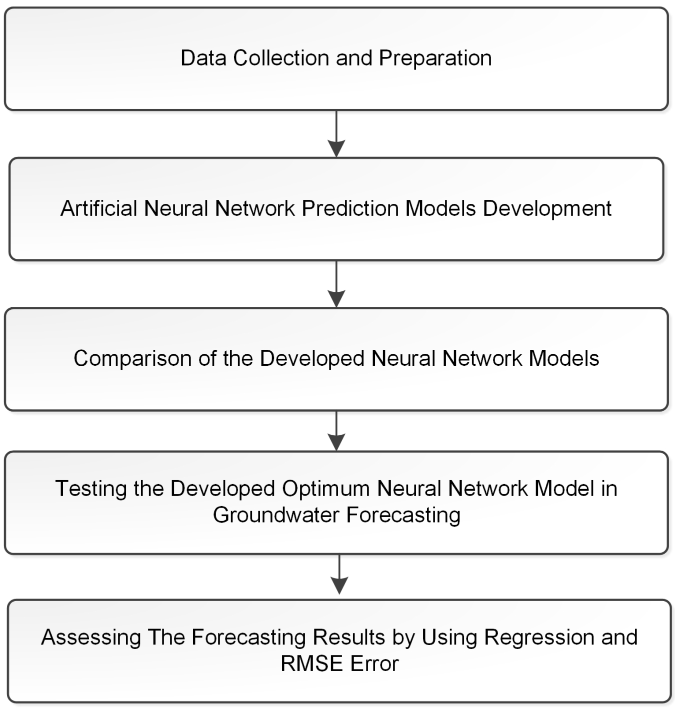

Figure 2. Overview of the research methodology.

4. RESULTS

4.1 Data collection and preparation

Groundwater data were retrieved from the official website of the United States Geological Survey (USGS) for the Montgomery country in Pennsylvania (https://www.usgs.gov). The climate data were retrieved from the official website of Pennsylvania State Climatologist office. Data regarding the daily values of the average temperature, the average relative humidity and the precipitation were pre-processed in order to be used as inputs in the developed Artificial Neural Network Models.

Groundwater data were collected for the Montgomery County Observation Well in Pennsylvania of USA. The data were collected and prepared (checked for incoherencies, duplicates, etc.) for the 365 days of the year 2014.

Figure 3 shows the values of the mean daily groundwater level (feet below the land surface) for the year 2014 for the Montgomery County Observation Well.

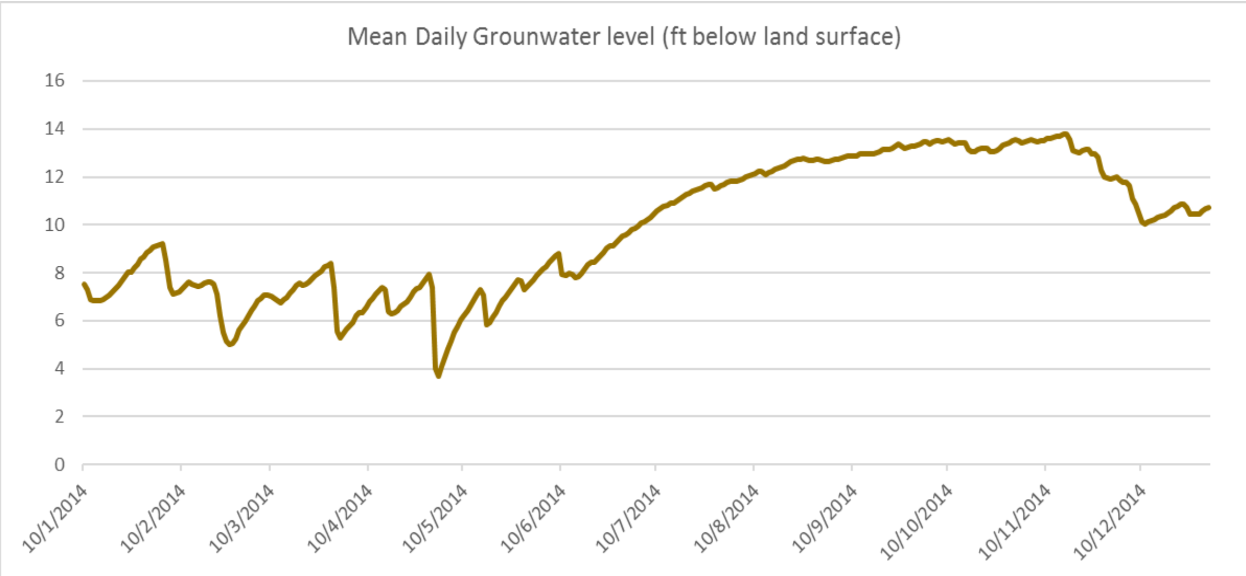

Figure 3. Daily groundwater level (feet below the land surface).

4.2 Artificial neural network models

The artificial neural network models were developed by using as input parameters several factors that influence the groundwater levels such as Average Temperature (AT), Average Relative Humidity (ARH), precipitation (PRE) (includes drizzle and rainfall) and the depth to groundwater level (DGL) below land surface.

The data was divided into three different parts. 70% of the primary data was used as the training set, 15% for the validation set and 15% for the test set. The training data set was used in order to train the neural network with historical data of the input factors. The validation set was used so as to assess the performance of the developed artificial neural network models.

4.3 Optimum neural network model implementation

The optimum forecasting model was developed by comparing the performance of the different neural network models that were investigated. Multiple neural network topologies were tested regarding the number of the hidden layers (one to two hidden layers), the number of the neurons of every hidden layer (one to fifty neurons) and the most common transfer functions for the hidden layers and also the most common learning algorithms.

The most common transfer functions that were tested for the hidden layers are the following: Tanh-Sigmoid Transfer Function (TSTF), Log-Sigmoid Transfer Function (LSTF) and Linear Transfer Function (LTF).

The most common training algorithms that were tested for training every neural network model were: Levenberg Marquardt (LM), Resilient Backpropagation (RB), Scaled Conjugate Gradient (SCG), BFGS Quasi-Newton (BFGS-QN).

The optimal neural network model was the one where the training algorithm was Levenberg Marquardt (LM) algorithm. The optimum model was evaluated according to the minimum Root Mean Squared Error (RMSE) among all the other constructed models. The Root Mean Squared Error (RMSE) of the optimal model was found to be 0.0742.

According to the results, the topology of the optimal neural network model consists of 22 neurons in the first hidden layer and 25 neurons in the second hidden layer. The transfer function of the first hidden layer is Tanh-Sigmoid Transfer Function (TSTF) and the transfer function of the second hidden layer is Tanh-Sigmoid Transfer Function (TSTF).

The linear regression analysis was utilized to evaluate the forecasting accuracy of the optimal artificial neural network model. The R value was used which shows the linear relationship between the outputs and the targets. If the R value is one, then it shows that there is an exact linear relationship between the predicted values (outputs) of the network and the target values (inputs) of groundwater levels.

The R value of the predicted and the actual groundwater levels was 0.99775. This R value shows that the predictions of the groundwater levels produced by the optimal forecasting model are very precise for the selected region.

5. CONCLUSIONS AND DISCUSSION

This research utilizes artificial intelligence and regression analysis in forecasting groundwater levels which is very vital in public administration, in natural resources management and in water management. In this paper, several artificial neural network models were developed by testing several parameters regarding the topology of the neural network and also the training algorithms and the transfer functions of the hidden layers.

The optimal artificial neural network model was developed by taking into consideration several factors that affect the groundwater levels. These parameters were used as input values in order to forecast the groundwater levels. Also, several neural network topologies were tested in order to find the optimum one according to their performance.

Compared to other researches this study have used artificial neural networks in combination with regression analysis by utilizing different input variables and a wider range of learning algorithms and topologies of ANNs by investigating also several factors that improved the performance of the optimal forecasting model.

The results showed a very precise prediction accuracy of the groundwater levels compared to other researches that have used artificial neural network for groundwater predictions (Adamowski and Chan, 2011; Giustolisi and Simeone, 2006; Sreekanth et al., 2009).

The developed optimal neural network forecasting model and the groundwater predictions can be very useful to the authorities in public and water management and also in taking proactive measures so as to apply effective management strategies in order to minimize the rapid losses of the amount of the water resources.

This research could contribute to the prevention of natural water loss which is very vital and could have negative impact on human health by adopting the appropriate water management strategies and is aiming at providing scientists and engineers with prediction models in order to be used for forecasting the water levels of groundwater with increased prediction accuracy.

ACKNOWLEDGEMENTS

United States Geological Survey (USGS) website for retrieving the groundwater data.

REFERENCES

Adamowski, J., Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology, 407(1), 28-40.

Almeida, J. S. (2002). Predictive non-linear modeling of complex data by artificial neural networks. Current opinion in Biotechnology, 13(1), 72-76.

Basheer, I. A., Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43(1), 3-31.

Blum, E. K., Li, L. K. (1991). Approximation theory and feedforward networks. Neural Networks, 4(4), 511-515.

Daliakopoulos, I. N., Coulibaly, P., Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. Journal of Hydrology, 309(1), 229-240.

Davies, H., Rosen, M. R., White, P. A. (2001). Groundwater and health. Groundwaters of New Zealand, 253-272.

Giustolisi, O., Simeone, V. (2006). Optimal design of artificial neural networks by a multi-objective strategy: groundwater level predictions. Hydrological Sciences Journal, 51(3), 502-523.

Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. Neural networks, 4(2), 251-257.

Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359-366.

Howard, G., Bartram, J., Pedley, S., Schmoll, O., Chorus, I., Berger, P. (2006). Groundwater and public health. Protecting groundwater for health: managing the quality of drinking-water sources. London, International Water Association Publishing, 3-19.

Jacobs, K. L., Holway, J. M. (2004). Managing for sustainability in an arid climate: lessons learned from 20 years of groundwater management in Arizona, USA. Hydrogeology Journal, 12(1), 52-65.

Kemper, K. (2004). Groundwater—from development to management. Hydrogeology Journal, 12(1), 3-5.

Kouziokas, G. N. (2016a). Artificial intelligence and crime prediction in public management of transportation safety in urban environment. Proceedings of the 3rd Conference on Sustainable Urban Mobility (pp. 534-539). Volos: University of Thessaly Greece.

Kouziokas, G. N. (2016b). Geospatial Based Information System Development in Public Administration for Sustainable Development and Planning in Urban Environment. European Journal of Sustainable Development, 5(4), 347-352.

Kouziokas, G. N. (2016c). Technology-based management of environmental organizations using an Environmental Management Information System (EMIS): Design and development. Environmental Technology and Innovation, 5, 106-116. doi:http://dx.doi.org/10.1016/j.eti.2016.01.006.

Kouziokas, G. N. (2016d, September). An Information System for Monitoring Environmental Indicators in Public Management for Sustainable Development. Paper presented at the 11th International Scientific Conference eRA-11, Piraeus University of Applied Sciences, Piraeus.

Kouziokas, G. N., Chatzigeorgiou, A., Perakis, K. (2016a, June). Applying Levenberg Marquardt Algorithmin Feedforward Neural Network Models for Predicting Crime in Public Management. Paper presented at the International Conference on Computational and Informational Sciences and Engineering, University of Thessaly, Portaria, Volos.

Kouziokas, G. N., Chatzigeorgiou, A., Perakis, K. (2016b, September). Predicting Environmental Data in Public Management by Using Artificial Intelligence. Paper presented at the 11th International Scientific Conference eRA-11, Piraeus University of Applied Sciences, Piraeus, Greece.

Liu, H. (2010, August). On the levenberg-marquardt training method for feed-forward neural networks. In 2010 Sixth International Conference on Natural Computation (Vol. 1, pp. 456-460). IEEE.

Lourakis, M. I. (2005). A brief description of the Levenberg-Marquardt algorithm implemented by levmar. Foundation of Research and Technology, 4, 1-6.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial and Applied Mathematics, 11(2), 431-441.

Metaxiotis, K., Ergazakis, K., Samouilidis, E., Psarras, J. (2003). Decision support through knowledge management: the role of the artificial intelligence. Information Management and Computer Security, 11(5), 216-221.

Poznyak, A. S., Yu, W., Sanchez, E. N., Perez, J. P. (1999). Nonlinear adaptive trajectory tracking using dynamic neural networks. IEEE Transactions on Neural Networks, 10(6), 1402-1411.

Sreekanth, P. D., Geethanjali, N., Sreedevi, P. D., Ahmed, S., Kumar, N. R., Jayanthi, P. K. (2009). Forecasting groundwater level using artificial neural networks. Current Science, 96(7), 933-939.

Schmoll, O. (2006). Protecting groundwater for health: managing the quality of drinking-water sources. World Health Organization.

Suykens, J. A., Vandewalle, J. P., de Moor, B. L. (2012). Artificial neural networks for modelling and control of non-linear systems. Springer Science & Business Media.

Tang, Z., Fishwick, P. A. (1993). Feedforward neural nets as models for time series forecasting. ORSA Journal on Computing, 5(4), 374-385.

Zhang, G. P., Patuwo, B. E., Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35-62.

Zhang, G. P., Patuwo, B. E., Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. Computers and Operations Research, 28(4), 381-396.