IJST (2015) 39A4: 463-472
Iranian Journal of Science & Technology

http://ijsts.shirazu.ac.ir

An artificial neural network model for the determination of leaky confined aquifer parameters: an accurate alternative to type curve matching methods

T. Azari[1], N. Samani[1]* and E. Mansoori[2]

[1]*Department of Earth Sciences, Shiraz University, Shiraz 7146713565, Iran*
[2]*School of Engineering, Shiraz University, Shiraz, Iran*
E-mail: samani@susc.ac.ir

Abstract

A neural network is developed for the determination of leaky confined aquifer parameters. Leakage into the aquifer takes place from the storage in the confining aquitard. The network is trained for the well function of leaky confined aquifers by the back propagation technique and adopting the Levenberg–Marquardt optimization algorithm. By applying the principal component analysis (PCA) on the adopted training input data and through a trial and error procedure the optimum structure of the network is fixed with the topology of [2×10×2]. The network generates the optimal match point coordinates for any individual real pumping test data set which are incorporated with Hantush's analytical solution and the aquifer parameter values are determined. The performance of the network is evaluated by real field data and its accuracy is compared with that of the type curve matching technique. The network eliminates graphical error inherent in the type curve matching technique and is recommended as a simple and reliable alternative to the type-curve matching technique.

Keywords: Aquifer parameters; aquitard; well function; artificial neural network; pumping test

1. Introduction

Figure 1 shows a well penetrating an artesian aquifer overlain by an aquitard and underlain by an aquiclude. Overlying the aquitard are permeable deposits (source bed) in which there is a water table. The aquifer is homogeneous, isotropic, infinite in areal extent and is of the same thickness throughout. The well completely penetrates the aquifer, and flow in the aquifer is radial throughout. Flow in the aquifer is augmented by vertical leakage through the aquitard. The flow lines are assumed to be refracted a full right angle as they cross the aquitard-aquifer interface. The discharge of water from a well in such an aquifer is supplied from storage within the aquifer and aquitard, as well as from leakage through the aquitard. Because of the presence of recharge in the form of leakage, water levels will stabilize when the entire discharge of the well is derived from leakage within the aquitard.

Assuming that water is released from storage instantaneously with a decline in head, the differential equation governing the unsteady-state flow in the leaky artesian aquifer in polar-coordinate notation is (Hantush, 1960):

$$\frac{\partial^2 s}{\partial r^2} + \frac{1}{r}\frac{\partial s}{\partial r} + \frac{K'\partial s}{T\partial z} = \frac{S}{T}\frac{\partial s}{\partial t} \quad (1)$$

where s is the drawdown measured at time t (T) in an observation well located at the distance r (L) from the pumping well, T (L^2T^{-1}) is transmissibility, S is aquifer storage coefficient, K' (LT^{-1}) is the hydraulic conductivity of the aquitard. Note that the third term in the left hand side of Equation (1) represents leakage into the aquifer through the aquitard. Hantush (1960) solved Equation (1) with appropriate boundary and initial conditions and derived an analytical solution for the drawdown due to the pumping of a fully penetrating well:

$$s = \frac{Q}{4\pi T} W(u,\psi) \quad (2)$$

where

$$u = \frac{r^2 S}{4Tt} \quad (3)$$

$$\psi = \frac{r}{4}\sqrt{\frac{S'K'}{TSb'}} \quad (4)$$

$$W(u,\psi) = \int_u^\infty \frac{e^{-y}}{y}\,\mathrm{erfc}\,\frac{\psi\sqrt{u}}{\sqrt{y(y-u)}}\,dy \quad (5)$$

*Corresponding author
Received: 9 March 2015/ Accepted: 16 June 2015

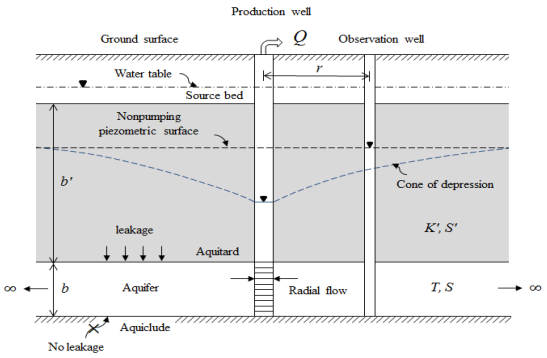


Fig. 1. Leaky confined aquifer with fully penetrating pumping well

S' and b' are the storativity and thickness of the aquitard, respectively and $\text{erfc}\,(x)$ is the complementary error function defined by:

$$\text{erfc(x)} = 1 - \text{erf(x)} = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-y^2)\,\text{dy} \qquad (6)$$

Hantush (1961) tabulated values of well function $W(u,\psi)$ that is plotted against $1/u$. Fig. 2 shows the resulting family of type curves, with each curve having its own ψ value.

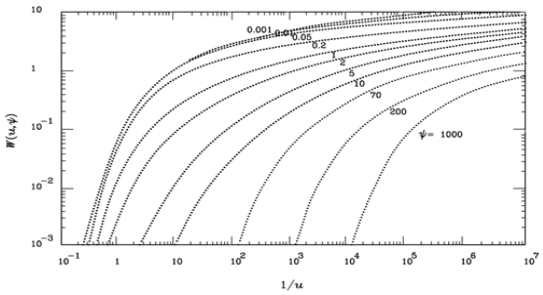


Fig. 2. Family of Hantush's type curves (1961), $W(u, \psi)$ versus $1/u$ for different values of ψ

To determine the leaky confined aquifer parameter values (T, S, K', S'), time-drawdown data recorded in a pumping test are plotted on logarithmic paper of the same scale as for the Hantush's type curves. The time-drawdown field data curve is superimposed on the type curves, keeping the coordinate axes of the two plots parallel and adjusted until plotted points of observed data match on one of the type curves. A match point is selected and its coordinates on both plots is recorded $[1 / u_{\text{m}}, W(u, \psi)_{\text{m}}, s_{\text{m}}, t_{\text{m}}]$. With values of match point coordinates thus determined, aquifer parameters are obtained from Equations (2-5).

The aquifer parameters obtained by the type-curve graphical method are rather subjective due to graphical and personal errors and the accurate results are usually not obtained. In recent years, some suitable approaches based on artificial neural networks (ANNs) have been developed as an alternative approach to model the well functions and remove the errors resulted from graphical type curve matching techniques (Lin and Chen, 2005; 2006; Samani et al., 2007; Lin et al., 2010).

The idea of ANN was first proposed in the 1940s (McCulloch and Pitts, 1943). An ANN is an interconnected group of simple processing elements (artificial neuron or nodes) analogous to the network of neurons in the human brain. ANN is a parallel processor that uses a mathematical model for information processing based on a connectionist approach to computation. This method can model complex non-linear relationships between inputs and desired outputs through a training process to find patterns in data that are not easily analyzed using conventional methods. Because of this property ANNs have been successfully used in hydrological problems. The ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000a; 2000b) described various aspects of ANNs and reviewed many articles on ANN applications in various branches of hydrology. In Groundwater hydrology, Lin and Chen (2005) proposed an ANN approach to estimate aquifer parameters of leaky confined aquifers based on a combination of a Radial Basis Function Network (RBFN) and Hantush and Jacob (1955) analytical solution. Lin and Chen (2006) also suggested the combination of an ANN and the Theis (1935) analytical solution in confined aquifers. The drawback of Lin and Chen networks is that as the number of time-drawdown data increases, the dimensionality of networks becomes larger and they also have to be trained and tested for each individual set of pumping test data (Samani et al., 2007) Accordingly, Samani et al. (2007) proposed a simple ANN by replacing the gradient descent algorithm with the faster Levenberg-Marquardt (LM) training algorithm and applying Principal Component Analysis (PCA) on the training data set to estimate confined aquifer parameters without the aforementioned limitations. Later Lin et al. (2010) applied the PCA on the training and testing data patterns in the development of an ANN for the parameter estimation of anisotropic aquifers.

The objective of this paper is to construct an ANN that is trained to model the Hantush's well function (Equation (5)) and predict the coordinates of match point from which the aquifer parameters are determined. The developed network is a single-hidden-layer feed-forward neural network with 2, 10 and 2 neurons in the input, hidden and output layers, respectively. The model development processes consist of a six step protocol similar to that suggested by Wu et al. (2014) and the methodology used in each step is documented. The accuracy of network was tested by 100000 synthetic error-free time-drawdown data sets. The

applicability, adequacy and validity of the developed network is evaluated using one set of real pumping test data and results are compared with that of type curve matching technique. The developed network is written in Matlab program called (ANN4LAPE, Artificial Neural Network for Leaky Aquifer Parameter Estimation) and is available from the corresponding author upon request. The program receives pumping test data and provides the user with the aquifer parameter values, i.e. S, T, K' and S'. The network appears to be an efficient, accurate and easily used alternative tool to the type-curve graphical method in the hand of practitioners for the determination of parameters of leaky confined aquifers with water released from storage in the aquitard.

2. Modeling Strategy

For the development of an ANN model for a hydrological system, by reviewing over 81 papers published from 2000 to 2012 in the field of drinking water quality Wu et al. (2014) argue "despite the recognition of the importance of the adoption and articulation of rational ANN model development procedures, there is no comprehensive protocol for the development of ANN models" and suggest the following six main steps extracted from ten steps proposed by Maier et al. (2010). These are: input selection, data splitting, selection of model architecture, determination of model structure, model calibration/training (optimization of model parameters) and model validation. By tracing two decades of neural network rainfall-runoff and stream flow modeling, Abrahart et al. (2012) similarly state that more objective and consistent protocol is needed for the development of ANNs. While these steps logically establish modeling procedure, depending on the nature of system that is to be modeled and the modeling objectives (i.e. forecasting, classification, function approximation, etc.) one or more steps may be combined or not carried out. In this paper we tried to follow the above logical steps with some minor modifications to design an ANN based on the combination of a feed-forward neural network and the exact Hantush's analytical solution for the determination of leaky confined aquifer parameters. The developed network generates the match points coordinates $\hat{y}_1 = \log\left(\frac{1}{u_\mathrm{m}}\right)$ and $\hat{y}_2 = (\psi)_\mathrm{m}$ as output for any individual real (field) pumping test data set received as input.

2.1. Generation and Selection of Input Data Patterns

To generate (to select) the training data patterns for the network design, first a logical range of $\log(1/u_\mathrm{m})$ (from -0.5 to 7.0), and of $(\psi)_\mathrm{m}$ (from 0.0 to 10.0) according to the type curve presented by Hantush (1961) and as inferred from Fig. 2 are selected as the ANN output targets that represent the aquifer behavior. Using interval values of 0.0073 and 0.2041 for $\log(1/u_\mathrm{m})$ and $(\psi)_m$, respectively, 51200 sets of training input pattern each of which constitutes N-1 elements (N is the number of drawdown-time records) were generated as illustrated in Fig. 3. Therefore, the size of the input data matrix is [(N-1) × 51200]. Then for all these sets of $[\log(1/u_\mathrm{m}),\ (\psi)_m]$, well function $W(u,\psi)$ is calculated by Equation (5) and then the training input patterns X_i are generated by the following equation as illustrated in Fig. 3:

$$X_\mathrm{i} = \log\left[\frac{W(u_\mathrm{m} t_1/t_{\mathrm{i}+1},\ (\psi)_\mathrm{m})}{W(u_\mathrm{m},\ (\psi)_\mathrm{m})}\right] \qquad (7)$$

Subscript m denotes the match point, i=1, 2,…, N-1 and N is the number of time-drawdown records.

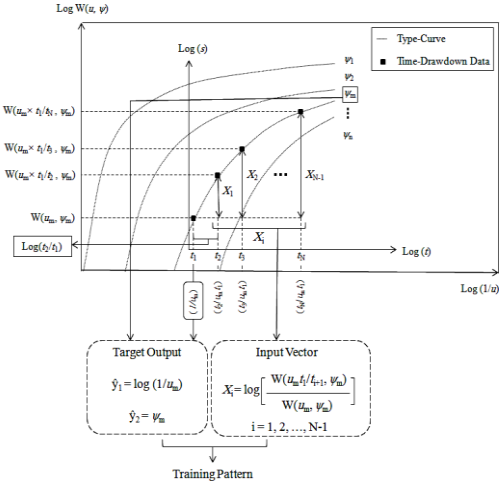


Fig. 3. Graphical presentation of input vectors and target outputs generation for the training of the ANN

Before using the generated training data patterns as input vector to the ANN, their significance and the independence needs to be examined (Maier et al., 2010). Therefore, the training data patterns are normalized and subjected to PCA to reduce their dimensionality and eliminate redundant data. PCA transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first PC), and the second PC explains the maximum variances of the residual data set and so on. When the analysis is complete, the resulting components will display varying degrees of correlation with the observed variables, but are completely uncorrelated

with one another. One of the parameters in PCA method is minimum fraction variance. By specifying a minimum fraction variance one can eliminate those principal components that contribute less than this value to the total variation in the data set and hence the dimensionality of the data set is reduced with no information lost (Davis, 2002). Samani et al. (2007) applied PCA successfully to reduce the dimension of the input vector of the original data and develop a network with fixed structure for accurate determination of confined aquifer parameters.

Table 1 provides PCA parameters for the training patterns X_i (generated by Equation (7)). The principal components were extracted by considering minimum fraction variance of 0.005 (0.5%). The result of PCA given in Table 1 shows that the first and second principle components together describe 99.996% of the variance of the training data sets. Because the variance accounted by the third component is smaller than minimum fraction variance (i.e., 0.0044 < 0.5%) it is ignored. Moreover, the results of the Scree test, Cattell (1966), as shown in Fig. 4, confirm that only the two first principle components should be considered as input vectors. This means that the elements of training input vector reduces to 2 and the number of neurons in the input-layer of the network to be designed can be fixed to 2 neurons instead of number of records in drawdown-time data which varies from one pumping test to another.

Table 1. The principal component parameters of the training set for the leaky confined aquifer

Principal component	Eigenvalue	Variance (%)	Cumulative variance (%)
PC1	35.5746	98.81833	98.81833
PC2	0.4238	1.177222	99.99556
PC3	0.0016	0.004444	100
PC4	2.4506e-06	0	100

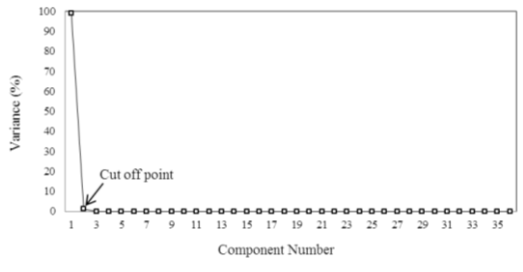


Fig. 4. The scree plot of the training patterns

2.2. Selection of the Network Architecture

A typical ANN consists of an input layer, a number of hidden layers and an output layer each having a number of processing neurons (nodes). The number of neurons in input and output layers is determined by the number of input and output variables, respectively. The number of hidden layers and their neurons are determined in the process of model training (calibration) usually by a trial-and-error procedure. The pattern of connection between nodes, the method of determining the connection weights, and the activation function characterizes the architectures of the network (Fausett, 1994; ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a). ANNs are also categorized based on the direction of information flow and processing (Maier et al., 2010). For instance, in a feed-forward neural network (Hornik et al., 1989), also known as multi-layer perceptrons (MLP) information passes from the input nodes to the output nodes only. This is in contrast to a recurrent ANN in which information flows through the nodes in both directions, from the input to the output nodes and vice versa (Karamouz et al., 2008). Feed forward neural networks are the most commonly used ANN architecture (ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a; Maier et al., 2010; Razavi and Tolson, 2011; Wu et al., 2014). A single-hidden-layer feed forward neural network is sufficient to approximate any continuous mapping from the input patterns to the outputs mainly because they are less susceptible to poor local minima (Razavi and Tolson, 2011). Therefore, we started with a single-hidden-layer feed forward neural network. The number of nodes in input and output layers was already determined in step 2.1 based on the two principal components of the training input data and the two variables in the Hantush's well function of leaky confined aquifers (i.e. the coordinates of match point, $\mathrm{Log}\left(\frac{1}{u_\mathrm{m}}\right)$ and $(\psi)_\mathrm{m}$). The inputs are fed through the input layer and, after being multiplied by synaptic weights, are delivered to the hidden layer. In the hidden and output layer an activation function such as threshold, sigmoid, hyperbolic tangent, Gaussian and linear are used for converting the weighted summation input to the output (Haykin, 1999). Based on our previous experience (Samani et al., 2007), the hyperbolic tangent function (tansig): $f(x) = (1 - e^{-2x})/(1 + e^{-2x})$ for the hidden layer and the linear transfer function (purelin): $[f(x) = x\,, \text{for all } x]$ for the output-layer are used. Figure 5 illustrates the architecture of our single-hidden-layer network before and after conducting PCA on the input data sets. The optimum number of neurons in the hidden-layer is determined in the process of network training (calibration) in the next sections where the optimum structure of the network is also determined.

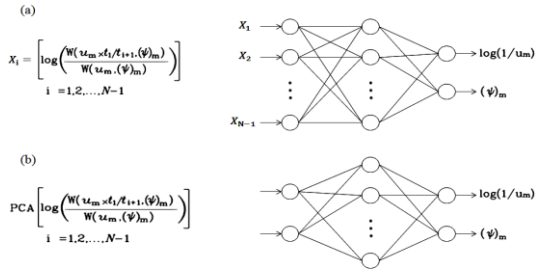


Fig. 5. Structure of the single-hidden-layer network before and after conducting PCA on the input training data sets

2.3. Network Training (Calibration)

Among various methods for neural network training or calibration, the classic back-propagation (BP) method developed by Rumelhart et al. (1986) among others is the most popular for it can learn the mapping of any linear and nonlinear relationship between the inputs and outputs (ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a; Maier and Dandy, 1999; 2000) by finding the optimum set of weights with the use of an optimization algorithm.

In the BP method, the inputs are delivered to the hidden-layer after being multiplied by synaptic weights. The output of the hidden-layer (h_l) can be obtained from:

$$Z_l = b_l + \sum_{i=1}^{n} X_i \, w_{il} \tag{8}$$

$$h_l = f(Z_l) \tag{9}$$

where X_i is the input to the processing nodes (determined in Equation (7) and subjected to PCA); b_l is the bias representing the threshold value associated with node l; w_{il} is the connection weight from the ith node in the preceding layer to node l for imitating the biological synapse strength; n is the total number of inputs applied to nodes i in the input layer and f is the activation function for converting the weighted summation input to the output.

Different implementations of the conjugate gradient approach and various quasi-Newton implementations have been incorporated into the BP algorithm to enhance the convergence speed of this algorithm. The Levenberg–Marquardt (LM) algorithm is probably the most efficient optimization method, which is the approximated Newton algorithm.

In this paper, the network training and the weights adjustment are implemented by the deterministic LM optimization method which is probably the most efficient optimization method for small and medium sized neural networks (Razavi and Tolson, 2011). Many researchers have successfully used this methodology for hydrology problems (Samani, 1990; Maier and Dandy, 1999; 2000; Toth et al., 2000; Coulibaly et al., 2001; Daliakopoulos et al., 2005; Samani et al., 2007). In the LM algorithm, the updated function of the weights $w(k+1)$ is estimated using:

$$w(k+1) = w(k) - [J^{T} \times J + \mu I]^{-1} \times J^{T} \times e \tag{10}$$

where J is the Jacobin matrix of the performance (error) criteria to be minimized, μ is the learning rate, k is the iteration during the optimization process, e is the vector of the residual value and I is the identity unit matrix.

Afterwards, the weights and biases are adjusted for all the interconnection neurons in different layers and the convergence criterion is reached (e.g. 10^{-6}), network training is complete. Our single-hidden-layer network was calibrated (trained) with ten nodes in its hidden-layer through a trial and error process. Therefore, the structure of the trained network gained the topology of [2×10×2], Fig. 6. 2, 10 and 2 refer to the number of neurons in the input, hidden and output layers, respectively. The parameters applied during the training process are shown in Table 2.

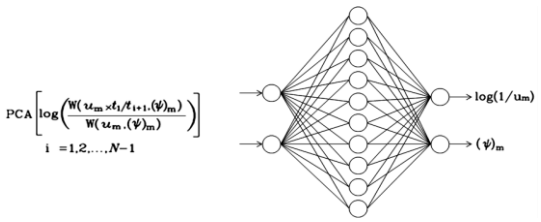


Fig. 6. Structure of the developed ANN in the training stage

Table 2. The ANN parameters applied during training for the leaky confined aquifers

Parameter	Value
Learning rate	0.5
Convergence criterion	1e-6
Maximum training cycle	10000
Number of training patterns	51200

2.4. Determination of Network Optimum Structure

To make sure that the network structure designed in the calibration section is the optimum structure of the network, a sensitivity analysis is conducted to determine the optimum number of hidden layers and nodes. This will put further confidence on generalization of the trained network (structural validity, Wu et al. (2014)) in predicting the match

point coordinates (network target). In the sensitivity analysis the following two efficiency criteria are used:
a) The relative root mean square error (RRMSE) of the calculated target:

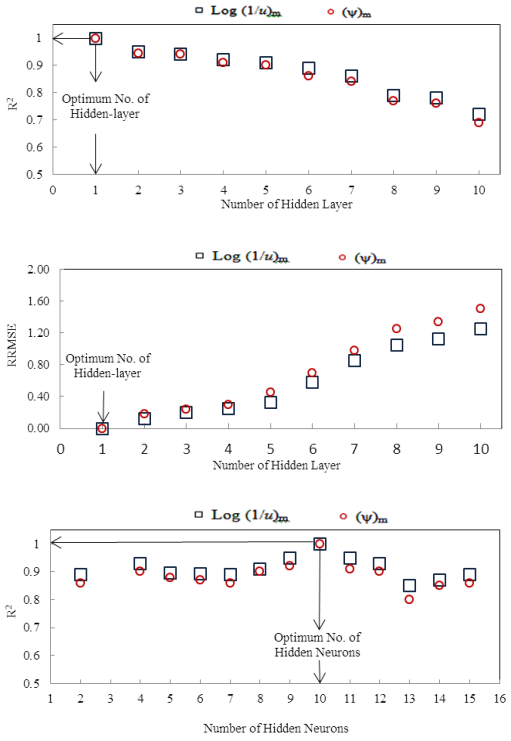
$$\mathrm{RRMSE} = 100 \times \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\hat{y}_j - y_j}{y_j}\right)^2} \qquad (11)$$

where $\hat{y}_j$ is the simulated target calculated by the network, y_j is the actual target and n is the number of data patterns. The accuracy of prediction increases as the value of RRMSE decreases. RRSME=0 indicates 100% precision.
b) The determination coefficient, R^2:

$$R^2 = 1 - \frac{\sum(y_j - \hat{y}_j)^2}{\sum y_j{}^2 - \frac{\sum \hat{y}_j}{n}} \qquad (12)$$

R^2=1 indicates 100% fit between network output values and expected target values.

Figure 7 shows variations of the error criteria (i.e. R^2 and RRMSE) with respect to number of hidden layers and number of nodes for the predicted the two network targets. The four plots in Fig. 7 collectively indicate that a single-hidden-layer feed forward network with the topology of [2×10×2] is the best and optimum ANN that efficiently model the well function of leaky confined aquifers and accurately predict the match point coordinates. The values of the above two criteria for the developed optimum network are also indicated on Fig. 7.



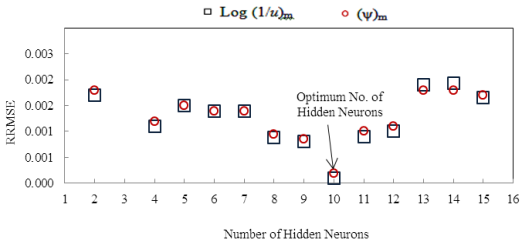


Fig. 7. Sensitivity plots for the model structure

Figure 8 is a plot of convergence of BP indicating that the network with the topology of [2×10×2] is trained smoothly and quickly without falling to local optima compare to networks with simpler and complicated topologies, i.e. [2×4×2], [2×8×2], [2×9×2], [2×11×2] and [2×12×2]. In Figure 8, the y-axis is the mean squared error (MSE) that is the difference between the network output and the actual target.

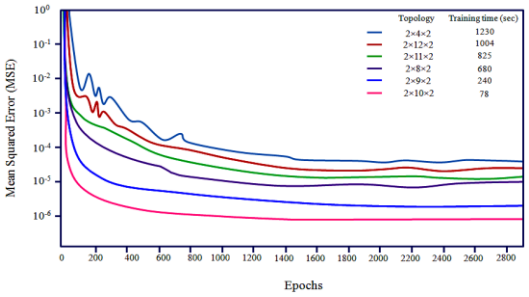


Fig. 8. Convergence plot of networks with different topology

2.5. Testing the Developed Network

Having developed the network with the optimum topology the test patterns are used to assess its performance in determining the match point coordinate and hence the aquifer parameters. The best result of network performance is obtained when the trained network produces the smallest prediction error on the test data sets which are different from the training sets.

The performance of the trained network was assessed by 100000 sets of synthetic error-free drawdown data. The synthetic data are generated by Hantush's analytical solution [Equations (2-5)] for the leaky confined aquifer from combinations of idealized T, S and ψ values ranging from 10^2 to 10^6 m^2/day, 10^{-2} to 10^{-6} and 0.0 to 10, respectively by selecting a number of time steps as in Fig. 3. These synthetic error-free data (they are error-free because Equations (2-5) are exact analytical solution of leaky confined aquifers response to pumping) were converted to drawdown ratios (Lin and Chen, 2005) using Equation (13) which forms the input testing vectors (X_i) to the developed ANN.

$$X_{\mathrm{i}} = \log(s_{\mathrm{i}+1}) - \log(s_1) = \log\left(\frac{s_{\mathrm{i}+1}}{s_1}\right) \quad (13)$$

where s_{i} is the drawdown recorded at time t_{i}, s_1 is the first drawdown record and i=1, 2,…N-1.

Applying the PCA to the input vectors (X_{i}), the reduced drawdown ratios $\mathrm{PCA}\left[\log\left(\frac{s_{\mathrm{i}+1}}{s_1}\right)\right]$ are generated. Figure 9 illustrates the structure of the developed network at the testing stage. The developed network receives the reduced drawdown ratios (drawdown ratios subjected to PCA) and calculates $\log(1/u_m)$, and $(\psi)_m$ which are converted to T, S, K' and $K'S'$ by Equations (2-5).

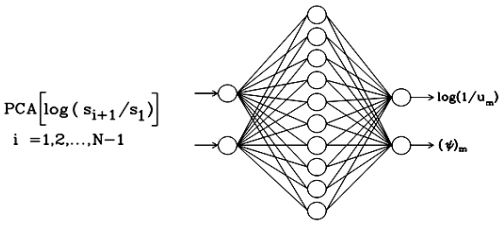


Fig. 9. Structure of the developed ANN in the testing stage

Figure 10 shows the scatter plots and the best-fitted line as well as the QQ (normal distribution) plot of residuals (Bennett et al., 2013) for the idealized and estimated aquifer parameters (T and S) by the proposed ANN for the leaky confined aquifer, respectively. As shown in these figures, the developed network can accurately estimate aquifer parameters over a wide tested range. In all plots the value of R^2 is equal to unity and RRMSE values are very close to zero (summarized in Table 3) indicating a high prediction precision of the developed network in simulating the exact theoretical response of the real system (Equation (2)) and hence its replicative validity (Wu et al., 2014).

2.6. Validation of the developed Network

The main goal of model validation is establishing greater confidence in the model in producing/predicting system actual response. In other words the model should be physically plausible as well as being predictive. Therefore, in this step, a set of real pumping test (time-drawdown) data adapted from Neuman and Witherspoon (1972) was used to evaluate the applicability and reliability of the developed ANN.

Table 3. R^2 and RRMSE (%) values of the estimated parameter values by the developed network during the testing process

RRMSE		
T	S	(ψ)
1.86 e-4	1.62 e-4	1.95 e-3
R^2=1 for the three estimated parameters		

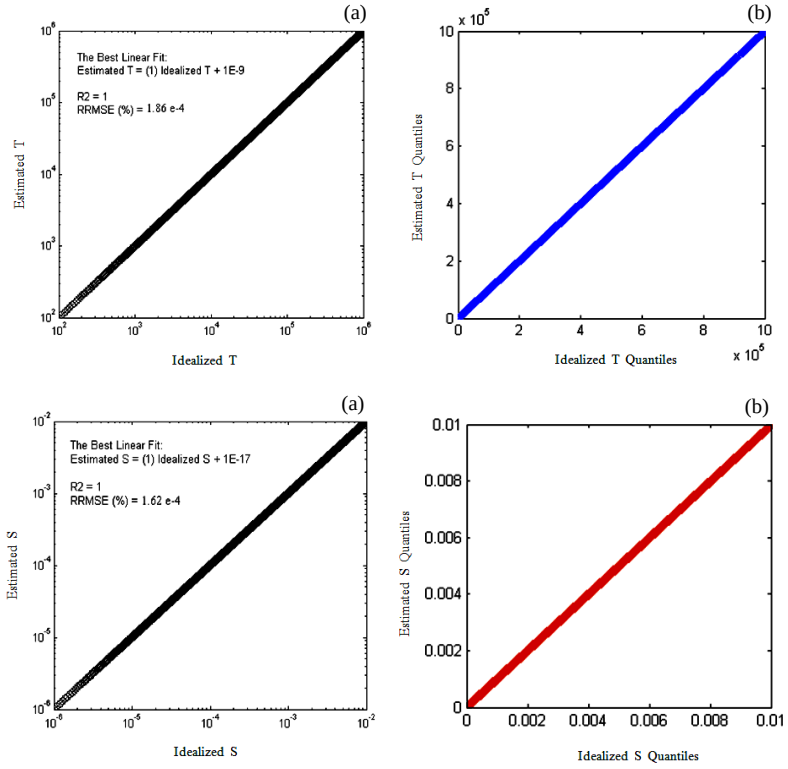


Fig. 10. a) Idealized versus calculated aquifer parameter values, b) QQ plot of residuals

In this test, a fully penetrating well discharges a leaky confined aquifer with a uniform pumping rate of 1000 gpm. The time-drawdown data are measured in a piezometer at 100 ft from the pumping well as given in Table S1 (supplementary materials). The aquitard releases water from storage (S'>0.0). The data set is subjected to the PCA and the reduced set is used as an input to the developed network and allowed the network to determine the match point coordinates values $[(1/u)_{\mathrm{m}}, (\psi)_{\mathrm{m}}]$.

Table S1. Time-drawdown data in the real pumping test (Neuman and Witherspoon, 1972)

Time (min)	Drawdown (ft)	Time (min)	Drawdown (ft)	Time (min)	Drawdown (ft)
10	6.30	130	8.53	1800	10.68
13	6.51	180	8.66	2000	10.80
18	6.72	220	8.84	2400	10.97
22	6.83	250	8.97	2900	11.14
26	7.01	300	9.13	3500	11.27
30	7.16	350	9.33	4000	11.41
38	7.29	430	9.47	4900	11.60
43	7.40	500	9.67	5900	11.78
53	7.65	600	9.88	6900	11.96
63	7.79	750	10.10	8100	12.14
73	8.08	880	10.19	10000	12.32
83	8.26	1100	10.35		
110	8.37	1400	10.51		

3. Determination of aquifer parameter values

Having determined the match point coordinates of the pumping test data by the developed network, they are adjusted and extended for all time-drawn records as illustrated in Fig. 3:

$$\hat{y}_1 = \log\left(\frac{1}{u_{\mathrm{j}}}\right) = \log\left(\frac{1}{u_{\mathrm{m}}}\frac{t_{\mathrm{j}}}{t_1}\right), \mathrm{j} = 1, 2, \ldots, \mathrm{N} \quad (14)$$

$$\left(\frac{1}{u_{\mathrm{j}}}\right) = \left(\frac{1}{u_{\mathrm{m}}}\frac{t_{\mathrm{j}}}{t_1}\right) = 10^{\hat{y}_1} \quad (15)$$

$$W\left[u_{\mathrm{j}}, (\psi)_{\mathrm{m}}\right] = W\left(\frac{1}{10^{\hat{y}_1}}, \hat{y}_2\right) \quad (16)$$

$$s_{\mathrm{m}} = s_{\mathrm{j}} \quad (17)$$

$$t_{\mathrm{m}} = t_{\mathrm{j}} \quad (18)$$

Substituting these values into the well functions [Equations (2-5)] the aquifer parameter values are determined. The drawdown record point $[1/u_{\mathrm{j}}, W_j, s_{\mathrm{j}}, t_{\mathrm{j}}]$ that yields the minimum RRMSE of the estimated drawdown is selected as the optimal match point which yields the aquifer parameter values with the greatest possible accuracy. The aquifer parameters values (T, S and $K'S'$) are determined by the developed network and compared with that of the type curve matching technique in Table 4 which shows a much lower value for RRMSE (i.e. 0.05% compared to 9.4%). In this pumping test the thirty first time-drawdown record was found to be the optimal point (Fig. 11). In Fig. 11 we also plotted the accuracy of each drawdown record relative to that of the optimum record in estimating the match point coordinates. The results in Table 4 demonstrate the predictive validity of the developed network as it has been able to simulate the response of the real system very accurately (Gass, 1983; Wu et al., 2014).

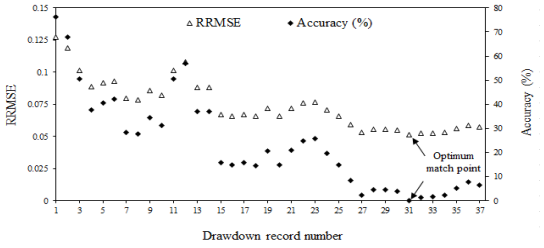


Fig. 11. RRMSE and accuracy plot locating the optimum drawdown-time record for a set of real pumping test data

Table 4. The estimated aquifer parameters and RRMSE (%) of the estimated drawdown using the developed ANN and the type-curve graphical method

Method	Aquifer parameters				RRMSE
	T (gpd/ft)	S	ψ	$K'S'$ (gpd/ft^2)	
Type-curve graphical Method (Neuman and Witherspoon, 1972)	130000	1.11×10^{-4}	0.005	1.73×10^{-5}	9.4
Developed ANN	130620	4.85×10^{-5}	0.006	1.09×10^{-5}	0.05

4. Summary and Conclusions

In this paper, a single-hidden-layer feed forward neural network with the Levenberg-Marquardt (LM) training algorithm was developed to imitate Hantush's well function for the determination of leaky confined aquifer parameters. Inspired by Wu et al. (2014) we followed a six step protocol to develop the network and documented the methodology used in each step. In the first step (input data selection), N-1 sets of input data patterns were generated using the exact analytical Hantush's solution of flow in leaky confined aquifers by considering a wide range of logical values for the network output targets namely,

$\mathrm{Log}\,(\frac{1}{u})_{\mathrm{m}}$ and $(\psi)_{\mathrm{m}}$ (Fig. 3 and Equation (7)). These two variables form neurons in the output-layer. Before designing the network structure the PCA was also applied on the training data sets to reduce the dimension of input training data patterns and eliminate the redundant data by considering a minimum fraction variance of 0.005%. It was found that the first two principle components together explain 99.996% of the variance of the training data. As a result the dimension of the input patterns and hence the number of neurons in the input-layer is reduced and fixed to 2. In the second step (selection of network architecture), we started with a single-hidden-layer network that is probably the most commonly used neural network in engineering applications. In this step activation functions were selected and an arbitrary number of neurons were set for the hidden-layer. In the third step (network calibration), the network training and weights adjustment were implemented by employing the classical BP training method and using the deterministic LM optimization algorithm. The network was trained with 10 neurons in the hidden-layer and gained the topology of [2×10×2] by maintaining the desired convergence criterion of 10^{-6} through a trial and error process. In the fourth step (determining the network optimum structure), a sensitivity analysis was conducted to determine the optimum topology of the trained network by evaluating the ability of networks to generate the target output using the two efficiency criteria of RRMSE and R^2. It was found that both increasing and decreasing number of nodes in the hidden-layer and also increasing number of hidden-layers reduce the accuracy of the network to estimate the expected target. As a consequence the optimum topology of the developed network was fixed to [2×10×2], regardless of the number of drawdown measurements. In the fifth step (testing the network), the accuracy of the developed network was tested by 100000 synthetic *error-free* drawdown-time data sets. The scatter diagram of output targets (estimated target values versus actual target values) and the QQ plot of residuals indicated the network replicative validity. In the last step (network validation) , the performance of the network in generating the match point coordinates is compared with the type curve matching technique using a set of real pumping test data. The comparison evaluated by efficiency indices, i.e. RRMSE and R^2 showed the predictive validity of the network structure.

In brief, the developed network receives the reduced drawdown ratios as inputs and provides the match point coordinates of the first record as the outputs. The outputs are incorporated with Hantush's analytical solution for each individual time-drawdown records and aquifer parameters are calculated. Using each set of aquifer parameters, drawdown records are generated and compared with the real drawdown records, in terms of the RRMSE. The parameter values that give the lowest RRSME value are selected as the most accurate estimate of aquifer parameter values.

The developed network is recommended as an efficient, accurate and easily used alternative tool to the graphical type-curve matching methods for the determination of leaky confined aquifer parameters. The developed network eliminates the graphical error resulted from type curve matching and greatly improves the accuracy of aquifer parameter values. The applied modeling procedure in this paper may be used to design ANN models for other well functions in groundwater hydrology.

References

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., Shamseldin, A. Y., Solomatine, D. P., Toth, E., & Wilby, R. L. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Prog. Phys. Geogr*, 36(4), 480–513.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000a). Artificial Neural Networks in hydrology. I: preliminary concepts. *Journal of Hydrologic Engineering*, 5(2), 115–123.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000b). Artificial Neural Networks in hydrology. II: hydrologic applications. *Journal of Hydrologic Engineering*, 5(2), 124–137.

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D., & Andreassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling and Software*, 40(0), 1–20.

Cattell, R. B. (1966). The Scree test for the number of factors. *Multivariate Behavior Research*, (1), 245–276.

Coulibaly, P., Anctil, F., Aravena, R., & Bobee, B. (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resource Research*, 37(4), 885–896.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309(1–4), 229–240.

Davis, J. C. (2002). *Statistics and data analysis in geology.* 3rd edn. Wiley, New York.

Fausett, L. (1994). *Fundamentals of neural networks*. Prentice Hall, Englewood Cliffs, N.J.

Gass, S. I. (1983). Decision-aiding models: validation, assessment, and related issues for policy analysis. *Oper. Res*, 31(4), 603–631.

Hantush, M. S., & Jacob, C. E. (1955). Non-steady radial

flow in an infinite leaky aquifer. *Transactions American Geophysical Union*, 36(1), 95–100.

Hantush, M. S. (1960). Modification of the theory of leaky aquifers. *Journal of Geophysical Research*, 66(11), 3713–3726.

Hantush, M. S. (1961). Tables of the function $H(u,\beta) = \int_u^{\infty} \frac{e^{-y}}{y}$ erfc $\left(\frac{\beta\sqrt{u}}{\sqrt{y(y-u)}}\right) dy$. *New Mexico Institute of Mining and Technology*, 103, pp. 12.

Haykin, S. (1999). *Neural networks: A comprehensive foundation.* Prentice-Hall: Englewood Cliffs, NJ.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed forward networks are universal approximators. *Neural Network*, 2(5), 359–366.

Karamouz, M., Razavi, S., & Araghinejad, S. (2008). Long-lead seasonal rainfall forecasting using time-delay recurrent neural networks: a case study. *Hydrological Process*, 22(2), 229–241.

Lin, G. F., & Chen, G. R. (2005). Determination of aquifer parameters using radial basis function network approach. *Journal of the Chinese Institute of Engineers*, 28(2), 241–249.

Lin, G. F., & Chen, G. R. (2006). An improved neural network approach to the determination of aquifer parameters. *Journal of Hydrology*, 316(1–4), 281–289.

Lin, H. T., Ke, K. Y., Chen, Ch. H., Wu, Sh. Ch., & Tan, Y. Ch. (2010). Estimating anisotropic aquifer parameters by artificial neural networks. *Hydrological Processes*, 24, 3237–3250.

Maier, H. R., & Dandy, G. C. (1999). Empirical comparison of various methods for training feed-forward neural networks for salinity forecasting. *Water Resource Research*, 32(8), 2591–2596.

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

Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. *Environmental Modelling and Software*, 25(8), 891–909.

McCulloch, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5, 113–115.

Neuman, Sh. P., & Witherspoon, P. A. (1972). Field determination of the hydraulic properties of leaky multiple aquifer systems. *Water Resource Research*, 8(5), 1284–1298.

Razavi, S., & Tolson, B. A. (2011). A new formulation for feed forward neural networks. *Neural Network, IEEE Trans*, 22(10), 1588–1598.

Rumelhart, D. E., Hinton, G. R., & Williams, R. J. (1986). Learning internal representations by error propagation. In: Rumelhart, D. E., David, E., (Eds.), *Parallel Distributed Processing. MIT Press, Massachusetts*, 318–362.

Samani, N. (1990). On the development and calibration of a parametric catchment sediment model. *Journal of Engineering IRI*, 2(3), 47–54.

Samani, N., Gohari-Moghadam, M., & Safavi, A. A. (2007). A simple neural network model for the determination of aquifer parameters. *Journal of Hydrology*, 340(1–2), 1–11.

Theis, C. V. (1935). The relationship between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Transactions American Geophysical Union*, 16, 519–524.

Toth, E., Brath, A., & Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, (239), 132–147.

Wu, W., Dandy, G. C., & Maier, H. R. (2014). Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modeling. *Environmental Modelling and Software*, (54), 108–127.