# Daily Suspended Sediment Discharge Prediction Using Multiple Linear Regression and Artificial Neural Network

To cite this article: Uca *et al* 2018 *J. Phys.: Conf. Ser.* **954** 012030

View the article online for updates and enhancements.

# Daily Suspended Sediment Discharge Prediction Using Multiple Linear Regression and Artificial Neural Network

**Uca[1*], Ekhwan Toriman[2], Othman Jaafar[3], Rosmini Maru[1], Amal Arfan[1], Ansari Saleh Ahmar[4,5]**

[1]Department of Geography, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Indonesia
[2]Faculty Sciences Social and Humanities, Universiti Kebangsaan Malaysia, Malaysia
[3]Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Malaysia
[4]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Indonesia
[5]AHMAR Institute, Makassar, Indonesia

*ucasideng@unm.ac.id

**Abstract**. Prediction of suspended sediment discharge in a catchments area is very important because it can be used to evaluation the erosion hazard, management of its water resources, water quality, hydrology project management (dams, reservoirs, and irrigation) and to determine the extent of the damage that occurred in the catchments. Multiple Linear Regression analysis and artificial neural network can be used to predict the amount of daily suspended sediment discharge. Regression analysis using the least square method, whereas artificial neural networks using Radial Basis Function (RBF) and feedforward multilayer perceptron with three learning algorithms namely Levenberg-Marquardt (LM), Scaled Conjugate Descent (SCD) and Broyden-Fletcher-Goldfarb-Shanno Quasi-Newton (BFGS). The number neuron of hidden layer is three to sixteen, while in output layer only one neuron because only one output target. The mean absolute error (MAE), root mean square error (RMSE), coefficient of determination ($R^2$) and coefficient of efficiency (CE) of the multiple linear regression (MLRg) value Model 2 (6 input variable independent) has the lowest the value of MAE and RMSE (0.0000002 and 13.6039) and highest $R^2$ and CE (0.9971 and 0.9971). When compared between LM, SCG and RBF, the BFGS model structure 3-7-1 is the better and more accurate to prediction suspended sediment discharge in Jenderam catchment. The performance value in testing process, MAE and RMSE (13.5769 and 17.9011) is smallest, meanwhile $R^2$ and CE (0.9999 and 0.9998) is the highest if it compared with the another BFGS Quasi-Newton model (6-3-1, 9-10-1 and 12-12-1). Based on the performance statistics value, MLRg, LM, SCG, BFGS and RBF suitable and accurately for prediction by modeling the non-linear complex behavior of suspended sediment responses to rainfall, water depth and discharge. The comparison between artificial neural network (ANN) and MLRg, the MLRg Model 2 accurately for to prediction suspended sediment discharge (kg/day) in Jenderan catchment area.

Keywords: Suspended sediment discharge, multiple linear regression, artificial neural network, Jenderam catchment

## 1. Introduction

The every time the flowing of water river always of bring a variety of material. The quantity of material flowing by Water River it influenced by the season. Concentration of suspended sediment load high during the rainy season, while the dissolved load high of concentration in dry season, because the water in a river source from soil water that contain a high concentration of chemical elements [1]. The flowing of water river brings a variety of materials such as suspended sediment load, bed load and dissolved load [2][3]. This material sourced from by the erosion process which occurs in the subsurface and or surface of soil. Erosion that occurs on the surface soil caused by transformation rainfall water became to be surface runoff. Suspended sediment concentration from different time it varies depending on the season condition [4]. The situation is strongly influenced by climatic events, such as rainfall events [5]. Furthermore, the occurrence of rainfall is influenced by other climate factors such as temperature and air humidity [6]-[8]. Climate change in a region is influenced by the conversion of green areas into constructed areas, such as protected forests into production and settlement forest areas [8].

Predict the amount of daily suspended sediment discharge which occurs in the catchments area is very important as an indicator to assess the level of erosion hazard, management of the water resources, water quality, hydrology project management (dams, reservoirs, and irrigation) and to determine the level of damage in the catchment area [9].

The models predict the amount of sediment in a catchment area has a lot of developed such as regression methods, InfoWork Rs, stormwater & wastewater management model (XPSWMM), watershed modeling systems (WMS) and artificial neural network (ANN) [10]. Prediction of sediment suspension using ANN has been widely carry out in the catchments area with the exact result [11][12]. In a study conducted in the Jenderam catchment using multiple linear regression (MLRg) and artificial neural network for to predict of suspended sediment discharge. The multilayer perceptron feed forward training algorithm is used, with three different learning algorithms to predict the amount of suspended sediment discharge is: Levenberg-Marquardt (LM), Scaled Conjugate Gradient ( SCD ) and Broyden-Fletcher-Goldfarb-Shanno Quasi-Newton (BFGS). Although many researchers have previously been using ANN [11][12], but only use one to two training algorithms. Therefore in this research using three training algorithms . The purpose of this research is to compare the accuracy of the regression analysis and ANN models in the small catchments area .

## 2. Multiple Linear Regression

Multiple linear regression (MLRg) analysis is one of the most widely used of all statistics [11]. A regression model that involves more than one regressos variable is called a multiple linear regression model [14]. Multiple linear regression modelling has been widely used for modeling such as urban runoff pollutant load [15], wash load sediment concentrations [16], suspended-sediment discharge [17], prediction of swell potential of clayey soils [18]. The general form of regression model for k independent variables with have two or more regressor variables is given by [19]:

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_x + E \qquad (1)$$

where $\beta_0, \beta_1, \beta_2, ...., \beta_k$ are the regression coefficients respectively, that need to be estimated, while $E$, $Y$ and $X_1, X_2, ..., X_k$ are the error, dependent and independent variable. The value of the independent variable ($X$) is always related with a value of the dependent variable ($Y$). The gooddnes performance of multiple linear regression model can be expressed by the value of the error (differences between observed and predicted values), coefficient correlation ($r$) and coefficient of determinations ($R^2$). In general determining the best estimate of the multiple regression equation using the least-squares method chooses as the best-fitting model the one that minimizes the sum of squares of the difference between the observed and predicted by the fitted model [14][19].

## 3. Artificial Neural Network

An artificial neural network is an information processing systems that has certain performance characteristics in common with biological neural network such as the human [20]. Artificial neural network has been developed as generalization of mathematical model of human cognition or neural biology, based on the assumptions that: Information processing occurs at many simple elements known as neurons, signals sent between neurons is through connection link, each connection link has an associated weight, which in a typical neural net, multiplies the signal transmitted, each neuron applies an activation function (usually non-linear) to its net input (sum of weighted input signals) to determine its output signal [21].

A artificial neural network is characterized by its pattern of connections between the neurons (called its architecture ), the method of determining the weights on the connections (called its training, or learning, algorithm), and its activation function. The multilayer neural network architecture consists of input layer, hidden layer and output layer. Input layer consists of nodes or neurons that will receive the data. Hidden layer consists of nodes or neurons that receive input from the input layer. The output layer consists of neurons receiving data from the output of the hidden layer [22].

### 3.1. Multilayer perceptron feedforward

Multilayer perceptrons feedforward (MLP) is a systematic method for training multilayer neural networks. This method has a good mathematical basics, objectives and get the shape formula algorithm and the coefficient in the equation by minimizing the sum of squares error value through models developed by the net. The input layer, hidden layer and output layer with the interconnection weights value $v$ and $w$ between layers of neurons [23]. Backpropagation artificial neural network architecture is given in Figure 1.
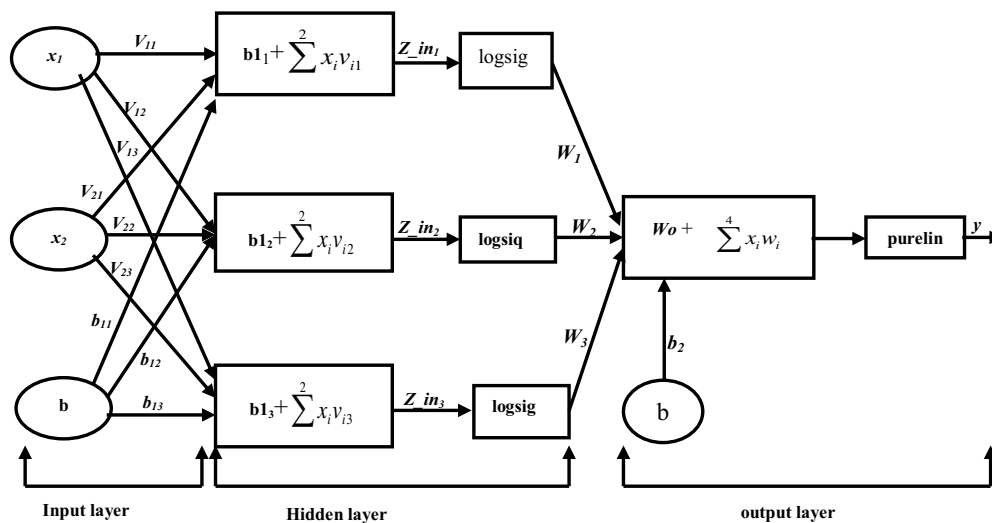


**Figure 1.** Schematic diagram of an MLP feed forward network

### 3.2. Selection of Training Algorithms

Some training in multilayer perceptrons feedforward algorithm in artificial neural networks that are often used: gradient descent, gradient descent with momentum, resilient backpropagation, powell-beale restarts, scaled conjugate gradient, BFGS quasi-newton, and levenberg-marquardt. Each training algorithms has different characteristics. In this research used training algorithms: gradient descent (GD), lavenberg-marquardt (LM), scaled conjugate descent (SCD) because not many researchers who

compare the three algorithms for the training of predicting the amount of daily suspended sediment discharge in a small basin in the tropical area. Results of the analysis of the training algorithm will be used for compare the performance of each MLP training algorithm, RBF and MLRg.

*3.2.1. Lavenberg-Marquardt.* In mathematical levenberg-marquardt (LM) algorithm, also called as damped least-squares method (DLS), as part of troubleshooting to minimize numerical problems are generally non-linear functions. Algoritmh levenberg-marquardt (LM) is they gauss newton interpolation algorithm (GNA) and gradient descent (GD) methods. LM method more robust than the GNA found means in many cases can solve this problem if started very far from the minimum end. LM can also describe, as a Gauss Newton using the trust region approach. LM is very popularly used in curve fitting algorithm to solve the problem of generic curve fitting. But only found in local minimum is not at the global minimum. Levenberg-marquardt algorithm is designed by using the second derivative approach without having to compute the hessian matrix. LM training algorithm to update the weights using the following equation:

$$w_{t+1} = w_t - \Delta w_t \tag{2}$$

$$\Delta w_t = \left[ J^T(w_t) J(w_t) + \mu I \right]^{-1} J^T(w_t) e(w_t) \tag{3}$$

where:

$w_{t+1}$     =   the update weight vector.

$w_t$     =   the weight vector before updating.

$I$     =   represents an identity matrix

$J$     =   the jacobian matrix that contains first derivatives from network error toward the weights value.

$\mu$     =   the learning rate

$e(w_t)$     =   the matrix error function evaluated at previous iteration.

When learning rate value 0, then this approach will be the same as Newton's method, while learning rate is large then this same approach with gradient descent. Newton's method is very fast and accurate for minimum error as the algorithm is expected to quickly change the learning rate to be equal to 0. For that, after a few iterations, the algorithm will reduce the learning rate. Increase the rate of learning will be done when it takes a step (temporarily) to reduce function performance.

*3.2.2. BFGS Quasi Newton.* BFGS Quasi-Newton algorithm is one of the conjugate gradient alternatives that can be used to get optimal value faster. Newton's method can work fast to get the optimum value completion, this method requires a large computer memory for each iteration calculating the second derivative. To overcome these obstacles are made improvements with quasi-newton method. In quasi-newton methods the hessian matrix does not need to be computed. The hessian is updated by analyzing successive gradient vectors instead. BFGS Quasi-Newton methods are a generalization of the secant method to find the root of the first derivative of multidimensional problems. In multi-dimensional the secant equation is under-determined, and quasi-newton methods differ in how they constrain the solution, typically by adding a simple low-rank update to the current estimate of the hessian. This method was developed by Broyden, Fletcher, Goldfarb, and Shanno so-called BFGS Quasi-Newton. BFGS. Quasi-Newton is the have the basic concepts:

$$w_{k+1} = w_k - A_k^{-1} \times gw_k \tag{4}$$

where:

$w_{k+1}$ = is the update weight vector

$w_k =$ is the weight vector before updating

$A_k =$ is the second derivative hessian matrix

$g =$ gradien

*3.2.3. Scaled Conjugate Gradient.* Scale conjugate gradient (SCG) is a supervised learning algorithm for feedforward neural networks, and is a number of the class of conjugate gradient methods. They are general purpose second order techniques that help minimize goal functions of several variables. Second order means that these methods make use of the second derivatives of the goal function, while first-order techniques like standard bacpropagation only use the first derivatives. A second order technique generally finds a better way to a local minimum than a first order technique, but at a higher computational cost. The development in the family of SC training algorithm for neural networks named with a SCG training algorithm [3]. All using the conjugate gradient algorithm will do the on lines search process continuously during the iteration process. For large amounts of data, then the iteration will also be large, so it takes a long time. Because the SCG will make improvements matter. In SCG training algorithm the weights are updated using the following set of equation [24]:

$$w_{m,t} = w_t + \gamma_t d_t \tag{5}$$

$$w_{t+1} = w_t + \alpha_t + d_t \tag{6}$$

$$s_t = E^{''}(w_t)d_t \approx \frac{E^{'}(w_{w,t}) - E^{'}(w_t)}{\gamma_t} \tag{7}$$

$$\alpha_t = \frac{-d_t^T E^{'}(w_t)}{d_t^T s_t} \tag{8}$$

where,

$w_{m,t} =$ is a temporal weight vector which lies between $w_{t+1}$ and $w_i$.

$d_t =$ is the conjugate direction vector of the temporal weight at $t^{th}$ iteration

$\gamma_t =$ is the temporal weight updating step size called the short step size such that $0 < \gamma_t < 1$

$w_{t+1} =$ is the next weight update vector.

$w_t =$ is the vector of current weight.

$\alpha_t =$ is the actual weight updating step size called the long-step size

$s_t =$ is the second order information

$E^{'} dan E^{''} =$ are the first and second derivative of error information with respect to respective weight vectors.

$d_t^T =$ is the transpose $d_t$.

In SCG algorithm in any iteration, the temporal weights $w_{m,t}$, is calculated fist using the short- step size $\gamma_t$ (equation 7). The temporal weight is then used to determine the long step-size size $\alpha_t$ (equations 6 and 8). The final weight update is computed using equation 8.

*3.3. Radial Basis Function*
Radial basis functions (RBF) network is training using a supervised training algorithm which can application for classification problems, function approximation, noisy interpolation and regularization [25][26]. The RBF applications in the field of hydrology has been widely used to non-linear rainfall-runoff model, suspended sediment load, stream flow forecasting to short-term and long-term and

prediction of pore-water pressure [27]-[29]. Radial basis function the generally consists of three layers, namely input layer, hidden layer (radial basis layer) and output layer. The input layer consist nodes, which contain the input variable in the form dimensionality of the input vector, while the hidden layer consists computation unit, where each unit is mathematically described by radial basis function. The hidden to output layer part operates like a standard feed-forward multilayer perceptron network, with the sum of the weighted hidden unit activations giving the output unit activations. The link between nodes input and nodes hidden layer are direct with no weight value, while connected between hidden layer and output layer using the weight value. The output layer consist nodes is connected with the previous nodes in the hidden layer by linear weight, no restriction on the size on the output layer, but that typically the size of the output layer has been always smaller than that of the hidden layer [3][20].
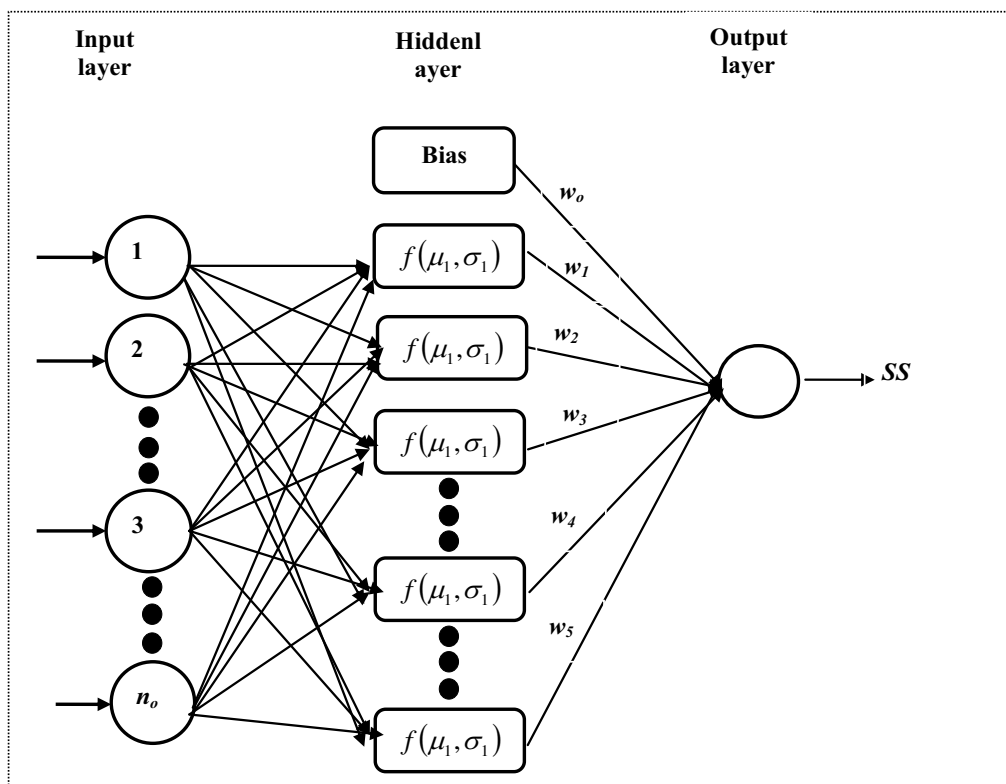


**Figure 2.** The structure of radial basis function neural network

In this study use the popular Guassian basis function as the transfer function at the hidden layer, most commonly used in supended sediment load [28], rainfall-runoff model [27] and streamflow forescasting [26].

$$\phi_j(x) = \exp\left( -\frac{1}{2\sigma_j^2} \parallel x_i - \mu_j \parallel^2 \right) \tag{9}$$

where:
$x_i$ : $n_o$ = dimensional input vector, $i = 1,2 \ldots P$

$\mu_j$ = mean (center), $j$ = 1,2 ...$n$

$\sigma_j$ = standard deviation (spread), $j$ = 1,2 ...$n$

$\phi_j$ = basis function value

$n$ = number of hidden nodes.

The linear mapping between the hidden layer and output layer is given by:

$$SSQ = \sum_{k=1}^{n} \phi_j w_k + w_o , \quad i = 1,2..., P \tag{10}$$

where:

$SSQ$ = output values, in this study suspended sediment discharge corresponding to $X_i$ input vector.

$w_k$ = connection weight;

$w_o$ = bias term;

### 4. Selection of Activation Function/Transfer Function

Multilayer networks often use the log-sigmoid transfer function (logsig) in the hidden layer, while in output layer use linear transfer function (purelin). The function logsig generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity [22]. Logistic sigmoid transfer function are more suitable for use on non-linear data trained with feedforward and more often used as the function between 0 and +1 and simple for derivatives. While activation function in the output layer transfer function using purelin, as desired a form of network outputs any number of real values, not on values between -1 and 1 or 0 and 1. The logistic sigmoid transfer function in hidden layer and linear transfer function in output layers are represented as [3][20][30].

$$f(x) = \frac{e^x - e^{-x}}{e^x - e^{-x}} , \text{ with derivative } f'(x) = [1 + f(x)][1 - f(x)] \tag{11}$$

$$f(x) = x, \text{ with derivative } f'(x) = 1$$

where, $x$ is the input data.

### 5. Material and methods

#### 5.1. Data source and study area

In this study water depth data is measured starting in January 1, 2011 until December 31, 2013. The changes of water depth are monitoring every 10 minutes using Omega-CPM instrument. The sampling of suspended sediment and flow discharge is taking during the dry period (no rain) and rainfall period. The number of suspended sediment sampling and flow of discharge is 1095 (2011-2013). The rainfall data is monitored using tree tipping-bucket rain gauge and every event drop rainfall data will the stored in Hobo Event Data Logger.

Jenderam catchment are is located in Selangor Malaysia . The catchments lies between at the $101^0$ 40 ' 00" to $101^0$ 50 ' 00" E longitude and $02^0$ 45'00 " to $02^0$ 50'00 " S latitude. Jenderam catchment is 20.51 $km^2$ in size. The texture characteristics of soils its a fine sandy loam, weak to moderate grade, the soils susceptible to erosion when rains occur mainly on slopes that disorders by human, such as land clearing and conversion of forest into residential and agricultural areas. Jenderam catchment has an elongated shape, with the elongation ratio of 0.44. The peak of hydrograph runoff is characterized by up and down a quickly. The number of stream order and drainage frequency is a 183 and 8.93 respectively. The study area consists of the Serdang - Kedah soil series , classified as: typic paleudult, fine loamy, kaolinitic, isohyperthermic (USDA Soil Taxonomy).
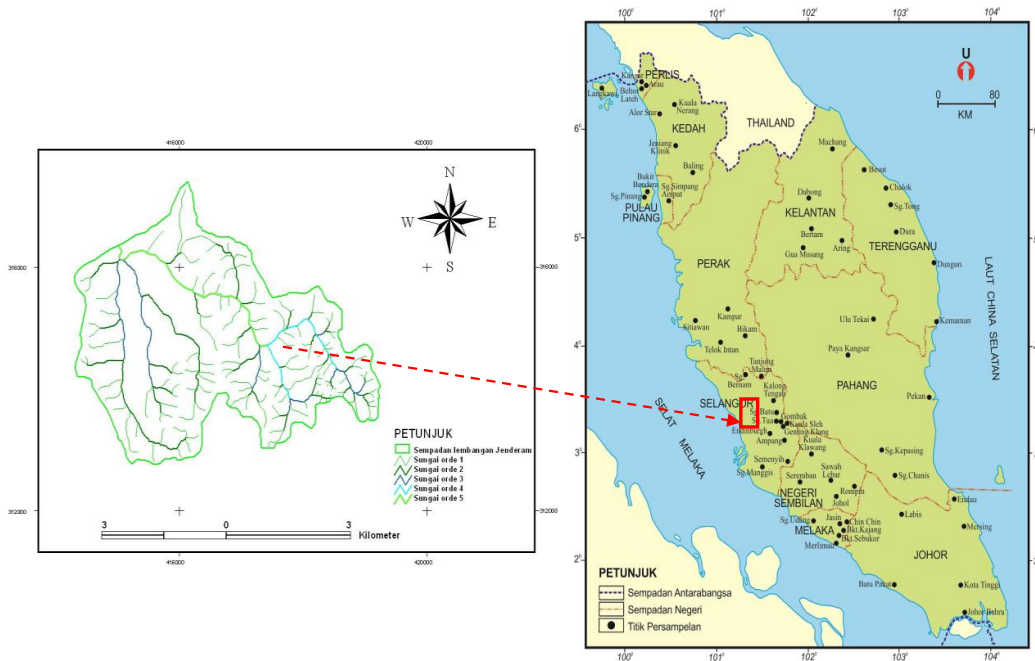
**Figure 3.** Location map of study area (catchment Jenderam) in Selangor Malaysia

### 5.2. Normalization of data

The normalization the raw data it is necessary to make neural network more efficient every step process on the network inputs and target [22]. Normalize data to ensure fast convergence and minimalization of global error during network training [31]. Most studies on the suspended sediment discharge prediction using ANN, have a used data normalization (scaling) relationship of the form [32]:

$$X_n = \left( \frac{X_i - X_{min}}{X_{max} - X_{min}} \right) \tag{12}$$

$X_n$: is the normalized value
$X_i$: is the original value
$X_{max}, X_{min}$: are the maximum and minimum original value data respectively.

Generally, the normalization step is applied to both the input vectors and the target vectors in the data set. The input and target data were normalized before is using for training and testing by transforming the data to the range of 0 to 1. The network output can then be reverse transformed back into the units of the original target data when the network is put to use in the field. Usually the input data should be normalized to commensurate with the limit of the activation function used in the network [33].

*5.3. Input data selection, input layer combination and statistical analyses*

Input data selection and input layer combination play a vital role in the development of an appropriate neural network [3][32]. The modeling using artificial neural network, when the using multilayer networks the generally data is divided into three subsets namely training, validation and testing data. The first subset is the for the training set, which is used for computing the gradient and updating the networks weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The test set error is not used during training, but it is used to compare different models [22]. For this study only use two subset data, about 66.7% (January 1, 2011 to December 31, 2012) have been selected for training and 33.3 % (January 1,  to Dismember 31, 2013) for testing the model.

The statistical parameters for the training and testing subsets, include maximum ($X_{max}$), minimum ($X_{min}$), mean ($X_{mean}$), standard deviation (SD), skewness (Skn) and coefficient of variation (CV).

*5.4. Selection of model performance evaluation criteria*

The performance of multiple linear regression, multilayer perceptron feedforward and radial basis function model evaluation criteria using the standard statistical measures, namely root mean square error (RMSE) [34][35], mean absolute error (MAE) [34][35], coefficient of determination ($R^2$) and coefficient of efficiency (CE) defined as:

Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(SSQ_{pi} - SSQ_{oi}\right) \qquad (13)$$

Root Mean Square Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(SSQ_{pi} - SSQ_{oi}\right)^2}{n}} \qquad (14)$$

Coefficient of Determination

$$R^2 = \left[\frac{\sum_{i=1}^{n}\left(SSQ_{oi} - \bar{SSQ_o}\right)\left(SSQ_{pi} - \bar{SSQ_p}\right)}{\sqrt{\left(\sum_{i=1}^{n}SSQ_{oi} - \bar{SSQ_o}\right)^2}\sqrt{\left(\sum_{i=1}^{n}SSQ_{pi} - \bar{SSQ_p}\right)^2}}\right]^2 \qquad (15)$$

Coefficient of Efficiency

$$CE = 1 - \frac{\sum_{i=1}^{n}\left(SSQ_{oi} - SSQ_{pi}\right)^2}{\sum_{i=1}^{n}\left(SSQ_{oi} - \bar{SSQ_p}\right)^2} \qquad (16)$$

where $SSQ_{oi}$ and $SSQ_{pi}$ are the observed and predicted values respectively, $\bar{SSQ_o}$ and $\bar{SSQ_p}$ are the mean observed and predicted value respectively, and *n* is the number of observations. Ideally , the value of MSE, RMSE and MAE should be zero, while $R^2$ and CE should be one [3]. The error value

can be zero and the value of the coefficient of determination can be one, if the observed and predicted value is the same.

**Table 1.** Model data input for MLRg, MLP and RBF

|  | Data Input | Total data | Model | Input layer variables | Target |
|---|---|---|---|---|---|
| Training | 1. Rainfall (mm/day), <br> 2. Water depth (m/day) <br> 3. Discharge (m$^3$/day) <br> 4. Suspended sediment discharge (kg/day) | January 1, 2011 – December 31, 2012 | MLP <br> RBF | 1. $P_t$, $H_t$, $Q_t$ <br> 2. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ <br> 3. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$ <br> 4. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | SSQ |
| Testing | 1. Rainfall (mm/day) <br> 2. Water depth (m/day) <br> 3. Discharge (m$^3$/day) <br> 4. Suspended sediment discharge (kg/day) | Januari 1, 2013 – Desember 31, 2013 | MLRg <br> MLP <br> RBF | 1. $P_t$, $H^t$, $Q_t$ <br> 2. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ <br> 3. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$ <br> 4. $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | SSQ |

Where:

$Q_t$ , $P_t$, $H_t$ : discharge, rainfall and  average of the water depth at the current time respectively.

$Q_{t-1}$, $Q_{t-2}$ , $Q_{t-3}$: discharge at one, two and three previous day.

$P_{t-1}$, $P_{t-2}$, $P_{t-3}$: rainfall at one, two and three previous day.

$H_{t-1}$, $H_{t-2}$, $H_{t-3}$:water depth at one, two and three previous day.

**Table 2.** Summary of statistics analysis on rainfall, discharge and suspended sediment discharge data.

| Data set | Data Type | Min | Max | Mean | SD | Skewness | CV |
|---|---|---|---|---|---|---|---|
| Training | Rainfall (mm/day) | 0.0000 | 93.9000 | 7.4321 | 14.5637 | 2.8528 | 1.9596 |
|  | Water depth (m/day) | 0.0046 | 0.5840 | 0.1350 | 0.0600 | 2.6549 | 0.4446 |
|  | Discharge (m$^3$/day) | 1,707.7532 | 1,2189,923.526 | 30,902.2424 | 70,005.8353 | 13.8342 | 0.4445 |
|  | Suspended sediment discharge (kg/day) | 4.3762 | 3,550.3815 | 190.9396 | 276.8457 | 6.4555 | 1.4499 |
| Testing | Rainfall (mm/day) | 0.0000 | 141.1000 | 5.1956 | 12.7295 | 5.4183 | 0.5629 |
|  | Water depth (m/day) | 0.0299 | 0.6530 | 0.0995 | 0.0560 | 3.2889 | 2.2581 |
|  | Discharge (m$^3$/day) | 1703.7532 | 1,218,992.353 | 30,952.9815 | 69,990.7580 | 13.8412 | 2.2612 |
|  | Suspended sediment discharge (kg/day) | 0.4723 | 4465.3437 | 87.8384 | 253.3504 | 14.5819 | 2.8843 |

## 6. Result and discussion

### 6.1. Multiple Linear Regressions

The analysis of the rainfall, water depth, discharge and daily suspended sediment discharge height the variability of the value the Jenderam catchment. The maximum, minimum, mean, standard deviation and coefficient of variation is show in Table 2. The general purpose of the multiple linear regressions (MLRg) is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable [18]. Multiple linear regression analysis was carried out to correlate the measured suspended sediment discharge to the variables independed, namely rainfall, water depth and discharge. The formula multiple linear regression models to predict the suspended sediment discharge (SSQ) are given below (Table 3). The performance analysis multiple linear regression using four models data independent with given in Table 4.

**Table 3.** Formula MLRg for prediction SSQ

| Variable Independent | Variable dependent | Formula |
|---|---|---|
| $P_t$, $H_t$, $Q_t$ | SSQ | $SSQ = 27.242 - 0.067P_t - 635.862H_t + 343.303Q_t$ |
| $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | SSQ | $SSQ = 33.074 - 0.033P_t - 550.436H_t + 343.394Q_t - 0.108P_{t-1} - 164.980H_{t-1} + 10.869Q_{t-1}$ |
| $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | SSQ | $SSQ = 35.336 - 0.004P_t - 534.859H_t + 342.920 Q_t - 0.092P_{t-1} - 133.771H_{t-1} + 9.279Q_{t-1} - 0.028 P_{t-2} - 81.602H_{t-2} + 5.236Q_{t-2}$, |
| $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | SSQ | $SSQ = 36.094 + 0.074P_t - 532.906H_t + 342.674Q_t - 0.056P_{t-1} - 119.269H_{t-1} + 9.043Q_{t-1} - 0.009P_{t-2} - 50.851H_{t-2} + 3.931Q_{t-2} - 0.121P_{t-3} - 51.285H_{t-3} + 0.676 Q_{t-3}$ |

The performance of the statistics value for to four MLRg models with the input variable independent 3 ($P_t$, $H_t$, $Q_t$), 6 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$), 9 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$,) and 12 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$) were established (Table 4). The data set training and testing used for ANN, while for MLRg model only using testing data to develop the model. The result predicted suspended sediment discharge by MLRg compared with the predicted by ANN models.

**Table 4.** The Performance of statistical value Multilinear Regression

| No | Input Variable | MAE | RMSE | $R^2$ | CE |
|---|---|---|---|---|---|
| Model 1 | $P_t$, $H_t$, $Q_t$ | 0.0000002 | 14.4113 | 0.9970 | 0.9967 |
| Model 2 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | 0.0000002 | 13.6039 | 0.9971 | 0.9971 |
| Model 3 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | -3.8356 | 72.8125 | 0.9774 | 0.9172 |
| Model 4 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | -4.0274 | 76.4106 | 0.9747 | 0.9088 |

The MAE, RMSE, $R^2$ and CE of the MLRg model are given in the Table 4. The MAE, RMSE, R2 and CE of the MLRg value Model 2 (6 input variable independent) has the lowest the value of MAE and RMSE (0.0000002 and 13.6039) and highest $R^2$ and CE (0.9970 and 0.9971). The comparison between observed and predicted suspended sediment discharge by MLRg are plotted in Figure a1 – d2. The perfect line using MLRg model 2 with a 6 variable independent produced $R^2$ value 0.9971, with highest than the model 1 ( $R^2$ 0.9970), while model 3 ($R^2$ 0.9774), and the model 4, $R^2$ and CE (0.9747 and 0.9088) only a slightly different. Even between MLRg model 1 and model 2 (3 and 6 variable independent) there was no significance difference in prediction, because the value MAE, RMSE, $R^2$ and CE difference only slightly (Table 4). Increasing the number of independent variables in the model MLRg, suspended sediment discharge prediction has a trend less accurate. Prediction obtained the overall MLRg model (four model) were completely free of negative prediction. The MAE of MLRg models 3 and 4 have a negative value, this means the total value of the prediction is lower than the total value of observation. While the MAE of MLRg models, has a positive value, this means the total value of prediction is higher than the total value of the observation. Based on result show that the MLRg model can be success applied to prediction daily suspended sediment discharge in Jenderam catchments with has size 20.51 km$^2$. This mean that suspended sediment concentration positively correlated with the daily rainfall, water depth and discharge the current time, likewise the one, two and three previous day.
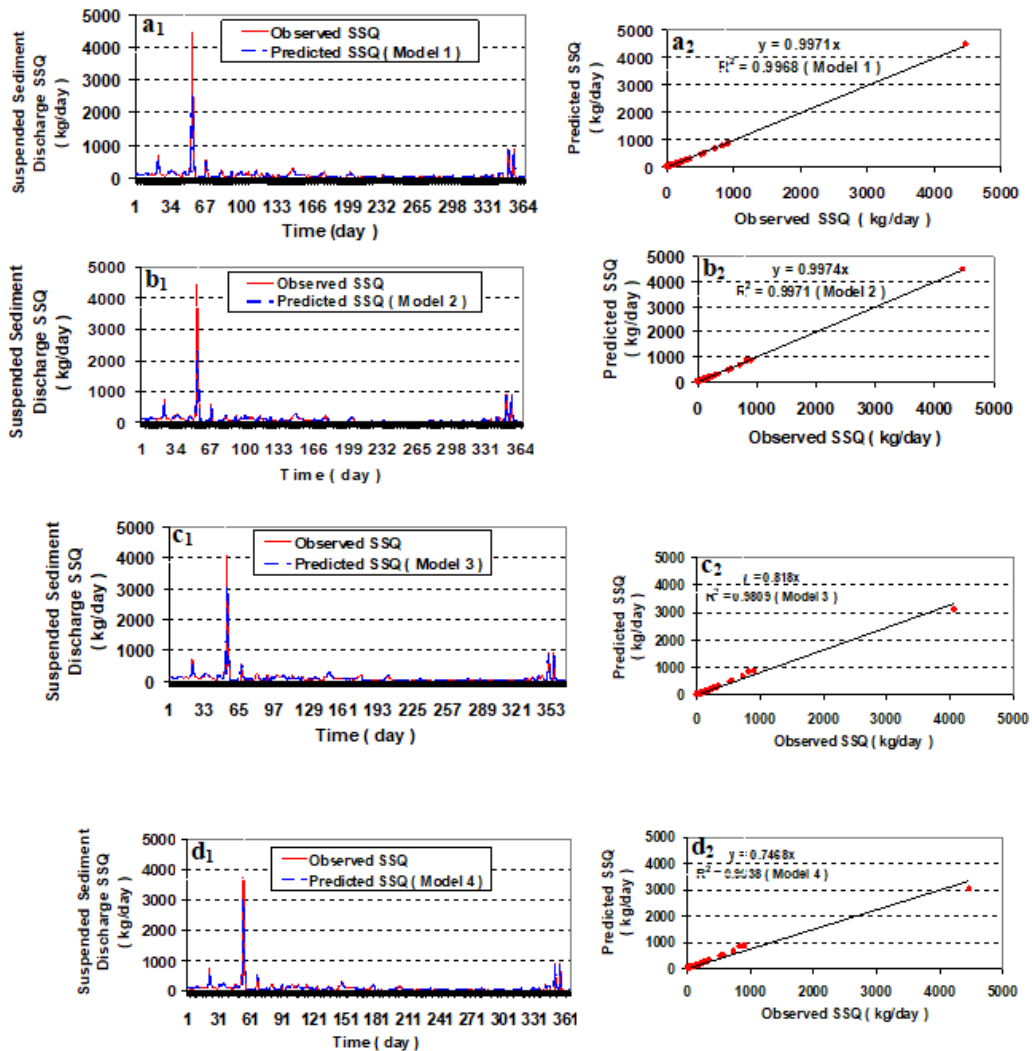
**Figure 4** Comparisons between observed and predicted SSQ using MLRg

*6.2. Multilayer Perceptron Feedforward*

The performance of the multilayer perceptron (MLP) feed forward using the algorithms LM, BFGS-Quasi Newton and SCG using the daily data is show in Table 5,6 and 7. Based on the selected input data structure and two year of record for training and one year for testing the comparison between observed and predicted suspended sediment discharge by LM, BFGS and SCG algorithms are plotted in Figure $a_1$-$s_2$. The number nodes (neuron) in input layer are 3 ($P_t$, $H_t$, $Q_t$), 6 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$), 9 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$,) and 12 ($P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$ , $H_{t-3}$, $Q_{t-3}$) respectively. Meanwhile the number nodes in hidden layer are determined by trial and error method. The number of output neurons only one because only one target output (suspended sediment discharge) from the ANN model.

The performance of the LM algorithm to prediction suspended sediment discharge based on number neuron in input, hidden and output layer is show in Table 5. The performance of statistics value (MAE, RMSE, $R^2$ and CE) at the processes training and testing, the best result on structure 6-10-1 (6 neurons in input layer, 10 neuron in hidden layer and 1 neuron in output layer). The performance value in testing process, MAE and RMSE (13.9107 and 18.9514) is smallest, meanwhile $R^2$ and CE (0.9953 and 0.9954) is the highest if it compared with the another LM model (3-7-1, 9-10-1 and 12-12-1). Testing result obtained by LM model 6-10-1 followed the observed suspended sediment discharge very closely the whole span of the testing data. The MAE LM 3-7-1 and LM 6-10-1Is a negative, this means the total value of the prediction is lower than the total value of observed. While the MAE of LM 9-10-1 and 12-12-1 is a positive, this means the total value of prediction is higher than the observed.

Table 5. The performance of statistical algorithm LM

| No | ANN Model Inputs | Lavenberg-Marquard (LM) | | | | | | | | |
| | | Structure | Training | | | | Testing | | | |
| | | | MAE | RMSE | $R^2$ | CE | MAE | RMSE | $R^2$ | CE |
| 1 | $P_t$, $H_t$, $Q_t$ | 3:7:1 | 0.5659 | 3.3976 | 0.9999 | 0.9998 | 15.9973 | 21.4476 | 0.9958 | 0.9928 |
| 2 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | 6:10:1 | 0.5503 | 3.3953 | 0.9999 | 0.9998 | 13.9107 | 18.9514 | 0.9953 | 0.9944 |
| 3 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | 9:10:1 | 0.8389 | 3.5256 | 0.9999 | 0.9998 | 18.6625 | 25.9216 | 0.9907 | 0.9895 |
| 4 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | 12:12:1 | 0.7056 | 3.3736 | 0.9999 | 0.9999 | 24.7100 | 25.6389 | 0.9915 | 0.9897 |

The results demonstrate the accuracy and reliability of the proposed LM model to prediction suspended sediment discharge in Jenderam catchments.

The performance of the BFGS Quasi-Newton algorithm to prediction suspended sediment discharge based on number neuron in input, hidden and output layer is show in Table 6. The performance of statistics value (MAE, RMSE, $R^2$ and CE) at the processes training and testing, the best result on structure 3-7-1 (3 neurons in input layer, 7 neuron in hidden layer and 1 neuron in output layer). The performance value in testing process, MAE and RMSE (13.5769 and 17.9011) is smallest, meanwhile $R^2$ and CE (0.9999 and 0.9998) is the highest if it compared with the another BFGS Quasi-Newton model (6-3-1, 9-10-1 and 12-12-1). Testing result obtained by BFGS Quasi-Newton model 3-7-1 followed the observed suspended sediment discharge very closely the whole span of the testing data.
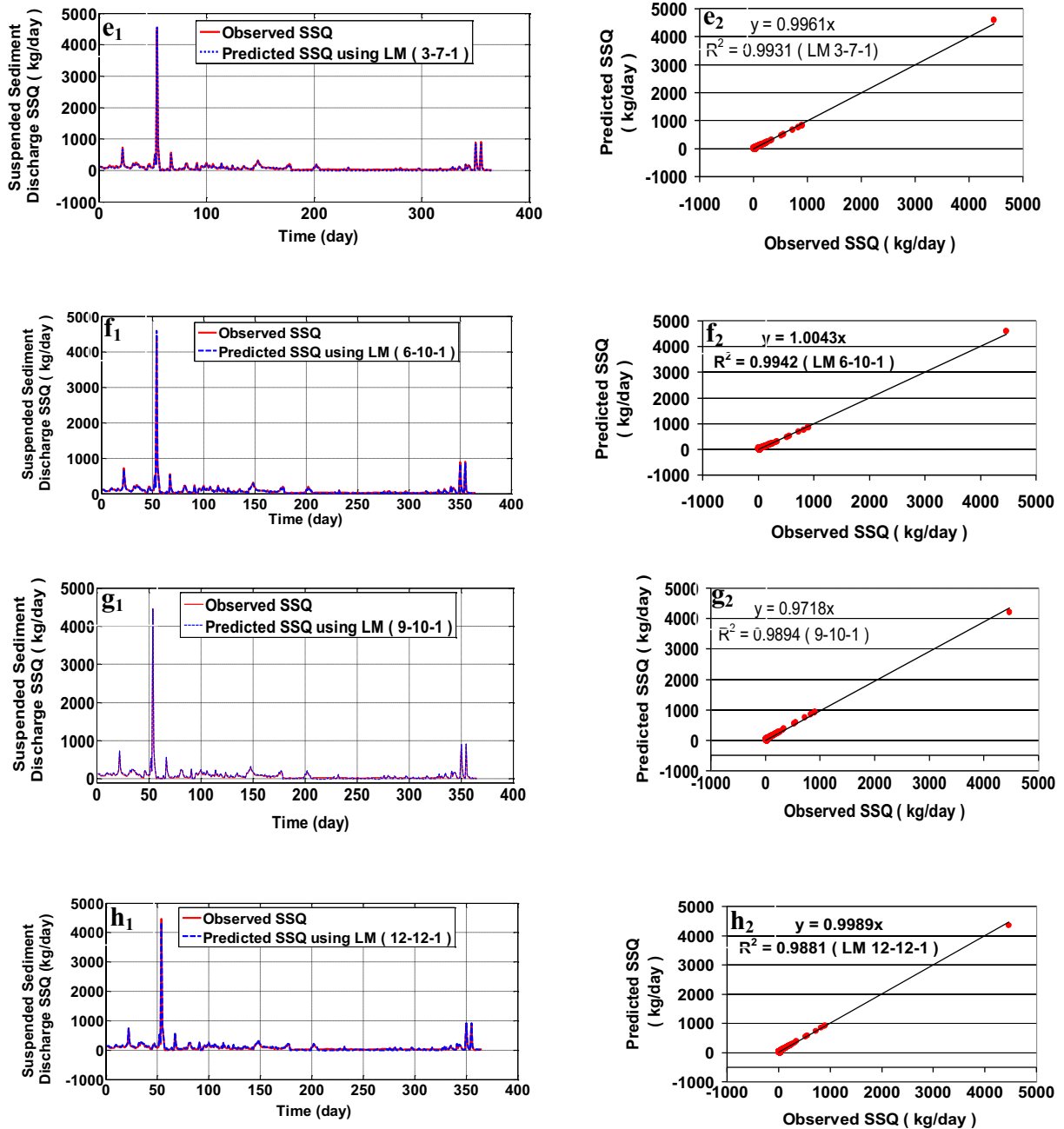
**Figure 5** Comparisons between observed and predicted SSQ using LM algorithm based on the testing data.
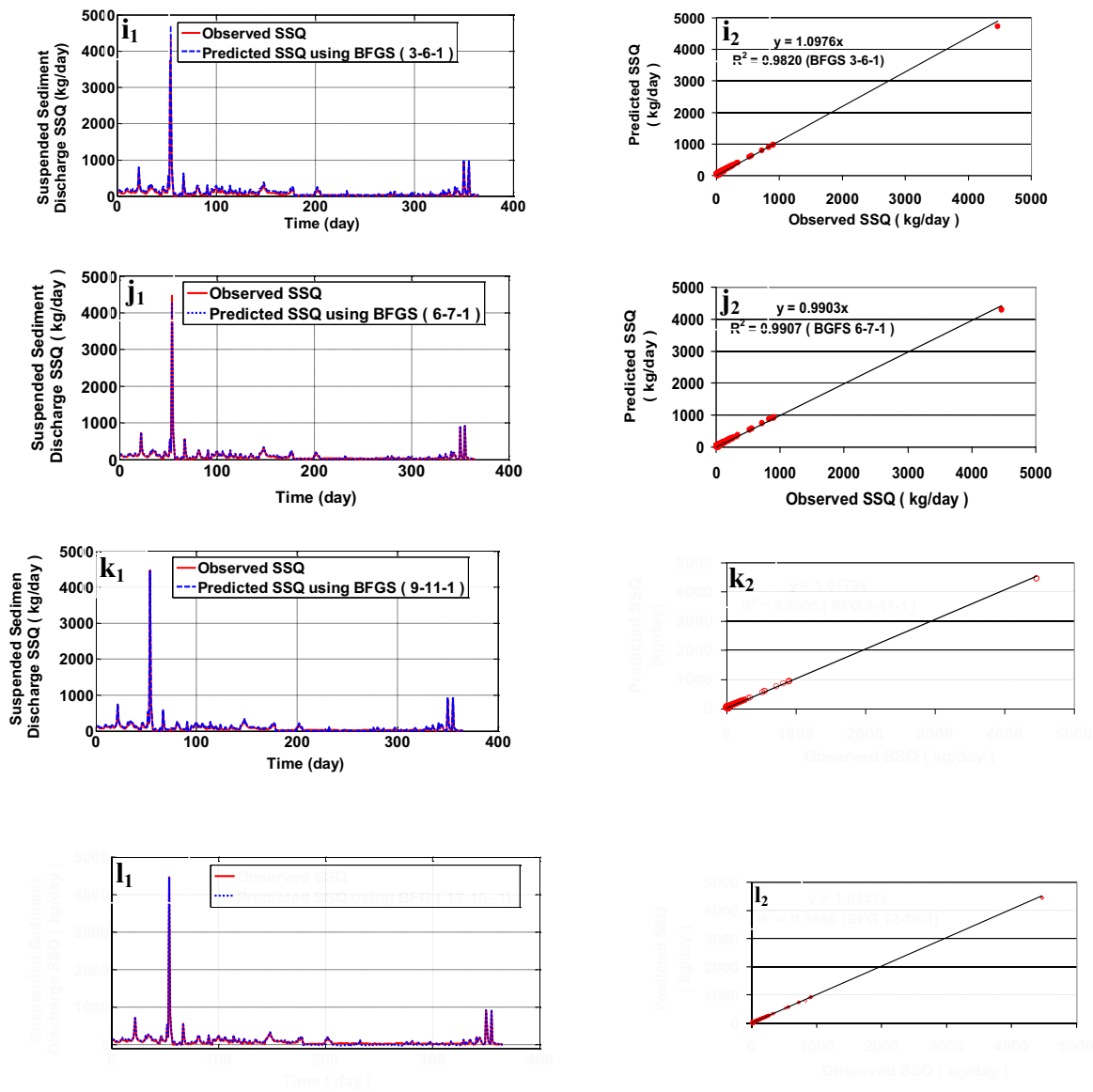
**Figure 6** Comparisons between observed and predicted SSQ using BFGS Quasi Neuton algorithm based on the testing data.

**Table 6.** The performance of statistical algorithm BFGS Quasi-Newton

| No | ANN model inputs | structure | BFGS Quasi-Newton | | | | | | | |
|----|------------------|-----------|------|------|-----|----|------|------|-----|----|
| | | | Training | | | | Testing | | | |
| | | | MAE | RMSE | $R^2$ | CE | MAE | RMSE | $R^2$ | CE |
| 1 | $P_t$, $H_t$, $Q_t$ | 3:7:1 | | | | | 13.5769 | 17.9011 | 0.9999 | 0.9998 |
| 2 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | 6:7:1 | 1.1414 | 3.8007 | 0.9999 | 0.9998 | 19.8777 | 23.1250 | 0.9950 | 0.9916 |
| 3 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | 9:11:1 | 0.6732 | 3.4310 | 0.9999 | 0.9998 | 25.6947 | 27.2228 | 0.9899 | 0.9884 |

| No | ANN model inputs | structure | BFGS Quasi-Newton | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Training | | | | Testing | | | |
| | | | MAE | RMSE | $R^2$ | CE | MAE | RMSE | $R^2$ | CE |
| 4 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | 12:16:1 | 1.2361 | 3.9058 | 0.9999 | 0.9998 | 24.9865 | 27.4468 | 0.9888 | 0.9882 |

The performance of the SCG algorithm to prediction suspended sediment discharge based on number neuron in input, hidden and output layer is show in Table 7. The performance of statistics value (MAE, RMSE, $R^2$ and CE) at the processes training and testing, the best result on structure 12-14-1 ( 12 neurons in input layer, 14 neuron in hidden layer and 1 neuron in output layer). The performance value in testing process, MAE and RMSE (13.5769 and 17.9011) is smallest, meanwhile $R^2$ and CE (0.9999 and 0.9998) is the highest if it compared with the another SCG model (3-4-1, 6-8-1 and 9-13-1). The MAE SCG 3-4-1 and SCG 12-14-1 Is a negative, this means the total value of the prediction is lower than the total value of observed. While the MAE of BFGS is all positive, this means the total value of prediction is higher than the observed. Testing result obtained by SCG model 12-14-1 followed the observed suspended sediment discharge very closely the whole span of the testing data.

**Table 7.** The performance of statistical scaled conjugate gradient

| No | ANN model inputs | structure | SCG | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Training | | | | Testing | | | |
| | | | MAE | RMSE | $R^2$ | CE | MAE | RMSE | $R^2$ | CE |
| 1 | $P_t$, $H_t$, $Q_t$ | 3:4:1 | 1.7751 | 4.4892 | 0.9999 | 0.9997 | 19.9337 | 38.8298 | 0.9927 | 0.9925 |
| 2 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | 6:8:1 | 1.8234 | 4.3843 | 0.9999 | 0.9997 | 24.5762 | 26.0567 | 0.9908 | 0.9894 |
| 3 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | 9:13:1 | 1.6950 | 4.3485 | 0.9999 | 0.9998 | 25.4747 | 27.8591 | 0.9888 | 0.9879 |
| 4 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | 12:14:1 | 1.8641 | 4.3932 | 0.9999 | 0.9997 | 18.2077 | 26.5019 | 0.9917 | 0.9890 |

*6.3. Radial Basis Function*

In this study, RBF models only apply four structures in the input, hidden and output layers, are namely: 3-3-1, 6-6-1, 9-9-1 and 12-12-1. To obtain the best performance of statistical value every RBF models structure, done method trial and error by the change spread value between 1 to 5. The best performance for structure 3-3-1 is a spread value 1, and to structure 6-6-1, 9-9-1 and 12-12-1 spread value is 2. The performance, four structure RBF models show and are summarized in Table 8. The comparison between observed and predicted is plotted in Figure (5-9). In the training and testing processes show the RBF structure 3-3-1 ( 3 neurons in input layer, 3 neurons in hidden layer and 1 neuron in output Layer) has a better and high prediction performance. The MAE, RMSE ( 0.6684 and 3.4646) training process has a smaller, meanwhile $R^2$ and CE (0.9998 and 0.9998) has a highest. Whereas in testing processes show the RBF model structure 3-3-1, MAE and RMSE (15.0265 and 19.6195) has a smaller, on the other hand $R^2$ and CE (0.9944 and 0.9940) has a highest. Based on performance the all RBF models structure input, hidden and output layer the different not significance, because difference only a slight and or performance relative similarly.

**Table 8.** The performance of statistical value RBF

| No | ANN model inputs | structure | Spread | RBF | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Training | | | | Testing | | | |
| | | | | MAE | RMSE | $R^2$ | CE | MAE | RMSE | $R^2$ | CE |
| 1 | $P_t$, $H_t$, $Q_t$ | 3:3:1 | 1 | 0.6684 | 3.4646 | 0.9998 | 0.9998 | 15.0265 | 19.6195 | 0.9944 | 0.9940 |
| 2 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$ | 6:6:1 | 2 | 1.0877 | 3.4817 | 0.9998 | 0.9998 | 29.4478 | 34.2146 | 0.9834 | 0.9817 |
| 3 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, | 9:9:1 | 2 | 1.2041 | 3.5195 | 0.9998 | 0.9998 | 28.5460 | 32.2658 | 0.9847 | 0.9837 |
| 4 | $P_t$, $H_t$, $Q_t$, $P_{t-1}$, $H_{t-1}$, $Q_{t-1}$, $P_{t-2}$, $H_{t-2}$, $Q_{t-2}$, $P_{t-3}$, $H_{t-3}$, $Q_{t-3}$ | 12:12:1 | 2 | 1.2886 | 3.5433 | 0.9998 | 0.9998 | 26.8016 | 31.0912 | 0.9857 | 0.9849 |

In this study the all performance of the LM, BFGS, SCG and RBF the training result in training process is the better than the result in testing process. The LM learning algorithm attained the required accuracy with less number of the iteration. The LM achieved the goal before 5000 epoch and the convergent faster and shortest time than SCG and BFGS. When compared between MLRg, LM, SCG, BFGS and RBF , the MLRg model 2 with the 6  input variable is the better and more  accurate to prediction suspended sediment discharge in Jenderam catchments.

The special for RBF, the  MAE negative only on structure data input 3-3-1, while for structure data input 6-6-1, 9-9-1 and 12-12-1 the all MAE is a positive. Based on the performance statistics value, MLRg, LM, SCG, BFGS and RBF suitable for modeling the non-linear complex behavior of suspended sediment discharge.

## 7. Conclusion
In this paper, the MLRg, MLP ( LM, SCG and BFGS) and RBF are used prediction daily suspended sediment discharge in Jenderam catchments Selangor Malaysia. The perfect line between observed and predicted, the MLRg model 2 with a 6 variable independent produced $R^2$ value 0.9974, with highest than the model 1 ( $R^2$ 0.9968), model 3 ($R^2$ 0.9809) and  model 4  ( $R^2$ 0.9638) only a slightly different. Even between MLRg model 1 and model 2 ( 3 and 6 input variable independent) there was no significance difference in prediction, because the value MAE, RMSE, $R^2$ and CE difference only slightly (Table 4). The all predicted value positive for four MLRg model. Increasing the number of independent variables in the model MLRg, suspended sediment discharge prediction have a trend less accurate, but the difference only slightly. In this study the all performance of the LM, BFGS, SCG and RBF the training result in training process is the better than the result in testing process. When compared between MLRg, LM, SCG, RBF, has a obtained the BFGS model structure 3-7-1  is the better and more  accurate to prediction suspended sediment discharge in Jenderam catchments. Based on performance of MLRg, LM, SCG, BFGS and RBF accurately for modeling the non-linear complex behavior of suspended sediment responses to rainfall, water depth and discharge in small catchments area .

## References
[1]    Kninghton, D. 1998. Fluvial Forms and Processes: *A New Perspective*. Hodder Arnold.
[2]    Mustafa, M.R., Rezaur, R.B., Saiedi, S., & Isa, M.H. 2012. River suspended sediment prediction using various multilayer percepton neural network training algorithms-A case studi in Malaysia. *Water Resour Manage*, **26**:1879-1897.
[3]    Mustafa, M.R., Rezaur, Raharhjo, H., & Isa, M.H. 2012. Prediction of pore-water pressure using radial basis function neural network. *Engineering Geology 135-136*, 40-47.
[4]    Welling, D.E., & Webb,B.W. 1986. *Solute in River Systems*. Chichester Wiley.

[5]   Sanusi, W., Jemain, A. A., Zin, W. Z. W., & Zahari, M. 2015. The drought characteristics using the first-order homogeneous Markov chain of monthly rainfall data in peninsular Malaysia. *Water Resources Management*, **29**, 5, 1523-1539.

[6]   Maru, R. and Ahmad, S. 2015. The relationship between land use changes and the urban heat island phenomenon in Jakarta, Indonesia. *J. of Adv. Sci. Letters*. **21**, 2, pp. 150–152(3).

[7]   Maru, R., Baharuddin, I.I., Umar, R., Rasyid., R., Uca, Sanusi, W., and Bayudin. 2015. Analysis of The Heat Island Phenomenon in Makassar, South Sulawesi, Indonesia. *American J. of Appl. Sci.*. **12**, 9, pp 616-626. DOI: 10.3844/ajassp.2015.

[8]   Maru, R. and Ahmad, S. 2015. The Relationship between Temperature Patterns and Urban Morfometri in the Jakarta City, Indonesia. *Asian J. of Atmospheric Env.* **9**, 2, pp. 128-136.

[9]   Kisi, O., & Guven, A. 2010. A machine code-based genetic programming for suspended sediment concentration estimation. *Adv. in Eng. Soft.* **41**.

[10]  Sadeghi, A.H.R., Mizuyam, T., Miyata, S.,Gomi ,T., Kosugi, K., Fukushima, T., Mizugaki, S., & Onda, Y. 2008. Development, evaluation and interpretation of sediment rating curves for a Japanese small mountainous reforested watershed. *Geoderma*, **144**, pp. 198-211.

[11]  Cigizoglu, H.K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Adv. in Water Resources*, **27**, 185-195.

[12]  Melesse, A.M., Ahmad, S., McClain, M.E., Wang, X. & Lim, Y.H. 2011. Suspended sediment load prediction of river systems: An artificial neural network approach. *Agricultural Water Management*. **98**, 5, pp. 855-866.

[13]  Kutner,M.H., Nachtsheim, C.J., Neter, J., & Li, W. 2005. *Applied linear statistics models*. McGraw-Hill Irwin.

[14]  Montgomery, D.C., Peck, E.A., & Vining. G.G. 2006. *Introduction to linear regression analysis*. Fourth Edition Wiley-Interscience A john Wiley & Sons, Inc Publication.

[15]  Maniquis, MC., Lee S., & Kim, L.H. 2010. Multiple kinear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables. *J. of Env. Sci.* **22**, 6, pp. 946-952.

[16]  Yuill, B.T., & Gasparini, N.M. 2011. Hydrologic controls on wash load sediment concentrations within a low-ordered, ephemeral watershed. *J. of Hydrology,* **410**, pp. 73-83.

[17]  Roman, D.C., Vogel, R.M., & Schwarz, G.E. 2012. Regional regression models of watershed suspended-sediment discharge for the eastern United States. *J. of Hydrology*, **472-473**, pp. 53-62.

[18]  Yilmaz, I., & Kaynar, O. 2011. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Systems with App.* **38**, pp. 5958-5966.

[19]  Kleinbaum, D.G., Kupper, L.L., Muller, K.E., & Nizam, A. 1998. *Applied regression analysis and other multivariable methods*. Duxbury Press

[20]  Haykin, S. 1999. *Neural Networks and Learning Machines*. Prentice Hall International, Inc

[21]  Fauset, L., 1994. Foundation of Neural Network. New Jersey: Prentice-Hall

[22]  Beale, M.H., Hagan, M.T., & Demuth, H.B. 2012. *Neural Netwok Toolbox^{TM} User's Guide Matlab*. The MathWorks, Inc.

[23]  Liu, Q.J., Shi, Z.H., Fang, N.F., Zhu, H.D., & Ai, L. 2013. Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet-ANN approach. *Geomorphology,* **186**, pp. 181-190.

[24]  Moller, M.F. 1993. A scaled conjugate gradient algorithms. Neural Netw. **12**, pp. 145-151.

[25]  Kegl, B., Krzyzak, A., & Niemann, H. 2000. Radial basis function networks and complexity regularization in function learning and classification, *International Conference in Function Learning and Classification*.

[26]  Moradkhani, H., Hsu, K., Gupta H.V., & Sorooshian, S. 2004. Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J. of Hydro.,* **295**, pp. 246-262.

[27]  Lin, G. F., & Chen, L. H. 2004. A non-linear rainfall-runoff model using radial basis function

network. *J. of Hydro.,* **289**, pp. 1-8.

[28]  Alp, M., & Cigizoglu, H. K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Env. Model Sof.* **22***,* pp. 2-13.

[29]  Kagoda, P. A., Ndiritu, J., Ntuli, C., & Mwaka, B. 2010. Application of radial basis function neural networks to short-term streamflow forescasting. *Phy. Chem. of Earth*.

[30]  Bishop, C.M., 2008. Training with noise is equivalent to Tikhonov regularization. *Training*, **7**, 1.

[31]  Rojas, R. 1996. *Neural networks: A systematic introduction*. Springer Verlag, Berlin pp 151-184.

[32]  Cigizoglu, H.K., & Kisi, O. 2006. Methods to improve the neural network performance is suspended sediment estimation. *J. of Hydro.* **3**, 17, pp. 221-238.

[33]  Maier, H.R. and Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Env. Model. Soft.*, **15**, 1, pp.101-124.

[34]  Rahman, A. and Ahmar, A.S., 2017, September. Forecasting of primary energy consumption data in the United States: A comparison between ARIMA and Holter-Winters models. In *AIP Conference Proceedings* (**1885**, 1, p. 020163). AIP Publishing.

[35]  Ahmar, A.S., Rahman, A., Arifin, A.N.M. and Ahmar, A.A., 2017. Predicting movement of stock of "Y" using Sutte Indicator. *Cogent Eco. & Fin.*, **5**, 1, p.1347123.