ZANCO Journal of Pure and Applied Sciences
The official scientific journal of Salahaddin University-Erbil
https://zancojournals.su.edu.krd/index.php/JPAS

ISSN (print):2218-0230**, ISSN (online):** 2412-3986**, DOI: http://dx.doi.org/10.21271/zjpas**

RESEARCH PAPER

Prediction of solar direct irradiance in Iraq by using artificial neural network

Zana Saleem Mohammed[1], Gzing Adil Mohammed[2].

[1] Lafargeholcim, Kurdistan region of Iraq, Sulaymaniyah.

[2] Department of Oil, Gas and Energy administration, Public Administration and Natural Resources, Charmo University.

A B S T R A C T:

Global solar irradiance is one of the main significant factors for designing and considering the volume of any solar station beside of it is usage in agricultural and building issue. Due of lack a precise information about the irradiance in Iraq metrological organization and seismology, this study is aimed to adopt the historical global data, build numerical analysis via using artificial neural network and predicting hourly irradiance. The test is applied over three locations Erbil, Bagdad, and Basra for being references to their closest locations. A foreword neural network (FNN) is the learning algorithm that is used in this study with relying on seven input variables consisting of Temperature, Precipitation, Humidity, Wind speed, Wind direction Sunshine duration and Date. After normalizing and standardizing data, an iteration method is used for determining the optimum number of neuron(s) in a hidden layer. It yields a least Root Mean square error (RMSE) between 2.5 to 3. The computed correlation coefficients are between 0.94 -0.96 for the mentioned locations.

KEY WORDS: Renewable energy, Solar system, Artificial neural network, Prediction.
DOI: http://dx.doi.org/10.21271/ZJPAS.33.5.5
ZJPAS (2021) , 33(5);43-50 .

1.INTRODUCTION :

Sun light is one of the most evolution resource among the renewable energy projects over the recent period. Irradiance measurements get a wide focus of data collection due of having an own merit in many fields such as crop growing, electricity resource, structure and building design (David, Gerrit, & R.W., 1994). Demanding of electricity is raised up in Iraq and it is depending dominantly on Fossil fuel for generating electricity (S. & M., 2009). Solar photovoltaic investment starts to be convinced after the reflection of shortage of generating national electricity source beside of increasing service cost for the commercial power that produce by local generators.

* **Corresponding Author:**
Zana Saleem Mohammed
E-mail: zana.saleem82@gmail.com
Article History:
Received: 10/04/2020
Accepted: 07/08/2020
Published: 20/10 /2021

Variant geographical aspects need to be considered to design the solar system applications. Distribution of solar irradiance is the one of the significant factors that need to be paid attention through being involved in variety projects, thus resulting least project cost impact.

Investing solar energy in Iraq need to be focused which will boost the national electricity production. In high electric demand season, supplying electrical power to many industries are limited due to shortage of electrical production. Accordingly these businesses are continuously looking for alternative resources to cover the gap.

Many climate variables have influential factor of photovoltaic (PV) module and their performance. Generally, the average daily sunshine in Iraq is around 7-8 hours in winter, while it is around 10-11 hours in summer. Getting an accurate predicted data for weather, irradiance and solar power are continuously demanded in many cities in Iraq. Build a prediction model is helping in designing the relevant application such as solar power generation system. Other factors have impact on the PV system such as temperature, precipitation, and wind.

A common algorithm used for modeling irradiance prediction is Artificial neural network (Adel & Soteris, Artificial intelligence techniques for photovoltaic applications: A review, 2008), this tool is learned from experience and historical data. The model will be precise as much as the data is available (Negnevitsky, 2002).

2.Review of literature

Many aspects in solar irradiance prediction have been developed over the last ten years, **(Alzahrani, Kimball, & Dagli, 2014)**. The model was formulated by considering the inputs as time, day of the year, sky cover, pressure and wind speed and used by many researchers. Increase the accuracy of irradiance prediction by using nonlinear autoregressive neural network resulting a better performance than liner regression method. In Zimbabwe **(Chiteka & Enweremadu, 2016)** try to predict the irradiance with rely on neural network by selecting altitude, latitude, longitude, humidity, pressure, clearness index and average temperature as input. The absolute percentage error was 2.56%.

Monthly average solar radiation has been predicted in four locations in Uganda by (Mubiru & Banda, 2007). The used method was depending on artificial neural network with 8 neurons and one layer. Sunshine duration, temperature, cloud cover and location have been used as an input. The mean square error was 0.38 $MJ.m^2$. Multi-layer neural network was used by (Adel & Alessandro, A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, 2010) to develop correlation coefficient for sunny and cloudy day. Try and error was used for finding the best activation function.

Solar irradiance corelates with various weather and geographical data, which considered as non-linear relation, artificial neural network is common method which used for such application (Adel M. , 2008).

In this article, the focus is on selecting a sufficient transfer function, getting the best number of neurons and improving the status of training ratio versus validation ratio via using forward neural network (FNN). This tool producing a high effective result to deploy prediction module.

3.Material and methods

3.1 Locations and data preparation

Three main locations are selected at Erbil, Baghdad, and Basra. The aim study is to test prediction algorithm and to get a reliable mathematical model and programming statements to build a rapid and accurate prediction architecture. Lacks of solar database in Iraq metrological organization and seismology in targeted sites in Iraq were the main struggle in this study. In addition, requiring a specific tool for measuring irradiance in different part of Iraq are costly and hard to be applied. The data are collected from MeteoBlue website as global weather database for one month period Feb, 2021, (meteoblue, 2006).

The collected wide range of dataset is on hourly bases. Each sites latitude and longitude are addressed in Table 1. The results can be reflecting into surrounded and closed area (Kais, Munya, & Zahraa, 2012).

Set of historical weather variables have been assigned as inputs. The model is mainly based on geographic, temperature, wind, and precipitation parameters. Table 2, represents the details of each input factor with hourly based data during one complete month.

Table 1 Geographic coordinate system for three cities.

Location	Latitude	Longitude
Erbil	36.199	44.0737
Baghdad	33.3793	44.4501
Basra	30.494	47.8398

Table 2 Corresponding input climate data.

Variable	Unit
Temperature	°C
Precipitation Total	Mm
Relative Humidity	%
Wind Speed	km/h
Wind Direction	°
Sunshine Duration	Min
Direct Shortwave Radiation	W/m²

3.2Standardizing and Normalizing data

This study is based on automated recognition method with combining and correlating wide range input values. This variety of input ranges will suppress the acceleration of such automated tools such as machine learning, this impact on the performance of the formulating model (Adel & Alessandro, A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, 2010). Initially, pre-processing data for such a huge range of inputs is the crucial step, accordingly Min -Max normalization tool are the most efficient way to improve input status, this will be speeding up the machine learning performance and time for having a positive influence in this study (Kevin & Paul, 2005). Min -Max normalization representing in Eq.1.

$$x = Target_{min} + (Target_{max} - Target_{min}) * \frac{x - x_{min}}{x_{max} - x_{min}}$$

$Target_{max}$, $Target_{min}$ is corresponding normalized value (which considered to be $Target_{max}$=1&$Target_{min}$ =0). While x v [x_{max}, x_{min}]. After completing the process of correlation, the de-normalizing should be applied for getting the real value.

3.3Forward neural network methodology

Artificial neural network (ANN) is the most popular branch of machine learning, it forms a mathematical model depending on historical data, this can be interpreted as a process that is taught from experience. ANN can represent the brain mechanism for analyzing data, improving the act of sophisticated model and take a precise approach. The relation between brain biological structure and the mathematical statement is developed tightly. Neurons (As data input) address within initial layer that interconnected to subsequent nerves (intermediate layer or Hidden layer) by a specific strengthen link. The rate of strength (weight) will be changed upon learning epoch. The relative nerves are triggering neighborhood nerves until they combined to their eventual destination (output) to perform action (Michael, 2011). Figure 1 show the platform of Simulink equivalent model and each of initial weight, bias, activation function can be corresponding as a block in MATLAB. Eq 2, explains a simple mathematical expression for ANN excluding posting hidden layer. Activation functions transform the input data to more reasonable range.

$$Y = Sigmoid\left(\sum_{n=1}^{m} Input_n * Weight_n - Bias_n\right)$$

$$Simoid(A) = \frac{1}{1 + e^{-A}}$$

3.4Optimizing model through iteration and prediction performance evaluation

Posting hidden layer between input and output layer is to characterized input, increase a chance of manipulating the weights and corelated with the actual behavior (Michael, 2011). Serval statistical tools are commonly used for evaluating prediction deployment, Eq 3 and Eq 4 are Root mean square error (RMSE) and correlation coefficient (R) consequently are the most popular indicators.

$$RMSE = \sqrt{\frac{\sum_{n=1}^{m}(Y_{predicted} - Y_{Desired})^2}{m}}$$

$$R = \frac{\sum_{n=1}^{m}(Y_{Desired(n)} - \bar{Y}_{Desired})(Y_{predicted(n)} - \bar{Y}_{predicted})}{\sqrt{\sum_{n=1}^{m}(Y_{Desired(n)} - \bar{Y}_{Desired})^2 \cdot \sum_{n=1}^{m}(Y_{predicted(n)} - \bar{Y}_{predicted})^2}}$$

Within FNN algorithm, data can be grouped into three sets randomly: trained (70% of data), validation (15% of data) and test (15% of data). Selecting optimum number of neurons in hidden layer has a direct influence of FNN. In this study, sets of neurons in hidden layer sequentially iterated and used for indicating the minimum RMSE between trained and validated sets (Jinchuan & Xinzhe, 2008).

4.Result and discussion

This study is performed by applying Matlab (Toolbox, 2019) and used neural network pattern. The input datasets, that including seven inputs, are recorded for one months over three location Erbil, Baghdad, and Basra. It is lunched with normalizing input through substitute it within Eq 1 and logarithmic the output to ensure that they have same weight and to accelerate the training process. The inputs transformation goes between zero and one. High commonly algorithm is Levenberg-Marquardt backpropagation that considered as recommended option to model prediction problem and used it in this study. 762 inputs dataset are gotten for each input, they are randomly set and partition into a specific group such as: choosing 70% as training set and 30% as validation set. The aim is to make the model in better condition. We have chosen the right number of hidden layers by training the model with different number of neurons (1 to 60 neurons). The optimum number of neurons were selected regarding to their matching of the least RMSE record between training and validation test. A very reasonable RMSE had gotten (between 2.5 – 3) and obtained among the three sites. As they are illustrated in Figure 2. After calculating the training coefficient regarding to each location, this model can be used easily for tracking the irradiation for each separated zone. It is notice that the greater number of neuron usage can't guarantee obtaining the minimum number of RMSE for training and validation set, thus it can't generalize well.

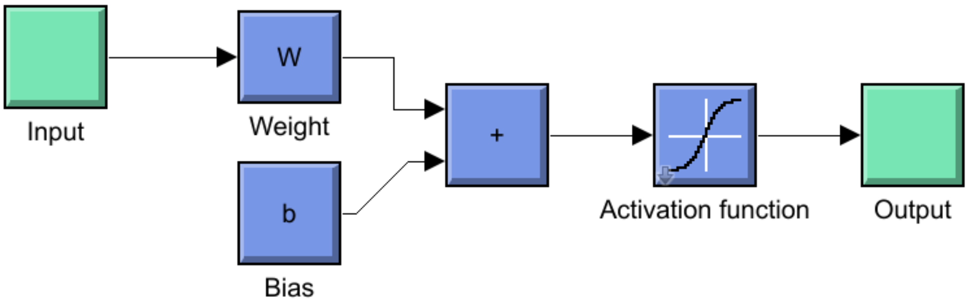
Figure 3 is plotting 70% of the predicted and desired output. A satisfied result with little of variance occurred between them and it is never far away from ideal case that represent as actual output value.

As shown in Figure 4, the statistical performance between estimated and desired value is presented. Since the output was normalized by logarithm their value, the range is transformed between 0 to 6.4. this transformation doesn't not change the property of data. The plotted regression (R) was greater than 0.94 for the three sits, that give the fit reasonably accurate model between actual and the desired value. The training and validation set both are well fitted.

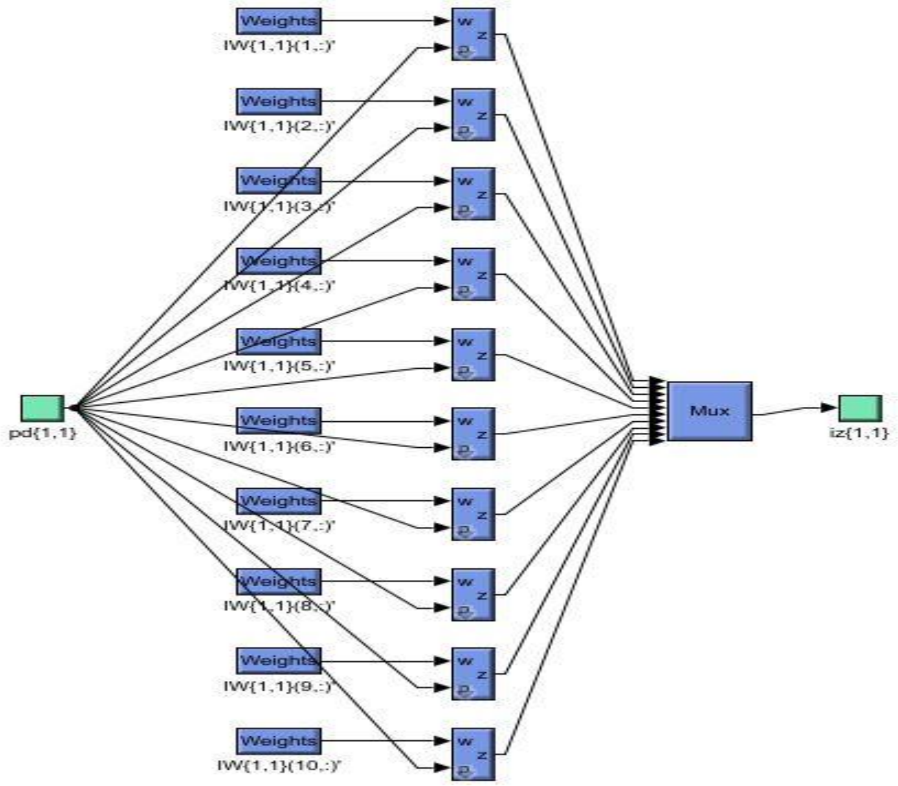
For evaluating neural network design and analyzing the model performance, as the dataset are randomly split into validation and trained groups, selecting optimum iteration point that can be appeared between underfitting and overfitting is need. Figure 5 shows the best validation performance which considered as a least Mean Square error over the three modeled cities that resulted between 0.8-1.4.

1 Conclusion

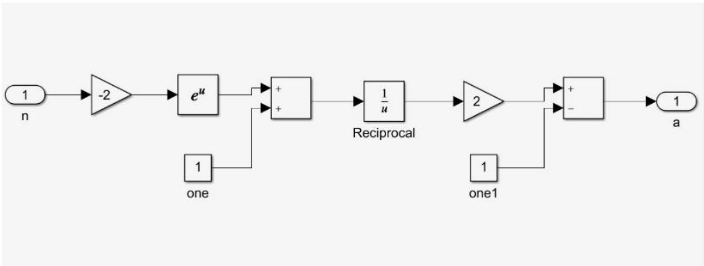
As a conclusion, hourly climate data are deployed to predict hourly irradiance in high accuracy range. The better performance for the model coming out, in a descending sort, Erbil and Basra then followed by Baghdad. Substituting the neuron in a sequential iteration in the hidden layer is an effective tool for generalizing the model between training and validation set. Increasing number of neurons is add a complicity of data and can't improve the fitness between training and validation set. It is realized that the neuron increment has a slight impact on making the model overfitted. The agreement between training and validation set in a minimum RMSE resulted for almost all the data indexes.



(a)



(b)



(c)

Figure 1 (a) Simple artificial neural network structure, (b) Generated simulink for one of the layer, (c) block for activation function.

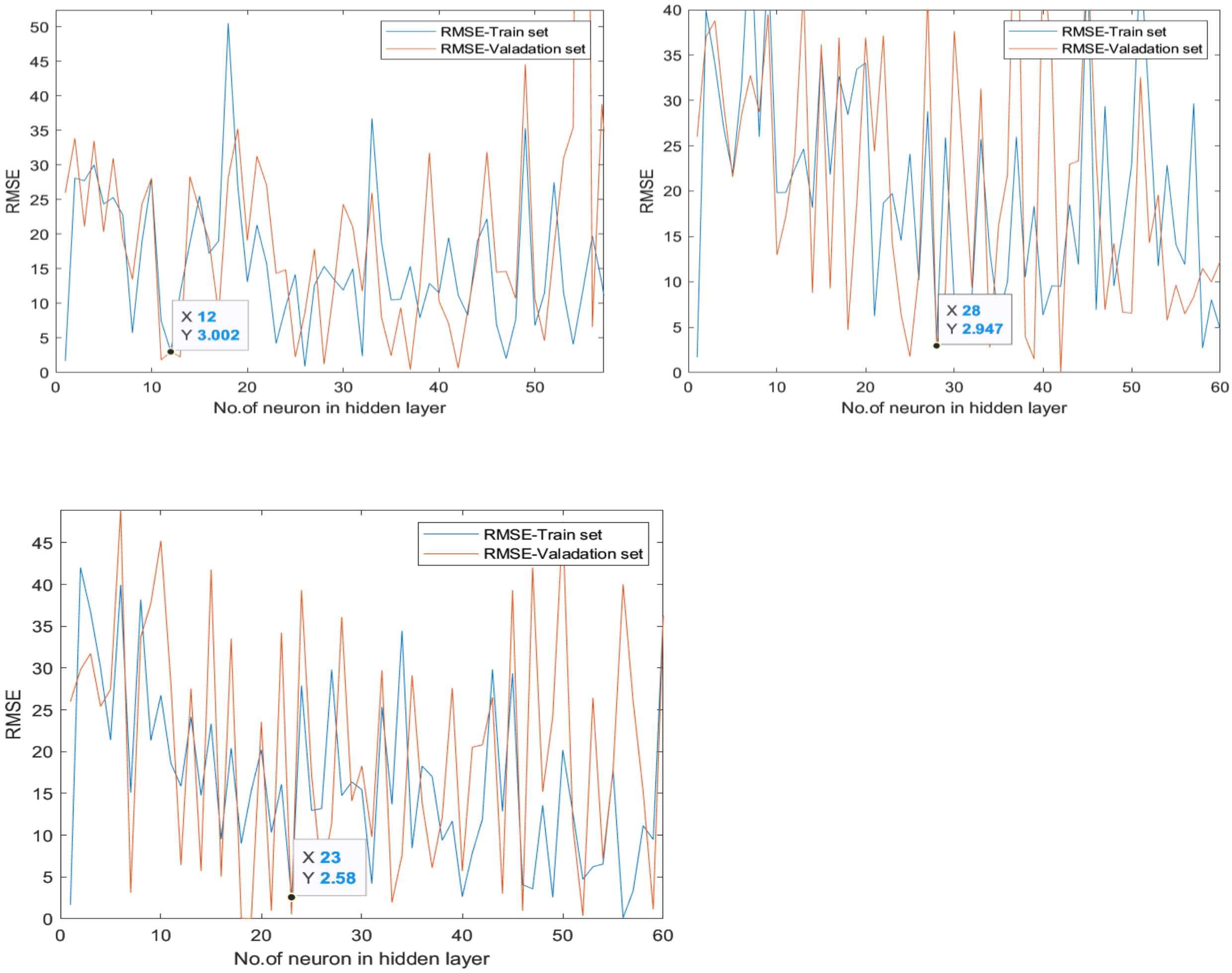


Figure 2 Different neurons in hidden layers Vs RMSE (For Erbil, Baghdad, and Basra)

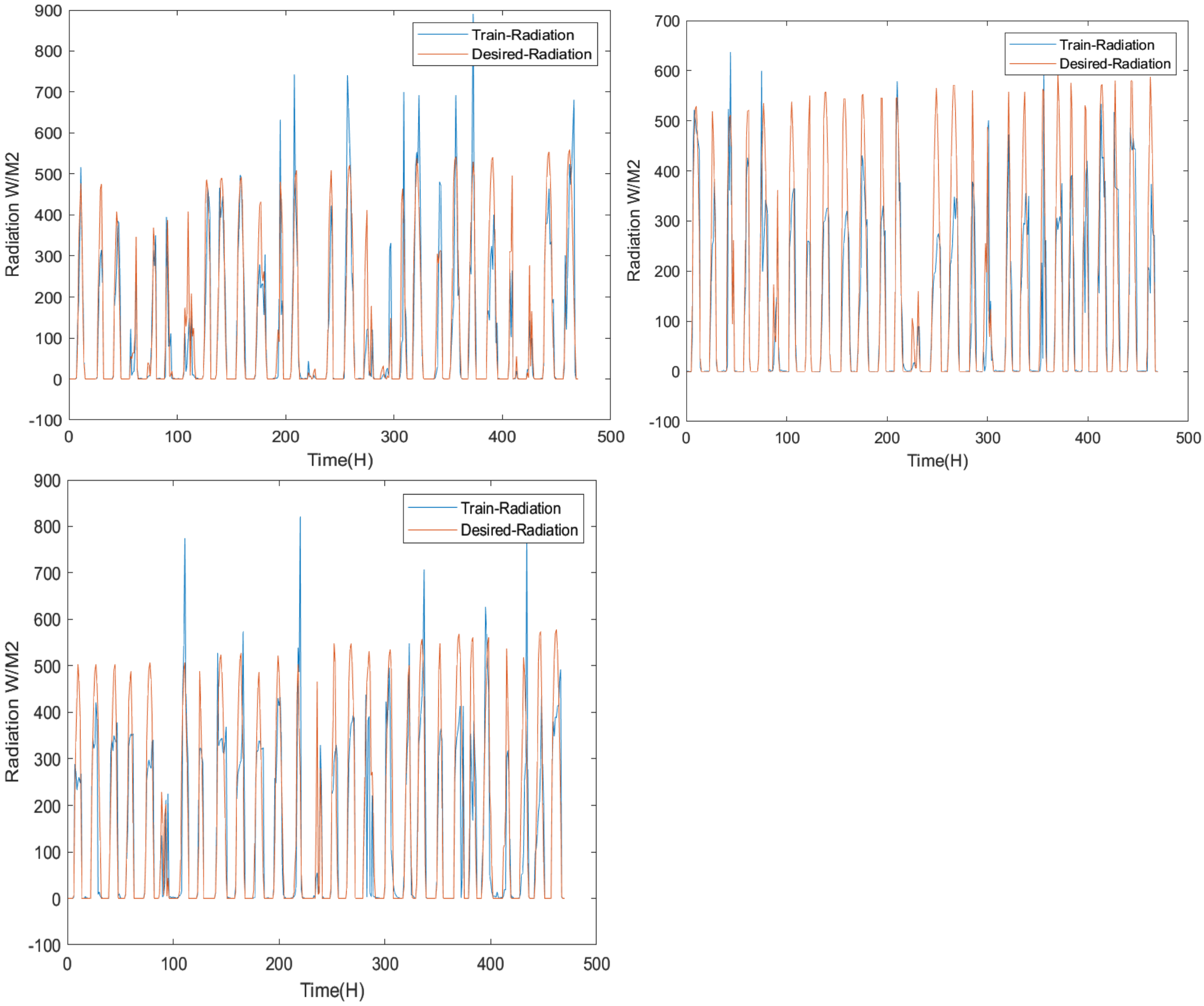


Figure 3 Actual Vs Predicted Irraidaince (For Erbil, Baghdad, and Basra)

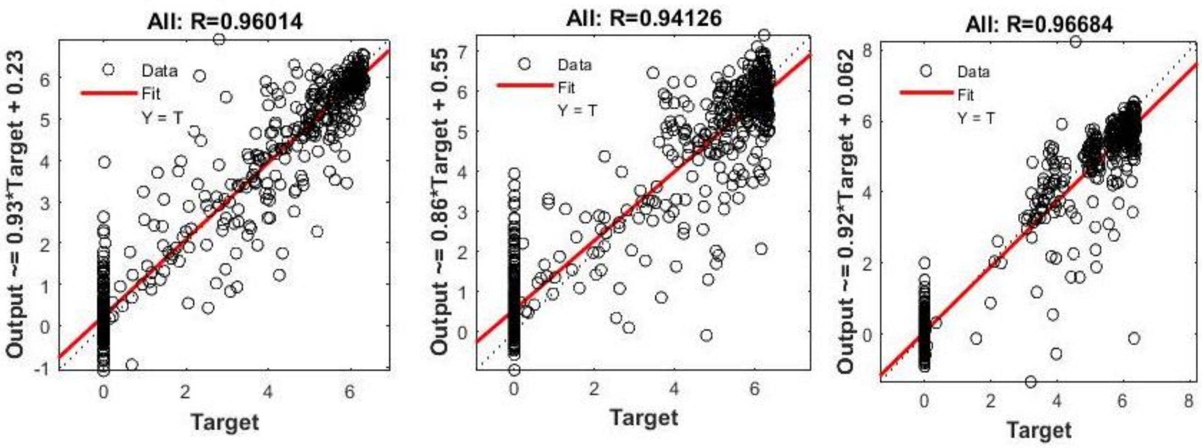


Figure 4 correlation coefficient for normalized actual Vs predicted output (Erbil, Baghdad, and Basra).

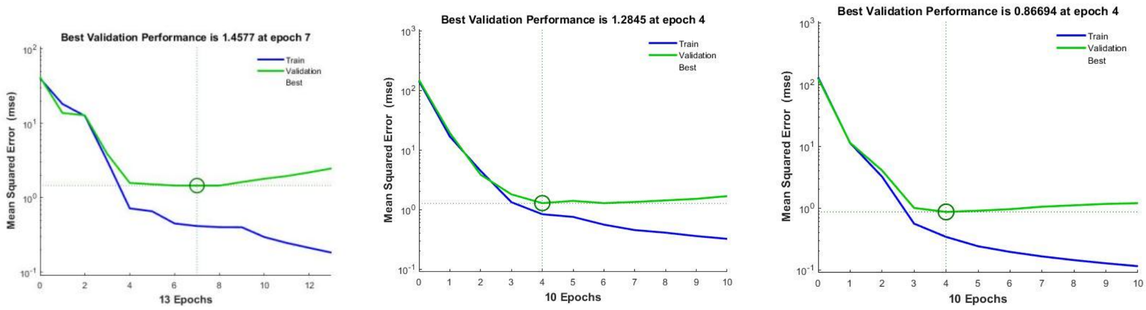


Figure 5 Mean square error Vs iteration with best validation performance (Erbil, Baghdad, and Basra).

Bibliography

Adel, M. (2008). Artificial Intelligence technique for modelling and forecasting of solar radiation data: a review. *Int. J. Artificial Intelligence and Soft Computing*.

Adel, M., & Alessandro, M. P. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste. *ELSEVIER*.

Adel, M., & Soteris, A. K. (2008). Artificial intelligence techniques for photovoltaic applications: A review. *ELSEVIER*.

Alzahrani, A., Kimball, J., & Dagli, C. (2014). Predicting Solar Irradiance Using Time Series Neural Networks. *ELSEVIER*.

Chiteka, K., & Enweremadu, C. (2016). Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks. *ELSEVIER*.

David, E., Gerrit, H., & R.W., M. (1994). Development of a neural network model to predict daily solar radiation. *ELSEVIER*.

Jinchuan, K., & Xinzhe, L. (2008). Empirical Analysis of Optimal Hidden Neurons in Neural Network Modeling for Stock Prediction. *IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*.

Kais, J. A.-J., Munya, F. A.-Z., & Zahraa, S. M. (2012). Estimation of clear sky hourly global solar radiation in Iraq. *INTERNATIONAL JOURNAL OF ENERGY AND ENVIRONMENT*.

Kevin, L. P., & Paul, E. K. (2005). *Artificial neural networks : an introduction.* The Society of Photo-Optical Instrumentation Engineers.

meteoblue. (2006). Retrieved from https://www.meteoblue.com/en/weather/week/basra_iraq_99532

Michael, N. (2011). *Artificial Intelligence: A Guide to Intelligent Systems.* Pearson Education Canada.

Mubiru, J., & Banda, E. (2007). Estimation of monthly average daily global solar irradiation. *elsevier*.

Negnevitsky, M. (2002). *Artificail intelligence a guide of intelligent system.* Pearson.

S., R., & M., M. (2009). Estimation of Diffuse Fraction of Global Solar Radiation Using Artificial Neural Networks. *Taylor & Francis Group, LLC*.

Toolbox, S. M. (2019). *MATLAB*. Retrieved from Mathworks Inc.