PREDICTION OF FUTURE GROUND WATER ELEVATION BY USING ARTIFICIAL NEURAL NETWORK SOUTH OF RIYADH, KSA (CASE STUDY)

*Yasser Hamed¹, Mohamed Elkiki², and Osama S. Al Gahtani³,

¹(Current position: visiting associate professor, Faculty of Engineering, Salman Bin Abdulaziz University, KSA), Civil Engineering Department, Faculty of Engineering, Port Said University, Port Fouad Port Said, Egypt - ²Civil Engineering Department, Faculty of Engineering, Port Said University, Port Fouad Port Said, Current position: visiting associate professor, Faculty of Engineering, AlGouf University, KSA,) ³Civil Engineering Department, Faculty of Engineering, Salman bin Abdulaziz University, Riyadh, Saudi Arabia

ABSTRACT

Saudi Arabia is considered an arid country with very limited water supplies. About 80% of water supplies come from nonrenewable water resources; ground water. Natural replenishment of the ground water is far less than the annual extraction of ground water. Consequently, ground waters elevations have sharply decreased through the last two decades. Neural Network has proved to be an effective tool for future prediction for a set of data. For a set of data of ground water elevations for a reasonable period of time Artificial Neural Network (ANN) can be effectively predict future ground water levels. The main objective of the current study is to predict the elevations of ground water in four wells south of Riyadh capital city for 20 years in future. In order to achieve that, nearly daily data of water level for period more than 30 years have been used for the prediction. A complete analysis of the 30 years data has been conducted. The results gave a warning to the governments in order to begin a general strategy in the area to reduce using excessive water for irrigation. The ANN was used to develop a prediction model to predict the water elevation in the well. About 80% of the data are used for training the network while the rest of the data are used for validating and testing the model. A sigmoid function is used at the hidden layer which consisted of 1-6-1. Results from both measurements and ANN were compared which gives a very good matching. The future prediction offered by the ANN says that the future reduction of the well water level in the coming 20 years will not exceed 30% of the reduction done during the last 30 years.

1. INTRODUCTION

Extensive use of groundwater including the non-renewable part has been heavily practiced to meet the increasing demands especially during the last three decades in several countries in arid and semi-arid regions. Among these, the USA, Australia, Spain, India, Jordan, Oman, Libya, Bahrain, the United Arab Emirates, Egypt and Saudi Arabia support domestic, agricultural and domestic activities. The impacts of these experiences on the sustainability of groundwater resources and on the economic and social sectors vary among countries (Abderrahman, W. 2005).

Countries within the Arab region are heavily depending on irrigation for food production and they are confronted by severe problems due to both climatic conditions and socioeconomic factors (ACSAD, 1999). From an ecoclimatic point of view, most of the region extends across semiarid and arid zones. The semiarid belts have been particularly affected by cycles of droughts and desertification in the past decades. Socioeconomically, the region is characterized by a fast increasing population,
which has resulted in a sharp decline of the per capita availability of water during latest decades. At the same time, supplies of good quality irrigation water are expected to sharply decrease in the near future because the development of new water supplies will not keep pace with the increasing water needs of industries and municipalities.

Saudi Arabia is considered the largest arid country in the Middle East which is located in extremely arid regions where the average annual rainfall ranges mostly from 25 to 150mm (Ministry of Agriculture and Water, 1988). The average annual evaporation ranges from 2500 to about 4500 mm. The country has an area of 2.25 million km², of which about 40% is desert lands. It lies within latitudes 16 and 32 N and longitudes 34 and 55 E.

With low annual precipitation and with no lakes, rivers or streams, Saudi Arabia faced the rising water demand by increasingly depending on ground water. About 80% of the total water supply comes from groundwater with receives very limited natural recharge and is, thus classified as fossil or nonrenewable water (Al-Ibrahim, A.A. 1991).

A variety of major problems including fast depletion of groundwater resources and deterioration of their quality can be occurred when the rate of water withdrawal exceeds the net recharge.

KSA followed a new strategy based on controlling aquifer development and demand management to use its groundwater resources. Corrective demand management measures including reduction in cultivated areas and modification in agricultural support policies in addition to the augmentation of water supplies by the reuse of treated wastewater have reduced the stress on groundwater (Abderrahman, W. 2005).

The ANN’s have been applied in almost all branches of science with good satisfaction during the last three decades. Both of Dibiki et al. 1999a,b and Dibiki and Abbott 1999 have explained basic information on ANN and their applications in the field of Hydraulic Engineering.

Liriano and Day 2001 have applied ANN to predict scour downstream of hydraulic structures such as culvert. Kheireldin 1999 predicted scour at bridge abutment where Negm et al.2002b predicted scour downstream of sudden expanding stilling basins by using ANN also Elkiki 2008a used the ANN to predict scour downstream skew siphon pipes.

Negm and Shouman 2002 and Negm et al.2002a have managed to Predict the hydraulic jump characteristics by using ANN where the characteristics of submerged H.J. in non-prismatic stilling basin has been predicted by Negm et al. 2003.

Owais et al. 2003 have succeeded to predict the characteristics of free radial H.B.J. formed at sudden drop. Application of ANN to predict discharge below sluice gate under free and submerged flow conditions was made by Negm 2002 and to predict discharge below submerged gate with sill was done by Negm 2001.

Application of ANN to predict the drawdown depth due to ship movement was introduced by Elkiki 2008a,b also the mean velocity through open channel with submerged aquatic weeds by using ANN was made by Mirdan 2012.

2. METHODOLOGY

2.1 Sites Location

The study area is considered one of the important agricultural areas in KSA. It is located south of Riyadh City. Around 26 % of the crops productions of the kingdom come from this area. It consists of
four mean wells. All of the four wells are distributed in the study area (Fig (1)). Table (1) shows the longitude and latitude of each well and its distance far from city Al Kharj (80Km south of Riyadh city).

<table>
<thead>
<tr>
<th>Well</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Distance to AlKharj city</th>
</tr>
</thead>
<tbody>
<tr>
<td>C (5-K-60)</td>
<td>47° 47’</td>
<td>24° 15’</td>
<td>45 Km</td>
</tr>
<tr>
<td>D (5-K-83)</td>
<td>47° 00’</td>
<td>24° 04’</td>
<td>30Km</td>
</tr>
<tr>
<td>E (5-K-84)</td>
<td>47° 10’</td>
<td>24° 04’</td>
<td>25 Km</td>
</tr>
<tr>
<td>F (5-k-87)</td>
<td>47° 32’</td>
<td>24° 14’</td>
<td>20 Km</td>
</tr>
</tbody>
</table>

2.2 Data Collection

The wells data used in this paper have been obtained by personal contact from water sector, ministry of water and electricity. The number of water elevation readings for each well is ranging from 6000 to more than 8000 for 23 years. Table (2) shows the data time range and the change in ground water surface elevations during this range.

<table>
<thead>
<tr>
<th>Well</th>
<th>Data time range</th>
<th>Change in ground water surface elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>C (5-K-60)</td>
<td>1978-2011</td>
<td>-70.23m to -116.13m</td>
</tr>
<tr>
<td>D (5-K-83)</td>
<td>1978-2011</td>
<td>-25.93m to -140.57m</td>
</tr>
<tr>
<td>E (5-K-84)</td>
<td>1978-2011</td>
<td>-29.85m to -69.12m</td>
</tr>
<tr>
<td>F (5-k-87)</td>
<td>1978-2010</td>
<td>-56.86 m to -106.88 m</td>
</tr>
</tbody>
</table>

2.3 Background of ANN Software

At their most fundamental level, neural networks are simply a new way of analyzing data. What makes them so useful is their ability to learn complex patterns and trends in data, an ability that is unique to neural networks. Conventional computing techniques are very good at arithmetic. They can add columns of numbers, check spellings against a dictionary, and so on. But when they are asked to do tasks such as recognizing a face, assessing a credit risk, or forecasting demand, they perform very badly. Most of these tasks are easy for humans, indeed some we do without conscious effort. We can recognize the face of a friend in a crowded, even when we haven’t seen that friend for many years; stoke market traders may make mistakes, but they are right often enough to make a profit. Unfortunately, many of the problems faced by businesses have more in common with subjective tasks, those at which computers traditionally perform badly, than with arithmetic tasks.

In order to emulate the human ability to solve this category of problem, neural computing has abandoned conventional computing techniques, and has instead concentrated on the way that biological systems, such as human brains, work. The human brain is made of many neurons, the brain cells, each of which is connected to many others in a network that adapts and changes as the brain learns. In neural computing, processing elements replace of the neurons, and these processing elements are linked together to form neural networks. Each processing element performs a simple task; it is the connections between the processing elements that give neural networks the ability to learn patterns and interrelationships in data.
By producing systems that learn the relationships between data and results, neural networks avoid many of the problems of conventional computing. Given new, unseen data, a neural network can make a decision based upon its experience. Until recently the benefits of neural networks were only available to people who could afford to spend time and effort becoming experts in neural computing. Neural connection has been designed so that one can harness this technology immediately, without needing to become a neural computing expert (ElKiki 2008b).

2.4 How is Neural Connection Structured?

Neural Connection Software (Neural Connections (1998)) is made up of 3 separate modular groups. The modules are the graphical user interface, the executive and the tool modules. The modular design of Neural Connection gives greater flexibility than standard analysis tools, and allows to adapt the application as required, in order to gain the best results from the data.

The graphical user interface provides the Neural Connection workspace, the program window where one builds problem-solving applications. The workspace allows to build, train, and run analysis applications. An application is a method for modeling the problem, which may include inputting the data, processing them in some way, and producing the output in a useful format. Tools are selected as icons from a palette, and moved onto the workspace, where they can be connected to other tools.

These connections determine an application’s topology, and the path along which data flows. The workspace has tools for inputting data, statistical analysis, problem modeling, and producing results. Each type of tool performs a specific type of task, and by combining tools in different ways, applications can be built that performs much more sophisticated analysis than any one technique could by itself. A tool which you have placed on the workspace has a series of parameters that can be set independently of other tools, even those of the same type.

To customize these parameters, one can open each tool. Each tool has a series of simple commands such as Run, Train, and Dialog, which enable it to be used in a straightforward manner.

The executive is the core component of Neural Connection. It creates instances of the analysis tools, displays them on the graphical user interface, and manages the data flow between them. It also loads, trains, and runs applications developed in the script language.

The tool modules used by Neural Connection consist of neural network and data analysis routines that have been extensively developed. They include tools for data input and output, data manipulation, statistical analysis, classification, and modeling and forecasting. In addition to providing the standard user-defined parameters, the neural network tools contain many advanced features, such as automatic configuration and training. The common characteristics of the analysis tool are:

- Tools store all their data variables in a data body in a computer’s memory. This allows different instances of a tool to be created by allocating new data bodies. The software associated with each tool is re-entrant, allowing it to be used by all instances of the tool.

- Tools pass data from one to another via a single logical connection. This connection passes both vector data and control information, such as the names of fields in the vector and their data types. Vector data can contain both input information and target information.

- Tools make no assumptions about which tools are connected in front of, or after them in the topology. If a tool uses the same data set many times (for example in Multi-Layer
Perception (MLP training), then it will automatically hold the data required in memory until it is no longer needed. (ElKiki 2008)

2.5 Building the ANN Models

The following systematic steps are followed to build each of the developed ANN model for the present study:

a – The data set of the specified problem are prepared in a file with the format suitable for the Neural Connection software such as Microsoft excel 5.

b – The Neural Connection is operated and the needed tools from the toolbar are connected together to form general topology of the network.

c – The input tool is used to specify and allocate the input data. In this stage the order of the data is randomized and divided into three subsets, training data (80%), validation data (10%) and test data (10%). All of them are selected randomly.

d – The MLP tool is used to specify the following parameters:

1. - A Zero mean unit standard deviation for the normalization method of the input fields.

2. - A Sigmoid activation function which is obtained by trail and error.

3. - 10 for the number of neurons in hidden layer which is obtained by trail and error so that the best size of the network is 1-6-2.

4. - A Zero mean unit standard deviation for the normalization method of the target fields.

5. - 0.1 for the initial weights with seed of 3 of the links between neurons obtained by trail and error.

6. - A conjugate gradient learning algorithm.

7. - 200 for the maximum updates.

e – The network is allowed to train and the validation system error is observed. Once the minimum validation system error is reached, the training is stopped. The output tool is used to save the output results for training, validation and test in three separate files.

f – The output files are transferred to the Microsoft Excel software to enable other plotting of results, comparisons and computations of the correlation coefficient (R) and the Root Mean Square Error (RMSE) between the measured and the predicted outputs.

g – The network is allowed to train 5 times at least with different starting point in each time until a global solution is obtained.

The procedure is repeated for each of the developed network. The results of optimal networks are used when comparing the ANN results with the measured and MLR results (ElKiki 2008b).
3. RESULT AND DISCUSSION

3.1 Effect of Size of Input Data Set.

The different percentages that are investigated to study the effect of number of training data on the variation of validation errors RMSE of the output variables are shown in figure (2). The network should be calculated several times with different percentages until the maximum R for the validation data occurs. The figure shows that a network of percentages 80-10-10 is suitable, i.e. 80% of data set for training the network; 10% for validation and 10% for test.

3.2 Effect of Number of Neurons of the Hidden Layer.

The best number of neurons of the hidden layer is found to be 6 neurons. To ensure about this size, the number of neurons in the hidden layer are changed several times and the errors from the network are investigated. Figure (3) shows the variations of the correlation coefficient (R) of the different outputs of the networks versus the number of neurons of the hidden layer. From these figures, it is clear that 6 neurons in hidden layer yields the maximum R for output variables considering the training, validation and test errors. This because, as the number of neurons increases, the errors decrease up to a minimum at a size of 6 neurons and then the errors start to increase again.

3.3 Effect of Single Output Neurons versus Multiple Output Neurons.

Study the effect of using output layer having one neuron representing one output variable compared to output layer having three neurons were performed. Two networks, each of a size 1-6-1 are used to model the two targets of the present study with the same learning parameters used in the network of size 1-6-2. Figures (4) and (5) represent the RMSE (SEE) and correlation coefficient (R) for the two variables using single and multiple output layers for two computer runs having the best results. The Figures show that the performance of the network of size 1-6-2; i.e the two variables taken together; is better than those of size 1-6-1 in predicting the different water levels in water wells.

3.4 Effect of Activation Function.

The effect of activation functions on the resulting error from the best network is studied by performing three computer runs using the best ANN of 1-6-2 utilizing different activation function assuming other parameters unchanged. Figure (6) presents the values of RMSE for each of the three data sets using three activation functions, hyperbolic tangent (tansh), sigmoid and linear functions. From the figure, it is clear that the sigmoid function is more suitable to predict the outputs of the problem.

3.5 Effect of Seed Number (Stability of the Network).

Multiple runs were conducted with training, validation and test sets in order to study the performance and stability of the ANN model. The data sets were chosen randomly. This process was repeated four times. The results in terms of the RMSE and in terms of R are presented in figures (7 a,b,c and d). The mean values of the RMSE of the 4 random sets are 1.923265 for training, 1.227626 for validation and 1.325382 for test with standard deviation of 0.038519, 0.076531 and 0.082124 respectively. Also, the mean correlation coefficients are 0.99453 for training, 0.996767 for validation and 0.993255 for test with standard deviation of 0.003058, 0.000372 and 0.00415 respectively. Clearly, it is indicated that the values of RMSE and R are mostly around the mean values with very
small standard deviation. This means that the developed ANN model is stable and hence the data are consistent.

3.6 Outputs of Artificial Neural Networks

Output results of ANN compared to measured ones are presented in figures (4,5,6, and 7). The figures show examples of the results of train, validation and test data for minimum and maximum reading for the four testing wells. An excellent matching is observed in all cases indicating good performance of the ANN model in predicting the ground water depth.

The correlation coefficients between the measured and predicted water depth using ANN model are presented in table 3 for training, validation and test data sets. Also, as shown in figures 12 (a,b) a very good matching between training, validation and test data for all 4 wells is observed which indicates a good performance of the ANN model in predicting the ground water depth.

Table 3: The values of correlation coefficients between the measured and predicted water depth for different wells

<table>
<thead>
<tr>
<th>Well Name</th>
<th>Data Value</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
</tr>
<tr>
<td>5 – k – 60</td>
<td>Min.</td>
<td>0.932741</td>
</tr>
<tr>
<td></td>
<td>Max.</td>
<td>0.939142</td>
</tr>
<tr>
<td>5 – k – 83</td>
<td>Min.</td>
<td>0.966414</td>
</tr>
<tr>
<td></td>
<td>Max.</td>
<td>0.965854</td>
</tr>
<tr>
<td>5 – k – 84</td>
<td>Min.</td>
<td>0.992405</td>
</tr>
<tr>
<td></td>
<td>Max.</td>
<td>0.989395</td>
</tr>
<tr>
<td>5 – k – 87</td>
<td>Min.</td>
<td>0.988355</td>
</tr>
<tr>
<td></td>
<td>Max.</td>
<td>0.982605</td>
</tr>
</tbody>
</table>

3.7 Future Prediction

Fig (13) shows the relation between first, last and estimated readings after 20 years with minimum and maximum water levels. The first readings were collected from about 32 years ago, while the last readings were collected at recent years. Future prediction of well level of well 5-k-87 is about (-119.00) while in wells 5-k-84, 5-k-83, 5-k-60 are (-80.00), (-160.00) and (-126.00) respectively. The parentage of water level reduction after 20 years comparing with the reduction during the last 30 years will be only 24 %, 29%, 17% and 23% for wells 5-k-87, 5-k-84, 5-k-83 and 5-k-60 respectively. It means that the future reduction of the well water level in the coming 20 years will not exceed 30% of the reduction done during the last 30 years.
4 SUMMARY AND CONCLUSION

A ground water level prediction has been conducted by using The Artificial Neural Networks ANN. The ANN is used to develop a network of size 1-6-2 to predict the water level after 20 years. Daily data of water level for period more than 30 years have been used for the prediction. The data was collected from four wells located south of Riyadh city. A complete analysis of the 30 years data has been conducted. Results of the prediction show that the reduction of the ground water level during the coming 20 years will be no more than 30% of the reduction done through the last 30 years. However, results still give warning to the local authorities in order to begin to conduct a reliable strategy in order to reduce the future consumption of ground water.

ACKNOWLEDGMENT

This work was funded by the Deanship of scientific research in Salman bin Abdulaziz University, KSA through a scientific research project under the title "Environmental assessment for the ground water in agricultural area, El kharj", No (2/T/33).

REFERENCES


ACSAD, 1999. Water resources division; objectives, functions and activities, The Arab Center for the Studies of Arid Zones and Dry Lands, Damascus, Syria.


Figure 1. Wells Locations

Figure 2. Effect of percentages of input data on R of the training, validation and test data outputs for max. and min. readings.
Figure 3. Effect of No. of neurons of hidden layer on R of the ANN outputs for max. and min. readings.

Figure 4. Effect of single and multiple output neurons on RMSE of the ANN outputs for max. and min. readings.

Figure 5. Effect of single and multiple output neurons on R of the ANN outputs for max. and min. readings.
Figure 6. Effect of activation function of hidden layer on RMSE of the ANN outputs for max/ and min. readings.

Figure 7. Results of studying the stability of the network, (a,b) RMSE and (c,d) R of the ANN outputs for max. and min. readings.
Figure 8. Examples of results of the predicted ANN model as compared with measured water level for different (training, validation) data sets for min readings. (5-k-87)

Figure 9. Examples of results of the predicted ANN model as compared with measured water level for different (training, test) data sets for max. readings. (5-k-84)

Figure 10. Examples of results of the predicted ANN model as compared with measured water level for different (validation, test) data sets for min. readings. (5-k-83)

Figure 11. Examples of results of the predicted ANN model as compared with measured water level for different (validation, test) data sets for max. readings. (5-k-60)
a - For min. readings

Figure 12. The value of R of the ANN outputs for the different well name for (a) min. and (b) max. readings.

a - For max. readings

Figure 13. The relation between first, last and estimated readings after 20 years with min. and max. water levels.