Research on Engineering Structures & Materials
www.jresm.org
MIM Research Group

Technical Note

The effect of increasing hidden layers on the performance of the deep neural network: Modelling, investigation, and evaluation

Abdel-Nasser Sharkawy *,a

Mechanical Eng. Dept., Faculty of Eng., South Valley University, Qena 83523, Egypt
Mechanical Eng. Dept., College of Eng., Fahad Bin Sultan University, Tabuk 47721, Saudi Arabia

Article Info

Article History:

Received 09 Sep 2024
Accepted 27 Nov 2024

Keywords:

Neural network performance;
Hidden layers;
Design and training processes;
Mean squared error;
Solar power station;
Investigation;
Assessment;
Comparison

Abstract

Neural networks have a profound impact on many real-life applications. In this paper, the influence of increasing the hidden layers of the neural network on its performance is investigated and presented. For this purpose, three structures of the neural network are designed. The first structure has only one hidden layer, the second structure has two hidden layers, and the third structure has three hidden layers. The inputs and the output of each neural network structure are the same. To design and train these structures, data are collected from a solar power station in Egypt. These data include the temperature of the solar photovoltaic module and the radiation which are the inputs of each neural network structure. In addition, the power of the photovoltaic module, which is the output of each neural network structure. The obtained data is 7200 samples and is divided into three different parts, the largest part for the structure train stage, part for the test stage, and the last part for the validating stage. The main aim of this division is to investigate the efficiency of the structure in different modes. The training of each structure is conducted by Levenberg-Marquardt technique. The mean squared error (MSE) value is the main parameter used to identify the completeness and the effectiveness of the train, test, and validating stages. In addition, the approximated error between the actual output and the predicted outputs by each neural network structure is calculated. Structures with four, five, and six hidden layers are also developed and investigated. The results show that the MSE value is decreasing with the increase of the hidden layers. The MSE values obtained using three, four, five, and six hidden layers are 0.01686, 0.01634, 0.01593, and 0.01586 respectively. Furthermore, the average value of the approximation error is very small and is 0.0396 using the three hidden layers. Therefore, the increase of the hidden layers of the neural network increases its accuracy and performance. The results of the proposed method are compared with previous related works from literature. The result of this comparison shows the superior accuracy of the proposed method.

1. Introduction

Neural network is a powerful mathematical technique, and nowadays is widely used with many real applications such as in industry [1, 2], robotics [3], medicine [4], finance [5, 6], renewable energy [7, 8], and so on. It can be used with prediction or estimation problems, automatic control processes, and classifications or pattern recognitions, social media, and weather forecasting, [9]. Neural network has desirable advantages such as its effectiveness in any linear/nonlinear function approximation, adaptivity, and ability for generalization [10-12]. Different types of neural

*Corresponding author: abdelnassersharkawy@eng.svu.edu.eg
[a]https://orcid.org/0000-0001-9733-221X
DOI: http://dx.doi.org/10.17515/resm2024.442st0909tn

networks are found such as multilayer perceptron, feedforward, recurrent, cascaded forward, radial basis function, and NARX network, [12, 13]. In the current paper, we shed light on the estimation or prediction processes particularly the PV solar station output using the neural network.

Neural networks are widely used by previous researchers in the prediction or estimation of parameters in different fields. Neural networks are used for prediction of the power of wind turbine. Liu et al [14] structured a neural network for predicting and estimating the power of a wind plant. Their structure was built based on the speed and the direction of wind which were used as the neural network inputs. Their structure used only one hidden layer. For performance assessment, they used the value of root mean squared error which was in range between 14.99 to 17.086. In [15], Koroglu and Ekici developed a neural network to estimate the speed and power of the wind plant. Their structure was developed based on twelve inputs such as temperature, pressure, radiation, cloudiness, sunshine hours, and direction of wind. Their structure was implemented using only one hidden layer. For performance investigating, the value of mean squared error was used and was in the range of 0.3653 and 0.4721. Neural networks are used with heat exchangers. Kouidri et al. [16] developed a neural network to estimate the fouling resistance in a heat exchanger. Their structure was built based on six inputs such as the temperature and mass flux. They used only one hidden layer in their structure. For performance investigation, the normalized value of the mean squared error value was considered, and it was in the range of 0.032591 to 0.009694. Neural networks are used with the field of batteries. Teixeira et al. [17] structured a neural network for estimating the health state of batteries of lithium-ion. The structure was based on a vector of temperature, current, and voltage. For performance investigation, the mean squared error value is considered, and it was from 0.0376 to 0.1278.

Neural networks are used to predict the power of a solar PV station which is the related topic to our current paper. In [18], Barrera et al. developed a neural network to estimate the energy of a solar module. The inputs of their system were three factors based on radiation. Their neural network structure had only one hidden layer. For investigating their neural network performance, they depended on the mean squared error value which was 0.040. In [19], Gumar and Demir developed a neural network to estimate the power of solar module. Their structures used nine inputs such as date, time, humidity, pressure, temperature, and radiation. In their structure, they used only one hidden layer. For performance investigation, they used the mean squared error value which was 0.4607. Sharkawy et al. [20] structured three types of the neural network for estimating the power of solar module. Their structures were based on the radiation and the temperature of the solar module which were the inputs of the neural network structures. In their structures, they only used one hidden layer. For the performance investigation, they depended on the mean squared error values which are 0.275, 0.612, and 0.887. In [21], Azka et al. used a neural network with two hidden layers to estimate and model the power of a solar PV station. In their system, they used seven inputs which were temperature, pressure, humidity, precipitation, wind speed, cloud cover, and solar attitude. They also depended on the mean squared error value to evaluate their system which was a high value. Jinyeong Oh et al. [22] proposed a deep neural network to forecast the power of a solar PV station. Their method was implemented by using from two to ten hidden layers. They used 22 input variables based on stamps of time and conditions of weather such as temperature, humidity, the speed of wind, and so on. The best mean squared value that obtained with their approach was 0.0408. In [23], Sulaiman and Mustaffa developed a deep neural network-based approach for predicting the generation of the power of solar PV. In their approach, two hidden layers were used, and they used 11 inputs based on temperature, irradiation, and time. Their obtained best mean squared value was 0.3744. Amer et al. [24] developed a neural network with two hidden layers for the estimation of the output power of the solar plant. In the design of their system, they used 7 inputs based on features of irradiation, temperature, and the speed of wind. Their achieved mean squared error was 4.094. In [25], a neural network structure was proposed and trained with two different learning algorithms which were error back-propagation and Levenberg-Marquardt. They used this structure for estimation of the power of the solar PV by considering only hidden layer and two inputs which were the temperature and radiation. Their obtained mean squared error values were 0.0238 and 0.03482. Rushdi et al. [26] developed a neural network with two hidden layers to estimate the power of solar-wind system. Their neural network was designed with many sets of features based on temperature, radiation, wind speed, and wind direction. The obtained mean squared error value was 0.01986.

From this discussion, the main challenge is to consider many hidden layers in the structure of the neural network to investigate its performance and accuracy. In addition, using the value of the mean squared error or the root mean squared error is recommended to assess the training and testing processes of the neural network. Using a few numbers of inputs can minimize the complexity of the neural network and the computations.

The main contribution and novelty of this paper can be presented in the following points:

- The effect of increasing the number of hidden layers in the structure of a neural network is investigated and analyzed.
- For this purpose, a neural network is designed with one hidden layer (NN1), two hidden layers (NN2), and three hidden layers (NN3). In addition, four, five, and six hidden layers were used. In each case, the performance and the accuracy of the neural network structure is determined and compared.
- For investigating the performance/accuracy of the neural network structure, the value of the mean squared error, and the approximation error are used. To reach very small values, many experiments and trials are conducted. Furthermore, very small size of inputs (two inputs) is considered.
- To design and train these structures, data are obtained from solar power plant in Egypt and these data are the temperature of the solar module and the radiation which are the inputs of each neural network structure. In addition, the power of the solar module, which is the output of each structure. The correlation between these variables is analyzed and drawn.
- For high investigation of the neural network performance and efficiency, the obtained data is divided into three parts: one part for the train case and the other two parts for the test and validating cases. After that, all these data are combined and used again to test the neural network structure.
- The process of training is conducted using the algorithm of Levenberg-Marquardt which is a very fast technique and achieves the desired performance easily.
- The experimental results reveal that increasing the number of hidden layers can effectively improve the performance and accuracy of neural networks. However, the time of training is increasing. For offline training, the time of training is not important.

The methodology conducted in the current paper is presented in Fig. 1. This methodology is discussed as follows:

- Six structures of neural networks are designed and developed by determining their inputs and output. Considering a plant of solar PV, the selected inputs are the temperature of the solar module and the radiation, whereas the output of each neural network structure is the output power.
- Data is collected from a solar PV plant in Egypt and the number of samples is 7200. These samples are divided into three parts. 5040 samples are used for the NN structure train, 1080 samples are used for the NN structure test, and 1080 samples which are used for the NN structure validating.
- In every stage, whether in training, testing, or validation, the mean squared error (MSE) is calculated. When the obtained MSE becomes very small value and close to zero, the training of the neural network structure is then finished.
- The MSE value of each NN structure is compared to determine which NN structure has the best performance and accuracy.
- Once the training process is completed, all data is combined again and the NN structure is tested, and the approximation error value is calculated. If this error is a very small value and close to zero, then the trained neural network structure is efficient in estimating the output correctly.
- The approximation error value resulted by each NN structure is also compared to determine the most preferable and accurate structure.

The rest parts of this paper are discussed as follows. Section 2 shows the design of the neural networks' structures using one, two, and three hidden layers. In section 3, the stages of the train, test, and validating are presented in detail. The mean squared error value for each stage and each structure is presented and compared. Section 4 shows a comparison between the actual output and the estimated outputs by the neural network structures. In addition, the approximated output is determined. Section 5 shows the use of four, five, and six hidden layers in brief. We do not concentrate on this part as their results are remarkably close to the neural network with three hidden layers. Section 6 compares the results from the current study with literature results. In section 7, the main keys of this paper are summarized and some recommendations for the future.

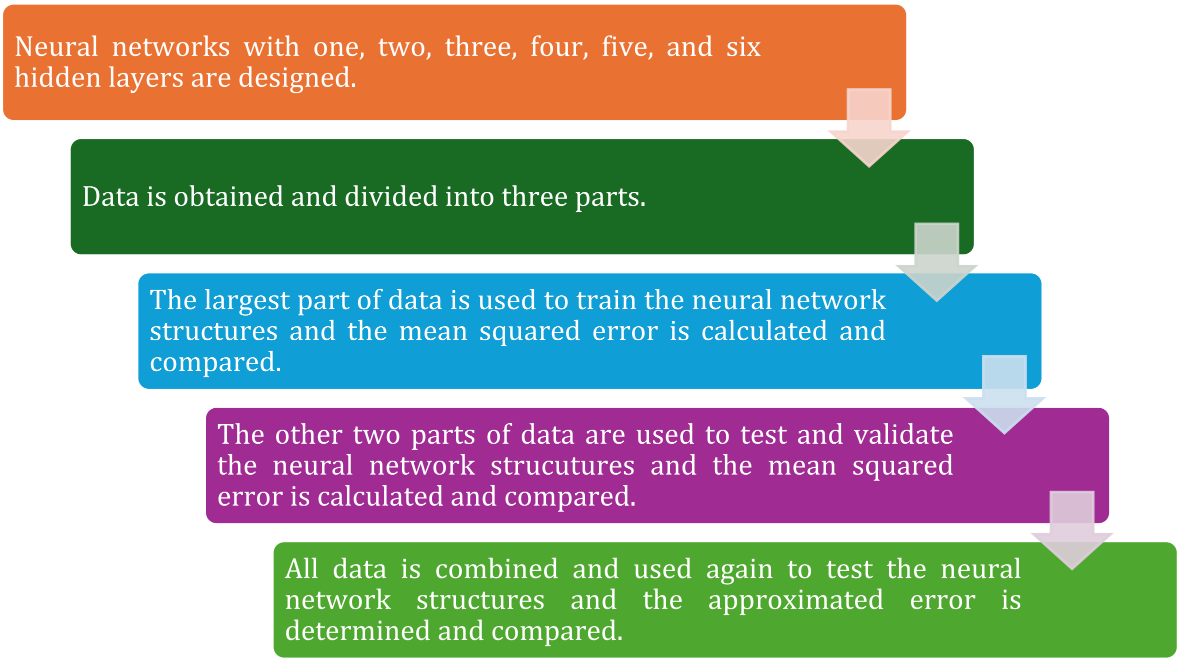


Fig. 1. The conducted methodology in the current paper

2. Structures' Design

In this section, the design of the proposed different structures of the NNs is discussed. To design and train and assess these structures, data from solar power station in Egypt is obtained. These data are obtained and available at the website [27]. This data is for Five days (7200 samples for 7200 minutes). These data are the temperature of the PV module and the radiation which are considered the inputs of the NN structure. In addition, the actual output power of the PV module, which is the output of the NN structure and used for the training process. This data is presented in Fig. 2. As shown, the signal of one day is close to the signal of another day because the data is collected from five consecutive days in summer and these days were approximately close together in conditions. Indeed, using different weather conditions is not important in our current study as we use the same data to train every neural network structure. In other meaning, the main aim of our study is to investigate the increase of hidden layers on the performance of the neural network. Therefore, it is recommended to use clear data without any outliers to show and clearly compare the performance of each structure.

Some analysis including the average, minimum, maximum, and standard deviation values about these data is presented in Table 1. The standard deviation is high whether for the input 1 (temperature), input 2 (radiation), or output (power). This means that the values are not close to each other. This is due to the fact that the measurements or the values of the variables during the day are different. This also can show the efficiency of the NN structures to work with different cases of data.

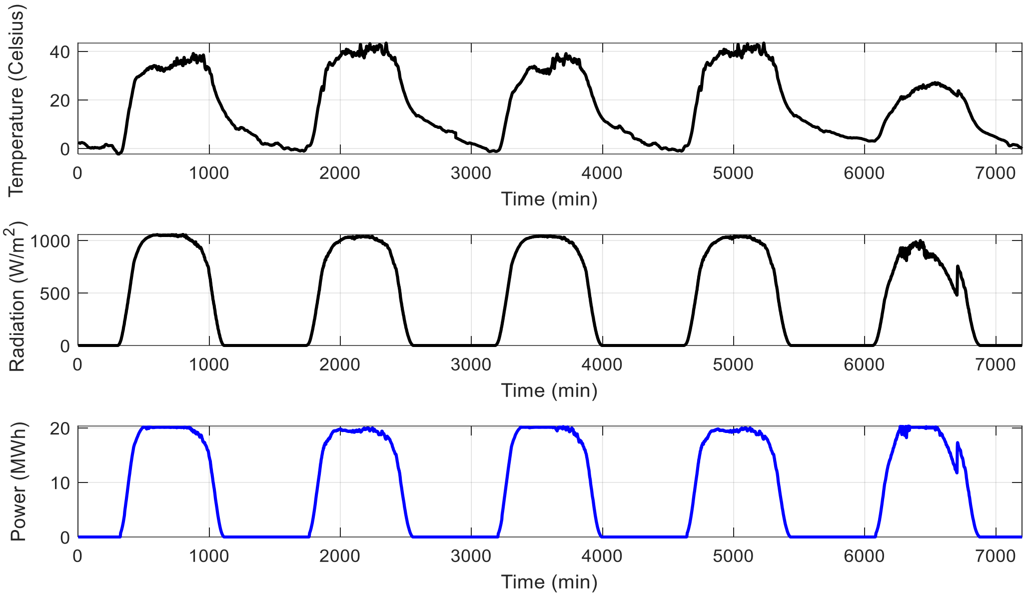


Fig. 2. The upper two diagrams are the inputs (temperature and radiation) for the designed NNs structures. The last diagram is the output (power) that is used for the train of the NNs structures

Table 1. Analysis for the inputs and the outputs used with the NN structures.

Parameter	Input 1 (Temperature - °C)	Input 2 (Radiation - W/m^2)	Output (Power - MWh)
Average value	17.7191	418.1952	8.5044
Maximum Value	43.3975	1056.8	20.3270
Minimum Value	−2.2750	0	0
Standard deviation	14.3713	448.5777	8.9727

The temperature and radiation are selected to be the inputs of the neural network structure due to these variables are mainly in effecting the power of the solar PV as shown from the following equation [28, 29]:

$$Power = \eta_s \tau_g \alpha_s RA[1 - \mu_s(T_s - T_r)] \tag{1}$$

where, η_s is the reference solar cell efficiency, τ_g is the transmissivity of glass, α_s is the absorptivity of the cell, R is the radiation in W/m^2 unit, A is the whole cell area in m^2, μ_s is the thermal factor or coefficient of the cell efficiency in $\%/°C$, T_s is the temperature of the solar PV in °C, and T_r is the referenced temperature in °C. From this equation, the temperature and radiation are mainly factors that affect the power.

The correlation (heatmap) between the input variables (temperature and radiation) and the output variable (power) is analyzed and drawn using MATLAB as shown in Fig. 3. This Figure shows that high correlation and relationship between the temperature and radiation with the solar power. The temperature has a positive correlation with the power which is 0.9236. This means the increase in temperature leads to an increase in power. Furthermore, this correlation is remarkably close to 1 which means the temperature and the power are very related. The radiation also has a positive correlation with the power which is 0.9955. This also means the increase in radiation leads to an increase in power. In addition, this correlation is remarkably close to 1 which means that the radiation and power are very related. The inputs' variables are also very related since the correlation between temperature and radiation is 0.9348 which is remarkably close to 1.

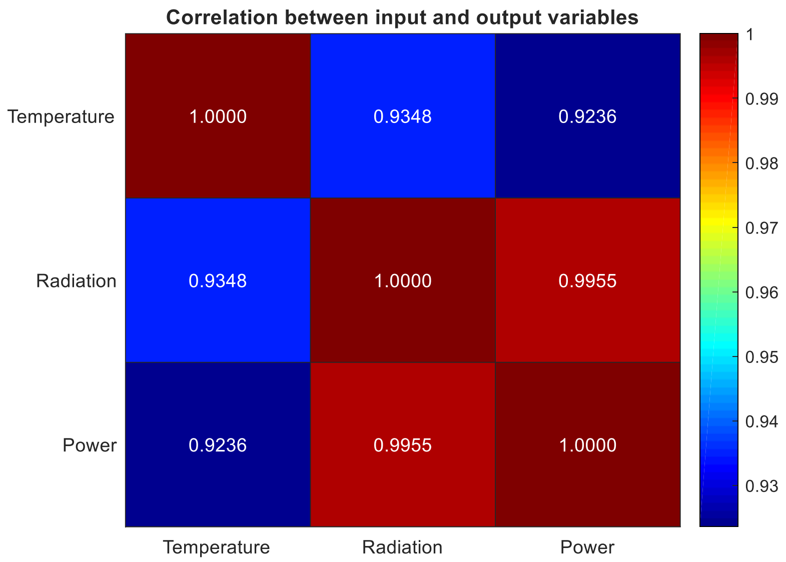


Fig. 3. The correlation (heatmap) between the input (temperature and radiation) and output (power) variables. This Figure is drawn using MATLAB

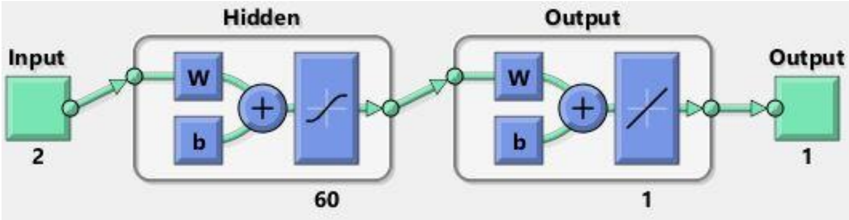
To investigate the effect of increasing the hidden layers of the NN architecture, three NN structures are designed: NN1(NN with one hidden layer), NN2 (NN with two hidden layers), and NN3 (NN with three hidden layers). The same inputs (temperature and radiation) and output (output power) are used with each NN structure. The specifications of each structure including the number of hidden layers, neurons for each layer, activation function for each layer, and inputs and outputs, are presented in Table 2.

Table 2. Comparison between the three structures of the NNs

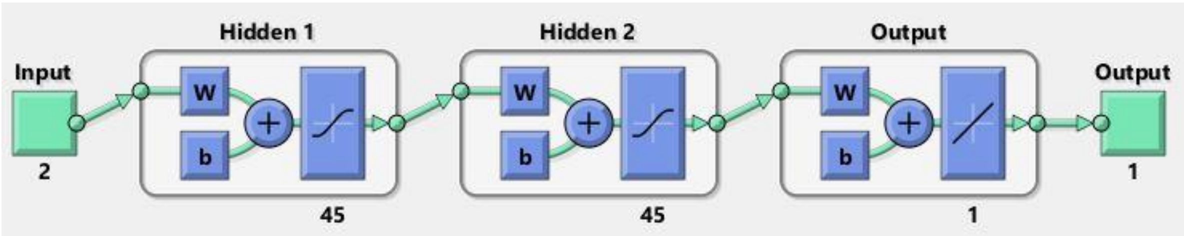
Parameter	NN1	NN2	NN3
Number of inputs	Two (temperature and radiation)	Two (temperature and radiation)	Two (temperature and radiation)
Number of hidden layers	One	Two	Three
Number of hidden neurons	60	First hidden layer: 45 Second hidden layer: 45	First hidden layer: 65 Second hidden layer: 65 Third hidden layer: 65
Activation function	Hidden layer: Tanh Output layer: Linear	First hidden layer: Tanh Second hidden layer: Tanh Output layer: Linear	First hidden layer: Tanh Second hidden layer: Tanh Third hidden layer: Tanh Output layer: Linear
Number of outputs	One (Power)	One (Power)	One (Power)

In Table 2, the number of neurons for each hidden layer (hidden neurons) are obtained after many experiments (trails and errors) until the highest performance of the NN is reached. The highest performance of the NN is the obtaining of the smallest mean squared error value or value close to zero. In addition, as the hidden layers are increased from NN1 to NN3, the complexity is increased.

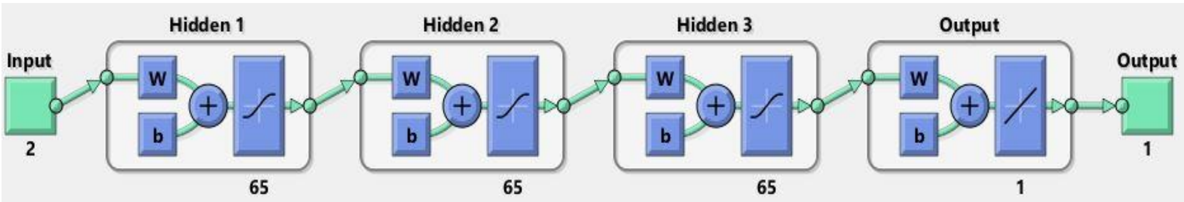
In Fig. 4, the shape of the designed structures is shown. These shapes are obtained from MATLAB software which is used in this work.



(a)



(b)



(c)

Fig. 4. The structure of the proposed NNs. (a) NN with one hidden-layer, (b) NN with two hidden-layers, and (c) NN with three-hidden layers

It should be noted that we use experimental (trial and error) methodology in determining the number of hidden layers. In other meaning, we use neural network structure with one, two, three, four, five, and six hidden layer and investigate the performance of each structure. The main aim of this methodology is to study the advantages and disadvantages of each structure to make it easy for future researchers to determine which structure can be used. The same methodology is followed for determining the number of neurons in each hidden layer. Many different neurons (from 2 to 120) are tried until the best performance of the structure is obtained. The best performance is achieving highest accuracy which is obtaining a very small value of mean squared error and approximation error. Indeed, this methodology is followed by previous researchers as seen in [30-35].

3. Structures' Training, Testing, and Validation

After the three NNs are structured, the second three stages are carried out which are the train and test and the validating. The obtained data mentioned in previous section is divided into three parts: the first part which is 5040 samples are used for the train, the second part is the 1080 samples which are used for the test, and the last part is the 1080 samples which are used for the validating.

Table 3. The division of the data for three stages (train, test, and validation) of the NNs structures.

Stage	Training	Testing	Validation	Total
Samples of Data	5040	1080	1080	7200

The division of these data is presented in Table 3. This division is also carried out randomly in MATLAB which is used in these stages. All these stages are conducted using processor of type

Intel(R) Core (TM) i5-8250U CPU @ 1.60GHz 1.80 GH and RAM of 8.00 GB. The crucial factor during these stages is obtaining the very small value of the mean squared error. The mean squared error value is calculated by the following equation:

$$Mean\ Squared\ Error\ Value = \sum_{i=1}^{n} \frac{(Actual\ output(i) - Predicted\ output(i))^2}{n} \tag{2}$$

where, n is the number of samples.

The training stage is carried out using the technique of Levenberg-Marquardt. This technique is a very fast and stable method and achieves the highest performance of the NN easily compared with other training algorithms, [36-38]. Many experiments which are trials and errors are executed to train the NN structures. These experiments include trying the use of many different hidden neurons and many different initializations of the weights until obtaining the very small value of the mean squared error. This methodology is followed by previous researchers as in ref. [32, 39]. The conducted methodology in the three stages (train, test, and validating) is shown in Fig. 5.

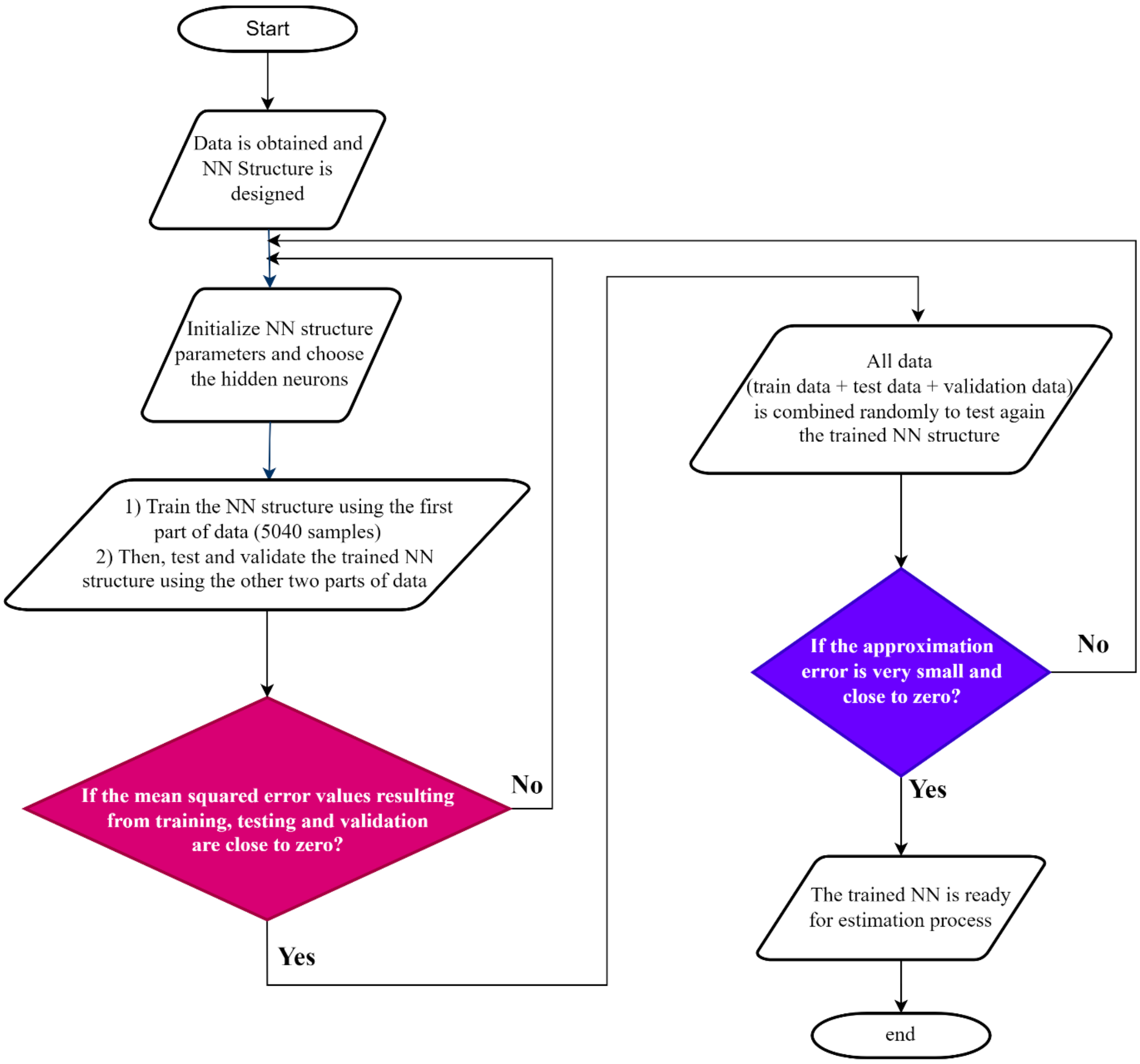
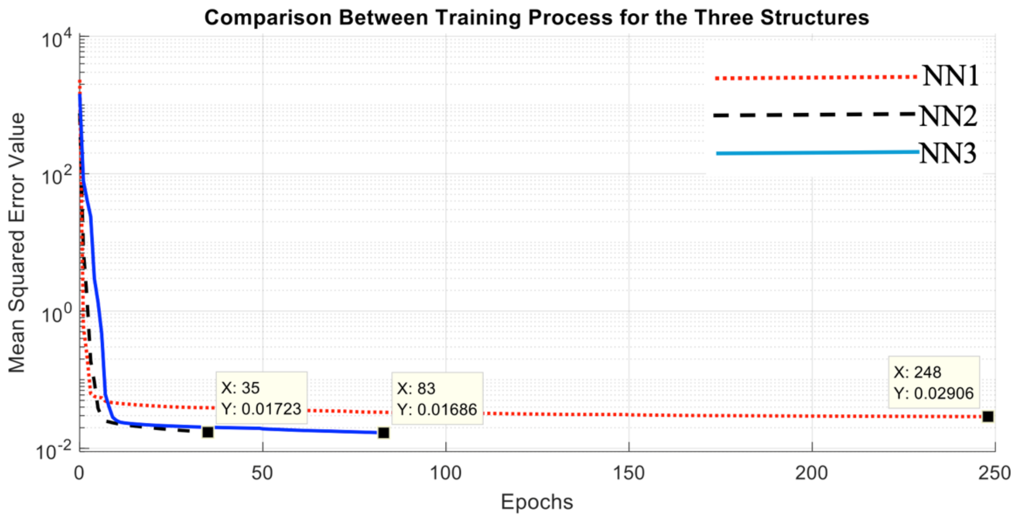
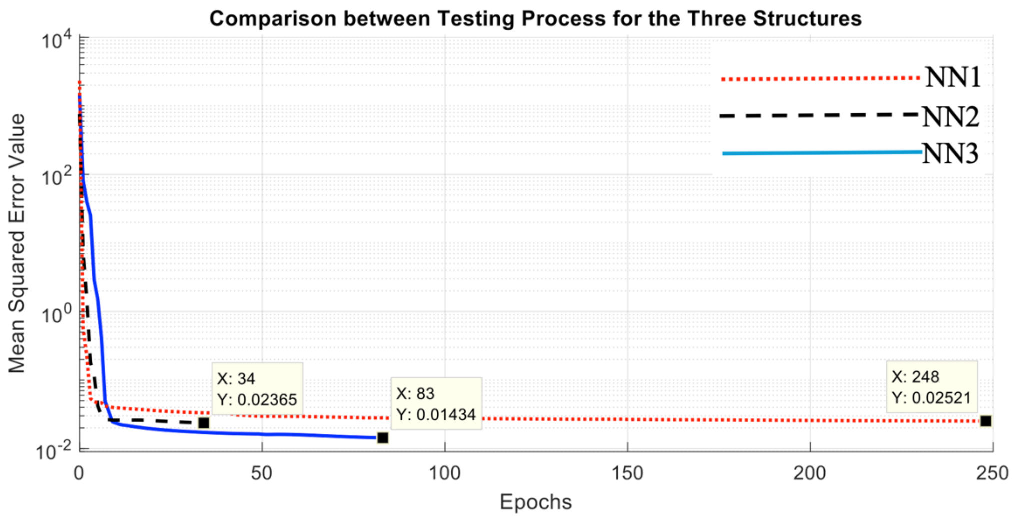


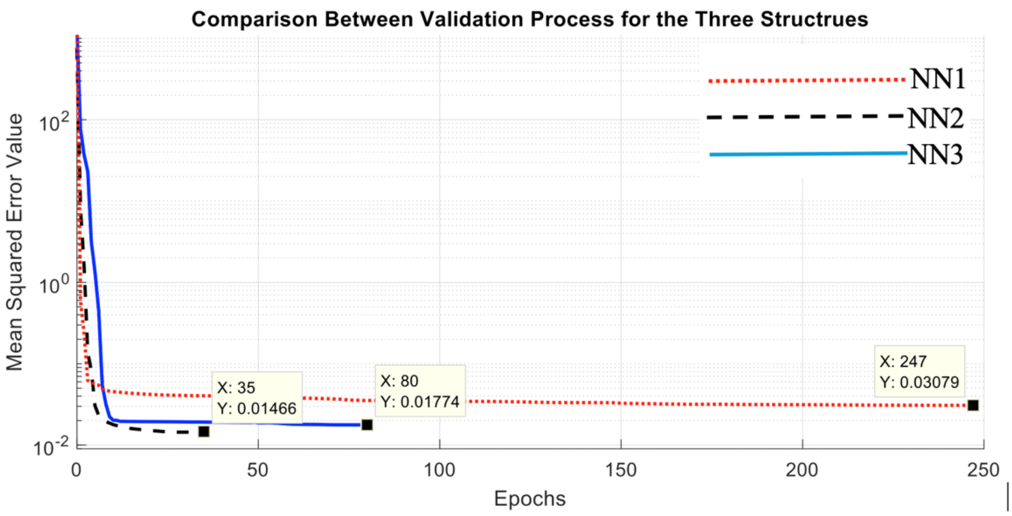
Fig. 5. A flowchart illustrating the conducted methodology for the train, test, and validating stages



(a)



(b)



(c)

Fig. 6. Comparing the mean squared error value deduced by the three structures of NNs in case of (a) training, (b) testing, and (c) validation

The mean squared error values obtained from the three stages (training, testing, and validation) are shown in Fig. 6. The mean squared error resulting from training the three structures is presented in Fig. 6(a). As shown in this Figure, the mean square error from the NN3 is the smallest one compared with the other structures. This means that the error in the case of the NN3 is the lowest one and its accuracy is the highest. Other parameters during the stage of training are illustrated in Table 4 such as the used epochs, the time of training, and the regression. The regression resulting from NN3 is the best compared with the other structures. This also supports the fact that NN3 has the best accuracy. However, the time used for training the structure of NN3 is the highest. This means that there is a need for longer time to train the NN3 compared with other structures. Indeed, this is not particularly important if offline training is executed. In the case of online training, longer time for training is not desired. The number of epochs used with NN3 is between NN1 and NN2. We conclude that increasing the number of hidden layers improves the performance of the NN but increases the time of training and complexity

Table 4. The parameters obtained from training, test, and validating of the three different NNs structures

Parameter	NN1	NN2	NN3
Method of training	Levenberg-Marquardt technique		
MSE from training	0.02906	0.01723	0.01686
MSE from testing	0.02521	0.02365	0.01434
MSE from validation	0.03079	0.01466	0.01774
Used Epochs in training	248	35	83
Time used in training	31 seconds	3 minutes and 20 seconds	28 minutes and 5 seconds
Regression	0.9972	0.9981	0.9998

The testing and validating processes are conducted simultaneously with the training case. At each epoch/iteration, once a NN structure is trained, it is tested and validated, and the mean squared error value is calculated. This methodology is repeated until the smallest mean squared error value is reached in the three stages (train, test, and validation) simultaneously. As mentioned in previous section that the data used in test and validating stages are different from the training case. This is very useful to investigate and assess the effectiveness and the efficiency of the trained NN structures. The mean squared error values deduced from the test and the validating processes are shown in Fig. 6(b) and (c) and Table 4. The mean squared value deduced by the NN3 during the testing stage is the best/lowest compared with the other structures. In the validating stage, the mean squared error by the NN3 is located between the other structures. However, it is also very small value. From these results, we deduce that all the NN structures are trained in an exceptionally good way and the mean squared error values deduced by them are small and near to zero. In addition, the results are improved by increasing the hidden layers. This is clear as the results obtained by NN2 are better than the corresponding ones by NN1, and the results deduced by NN3 are the best compared with NN1 and NN2. On the other hand, the complexity and the time of training is increasing with NN3.

4. Output Estimation and Prediction

In this section, another stage of NN structure testing is conducted. All obtained data (train data + test data + validating data) are combined randomly in MATLAB to test the three designed structures. Then, the output of each NN structure is determined and compared with the actual output that is used for the training stage. Figure 7 shows a comparison between the actual output and the predicted ones using the trained NN1, NN2, and NN3. As shown in the Figure, the predicted outputs by NNs coincide with the actual output. This illustrates that all NN structures are trained in a very well and effective way. To show which NN structure is the most accurate compared with the others, a zoomed area from Fig. 7 is presented in Fig. 8.

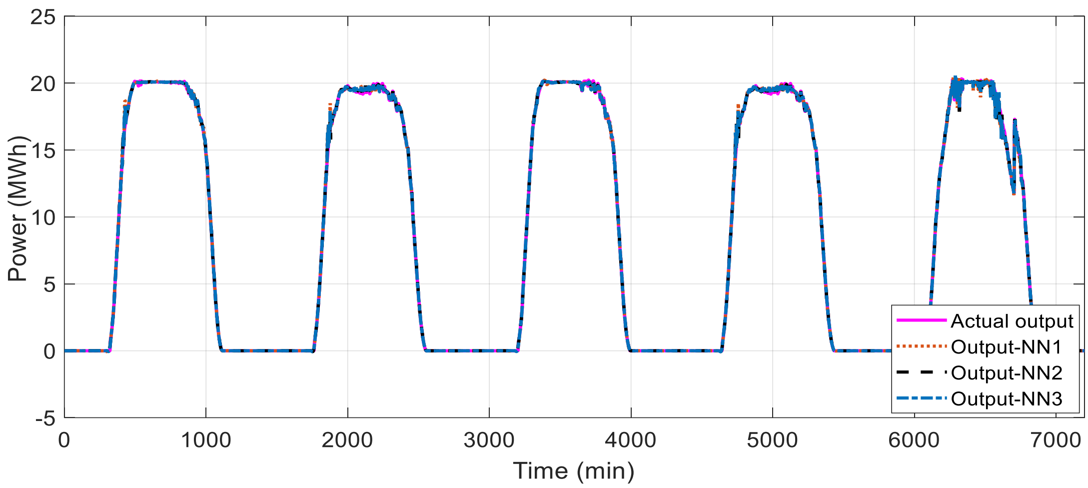
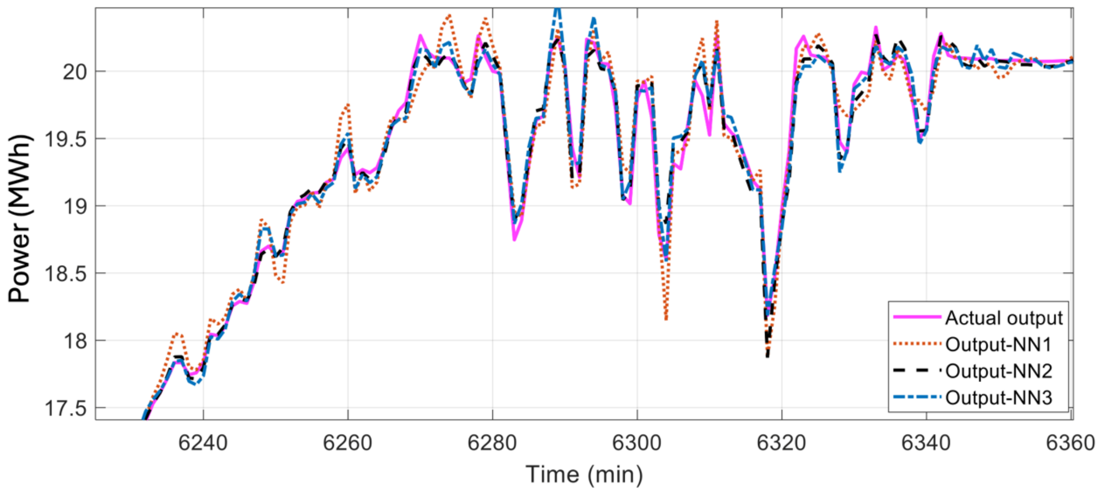
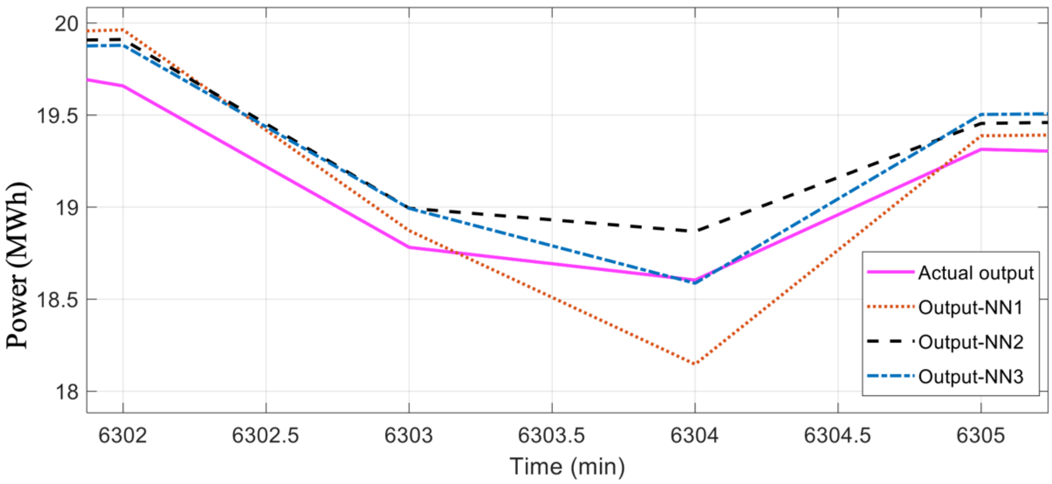


Fig. 7. The comparison between the actual output and the predicted outputs by NN1, NN2, and NN3. The colored curves are coinciding together



(a)



(b)

Fig. 8. Zoomed area from Fig. 7: the comparison between the actual output and the predicted outputs by NN1, NN2, and NN3. (a) Zoomed area during the 120 minutes between 6240 and 6360 minutes. (b) Zoomed area during the 3 minutes between 6302 and 6305 minutes

Zoomed area during 120 minutes between 6240 and 6360 minutes, as shown in Fig. 8(a) and zoomed area during 3 minutes between 6302 and 6305 minutes, as shown in Fig. 8(b). From Fig. 8, the estimated output by NN3 is the closest to the actual output and the estimated output by NN1 is the most distant to the actual output. This means that the NN3 is the most accurate compared with other structures and the NN1 is the lowest accurate. NN2's performance is better than NN1's performance and lower than NN3's performance. For more clarification, the error or the difference between the actual output and the predicted output by each structure is obtained. This error is calculated by the following equation:

$$Approximation\ Error = Actual\ output - Predicted\ output \tag{3}$$

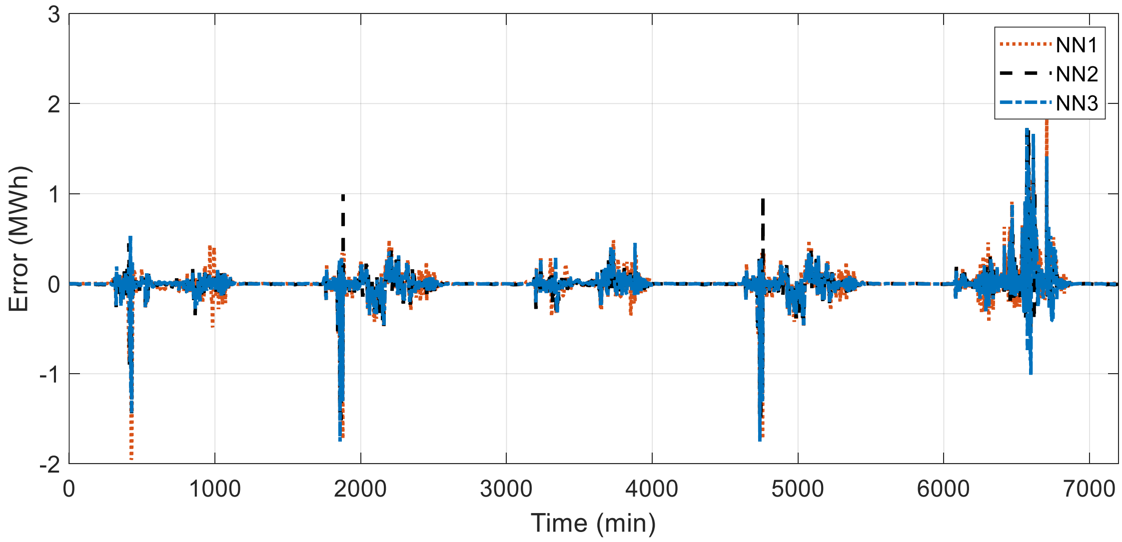
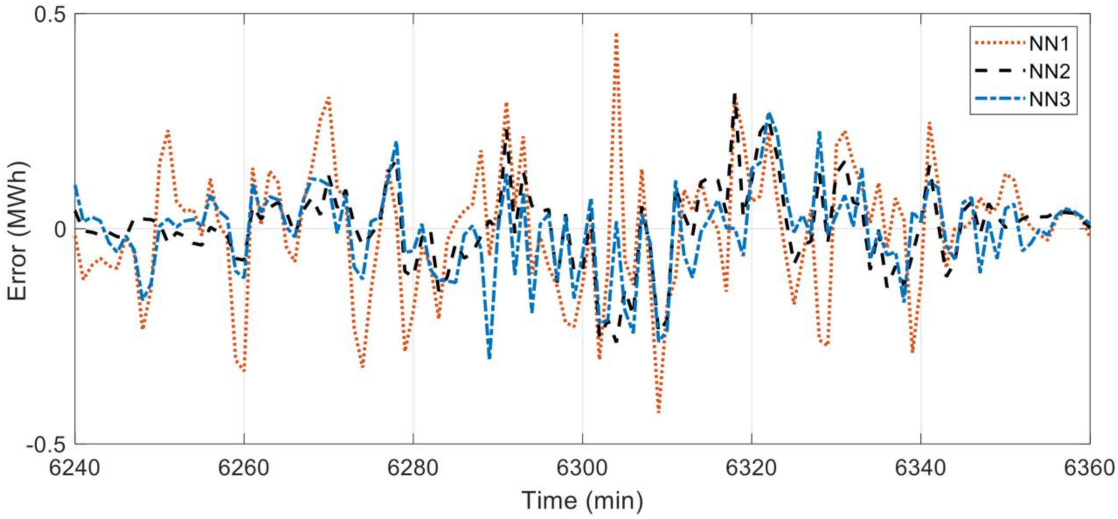


Fig. 9. The error or difference between the actual output and the predicted outputs by NN1, NN2, and NN3

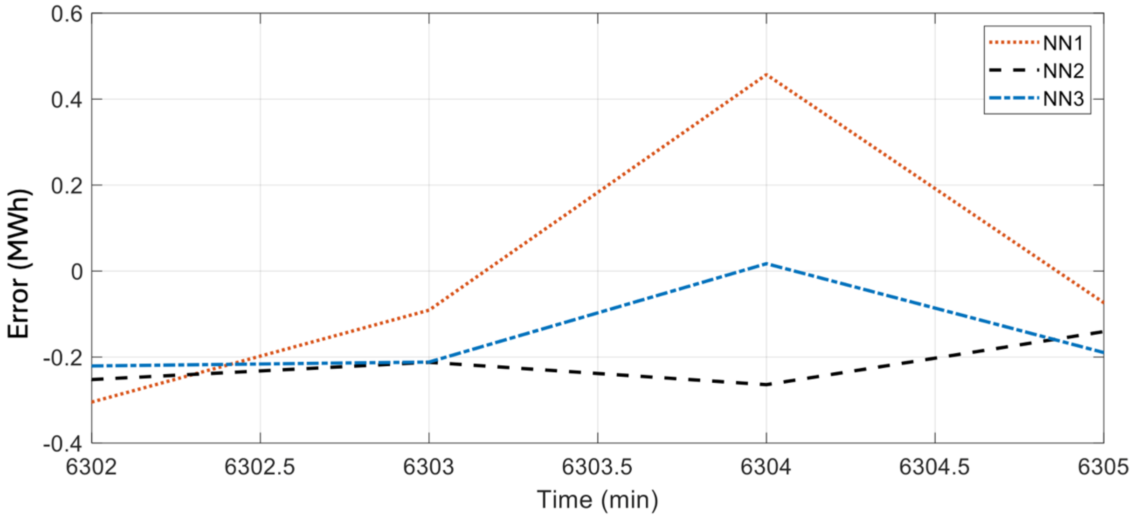
Figure 9 shows the errors resulting from the three NN structures and Fig. 10 shows zoomed area from Fig. 9. From Fig. 9, the errors resulting from all structures are satisfactory and close to zero. This supports that the NN structures are effectively trained. From Fig. 10, the error resulting from NN3 is the smallest and the closest to zero compared to the corresponding ones by NN1 and NN2. The error obtained by NN1 is the worst case. This reveals that NN3 is the most accurate and NN1 is the least reliable. Analysis about these errors is presented in Table 5 such as the average value, minimum value, maximum value, and standard deviation. The maximum value of the error for each structure is small and this is good and supports that the structures are effectively trained. In addition, the standard deviation value is small for each structure which reveals that the deviation between the values is small. From this discussion, we conclude that increasing the hidden layers for the NN structure minimizes the error between the predicted and actual outputs. Therefore, the accuracy or the performance of the NN structure is increasing.

Table 5. Analysis of the error resulted by the developed NN1, NN2, and NN3.

Parameter	NN1	NN2	NN3
Average Value	0.0681	0.0474	0.0396
Maximum Value	2.3124	1.7753	1.7513
Minimum Value	3.8724e-06	4.4804e-07	2.5668e-06
Standard Deviation	0.1553	0.1212	0.1235



(a)

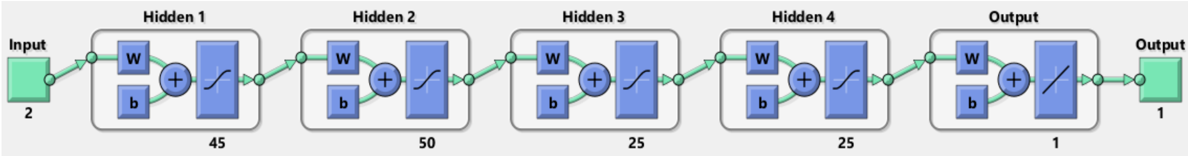


(b)

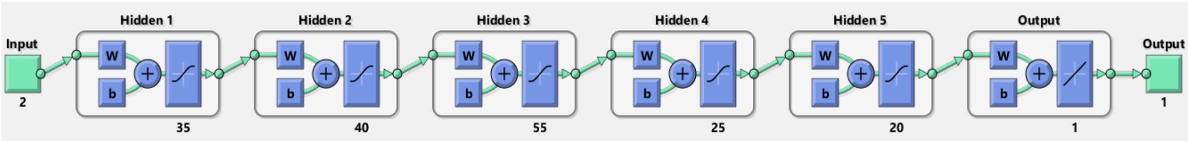
Fig. 10. Zoomed area from Fig. 9: the error or difference between the actual output and the predicted outputs by NN1, NN2, and NN3. (a) Zoomed area during the 120 minutes between 6240 and 6360 minutes. (b) Zoomed area during the 3 minutes between 6302 and 6305 minutes

5. Using More than Three Hidden Layers

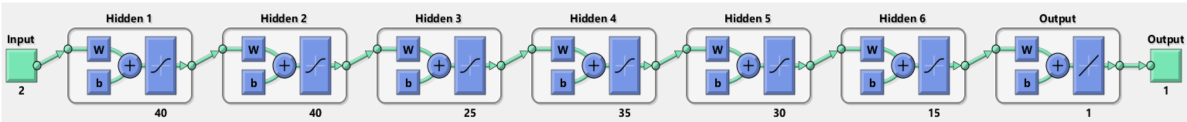
In the previous sections, we presented in detail only neural networks structures using from one hidden layer up to three hidden layers. Using more than three hidden layers (see Fig. 11) is presented briefly in this section. Structures with four, five, and six hidden layers are also designed, and the same data is used to train, test, and validate these structures. To train these structures, few trials are conducted because these structures consume longer time compared with previous structures. The best results including the number of hidden neurons in each layer, the time of training, and the mean squared error value obtained from these structures are presented in Table 6. It should be noted that we do not use a remarkably high number of hidden neurons in each layer to avoid the longer time of training and the more complexity. In addition, we do not have a powerful computer or processor to do the calculations more easily and in a short time.



(a)



(b)



(c)

Fig. 11. The design of the NN with (a) four hidden layers, (b) five hidden layers, and (c) six hidden layers

Table 6. Results from neural networks structure using four, five, and six hidden layers.

Parameter	NN with four hidden layers	NN with five hidden layers	NN with six hidden layers
Number of hidden neurons in each layer	First hidden layer: 45 Second hidden layer: 50 Third hidden layer: 25 Fourth hidden layer: 25	First hidden layer: 35 Second hidden layer: 40 Third hidden layer: 55 Fourth hidden layer: 25 Fifth hidden layer: 20	First hidden layer: 40 Second hidden layer: 40 Third hidden layer: 25 Fourth hidden layer: 35 Fifth hidden layer: 30 Sixth hidden layer: 15
Mean squared error value	0.01634	0.01593	0.01586
Time of training	50 minutes and 15 seconds	1 hour and 20 minutes and 31 seconds	2 hours and 22 seconds

As shown in Table 6, the mean squared error values obtained from these structures is less than the corresponding ones obtained by NN3 (neural network with three hidden layers). The neural network structure with six hidden layers achieves the smallest MSE compared with other structures which is 0.01586 but with the highest computational time which is 2 hours and 22 seconds. This means that increasing the hidden layers of the neural network leads to minimizing the mean squared error and therefore increasing the accuracy and performance. However, the time of training and the complexity are higher compared with NN3. Therefore, we recommend using the neural network structure with six hidden layers or more if the high performance is the only factor that needs to be achieved regardless of the computational time and complexity. For who need to achieve high performance with less complexity and time, we recommend using the neural network structure with three hidden layers which is considered a trade-off between performance and complexity.

6. Comparison With Related Previous Works

This section compares the results of the presented study with other previous and closely related works which used the neural network approach in estimating solar PV power. The main parameters considered in this comparison are the MSE value, the number of inputs (input size), and the used number of hidden layers. The previous related works considered in this comparison are as follows: Barrera et al., 2020 [18], Gumar and Demir, 2022 [19], Sharkawy et al., 2023 [20], Jinyeong Oh et al., 2024 [22], Sulaiman and Mustaffa, 2024 [23], Rushdi et al., 2024 [26]. The results from this comparison are shown from Fig. 12 to Fig. 14. For this comparison, the proposed neural network structure using three hidden layers which makes trade-off between the high performance and the complexity and the neural network with six hidden layers which gives the best performance are included.

Figure 12 compares the number of hidden layers used by the proposed approach and other previous ones. From Fig. 12, previous researchers used different numbers of hidden layers from one hidden layer to ten hidden layers. This means that investigating the use of more than one hidden layer on the performance of the neural network is worth studying and analysis. The proposed neural network with six hidden layers and the previous approach of Jinyeong Oh et al., 2024 [22] have the highest number of hidden layers. Therefore, the complexity and the computational time of these approaches may be higher.

Increasing the number of inputs or the input size of the neural network structure can lead to increased complexity and calculations or computational time. Figure 13 shows a comparison between the number of inputs used by the proposed neural network structures and other approaches by previous researchers. From Fig. 13, the inputs used by our proposed approach particularly using three hidden layers and previous approaches of Sharkawy et al., 2023 [20] and Rushdi et al., 2024 [26] are only two inputs and less than the other approaches. Therefore, the complexity and the computational time of these approaches are lower.

The main important parameter that must be compared is the accuracy of the developed neural network. Achieving a highly accurate method is considered the main objective of any proposed method. The accuracy of the method can be investigated by comparing the mean squared error value (MSE). The method has less MSE value, has higher accuracy. A comparison between the MSE value of our proposed neural network structures and the pervious ones are presented in Fig. 14. From Fig. 14, our proposed neural structure, whether with three or six hidden layers, has the least MSE value compared with other previous structures. Therefore, the proposed neural network structures in this paper have the highest accuracy. From this comparison, the proposed neural network structure with three hidden layers has less complexity and error and higher accuracy compared to other neural network structures proposed by previous researchers.

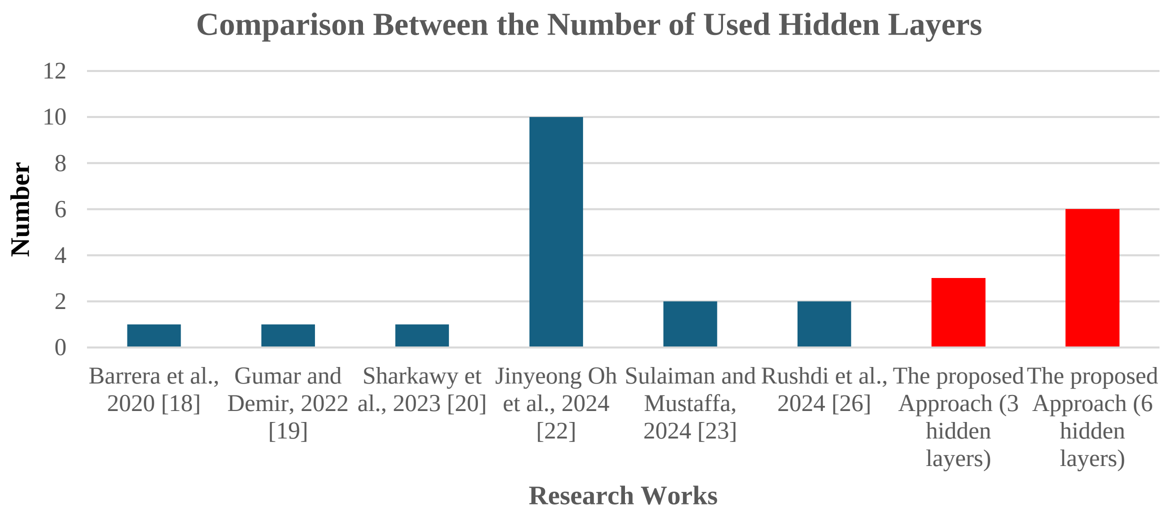


Fig. 12. Comparison between the number of hidden layers used by previous researchers

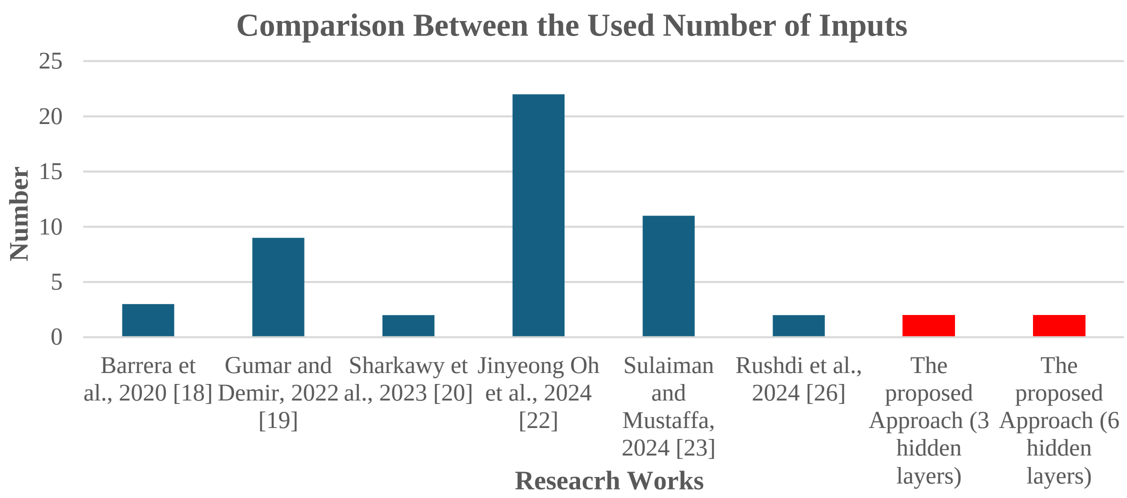


Fig. 13. Comparison between the number of inputs used by previous researchers

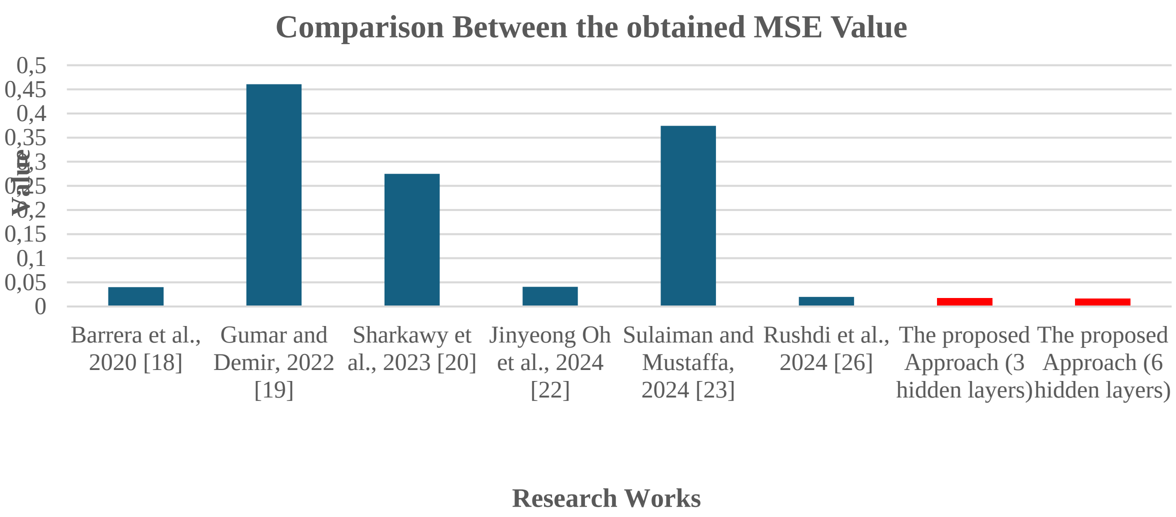


Fig. 14. Comparison between the obtained MSE value by previous researchers.

7. Conclusion and Future Work

In this paper, three neural networks structures are designed based on the principle of increasing the number of hidden layers to assess their accuracy and performance. The first structure has one hidden layer, the second structure has two hidden layers, and the last structure has three hidden layers. To complete the design of these structures, data is obtained from solar power station in Egypt. These data are the temperature of the photovoltaic module and the radiation which are used to be inputs to each NN structure. Also, the power of the photovoltaic module which is the output of each NN structure. The data is divided into three parts: one part for training and the other two parts for the test and validating. The mean squared error obtained by the neural network structure with three hidden layers is the smallest compared with other structures. However, the time used for the train and the complexity of the structure are highest. The mean squared error value by the neural network with one hidden layer is the highest, but the time used for the train and the complexity is the lowest. The training time and complexity of the neural network are not important if offline training is used. For more investigation, all data (train data + test data + validating data) are combined and used for again assessing each neural network structure and the approximated error between the actual output and the predicted output by each neural network structure is determined. The results reveal that the neural network with three hidden layers has the smallest error compared to other structures. The performance of using neural network structure with four, five, and six hidden layers is also investigated and analysis. The results from this investigation show that the neural network with six hidden layers gives the best performance but with the highest computational time and complexity. In conclusion, increasing the hidden layers of the neural network structure increases its accuracy and performance but the computational time is increased. The results of the proposed approach are compared with other previous related works from literature. The result from this comparison reveals that the proposed method achieves the highest accuracy.

Some future works include investigation of the same methodology with other different data from different fields such as robotics, wind power plant, and agriculture, and so on. In addition, applying the same methodology with distinct types of neural networks is highly recommended. Using more samples with different conditions is worth investigating.

List of Abbreviations

Abbreviation	Meaning
NN	Neural network
NN1	Neural network with one hidden layer
NN2	Neural network with two hidden layers
NN3	Neural network with three hidden layers
MSE	Mean squared error
NARX	Nonlinear autoregressive with external input

References

[1] Zhang HC, Huang SH. Applications of neural networks in manufacturing: A state-of-the-art survey. Int J Prod Res. 2007;33(3):705-28. https://doi.org/10.1080/00207549508930175

[2] Lennox B, Montague GA, Frith AM, Gent C, Bevan V. Industrial application of neural networks-an investigation. J Process Control. 2001;11(5):497-507. https://doi.org/10.1016/S0959-1524(00)00027-5

[3] Jiang Y, Yang C, Na J, Li G, Li Y, Zhong J. A brief review of neural networks based learning and control and their applications for robots. Complexity. 2017;2017:14. https://doi.org/10.1155/2017/1895897

[4] Patel J, Goyal R. Applications of artificial neural networks in medical science. Curr Clin Pharmacol. 2008;2(3):217-26. https://doi.org/10.2174/157488407781668811

[5] Maniati M, Evangelos ES, Sklavos S. A neural network approach for integrating banks' decision in shipping finance. Cogent Econ Financ. 2022;10(1):1-20. https://doi.org/10.1080/23322039.2022.2150134

[6] Dube F, Nzimande N, Muzindutsi PF. Application of artificial neural networks in predicting financial distress in the JSE financial services and manufacturing companies. J Sustain Financ Invest. 2023;13(1):723-43. https://doi.org/10.1080/20430795.2021.2017257

[7] Kalogirou SA. Artificial neural networks in renewable energy systems applications: A review. Renew Sustain Energy Rev. 2000;5(4):373-401. https://doi.org/10.1016/S1364-0321(01)00006-5

[8] Iglesias-Sanfeliz Cubero ÍM, Meana-Fernández A, Ríos-Fernández JC, Ackermann T, Gutiérrez-Trashorras AJ. Analysis of neural networks used by artificial intelligence in the energy transition with renewable energies. Appl Sci. 2023;14(1):389. https://doi.org/10.3390/app14010389

[9] Kaushik V. 8 Applications of Neural Networks. Analytic Steps. Available from: https://www.analyticssteps.com/blogs/8-applications-neural-networks

[10] Haykin S. Neural Networks and Learning Machines. 3rd ed. Pearson; 2009. Available from: http://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf

[11] Nielsen MA. Neural Networks and Deep Learning. Determination Press; 2015. Available from: http://neuralnetworksanddeeplearning.com/

[12] Sharkawy AN. Principle of neural network and its main types: Review. J Adv Appl Comput Math. 2020;7:8-19. https://doi.org/10.15377/2409-5761.2020.07.2

[13] Pan Y. Different types of neural networks and applications: Evidence from feedforward, convolutional and recurrent neural networks. Highlights Sci Eng Technol. 2024;85:247-55. https://doi.org/10.54097/6rn1wd81

[14] Liu Z, Gao W, Wan YH, Muljadi E. Wind power plant prediction by using neural networks. In: 2012 IEEE Energy Conversion Congress and Exposition (ECCE). Raleigh, NC, USA: IEEE; 2012. p. 3154-60. https://doi.org/10.1109/ECCE.2012.6342351

[15] Koroglu T, Ekici E. A comparative study on the estimation of wind speed and wind power density using statistical distribution approaches and artificial neural network-based hybrid techniques in Çanakkale, Türkiye. Appl Sci. 2024;14(3):1267. https://doi.org/10.3390/app14031267

[16] Ikram K, Djilali K, Abdennasser D, Al-Sabur R, Ahmed B, Sharkawy AN. Comparative analysis of fouling resistance prediction in shell and tube heat exchangers using advanced machine learning techniques. Res Eng Struct Mater. 2024;10(1):253-70. http://dx.doi.org/10.17515/resm2023.858en0816

[17] Teixeira RSD, Calili RF, Almeida MF, Louzada DR. Recurrent neural networks for estimating the state of health of lithium-ion batteries. Batteries. 2024;10(3):1-28. https://doi.org/10.3390/batteries10030111

[18] Barrera JM, Reina A, Maté A, Trujillo JC. Solar energy prediction model based on artificial neural networks and open data. Sustain. 2020;12(17):1-21. https://doi.org/10.3390/su12176915

[19] Gumar AK, Demir F. Solar photovoltaic power estimation using meta-optimized neural networks. Energies. 2022;15(22):1-15. https://doi.org/10.3390/en15228669

[20] Sharkawy AN, Ali MM, Mousa HHH, Ali AS. Solar PV power estimation and upscaling forecast using different artificial neural networks types: Assessment, validation, and comparison. IEEE Access. 2023;11(February):19279-300. https://doi.org/10.1109/ACCESS.2023.3249108

[21] Azka R, Soefian W, Aryani DR, Jufri FH, Utomo AR. Modelling of photovoltaic system power prediction based on environmental conditions using neural network single and multiple hidden layers. IOP Conf Ser Earth Environ Sci. 2020;599(12032):1-7. https://doi.org/10.1088/1755-1315/599/1/012032

[22] Oh J, So D, Jo J, Kang N, Hwang E, Moon J. Two-stage neural network optimization for robust solar photovoltaic forecasting. Electron. 2024;13(9):1-30. https://doi.org/10.3390/electronics13091659

[23] Sulaiman MH, Mustaffa Z. Forecasting solar power generation using evolutionary mating algorithm-deep neural networks. Energy AI. 2024;16(April):100371. https://doi.org/10.1016/j.egyai.2024.100371

[24] Amer HN, Dahlan NY, Azmi AM, Latip MFA, Onn MS, Tumian A. Solar power prediction based on artificial neural network guided by feature selection for large-scale solar photovoltaic plant. Energy Reports. 2023;9(S12):262-6. https://doi.org/10.1016/j.egyr.2023.09.141

[25] Sharkawy A, Ali MM, Mousa HHH, Ali AS. Short-Term Solar PV Power Generation Day-Ahead Forecasting. International Journal of Robotics and Control Systems. 2022; 2(3):562–580. https://doi.org/10.31763/ijrcs.v2i3.780

[26] Rushdi MA, Yoshida S, Watanabe K, Ohya Y, Ismaiel A. Deep learning approaches for power prediction in wind-solar tower systems. Energies. 2024;17(15):1-23. https://doi.org/10.3390/en17153630

[27] NASA. POWER Data Access Viewer: Prediction of Worldwide Energy Resource. Available from: https://power.larc.nasa.gov/data-access-viewer/

[28] Nahar A, Hasanuzzaman M, Rahim NA. Numerical and experimental investigation on the performance of a photovoltaic thermal collector with parallel plate flow channel under different operating conditions in Malaysia. Sol Energy. 2017;144:517-28. https://doi.org/10.1016/j.solener.2017.01.041

[29] Nasrin R, Hasanuzzaman M, Rahim NA. Effect of high irradiation on photovoltaic power and energy. Int J Energy Res. 2018;42(3):1115-31. https://doi.org/10.1002/er.3907

[30] Mahdi BH, Yousif KM, Melhum AI. Application of artificial neural network to predict wind speed: Case study in Duhok City, Iraq. J Phys Conf Ser. 2021;1829(12002):1-8. https://doi.org/10.1088/1742-6596/1829/1/012002

[31] Rao KP, Babu TV, Anuradha G, Rao BVA. IDI diesel engine performance and exhaust emission analysis using biodiesel with an artificial neural network (ANN). Egypt J Pet. 2017;26(3):593-600. https://doi.org/10.1016/j.ejpe.2016.08.006

[32] Kwon S, Kim J. Real-time upper limb motion estimation from surface electromyography and joint angular velocities using an artificial neural network for human-machine cooperation. IEEE Trans Inf Technol Biomed. 2011;15(4):522-30. https://doi.org/10.1109/TITB.2011.2151869

[33] Zhang YF, Nee AYC, Fuh JYH, Neo KS, Loy HK. A neural network approach to determining optimal inspection sampling size for CMM. Comput Integr Manuf Syst. 1996;9(3):161-9. https://doi.org/10.1016/S0951-5240(96)00005-5

[34] Sharkawy AN, Ameen AG, Mohamed S, Abdel-Jaber GT, Hamdan I. Design, assessment, and modeling of multi-input single-output neural network types for the output power estimation in wind turbine farms. Automation. 2024;5(2):190-212. https://doi.org/10.3390/automation5020012

[35] Madhiarasan M, Deepa SN. Comparative analysis on hidden neurons estimation in multi-layer perceptron neural networks for wind speed forecasting. Artif Intell Rev. 2017;48(4):449-71. https://doi.org/10.1007/s10462-016-9506-6

[36] Du K, Swamy MNS. Neural Networks and Statistical Learning. Springer; 2014. https://doi.org/10.1007/978-1-4471-5571-3
[37] Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math. 1963;11(2):431-41. https://doi.org/10.1137/0111030
[38] Hagan MT, Menhaj MB. Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw. 1994;5(6):2-6. https://doi.org/10.1109/72.329697
[39] Sharkawy AN, Koustoumpardis PN, Aspragathos N. A neural network-based approach for variable admittance control in human-robot cooperation: Online adjustment of the virtual inertia. Intell Serv Robot. 2020;13:495-519. https://doi.org/10.1007/s11370-020-00337-4