The Islamic University Journal (Series of Natural Studies and Engineering)
Vol.19, No.1, pp 157-177 , **2011**, ISSN 1726-6807, http://www.iugaza.edu.ps/ar/periodical/

A NEW APPROACH FOR GROUNDWATER QUALITY MANAGEMENT

Mohammed Seyam[1], and Yunes Mogheir[2]
1:Ministry of Public Works & Housing, Gaza, Palestine, mseyam@gov.ps
2:Environmental Engineering Dept., IUG, Gaza, Palestine, ymogheir@iugaza.edu.ps.

Abstract: The main source of water in Gaza Strip is the shallow aquifer, the quality of the aquifer's groundwater is extremely deteriorated in terms of salinity. Salinization of groundwater may be caused and influenced by many variables. Studying the relation of between these variables and salinity is often a complex and nonlinear process, making it suitable to model by Artificial Neural Networks (ANN) .

In order to model groundwater salinity in Gaza Strip using ANN it is necessary to gather data for training purposes. Initially, it is assumed that the groundwater salinity (represented by chloride concentration, mg/l) may be affected by some variables as: recharge rate (R), abstraction (Q), abstraction average rate (Qr), life time (Lt), groundwater level Wl, aquifer thickness (Th), depth from surface to well screen (Dw), and distance from sea shore line (Ds). Data were extracted from 56 wells, most of them are municipal wells and they almost cover the total area of Gaza Strip.

After a number of trials, the best neural network was determined to be Multilayer Perceptron network (MLP) with four layers: an input layer of 6 neurons, first hidden layer with 10 neurons, second hidden layer with 7 neurons and the output layer with 1 neuron. The ANN model generated very good results depending on the high correlation between the observed and simulated values of chloride concentration. The correlation coefficient (r) was 0.9848. The high value of (r) showed that the simulated chloride concentration values using the ANN model were in very good agreement with the observed chloride concentration which mean that ANN model is useful and applicable for groundwater salinity modeling. The ANN model proved that chloride concentration in groundwater is directly affected by abstraction (Q), abstraction average rate (Qr) and life time (Lt) and it was inversely affected by recharge rate (R) and aquifer thickness (Th). The approach is reasonable for the new planning and management of water resources through the attended reconstruction process in Gaza.

Keywords: Groundwater, Quality, ANN, Modeling.

طريقة جديدة لإدارة نوعية المياه الجوفية

الملخص: تعتبر المياه الجوفية المصدر الرئيسي للمياه في قطاع غزة و هي معرضة للتلوث وخصوصاً فيما يتعلق بازدياد معدلات الملوحة التي تتواجد وتتأثر بالعديد من العوامل. دراسة هذه

العوامل عادة ما تكون عملية معقدة مما يجلها مناسبة لتدرس من خلال نظام الشبكات العصبية الصناعية.

إن نمذجة ملوحة المياه الجوفية من خلال الشبكات العصبية الصناعية تتطلب جمع البيانات اللازمة لعملية التدريب التي تقوم بها الشبكة العصبية. في البداية أُفترض أن ملوحة المياه الجوفية المتمثلة بكمية الكلوريد في المياه الجوفية تتأثر بتسعة عوامل هي معدل تسرب مياه الأمطار للخزان الجوفي و كمية السحب الخاصة بكل بئر ومعدل السحب من الخزان الجوفي و المدة الزمنية التي تعرض فيها الخزان الجوفي للسحب و منسوب المياه الجوفية و سمك الخزان الجوفي و عمق الخزان الجوفي و المسافة بين منطقة السحب و البحر ولقد استخرجت هذه البيانات من 56 بئر مياه تغطي معظم مساحة قطاع غزة.

تم تنفيذ عدة محاولات للحصول على نموذج يعطى نتائج جيدة. في البداية تمت عملية النمذجة باستخدام جميع العوامل المفترضة و من النماذج التي تم تطويرها تم دراسة تأثير العوامل على تركيز الكلوريد في المياه الجوفية و بناء على الدراسة تبين أنه يمكن تجاهل بعض العوامل و تم عمل محاولات أخرى تبين من خلالها أن أفضل شبكة عصبية تم التوصل إليها هي Multilayer Perceptron network (MLP) و تتكون من أربع طبقات هي طبقة المدخلات و يوجد بها 6 نيورن و الطبقة المخفية الأولى و يوجد بها 10 نيورن و الطبقة المخفية الثانية و يوجد بها 7 نيورن وطبقة المخرجات و يوجد بها نيرون واحد. طبقة المدخلات تمثل العوامل التالية تركيز الكلوريد الابتدائي و معدل تسرب مياه الأمطار للخزان الجوفي و كمية السحب الخاصة بكل بئر ومعدل السحب من الخزان الجوفي و المدة الزمنية التي تعرض فيها الخزان الجوفي للسحب و سمك الخزان الجوفي أما طبقة المخرجات فتمثل تركيز الكلوريد النهائي.

لقد أعطت الشبكة العصبية نتائج ممتازة اعتماداً على التقارب الكبير بين القيم الحقيقة و القيم المستخرجة من النموذج حيث بلغت قيمة معامل الارتباط 0.9848 و هذا يعني أن هناك توافقا كبيراً بين القيم الحقيقة و القيم المستخرجة من النموذج مما يجعل النموذج صالحاً للاستخدام و التطبيق. تم استخدام النموذج بنجاح كأداة لدراسة تأثير العوامل على تركيز الكلوريد حيث تبين أن تركيز الكلوريد يتناسب طردياً مع كمية السحب الخاصة بكل بئر ومعدل السحب من الخزان الجوفي والمدة الزمنية التي تعرض فيها الخزان الجوفي للسحب و أنها تتناسب عكسياً مع معدل تسرب مياه الأمطار للخزان الجوفي و سمك الخزان الجوفي واستخدم النموذج كوسيلة للتنبؤ بتركيز الكلوريد من الخزان الجوفي في المستقبل وذلك في حالة إعادة إعمار غزة أيضاً.

A New Approach for Groundwater Quality Management

1. INTRODUCTION

The main source of water in Gaza Strip is the shallow aquifer which is part of the coastal aquifer. The quality of the groundwater is extremely deteriorated in terms of salinity and nitrates. Salinity in the Gaza coastal aquifer is often described by the chloride concentration in groundwater. Depending on location and hydrochemical processes, rates of salinization may be gradual or sudden [1].

Salinization of groundwater may be caused by a number and/or combination of different processes, including: seawater intrusion, migration of brines from the deeper parts of the aquifer, dissolution of soluble salts in the aquifer (water-rock interaction), and contribution from discharges from older formations surrounding the coastal aquifer. In addition, potential man-induced (anthropogenic) sources include agricultural return flows, wastewater seepage, and disposal of industrial wastes [2]. In addition, water quality (eg - salinization) is influenced by many factors such as flow rate, contaminant load, medium of transport, water levels, initial conditions and other site-specific parameters. The estimation of such variables is often a complex and nonlinear process, making it suitable for Artificial Neural Networks (ANN) application [3].

The importance of this research is to develop ANN model studying the relation between groundwater salinity (represented by chloride concentration mg/l) and some variables as: recharge rate (R), abstraction (Q), abstraction average rate (Qr), life time (Lt), groundwater level (Wl), aquifer thickness (Th), depth from surface to well screen (Dw), and distance from sea shore line (Ds). Understanding spatial relations between hydrological variables and salinity of groundwater can contribute in an integration of water resources management. Modeling groundwater salinity using traditional modeling softwares consume a lot of efforts and required huge quantity of data while ANN could provide an easy and efficient tool for modeling and prediction that help in water resources management. This research might be considered as one of the few contributions in quantitatively modeling of the relation between groundwater salinity and the hydrological variables in spatial scale using ANN.

2. OBJECTIVE

The primary objective of this research is to develop ANN model studying the relation between groundwater salinity represented by chloride concentration in groundwater and some hydrological variables as: recharge rate (R), abstraction (Q), abstraction average rate (Qr), life time (Lt), groundwater level (Wl), aquifer thickness (Th), depth from surface to well screen (Dw), and distance from sea shore line (Ds).

3. GROUNDWATER SALINITY IN GAZA STRIP

The Gaza Strip is a narrow strip of land on the Mediterranean coast. It is situated in the southeastern coast of Palestine. The area is bounded by the Mediterranean in the west, the 1948 cease-fire line in the north and east and by Egypt in the south The total area of the Gaza Strip is 365 km^2 with approximately 40 km long and the width varies from 8 km in the north to 14 km in the south [4]. Figure (1) showed regional and location map of Gaza Strip.

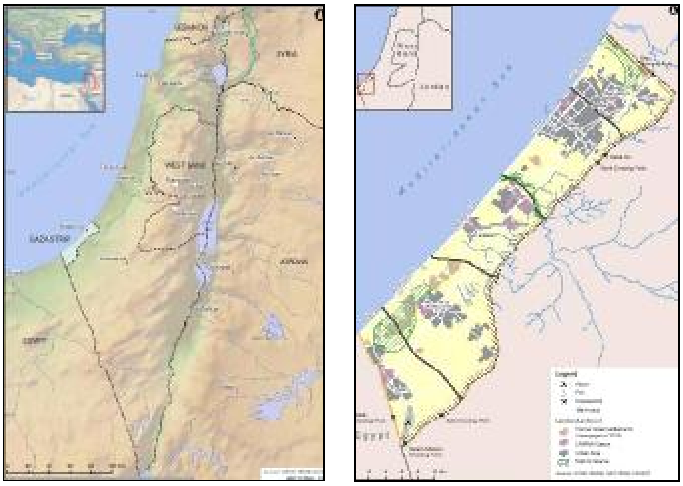

Figure (1): Regional and location map of Gaza Strip [5].

Gaza Strip is one of the places where the exploitation level of recourses exceeds the carrying capacity of the environment. This is especially true for the water and land resources, which are under high pressure and subject to sever over exploitation, pollution and degradation. Quality of the groundwater is a major problem in Gaza strip. The aquifer is highly vulnerable to pollution. The domestic water is becoming more saline every year and average chloride concentrations of 500 mg/l or more is no longer an exception. Most of the public water supply wells don’t comply with the drinking water quality standards and concentrations of chloride and nitrate of the water exceed the World Health Organization (WHO) standards in most drinking water wells of the area and represent the main problem of groundwater quality. Over pumping of groundwater and salt water intrusion are the main reasons behind high chloride concentration [2].

It is clearly noticed that the chloride concentration increases significantly over all Gaza Strip especially in southern east and middle area. The best water quality according to chloride concentration is found in the sand dune areas in the north, mainly in the range of 50 – 250 mg/l. Figures (2) and (3) present average chloride concentration of pumped Groundwater of Gaza Strip for the year 2002 and 2007.

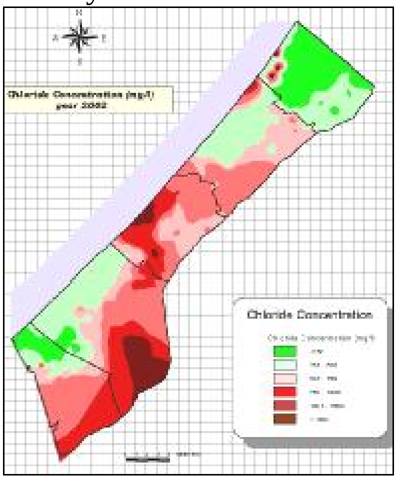

Figure (2): Average chloride concentration of pumped groundwater of Gaza Strip for the year 2002 [6].

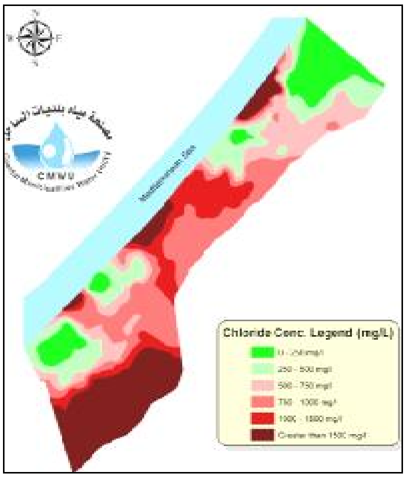

Figure (3): Average chloride concentration of pumped groundwater of Gaza Strip for the year 2007 [7].

4. GROUNDWATER MODELING IN GAZA STRIP

Since its establishment in 1995, Palestinian Water Authority (PWA) has forced to use this modern technique - modeling - in water resources management program in order to simplify the complex hydrogeologic situation of groundwater aquifers and tries to understand the water regime within the entire aquifers. The ultimate goals of the (PWA) is to produce a long-term management plan that will provide rational and practical tools for management of groundwater extraction in Gaza Strip and West Bank aquifers and to identify the most potential zones that are suitable for future development [8]. In recognition the worsening situation of the water in Gaza Strip, PWA and United State Agency for International Development (USAID) have jointly developed the implementation of an Integrated Aquifer Management Plan (IAMP). The IAMP presented overall planning guidelines for water supply and usage through year 2020. As a result to this jointly, new model depending on Coupled Flow-Transport Modeling Code (DYNCFT) was conducted to simulate the effect of IAMP [9].

There are some researchers attempted to model the groundwater in Gaza Strip. Aish (2004) used GIS and MODFLOW in Artificial Recharge Modeling of the Gaza Coastal Aquifer. This research work investigates the first phase of a feasibility study on the impact of artificial recharge from a planned wastewater treatment plant on the groundwater quantity and quality of the coastal aquifer in the Gaza Strip [10]. Ghabayen (2004) developed a model using Bayesian belief networks (BBNs), for Identification of salinity origin. The BBN model incorporates the theoretical background of salinity sources, area specific monitoring data that are characteristically incomplete in their coverage, expert judgment, and common sense reasoning to produce a geographic distribution for the most probable sources of salinization [11]. Qahman (2004) achieve a numerical assessment of seawater intrusion in Gaza Strip by applying a 3-D variable density groundwater flow model [12]. Barakat (2005) developed a model to find optimal values of water quantities from different resources in in the southern Gaza Strip. Visual Modflow (VMF) and its integrated modules, was developed to quantify, and analyze the raw input data [13].

The use of ANN in groundwater quality modeling in Gaza Strip doesn't found in large scale, Till now, only one researcher used it. Al Mahalawi used ANN in Modeling Groundwater Nitrate Concentration of the Gaza Strip [14]. My new research (Groundwater Salinity Modeling Using Artificial Neural Networks - Gaza Strip case study) might be considered as one of the few contributions in quantitatively modeling of the relation between groundwater salinity and the hydrological variables in spatial scale using ANN.

5. BRIEF DESCRIPTION OF ARTIFICIAL NEURAL NETWORKS

ANN refer to computing systems whose central theme is borrowed from the analogy of biological neural networks. They represent highly simplified mathematical models of biological neural networks. They include the ability to learn and generalize from examples to produce meaningful solutions to problems even when input data contain errors or are incomplete, and to adapt solutions over time to compensate for changing circumstances and to process information rapidly [15].

The brain consists of a large number of neurons, connected with each other by synapses. These networks of neurons are called neural networks, or natural neural networks. ANN is a simplified mathematical model of a natural neural network. ANN are a new information-processing and computing technique inspired by biological neuron processing[16]. The human brain is a collection of more than 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive, process, and transmit information [17]. Figure (4) presented mammalian neuron. Treelike networks of nerve fibers called dendrites are connected to the cell body or soma, where the cell nucleus is located. Extending from the cell body is a single long fiber called the axon, which eventually branches into strands and sub strands, and are connected to other neurons through synaptic terminals or synapses. The transmission of signals from one neuron to another at synapses is a complex chemical process in which specific transmitter substances are released from the sending end of the junction[17].

Artificial neurons connected together form a network. The structure of ANN is, as rule, layered. Three functional group can be distinguished in the ANN ie the inputs receiving signals from the network's outside and introducing them into its inside, the neuron which process information and the neurons which generate results. A model of the artificial neuron is shown in the Figure (5).

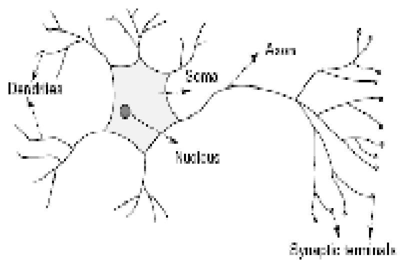

Figure (4): Mammalian neuron [17].

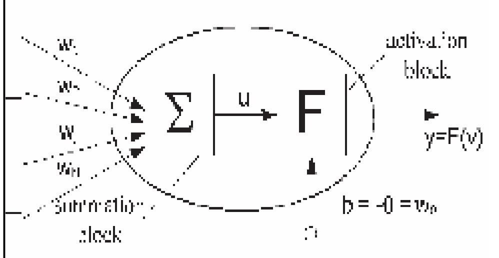

Figure (5): Model of artificial neurons [18].

ANN is an informational system simulating the ability of a biological neural network by interconnecting many simple artificial neurons . The neuron accepts inputs from a single or multiple sources and produces outputs by simple calculations, processing with a predetermined non-linear function [19].

Most ANN has three layers or more: an input layer, which is used to present data to the network; an output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors. Determination of appropriate network *architecture is one of the most important*, but also one of the most difficult, tasks in the model-building process. Unless carefully designed an ANN model can lead to over

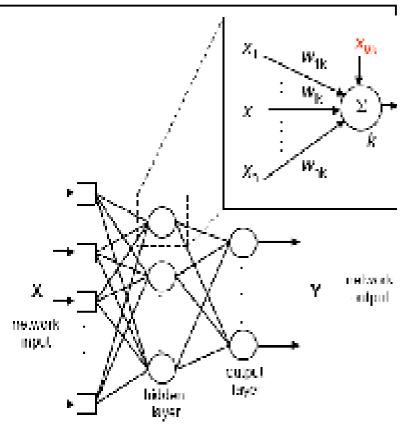

Figure (6): Schematic description of a three layer ANN and of the elements of its (mathematical) neurons [21].

parameterization, resulting in an unnecessarily large network [20]. Figure (6) demonstrated schematic description of a general ANN model of three layers.

6.METHODOLOGY

To achieve the objectives of this research, the following methodology has been applied:

- Data gathering from relevant institution and ministries and revision of accessible references as books, studies, papers and researches relative to the topic of this research which may include ANN, groundwater hydrology and groundwater salinity on Gaza Strip.
- Data analysis using software as Ms. Excel and Access. The analysis is required to construct data matrix which includes some hundreds of data cases of input and output variables. Data cases are considered as row material to ANN model.
- Building ANN model utilization STATISTICA Neural Networks (SNN) which built in STATISTICA program version 7. This step includes training, validation and testing ANN model. The validation and testing is achieved using SNN directly after training process.
- The practical steps to echieve the methodology was as folows:

6.1 Construction Data Matrix for ANN Model training

In order to model the groundwater salinity in Gaza strip using ANN it is necessary to gather data for training purposes. The training data must include a number of cases, each containing values for input and output variables. The first decisions needed are: which are variables to use, and how many (and which) cases to gather? The choice of variables (at least initially) is guided by intuition. Understanding and experience in the problem domain and conditions give initial idea of which input variables are likely to be influential. Once in ANN, variables can be select and deselect, ANN can also experimentally determine useful variables [22]. As a first pass, any variables which could have an influence on groundwater salinity should be included on initial studies.

The required data were extracted mainly from the domestic wells in Gaza Strip because it usually have quality test twice a year in February and October periodically. The quality test includes the chloride concentration test which gives us a great chance to monitor groundwater salinity in Gaza Strip and it's changes two times per year. The previous assumed variables were gathered, studied, validated and rearranged to create training data matrix which should contain many hundreds of cases each containing values for input variables and output.

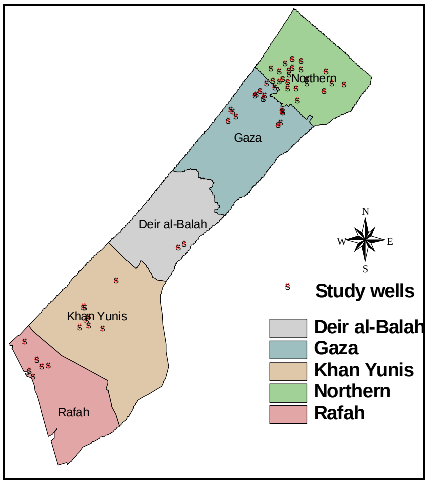

Figure (7): Study wells location in Gaza Strip

In this research, it is necessary to deal with regular time series data to construct data training matrix so many sources of data have been neglected because of the deficiency of complete required data. Since that the detailed abstraction records have not been obtained for years prior to 1996, the period of model which include the modeling and calibration starts from 1997 to 2006. There are an estimated 4000 wells within the Gaza Strip, almost all of these wells are privately owned and used for agricultural purposes. Approximately 100 wells are owned and operated by municipalities and are used for domestic supply [23]. In this research, data were extracted from 56 wells, most of them are municipal wells and they almost cover the total area of Gaza Strip as represented in Figure (7) The choice of these wells depends only on the availability of required data.

1.Selection the Variables of ANN Model

Hydrogeologically, the change of chloride concentration (salinity) was assumed to be depend on many variables such as infiltration, abstraction, life time of abstraction from aquifer, groundwater level, aquifer depth, aquifer thickness, and distance from sea shore line. The variables are described in Table (1).

2.Time Distribution Phases of ANN Model Data

The model data were extracted mainly from domestic wells in Gaza Strip because they usually have records of chloride concentration twice a year in February and October periodically. The time distribution divides the year in two phases A and B. The phase A starts from April to September and the phase B starts from October to March in next year. For example, time phase 1997-A extends from April 1997 to September 1997 , time phase 1997-B extends from October 1997 to March 1998 and time phase 1998-A extends from April 1998 to September 1998, etc. So all other factors were organized according to this time distribution.

3.Organizing of ANN Model Data

The organizing of ANN model data are required to construct some hundreds of data cases of input and output variables. These cases construct data matrix. Data organizing was carried out using software Ms. Excel and Access software. The data matrix is considered as row material to ANN model.

6.2 Analysis of ANN Model Data

Considering only those cases that have complete numeric values for all variables without any missing data, only 499 cases satisfy the above-mentioned criteria from 1997 to 2004. ANN model might perform well over an entire space only when the training data are evenly distributed in the space. As the current data are collected from limited sources (56 municipal wells), they may constitute clusters. Therefore, the distribution of each variables across its range in the database is examined. The mean, standard deviation and ranges of different variables used to train the ANN is shown in Table (1).

Table (1): Mean , standard deviation and ranges of variables used to train the ANN model

Variable	Sym.	Unit	Mean	Std. Dev	Range	
					Min.	Max.
Initial chloride concentration	Cl_o	mg/l	333.07	253.94	28.00	1412.0
Recharge rate	R	mm/m^2/month	18.19	24.44	0.00	83.07
Abstraction	Q	m^3/hour	105.55	57.99	0.00	254.94
Abstraction average rate	Qr	mm/m^2/month	22.50	5.80	11.37	33.94
Life time	Lt	y	22.02	13.94	0.00	60.00
Groundwater level	Wl	m	-1.16	1.15	-4.00	1.00

Depth from surface to well screen	Dw	m	65.26	22.05	18.00	110.00
Aquifer thickness	Th	m	64.17	27.25	30.00	124.00
Distance from sea shore line	Ds	km	3.38	1.47	0.60	7.70
Final chloride concentration	Cl_f	mg/l	341.11	261.09	35.00	1744.1

6.3 Building ANN Model

The ANN model was designed using the STATISTICA Neural Networks [24]. (SNN) which built in STATISTICA program version 7. The procedural steps in building and applying for ANN model varies according to the tool used in building ANN models. Using SNN, the procedural steps involves the following procedures:

1. Data Importation

Feed the data matrix for SNN to train the Network by “importing” or through the data entry process. The data must be in acceptable format such as spreadsheet. The input data is the cases that the network use to train itself.

2. Problem Definition

Specify the inputs (Independent) and the output (Dependent) variable for the ANN model. Initially, there are nine inputs variables and one output variables as mentioned in Table (1).

3. Extraction of the Test Set

In SNN, The test set extraction is about 50% of cases for training. 25% for calibration and 25% for testing and it is randomly selected and the user can change these percentage. Test set provide a means by which the network knows when to stop training and used for calibration and Testing.

4. Network Design

Choose the appropriate architecture of network among the available networks based on the type of the data and the problem. After many trials, Multilayer Perceptron network (MLP) has been chosen because of its high capabilities to generalize well in problems plagued with significant heterogeneity and nonlinearity.

5. Network Training

Once the type of network has been chosen, the conditions to stop training processes was set before the network is trained. Training was controlled by some of conditions as: the maximum number of iterations, target performance which specifies the tolerance between the neural network prediction and actual output, the maximum run time and the minimum allowed gradient.

6. Network Calibration

A trained network will continuously train in order to make a model perform best on the training set. However, after some time, it is very possible for the network to "memorize" the training set instead of learning it. In order to prevent the possibility of memorization to occur, calibration is utilized. Calibration is a parameter, which indicates that the network has trained enough thus stopping the iteration process.

7. Testing of Network

After the network has been successfully trained well, it is then tested against a set of cases withheld from it during its training session. The ANN is then ready to be applied to any other values of variables. The results are then presented in statistical manner. Regression analysis is utilized to measure the degree of correlation between the actual output and the network output. Correlation factor (r) of 1 gives an indication of a perfect model while an (r) of 0 indicates a very bad model. Mathematically the values of (r) represented in equation (1). Detailed description for new built ANN model are presented in results and discussion.

$$R^2 = 1 \quad \frac{\sum_{i=1}^{n}(actual_i - predicted_i)^2}{\sum_{i=1}^{n}(actual_i - mean)^2} \tag{1}$$

7. RESULTS AND DISCUSSION

Many trials were applied to get best performance model. The initial modeling trials were made using all input variables. From created ANN models, the sensitivity analysis was applied, also the importance and effect of each variables was studied and represented. Depending on the results of ANN models some input variables were neglected and new modeling trials are made without using neglected input variables.

7.1 Sensitivity Analysis

SNN conducts a sensitivity analysis on the inputs to a neural network. This indicates which input variables are considered most important by that particular neural network. Sensitivity analysis can be used purely for informative purposes. Sensitivity analysis can give important insights into the usefulness of individual variables. It often identifies variables that can be safely ignored in subsequent analyses, and key variables that must always be retained. However, it must be deployed with some care, for reasons that are explained below.

Input variables are not, in general, independent - that is, there are interdependencies between variables. Sensitivity analysis rates variables according to the deterioration in modeling performance that occurs if that variable is no longer available to the model. In so doing, it assigns a single rating value to each variable. However, the interdependence between variables means that no scheme of single ratings per variable can ever reflect the subtlety of the true situation.

Consider, for example, the case where two input variables encode the same information (they might even be copies of the same variable). A particular model might depend wholly on one, wholly on the other, or on some arbitrary combination of them. Then sensitivity analysis produces an arbitrary relative sensitivity to them. Moreover, if either is eliminated the model may compensate adequately because the other still provides the key information. It may therefore rate the variables as of low sensitivity, even though they might encode key information. Similarly, a variable that encodes relatively unimportant information, but is the only variable to do so, may have higher sensitivity than any number of variables that mutually encode more important information.

SNN conducts sensitivity analysis by treating each input variable in turn as if it were "unavailable". SNN has defined a missing value substitution procedure, which is used to allow predictions to be made in the absence of values for one or more inputs. To define the sensitivity of a particular variable, v, the network first was run on a set of test cases, and the network error was accumulated. Then the network was run again using the same cases, but this time replacing the observed values of v with the value estimated by the missing value procedure, and again the network error was accumulated [24].

After that, It is expected some deterioration in error to occur. The basic measure of sensitivity is the ratio of the error with missing value substitution to the original error. The more sensitive the network is to a particular input, the greater the deterioration we can expect, and therefore the greater the ratio. Once sensitivities have been calculated for all variables, they may be ranked in order. SNN provides these rankings, for convenience in interpreting the sensitivities.

7.2 Determination of the Real Importance of Input Variables

Sensitivity analysis does not rate the "usefulness" of variables in modeling in a reliable or absolute manner. The cautious is needed during drawing the conclusions about the importance of variables. Nonetheless, in practice it is extremely useful. If a number of models are studied, it is often possible to identify key variables that are always of high sensitivity, others that are always of low sensitivity and "ambiguous" variables that change ratings and probably carry mutually redundant information [24].

To determine the real importance and ranking of variables, twenty ANN models were created and sensitivity analysis was applied and results of sensitivity analysis were presented in Table (2).

Table (2): Sensitivity analysis of twenty ANN models

Model No.	Cl_o	R	Q	Qr	Lt	Wl	Dw	Th	Ds
1	2.95	1.02	1.00	0.98	0.98	0.98	1.02	0.98	1.01
2	3.22	1.02	0.98	1.06	1.01	0.99	1.02	1.00	1.04
3	3.52	1.00	1.04	1.01	1.02	1.03	0.98	1.04	1.04
4	3.00	0.99	1.00	0.99	1.01	1.01	0.98	1.02	0.99
5	3.14	0.99	1.00	0.98	0.99	1.00	0.98	0.99	0.96
6	3.06	1.00	1.00	1.01	1.01	1.00	0.98	1.00	0.95
7	3.51	0.99	1.00	1.02	1.03	1.00	1.00	1.02	0.96
8	3.27	1.00	1.00	1.00	1.01	1.00	0.99	0.98	0.97
9	3.54	0.99	1.01	0.99	1.02	1.01	1.00	1.01	0.97
10	3.51	1.01	1.01	1.02	1.04	1.00	0.99	0.99	0.99
11	2.72	1.01	1.00	1.04	0.98	1.02	0.97	1.04	0.98
12	2.96	0.99	1.05	1.08	0.98	1.04	0.99	1.02	0.98
13	3.09	1.00	1.01	1.12	1.03	1.03	0.97	1.15	1.03
14	3.08	1.04	1.05	1.09	1.10	1.02	1.07	1.08	0.99
15	3.23	1.02	1.01	1.02	0.98	1.01	1.00	0.98	0.97
16	3.57	1.01	1.01	1.09	1.00	1.01	1.04	1.06	1.01
17	3.65	1.04	1.00	1.04	1.01	1.01	1.03	1.04	0.98
18	3.74	1.02	1.01	1.08	1.01	1.02	1.02	1.06	1.00
19	4.10	1.05	1.01	1.05	1.05	1.04	1.01	1.03	1.00
20	4.24	1.05	1.04	1.08	1.05	1.01	1.02	1.01	1.02
Total	**67.10**	**20.26**	**20.23**	**20.73**	**20.32**	**20.23**	**20.07**	**20.49**	**19.85**
Average	**3.355**	**1.013**	**1.012**	**1.037**	**1.016**	**1.011**	**1.004**	**1.024**	**0.992**
Rank	**1**	**6**	**5**	**2**	**4**	**7**	**8**	**3**	**9**

From the results in Table (2), the variables can be ranked. This rank is needed to identify variables that can be safely ignored in subsequent modeling, and key variables that must always be retained. From the result three variables with lowest sensitivity which is groundwater level (Wl), depth from surface to well screen (Dw), and distance from sea shore line (Ds) can be ignored in subsequent modeling.

Depending on the results, new training trials are made without using neglected input variables. The neglected input variables is groundwater level, depth from surface to well screen, and distance from sea shore line.

7.3 Characteristic of ANN model

- **Topology of ANN**

After a number of trials, the best neural network was determined to be **<u>Multilayer Perceptron network (MLP)</u>** with four layers: an input layer of 6 neurons, first hidden layer with 10 neurons, second hidden layer with 7 neurons and the output layer with 1 neuron as shown in Figure (8). The six input neurons are: initial chloride concentration (Cl_o), recharge rate (R), abstraction (Q), abstraction average rate of area (Qr), life time (Lt), aquifer thickness (Th). The output neuron gives the final chloride concentration (Cl_f).

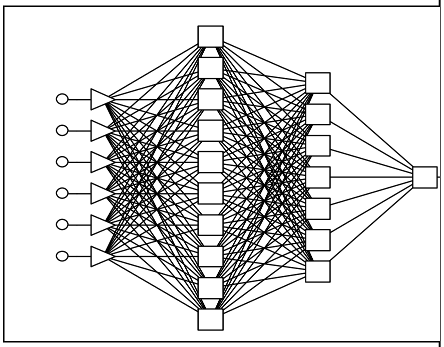

Figure (8): Topology of final ANN model

- **Performance of ANN**

The progress of the training was checked by plotting the training, and test mean square errors versus the performed number of iterations, as presented in Figure (9).

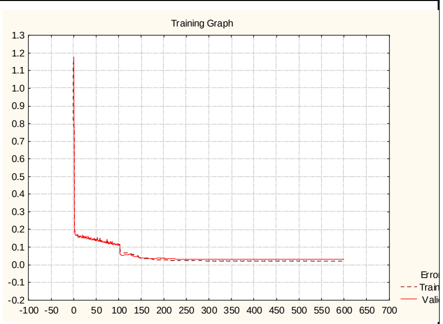

Figure (9): Training progress of ANN

Figure (10) presented a comparison of simulated chloride concentration using ANN and the observed chloride concentration. The Figure (10) showed a very high correlation between the observed and predicted values of chloride concentration. The correlation coefficient (r) between the predicted and observed output values of the ANN model is 0.9848. The high value of correlation coefficient (r) showed that the simulated chloride concentration values using the ANN model were in very good agreement with the observed chloride concentration which gave initial impression that ANN model are useful and applicable. Simulated chloride concentration using ANN model and observed chloride concentration on 1/10/2000 are presented in Figure (11).

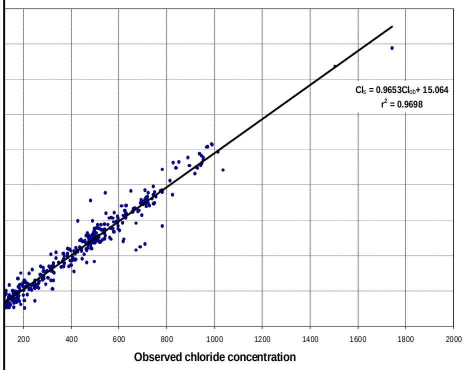

Figure (10): Comparison of simulated chloride concentration using ANN model and the observed chloride concentration

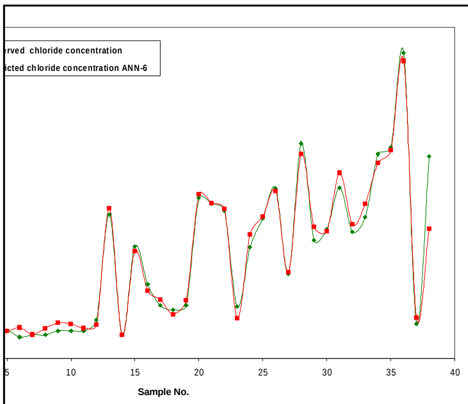

Figure (11): Comparison of simulated chloride concentration using ANN and the observed chloride concentration on 1/10/2000

- **Regression Statistics of ANN Model**

In regression problems, the purpose of the neural network is to learn a mapping from the input variables to a continuous output variable. A network is successful at regression if it makes predictions with accepted accuracy. SNN automatically calculates correlation coefficient (r) between the actual and predicted outputs. A perfect prediction will have a correlation coefficient of 1.0. A correlation of 1.0 does not necessarily indicate a perfect prediction (only a prediction which is perfectly linearly correlated with the actual outputs), although in practice the correlation coefficient is a good indicator of performance. It also provides a simple and familiar way to compare the performance of neural networks with standard least squares linear fitting procedures. The degree of predictive accuracy needed varies from application to application.

Regression statistics are listed as follows:

- **Data Mean**: Average value of the target output variable.
- **Data S.D.**: Standard deviation of the target output variable.
- **Error Mean**: Average error (residual between target and actual output values) of the output variable.
- **Abs. E. Mean**: Average absolute error (difference between target and actual output values) of the output variable.
- **Error S.D.**: Standard deviation of errors for the output variable.
- **S.D. Ratio**: The error/data standard deviation ratio.

- **Correlation**: The correlation coefficient (r) between the predicted and observed output values.

Table (2) present the values of regression statistics for the ANN model.

Table (2): The values of regression statistics for final ANN model

Regression statistics	All model data	Training data set	Validation data set	Test data set
Data Mean	341.105	**295.877**	345.200	361.427
Data S.D.	260.827	247.433	262.657	263.607
Error Mean	3.242	5.016	8.428	-0.196
Error S.D.	45.371	45.125	47.312	44.204
Abs E. Mean	29.798	29.262	32.128	28.911
S.D. Ratio	0.174	0.182	0.180	0.168
Correlation (r)	0.9848	0.9832	0.9837	0.9860

Notes

- Low value of **Error Mean**, **Abs E. Mean** and **S.D. Ratio** showed that the error between observed and simulated chloride concentration values using the ANN model are small.
- High value of **correlation coefficient (r)** showed that the simulated chloride concentration values using the ANN model are in good agreement with the observed chloride concentration.

7.4 Response Graph

Response graph shows the effect on the output variable prediction of adjusting input (independent) variables. The ANN model was utilized to study the influence of the input variables on output variable which is chloride concentration. Figure (12) presented a response graph of each input variables of final ANN model.

Figures (12.a,c,d,e) indicated that chloride concentration increases nonlinearly as chloride concentration initial, abstraction, abstraction average rate and life time increase. Figures (6.10.b,f) indicated that chloride concentration decreases nonlinearly as recharge rate and aquifer thickness

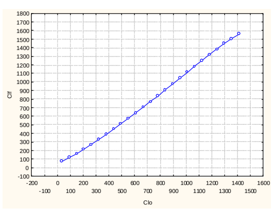

Figure (12.a): Response graph of Cl_o

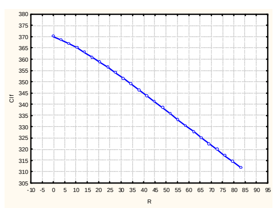

Figure (12.b): Response graph of R

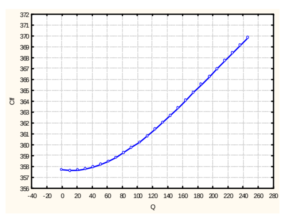

Figure (12.c): Response graph of Q

increase.

8.CONCLUSIONS

The following conclusions were made based on the results obtained from the current study:

1. A new approach for Groundwater salinity modelling in Gaza Strip utilizing ANN was successfully developed and applied. ANN model was developed to study the relation between groundwater salinity (represented by chloride concentration in groundwater) and some related hydrological factors such as recharge rate (R), abstraction (Q), abstraction average rate (Qr), life time (Lt), groundwater level (Wl), aquifer thickness (Th), depth from surface to well screen (Dw), and distance from sea shore line (Ds).

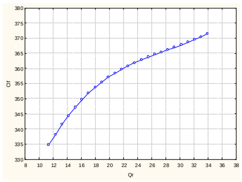

Figure (12.d): Response graph of Qr

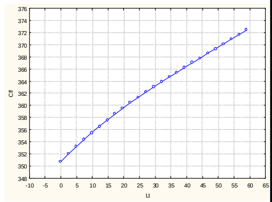

Figure (12.e): Response graph of Lt

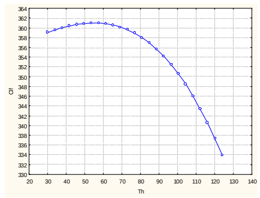

Figure (12.f): Response graph of Th

2. After a number of modelling trials, the best neural network was Multilayer Perceptron network (MLP) with four layers: an input layer of 6 neurons, first hidden layer with 10 neurons, second hidden layer with 7 neurons and the output layer with 1 neuron . The six input neurons represented the input variables which are: initial chloride concentration (Cl_o), recharge rate (R), abstraction (Q), abstraction average rate (Qr), life time (Lt) and aquifer thickness (Th). The output neuron gives the final chloride concentration (Cl_f).
3. The new approach generated very good results depending high correlation between the observed and predicted values of chloride concentration. The correlation coefficient (r) between the predicted and the observed output values of the ANN model was 0.9848. The high value of correlation coefficient (r) showed that the simulated chloride concentration values using the ANN model were in very good agreement with the observed chloride concentration which mean that ANN model are useful and applicable.
4. The ANN model proved that chloride concentration in groundwater is directly affected by abstraction (Q), abstraction average rate (Qr) and

life time (Lt). Furthermore, it was adversely affected by recharge rate (R) and aquifer thickness (Th).

5. Therefore, the current research showed that ANN model can be used in groundwater quality management and it is comparable to other used approaches such as groundwater modelling and statistical modelling. It showed that the strong remedial actions for solving the groundwater deterioration problem in the aquifer of Gaza Strip (salinity) are reducing the abstraction rate and increasing the recharge quantities to the aquifer.

9.RECOMMENDATIONS

The following recommendations were made based on the results obtained from the study:

1. New water sources should be found and the abstraction from Gaza Strip aquifer should be reduced to solve groundwater salinity problem.
2. New abstraction wells should be constructed in appropriate areas that have large aquifer thickness and high ability to infiltrate rainfall to groundwater depending on their adversel reletion with chloride concentration.
3. It is recommended to close old constructed wells to reclaim groundwater aquifer around the well area. In addition, some wells with very high chloride concentration especially at the west area in Gaza city should be closed.
4. Although, the ANN model performed well, further studies about hydrological processes using ANN in Gaza strip will enhance the utilizing ANN as modelling and management approach.
5. Although, the ANN model performed well, further studies about using ANN model in groundwater management approach is recommended. An example of these studies is the effect of increasing recharge areas such as sormwater and treated wastewater infilitration basins, on salinity. In contrast, the extention of urbanized areas and their influence on slainity can be also a future study.

10. REFERENCES

[1] Metcalf & Eddy, (2000). Costal Aquifer Management Program, "Final Report: Modeling of Gaza Strip Aquifer". The program is funded by US Agency for International Development (USAID) and owned by the Palestinian Water Authority (PWA). Gaza, Palestine.

[2] Coastal aquifer management plan (CAMP) (2000). Gaza Coastal Aquifer Management Program, Volume 1, Task 3.

[3] Govindaraju R. S., (2000). "Artificial neural network in hydrology". Journal of Hydrologic Engineering, 5(2), 124-137.

[4] United Nations Environment Programm (UNEP) (2009). "Environmental Assessment of the Gaza Strip following the escalation of hostilities in December 2008 -January 2009".

[5] United Nations Environment Programm (UNEP) (2003). "Desk study on the Environment in the Occupied Palestinian Territories". Switzerland.

[6] Palestinian Water Authority (PWA), (2003). "Groundwater Levels Decline Phenomena in Gaza Strip Final Report". Water Resources and Planning Department-Hydrology Section. February, 2003. Gaza, Palestinian National Authority.

[7] Coastal Municipality Water Utility (CMWU) (2007). "Annual Report of Wastewater Quality in Gaza Strip for years 2007 and 2008". Gaza, Palestine.

[8] Palestinian water Authority PWA, (2005). "Agricultural and Municipal Water Demand in Gaza Governorates for 2004". Strategic Planning Department ,Water Resource and Planning Directorate, Gaza, Palestine.

[9] Moe H., Hossain R., Fitzgerald R., Banna M., Mushtaha A., Yaqubi A. (2001) "Application of a 3-dimensional coupled flow and transport model". First International Conference on Saltwater Intrusion and Coastal aquifers-Monitoring, Modeling, and Mangement. Essaouira, Marocco.

[10] Aish A. (2004). "Hydrogeological Study and Artifitial Recharge Modeling of the Gaza Coastal Aquifer using GIS and MODFLOW". PhD thesis, Vrije Universiteit Brussel (VUB), Belgium.

[11] Ghabayen Said, Mac McKee, and Mariush Kemblowski.(2004). "Using Bayesian Networks, Ionic Ratios, and Isotopes for Identification of Salinity Origin and Data Requirements in the Gaza Coastal Aquifer". USU Water Initiative 2004 Spring Runoff Conference.

[12] Qahman K. and Larabi A. (2004)."Three dimensional numerical models of sea water intrusion in Gaza aquifer". LIMEN, Ecole Mohammadia D'Ingenieurs, Université Mohammed V. Rabat, Morocco.

[13] Barakat M. (2005). "Optimal water resources management in the southern Gaza Strip". M.Sc. thesis, The Islamic University, Gaza, Palestine.

[14] Mahalawi K.(2007) "Modeling Groundwater Nitrate Concentration of the Gaza Strip by Using Artificial Neural Network". Intercontinental conference - Sustainable development and management of water in Palestine. Amman, Jordan.

[15] Jain, S. K., Singh, V. P., and Van Genuchten, M.T. (2004). "Analysis of soil water retention data using artificial neural networks". Journal Hydrologic Engineering. 9(5), 415-420.

[16] Lee S., Cho S. and Wong M.. (1998). "Rainfall Prediction Using Artificial Neural Networks". Journal of Geographic Information and Decision Analysis 2 (2), 233 – 242.

[17] Ajith A. (2005) "Artificial Neural Networks Oklahoma State" University, Stillwater, OK, USA. Handbook of Measuring System Design, John Wiley & Sons, Ltd.

[18] Hola J. and Schabowicz K. (2005). "Application of Artificial Neural Networks to determine concrete compressive strength based on non-destructive test". Journal of civil engineering and management, 11 (1), 23-32.

[19] Jeng D. and Cha H. (2003) "Application of Neural Network in Civil Engineering" school of Engineering, Griffith University Gold Coast Campus. Australia.

[20] Sudheer K. P., Gosain A. K., Mohana R. D. and Saheb S. M. (2002) "Modeling evaporation using an artificial neural network Algorithm" Hydrol. Process. 16, 3189–3202.

[21] Claudius M. and Olaf K. (2005). "Progress report on trend analysis methods, tools and data preparation (in particular artificial neural networks (ANN))". University of Tubingen, Center for Applied Geoscience.

[22] Jiang H. and Cotton W. R. (2004) "Soil moisture estimation using an artificial neural network: a feasibility study". Canadian Journal of Remote Sensing. Vol. 30, No. 5, pp. 827–839.

[23] Palestinian Water Authority (PWA), (2000). "National Water Plan", Final Report. The study is funded by United Nation Development Program (UNDP) and owned by the Palestinian Water Authority (PWA). Gaza, Palestine.

[24] STATISTICA Electronic Manual (2004) StatSoft, Inc. STATISTICA (data analysis software system), version 7.