The aim of the book is to develop a model equation that describes the spread of contamination through soils which can be used to determine the rate of environmental contamination by estimate the concentration of heavy metals in soil. The developed model equation can be considered as a good representation for a problem of environmental contamination. The other aim of this book is to design two feed forward neural networks (FFNN) as an alternative accurate technique to determine the rate of environmental contamination which can be used to solve the model equation. The first network is to simulate the soil parameters which can be used as input data in the second suggested network, while the second network simulates to estimate the concentration of heavy metals.

Luma Naji Mohammed Tawfiq received her Ph.D. in Applied Mathematics in the ANNs and its applications in 2004, from the University of Baghdad. She is currently a professor and research member leader at the Mathematical science. Farah Faisal Ghazi received her B.E. and M.E. from the University of Baghdad in 2010 and 2016, respectively.

978-3-659-87155-9

Heavy Metals in Soil

Luma Naji Mohammed Tawfiq
Farah Faisal Ghazi

New Technique to Estimate the Concentration of Heavy Metals in Soil

On Contamination in Soil

LAP LAMBERT Academic Publishing

Luma Naji Mohammed Tawfiq
Farah Faisal Ghazi

New Technique to Estimate the Concentration of Heavy Metals in Soil

Luma Naji Mohammed Tawfiq
Farah Faisal Ghazi

New Technique to Estimate the Concentration of Heavy Metals in Soil

On Contamination in Soil

LAP LAMBERT Academic Publishing

Impressum / Imprint
Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek: The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Coverbild / Cover image: www.ingimage.com

Verlag / Publisher:
LAP LAMBERT Academic Publishing
ist ein Imprint der / is a trademark of
OmniScriptum GmbH & Co. KG
Bahnhofstraße 28, 66111 Saarbrücken, Deutschland / Germany
Email: info@lap-publishing.com

Herstellung: siehe letzte Seite /
Printed at: see last page
ISBN: 978-3-659-87155-9

New Technique to Estimate the Concentration of Heavy Metals in Soil

By

Luma Naji Mohammed Tawfiq & Farah Feasal Ghazi

Acknowledgements

Praise be to Allah and thanks for his blessing and help me to complete this piece of book.

We would like to thanks Mr. Majed Hameed Rasheed who helped to get the data of the book. Also we introduce our gratefulness to Iraqi Ministry of Agriculture.

List of Symbol and Abbreviation

Symbol	Definition
AAS	Atomic Absorption Spectrophotometer
XRF	X- Ray Fluorescence analysis
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
w^*	Optimal Weight
W^*	Optimal Weight Matrix
ANN	Artificial Neural Network
w_{ij}	Weight Connecting the Input Unit j to the Hidden Unit i
v_i	Weight Connecting the Hidden Unit i to the Output Unit
y	Output Value
y_d	Desired Output
x_i	Input Value
w_i	Weight Value
b	Bias
φ	Activation Function
logsig	Log-Sigmoid Transfer Function
tansig	Hyperbolic Tangent Transfer Function
FFNN	Feed Forward Neural Network
BP	Back propagation
SSE	Sum of Squared Errors
MSE	Mean Squared Errors
e	Vector of Network Errors
y_t	Trial Solution

Zn	Zinc
Ni	Nickel
Pb	Lead
Cu	Copper
Cd	Cadmium
Cr	Chromium
As	Arsenic
Co	Cobalt
γ	Performance Ratio
η	Learning Rate
ρ_k	Search Direction
LM	Levenberg–Marquardt algorithms
trainlm	Levenberg - Marquardt Training Algorithm
J	Jacobian Matrix
H	Hessian Matrix
g	Gradient
E	Error Function
μ	Combination Coefficient
I	Identity Matrix
$\|\|\cdot\|\|$	2^{nd} norm
D	diagonal matrix
L	Lower triangular matrix
U	Higher triangular matrix
y_a	Analytic Solution

C_0	Initial Concentration
C_x	Concentration for depth x
ρ	Retardation factor.
PH	power of hydrogen
EC	Electric Conductivity
C	Solute concentration
v	Darcy's flux
mg	Milligram
cm	Centimeter
hr	Hour
L	Liter
D_L	Hydrodynamic dispersion coefficient
x	Soil depth
t	Time
C_a	Concentration of adsorbed chemical
f	Irreversible reaction decay rate
K_d	Distribution coefficient
kg	Kilogram
θ	Soil water content
ρ_b	Soil's bulk density
D_L	Hydrodynamic dispersion coefficient
t	Time
ρ	Solid density
PDE	Partial Differential Equation

Abstract

There are many aims of this book:

The first aim is to develop a model equation that describes the spread of contamination through soils which can be used to determine the rate of environmental contamination by estimate the concentration of heavy metals (HMs) in soil. The developed model equation can be considered as a good representation for a problem of environmental contamination.

The second aim of this work is to design two feed forward neural networks (FFNN) as an alternative accurate technique to determine the rate of environmental contamination which can be used to solve the model equation. The first network is to simulate the soil parameters which can be used as input data in the second suggested network, while the second network simulates to estimate the concentration of heavy metals.

The third aim is to develop a conceptual theory of training stage of neural networks from the perspective of functional analysis and optimization methods. Within this formulation, learning means to solve a variational problem by minimizing a performance function associated to the neural network. The choice of the objective functional depends on the particular application. On the other side, we suggest modification of the performance function to improve the generalization of the suggested networks and to treat the early stopping and local minima problems.

The fourth aim is to compare the performance of aforementioned algorithms with regard to predicting ability. Then applied the suggested technique to estimate the concentration of heavy metals such as: Copper(Cu), Lead(Pb), Cadmium(Cd), Cobalt(Co), Zinc(Zn) and Nickel(Ni) in Baghdad soils. First, sixty four soil samples were selected from a phytoremediated contaminated site located in some zones in Baghdad city (residential, industrial, commercial, agricultural

and main roads). Second, a series of measurements were performed on the soil samples and analyzed measuring of concentrations for heavy metals using devices such as : Atomic Absorption Spectrophotometer (AAS), X-Ray Fluorescence (XRF) and Inductively Coupled Plasma-Mass Spectrometry (ICP- MS) to get initial concentrations for those heavy metals. Third, simulate and train the suggested networks to get the concentration of heavy metals. The performance of the suggested networks was compared with the traditional laboratory inspecting using the training and test data sets. The results of this book show that the suggested networks trained on experimental measurements can be successfully applied to the rapid and accuracy estimation of concentration of heavy metals.

Finally, we suggest some methods for the treatment of contaminated soil by using some herbal plants.

Contents

Introduction

Mathematical modeling is important to describe the different aspects of the real world and the solution of this modeling represents the solution for the problems in life [36]. This work studies and modifies a model equation of convective – diffusive to represent a suitable form which describes the environment contamination. especially, soil contamination by heavy metals. Then using this model to estimate the concentration of heavy metals in soil. This work suggests the artificial neural network as a tool to solve the modified model.

Artificial neural network (ANN) is a computing system which can be trained to learn a complex relationship between two or many variables or data sets. Having the structures similar to their biological counterparts, neural networks are representational and computational models processing information in a parallel distributed fashion composed of interconnecting simple processing nodes [1].

There are many studies for solving the soil contamination using ANN. For instance, Buszewski and Kowalkowski, 2006 [10] present the results of chemo metric treatment of data from experiment of column leaching to find dependencies between physicochemical parameters of soil and heavy metals concentration by using ANN model. Yetilmezsoy and Demirel, 2008 [47] used ANN model to predict the removal efficiency of Pb from aqueous solution. Kardam et al., 2010 [25] presented ANN model to describe the removal efficiency of Cd from aqueous solution using Shelled Moringa Oleifera Seed (SMOS) powder. Yin Li et al., 2011 [48] employed BP-NN model and Geogrophic Information System (GIS) for heavy metals descripted the spatial dynamics of distribution in Huizhou city. El-Badaoui et al., 2013 [15] predict the concentration of heavy metals in Moroccan river sediments relying on a number of physico-chemical parameters using ANNs. Pandhuripande et al., 2013 [34]

estimated the concentration of Ni and Cr in aqueous solution with its physical properties using ANN. Krishna and Sree, 2014 [28] suggested the ANN for remove efficiency of Cr from aqueous solution used a Borasus flabellifer coir powder as adsorbent. Zongshu Wu et al., 2015 [49] explained the optimization way for soil sampling to spatial distribution of heavy metals concentrations using ANN with genetic algorithm. Madhloom 2015[30] suggested using ANN to describe the removal efficiency of Cu from wastewater by fungal biomass.

This work suggested an effective, low cost and easily accessible design of ANN to estimate soil contamination problems. Two neural networks are suggested to estimate the rate of contamination in soil. The first network is to simulate suitable soil parameters which is needed in the model equation that describes the contamination problem. While the second ANN is used to estimate the concentration of heavy metals in soil The results are compared with those obtained by the traditional laboratory devices such as: Atomic Absorption Spectrophotometer (AAS), X-Ray Fluorescence (XRF) and Inductively Coupled Plasma-Mass Spectrometry (ICP- MS) to illustrate the accuracy and the efficiency of the suggested technique.

This book is organized as follows: Chapter one contains three sections, section one represents mathematical concepts related to some definitions, hypotheses, axioms and theorems which we need in this book. Section two describes a few basic concepts about the ANNs, e.g., what is ANN, Transfer functions, Architecture of ANNs, FFNN, Training of ANNs, Back propagation training algorithm, Performance functions. Section three contains a brief overview of environmental contamination, soil, heavy metals which we need throughout the work. Chapter two describes the training algorithm for ANNs and focus on Levenberg – Marquardt (LM) method with tells cost of using this method. Then suggests an improvement of LM training algorithm and prove the rate of convergence of the suggested improvement. We also describe how to

choose a suitable learning rate of ANN and how to improve the generalization of the suggested network. In the third chapter, we describe and develop a mathematical model for spread of contamination through soils which can be used to determine the rate of contamination by estimating the concentration of heavy metals in soil. Then design suitable ANNs to solve this model equation where the solution of the equation represents the concentration of heavy metals. Chapter four introduces a practical application of the suggested problem to estimate the contamination in soil of Baghdad. Finally, chapter five contains the conclusions and future works. Note that all algorithms in this work implemented using MATLAB version 7.12.

Chapter One

Preliminaries

Chapter One

Preliminaries

1. Introduction

This chapter gives a brief introduction to some background ideas that are used in various places throughout the work. In many cases this material may be familiar, however a limited discussion is provided here in an attempt to make the work self-contained.

Section one consists of some mathematical concepts which we will need in this book. In section two, we introduce a brief review of artificial neural networks to choose one, that has suitable design and architecture to implement the problem of this book.

In section three, we introduce a general idea about contamination of environment, especially soil contamination by heavy metals that we need in the formulation of the model equation. The model equation can be then used to determine the rate of environmental contamination by estimating the concentration of heavy metals in soil.

Section One: Mathematical Concepts

This section consists of some of the definitions, hypotheses, axioms and theorems which we need in this book.

1.1.1. What is Mathematical Model?

Mathematical models are simplified representations of some real world entity. It can be formulated as equations or computer code that intends to mimic essential features while leaving out inessentials. Mathematical models describe our beliefs about how the world is functioning. It aims to describe the different aspects of the real world, their interaction, and their dynamics through mathematics [36].

Mathematical models are characterized by assumptions about:

- Variables (the things which change);
- Parameters (the things which do not change);
- Functional forms (the relationship between the two).

Now, we give a brief discussion of positive definite matrices.

1.1.2. Positive-Definite Matrices

A symmetric matrix A is said to be positive definite if and only if

$$x^T Ax > 0, \forall \text{ Nonzero vectors } x \neq 0.$$

This can be a difficult condition to verify, but there are equivalent definitions that are sometimes more practical. For example, A will be positive definite if all of its eigenvalues are positive. Also, if Gaussian elimination is applied to A without pivoting to transform A into upper triangular matrix:

$A \to \mathrm{U} = \left[u_{i,j}\right]_{i,j}$ and if $u_{i,i} > 0 \ \forall$ i, then A is positive definite.

Similarly a symmetric matrix A is

- Positive semi definite if $\mathrm{x}^{\mathrm{T}}A\mathrm{x} \geq 0, \forall\ x$ or equivalently, all eigen-values of A are nonnegative;
- Negative definite if $x^T Ax < 0, \forall\, x \neq 0$ or all the eigenvalues of A are negative;
- Negative semi definite if $x^T Ax \leq 0$, $\forall\ x$ or all the eigenvalues of A are non-positive;
- Indefinite if $x^T Ax$ can take on both positive and negative values or A has both positive and negative eigenvalues.

A positive definite matrix is automatically positive semi definite, and likewise a negative definite matrix is negative semi definite.

1.1.3. Global Minimizer

A point w^* is said to be a global minimizer if $\mathrm{f}(\mathrm{w}^*) \leq \mathrm{f}(\mathrm{w}), \forall\, w$, where w ranges over all of R^{n} or at least over the domain of interest to the model.

Note, the global minimizer can be difficult to find, because our knowledge of f is usually only local, then at the least we would like to find a local minimizer.

1.1.4. Local Minimizer

A point w^* is said to be a local minimizer if there is a neighborhood $\mathcal{N}$ of w^* such that $\mathrm{f}(\mathrm{w}^*) \leq \mathrm{f}(\mathrm{w}), \forall\ \mathrm{w} \in \mathcal{N}$.

1.1.5. Strict Local Minimizer

A point w^* is a strict (also called strong) local minimizer if there is a neighborhood $\mathcal{N}$ of w^* such that $\mathrm{f}(\mathrm{w}^*) < \mathrm{f}(\mathrm{w}), \forall \mathrm{w} \in \mathcal{N}$ with $\mathrm{w} \neq \mathrm{w}^*$.

Various one-dimensional examples are illustrated in Figure (1.1).

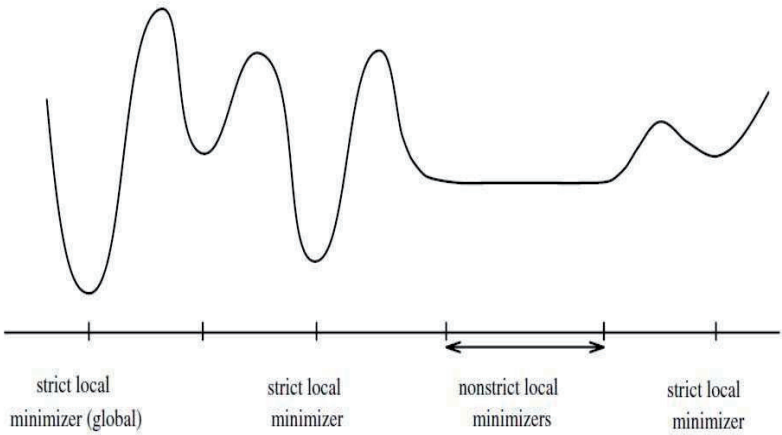

Figure 1.1: Examples of local minimizers

1.1.6. Rates of Convergence

For many optimization methods, the number of operations or steps required to find an exact solution will be infinite, so some other measures of efficiency must be used. The rate of convergence is one such measure. It describes how quickly the estimates of the solution approach the exact solution.

Let us assume that we have a sequence of point's w_k converging to a solution w*, that is $\lim_{k\to\infty} w_k = w*$.

We say that the sequence $\{w_k\}$ converges to w* with rate r and rate constant C if; $\lim_{k\to\infty} \frac{\|w_{k+1} - w^*\|}{\|w_k - w^*\|^r} = C, \text{and } C < \infty.$

When r = 1 this is referred to as **linear convergence.**

If $C \in (0, 1)$, then the distance to the solution w^* decreases at each iteration by at least a constant factor.

We say that the convergence is super linear if, r = 1 and C = 0, i.e.,

$$\lim_{k \to \infty} \frac{\|w_{k+1} - w^*\|}{\|w_k - w^*\|} = 0$$

When r = 2, the convergence is called quadratic.

Section Two: Artificial Neural Networks

1.2.1. Introduction

This section gives a brief overview of various topics in artificial neural networks (ANN)

1.2.2. What is Artificial Neural Network?

An artificial neural network (ANN) is composed of simple elements operating in parallel. These elements are inspired by biological nervous systems built to perform tasks in a similar manner that human brain works.

An ANN consists of a number of very simple and highly interconnected processors which are called neurons. These neurons are like the biological neurons in the brain. The neurons are connected by weighted links passing signals from one neuron to another. Figure (1.2) represents connections of a typical ANN.

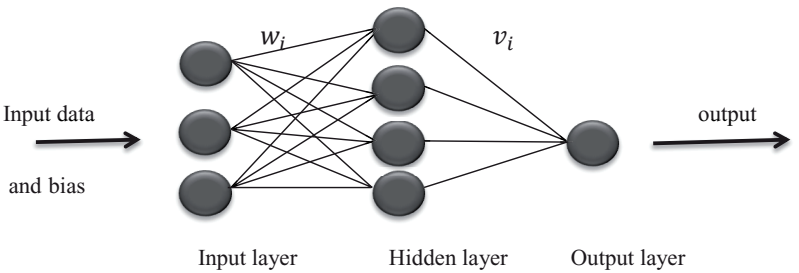

Figure 1.2: Topology of ANN

The input comes into the neuron that is weighted (each input can be individually multiplied with a weight). The neuron then sums the weighted inputs, bias and processes the sum with a transfer function. At the end a neuron

passes the processed information via output. Benefit of neuron model simplicity can be seen in its mathematical description below:

$$y_d = \varphi\left(\sum_{i=0}^{m} w_i x_i + b\right) \qquad (1.1)$$

When;

y_d is desird value, w_i is weight value, b is bias,

φ is the transfer function , x_i is input value.

1.2.3. Transfer Function

The necessary and important unit in ANN structure is a function that is a scalar for net inputs which called transfer function, or activation function and the result value of the output is called the unit's activation. A transfer functions for limiting the amplitude of the output of a neuron. Some of the most commonly used transfer functions are to solve non-linear problems.

The transfer function denoted by $\varphi: R \rightarrow R$ defines the output of a neuron, which is bounded, monotonically increasing, continuous, differentiable and satisfies:

$$\lim_{x \to +\infty} \varphi(x) = 1 \text{ and } \lim_{x \to -\infty} \varphi(x) = 0$$

The sigmoid function is by far the most common form of transfer function used in construction of ANNs. An example of the sigmoid function is the logistic function defined the range from 0 to +1 (see Figure (1.3b)), an important feature of the sigmoid function that it is differentiable [14]. Dorofki et al pointed that, one of the most commonly used functions is the log-sigmoid transfer function (logsig.), this transfer function takes the input (which may have any value between $+\infty$ and $-\infty$) and squashes the output into the range 0 to 1.

It is sometimes desirable to have the transfer function range from –1 to +1 allowing the transfer function of the sigmoid type to take negative values, in which case the transfer function assumes an anti-symmetric form with respect to

the origin, i.e., one that satisfies $\varphi(-x) = -\varphi(x)$. For example, the hyperbolic tangent transfer function (tansig.), which has a particularly simple derivative:

$\varphi(x) = \tanh(x)$, $\varphi'(x) = 1 - \varphi^2(x)$.

Throughout this work, we take tansig. as a transfer function, depending on the results of [16], which evidence that an anti-symmetric transfer function tansig. enables the training algorithm to learn faster, and in the term of ANN, is related to a bipolar sigmoid which has an output in the range of -1 to +1. As can be seen in Figure (1.3)(a), this function is, mathematically equivalent to tanh(n). It differs in that it runs faster than logsig., but the results can have very small numerical differences. This function is a good tradeoff for ANN in which the speed is more important than the exact shape of the transfer function.

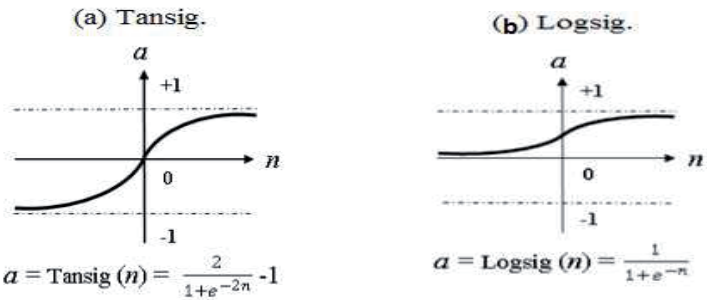

Figure 1.3: Sigmoid functions (a) tansig. and (b)logsig.

1.2.4. Architecture of ANNs

Architecture of network defined its structure including number of hidden layers, number of hidden nodes (neurons) and number of input and output nodes**.**

Typically, the network consists of an input layer of source neurons, at least one middle or hidden layer of computational neurons, and an output layer of computational neurons. The input signals are propagated in a forward direction on a layer-by-layer basis.

The hidden layers provide the network with its ability to generalize. In theory, ANN with one hidden layer and a sufficient number of hidden neurons is capable of approximating any continuous function. In practice, ANN with one and occasionally two hidden layers are widely used and has to perform very well. In general, ANNs with one hidden layer, a nonlinear transfer function and a sufficient number of hidden neurons are able to approximate any function with arbitrary precision [42]. Eqhaar [16] and Jabber [23] determined the optimal network architecture for approximation problems.

There is no magic formula for selecting the optimum number of hidden neurons. Let n be the number of input nodes. From Kolmogorov existence theorem we know that a three-layers perceptron network with n(2n+1) nodes can compute any continuous function of n variables. However, Eqhaar provides the number of hidden nodes in FFNN. He proved that any bounded continuous function in R^n can be uniformly approximated within any error by using one hidden layer network having 2n+1 units (nodes). For the architecture of FFNN with multi hidden layers Eghaar show that the number of hidden nodes in the first hidden layer is 2n+1 nodes, while the number of hidden nodes in the second hidden layers is 2(2n+1) +1 units. In general, the number of hidden nodes equal to double hidden nodes which contained in the previous layer plus one.

The numbers of input neurons are equal to the dimension of domain. The numbers of output neurons are equal to the dimension of the range.

1.2.5. Feed Forward Neural Networks

Artificial neural network with feed forward topology is called feed forward neural network (FFNN) and as such has only one condition: information must flow from input to output in only one direction with no back-loops. That is, the output from one layer of neurons feeds forward into the next layer of neurons.

There are never any backward connections and connections never skip a layer (see Figure 1.4).

In this book, the fully connected multilayer FFNN consists of a one hidden layer and one output layer. Typically, the layers are fully connected, meaning that all neurons at one layer are connected with all neurons at the next layer (see Figure 1.4). Each connection between neurons has a weight associated with it. In addition, there is a special weight that feeds into every neuron at the hidden layer and output layer, these weights are called the bias (b).

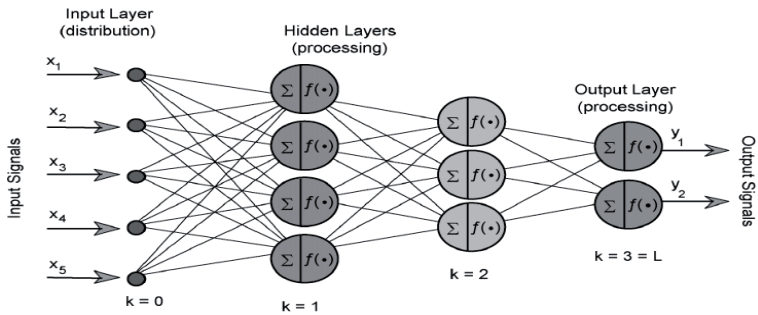

Figure 1.4: Multilayer FFNN.

1.2.6. Training of Neural Networks

The training of any process of ANNs operations is an adaptive process of the network by adjusting the weights associated with each neuron. The training is conducted by specialized algorithms to change the weights. This is called conditioned weight which happens through a process called training phase. The initial weights are selected randomly.

All training algorithms used in ANNs classified into three distinct types:

- **Supervised Training**: which incorporates an external teacher, so that each output neuron is told what its desired response to input signals ought to be. During the training process global information may be required.

- **Unsupervised Training**: uses no external teacher and is based upon only local information. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties.
- **Semi-Supervised Training**: combines supervised and unsupervised training. Part of the weights is usually determined through supervised training while the others are obtained through unsupervised training.

Throughout this work we use supervised training algorithm.

1.2.7. Back Propagation Training Algorithm

The back propagation (BP) algorithm is a mathematical technique which requires iterative processes and thus can be easily implemented on the computer. The back propagation algorithm is one of the most useful algorithms of ANN training; it is powerful for training.

The term back propagation refers to the process by which derivatives of network error with respect to network weights and biases can be computed. This process can be used with a number of different optimization strategies. The central idea is that the errors for the neurons of the hidden layer are determined by BP of the errors from the neurons of the output layer. For this reason the method is often called the back propagation learning rule. The application of the generalization BP thus involves two phases: during the first phase the input x is presented and propagated forward through the network to compute the output values y for each output unit. This output is compared with its desired value y_d, resulting in an error signal E for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated.

1.2.8. Performance Functions

Performance functions are used in supervised training to help update the network weights and biases. In supervised training, a network is provided with the desired output y_d for each input. The error is defined as the difference between the desired output and the actual network output. Network weights are updated according to one of two performance functions to reduce the network error:

1. **Least mean of squared errors (MSE)**: the average of the squared network errors.
2. **Least sum of squared errors (SSE)**: the sums of the squared network errors. it is calculated by:

$$\mathrm{E(x, w)} = \frac{1}{2}\sum_{p=1}^{P}\sum_{m=1}^{M} e_{p,m}^{2} \qquad (1.2)$$

Where,

X: is the input vector;

W: is the weight vector;

$e_{p,m}$: is the training error at output m when applying pattern p and is defined as:

$$\mathrm{e_{p,m}} = (\mathrm{y_d})_{\mathrm{p,m}} - (\mathrm{y_t})_{\mathrm{p,m}} \qquad (1.3)$$

Where,

$\mathrm{y_d}$: is the desired output;

$\mathrm{y_t}$: is the trial solution.

Section Three: Contamination

This section consists of a brief overview of environmental contamination, soil, heavy metals which we need throughout book.

1.3.1. Environment Contamination

Contamination is known as that everything that affects all elements of the environment, including the plant, animal and human being as well as everything that affects the installation of third type elements such as vivid nature: air, water and soil. This book focus on contamination of soil.

1.3.2. Soil Contamination

Soil contamination is defined as the build-up in soils of persistent toxic compounds, chemicals, salts, radioactive materials, or disease causing agents, which have adverse effects on plant growth and animal health. Soil is the thin layer of organic and inorganic materials that covers the earth's rocky surface.

The organic portion, which is derived from the decayed remains of plants and animals, is concentrated in the dark uppermost topsoil. The inorganic portion made up of rock fragments was formed over thousands of years by physical and chemical weathering of bedrock. Productive soils are necessary for agriculture to supply the world with sufficient food.

There are many different ways that the soil can be contaminated for example: seepage from a landfill, discharge of industrial waste into the soil, percolation of contaminated water into the soil, rupture of underground storage tanks, excess application of pesticides, herbicides or fertilizer, and solid waste seepage.

The most common chemicals which causing in soil contamination is: petroleum, hydrocarbons, heavy metals, pesticides, and solvents.

In this book, we focus on soil contamination by heavy metals.

1.3.3. What are the Heavy Metals?

Means all heavy metals minerals that increase density 5 g/cm^3, and at least it is called light metals. Some of these metals play an important role in the lives of the living and have different biological effects. For example the iron well-known enzymes in the blood and the installation of importance are all of the elements manganese, zinc, copper enzymatic catalysts. However, these metals are toxic and dangerous to be in certain concentrations. Adding to that the seriousness of these metals, it is not possible to be analyzed by bacteria and other natural processes as well as the authenticity of which enable them to spread over long distances for ton or sources sites. Perhaps the most dangerous thing is due to susceptibility to each bio-accumulate in the tissues and organs of living organisms in the environment water or land. In addition, some heavy metals serve as radioactive isotopes, therefore, these metals will be charged double the risk to the environment in terms of being toxic and radioactive at the same time, as is the case in 65 of radioactive zinc, uranium 235 [46, 29].

Heavy metal contamination in soil may pose risks and hazards to human beings. Excessive concentrations of some heavy metals in biological systems, especially animals (human beings in particular) are highly dangerous to human health, and may even cause death. For example, heavy metals such as: Cadmium (Cd), Nickel (Ni) and Arsenic (As) are carcinogenic. Table 1.1 gives a summary of some dangerous heavy metals that are commonly present in farm soils and their health impacts to human beings [38].

Table 1.1: Dangerous heavy metals and their health impacts to human beings[7]

Heavy Metal	Health Impact/s
Pb	Mental lapse or even death
Cr	Allergic dermatitis
As	Skin damage, cancer, affects kidney and central nervous system

Zn	Zinc shortages can cause birth defects
Cd	Affects kidney, liver and GI tract
Cu	Anemia, liver and kidney damage, and stomach/intestinal irritation
Ni	Various kinds of cancer

The sources of some heavy metals in soils are:

Cadmium: metal plating - Paints - Sticky plastic - the battery industry.

Copper: coal mining and its aftermath - fertilizer.

Lead: fuel and coal combustion - iron and steel production plants.

Nickel: oil and coal - alloying - coating of metal combustion.

Zinc: iron and galvanized steel - alloys - Batteries – rubber factories .

Chapter Two

Improve Levenberg - Marquardt Training Algorithm

Chapter Two

Improve Levenberg - Marquardt Training Algorithm

2.1. Introduction

In the last years with the development of tutorials and applications, the ANN technology has been evaluated and enhanced. Many researchers have modified the conventional learning algorithm in order to modify the efficiency of the training technique of ANN. Many works have been devoted to modify the time process and others modified the ability of generalization of the networks. Many researches have been devoted to develop the heuristic methods based on studying the properties of the conventional training algorithm. These researches include such idea as choosing the learning rate by using momentum and well chose of the activation function [42], [44],[45]. Various acceleration techniques have been proposed in heuristic technique. In this chapter, we will explain how we can improve Levenberg-Marquardt training algorithm (LMTA) and reduce its limitation.

2.2. Training Algorithm for Neural Networks

Any optimization method a local or global one can be applied to the get the optimal value of neural networks parameters such as the weights. Naturally, local searches give local solutions; once attempt to avoid this limitation. The performance of training varies depending on the network configuration and error surface for a given problem. Since the gradient information of error surface is available for the most widely applied network configurations, the most popular optimization methods have been variants of gradient based BP algorithms. Of course, this is sometimes the result of an inseparable combination of network

configuration and training algorithm which limits the freedom to choose the optimization method.

Nevertheless, global optimization algorithms may be useful for validation of an optimal solution achieved by BP training algorithm.

In the training process the parameters of the network are updated as the following rule:

$$w_{k+1} = w_k + \eta\rho_k \qquad (2.1)$$

Where;

η: is the learning rate;

ρ_k: Search direction computed by training algorithm.

The learning rate η is used to determine the length of the weight update in most of the training algorithms.

The training progress of the standard BP algorithm is given as follows. Here, assume that there are *n* nodes in the first layer (input layer), the second layer (hidden layer) consist *m* nodes, and the third layer (output layer) consist one node.

Step 0: ***Weights initialization*****.**

Randomly initialize weights and biases, respectively, on all the connections (paths) in the ANN.

Feed forward*:*

Step 1: ***Distributing inputs feed forwarding***.

Each inputs (***x***: x_1, x_2, x_3,…, x_n) and (***t***: t_1,t_2,t_3,….,t_n) are sent to all neurons in the hidden layer.

Step 2: ***Summation in each neuron in hidden layer*****.**

Each neuron in this layer computes the summing for the inner product of its weights and bias with the inputs.

Step 3: ***The inner product transferring in the hidden layer*****.**

Take suitable transfer function for summing the value in step 2. Then its output is sent to the output layer as an input.

Step 4: ***Summation at the adder in each neuron in output layer*****.**

The neurons in this layer, sums its weighted inputs.

Step 5: ***The inner product transferring in extern layer***.

The transfer function of the neuron in this layer computes its output which is also the output of overall network.

Step 6: ***Squared error computation*****.**

After the output neuron computed the output, then the error can be calculated.

Step 7: ***Checking halt condition.***

If the error is less than a specific value, the training stops and the weights and biases are sent. Otherwise, the training proceeds to the next step.

Back propagation**:**

Step 8: ***Sensitivity computation for the output layer*****.**

Sensitivity is the change of the index of performance responding to the input in the output layer.

Step 9: ***The correction law for the weights calculation in the output layer*****.**

According to the choosing training rule, the correction laws of the weights between the hidden layer and the output layer are calculated.

Step 10: ***The sensitivity propagation*****.**

The sensitivity term of the output is sent back to the hidden layer. Each hidden unit gets its input from the output layer, and then the sensitivity is computed for updating the weights in the hidden layer. Because there are *m* nodes in the hidden layer, the sensitivity term is a vector.

Step11: The correction for the weights calculation in the output layer

Now, the correction laws for the weight and bias are calculated.

Update weights and biases**:**

Step 12: ***The output unit updates the parameters*****.**

At the end of each iteration, the weights and bias will be renewed. The renewed weights and bias are as follows:

$v_{kcurrent} = v_{kprevious} + \Delta v_k$

$b_{current} = b_{previous} + \Delta v_k$,

where, "current" means the current iteration, and "previous" means the previous iteration, v is the weight of the output layer, and b is the bias.

Step 13: ***Each hidden unit updates its bias and weights as follows*****:**

$w_{kcurrent} = w_{kprevious} + \Delta w_k$

$b_{kcurrent} = b_{kprevious} + \Delta w_k$

where, w is the weight of the hidden layer, and b is the bias.

Step 14: ***Return to step 2 for next iteration*****.**

2.3. Levenberg – Marquardt Method

Levenberg – Marquardt method (LM) was formulated to approach 2nd order training speed without having to compute the Hessian matrix, when the performance function has the form of a sum of squares. In this case, Hessian matrix can be approximated as $H=J^TJ$ and the gradient can be computed as g= $J^TE(w_k)$, where J is the Jacobian matrix which contains the first derivatives of the error function (performance function) with respect to the weights and $E(w_k)$, is a vector of network errors. The LM method uses the following approximation to the Hessian matrix:

$$H \approx J^TJ + \mu I \qquad (2.2)$$

Where;

μ: is always non negative, called the combination coefficient,

I: is the identity matrix.

From equation (2.2), one may notice that the elements on the main diagonal of the approximated Hessian matrix will be larger than zero.

We use this approximation of H (equation (2.2)) in the Newton update as the following:

$$w_{k+1} = w_k - \left(J_k{}^T J_k + \mu_k I\right)^{-1} J_k^T \, e(w_k) \qquad (2.3)$$

When the scalar $\mu = 0$, this is just Newton's method, when μ is large, this becomes gradient descent with a small step size.

2.4. The cost of Using Levenberg - Marquardt's Method

There are many types of costs associated with using LM method: derivatives, computation and storage. In its classical form, LM method requires second derivatives, the solution of a linear system, and the storage of a matrix. For an n-variable problem, there are $O(n^2)$ entries in the Hessian matrix (H), meaning that $O(n^2)$ expressions must be programmed to compute these derivatives that is the exact evaluation of the H is computationally intensive. The computation of the inverse Hessian (H^{-1}) is even more computationally intensive. The other costs of LM method are the computational costs of applying the method. If there are n variables, and if the problem is not sparse, calculating the H involves calculating and storing about n^2entries, and solving a linear system requires about n^3arithmetic operations. If n is small, these costs might be acceptable. For n of moderate size (say, $n < 200$) storing this matrix might be acceptable, but solving a linear system might not be. Also for large n (say, $n > 1000$), even storing this matrix might be undesirable.

For this limitation, LM method can fail to converge, or it can converge to a point that is not a minimum.

To overcoming these disadvantages we suggest the following improvement.

2.5. Improve Levenberg-Marquardt Training Algorithm

Consider the nonlinear performance equations:

$$E(w) = 0 \qquad (2.4)$$

where $E(w): R^m \rightarrow R^n$ is continuously differentiable Lipschitz function. The latter assumption means that $\|E(w_2) - E(w_1)\| \leq L\|w_2 - w_1\|$ where L is called Lipschitz constant. Suppose that the equation (2.4) has nonempty solution w^*. We refer $\|\cdot\|$ to the 2nd norm in all cases. With a suitable choose of the parameter μ we prove that if $\|E(w)\|$ gives the error bound for some $w^* \in W$, then the sequence $\{w_k\}$ generated using the modified LM algorithm converges super linearly to the solution of equation (2.4).

Now, we will introduce some of the definitions, hypotheses and theorems to help the prove the convergence of modified LM algorithm.

Definition 2.1

A **local minimizer** is a point **w∗** that satisfies the condition:

$f(w*) \leq f(w)$, $\forall$ w, such that $\|w - w*\| < \varepsilon$,

where ε is some small positive number whose value may depend on w∗. Similarly, defined is a **strict local minimizer**:

$f(w*) < f(w)$ for all w such that $0 < \|w - w*\| < \varepsilon$.

Note that, it is possible for a function to have a local minimizer and yet have no global minimizer. It is also possible to have neither global nor local minimizers, to have both global and local minimizers, to have multiple global minimizers, and various other combinations [19]. In this chapter, we will attempt to get network parameters which are local minimizers for the performance function (2.4). The main aim is to get optimal parameters which are both global and local

minimizers. This is hold when the performance function in (2.4) satisfies the first-order necessary condition for a minimizer.

Note here $\nabla E(w*) = 0$, is referred to as the first-order necessary condition for a minimizer. That is the gradient of performance function at the optimal weights are equal to zero, i.e., $g(w*) = 0$.

Throughout this work we use**:** $\nabla E(w^*) = g(w^*)$.

It is a "necessary" condition since if $w*$ is a local minimizer, then it "necessarily" satisfies $g(w*) = 0$. The condition is not "sufficient" to determine a local minimizer since a point satisfying $g(w*) = 0$, could be a local minimizer, a local maximizer, or a saddle point (a stationary point that is neither a minimizer nor a maximizer).

Local minimizers can be distinguished from other stationary points by examining 2nd derivatives depending on Taylor series expansion at $w = w* + p$, and using the result that $g(w*) = 0$. We see that $\nabla^2 E(w*) = H$ must be positive semi definite. If not, then $v^T H(w*)v < 0$, for some v. Then it is also true that $v^T H(\xi)v < 0$ if $\| \xi - w* \|$ is small. This is because H is assumed to be continuous at $w*$. If p is chosen as some sufficiently small multiple of v, then the point ξ will be close enough to $w*$ to guarantee (via the Taylor series) that $E(w) < E(w*)$, a contradiction. Hence if $w*$ is a local minimizer, then $H(w*)$ is positive semi definite. This is referred to as the second-order necessary condition for a minimizer, with the "second-order" referring to the use of 2nd derivatives or the 2nd order term in the Taylor series.

There is also a 2nd order sufficient condition, "sufficient" to guarantee that $w*$ is a local minimizer:

If $g(w*) = 0$ and $H(w*)$ is positive definite,

We need the following theorem:

Theorem 2.2

Let F(x) be an m-vector of functions of $x = (x_1, \dots, x_n)^T$. Assume that the Jacobian of F is Lipschitz continuous on an open convex set S with constant L. Then for any $x, y \in S$,

$\|F(y) - F(x) - \nabla F(x)^T (y - x)\| \leq \frac{L}{2} \|y - x\|^2$ [19].

In this section we show how to use a "modified matrix factorization" to guarantee a descend direction for LM method. (Additional requirements needed to guarantee convergence).

In the classical LM method the search direction is defined by:

$\rho_k = -\left(J_k^T J_k + \mu_k I\right)^{-1} g(w_k)$, i.e.,

$$\left(J_k^T J_k + \mu_k I\right)\rho_k = -g(w_k) \qquad (2.5)$$

Where ρ_k is the search direction calculated in step k, $g(w_k)$ is the gradient of the performance function with respect to the weights for the step k, which can be computed as $g(w_k) = J^T(w_k).\, e(w_k)$, and J is the Jacobian matrix[19].

Now, to guarantee the convergence of LM algorithm, we must satisfy the following condition (descent condition):

$$E(w_k + \eta\rho_k) < E(w_k), \quad \eta > 0 \qquad (2.6)$$

This is possible if the search direction is a descent direction, that is, if $\rho_k^T g(w_k) < 0$, use a "modified matrix factorization" to guarantee this for LM algorithm, so we suggest the following procedures:

In the LM algorithm the search direction is defined by (depending on (2.5):

$$\rho_k = -\left(J_k^T J_k + \mu_k I\right)^{-1} g(w_k), \qquad (2.7)$$

If ρ is to be a descent direction at the point w, it must satisfy:

$$\rho_k{}^{\mathrm{T}} g(w_k) = -g(w_k)\left(J_k{}^T J_k + \mu_k I\right)^{-1} g(w_k) < 0, \qquad (2.8)$$

Or

$$g(w_k)\left(J_k{}^T J_k + \mu_k I\right)^{-1} g(w_k) > 0, \qquad (2.9)$$

This condition will be satisfied if $(J^T J + \mu I)^{-1}$ or equivalently $(J^T J + \mu I)$, is positive definite. Requiring that $(J^T J + \mu I)$, be positive definite is a stronger condition than $\rho_k{}^{\mathrm{T}} g(w_k) < 0$,

To motivate this, recall that LM algorithm can be interpreted as approximating $E(w_k + \rho)$ by a quadratic form:

$$E(w_k + \rho) \approx E(w_k) + \rho^{\mathrm{T}} g(w_k) + \frac{1}{2}\ \rho^{\mathrm{T}}\left(J_k{}^T J_k + \mu_k I\right)\rho, \qquad (2.10)$$

An alternative view is to minimize the quadratic as a function of ρ. If $J^T J + \mu I$ is positive definite, then the minimum is obtained by setting the derivative equal to zero. If $J^T J + \mu I$is indefinite, then the quadratic function does not have a finite minimum.

If the Hessian matrix is indefinite, then one possible strategy is to replace the Hessian by some related positive definite matrix. This guarantees that the ρ is a descent direction. It also implies that the ρ corresponds to the minimization of a quadratic approximation of the objective function E, this quadratic approximation is obtained from the Taylor series by replacing $J^T J + \mu I$ with the related positive definite matrix.

Now, in equation (2.5), if $(J^{\mathrm{T}}J + \mu I)$ is indefinite, then the factorization

$$J^T J + \mu I = LU, \qquad (2.11)$$

Where L is a lower triangular matrix and U is upper triangular matrix. Let D be the diagonal matrix whose entries are the diagonal entries of U, that is $d_{i,i} = u_{i,i}$, and define $\tilde{U} = D^{-1}U$, so that, $D\tilde{U} = U$. Hence:

$$J^T J + \mu I = L\, D\tilde{U} \tag{2.12}$$

If ($J^T J + \mu I$) is positive definite, then it is also symmetric, so that:

$$(J^T J + \mu I)^T = \tilde{U}^T D^T L^T = \tilde{U}^T D L^T = J^T J + \mu I = L\, D\tilde{U} \tag{2.13}$$

Then, it is easy to verify that $\tilde{U} = L^T$, so that:

$$J^T J + \mu I = L\, D L^T \tag{2.14}$$

This is the new factorizations when $(J^T J + \mu I)$ is a positive definite and the diagonal matrix D has a positive diagonal entries.

If $(J^T J + \mu I)$ is not positive definite, then at some point during the computation of the factorization some diagonal entry of D will satisfy $d_{i,i} \leq 0$, if this happens, then $d_{i,i}$ should be replaced by some positive entry, perhaps $|d_{i,i}|$ or some small positive number.

It can be shown (via the formulas for the matrix factorization) that modifying the entries of D is equivalent to replacing ($J^T J + \mu I$) by:

$$J^T J + \mu I \rightarrow J^T J + \mu I + A, \tag{2.15}$$

Where A is a diagonal matrix, and then factoring this matrix,

$$J^T J + \mu I + A = L\, D L^T, \tag{2.16}$$

and so the modified Hessian matrix is positive definite. Then, this factorization is used to compute the search direction ρ (in equation (2.5)) as the following:

$$(\mathrm{L\,DL^T})\rho_K = -\mathrm{g}(\mathrm{w}_k), \qquad (2.17)$$

and hence the overall technique corresponds to replacing $(\mathrm{J^TJ} + \mu\mathrm{I})$ by the related positive definite matrix $(\mathrm{J^TJ} + \mu\mathrm{I} + \mathrm{A})$. Even if $(\mathrm{J^TJ} + \mu\mathrm{I})$ were always positive definite, this matrix would still be factored to compute the ρ and so this "modified" matrix factorization is obtained with little effort, just the effort of changing any negative (or zero) $\mathrm{d_{i,i}}$ to a suitable positive number. Figure (2.1), illustrate the flowchart of modified LM algorithm.

The following theorem guarantees that, if $(\mathrm{J^TJ} + \mu\mathrm{I})$ or its modification $(\mathrm{J^TJ} + \mu\mathrm{I} + \mathrm{A})$ is positive definite, then the sequence $\{w_k\}$ generated by using the modified LM algorithm converges super linearly to the solution of equation (2.4).

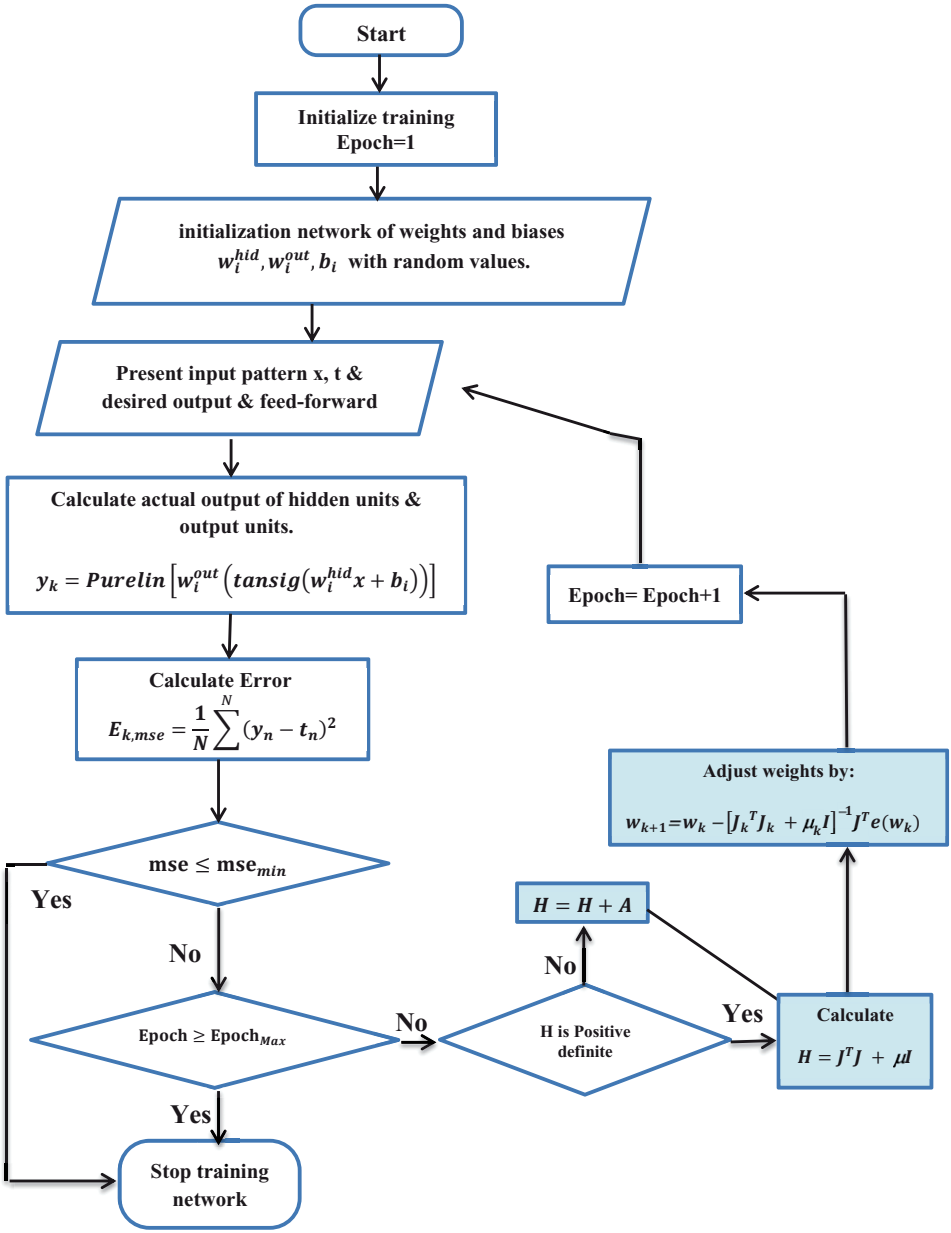

Figure 2.1: Flowchart of modified Levenberg-Marquardt algorithm.

Theorem 2.3 (Super Linear Convergence).

Let E be a real-valued function of n variables defined on an open convex set S. Assume that $J^TJ + \mu I$ is positive definite Lipschitz continuous. Consider the sequence $\{ w_k\}$ generated by: $w_{k+1} = w_k + \rho_k$, where ρ_k is the search direction computed by LM algorithm, suppose that $\{w_k\} \subset S$, $\lim_{k\to\infty} w_k = w_* \in S$, and that $w_k \neq w_*$ for all k. Then

i) $\{w_k\}$ converges to w_* superlinearly

ii) $g(w_*) = 0$

If and only if,

$$\lim_{k\to\infty} \frac{\|\rho_k - (\rho_N)_k\|}{\|\rho_k\|} = 0$$

where $(\rho_N)_k$ is computed by improved LM algorithm at w_k.

Proof

(1) Assuming that $\lim_{k\to\infty} \frac{\|\rho_k-(\rho_N)_k\|}{\|\rho_k\|} = 0$, is true. We must show:

i) $\{w_k\}$ Converges to w_* superlinearly.

ii) $g(w_*) = 0$.

Now, we prove (i):

From the assumptions on g and its derivatives, there exist an $L > 0$, such that:

$$\|g(w_{k+1})\| = \|g(w_{k+1}) - 0\|$$

$$= \|g(w_{k+1}) - g(w^*)\| \leq L\|w_{k+1} - w^*\| \text{ (by Lipschitz continuous)}$$

for all sufficiently large values of k hence:

$$\frac{\|g(w_{k+1})\|}{\|\rho_k\|} \leq \frac{L\|w_{k+1} - w^*\|}{\|\rho_k\|}$$

$$= \frac{L\|w_{k+1} - w^*\|}{\|w_{k+1} - w_k\|} \quad (\text{since, } \|\rho_k\| = \|w_{k+1} - w_k\|)$$

$$= \frac{L\|w_{k+1} - w^*\|}{\|w_{k+1} - w^* + w^* - w_k\|}$$

$$= \frac{L\|w_{k+1} - w^*\|}{\|w_{k+1} - w^* - (w_k - w^*)\|}$$

$$\geq \frac{L\|w_{k+1} - w^*\|}{\|w_{k+1} - w^*\| + \|w_k - w^*\|}$$

$$= \frac{\frac{L\|w_{k+1} - w^*\|}{\|w_k - w^*\|}}{\frac{\|w_{k+1} - w^*\| + \|w_k - w^*\|}{\|w_k - w^*\|}}$$

$$= \frac{\frac{L\|w_{k+1} - w^*\|}{\|w_k - w^*\|}}{\frac{\|w_{k+1} - w^*\|}{\|w_k - w^*\|} + 1}$$

But since,

$$\lim_{k \to \infty} \frac{\|g(w_{k+1})\|}{\|\rho_k\|} = 0,$$

it follows that:

$$\lim_{k \to \infty} \frac{L\|w_{k+1} - w^*\|}{\|w_k - w^*\|} = 0,$$

That is, $\{w_k\}$ converge superlinearly to w^*.

Now to prove (ii) $g(w^*) = 0$

$$g(w_{k+1}) = g(w_{k+1}) - H(w_k)[\rho_k - (\rho_N)_k] + H(w_k)[\rho_k - (\rho_N)_k]$$

$$= g(w_{k+1}) - H(w_k)\rho_k + H(w_k)(\rho_N)_k + H(w_k)[\rho_k - (\rho_N)_k]$$

and since, $-g(w_k) = H(w_k)(\rho_N)_k$ and $w_{k+1} = w_k + \rho_k$, we have:

$$= g(w_{k+1}) - H(w_k)(w_{k+1} - w_k) - g(w_k) + H(w_k)[\rho_k - (\rho_N)_k]$$

Then,

$$\frac{g(w_{k+1})}{\|\rho_k\|} = \frac{g(w_{k+1}) - H(w_k)(w_{k+1} - w_k) - g(w_k) + H(w_k)[\rho_k - (\rho_N)_k]}{\|\rho_k\|}$$

$$\frac{\|g(w_{k+1})\|}{\|\rho_k\|} = \frac{\|g(w_{k+1}) - H(w_k)(w_{k+1} - w_k) - g(w_k) + H(w_k)[\rho_k - (\rho_N)_k]\|}{\|\rho_k\|}$$

$$\frac{\|g(w_{k+1})\|}{\|\rho_k\|} \leq \frac{\|g(w_{k+1}) - g(w_k) - H(w_k)(w_{k+1} - w_k)\|}{\|\rho_k\|} + \|H(w_k)\| \frac{\|[\rho_k - (\rho_N)_k]\|}{\|\rho_k\|}$$

But $\|g(w_{k+1}) - g(w_k) - H(w_k)(w_{k+1} - w_k)\| \leq \frac{L}{2}\|\rho_k\|^2$ (by theorem(2.2))

Then:

$$\lim_{k\to\infty} \frac{\|g(w_{k+1})\|}{\|\rho_k\|} \leq \lim_{k\to\infty} \frac{L\|\rho_k\|^2}{2\|\rho_k\|} + \lim_{k\to\infty} \|H(w_k)\| \frac{\|[\rho_k - (\rho_N)_k]\|}{\|\rho_k\|} = 0$$

The right hand side of the above inequality is zero because $\|H(w_k)\|$ has a finite limit because of the continuity assumptions and since $\lim_{k\to\infty} \|\rho_k\| = 0$, then

$$g(w^*) = \lim_{k\to\infty} g(w_k) = 0.$$

Then, the prove of the second part complete.

(2) Assume (i) and (ii), to prove:

$$\lim_{k\to\infty} \frac{\|\rho_K - (\rho_N)_K\|}{\|\rho_K\|} = 0$$

Now, there exists a constant $L > 0$ such that:

$\|g(w_{k+1})\| = \|g(w_{k+1}) - 0\|$ (from (ii))

$= \|g(w_{k+1}) - g(w_*)\| \leq L\|w_{k+1} - w_*\|$ (by Lipschitz continuous), for all sufficiently large k.

$$\lim_{k\to\infty} \frac{\|w_{k+1} - w^*\|}{\|w_k - w^*\|} = 0 \qquad \text{(from (i))}$$

$$\geq \lim_{k\to\infty} \frac{1}{L} \frac{\|g(w_{k+1})\|}{\|w_k - w^*\|}$$

$$= \lim_{k\to\infty} \frac{1}{L} \frac{\|g(w_{k+1})\|}{\|w_k - w^*\|} \frac{\|\rho_K\|}{\|\rho_K\|}$$

$$= \lim_{k\to\infty} \frac{1}{L} \frac{\|g(w_{k+1})\|}{\|\rho_K\|} \frac{\|w_{k+1} - w_k\|}{\|w_k - w^*\|}$$

Since, $\lim_{k\to\infty} \frac{\|w_{k+1} - w_k\|}{\|w_k - w^*\|} = 1,$

We obtain that

$$\lim_{k\to\infty} \frac{\|g(w_{k+1})\|}{\|\rho_K\|} = 0\,.$$

Now, by an argument similar to that used in step (ii) of the first half of the proof,

$$\frac{\|H(w_k)[\rho_k - (\rho_N)_k]\|}{\|\rho_K\|} \leq \frac{\|g(w_{k+1}) - g(w_k) - H(w_k)(w_{k+1} - w_k)\|}{\|\rho_k\|}$$

$$+ \frac{\|g(w_{k+1})\|}{\|\rho_k\|}\,.$$

Since the limit of the right-hand side is zero, we obtain that

$$\lim_{k\to\infty} \frac{\|H(w_k)[\rho_k - (\rho_N)_k]\|}{\|\rho_K\|} = 0\,.$$

Since $H(w^*)$ is positive definite, then $H(w_k)$ is positive definite for large values of k and hence:

$$\lim_{k\to\infty} \frac{\|[\rho_k - (\rho_N)_k]\|}{\|\rho_K\|} = 0$$

Now, in the following we study how to choose suitable learning rate.

2.6. Suitable Choose of Learning Rate

The learning rate (η) is a parameter governing the speed of learning and controlling the distance between the weights at different iterations.
Choosing a suitable value of the η perform to speed the training and then speed the convergence, also it's help the network to avoid the local minima problem.

In this book, we choose η that satisfy the following sufficient decrease condition:

$$E(w_k + \eta_k \Delta w_k) \le E(w_k) + \mu \eta_k \Delta w_k \qquad (2.18)$$

where μ is the combination coefficient in LM algorithm, $0 < \mu < 1$,

2.7. Improve the Generalization of Suggested Network

One of the problems that occur during the training process is called over fitting. This occur when the error on the training set become very small value, but when input new data to the ANN the error is large.

There are two methods for modifying generalization of ANN which are implemented in this work: regularization and early stopping.

The next subsections will describe these two techniques, and the routines to implement them.

2.7.1. Regularization

This involves modifying the performance function, which is normally chosen to be the sum of squares of the network errors on the training set:

$$E = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}((y_a)_i - (y_t)_i)^2 \qquad (2.19)$$

The most method to avoid over fitting problem is to change or modify the performance function to achieve best generalization.

In this book, we modified the performance function by adding a term that consists of the mean of the sum of squares of the network weights:

$$msereg = \gamma\ mse + (1-\gamma)\ msw \quad (2.20)$$

where γ is the performance ratio, and

$$msw = \frac{1}{n}\sum_{j=1}^{n} w_j^2 \quad (2.21)$$

2.7.2. Early Stopping

Another method for modifying generalization is called early stopping, in this technique the available data is divided into three subsets. The first subset is the training set which is used for computing the gradient and updating the network parameters. The second subset is the validation set; the error on the validation set is observed during the training process. The validation error will normally decrease during the initial phase of training, as does the training set error. However, when the network begins to over fit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned.

The third subset is test set, the test set error is not used during the training, but it is used to compare different models.

Chapter Three

Formulation Mathematical Model and Design Neural Networks

Chapter Three

Formulation Mathematical Model and Design Neural Networks

3.1. Introduction

In this chapter, we describe and develop a mathematical model for spread of contamination through soils which can be used to determine the rate of contamination by estimating the concentration of heavy metals in soil. We then design suitable networks to solve this model equation where the solution of the equation represents the concentration of heavy metals in soil for any depth and time without spending in traditional laboratory inspecting.

3.2. Mathematical Model

This research book aims at developing a mathematical model equation that estimate the rate of concentration of heavy metals in soil. This aim can be achieved through the realization of the following objectives:

- Collection of data showing the concentration of heavy metals at different percentage of the soil with respect to distance and time.
- Development of mathematical model equations for the concentration of heavy metals in the soil.
- Simulation of the model equation using computer software program Mat lab 2014 professional.
- Compare the simulated result with the experimental data.

3.3. Convective – Diffusive Equation

For many years, potentially harmful substances have been added to the soil through land application of agricultural chemicals, industrial wastewater and sludge disposal, landfills, and leaking hazardous waste storage sites. The

potentially harmful substances including heavy metals, pesticides and other industrial organic chemicals, and even plant nutrient supplements may contaminate soils, surface water bodies, and subsurface aquifers. Thus, concerns about soil and water quality have led to an increase interest in understanding the processes of solute reactions and transport in soils.

Models that include retention and release reactions of solutes with the soil matrix are needed. Retention and release reactions in soils include precipitation / dissolution, ion exchange, and adsorption / desorption reactions (Amacher et al. 1986 [5]). Retention and release are influenced by number of soil properties including texture, bulk density, power of hydrogen (PH), Electric Conductivity (EC), organic matter, and type and amount of clay minerals. Then the model equation which describes this problem is said to be convective – diffusive equation and has the form:

$$\frac{\partial C}{\partial t} + v\frac{\partial C}{\partial x} = \frac{\partial}{\partial x}\left(D_L \frac{\partial C}{\partial x}\right) - f - \frac{\partial C_a}{\partial t} \qquad (3.1)$$

Where;

C: Solute concentration ($\frac{mg}{L}$)

v: Darcy's flux ($\frac{cm}{hr}$)

D_L: Hydrodynamic dispersion coefficient ($\frac{cm^2}{hr}$)

x: Soil depth (cm)

t: Time (day^{-1})

C_a: Concentration of adsorbed chemical

f: Irreversible reaction decay rate.

3.4. Modification of Mathematical Model

The convective – diffusive equation described in the previous section consists of description of the contamination of soils by sewage sludge. Heavy metals released from sewage sludge are distributed throughout the soil system

remain in the soil solution as iron, organic and inorganic complexes. Some of these heavy metals are mobile for uptake by plants. This mobility and availability depends on several factor including soil texture and pH (Nouri., 1980 [33] and Alloway, 1995 [3]). Several researchers have already investigated the mobility of heavy metal in the soil amended with sewage sludge and concluded that only relatively small amount of metal were available for transport in the soil's water immediately after sludge application which add to the model equation (3.1) as a modification. Also, in the industrial regions, where some types of factories are active, several chemical and petrochemical processes would be also active. As a result, industrial water becomes contaminated with various substances which are harmful; these are sources of environmental contamination, add in the description the modify model equation. Then model equation (3.1), which can be written as:

$$\frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x} - \frac{\rho_b}{\theta} \frac{\partial S}{\partial t} \qquad (3.2)$$

Where;

$S = k_d . C$

C: Solute concentration ($\frac{mg}{L}$)

K_d : Distribution coefficient ($\frac{cm^2}{Kg}$)

θ: Soil's water content

ρ_b: Soil's bulk density

V_x: Darcy's flux ($\frac{cm}{hr}$)

D_L: Hydrodynamic dispersion coefficient ($\frac{cm^2}{hr}$)

x: Soil depth (cm)

t: Time (day^{-1})

Simply, we can write equation (3.2) as following:

$$\frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x} - \frac{\rho_b}{\theta} \frac{\partial}{\partial t} (K_d C) \qquad (3.3)$$

Since K_d is a constant, so rewrite equation (3.3) as following:

$$\frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x} - \frac{K_d \rho_b}{\theta} \frac{\partial C}{\partial t} \quad (3.4)$$

And rearranging we get:

$$\frac{\partial C}{\partial t} + \frac{K_d \rho_b}{\theta} \frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x} \quad (3.5)$$

Then we get:

$$\frac{\partial}{\partial t} C(1 + \frac{K_d \rho_b}{\theta}) = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x} \quad (3.6)$$

That is:

$$(1 + \frac{K_d \rho_b}{\theta}) \frac{\partial C}{\partial t} = D_L \frac{\partial^2}{\partial x^2} C - V_x \frac{\partial}{\partial x} C \quad (3.7)$$

Let $\rho = (1 + \frac{K_d \rho_b}{\theta})$ (3.8)

So, equation (3.7) can be written as:

$$\rho \frac{\partial C}{\partial t} = D_L \frac{\partial^2 C}{\partial x^2} - V_x \frac{\partial C}{\partial x}, \quad 0 < x < \infty, \quad t > 0 \quad (3.9)$$

Which is a second order linear PDE, with initial - boundary conditions:

$C(0, t) = C_0$ and $\frac{\partial C}{\partial x}(\infty, t) = 0.$

$C(x, 0) = C_x$

Where;

C_0: Initial concentration.

C_x: Concentration for depth x ($\frac{mg}{L}$)

V_x: The average pore – water velocity, ($\frac{cm}{hr}$)

x: Soil depth (distance) (cm).

t : Time (day^{-1}).

ρ: Retardation factor. ($\frac{g}{cm^3}$)

The amount of each element retained by each soil ($\frac{mg}{kg}$) was calculated from the initial concentration in solution ($\frac{mg}{L}$) and the final concentration C in

solution ($\frac{\mathrm{mg}}{\mathrm{L}}$). Equation (3.9), which can be represented as a mathematical model for spread of contamination through soils which can be used to determine the rate of contamination. The solution of model equation gives the concentration of the heavy metals in soil at any distances and time. Consider this as an important model to give concentrations of heavy metals without spending in traditional laboratory inspecting.

3.5. Solution of Model Equation

In this section, we suggest suitable design for neural network to solve equation (3.9), that is we suggest two ANN of type FFNN. The first FFNN is used to find the suitable soil parameters which are needed in equation (3.9). While the second FFNN used to estimate the concentration of heavy metals in soil.

3.6. Design the First Network

This section illustrates the design of first ANN. we suggest a multilayer FFNN which consists of three layers: input layer consists of seven input nodes, one hidden layer consists of fifteen hidden nodes with tansig. transfer function and output layer consists of three output nodes with purelin. transfer function. The output of the suggested FFNN, which represents the trial solution, can be written as:

$$\mathrm{C_d} = \mathrm{N}\,(x, t, \mathrm{c_0}, \mathrm{D}_L, \mathrm{V}_x, \rho, \mathrm{p})$$

Where; x, t, D_L, V_x, ρ, $\mathrm{c_0}$ are defined in equation (3.9), and p represent the weights.

Then;

$$\mathrm{N} = \sum_{\mathrm{i}=1}^{15} \mathrm{v_i}\,\mathrm{tansig.}\,(\mathrm{z_i}), \quad \text{where } \mathrm{z_i} = \sum_{\mathrm{j}=1}^{7} \mathrm{w_{ij}}(x_\mathrm{j}, t_\mathrm{j}, \mathrm{c_{0j}}, \mathrm{D}_{L_\mathrm{j}}, \mathrm{V}_{x_\mathrm{j}}, \rho_j) + \mathrm{b_i} \qquad (3.10)$$

$\mathrm{w_{ij}}$: denotes the weight connecting input node j to hidden node i,

v_i : denotes the weight connecting hidden node i to output node,

b_i : denotes the bias of hidden node i.

The quantity of error which be minimized is given as:

$$E(p) = \sum_{i=1}^{n} \left\{ \frac{\partial^2 C_d}{\partial x^2} - f\left(x_i , t_i , \frac{\partial C_d}{\partial t}, \frac{\partial C_d}{\partial x}, \rho, D_L, V_x \right) \right\}^2 \qquad (3.11)$$

Where; $0 < x < \infty, t > 0$, and:

$$\frac{\partial^2 C_d}{\partial x^2} = \frac{\partial^2 N\,(x, t, c_0, \rho, D_L, V_x, p)}{\partial x^2};$$

$$\frac{\partial C_d}{\partial t} = \frac{\partial N\,(x, t, c_0, \rho, D_L, V_x, p)}{\partial t}$$

$$\frac{\partial C_d}{\partial x} = \frac{\partial N\,(x, t, c_0, \rho, D_L, V_x, p)}{\partial x};$$

$$\frac{\partial^2 C}{\partial x^2} = \frac{\partial f\left(x_i , t_i , \frac{\partial C}{\partial t}, \frac{\partial C}{\partial x}, \rho, D_L, V_x \right)}{\partial x^2} = \frac{\rho \frac{\partial C}{\partial t} + V_x \frac{\partial C}{\partial x}}{D_L};$$

The architecture of the suggested design illustrate in Figure (3.1).

3.7. Design the Second Network

This section illustrates the design of the second ANN. we suggest a multilayer FFNN which consist three layers: input layer consist five input nodes, one hidden layer consist eleven hidden nodes with tansig. transfer function and output layer consist one output node with purelin. transfer function. The output of the suggest FFNN which represent the trial solution can be written as:

$$C = N\,(x, t, c_0, s, p)$$

Where; *x, t*, c_0 are defined in equation (3.9), and s is the soil parameters defined in equation (3.9), p represent the weights.

Then;

$$N = \sum_{i=1}^{11} v_i \text{tansig.}(z_i), \quad \text{where } z_i = \sum_{j=1}^{5} w_{ij}(x_j, t_j, c_{0j}, s_j) + b_i \qquad (3.12)$$

Where; w_{ij}, v_i, b_i defined as in equation (3.10).

The quantity of error which can be minimized is given as:

$$E(p) = \sum_{i=1}^{n} \left\{ \frac{\partial^2 C}{\partial x^2} - f\left(x_i , t_i , \frac{\partial C}{\partial t}, \frac{\partial C}{\partial x}, s \right) \right\}^2 \qquad (3.13)$$

Where; $0 < x < \infty, t > 0$, so:

$$\frac{\partial^2 C}{\partial x^2} = \frac{\partial^2 N\,(x, t, c_0, s, p)}{\partial x^2};$$

$$\frac{\partial C}{\partial t} = \frac{\partial N\,(x, t, c_0, s, p)}{\partial t};$$

$$\frac{\partial C}{\partial x} = \frac{\partial N\,(x, t, c_0, s, p)}{\partial x};$$

$$\frac{\partial^2 C}{\partial x^2} = \frac{\partial f\left(x_i , t_i , \frac{\partial C}{\partial t} , \frac{\partial C}{\partial x} , s \right)}{\partial x^2} = \frac{\rho \frac{\partial C}{\partial t} + V_x \frac{\partial C}{\partial x}}{D_L};$$

The architecture of the suggested design illustrate in Figure (3.2).

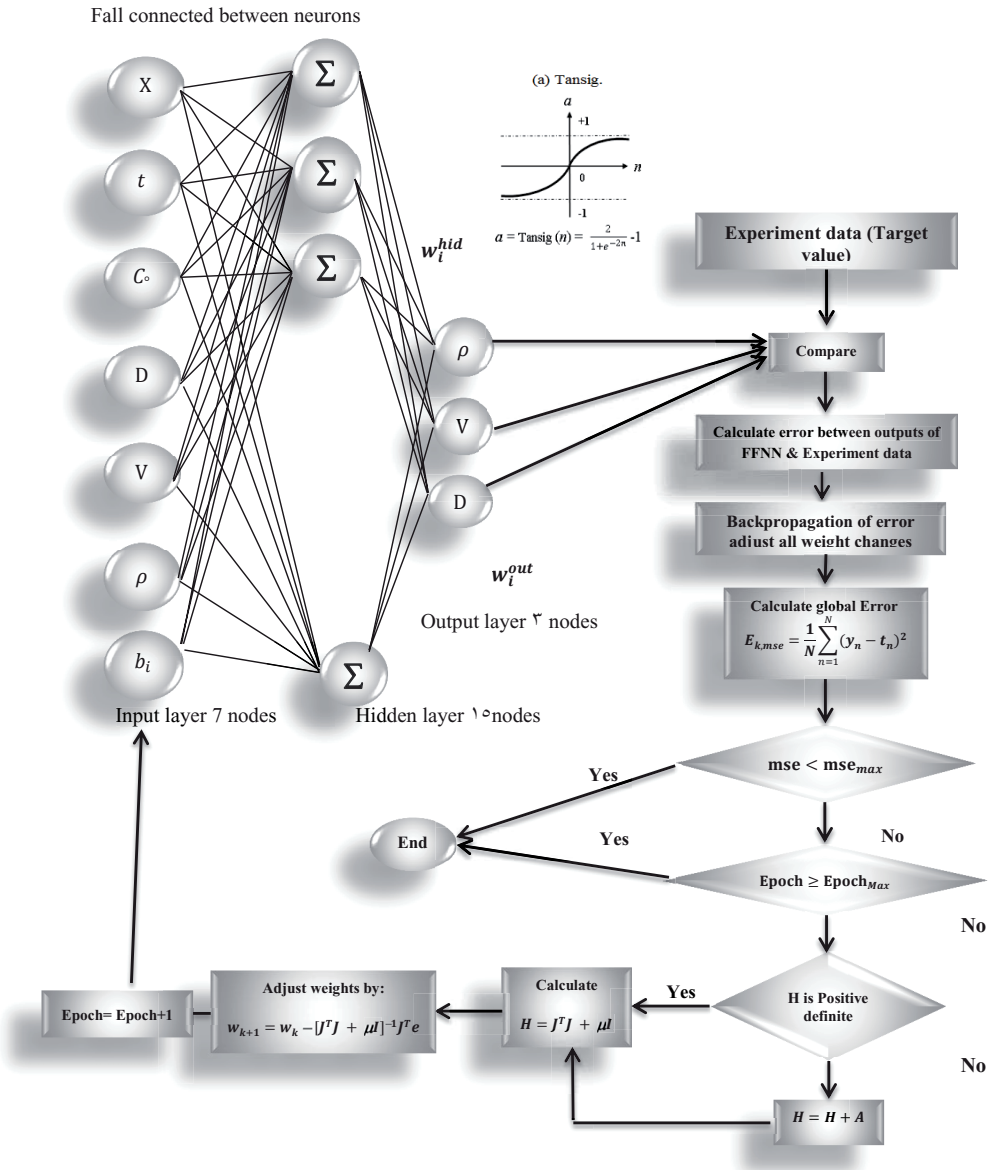

Figure 3.1: Design first FFNN

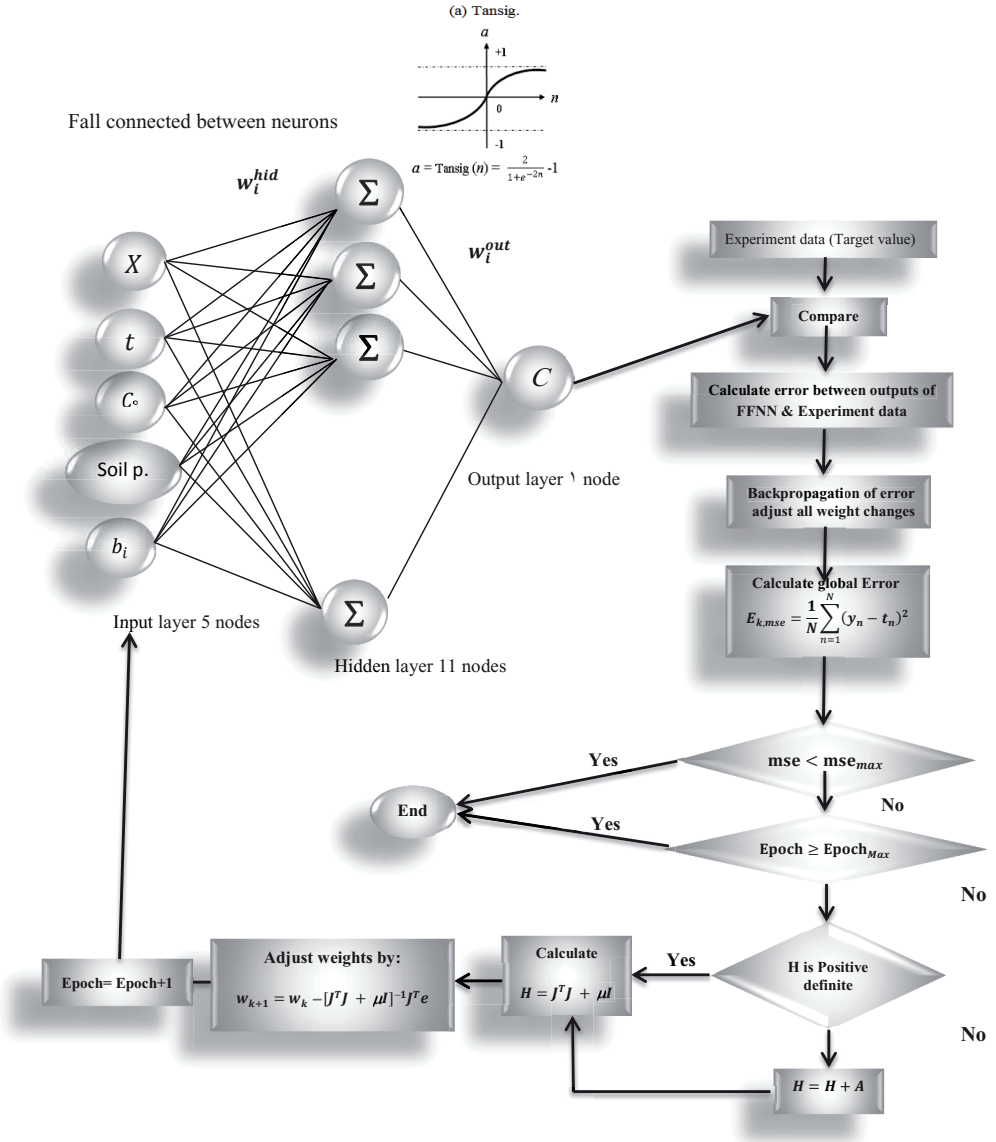

Figure 3.2: Design second FFNN

2.8. When to Stop the Training

Training stops when any of these conditions occurs:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time has been exceeded.
- Performance has been minimized to the goal.
- The performance gradient falls below "min_grad".
- Validation performance has increased more than max_fail time since the last time it decreased (when using validation).

Chapter Four

Applications

Chapter Four

Applications

4.1. Introduction

There are many applications of ANN, that include, estimation of composition of a ternary liquid mixture, mass transfer predictions in a fast fluidized bed of fine solids, modeling for estimation of hydrodynamics of packed, fault diagnosis in complex chemical plants, adsorption study, and similar other applications are reported in [6],[8],[10],[18],[38],[21], [32],[37].

In this chapter, we introduce an application for the suggested design of model equation and ANN which described dilation in previous chapters about the environment contamination specially, soil contamination by heavy metals for different regions in Baghdad city . We aim to estimate the concentration of these heavy metals.

4.2. Sampling

The Capital of Iraq Baghdad City (33°14'-33°25'N, 44°31'-44°17'E), is characterized by arid to semi-arid climate with dry hot summers and cold winters; the mean annual rainfall is about 151.8 mm. For the purpose of collection of soil samples, the study area was divided in to five main types of land use viz. residential, commercial, agricultural, main roads and industrial; and two main source areas, within each land use type viz. roadside and open areas . The sample areas are illustrated by geographic information system (GIS) give in Figure 4.1–4.3.

Soil samples were collected during winter season of 2015. Sixty four soil samples with depth (0 – 40 cm) were carefully collected from each source area in different land using types with a stainless steel spatula. They were air– dried in

the laboratory, homogenized and sieved through a 2mm polyethylene sieve to remove large debris, stones and pebbles, after they were disaggregated with a porcelain pestle and mortar. Then these samples were stored in clean self–sealing plastic bags for further analysis. Metal determinations were done by Atomic Absorption Spectrometry (AAS 6300, Shimadzu, Japan), X– ray fluorescence analysis (XRF) or Inductively Coupled Plasma-Mass Spectrometry (ICP-MS). The laboratory results for many zones in Baghdad are given in Table (4.1 – 4.4).

The data for previous years is obtained from the Ministry of Agriculture which included information about different zones in Baghdad and for different years (1998, 2002, 2009, 2010 and 2012) (see Table 4.5 – 4.6).

Figure 4.1: Illustrate the study areas in Baghdad city- Qanat Aljaeesh by GIS

Figure 4.2: Illustrate the study areas in Baghdad city- Zafaraniyah by GIS

Figure 4.3: Illustrate the different location of study areas in Baghdad city by GIS

Table 4.1: The laboratory results for some zones in Baghdad with depth of 0-20 cm

Zones	Zn	Pd	Ni	Cd
Kasra wa Atash	8.35	_	28.85	0.15
Naseb Alshaheed zone	52.5	15.65	14	0.7
Rashidiya zone	15	6.3	47.75	0.4
Mahmudiyah zone	130	95	38.8	3
Abu Ghraib zone	37.7	12	20.5	-
Zuwarah Park	20	10.4	43.5	0.8
Al-Obeidi zone	135	18.35	25.6	0.05
Shaab zone	30.5	17.5	67.5	0.3

Table 4.2: The laboratory results for both sides of Qanat Al jaeesh in Baghdad

Samples	Depth	Cu	Zn	Pd	Ni	Cd
State 1	0 – 20 cm	10	24	18.8	62.5	0.2
	20 – 40 cm	4.25	9	11	12.5	0.1
State 2	0 – 20 cm	8.5	14.5	11.6	57.5	0.1
	20 – 40 cm	6	14	10	47.5	0.1
State 3	0 – 20 cm	10	16	4.8	27.3	0.1
	20 – 40 cm	9.2	16.1	3.45	25	0.1
State 4	0 – 20 cm	8.6	15	24.05	43.5	0.6
	20 – 40 cm	8.7	10	10	13.75	0.35
State 5	0 – 20 cm	0.2	0.8	2.25	36.5	0.25
	20 – 40 cm	5.5	11.2	8	36	0.45
Average	0 – 20 cm	7.46	14.06	12.3	45.46	0.25
	20 – 40 cm	6.73	12.06	8.49	26.95	0.22

Table 4.3: The laboratory results for Al-Jadriya zone in Baghdad

Samples	Depth	Cu	Zn	Pd	Ni	Cd
State 1	0 – 20 cm	12.5	85	12	55	0.2
	20 – 40 cm	12	32.5	10	37.5	0.2
State 2	0 – 20 cm	8	9.5	10	20.5	0.05
	20 – 40 cm	7	7	8	13	0.05
State 3	0 – 20 cm	8.5	11	11	26	0.05
	20 – 40 cm	10.5	14.5	10	13	0.05
State 4	0 – 20 cm	17.5	15.5	17.5	27	0.05
	20 – 40 cm	12.35	10	10	54	0.6
State 5	0 – 20 cm	9.35	10	13	54	0.75
	20 – 40 cm	12.6	12.5	11.5	55	0.8
Average	0 – 20 cm	11.17	26.2	12.7	36.5	0.22
	20 – 40 cm	10.89	15.3	9.9	34.5	0.34

Table 4.4: The laboratory results for Al- Zafaraniyah zone in Baghdad

Samples	Depth	Cu	Zn	Pd	Ni	Cd
State 1	0 – 20 cm	9.5	25	34.5	52	0.25
	20 – 40 cm	6	10	6.1	55	0.15
State 2	0 – 20 cm	11	18.8	9.7	66	0.7
	20 – 40 cm	12	25	8.8	80	0.2
State 3	0 – 20 cm	4.0	18	7.65	9.5	0.65
	20 – 40 cm	1.00	7	2.75	2.5	0.1
Average	0 – 20 cm	8.167	20.6	17.283	42.5	0.533
	20 – 40 cm	6.333	14	5.8833	45.833	0.15

Table 4.5: Average of the concentrations of heavy metals in Baghdad soil.

HM	١٩٩٨ Depth 10 cm	٢٠٠٢ Depth 10 cm	٢٠٠٩ Depth 15 cm	٢٠١٠ Depth 20 cm	٢٠١٢ Depth 10 cm
Pb	٥٠	١٠٥	67.52	١٥١	٤٣
Zn	١٠٠	١٢٥	77.9	١٣٠	-
Cd	٤	١٠	4.11	٥	١٩
Cu	٣٥	٢٥	-	٩٠	-
Ni	١٧٥	٦٠	-	١١١	١٧٢
Co	٤٨	٢٢	-	٢٨	٣٩

4.3. Simulate Suggested Networks

We suggest a suitable design for two FFNN to solve equation (3.9. The first FFNN is used to find the suitable soil parameters which is needed in equation (3.9). While the second FFNN used to estimate the concentration of heavy metals in soil.

The architecture of first network is three layers FFNN with input layer consists of seven input nodes, one hidden layer consist fifteen hidden nodes with tansig.

transfer function and output layer consist three output nodes with purelin. transfer function. The data in Table 4.1 – 4.6 represent the input data which is divided into three sets: training set, testing set and validation set. When we simulate the suggested FFNN we get the parameters of soil as the following: $D_L = 0.5 \frac{m^2}{d}$ and $V_X = 5.14 \times 10^{-6}$ ms^{-1} $= 44.4096 \times 10^{-2} \frac{m}{d}$, and $\rho = 1.3 \frac{g}{cm^3}$.

Table 4.6: The concentrations of HMs in Baghdad soil for 2012 and 20 cm depth

No.	site	Co	Cd	Ni	Pb
1.	**Shaab**	**41**	**19**	١٤٤	٣٩
2.	**Talbiya**	**37**	**15**	١٦٩	٣٣
3.	**Jamela**	**44**	**18**	١٨١	٤١
4.	**Adhmiya**	**36**	**11**	١٨٣	٣٢
5.	**Al-Salam**	**38**	**18**	١٨٣	٣٨
6.	**Ataifia**	**41**	**13**	١٥٦	٣٤
7.	**Haifa**	**58**	**41**	١٨٠	١٨٣
8.	**Zayuna**	**32**	**17**	١٥٦	٣٦
9.	**Baghdad jadeda**	**48**	**33**	١٣٨	٥٨
10.	**Mashtal**	**36**	**15**	١٣٠	٣٦
11.	**Zufaraniya**	**34**	**13**	٢٠٨	٣٩
12.	**Dora**	**33**	**17**	١٠٥	٣٧
13.	**Karada**	**35**	**16**	١٨٦	٣٩
14.	**Jaderiya**	**37**	**14**	١٩٥	٤٩
15.	**Abu Dsheer**	**51**	**32**	١٦٩	٧٥
16.	**Saidiya**	**38**	**18**	١٦٩	٣٨
17.	**Qadisiya**	**44**	**29**	١٥٦	٥٩
18.	**Mansor**	**38**	**14**	١٨١	٣٦
19.	**Qazaliya**	**39**	**14**	١٢٩	٣٣
20.	**Al jameaa**	**38**	**17**	٢٠٨	٣١
21.	**Khadhraa**	**33**	**15**	١٥٧	٢٩

22.	**Nafaq shurta**	**31**	**16**	١٤٤	٣٧
23.	**Amriya**	**33**	**17**	١٥٦	٢٩
24.	**Jehad**	**40**	**19**	٢٠٨	٣٧
25.	**Furat**	**31**	**17**	١٣٠	٣٣

Figure (4.4) illustrates the training, testing and validation results when we used modified LM algorithm. Figure (4.5) illustrates the performance; epoch and time of the first FFNN. Figure (4.6) illustrates the training, testing and validation results with regularization.

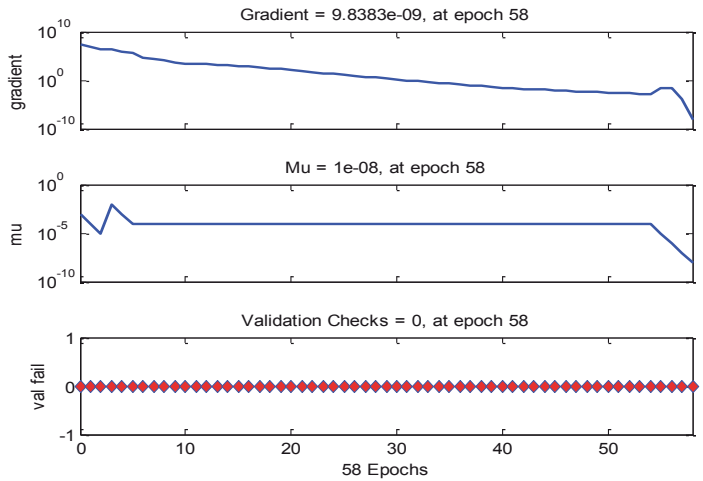

Figure 4.4: Training, testing and validation first FFNN by modify LM

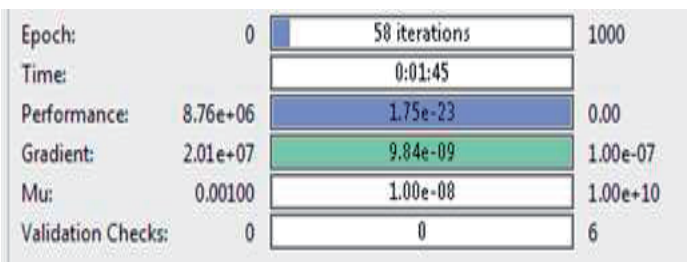

Figure (4.5) illustrate the performance, epoch and time of first FFNN.

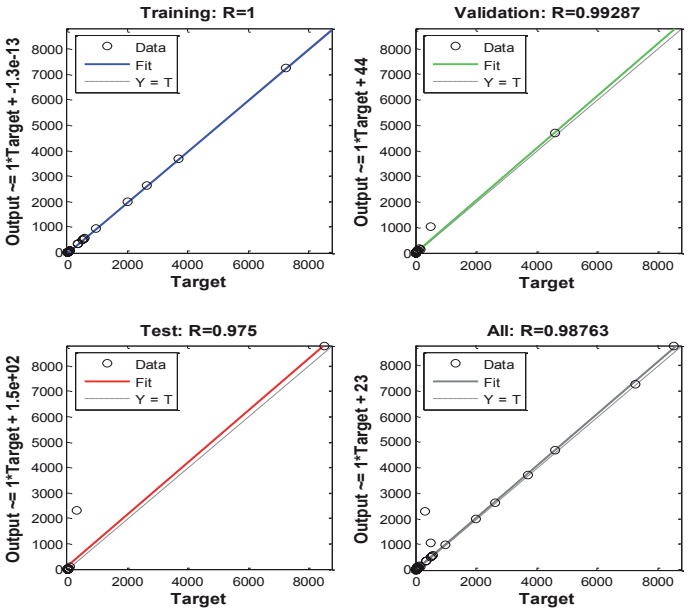

Figure 4.6: Training, testing and validation first FFNN with regularization.

Then, we used the parameters which are obtained from first FFNN in the simulation of the 2nd network, when the second network is of type three layers FFNN and we used the same data in the first FFNN for training, testing and validation. Figure (4.7) illustrates the architecture of the second FFNN. Figure (4.8) illustrates in general, the performance of the 2nd FFNN for training, testing and validation when we used modified LM algorithm. Now, we apply the suggested design to determine the rate of contamination in soil by heavy metals especially by cadmium (Cd), nickel (Ni), copper (Cu), cobalt (Co), zinc (Zn) and lead (Pb) in Baghdad.

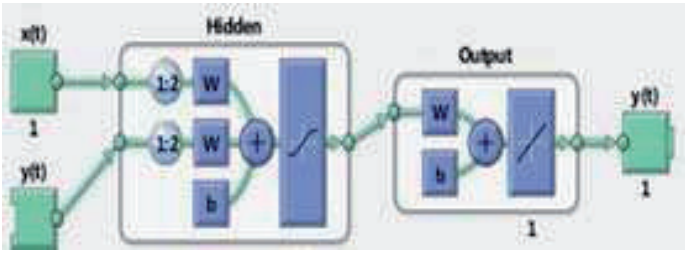

Figure 4.7: Architecture of the second FFNN.

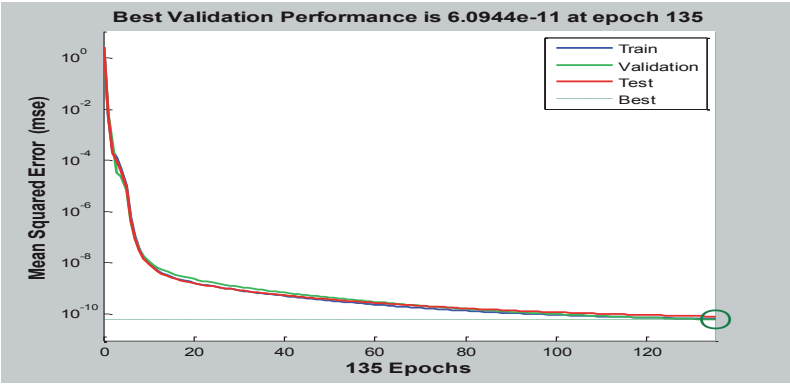

Figure 4.8: Performance of the 2nd FFNN for training, testing and validation

We start with cadmium (Cd), simulate the FFNN with the measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.7) gives the target values for training, testing and validation samples and regularization parameter γ of Cd. Table (4.8) gives the accuracy of the train for epoch and time. Figure (4.9) illustrates the performance of the suggested design for concentration of Cd.

Table 4.7: Target values of Cd.

Type	Target values	MSE	Γ
Training	68	1.73716e -10	9.999999 e-1
Validation	14	2.61731e -10	9.999999 e-1
Testing	14	1.35569e -10	9.999999 e-1

Table 4.8; The accuracy of the train for epoch and time.

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	1.74e-10	396	٠:٠٠:٠4	1.00e-07

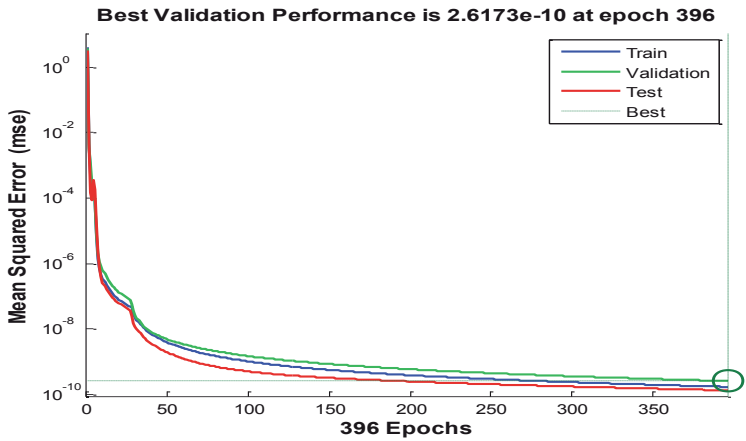

Figure 4.9: Performance of suggested FFNN for concentration of Cd

Now, we simulate suggested FFNN for Zinc (Zn) with the measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.9) gives the target values for training, testing and validation samples and regularization parameter γ of Zn. Table (4.10) gives the accuracy of the train for epoch and time. Figure (4.10) illustrates the performance of suggested design for concentration of Zn.

Table 4.9: The target values for concentration of Zn.

Type	Target values	MSE	γ
Training	68	7. 34209e -11	9.9999999 e-1
Validation	14	1.02701e -9	9.9999999 e-1
Testing	14	1.67711e -10	9.9999999 e-1

Table 4.10: The accuracy of the train for epoch and time.

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	7.34e-11	٥٤٠	٠:٠٠:٠٥	1.00e-08

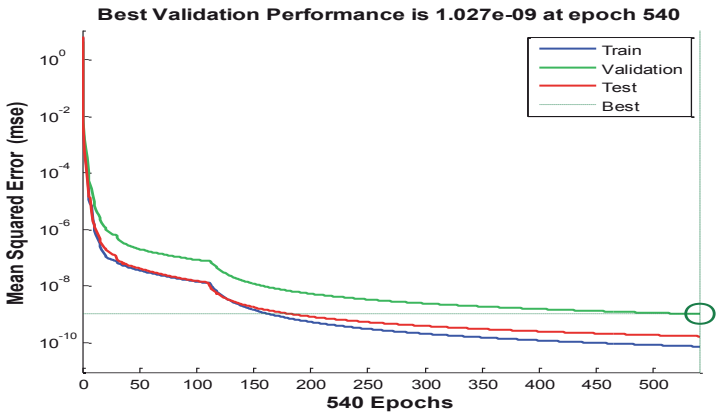

Figure 4.10: Performance of suggested FFNN for concentration of Zn

Now, we simulate suggested FFNN for copper (Cu) with measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.11) gives the target values for training, testing and validation samples and regularization parameter γ of Cu. Table (4.12) gives the accuracy of the train for epoch and time. Figure (4.11) illustrates the performance of suggested design for concentration of Cu.

Table 4.11: Target values for concentration of Cu.

Type	Target values	MSE	γ
Training	68	1.60748e -8	9.9999999 e-1
Validation	14	7.11480e -9	9.9999999 e-1
Testing	14	1.92286e -8	9.9999999 e-1

Table 4.12: The accuracy of the train for epoch and time.

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	1.24e-٠8	٥١	٠:٠٠:٠٦	1.00e-08

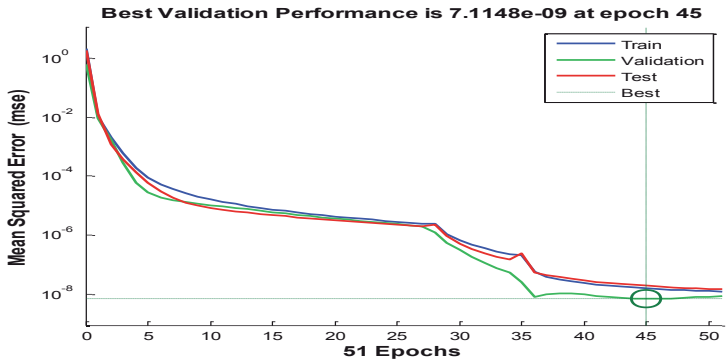

Figure 4.11: Performance of suggested FFNN for concentration of Cu

Now, we simulate suggested FFNN for nickel (Ni) with measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.13) gives the target values for training, testing and validation samples and regularization parameter γ of Ni. Table (4.14) gives the accuracy of the train for epoch and time. Figure (4.12) illustrates the performance of suggested design for concentration of Ni.

Table 4.13: Target values for concentration of Ni.

Type	Target values	MSE	γ
Training	68	1.63755e -10	9.9999999 e-1
Validation	14	2.21215e -10	9.9999999 e-1
Testing	14	2.77821e -10	9.9999999 e-1

Table 4.14: The accuracy of the train for epoch and time.

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	1.٦4e-10	٥٠٠	٠:٠٠:١٣	1.00e-07

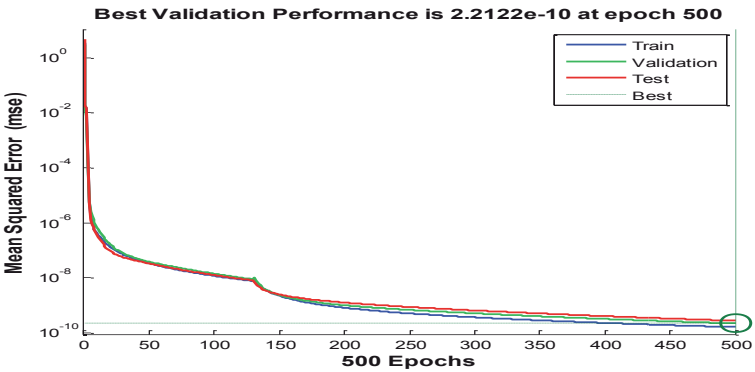

Figure 4.12: Performance of suggested FFNN for concentration of Ni.

Now, we simulate suggested FFNN for cobalt (Co) with measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.15) gives the target values for training, testing and validation samples and regularization parameter γ of Co. Table (4.16) gives the accuracy of the train for epoch and time. Figure (4.13) illustrates the performance of suggested design for concentration of Co.

Table 4.15: Target values for concentration of Co.

Type	Target values	MSE	γ
Training	68	5.13955e -11	9.9999999 e-1
Validation	14	7.59719e -11	9.9999999 e-1
Testing	14	1.13493e -10	9.9999999 e-1

Table 4.16: The accuracy of the train for epoch and time.

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	5.14e-11	852	٠:٠٠:٠٨	1.00e-08

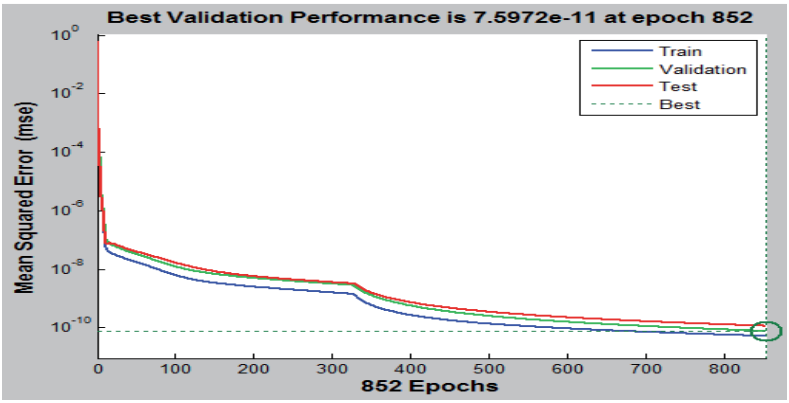

Figure 4.13: Performance of suggested FFNN for concentration of Co.

Finally, we simulate suggested FFNN for lead (Pb) with measured data. The comparison between the predicted concentrations and the measured data resulted in the performance function. Table (4.17) gives the target values for training, testing and validation samples and regularization parameter γ of Pb. Table (4.18) gives the accuracy of the train for epoch and time. Figure (4.14) illustrates the performance of suggested design for concentration of Pb.

Table 4.17: Target values for concentration of Pb.

Type	Target values	MSE	γ
Training	68	5.12985e -11	9.9999999 e-1
Validation	14	7.55379e -12	9.9999999 e-1
Testing	14	1.13893e -10	9.9999999 e-1

Table 4.18: The accuracy of the train for epoch and time

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	5.14e-11	852	٠:٠٠:٠٨	1.00e-08

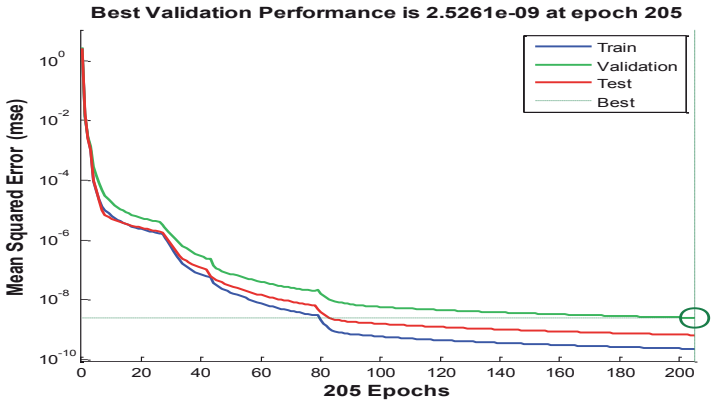

Figure 4.14: Performance of suggested FFNN for concentration of Pb

Now, we applied the suggested design in Qanat Aljaeesh zone in Baghdad that is, we take the data sample from this zone only see Figure (4.1) and the laboratory results gave in Table (4.2) of depth 0 – 40 cm for five state, according to experimental process described in section 4.2, the concentration of heavy metals can be obtained for any depth and time by suggested FFNN. Then comparing these results with that obtained by XRF which are illustrated in Figure (4.15) and Table (4.19) gives the target values for training, testing and validation samples and γ. Table (4.20) gives the accuracy of the train for epoch and time.

Figure (4.16) illustrates the concentrations of heavy metals in soil of Qanat Aljaeesh zone in Baghdad.

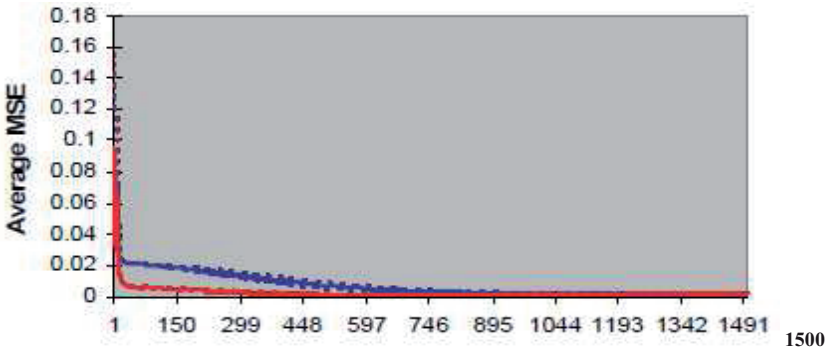

Figure 4.15: A Comparing of the concentration calculated by FFNN &by XRF

Table 4.19: Target values for concentration of HM.

Type	Target values	MSE	γ
Training	68	2.31128e -11	9.9999999 e-1
Validation	16	2.52614e -12	9.9999999 e-1
Testing	16	6.69947e -12	9.9999999 e-1

Table 4. 20: The accuracy of the train for epoch and time

Train Function	Performance of train	Epoch	Time	μ
Modify Trainlm	2.31e-11	252	٠:٠٠:٠6	1.00e-08

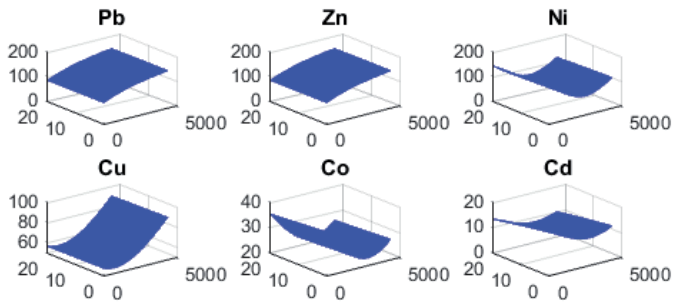

Figure 4.16: Concentrations of HMs in soil of Qanat Aljaeesh zone

4.4. Discussion

The practical results show the following:

- The average of the concentrations of HMs in soil for different time and zones in Baghdad are increase with time, posing a great risk to the environment contamination.
- For the comparison among the concentrations of different regions: residential, industrial, commercial and agricultural regions, we see that:

soils agricultural < soils residential < soils commercial < soils industrial that is, the agricultural regions are the lowest while the 76industrial regions are the highest for the concentrations of HMs.

- For comparing upon the depth of soil, we see the effect of depth on the concentration of HMs that is when the depth increase the concentrations of HMs are decrease, i.e., the concentration of HMs in depth 0 – 20 cm is larger than the concentration in depth 20 – 40 cm for the same soil.
- There are different causes for increasing the concentrations of HMs in soil such as: the big traffic jams resulting from the great number of cars lately which use gasoline that contains a lot of fourth lead Ethylene which cause big problems to the environment .This creates dangers to human beings. In addition, the increase in the amount of litter and how to get rid of industry waste in sewerage and the decrease in the green region which participate in lessening the damage of HMs on the environment. As a result of increase in the population during the late years which results in converting the regions of vegetation to residential regions and the technological development which causes contamination because of the prolife ration of plants and workshops scattered everywhere. Add to all this, wars and their great contamination which are considered the most dangerous contaminates of the soil and environment. All these types of contaminates cause high rate of concentration

of HMs which exceeds the normal amount in soil, the increase of these metals has different types of danger on human health. The plants absorb these dangerous materials which in its turn go to human being through food consumption which they acquire because of eating these plants that have the dangerous metals.

4.5. Suggestion Treatment of the Soil

There are many techniques for treatment the soil from the contamination by HMs which are: volatilization, biodegradation, leaching, isolation / containment and phytoremediation.

We suggest the treatment by phytoremediation because this mothed is the best to treatment the soil in sped of taken a large time to treatment, it is easy implemented and it was not costly.

Phytoremediation is one of the important processes for treatment the contaminated soil by using some of plants that the ability high-absorption to accumulation the concentration of metals in that plants to treatment contaminated land. Plant can be translocate the metals into aboveground by accumulation this metals in the roots then to other parts of plant. After plants have grown for some time, they are harvested and incinerated or composted to recycle the metals. Several crop growth cycles may be needed to decrease contaminant levels to allowable limits. If the plants are incinerated, the ash must be disposed of in a hazardous waste landfill, but amount of the ash is much smaller than the amount of contaminated soil if dug out and removed for cleanup [13,24].

The phenomenon of accumulation of HMs in the accumulation of plants has received a significant attention of researchers because of its important applications in vegetable processors phytoremediation can exploit these plants

and used to extract contaminants (HMs) from the soil, if it is the use of plants valuable absorption of these metals from the soil solution and the transition to the total vegetative and also transfer technology to the volatile materials Phytovolatilization where exploiting this technology in the ability of some plants on the introduction of some HMs in rechargeable volatility compounds for disposal. Studies to indicate that there are many plants that can capture and accumulate HMs from contaminated regions. The ideal plant for this process must be of certain advantages such as fast growth, roots dense and deep large and easy harvesting and cutting and accumulation of a wide range of metals and live mass, are also available unsustainable high levels of those metals (see Figure (4.17)) [4].

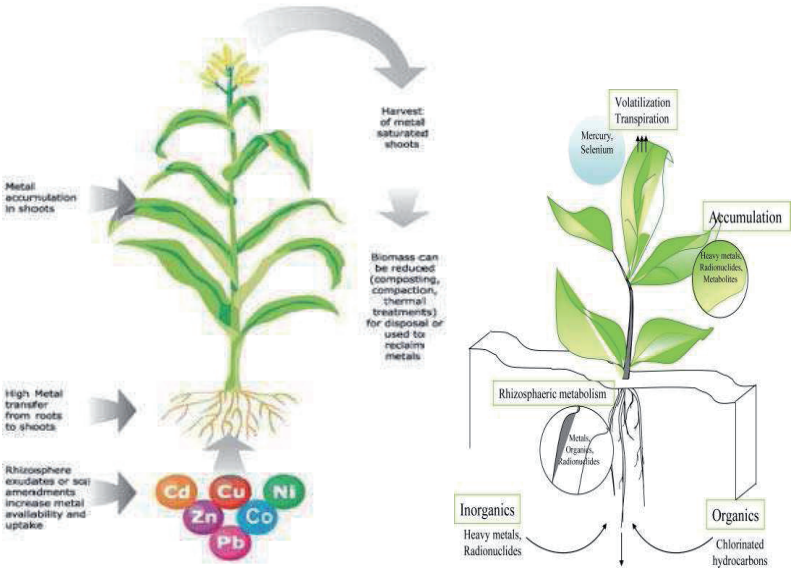

Figure 4.17: Phytoremediation processes

There are many plants with high capacity to absorb dangerous substances released by the label plants hyper accumulation. Now, we introduce some of these plants:

- **Brassica Oleracea L.**

Where composed of high economic importance such as cabbage, cauliflower, broccoli and green varieties. It used to treat inorganic contaminant in soil such as: Cu, Pb, Ni and Zn, see Figure (4.18).

Figure 4.18: Brassica Oleracea L.

- **Brassica Nigra**

Herbaceous plants grow up to a height of 60-80 cm. Grain sharp taste lobed leaves. Leg many branching List bluish-green. Clustered flowers in bunches of a shiny yellow color. Fruit with whit long beak. Black seeds or small-scale structure and used to treat inorganic contaminants in soil such as: Cr and Cd, see Figure (4.19).

Figure 4.19: Brassica Nigra

- **Sunflower (Helianthus annulus L)**

Featuring sunflower large radial flowers, which revolve with the sun wherever it took place; therefore called so, and cultured sofa accessories, and eaten the seeds of this plant have the ability to absorb HMs and radioactive

materials, and its concentration in the leaves, so used in treat the soil from HMs such: Cu, Ni, Pb and Zn, Figure (4.20).

Figure 4.20: Sunflower (Helianthus annulus L)

- **Sorghum Bicolor**

  It can absorb some HMs such as : Cu, Ni, Pb and Zn, see Figure (4.21).

Figure 4.21: Sorghum Bicolor

- **Zea Mays L.**

It can absorb some HMs such as : Pb , Co and other contaminants, see Figure (4.22).

Figure 4.22: Zea Mays L.

- **Tritium Aestivum L.**

The characterize of this plants is rows of bright green, and look like grass, and plant height ranges from 0.6-1.5 m. They turn into brown color slant to yellowing when mature. It's used to treat the soil from Pb, Co and other contaminants, Figure (4.23).

Figure 4.23: Tritium Aestivum L.

Chapter Five

Conclusions & Future works

Chapter Five

Conclusions & Future works

5.1. Conclusions

This work demonstrates the development of a model equation that describes the spread of contamination through soils which can be used to determine the rate of environmental contamination by estimates the concentration of heavy metals in soil. Then solving this problem by FFNN's. The results provide many of the features, such as:

1- The learning problem in the multilayer FFNN has been stated from the point of view of functional analysis. The solving approach here consists of in three steps: The first step is to choose a suitable parameterized function space in which the solution to the problem is to be approximated, then spanned by a multilayer FFNN. In the second step the contamination problem is formulated by selecting an appropriate performance function. The third step is to solve the problem by designing a suitable FFNN with a training algorithm which depends on optimization problem capable of finding an optimal set of the parameters.

2- The LM training algorithm described was developed to reduce the memory requirements and computational overhead of training algorithm.

3-The training algorithm and number of iterations (epoch) determines how good the error on the training set is minimized.

4-The number of training samples determines how good training samples represent the actual function.

5- The model developed can be considered to be a good representation of that estimate the concentrations of heavy metals in the soil.

6- The concentration of heavy metals can be estimated by FFNN. The sixty four soil samples from the different sections of Baghdad city were estimated by this technique. The prediction errors of this technique are less than 1% compared with those of AAS or XRF. This technique is fast, convenient, sensitive, and can eliminate the interference among various species. The determination of sample by AAS is finished in seven days; while the determination of sample by the proposed technique is finished in one day.

5.2. Future Works

The studies for future investigations are suggested below:

1- Develop a model equation that describes other types of environment contamination problems such as air contamination, water contamination and food contamination.
2- Use or design other types of ANN or use other architectural for ANN.
3- The impact of wastewater and irrigation water contaminated with HMs on soil contamination and the transfer of these metals into groundwater can be studied by the same manner.

References

References

[1] Abdul Samath, J., Kumar, P.S. and Begum, A., (2010), "Solution of Linear Electrical Circuit Problem using Neural Networks, International Journal of Computer Applications, 2 (1), pp: 6-13.

[2] Abu-Saleman, M.S., Hassnin, S.A. and Saeed Kandel, N.F., 2007, Integrating the concepts of culture population and environmental and food security in the programs and activities of agricultural extension services environment and sustainable agricultural development project, The Ministry of Agriculture and Land Reclamation Egyptian territory Agricultural Research Center of the Central Administration for Agricultural Extension, NO.1080.

[3] Alloway, B. J., (1998), "The mobilization of trace elements in soil", in prost, R.(ed.), contaminated soils, Les colloques, No.85, INRA, Paris, p. 133-145.

[4] Al-Whaibi, M. H., (2007), "Accumulation phenomenon of heavy metals in plants", Saudi Journal. Biol. Sc., Vol. 14, No. 2.

[5] Amacher, M.C., Kotuby-Amacher J., Selim H.M. and Iskandar I.K., (1986), Retention and release of metals by soils--evaluation of several models. Geoderma, 38: 131-154.

[6] Anagu, I., Ingwersen, J., Utermann, J.and Streck, T., (2009), "Estimation of heavy metal sorption in German soils using artificial neural networks. Geoderma 152, 104–112.

[7] Ayeni, O. O., Ndakidemi, P. A., Snyman, R. G. and Odendaal, J. P., (2010), Chemical, biological and physiological indicators of metal pollution in wetlands. Scientific Research and Essays, 5 (15), pp.1938-1949.

[8] Behrens, T., Förster, H., Scholten, T., Steinrüken, U., Spies, E. and Goldschmitt, M., (2005), "Digital soil mapping using artificial neural networks", Journal of plant nutrition and soil science, Vol. 168, pp: 21-33.

[9] Bokil, V. A., "Introduction to Mathematical Modeling", spring (2009).

[10] Buszewski, B. and Kowalkowski, T., (2006), "A new model of heavy metal transport in the soil using non-linear artificial neural networks", Journal of environmental engineering science 23 (4), 589-595.

[11] Chahar, V., (2013), "Analysis of Back Propagation Algorithm", International Conference on Service Systems and Service Management (ICSSSM), 2 (8),ISBN:978-1-4673-4434-0

[12] Demuth, H. and Beale , M., (1998), "Neural Network Toolbox For Use with MATLAB", Version 3, The Math Works. Inc.

[13] Donahue, S. and Auburn, September(2000), SOIL QUALITY – Heavy Metal Soil Contamination,(USDA) United State Department of Agriculture, Urban Technical Note No. 3, AL 36832 ,334-844-4741 X-177.

[14] Dorofki, M., Elshafie, A. H., Jaafar, O., Karim, O. A. and Mastura, Sh., (2012), "Comparison of Artificial Neural Network Transfer Functions Abilities to Simulate Extreme Runoff Data", International Conference on Environment, Energy and Biotechnology, IACSIT Press, IPCBEE, Singapore, 33, pp: 39- 44.

[15] El Badaoui, H., Abdallaoui, A., Manssouri, I. and Lancelot, L., (2013), "Application of Artificial Neural Networks of MLP type for the prediction of the levels of Heavy Metals in Moroccan Aquatic Sediments", 3 (6), pp:75-81.

[16] Eqhaar, Q. H., (2007), "Comparison Between Artificial Neural Networks with Radial Basis Function and Ridge Basis Function for Interpolation Problems", MSc. Thesis, University of Baghdad, College of Education – Ibn Al-Haitham.

[17] Galushkin, I. A., (2007), "Neural Networks Theory", Berlin Heidelberg.

[18] Gandhimathi, A. and Meenambal, T., (2012), "Analysis of Heavy Metal for Soil in Coimbatore by using ANN Model", European Journal of Scientific Research 68, (4), 462-474.

[19] Griva, I., Nash, S. G. and Sofer, A., (2009), "Linear and Nonlinear Optimization", 2nd edition, USA ,ISBN 978-0-898716-61-0.

[20] Hadithi, F. H.A. and Abdul-Jabbar, B. "Environment Pollution", (2013), Baghdad university.

[21] Hambli, R., Chamekh, A. and Bel Hadj Salah, H., (2006), "Real-time deformation of structure using finite element and neural networks in virtual reality applications", Finite Elements in Analysis and Design 42, (11), 985-991.

[22] Hristev, R. M., (1998), "The ANN Book", Edition 1, GNU Public License, ver. 2.

[23] Jabber, A. K., (2009), "On Training Feed Forward Neural Networks for Approximation Problem", MSc. Thesis, University of Baghdad, College of Education – Ibn Al-Haitham.

[24] Jankaitė, A., (2009), "Soil remediation from heavy metals using mathematical modelling" , Journal of Environmental Engineering and Landscape Management, 17(2): 121–129.

[25] Kardam, A., Raj, K. R., Arora, J. K., Srivastava, M. M. and Srivastava, S., (2010), "Artificial Neural Network Modeling for Sorption of Cadmium from Aqueous system by Shelled Moringa Oleifera Seed (SMOS)powder as an Agricultural waste", Journal of Water Resource and Protection, 2 (4), 339-344.

[26] Karlik, B. and Olgac, A. V., (2011), " Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks", International Journal of Artificial Intelligence And Expert Systems (IJAE), 1 (4), pp: 111-122.

[27] Krenker, A., Bešter, J. and Kos, A., (2011), "eBook of : Introduction to the Artificial Neural Networks, Artificial Neural Networks – Methodological Advances and Biomedical Applications", Prof. Kenji Suzuki (Ed.), ISBN: 978-953-307-243- 2, InTech.

[28] Krishnaa, D. and Sree, R. P., (2014), " Artificial Neural Network (ANN) Approach for Modeling Chromium (VI) Adsorption From Aqueous Solution Using a Borasus Flabellifer Coir Powder", International Journal of Applied Science and Engineering, 12 (3), 177-192.

[29] Kukkola, E., Rautio, P. and Huttunen, S., (2000), Stress indications in Copper and Nickel exposed Scots pine seedlings, Environ. Exp. Bot., Vol. 43, pp: 197-210.

[30] Madhloom, H. M., (2015), "Modeling of Copper removal from simulated wastewater by adsorption on to fungal biomass using artificial neural network", Global Journal on Advances in Pure & Applied Sciences, 05, pp: 35-44.

[31] Mahmood, S. H. and Jalil, T. S. ,(2012),"Extended Kalman Filter of Training Feed Forward Artificial Neural Network", journal of kirkuk University For Administrative and Economic Sciences, 2 (1), 22222995, Pages: 183-201.

[32] Minasny, B., Hopmans, J.W., Harter, T., Eching, S.O., Tuli, A. and Denton, M.A., (2004), "Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data", Soil Science Society of America Journal 68, 417- 429.

[33] Nouri, J., (1980), "Heavy metals in sewage sludge, soils amended with sludge and their uptake by crop plants", Ph.D. Thesis, University of London,.

[34] Pandharipande, S. L., Deshmukh, A. R. and Kalnake, R., Mar(2013),"Artificial Neural Network Modeling for Estimation of concentration of Ni(ll) and Cr(Vl) present in Aqueous Solution", International Journal of Advances in Engineering & Technology, 5 (2), pp: 122-131, 2231-1963.

[35] Pradeep, T., Srinivasu, P., Avadhani, P. S. and Murthy, Y. V. S., (2011), "Comparison of variable learning rate and Levenberg-Marquardt back-propagation training algorithms for detecting attacks in Intrusion Detection Systems", International Journal on Computer Science and Engineering (IJCSE), Vol. 3, No. 11.

[36] Quarteroni, A., (2009), "Mathematical Models in Science and Engineering", Notices of the AMS, 56(1),.

[37] Sarmadian, F. and Taghizadeh Mehrjardi, R., (2008)," Modeling of Some Soil Properties Using Artificial Neural Network and Multivariate Regression in Gorgan Province, North of Iran", Global Journal of Environmental Research 2 (1), 30-35.

[38] Schug, B., Düring, R.A. and Gäth, S., (2000), "Improved cadmium sorption isotherms by the determination of initial contents using the radioisotope 109 Cd", Journal of plant nutrition and soil science, Vol. 163, 197–202.

[39] Sekhar, P. C., Sanker Ram, B. V. and Sarma, K. S., (2010), "Fast Computing Neural Network Modeling for Fault Diagnosis in Power Systems", ARPN Journal of Engineering and Applied Sciences, 5 (9), 1819-6608.

[40] Selim, H. M., Amacher, M. C., and Iskandar, L. K., (1990), Modeling the Transport of Heavy Metals in Soils, Monograph 90-2,U.S. army corps of engineering.

[41] Shimi, H. M., (2001),"Management and maintenance of land and water in the desert crops", Dare Lfiker El-Arabi, Frist Edition, I.S.B.N.977-10-1470-6.

[42] Tawfiq, L . N. M. and Eqhaar, Q. H., (2009), "On Multilayer Neural Networks And Its Application For Approximation Problem, 3rd scientific conference of the College of Science ", University of Baghdad. 24 to 26 March.

[43] Tawfiq, L. N. M., (2007), "On Training of Feed Forward Neural Networks", Journal of Um-Salama for Scince, Vol. 4, No.1.

[44] Tawfiq, L. N. M., (2013), "Improving Gradient Descent method For Training Feed Forward Neural Networks", International Journal of Modern Computer Science & Engineering, Vol. 2, No. 1, pp: 1-25.

[45] Wang, X. G., Tang, Z., Tamura, H., Ishii, M., and Sun, W. D., (2004), "An improved back propagation algorithm to avoid the local minima problem", Neurocomputing, Vol. 56, pp: 455-460.

[46] Wuana, R. A. and Okieimen, F. E., (2011), Heavy metals in contaminated soils: A review of sources, chemistry, risks and best available strategies for remediation, ISRN Ecology, l. 2011, Article ID 402647, 20 pages.

[47] Yetilmezsoy, K. and Demirel, S., (2008), "Artificial Neural Network approach for modeling of Pb adsorption from aqueous solution by Antep pistachio (pistacia vera L.) shells", Journal of Hazardous Materials, 153 (3), 1288-1300.

[48] Yin Li, Chao-kui Li ,Jian-junTao and Li-dong Wang, (2011), "Study on Spatial Distribution of Soil Heavy metals in Huizhou city Based on BP- Artificial Neural Network Modeling and GIS", Procedia Environmental Sciences Journal, doi:10.1016/j.proenv.2011.09.306, 1953 – 1960.

[49] Zongshu, Wu, Jiaoyan, Ai, Chaobing Deng, Yajuan Cai and Zongming Wei, (2015), " Method for Optimal Arrangement of Soil Sampling Based on Neural Networks and Genetic Algorithms", International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE), 124, 1951-6851.

[50] Tawfiq, L. N. M. and Ghazi, F. F., (2015), "Using Artificial Neural Network Technique For The Estimation Of Cd Concentration In contaminated Soils", International Journal of Innovations in Scientific Engineering, Volume 1, No. 1, pp: 58-68.

[51] Tawfiq, L. N. M. and Ghazi, F. F., (2015)," Heavy metals pollution in soil and its influence in Iraq", Merit Research Journal of Agricultural Science and Soil Sciences, Volume 3, No. 5, pp: 82 - 87.

[52] Tawfiq, L. N. M. and Ghazi, F. F., (2015), "Contaminated Soils by Heavy Metals in South of Iraq", International Journal of Environment and Bioenergy, Volume 10, No. 3, pp: 41-47.

[53] Tawfiq, L. N. M., Kareem A. Jasim and Abdul hmeed, E. O., (2015), Pollution of Soils by Heavy Metals in East Baghdad in Iraq, International Journal of Innovative Science, Engineering & Technology, Vol. 2, Issue 6, pp: 181- 187.

[54] Tawfiq, L. N. M., Kareem A. Jasim and Abdul hmeed, E. O., (2015), Mathematical Model for Estimation the Concentration of Heavy Metals in Soil for Any Depth and Time and its Application in Iraq, International Journal of Advanced Scientific and Technical Research, Vol. 4, Issue 5, pp:718 – 726.

[55] Tawfiq, L. N. M., Kareem A. Jasim and Abdul hmeed, E. O., (2016), "Numerical Model for Estimation the Concentration of Heavy Metals in Soil and Its Application in Iraq", Global Journal Of Engineering Science And Researches, Vol. 3, Issue 3, pp: 75 – 81.