Improving groundwater level forecasting with a feedforward neural network and linearly regressed projected precipitation

Ioannis K. Tsanis, Paulin Coulibaly and Ioannis N. Daliakopoulos

ABSTRACT

A module that uses neural networks was developed for forecasting the groundwater changes in an aquifer. A modified standard Feedforward Neural Network (FNN), trained with the Levenberg–Marquardt (LM) algorithm with five input variables (precipitation, temperature, runoff, groundwater level and specific yield) with a deterministic component, is used. The deterministic component links precipitation with the seasonal recharge of the aquifer and projects the seasonal average precipitations. A new algorithm is applied to forecast the groundwater level changes in Messara Valley, Crete, Greece, where groundwater level has been steadily decreasing due to overexploitation during the last 20 years. Results from the new algorithm show that the introduction of specific yield improved the groundwater level forecasting marginally but the linearly projected precipitation component drastically increased the window of forecasting up to 30 months, equivalent to five biannual time-steps.

Key words | aquifer overexploitation, artificial neural networks, forecasting, groundwater, Messara Valley, specific yield

Ioannis K. Tsanis (corresponding author)
Paulin Coulibaly
Department of Civil Engineering,
McMaster University,
Hamilton, Ontario,
Canada
E-mail: *tsanis@mcmaster.ca*

Ioannis N. Daliakopoulos
Department of Environmental Engineering,
Technical University of Crete,
Chania,
Greece

NOMENCLATURE

AI	Artificial Intelligence
BPTT	Backpropagation Through Time
FNN	Feedforward Neural Network
IDNN	Input Delay Neural Network
LM	Levenberg–Marquardt (algorithm)
MLP	MultiLayer Perceptron
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
A	evaluation criterion
A_{forecast}	forecast criterion
$A_{\mathrm{reference}}$	reference criterion
A_{perfect}	perfect fit of criterion
E_i	residual error in results
ϵ_i	percentage error
H	groundwater level
N	number of observations
R	annual aquifer recharge
S_{y}	specific yield
skill	skill score
t	time-lag
t_{m}	monthly time-step
t_{b}	biannual time-step
W	annual aquifer withdrawal
y_i	observed variable
$\hat{y}_i$	calculated variable

doi: 10.2166/hydro.2008.006

INTRODUCTION

Groundwater is an inherent part of the hydrological cycle. While precipitation and surface water bodies recharge the underground water bodies, groundwater steadily flows towards a discharge point or is stored in underground geological formations. Provided that the groundwater is primarily influenced by hydro-meteorological processes, the water table fluctuates periodically. Under natural conditions, the aquifer fluctuates around a multi-annual average and there is a balance between annual recharge and withdrawal.

Hydrologically wet years can cause an upward trend to the groundwater level over a series of years but are generally balanced by a succession of dry years where precipitation is below the recorded average. When the natural equilibrium is affected by human activities as excess infiltration from irrigation, continuous depletion by pumping, construction of a canal which either recharges or drains the groundwater, etc., sharp changes can be observed (Kovacs 1981). Exceptions can also occur depending on precipitation patterns. For example, less frequent but heavier precipitation events during the annual cycle could lead to longer dry periods, even though annual rainfall amounts might have increased (Andreadis & Lettenmaier 2006).

The overexploitation of groundwater resources due to the increased demand from agriculture, industry and domestic use is one of the largest problems in Europe's arid Mediterranean areas (Lehner *et al.* 2001). The lack of sufficient precipitation and permanent streams has increased the stress on groundwater resources. Development brought an increase of the water resources exploitation, often in a non-sustainable way. As a result, many watersheds already face severe financial, environmental and social consequences, which in turn exercise financial, social and political pressure for the solution of the conflict for the exploitation of groundwater resources. Many areas of the Mediterranean that face a problem of groundwater overexploitation are located in Greece. A characteristic example is the western part of the Valley of Messara in the island of Crete, Greece (Balabanis 1999). Groundwater level predictions for an area that faces a groundwater depletion problem are a useful tool for extreme event management, environmental control and protection and even a water price policy tool. Until now, the authorities of the region have not been able to obtain accurate predictions or successfully setup a conceptual model since data availability and the complex geology limit model effectiveness and accuracy.

Data-driven models, primarily developed for Artificial Intelligence (AI) applications, are based on a limited knowledge of the modelling process and rely on the data describing input and output characteristics. These methods, however, are able to make generalizations of the process and often play a complementary role to physically based models. Artificial Neural Networks (ANNs) are information processing systems that roughly replicate the behaviour of a human brain by emulating the operations and connectivity of biological neurons. As such, they can be used for data driven modelling. From a mathematical point of view, ANN is a complex nonlinear function with many parameters that are adjusted (calibrated or trained) in such a way that the ANN output becomes similar to the measured output of a known dataset (Solomatine 2002). Sometimes 'hybrid models' are built combining both types of approaches: data driven and knowledge based (Jain & Kumar 2007).

The challenge of this research is to overcome the problems of conventional groundwater modelling techniques by applying innovative solutions. Neural networks are a relative new technique in hydrological modelling and generally accepted validation techniques do not exist, nor a defined approach of neural network result justification (ASCE 2000*a*,*b*). The innovation of this tool is that it can be applied in areas characterized by scarcity of hydrogeological data, such as Messara Valley, rendering traditional modelling techniques difficult, if not impossible. ANN models have had limited real life engineering applications, and our objective is to bring this mostly experimental technique a step closer to real-world applications.

Although conceptual models are the main tool for depicting hydrological variables and understanding the physical processes taking place in a system, they do have practical limitations. When data are not sufficient and getting accurate predictions is more important than explaining the actual physics, empirical models remain a good alternative method and can provide useful results without the sometimes costly calibration time.

ANN models are 'black box' models with particular properties that are well suited to dynamic nonlinear system modelling. The advantages of ANN models over conventional simulation methods have been discussed in detail by French *et al.* (1992). ANN applications in hydrology vary from real-time to event-based modelling. They have been used for rainfall–runoff modelling, precipitation forecasting and water quality modelling (Govindaraju & Ramachandra Rao 2000). One of the most important features of ANN models is their ability to adapt to recurrent changes and detect patterns in a complex natural system. More concepts and applications of ANN models in hydrology have been discussed by Govindaraju & Ramachandra Rao (2000) and by the ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (ASCE 2000*a*,*b*). ANNs have also been applied with success in groundwater level prediction with limited data (Coulibaly *et al.* 2001*b*).

Feedforward Neural Networks (FNNs) have been applied successfully in many different problems since the advent of the error backpropagation learning algorithm and have been thoroughly discussed. A multilayer perceptron network consists of an input layer, one or more hidden layers of computation nodes, and an output layer. Figure 1 shows a typical FNN with one hidden layer consisting of three nodes, four input neurons and one output along with a schematic of training and cross-validation procedure. The input signal propagates through the network in a forward direction, layer by layer. After the network has been fully trained, it can produce predictions. The Levenberg–Marquardt (LM) method is a modification of the classic Newton algorithm for finding an optimum solution to a minimization problem and has proven to be a fast and efficient training algorithm for ANN applications (Coulibaly *et al.* 2000; Daliakopoulos *et al.* 2005).

The importance of data availability

Artificial neural networks are traditionally data demanding in the sense that long datasets are usually needed in order for the network to be adequately trained. The amount of data available for training the network is important, giving a better chance for locating global minimum of the error function. Nevertheless, the quality of the available data and the relevance of the input data with the desired output are also important since a large amount of irrelevant data can hinder the network's performance by confusing the training process (ASCE 2000*a*). There therefore has to be a balance between the quantity of data and the relevancy to the output or else a sensitivity analysis has to be conducted.

Since data availability is already a constraining factor in the case study, all available data were used. For this reason, the input variables of precipitation, temperature, discharge and groundwater level were used for most of the applied methods. When more data were collected from the study area, namely specific yield, it was possible to run the model with one more input variable and compare the results.

Recent applications

Literature provides a good reference in the use of ANNs as a forecasting tool for hydrogeological variables. In the majority of cases, the MultiLayer Perceptron (MLP) or variations are used. Variations of the MLP include the use

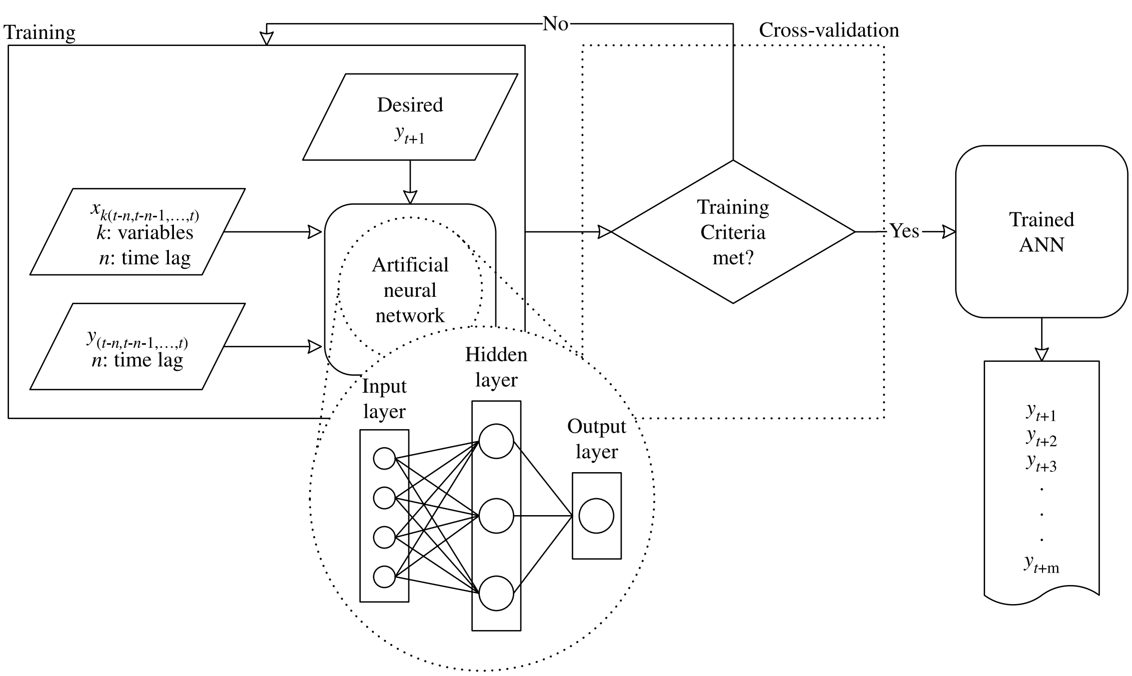

Figure 1 | ANN model calibration and prediction schematic.

of time-delayed variables. This remark is also backed by Coulibaly *et al.* (1999) who note that, on average, 90% of the application cases in hydrology have used the MLP technique. When compared with other network architectures, MLPs are generally considered superior and faster to train.

Coulibaly *et al.* (2001*a*) report that Recurrent Neural Networks (RNNs) may perform better in predicting monthly runoff at the cost of a longer training time. In order to investigate this further, four different networks were used: an MLP, an RNN and their adaptive analogs, an Input Delay Neural Network (IDNN) and a Time Delay Neural Network (TDNN). Both feedforward networks were trained with the LM algorithm and both recurrent networks with the Backpropagation Through Time (BPTT) method. Adding an input delay to the MLP dramatically improved its performance; however, recurrent configurations appear to have better results in short-term forecasting.

Coppola *et al.* (2003*a*,*b*, 2005) and Feng *et al.* (2008) also developed neural network models for predicting groundwater levels in real-world aquifer systems. Coppola *et al.* explicitly accounted for pumping rates, where groundwater withdrawals of wells were used as inputs. Feng *et al.* considered a regional aquifer system and implicitly accounted for pumping rates by estimating withdrawals based upon surrogate variables (e.g. total irrigation area).

Various architectures and variable configurations of ANN models were studied by Daliakopoulos *et al.* (2005) in order to determine which gives better predictions of the behaviour of the groundwater level of the study area. A total of seven different ANN configurations were tested in terms of optimum results for a prediction limited to 18 months. The most suitable configuration for this task proved to be a 20-3-1 Feedforward Neural Network trained with the LM method as it showed the most accurate predictions of the decreasing groundwater levels. It was inferred that ANN can be applied in cases where the datasets manifest trends and shifts and the desired output lies outside of the range of previously introduced input.

Summarizing, networks that use input time delay and stopped training seem to have a better performance. Most authors comment on the significance of input data and propose input selection for every different scenario and, when possible, the use of different input parameters for different variable classes in the same scenario. Longer training sets with a focus on more extreme events and the use of non-traditional variables such as snow accumulation (Zealand *et al.* 1999) or radar data (Kim & Barros 2001) as input variables can have a positive effect on results. It is also evident that there is no standard method of training neural networks in order to achieve good results. In most cases, a different approach is being taken leading to radically different results. Neural networks also offer a large range of possible configurations and training techniques so a lot of research is still required.

A new methodology is proposed in this paper which uses the FNN with five input variables (specific yield, precipitation, temperature, discharge and groundwater level) trained with the LM algorithm in conjunction with a deterministic component of linearly regressed projected precipitation for groundwater level forecasting. In each time-step, the projected precipitation is fed to the model to predict recharge, which is corrected. A new value of projected precipitation is produced in order to extend the period of the groundwater level forecasting.

METHODOLOGY

Initial model

A 20-3-1 FNN trained with the LM algorithm proved to be the most suitable Neural Network configuration for predicting groundwater levels when working with monthly data as showed by Daliakopoulos *et al.* (2005). During calibration, the values that correlated better with all network outputs where those of a 5 month moving window through the data series. Thus the input layer in all networks consisted of 20 input nodes. Time-lags of up to 5 months were included (i.e. time-lags t_m, $t_m - 1$, $t_m - 2$, $t_m - 3$ and $t_m - 4$ considering x_{t_m} is the value of a given variable at the present monthly time-step t_m) for precipitation, temperature, stream flow and groundwater level. The output of the network is a prediction of the well level at monthly time-step $t_m + 1$.

New approach

The study area this work focuses on has only two distinct seasons (wet and dry). This observation has led to the inference that annual aquifer recharge R can be linked to the annually available water for infiltration which is a fraction of the annual precipitation. Also, annual aquifer withdrawal W is mainly caused by pumping for irrigation which only occurs during the dry season. Other overland and groundwater sources and sinks (as stream discharge inflow and runoff or groundwater leakage) are in this case either negligible or have been taken into account in the calculation through a balance equation. For a biannual time-step t_b, R and W of the aquifer are calculated using the following equations

$$R_{(t_b \to t_b+1)} = H_{t_b+1} - H_{t_b} \tag{1}$$

where $t_b \to t_b + 1$ represents a biannual time interval in the wet season, and

$$W_{(t_b \to t_b+1)} = H_{t_b} - H_{t_b+1} \tag{2}$$

where $t_b \to t_b + 1$ represents a biannual time interval in the dry season. H_{t_b} and H_{t_b+1} are groundwater level values in corresponding time-steps (Figure 2). When R and W are multiplied by the average specific yield S_y at the current groundwater level, the actual amount of water that has infiltrated or lost can be estimated respectively. S_y is the ratio of the volume of water that drains from a saturated rock driven only by gravity to the total volume of the rock (Meinzer 1923) and characterizes the storage capacity of unconfined aquifers (Kovacs 1981). Intuitively, this amount of water is a more appropriate input for our ANN given the fact that it is an empirical model without further knowledge of the physical system.

This was also the motivation for focusing on a new approach which would aim to predict biannual R and W for only one time-step per forecast and subsequently update the input variables according to deterministic equations. During the dry season, precipitation is assumed equal to the long-term dry seasonal average. Temperature and discharge for the seasonal stream are also assumed equal to the dry seasonal average. During the wet season when the aquifer is recharged, precipitation is assumed to follow a historically based trend that depends on the relationship between corrected recharge and precipitation. This relationship can be updated depending on the available data.

The iterative algorithm that estimates the variables is also graphically represented in the schematic of Figure 3. Initially four successive time-steps are used as input for the trained network. Two successive groundwater levels are predicted and, depending on the season, the algorithm flow is lead to the next iteration through one of two states: (a) For the dry season, the corresponding precipitation is estimated from the seasonal average and (b) for the wet season R is multiplied by specific yield to estimate the actual water that has infiltrated and therefore the precipitation that would correspond to this

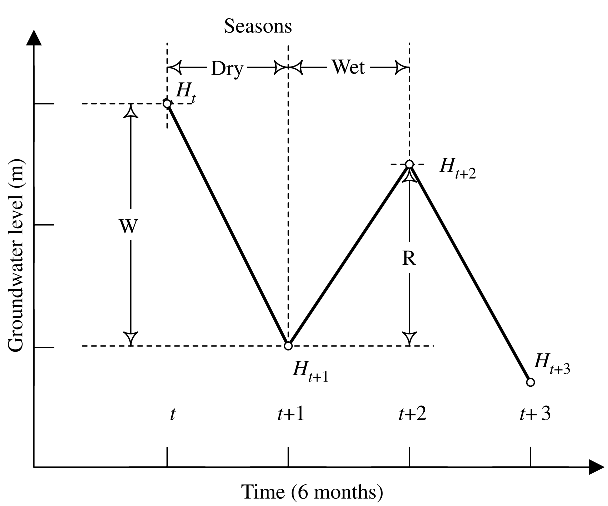

Figure 2 | Graphical representation of withdrawal *W* and recharge *R*.

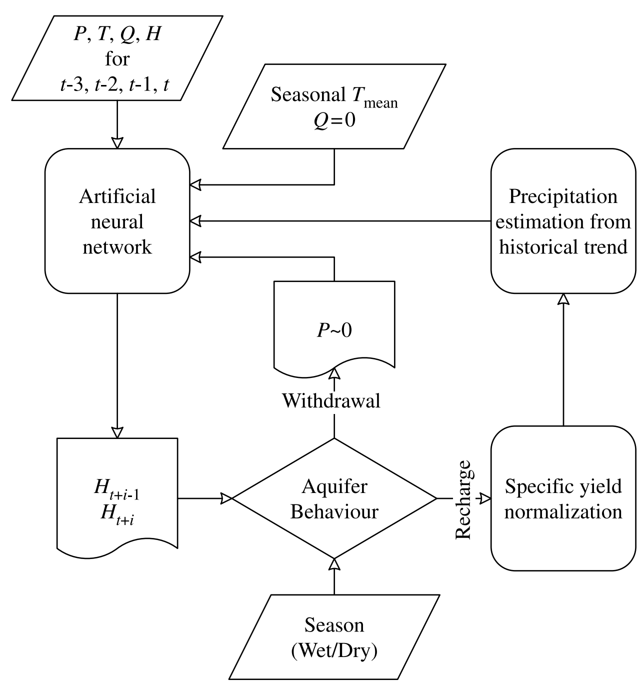

Figure 3 | Predictions schematic for the new approach.

infiltration. The estimated precipitation, together with the seasonal averages of temperature and stream flow, are used as input for the successive iteration.

CRITERIA OF EVALUATION

Several criteria were used in order to evaluate the effectiveness of each method and its ability to make precise predictions. The Root Mean Square Error (RMSE) is calculated by:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \tag{3}$$

where y_i is the observed variable, $\hat{y}_i$ the calculated variable and N is the number of observations. RMSE indicates the discrepancy between the observed and calculated values. The lower the RMSE, the more accurate is the prediction.

The R^2 efficiency criterion is calculated as

$$R^2 = \left(\frac{N\sum y_i\hat{y}_i - (\sum y_i)(\sum \hat{y}_i)}{\sqrt{N\sum y_i^2 - (\sum y_i)^2}\sqrt{N\sum \hat{y}_i^2 - (\sum \hat{y}_i)^2}}\right)^2 \tag{4}$$

representing the percentage of the initial uncertainty explained by the model. The perfect fit between observed and calculated values, which is unlikely to occur, would have RMSE $= 0$ and $R^2 = 1$.

The residual error in the results is given by

$$E_i = y_i - \hat{y}_i \tag{5}$$

where y_i is the observed and $\hat{y}_i$ the calculated groundwater level. The percentage error of a variable is given by

$$\varepsilon_i = \frac{y_i - \hat{y}}{y_i}100\% \tag{6}$$

Finally, the skill score (WWRP/WGNE 2005) implies information about the relative improvement of a forecast over an alternative reference forecast, thus comparing two different forecasting methods. In

$$\mathrm{skill} = \frac{A_{\mathrm{forecast}} - A_{\mathrm{reference}}}{A_{\mathrm{perfect}} - A_{\mathrm{reference}}} \tag{7}$$

A represents an evaluation criterion such as those presented above. A_{forecast} and $A_{\mathrm{reference}}$ are compared with respect to the value that represents the perfect fit of the given criterion, A_{perfect}. A skill value of 1 means that the current forecast is perfect, whereas a value of 0 indicates no improvement over the reference.

CASE STUDY

The developed models were tested with data acquired at Messara Valley, which covers an area of 398 km^2 at the southern part of the island of Crete in Greece (Figure 4). Messara Valley faces a severe problem of depletion of groundwater resources, mainly used in agriculture. Many wells in the region are illegitimate and pumping is often unregulated. This results in the overexploitation of the aquifer and, as a consequence, the sinking of the groundwater level over the years. During the period 1984–2001, mean annual precipitation was 516 mm yr^{-1}, well below the long-term average of 588 mm yr^{-1}. With an estimated 65% of total evaporation and a measured discharge of 21 mm yr^{-1}, the annual recharge of the aquifer can be calculated as

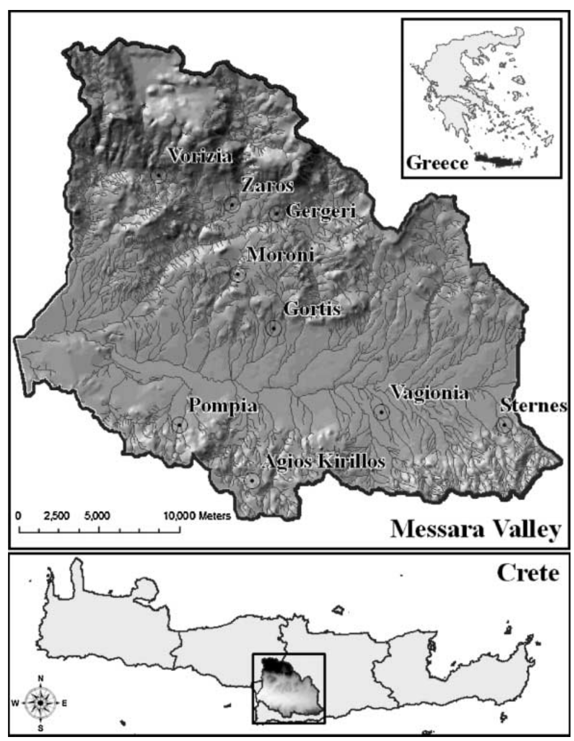

Figure 4 | Area of study, Messara Valley, Crete, Greece, with the location of the rain gauges in the area.

159 mm yr^{-1}. About 108 mm yr^{-1} are lost from subsurface outflow and if constant abstractions for irrigation are assumed to be 97 mm yr^{-1}, then the average net loss of water resources totals 46 mm yr^{-1}. Taking into account the porosity and area of Messara Valley, the annual decrease of the groundwater level can be calculated as no less than 1.5 m yr^{-1} (Vardavas *et al.* 1997) which is also implied by the level of a representative well in the area (Figure 5).

The karstic limestone formations constitute a significant groundwater loss mechanism (Stringfield & LeGrand 1971) in Crete. This makes the water resources of Messara Valley more susceptible to droughts, as demonstrated by the natural drop in the groundwater level by about 10 m in the early 1970s as reported by Croke *et al.* (2000). During that period, precipitation was well below average and there was also little pressure from agricultural pumping. Alluviual deposits fill erosion troughs within the Lower Pleistocene and comprise upper Pleistocene reddish and brown clay, silt and gravel beds and grey Holocene deposits of gravel, sand, silt and clay, often with organic matter. At the Geropotamos stream, the alluvial deposits extend from a few hundred metres to about 1.5 km and their corresponding thickness increases from about 60 m to 100 m.

The data acquired from the area consists of rainfall and temperature timelines, the discharge of the main stream of the valley, the depth of a representative well located in Pompia (Figure 5) and specific yield measurements and estimations.

Precipitation

Monthly average areal precipitation P for the whole study area shows the typical characteristics of the Mediterranean climate comprising of high rainfall during the winter months and no rainfall during the summer months, as shown in Figure 6. The average areal precipitation is about 585 mm yr^{-1}. Wet (1982, 1984, 1985, 1988, 1996 and 2001) and dry years (1986, 1990, 1993 and 2000) can be distinguished and appear to be equally distributed with a period of 2 to 3 years. The maximum precipitation of this 20-year period occurred in 1995–1996, giving a total value of 798 mm. It should be noted that, for these statistics, the hydrological year was considered from September to August of the next calendar year.

Stream flow

Figure 7 represents the seasonal discharge Q of Geropotamos Stream, the main seasonal stream that runs through Messara Valley. The water level of Geropotamos Stream has been steadily decreasing for the past 20 years due to the overexploitation of the water resources, part of which would be otherwise discharged into the sea. In Table 1, runoff can be seen both as Mm^3 and mm. The groundwater level reduction has caused wetlands in the area to dry up (Croke *et al.* 2000) and the stream to have no flow during most of the year. This steadily decreasing trend is also

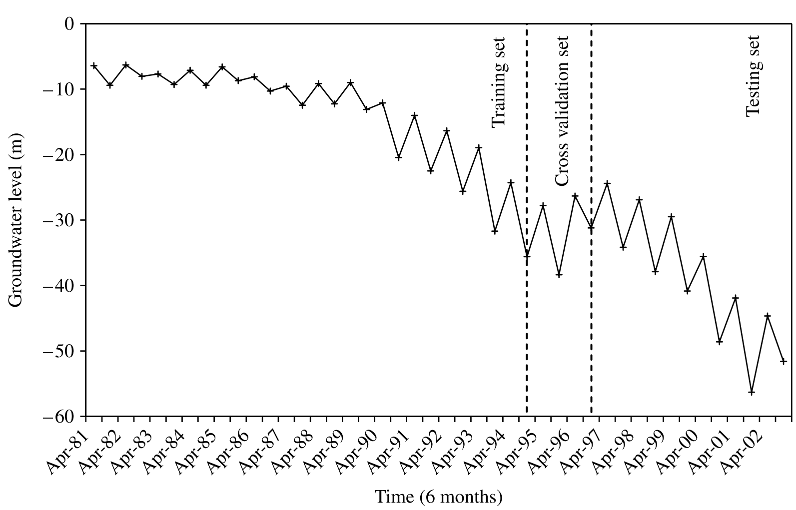

Figure 5 | Seasonal data from well station 207.

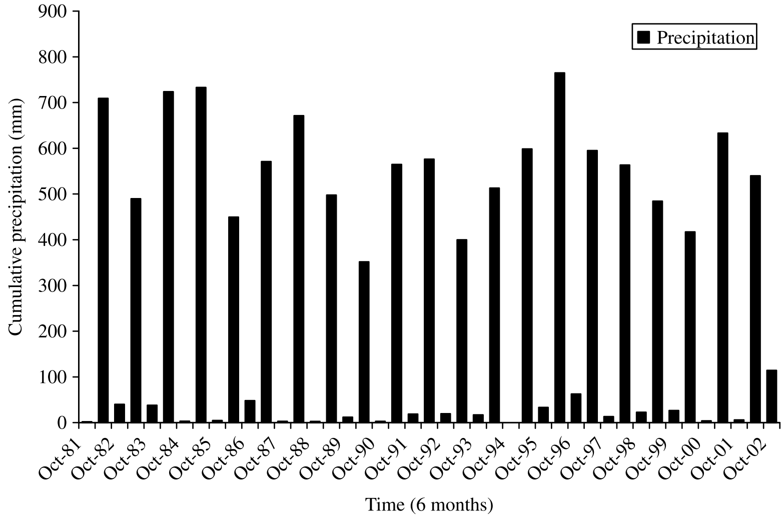

Figure 6 | Average seasonal precipitation, estimated with the Thiessen method for the whole watershed of Messara.

depicted in Figure 7, where from a maximum of 43 Mm^3 in 1985 the discharge was diminished to zero during the wet period of 2000.

Temperature

Temperature T also plays an important role in the water budget as it affects evapotranspiration and irrigation withdrawals via well extractions. Temperature values vary almost steadily throughout the available dataset with a wet season average of 12.7°C and a dry season average of 24.2°C. There is a minor increasing trend which is neutralized when compared to a longer timescale, showing no significant long-term temperature variation.

Specific yield

Specific yield S_y is very difficult to measure mainly because the properties of soil in larger depths cannot be retained and studied easily in the lab. The values that were used in the case study came from both literature and communication with expert scientists conducting experiments at the area (Personal communication with M. Kritsotakis 2004). Figure 8 shows field measurements (denoted $\times$) and expert

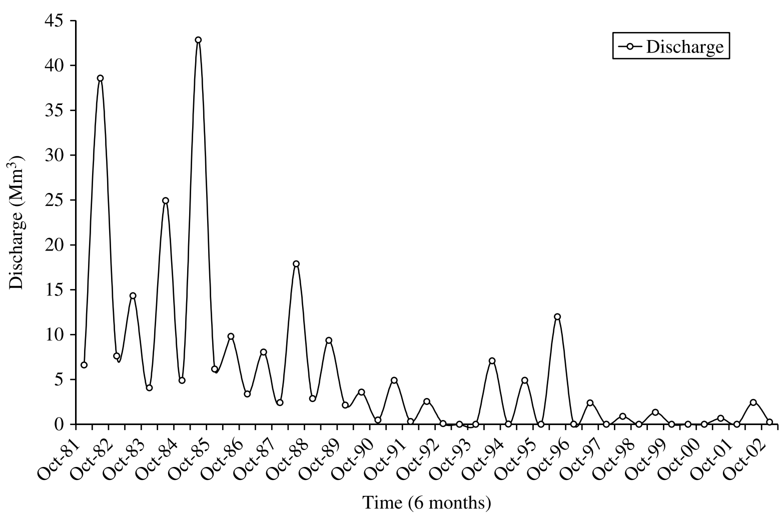

Figure 7 | Cumulative seasonal runoff of Messara Valley as measured at the location Phaistos.

Table 1 | Detailed presentation of groundwater level withdrawal, precipitation and discharge values during the dry season from 1981 to 2001

Years	Discharge (Mm^3)	Dry season precipitation (mm)	Reference withdrawal (m)	Corrected withdrawal (m)
1981	6.61	2.2	3.0	3.30
1982	7.61	40.1	1.8	1.98
1983	4.06	38.3	1.6	1.76
1984	4.89	3.3	2.3	2.53
1985	6.16	4.9	2.1	2.31
1986	3.38	48.2	2.2	2.42
1987	2.42	3.0	3.0	3.25
1988	2.87	2.8	3.1	3.38
1989	2.14	12.1	4.1	4.36
1990	0.48	3.2	8.4	4.68
1991	0.31	19.0	8.5	6.02
1992	0.07	19.8	9.2	5.82
1993	0.00	17.2	12.7	7.21
1994	0.03	0.0	11.3	6.22
1995	0.00	33.5	10.6	5.83
1996	0.01	62.8	4.9	2.70
1997	0.00	13.5	9.8	5.39
1998	0.00	22.9	11.0	6.05
1999	0.00	26.7	11.4	6.27
2000	0.00	4.0	13.0	7.15
2001	0.00	6.2	14.4	7.92

estimations (denoted o). This information was taken into account in order to derive a specific yield profile with respect to the depth of the aquifer.

For shallow depths (0 to -12 m) the specific yield of the ground lies between 0.1 and 0.12 and for larger depths (deeper than -20 m) it lies between 0.04 and 0.07 (Figure 8). A specific yield value of 0.1 implies that from the actual value of observed groundwater level, only 10% is water. For example, if in one year the groundwater level changes from -7 m to -8 m, i.e. decreases by 1 m, then the actual water lost is equivalent to 0.1 m. For medium depths (-12 to -20 m), specific yield decreases with an uncertain manner.

After 1989, measurements show a large difference in the groundwater level behaviour, as illustrated in Figure 9. A decreasing trend due to overexploitation but also a simultaneous increase in the amplitude of the oscillation can be observed. The latter increase can only be justified by a change in the properties of the rock around the depth of 10–20 m. The corrected groundwater depths retain the decreasing trend due to excessive pumping, but have lost the characteristic of increased oscillation to some extent. An individual analysis of the annual aquifer withdrawal with respect to summer rainfall and stream discharge, and

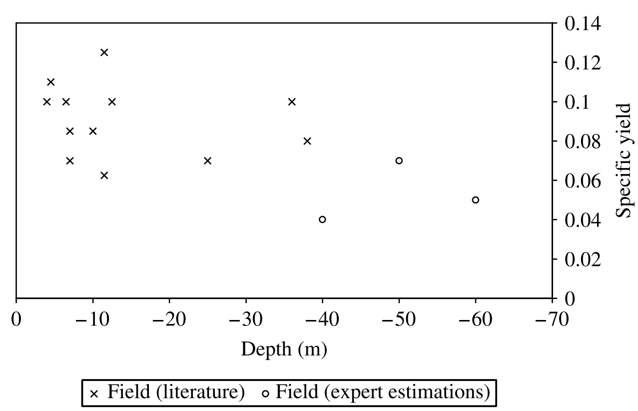

Figure 8 | Assumed specific yield profile (× denote field data taken from literature and o denote field data according to personal communication with M. Kritsotakis 2004).

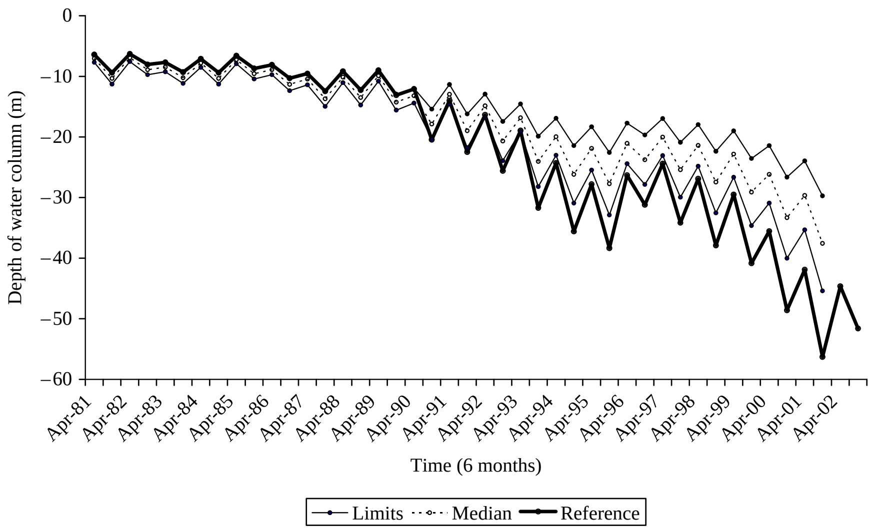

Figure 9 | Groundwater level measurements (reference) compared with the estimated average infiltrated water for a given specific yield uncertainty interval.

recharge with respect to precipitation during the wet season and stream discharge, justifies the assumptions.

There is a direct relation between the annual recharge and the rainfall that takes place during the wet season for years before 1989, as shown in Figure 10. Especially dry years such as 1982–1983, 1985–1986 and 1989–1990 contribute very little to the aquifer level. On the other hand, wet years such as 1981–1982, 1983–1984, 1984–1985 and 1987–1988 contributed as much as 3.6 m depending on discharge and precipitation patterns.

After 1989, a relatively large change takes place in the aquifer behaviour. This can be observed both from the depth measurements (Figure 5) and from the individual recharge (Figure 10). This significant change is due to the change of irrigation methods, the over-cultivation of Messara Valley and the resulting larger demand. Nevertheless, the augmentation of the demand by itself does not justify the apparent larger recharge events (Figure 10) that take place after 1989, even though precipitation load remains the same.

During this period, the annual aquifer withdraw becomes larger with a maximum of 12 m in 2001. Even though 2001 was indeed a very dry season, for comparable events before 1989 (for example in 1984 and 1985, withdrawal was only as much as 2.3 m; Table 1) irrigation alone

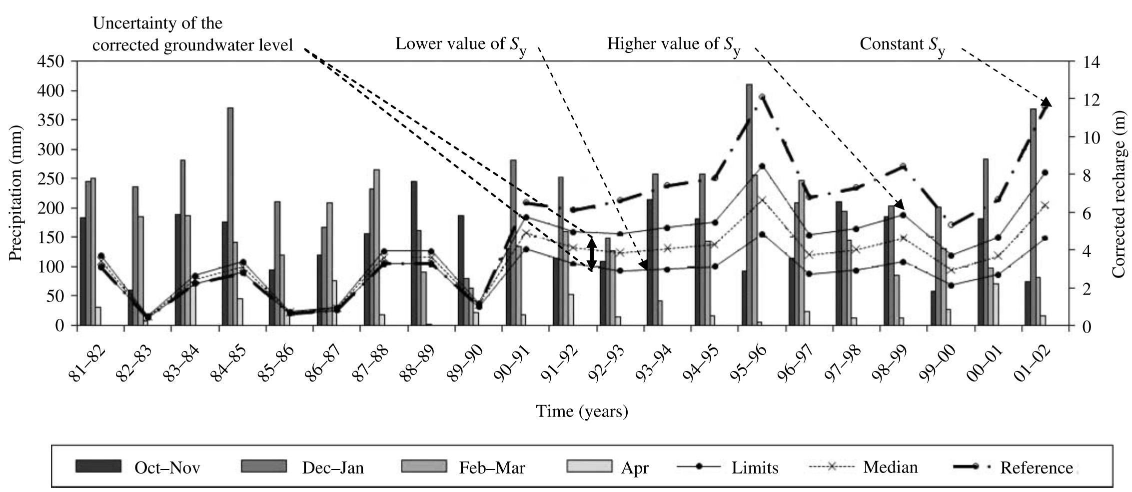

Figure 10 | Representation of the monthly wet season precipitation (mm) and annual recharge of the water column assuming a uniform specific yield profile throughout the aquifer depth (reference line refers to the groundwater level assuming constant specific yield; limit lines refer to groundwater level corrected with the corresponding specific yield intervals).

is not sufficient to justify a difference of 9.7 m. Inconsistencies such as this were partially corrected after incorporating the specific yield effect into the groundwater level timeline.

Regarding recharge, after the parameters of the natural system started changing in 1989, the previously observed correlation decreased. For example, the driest two winter periods of 1992–1993 and 1999–2000 in the dataset (400 mm and 417 mm of precipitation, respectively) contributed 6.6 m and 5.3 m of water to the aquifer. This observation seems to justify the assumption of a specific yield scaling even more. Actually, when the groundwater depth is corrected with the assumed specific yield profile, rainfall and recharge were strongly correlated.

RESULTS AND DISCUSSION

Method I

In the first method, more field measurements were incorporated in the model in order to provide a more global description of the natural system. The dataset from 1981 to 2002 comprised seasonal data (two values for each year for each variable). Since one of the most important factors that affect groundwater level is specific yield, one more input variable was added to the network. Practically, five input nodes were included for every variable, given that five biannual time delays are used for each input variable. The additional node was reserved for specific yield values. Because of the increase in the number of neurons (variables × time-lag) from 20 to 25, the network parameters had to be recalibrated. Through trial and error, the architecture of the FNN was specified to 25-5-1 (input-hidden-output), increasing the hidden nodes from three in the original method to five.

Table 2 | Comparison of RMSE statistics for Methods I with four and five inputs for different time-steps

		Four inputs (P, T, Q, H)		**Five inputs (P, T, Q, H, S_y)**	
Time-steps	**Criterion**	**Test**	**Validation**	**Test**	**Validation**
3	RMSE (m)	0.2	0.5	0.4	0.3
5	RMSE (m)	1.0	1.1	0.9	0.5
12	RMSE (m)	10.6	1.0	10.9	0.6

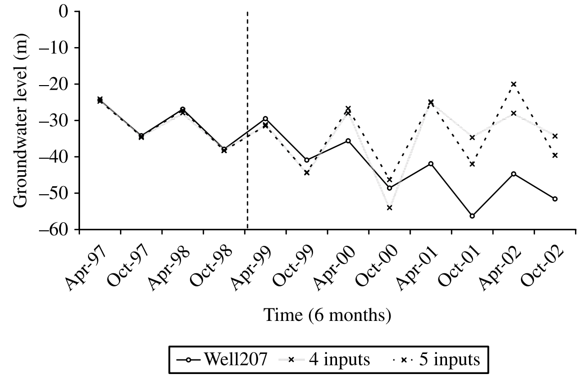

Figure 11 | Method I results for model predictions with four input variables (grey line) and five input variables (dotted line), compared with the observed data (black line).

The rest of the variables remained the same, and the dataset was divided into the following subsets:

- 1981–1995: Training/Calibration set
- 1995–1997: Cross-validation set
- 1997–2002: Testing/Predictions set

Detailed results and RMSE results for Method I for four and five variables are listed in Table 2. It can be seen that until the fifth time-step, predictions produce very good results with an RMSE of 1 m and an R^2 equal to 0.98. On the same table, Method I consistently scored an absolute error lower than 0.5 m for the first four time-steps, which is substantially better than that achieved by Daliakopoulos *et al.* (2005). This could be attributed to the earlier years added to the dataset which included information about the behaviour of the aquifer. The fact that only seasonal values were included in the input probably guided the network to

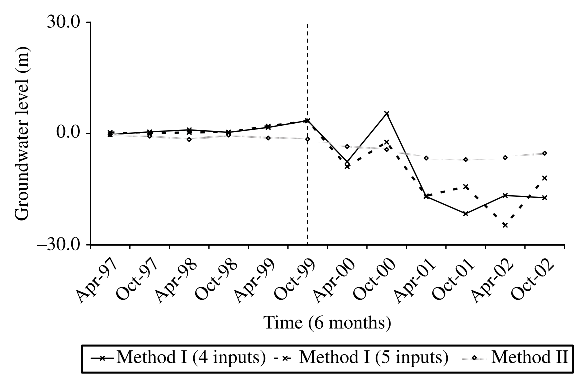

Figure 12 | Method I and II results: residuals for Method I predictions with four input variables (black line) and five input variables (dotted line) and residuals for Method II predictions (grey line).

Table 3 | Method II precipitation estimation and groundwater level predictions for the wet seasons for the period April 1998–April 2002

Time	Predicted recharge (m)	Corrected recharge (m)	Predicted precipitation (mm)	Actual precipitation (mm)	Precipitation error (mm)	% R/W error
Apr. 1998	8.3	4.5	557	564	7	1
Apr. 1999	9.3	5.1	584	485	−99	−21
Apr. 2000	7.3	4.0	533	417	−116	−28
Apr. 2001	8.8	5.0	579	633	54	9
Apr. 2002	11.7	6.1	635	540	−95	−18

focus on the groundwater level highs and lows instead of attempting to depict intermediate values which have no particular importance for our case study.

The predictions start to deteriorate after the fifth time-step or a 30-month period, increasing the RMSE to 10.6 m. Figure 11 shows the predictions results for four and five variables. In the figure, the solid black line represents the observed groundwater levels and the solid grey line the predicted values for four input variables. The dotted line represents the results when S_y was included in the input. The deterioration of Method I for longer prediction periods is also depicted in Figure 12.

It is interesting to note that for five time-step predictions, Method I with five inputs performed slightly better giving an RMSE of 0.9 m and an R^2 of 0.982. Even though in most cases the four-input application produced smaller absolute and relative errors, the overall conclusion was that the results were more consistent when five input variables were used. This can also be inferred from the validation of the method where for three, five and twelve time-steps, the R^2 scores were almost perfect. Moreover, the RMSE was kept under 0.6 m which is an insignificant quantity in our case study, where the groundwater levels at the particular time-steps were around the range of −25 m to −50 m.

Method II: new approach

The second method uses projected precipitation values based on a linear regression of past values in order to obtain more relevant data for one time-step predictions of the neural network. A description of the precipitation estimation for the wet seasons for the period April 1998–April 2002 is given in Table 3. The corrected recharge-precipitation trend line (Figure 13) is used to make rough estimations of the precipitation that caused the respective recharge. The derived precipitation value is then fed back into the ANN in order to come up with a new groundwater level prediction. The precipitation estimation has an overall standard deviation in the range of 70 mm. This can be considered acceptable since the goal of this method was to give the neural network an indication of the corresponding precipitation for each predicted aquifer recharge and not a precise value. Little

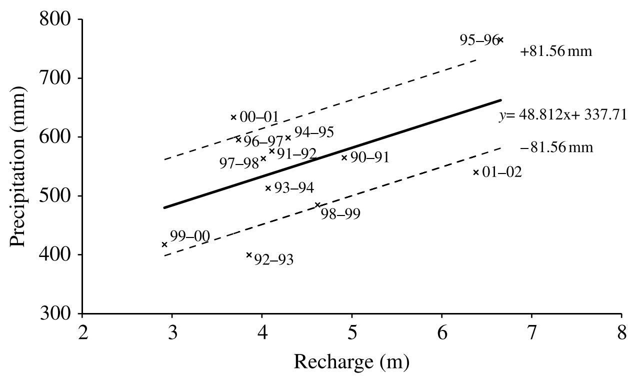

Figure 13 | Precipitation trend for the period 1992–2002 and one standard deviation (81.56 mm) above and below the precipitation trend.

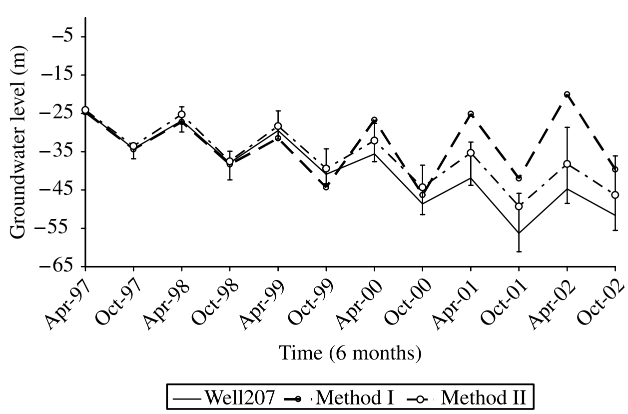

Figure 14 | Comparison of field data with the two forecasting methods: Method I predictions with five inputs (dashed line) and Method II with projected precipitation (dot-dash line).

Table 4 | Comparison of the results of Methods I with five inputs and Method II from 1997 to 2002

Time	Observed (m) Depth	R/W	Method I (five inputs) Depth	R/W	Method II Depth	R/W
Apr. 1997	−24.4		−24.7		−24.1	
Oct. 1997	−34.2	−9.8	−34.4	−9.65	−33.5	−9.4
Apr. 1998	−26.9	7.3	−27.2	7.15	−25.3	8.2
Oct. 1998	−37.9	−11.0	−38.3	−11.1	−37.5	−12.2
Apr. 1999	−29.5	8.4	−31.5	6.8	−28.3	9.2
Oct. 1999	−40.9	−11.4	−44.3	−12.8	−39.4	−11.1
Apr. 2000	−35.6	5.3	−26.7	17.6	−32.1	7.3
Oct. 2000	−48.6	−13.0	−46.3	−19.6	−44.3	−12.2
Apr. 2001	−41.9	6.7	−25.1	21.2	−35.3	9.0
Oct. 2001	−56.3	−14.4	−42.0	−16.9	−49.3	−14.0
Apr. 2002	−44.7	11.6	−20.0	22	−38.2	11.1
Oct. 2002	−51.6	−6.9	−39.6	−19.6	−46.3	−8.1

significance was given to aquifer withdrawals as they were considered relatively independent of the input of the model. The annual withdrawal of the aquifer is more relevant with the amount of irrigation in the valley; this can be considered fairly constant during the dry seasons after 1989. Predictions with Method II are compared with results obtained via Method I and the field data in Figure 14.

The observed and calculated values for the two methods are presented in Table 4. Values for recharge and withdrawal of the aquifer for each season are also listed. The results for the first five time-steps appear to be comparable. Predictions of Method II are also significantly more accurate than those of Method I after the first three–five time intervals. The comparison is also illustrated in Table 5. The skill scores and the two presented evaluation criteria show that for a few time-steps, the two approaches perform well and results are comparable although Method I has an advantage. Nevertheless, for longer prediction horizons, Method II has an obvious advantage giving an improved forecast.

Table 5 | Comparison of R^2 and RMSE results of Methods I with five inputs and Method II and the respective skill scores for different time-steps

Time-steps	Criterion	Method I (five inputs) Test		Method II Test
3	R^2	0.995		0.983
	R^2 skill		−2.4	
	RMSE (m)	0.4		1.0
	RMSE skill		−1.5	
5	R^2	0.982		0.991
	R^2 skill		0.5	
	RMSE (m)	0.9		1.0
	RMSE skill		−0.1	
12	R^2	0.205		0.960
	R^2 skill		0.9	
	RMSE (m)	10.9		4.1
	RMSE skill		0.6	

CONCLUSIONS

Using the FNN trained with the LM algorithm with five monthly variables (specific yield, precipitation, temperature, stream flow and groundwater level) from one well station of the study area, acceptable 30-month predictions of the groundwater level were achieved. Based on these methods and results, different data configurations were tested. All methods were tested for short (6 months) and extended prediction periods (30 months). The incorporation of specific yield as additional input to the model, and thus the increase of the number of input variables of the model from four to five, also had positive results. In addition, the fact that the validation of this method shows significantly better results supports the assumption that it produces equally good results for different input variables. After 30 months, the FNN method started to deteriorate producing large errors.

The new approach is a combination of the FNN with a simple deterministic component of rainfall prediction where the projected precipitation is fed into the model to predict a corrected recharge that led to a new value of projected precipitation. The method developed with this approach performed significantly better than all previous methods on longer period predictions, bridging the gap between validation results (produced through iterative training) and testing results (produced with a limited training set). The deterministic component reduced some of the uncertainty resulting from the stochastic nature of the neural network as the new approach relied on the FNN method for only one groundwater level prediction at a time.

In an area that is characterized by a scarcity of hydrogeological data, such as the Valley of Messara, a prediction with such a level of accuracy can be considered very useful for water resources management. Furthermore, the model can be used to provide predictions with a longer time horizon. This could potentially have an engineering application such as an integrated study for the fate of the groundwater level of the area under different pumping scenarios.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Ministry of Agriculture of Greece for providing the meteorological and hydrometric data. The work was supported by the Canadian National Science and Engineering Research Council (NSERC) through grants RGP157914-02 and RGP249582-02.

REFERENCES

Andreadis, K. M. & Lettenmaier, D. P. 2006 Trends in 20th century drought over the continental United States. *Geophys. Res. Lett.* **33**, L10403.

ASCE, Task 2000*a* Committee on applications of artificial neural networks in hydrology. Artificial neural networks in hydrology. I: preliminary concepts. *Hydrol. Eng.* **5** (2), 115–123.

ASCE, Task 2000*b* Committee on applications of artificial neural networks in hydrology. Artificial neural networks in hydrology. II: hydrologic applications. *Hydrol. Eng.* **5** (2), 124–137.

Balabanis, P. 1999 Water in Europe. Research achievements and future perspectives within the framework of European research activities in the field of environment, European Commission, Directorate General for Science, Brussels, Belgium. Retrieved October 2004 from www.cidob.org/ingles/publicaciones/Afers/45-46 balabanis.cfm

Coulibaly, P., Anctil, F. & Bobée, B. 1999 Hydrological forecasting using artificial neural networks: the state of the art (in French). *Can. J. Civil Eng.* **26** (3), 293–304.

Coulibaly, P., Anctil, F. & Bobée, B. 2000 Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J. Hydrol.* **230** (3–4), 244–257.

Coulibaly, P., Anctil, F. & Bobée, B. 2001*a* Multivariate reservoir inflow forecasting using temporal neural networks. *Hydrol. Eng.* **6** (5), 367–376.

Coulibaly, P., Anctil, F., Aravena, R. & Bobée, B. 2001*b* Artificial neural network modelling of water table depth fluctuations. *Water Resour. Res.* **37** (4), 885–896.

Coppola, E., Poulton, M., Charles, E., Dustman, J. & Szidarovszky, F. 2003*a* Application of artificial neural networks to complex ground water management problems. *J. Nat. Resour. Res.* **12** (4), 303–320.

Coppola, E., Szidarovszky, F., Poulton, M. & Charles, E. 2003*b* Artificial neural network approach for predicting transient water levels in a multilayered ground water system under variable state, pumping, and climate conditions. *J. Hydrol. Eng.* **8** (6), 348–359.

Coppola, E., Rana, A., Poulton, M., Szidarovszky, F. & Uhl, V. 2005 A neural network model for predicting water table elevations. *J. Ground Water* **43** (2), 231–241.

Croke, B., Cleridou, N., Kolovos, A., Vardavas, I. & Papamastorakis, J. 2000 Water resources is the desertification-threatened Messara valley of Crete: estimation of the annual water budget using a rainfall-runoff model. *Environ. Model Software* **15**, 387–402.

Daliakopoulos, I. N., Coulibaly, P. & Tsanis, I. K. 2005 Groundwater level forecasting using artificial neural networks. *J. Hydrol.* **309**, 229–240.

Feng, S., Kang, S., Huo, Z., Li, W., Chen, S. & Mao, X. 2008 Using neural network to simulate regional ground water affected human activities. *J. Ground Water* **46** (1), 80–90.

French, M. N., Krajewski, W. F. & Cuykendall, R. R. 1992 Rainfall forecasting in space and time using a neural network. *J. Hydrol.* **137**, 1–31.

Govindaraju, R. S. & Ramachandra Rao, A. 2000 *Artificial Neural Networks in Hydrology*. Kluwer Academic Publishing. The Netherlands.

Jain, A. & Kumar, A. M. 2007 Hybrid neural network models for hydrologic time series forecasting. *Appl. Soft Comp.* **7**, 585–592.

Kim, G. & Barros, A. P. 2001 Quantitative flood forecasting using multisensor data and neural networks. *J. Hydrol.* **246**, 45–62.

Kovacs, G. 1981 *Subterranean Hydrology*. Water Resources Publications, Colorado.

Lehner, B., Henrichs, T., Döll, P. & Alcamo, J. 2001 EuroWasser–Model-based assessment of European water resources and hydrology in the face of global change. Kassel World Water Series 5, Center for Environmental Systems Research, University of Kassel, Kurt-Wolters-Strasse 3, 34109 Kassel, Germany.

Meinzer, O. E. 1923 Outline of the groundwater hydrology, with definitions. U.S. Geological Survey Water-Supply Paper 494.

Solomatine, D. P. 2002 Data-driven modelling: paradigm, methods, experiences. In: *Proceedings of 5th International Conference on Hydroinformatics*. Cardiff, UK, July 2002, pp.757–763.

Stringfield, V. T. & LeGrand, H. E. 1971 Effects of Karst features on circulation of water in carbonate rocks in coastal areas. *J. Hydrol.* **14**, 139–157.

Vardavas, I. M., Papamastorakis, J., Fountoulakis, A. & Manousakis, M. 1997 Water resources in the desertification-threatened Messara Valley of Crete: estimation of potential lake evaporation. *Ecol. Model* **102**, 363–374.

WWRP/WGNE Joint Working Group on Verification 2005 Forecast Verification - Issues, Methods and FAQ. Retrieved May 2006 from www.bom.gov.au

Zealand, C., Burn, D. & Simonovic, S. 1999 Short term streamflow forecasting using artificial neural networks. *J. Hydrol.* **214**, 32–48.

First received 14 December 2007; accepted in revised form 23 June 2008