The final publication is available at
https://link.springer.com/article/10.1007/s11269-018-2126-y

Multilayer Feed Forward Models in Groundwater Level Forecasting Using Meteorological Data in Public Management

Georgios N. Kouziokas[*1], Alexander Chatzigeorgiou[2], Konstantinos Perakis[1]

Abstract
Managing the groundwater resources is very vital for human life. This research proposes a methodology for predicting the groundwater levels which can be very valuable in water resources management. This study investigates the application of multilayer feed forward network models for forecasting the groundwater values in the region of Montgomery country in Pennsylvania. Multiple training algorithms and network structures were investigated to develop the best model in order to forecast the groundwater levels. Several multilayer feed forward models were created in order to be tested for their performance by changing the network topology parameters so as to find the optimal prediction model. The forecasting models were developed by applying different structures regarding the number of the neurons in every hidden layer and the number of the hidden network layers. The final results have shown a very good forecasting accuracy of the predicted groundwater levels. This research can be very valuable in water resources and environmental management.

Keywords Artificial intelligence, Neural networks, Environmental management, Public management, Water level prediction, Water resources management

1 Introduction

The protection and management of the natural resources has a great value for the human life. Considering the climate change and the increasing numbers of droughts, a sustainable groundwater level management can play an important role in preventing Groundwater Depletion and in reserving the health of the natural habitats (Biswas 2004; Dean et al. 2018; Langridge and Daniels 2017; Thomas and Gibbons 2018).

*Corresponding author: Georgios N. Kouziokas, E-mail: gekouzio@uth.gr

[1] School of Engineering, University of Thessaly, Pedion Areos, 383 34 Volos, Greece

[2] Department of Applied Informatics, University of Macedonia, 156 Egnatia Str, 54006 Thessaloniki, Greece

Predicting the levels of ground water could play an important role by leading to more sustainable solutions and water supply management practices since the stakeholders can make plans, apply strategies and take proactive measures towards public health protection when groundwater levels is predicted to be decreased below the normal levels (Famiglietti 2014; Howard et al. 2006; MacKay, 2006).

The application of Information and Communication technology (ICT) have been enormously increased in environmental and water resources management (Kouziokas 2016b, 2016c; Mysiak et al. 2005). Furthermore, technology based techniques such as artificial neural networks have been applied by several researchers in environmental and water resources management the last decades (Adamowski and Chan 2011; Giustolisi and Simeone 2006; Kouziokas 2016a, 2017a, 2017b; Kouziokas and Perakis 2017).

The implementation of artificial neural networks has been mainly focused on solving forecasting-based problems. The main advantage of a neural network-based methodology is that it can model also nonlinear and very complex relationships between input variables and output-forecasting values in many kinds of time series forecasting cases (Zhang et al. 1998).

Artificial neural network based techniques have been implemented by several researchers as a sophisticated tool in order to forecast groundwater level values in various study areas (Adamowski and Chan 2011; Babu et al. 2016; Barzegar et al. 2017; Chang et al. 2016; Giustolisi and Simeone 2006; Guzman et al. 2017; Hong 2017; Khalil et al. 2015; Mohanty et al. 2015; Supreetha et al. 2015; Tapoglou et al. 2014).

Adamowski and Chan (2011), have developed coupling discrete wavelet transforms based on artificial neural networks for forecasting the groundwater levels. The constructed forecasting neural network models have yielded results which showed very precise forecasts for the monthly values of the groundwater levels.

Giustolisi and Simeone (2006), have investigated the implementation of a neural network based methodology for forecasting the groundwater levels by applying a multi-objective technique. The final results have illustrated a very good forecasting accuracy by implementing a linear model and also a nonlinear model in a short time series forecasting.

In this research, artificial intelligence is used for building feed forward multilayer models in order to predict the groundwater levels in the selected study area by using also several meteorological data. The groundwater levels are very crucial to be predicted, since the management of the groundwater resources is very valuable for the prosperity of human life and also for the health of the natural environment.

Several feed forward multilayer models have been constructed so as to investigate the development of the optimal prediction neural architecture. The developed network models can supply predictions of groundwater levels to the stakeholders which can be used in the decision-making process in water resources management and planning. In the next sections, the followed methodology, and also the discussion and the results are

described.

2 Theoretical Background

2.1 Feed Forward Artificial Neural Networks

The Artificial Neural Networks (ANNs) are computing systems that their structure can simulate the human brain neural structure (Basheer and Hajmeer 2000; Suykens et al. 2012). The Artificial Neural Networks (ANNs) were implemented in this study in order to develop and compare the most suitable prediction models to achieve the optimal groundwater prediction results. The main advantage of the neural networks is that they can model also non-linear relationships of the input and output parameters (Almeida 2002a, 2002b; Zhang et al. 1998). In this study, a Multilayer Perceptron with a Feedforward structure was utilized, so as to develop the neural network based forecasting models (Blum and Li 1991; Hornik 1991; Hornik et al. 1989). The feed forward neural networks were chosen to be implemented in this research, since the literature have illustrated that the feed forward neural networks are the most suitable when the research problems have to deal with time series predictions (Kouziokas 2017c, 2017d; Kouziokas et al. 2016, 2017; Tang and Fishwick 1993; Zhang et al. 2001). Figure 1 illustrates a typical neural network with a feed forward structure were the neurons are connected in only a forward direction.

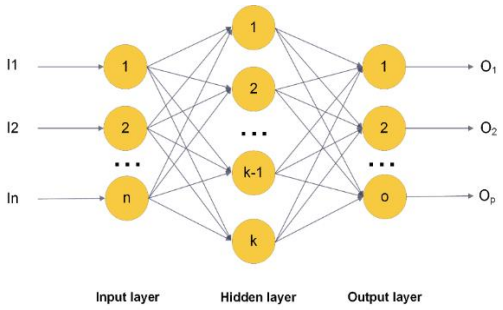

Fig. 1 A typical neural network with a feedforward structure.

2.2 The Levenberg Marquardt Algorithm

According to the results of the research, the Levenberg Marquardt algorithm produced the optimal forecasting results when used as the training algorithm in the Feed Forward Multilayer Neural Network models, compared to the other training algorithms.

The Levenberg Marquardt algorithm was presented in 1963 as one of the fastest training algorithms (Marquardt 1963). In this study, the Levenberg Marquardt algorithm produced better output results than the other tested algorithms in the neural network topologies of the constructed models, such as the Scaled Conjugate Gradient, the Resilient

Backpropagation and the BFGS Quasi-Newton.

The Levenberg Marquardt algorithm operates by combining the Gauss-Newton and the steepest descent algorithms and can be utilized for solving problems of nonlinear nature (Liu 2010; Lourakis 2005).

2.3 Research Methodology

The methodology consists of five main stages: the data collection and preparation, the feed forward neural network forecasting model creation, the comparison of the constructed network models in order to discover the optimal model according to the performance, the application of the optimal neural network model in predicting the groundwater levels and at last evaluating the prediction results by using the root mean squared error.

At the first stage, groundwater information and meteorological data were collected and preprocessed in order to feed the developed neural network models. At the second stage, the neural network forecasting models were constructed by investigating different network topologies, training algorithms and transfer functions so as to develop the optimal neural network prediction model. In the next stage, a comparison of the constructed neural network models was implemented, so as to discover the optimal forecasting model by measuring the performance. In the fourth stage, the optimal neural network model was implemented in the test set in order to forecast the groundwater levels. In the final phase, the prediction results were assessed by comparing the Root Mean Squared Errors of the models. Fig. 2 shows an overview of the phases of the followed methodology.

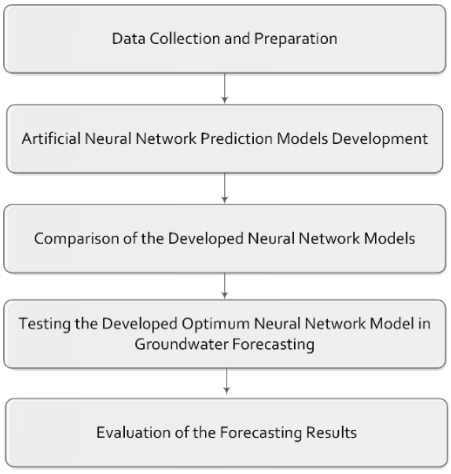

Fig. 2 Overview of the research methodology

3 Results

3.1 Data Collection and Preparation

The data regarding groundwater levels were downloaded from the official website of the United States Geological Survey (USGS) regarding the region of Montgomery country in Pennsylvania[1]. The meteorological data were downloaded from the official website of Pennsylvania State Climatologist office. At the initial stage, the data were pre-processed regarding the daily data of the average relative humidity, the precipitation, and the average temperature in order to be used as inputs in the constructed feed forward multilayer models.

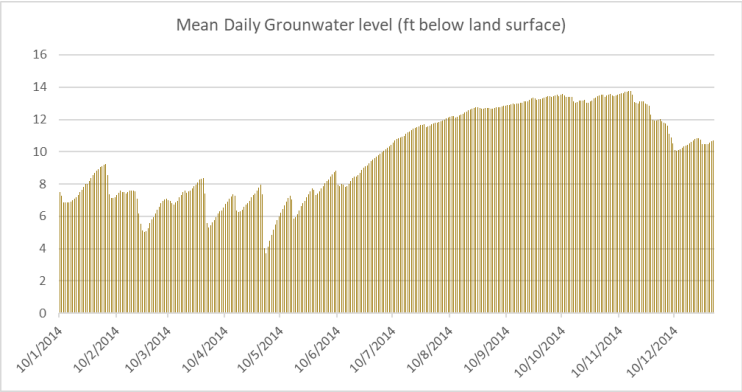

Fig. 3 Daily groundwater level (feet below the land surface)

The groundwater data that were collected for the Montgomery County Observation Well in Pennsylvania of USA for the 365 days of the year 2014, were prepared by checking them for duplicates, gaps, incoherencies, etc.). Figure 3 illustrates the values of the mean daily groundwater level (expressed in feet below the land surface) for the Montgomery County Observation Well regarding the year 2014.

3.2 Artificial Neural Network Models

Several input parameters were chosen in order to build the artificial neural network models. Several factors that influence the groundwater levels were taken into consideration such as the Average Relative Humidity (ARH), the Average Temperature (AT), the precipitation (PRE) which includes drizzle and rainfall and also the depth to the groundwater level (DGL) below land surface.

The data were separated into three different parts. The 70% of the collected and pre-processed data were utilized as the training set, the 15% for the validation set and the 15%

[1] https://www.usgs.gov/

for the test set. The training data set was used so as to train the feed forward neural network models with the collected data of the input factors. The validation set was utilized, so as to evaluate the network performance of every constructed neural network model. Also, the 10-Fold Cross-Validation method was utilized for evaluating the best prediction model. The 10-fold cross-validation is the most commonly used of the k-fold methods (McLachlan et al. 2005). In the *k*-fold cross-validation, the data sample is randomly separated into equal samples with size *k*. One of the subsamples is used for validating the model, and the remaining $k-1$ subsamples for the training set (Kohavi 1995).

3.3 Optimum Neural Network Model Implementation

In order to build the optimum prediction model, the neural network performance of the forecasting models that were tested. Several neural network topologies were investigated by implementing one to two hidden layers. Furthermore, the number of the nodes (neurons) in of every hidden layer was investigated by testing one to fifty neurons in every hidden layer and also the most common transfer functions for the hidden layers were tested. The training algorithms that were tested for every network model include the most common learning algorithms: The Resilient Backpropagation (RB), the Levenberg Marquardt (LM), the Scaled Conjugate Gradient (SCG) and the BFGS Quasi-Newton (BFGS-QN).

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

Table 1, shows the detailed information about the developed optimal artificial neural network models for every tested training algorithm (the two most accurate models for every teste training algorithm) by using several topologies and transfer functions in the hidden layers. Figure 4, illustrates a comparison chart of the optimal developed neural network models.

Table 1 The best neural network models constructed by using different algorithms and topologies

MODEL	Training Algorithm	Number of hidden layers	Number of hidden nodes 1st layer	Number of hidden nodes 2nd layer	Transfer Functions	RMSE Training Error	RMSE Validation Error	10-Fold Cross-Validation RMSE Error
ANN1	RB	2	15	17	TSTF – LTF	1.1851	0.4469	0.3587
ANN1A	RB	2	19	18	TSTF – LTF	1.2189	0.3192	0.2986
ANN2	LM	2	22	25	TSTF – TSTF	0.0001	0.0742	0.0562
ANN2A	LM	2	16	27	TSTF – TSTF	0.0025	0.1787	0.1289
ANN3	SCG	2	15	26	TSTF – LTF	0.4944	0.6321	0.6298
ANN3A	SCG	2	12	23	TSTF – LTF	0.3589	0.5384	0.4985
ANN4	BFGS-QN	2	17	22	TSTF – LTF	0.0673	0.2820	0.2569
ANN4A	BFGS-QN	2	15	31	TSTF – LTF	0.0948	0.3756	0.3148

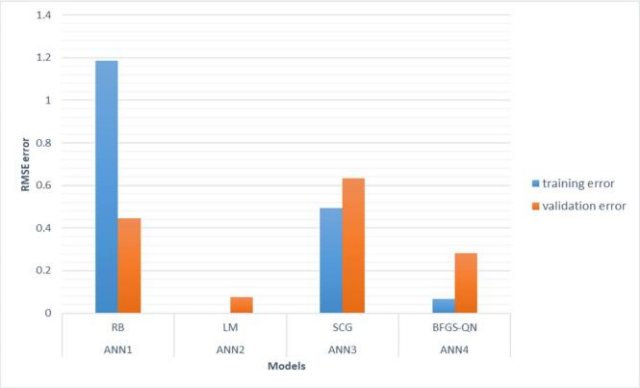

Fig. 4 Comparison chart of the best neural network models

The Root Mean Square Error (RMSE) was used so as to estimate the forecasting error and to assess the performance of every developed feed forward neural network model. The function for calculating RMSE is given by the following equation:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{p_i} - y_{r_i})^2} \tag{1}$$

where y_{p_i} is the predicted value, y_{r_i} is the real value and N represents the sample data. After testing different topologies and training algorithms the optimal neural network model was found to be the one with the Levenberg Marquardt (LM) algorithm. The optimal model was assessed according to the minimum Root Mean Squared Error (RMSE) that was found among all the other developed network models. The Root Mean Squared Error (RMSE) of the optimal model was found to be 0.0742 and the 10-Fold Cross-Validation RMSE Error was found 0.0562.

The results have shown that the optimal topology of the neural network model that have yielded the best predictions had 22 neurons in the first hidden layer and 25 neurons in the second hidden layer. The transfer function of the best model was the Tanh-Sigmoid Transfer Function (TSTF) in the first hidden layer and also the Tanh-Sigmoid Transfer Function (TSTF in the second hidden layer. In Fig. 5 the topology of the optimal constructed neural network model is illustrated.

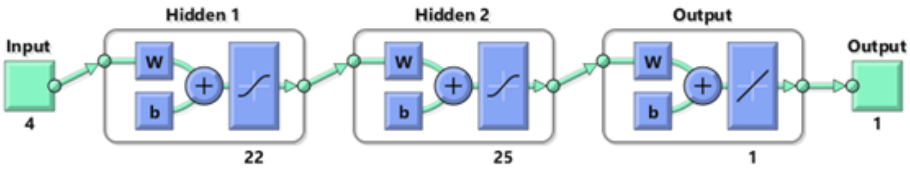

Fig. 5 The topology of the optimal artificial neural network model

In Fig. 6, the plots of the training parameters are illustrated: the gradient, the Mu parameter which affects the error convergence, and the validation checks performed during the training process of the optimal neural network topology.

In Fig. 7, the three-dimensional (surface) plot of the Mean Squared Error is illustrated.

The linear regression was implemented in order to assess the forecasting accuracy of the optimal neural network model. The R value was utilized which represents the linear relationship between the predicted and the target values. When the R value is one, it illustrates that an exact linear relationship exists between the predicted values of the neural network model and the target values regarding the groundwater levels.

The R value of the predicted levels and the real groundwater levels was found to be 0.99775. The R value illustrates that the groundwater level predictions when using the optimal feed forward prediction model are very accurate for the selected study area. In Fig. 8, the linear regression plot of the test set of the optimal neural network model between the predicted values (output) and the real values (target) of groundwater levels, is illustrated. Fig. 9, illustrates the comparison plot for the real and predicted values in the test set.

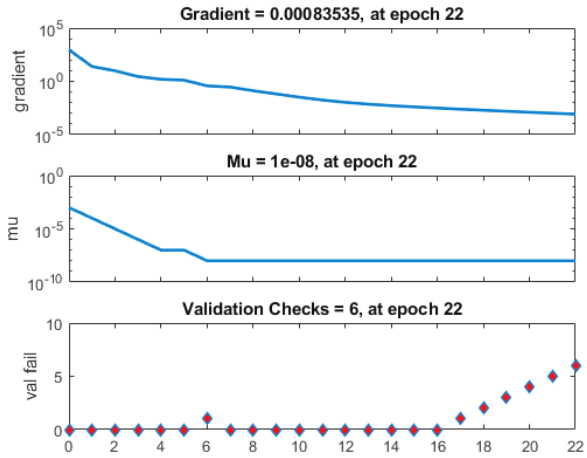

Fig. 6 The plots of the training parameters: gradient. Mu, validation checks for the optimal neural topology

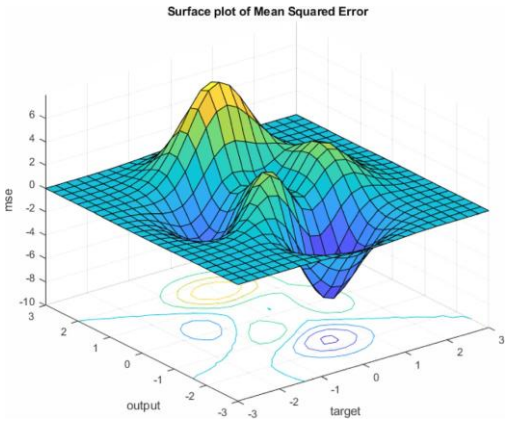

Fig. 7 The surface plot of mean squared error

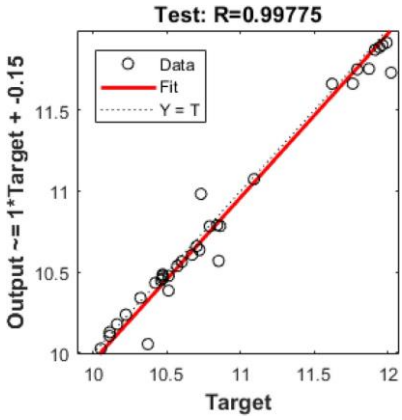

Fig. 8 The linear regression plot of the test set of the optimal neural network model between the predicted values (output) and the real values (target) of groundwater levels

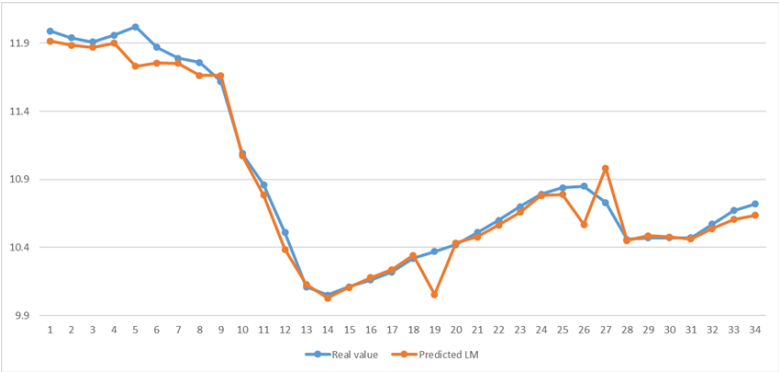

Fig. 9 The comparison plot for the real and predicted values of a time span of the test set

4 Conclusions and Discussion

In this research, feedforward neural network models have been applied in order to forecast the groundwater levels which are considered of high importance and affect the prosperity of human life and also of the natural environment. The proposed methodology can be very useful in public management and also at improving the environmental and water management strategies. In this paper, several feed forward neural models were constructed by investigating several factors and network parameters and also by testing several learning algorithms and the nature of the activation functions used in every hidden layer.

The best neural network prediction model was constructed by using meteorological data as input factors, that affect the amounts of the groundwater. The meteorological data were utilized as input parameters so as to predict the ground water values. Furthermore, several neural network architectures were constructed so as to find the optimal model according to the network performance.

Compared to other studies, this research has used feed forward neural networks which were fed by meteorological information as input parameters such as the Average Relative Humidity (ARH), the Average Temperature (AT), the precipitation (PRE) which includes drizzle and rainfall. The methodology is based on applying multiple training algorithms and also on tuning the neural network parameters in order to discover the optimal network structure that will produce the most accurate results. The main goal of testing different topologies was to improve the performance of the optimal feed forward prediction model. The results have illustrated a high forecasting accuracy regarding the predicted groundwater levels compared to other researches that have used artificial neural networks to produce groundwater predictions (Adamowski and Chan 2011; Chang et al. 2016; Giustolisi and Simeone 2006).

The methodology for developing the optimal neural network prediction model can be very valuable to the managers and stakeholders that deal with problems related to ground water management and also in adopting proactive measures in order to apply the best management and planning strategies so as to minimize the forthcoming dangers regarding the possible rapid losses of the amounts of the groundwater resources. This research can help at preventing the natural groundwater losses since drinking water is a very valuable good in people's life, by facilitating the decision makers to adopt the most suitable water management strategies and also is aiming at providing the environmental and water managers with a methodology for constructing the optimal prediction models.

Acknowledgements: United States Geological Survey (USGS) website for retrieving the groundwater data. A previous shorter version of the paper has been presented in the 10th World Congress of EWRA BPanta Rei^ Athens, Greece, 5-9 July 2017.

Compliance with Ethical Standards

Conflict of Interest None.

References

Almeida JS (2002) Predictive non-linear modeling of complex data by artificial neural networks Curr Opin Biotechnol 13:72-76

Adamowski J, Chan HF (2011) A wavelet neural network conjunction model for groundwater level forecasting JHyd 407:28-40

Almeida JS (2002) Predictive non-linear modeling of complex data by artificial neural networks Curr Opin Biotechnol 13:72-76

Babu BM, Kumar US, Reddy G (2016) Groundwater Level Forecasting Using Artificial Neural Network for Devasugur Nala Watershed in Raichur District, Karnataka Journal of Agricultural Engineering 53:19-27

Barzegar R, Fijani E, Moghaddam AA, Tziritis E (2017) Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models ScTEn 599:20-31

Basheer I, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application J Microbiol Methods 43:3-31

Biswas AK (2004) Integrated water resources management: a reassessment: a water forum contribution Water international 29:248-256

Blum EK, Li LK (1991) Approximation theory and feedforward networks Neural Netw 4:511-515

Chang F-J, Chang L-C, Huang C-W, Kao I-F (2016) Prediction of monthly regional groundwater levels through hybrid soft-computing techniques JHyd 541:965-976

Dean AJ, Fielding KS, Jamalludin E, Newton FJ, Ross H (2018) Communicating about sustainable urban water management: community and professional perspectives on water-related terminology Urban Water Journal:1-10

Famiglietti JS (2014) The global groundwater crisis Nature Climate Change 4:945

Giustolisi O, Simeone V (2006) Optimal design of artificial neural networks by a multi-objective strategy: groundwater level predictions Hydrological Sciences Journal 51:502-523 doi:10.1623/hysj.51.3.502

Guzman SM, Paz JO, Tagert MLM (2017) The use of NARX neural networks to forecast daily groundwater levels Water Resour Manage 31:1591-1603

Hong Y-M (2017) Feasibility of using artificial neural networks to forecast groundwater levels in real time Landslides 14:1815-1826

Hornik K (1991) Approximation capabilities of multilayer feedforward networks Neural Netw 4:251-257

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

Howard G, Bartram J, Pedley S, Schmoll O, Chorus I, Berger P, Chilton J (2006) Groundwater and Public Health. In: Protecting Groundwater for Health: Managing the Quality of Drinking-water Sources. Edited by O. Schmoll, G. Howard, J. Chilton and I. Chorus, IWA Publishing, London, UK, pp 3-19

Khalil B, Broda S, Adamowski J, Ozga-Zielinski B, Donohoe A (2015) Short-term forecasting of groundwater levels under conditions of mine-tailings recharge using wavelet ensemble neural network models HydJ 23:121-141

Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai, 1995. vol 2. Montreal, Canada, pp 1137-1145

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

Kouziokas GN (2016b) Geospatial Based Information System Development in Public Administration for Sustainable Development and Planning in Urban Environment European Journal of Sustainable Development 5:347-352 doi:http://dx.doi.org/10.14207/ejsd.2016.v5n4p347

Kouziokas GN (2016c) Technology-based management of environmental organizations using an Environmental

Management Information System (EMIS): Design and development Environmental Technology & Innovation 5:106-116. doi:http://dx.doi.org/10.1016/j.eti.2016.01.006.

Kouziokas GN (2017a) An Information System for Judicial and Public Administration Using Artificial Intelligence and Geospatial Data. In: Proceedings of the 21st Pan-Hellenic Conference on Informatics, Larissa, Greece, 2017. ACM, 3139402, pp 1-2. doi:10.1145/3139367.3139402

Kouziokas GN (2017b) Decision Support System Based on Indicators and Artificial Intelligence in Public and Environmental Management. In: Proceedings of the Sixth International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE) and the SECOTOX Conference, Thessaloniki, Greece, pp 649-657

Kouziokas GN (2017c) Machine Learning Technique in Time Series Prediction of Gross Domestic Product. In: Proceedings of the 21st Pan-Hellenic Conference on Informatics, Larissa, Greece, 2017. ACM, 3139443, pp 1-2. doi:10.1145/3139367.3139443

Kouziokas GN (2017d) The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment Transportation Research Procedia 24:467-473 doi:https://doi.org/10.1016/j.trpro.2017.05.083

Kouziokas GN, Perakis K (2017) Decision Support System Based on Artificial Intelligence, GIS and Remote Sensing for Sustainable Public and Judicial Management European Journal of Sustainable Development 6:397-404 doi:https://dx.doi.org/10.14207/ejsd.2017.v6n3p397

Kouziokas GN, Chatzigeorgiou A, Perakis K (2016) Predicting Environmental Data in Public Management by Using Artificial Intelligence. In: Proceedings of the 11th International Scientific Conference eRA-11, Piraeus, Greece, September 2016. Piraeus University of Applied Sciences, pp 39-46

Kouziokas GN, Chatzigeorgiou A, Perakis K (2017) Artificial Intelligence and Regression in Predicting Ground Water Levels in Public Administration European Water:361-366

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

Langridge R, Daniels B (2017) Accounting for climate change and drought in implementing sustainable groundwater management Water Resour Manage 31:3287-3298

Lourakis MIA (2005) A brief description of the Levenberg-Marquardt algorithm implemented by levmar Foundation of Research and Technology 4:1-6

Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters Journal of the society for Industrial and Applied Mathematics 11:431-441

McLachlan G, Do K-A, Ambroise C (2005) Analyzing microarray gene expression data vol 422. John Wiley & Sons.

Mohanty S, Jha MK, Raul S, Panda R, Sudheer K (2015) Using artificial neural network approach for simultaneous forecasting of weekly groundwater levels at multiple sites Water Resour Manage 29:5521-5532

Mysiak J, Giupponi C, Rosato P (2005) Towards the development of a decision support system for water resource management Environ Model Software 20:203-214 doi:10.1016/j.envsoft.2003.12.019

Supreetha B, Nayak PK, Shenoy NK (2015) Groundwater level prediction using hybrid artificial neural network with genetic algorithm International Journal of Earth Sciences and Engineering 8:2609-2615

Suykens JA, Vandewalle JP, de Moor BL (2012) Artificial neural networks for modelling and control of non-linear systems. Springer Science & Business Media

Tang Z, Fishwick PA (1993) Feedforward neural nets as models for time series forecasting ORSA journal on computing 5:374-385

Tapoglou E, Trichakis IC, Dokou Z, Nikolos IK, Karatzas GP (2014) Groundwater-level forecasting under climate change scenarios using an artificial neural network trained with particle swarm optimization Hydrological Sciences Journal 59:1225-1239

Thomas BF, Gibbons AC (2018) Sustainable Water Resources Management: Groundwater Depletion. In: The Palgrave Handbook of Sustainability. Springer, pp 53-77

Zhang G, Patuwo BE, Hu MY (1998) Forecasting with artificial neural networks: The state of the art Int J Forecasting 14:35-62

Zhang GP, Patuwo BE, Hu MY (2001) A simulation study of artificial neural networks for nonlinear time-series forecasting Computers & Operations Research 28:381-396