Modeling of Solar Radiation Using the Wavelet Neural Network Model in Mataram City Lombok Island

Syamsul Bahri

*a* Program Studi Matematika FMIPA, Mataram University
Majapahit Street No. 62-nd, Mataram-Indonesia

\textsuperscript{1}syamsul.math@unram.ac.id

Abstract

Sunlight is a source of energy for living things in general. In reality, the intensity of solar radiation is an environmental parameter that has positive and negative impacts on human life in particular. Furthermore, the knowledge on the characteristics of solar radiation, including its distribution pattern, is considered by many circles, both policy-makers and researchers in the environmental field. This study aims to create a solar radiation model in response to meteorological factors such as wind speed, air pressure and temperature, humidity, and rainfall using the Wavelet Neural Network (WNN). The modeling of solar radiation in this study is carried out by simultaneously utilizing its advantages as a hybrid model that combines the neural network model and the wavelet method. These advantages through the learning process (supervised learning) are multiplied through the use of the wavelet transform as a pre-processing data method and two type wavelets function, B-spline and Morlet wavelets, as an activation function in the neural network learning process. The WNN model was analyzed in two cases of meteorological variables, which are with and without rainfall. The results based on the root of the mean square error (RMSE) indicator show that the WNN model in these two cases is quite accurate. Meanwhile, the other indicator shows that the interval of the data distribution from the model is within the actual range. This implies that the predicted intensity of the solar radiation will be in a safe position in its adverse effect when the model is used as a reference.

Keywords: Models, Solar Radiation, Meteorology, Wavelets, Neural Networks.

1. Introduction

The world, in general, including Indonesia, is struggling with the emergence of the Covid-19 Pandemic this year. Several problems have attracted the attention of various parties, both regulators and researchers, including the model of virus transmission, strategies for preventing its development and the location of spread, various social, economic, educational, and cultural impacts, as well as the enhancement of human immunity.

Consequently, various methods were adopted in an attempt to increase human immunity. These include the consumption of various vitamins that increases endurance as well as basking in the sun at certain times as recommended by doctors and health experts. Furthermore, the immunity is required to survive and prevent the diseases caused by infection are included the Covid-19 infection \cite{1}. Besides bone formation, especially in children, vitamin D has a role in increasing endurance. Its main source is sunlight in the form of ultraviolet B rays, which is obtained when the body, especially the skin surface, is directly exposed to sunlight. Consequently, individuals that are exposed to sunlight fulfill their body's vitamin D needs, which strengthens the immune system to fight the coronavirus. However, sun exposure also has negative impacts, including sunburn, triggering signs of aging such as sagging, stretched, coarse, and dry skin, increasing the risk of cancer, and damaging eyes as well as the hair color \cite{2}.

Due to its benefits and negative effects, it is important to investigate the characteristics and distribution patterns of solar radiation. Furthermore, the distribution of its data pattern is identified using mathematical modeling. There are several techniques in the modeling of solar radiation. These include the use of statistics \cite{3}, estimation methods \cite{4}, air pressure parameters \cite{5}, multi-
layer neural network perceptron based [6], time series method based on ARMA [7] and machine learning [8], as well as the non-linear time series [9].

Some research that studies the effect of meteorological variables on the intensity of solar radiation has been carried out by some previous researchers. Sousa et al. [10] predicted the concentration of Ozone (O3) using meteorological variables as predictor variables with the adaptive neural network method based on principal analysis (statistical method). In another study, [11] have used a multi-linear regression empirical formula to predict the intensity of solar radiation using a variable magnitude of latitude in radians (L), the ratio of sunlight relative to 12 hours (D), relative humidity (R), maximum air temperature in ° C (t) and the latitude and altitude factor of the place (K). Falayi et al. [12] predicting solar radiation based on a combination of weather parameters, namely clarity index, average daily temperature, the ratio of maximum and minimum daily temperature, relative humidity, and relative duration of sunlight.

However, the modeling of solar radiation in this study is carried out by simultaneously utilizing its advantages as a hybrid model that combines the neural network model and the wavelet method. These advantages through the learning process (supervised learning) are multiplied through the use of the wavelet transform as a pre-processing data method and two type wavelets function, B-spline and Morlet wavelets, as an activation function in the neural network learning process. The method in this study was developed by Bahri [13] to model solar radiation based on the influence of meteorological factors, such as air temperature and pressure, wind speed, humidity, and rainfall, as well as the influence of previous solar radiation intensity in Mataram City, Lombok, West Nusa Tenggara.

2. Theoretical background

2.1. Solar Radiation

The sun as an energy source emits radiation in the form of ultraviolet light (UV) in various wavelengths consisting of UV-A, UV-B, and UV-C rays [14] and [15]. The UV-A rays with a wavelength of 315-400 nm have the lowest energy and are capable of causing rapid aging to skin and indirectly damage the DNA cells. They are generally associated with long-term skin damage, such as wrinkles. Meanwhile, UV-B rays with a wavelength of 280-315 nm have slightly greater energy than UV-A rays. They are also capable of directly causing DNA damage to cells and trigger sunburn and skin cancer. Furthermore, UV-C rays with a wavelength of 180-280 nm have the greatest energy.

Solar radiation to the earth's surface occurs with two events, namely (i) direct radiation, that is solar radiation that the earth receives directly without a change in direction, and (ii) diffuse radiation, that is solar radiation received by the earth after experiencing a change in direction due to reflection or scattering [16], [17], and [17]. The process of reflection and scattering of solar radiation in the earth's atmosphere is influenced by water vapor, carbon dioxide (CO2), and ozone (O3). Mairisdawenti et al. [18] show that ozone concentration follows the pattern of solar radiation intensity and surface temperature and is inversely related to air humidity. On the other hand, Sousa et al. [10] have shown that ozone concentration can be influenced by several meteorological variables such as temperature, wind speed, air pressure, air humidity, and NO2 concentration. Research conducted by [19] shows that high rainfall intensity will reduce the intensity of solar radiation. Therefore, the meteorological variables indirectly affect the intensity of solar radiation received by the earth.

2.2. Neural Network

The Neural network (NN) is one of the tools in the soft-computing category whose development is motivated by the working principle of the human brain. This model was originally developed by Warren McCulloch and Walter Pitts in 1943 [20], and it is used as one of the methods of time series analysis because it has a simple calculation process, high-level interconnection, and adaptive interactions between elements [21].

The difference in its application is determined by several indicators, including the architectural model, the feed-forward, back-forward, or recurrent nature of learning, type, activation function used, the number of hidden layers, and the neurons in each [13].
2.3. Wavelet

The wavelet is used to localize a function in two ways, namely position (time) and scale (frequency), and is widely applied in data processing, such as in signal and time series analysis.

Mathematically, wavelets are constructed from the process of translation and dilation on a particular function, which is defined as follows [22]:

$$\psi_{a,b}(t) = \sqrt{|a|}^{-1/2} \psi \left( \frac{t-b}{a} \right), \quad a, b \in \mathbb{R} \text{ and } a \neq 0,$$

(1)

Where \( \psi \) represents a function of mother wavelet, the constants \( a \) and \( b \) represent the dilatation (scale) and the translation (time) parameter, respectively.

Furthermore, wavelets have been widely used in signal processing since 1900 and have developed into a time-series analysis tool. This is an alternative to time series analysis compared to statistical methods that previously existed due to its advantage in denoising processes, data compression, and multi-resolution [13]. In line with this, wavelets are used as a tool for data pre-processing and as an alternative to the activation function in the NN model. Its application as an activation function in NN, known as the WNN model, is able to increase accuracy in pattern recognition or time-series data forecasting [23], [24], and [25].

Some of the wavelet function families often used as the basis of analysis include Daubechies, B-Spline, and Morlet wavelets. Furthermore, Daubechies has compact support and was developed by Ingrid Daubechies around the 1990s. It is a family of functions called N-order Daubechies (DbN) wavelet, for an N natural number. The N-order wavelet built based on the N-1 order Daubechies polynomials are defined as follows:

$$P_{N-1}(x) = \sum_{k=0}^{N-1} \left( \frac{2N-1}{k} \right) x^k (1-x)^{N-1-k}. $$

(2)

Daubechies wavelet is constructed by N-order Daubechies polynomial in Equation (2).

The next type of wavelet is the B-spline, which is obtained recursively from the convolution of the Haar scale function [26]. For natural numbers, \( n \geq 1 \), the n-order B-spline is defined as follows

$$N_n(x) = \underbrace{N_1 \ast N_1 \ast \ldots \ast N_1}_{n \text{ factors}}(x) = \left( N_{n-1} \ast N_1 \right)(x)$$

(3)

$$= \int_0^1 N_{n-1}(x-t) \, dt,$$

where \( N_1(x) \) is Haar’s scaling function,

$$N_1(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}.$$  

(4)

Based on this convolution, Equation (3) is written as follows:

$$N_n(x) = \frac{1}{(n-1)!} \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x-k)^{n-1}_+, $$

(5)
with \((x)_{p}^{n} = \max\{0, x\}^{n}\).

For \(n \geq 2\), the first derivative of \(N_{n}(x)\) in Equation (5) is called an \(n-1\) order B-spline wavelet. [27], using the Gabor (cosine-modulated Gaussian) function approximation, a formula for the \(n\)-order B-spline wavelet approximation was produced as shown in Equation (5).

\[
\psi^{n}(x) = \frac{4b^{n+1}}{\sqrt{2\pi(n+1)\sigma_{W}^{2}}} \cos(2\pi f_{0}(2x-1)) \exp \left(\frac{-(2x-1)^{2}}{2\sigma_{W}^{2}(n+1)}\right),
\]

(6)

where \(n\) represents the order of B-spline wavelet, constant \(b = 0.657066\), \(f_{0} = 0.409177\), and \(\sigma_{W}^{2} = 0.561145\).

The B-Spline has several advantages, which form the basis for consideration when choosing wavelet as an analytical tool. These include having an explicit formula in the time and frequency domain, a symmetrical shape, compact support, and a simple manipulation process [28].

Furthermore, Morlet is a continuous wavelet which is derived proportionally from the Gaussian cosine and probability density function (pdf) given by the following equation:

\[
\psi(x) = \exp(-x^{2}) \cos(5x)
\]

(7)

The characteristics of the Morlet wavelet function expressed in Equation (7) include non-orthogonality, have infinite support, and maximum energy around the origin point [28].

3. Research Methods

3.1. Data and organizing data

The secondary and rainfall data used were obtained from West Nusa Tenggara Environment and Forestry Service (DLHK) and the Nusa Tenggara I River Basin Office, Directorate General of Water Resources, Ministry of Public Works and Public Housing (PUPR), respectively. The data collected is meteorological, consisting of wind speed \((x_{1})\), humidity \((x_{2})\), air temperature \((x_{3})\), air pressure \((x_{4})\), and rainfall \((x_{5})\) as the predictor and solar radiation \((y)\) as response variable from June 2018 to May 2019.

3.2. The architecture of The Wavelet Neural Network

The application of the wavelet neural network is based on the architecture proposed by [29] as shown in Figure 1. The architecture consists of 5 layers, the first is used as the input (original data), and the second serves as the pre-processing stage for input data from layer 1 using data normalization and the 8-order Daubechies wavelet transformation level 3. The data normalization uses the following formula:

\[
X_{i} = \frac{x_{i} - \min_{1 \leq i \leq n} x_{i}}{\max_{1 \leq i \leq n} x_{i} - \min_{1 \leq i \leq n} x_{i}}
\]

(8)

Where \(x_{i}\) states \(i\) original data, \(x_{i}\) is the normalization result of the data \(x_{i}, x_{i} = \min_{1 \leq i \leq n} x_{i}\), and \(\max_{1 \leq i \leq n} x_{i}\) is transformed using 8-order Daubechies wavelet transformation level 3.
In the third layer, the pre-processed data were summed in parallel with the $w_i$ and $v_i$ weights for $i = 1, 2, \cdots n$,

$$X_w = \sum_{i=1}^{n} w_i x_i \quad \text{and} \quad X_v = \sum_{i=1}^{n} v_i x_i. \quad (9)$$

Furthermore, in the fourth, the $X_w$ and $X_v$ values are activated using 3-order B-spline and Morlet wavelet, respectively, given by equations (6) and (7).

At the fifth, the output of the WNN model is obtained based on Equation (10) below.

$$y_{\text{WNN}} = \alpha \times y^{B^m}(X_w) \times y^{M}(X_v) + \beta \quad (10)$$

for a constant $\alpha, \beta \in \mathbb{R}$.

### 3.3. Learning Parameters of WNN proposed

The proposed model is a WNN feed-forward type using supervised training as the learning process carried out to minimize the following cost functions:

$$E = \frac{1}{2NJ} \sum_{i=1}^{N} \left( y_i - y_i^j \right) \quad (11)$$

where $N$ represents the amount of data, $J_r = \max_{1 \leq i \leq N} \left\{ y_i^j \right\} - \min_{1 \leq i \leq N} \left\{ y_i^j \right\}$, $y_i$ is the output of the $i$th WNN model, and $y_i^j$ is the $i$th actual data (target output). The $w_i$ and $v_i$, weight parameters for $i = 1, 2, \cdots N$ were optimized using the gradient descent with momentum (GDM) algorithm. Parameters $w_i$ and $v_i$, for $i = 1, 2, \cdots N$ are updated based on Equations (12) and (13).

$$w_i(t + 1) = m_c w_i(t) + (1 - m_c) \eta_w \frac{\partial E}{\partial w_i}, \quad (12)$$

$$v_i(t + 1) = m_c v_i(t) + (1 - m_c) \eta_v \frac{\partial E}{\partial v_i}, \quad (13)$$
where $\eta_w$ dan $\eta_v$ represent the learning parameters for the weights W and V, $m_t$ represents the momentum coefficient on GDM, and the values of partial derivatives $\frac{\partial E}{\partial \omega_i}$ and $\frac{\partial E}{\partial \nu_i}$ are given by Equations (14) and (15) below.

$$\frac{\partial E}{\partial \omega_i} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial y_{B}} \frac{\partial y_{B}}{\partial X_w} \frac{\partial X_w}{\partial \omega_i}. \quad (14)$$

$$\frac{\partial E}{\partial \nu_i} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial y_{M}} \frac{\partial y_{M}}{\partial X_v} \frac{\partial X_v}{\partial \nu_i}. \quad (15)$$

The selection of values $\eta_w$ dan $\eta_v$ to ensure the convergence of FF-WNN model is determined based on the results of [25] as follows:

$$0 \leq \eta_\rho \leq \frac{2}{\max_i \left\{ \frac{\partial y(t)}{\partial \eta_\rho} \right\}}, \quad \rho = w \text{ dan } v. \quad (16)$$

4. Result and Discussion

4.1. Data Organization

The intensity of solar radiation is one of the parameters influenced by several meteorological variables such as wind speed, air pressure, humidity, and rainfall. Furthermore, this intensity as time-series data is analyzed based on the effect of solar radiation at previous times. In line with this, the intensity of solar radiation y is formulated based on the following relationships:

$$y(t) = f \left( x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), y(t-k) \right), \quad \text{for } k \in \mathbb{N}. \quad (17)$$

where $x_i(t)$, $i = 1, 2, \cdots, 5$ represent wind speed, air humidity, temperature, and pressure, as well as rainfall at time t and y states the intensity of solar radiation.

4.2. Solar Radiation Model Based on Meteorological Variables Without Rainfall

Based on Equation (17), the relationship between the intensity of solar radiation (a response variable) and the meteorological variables without rainfall (as a predictor) is given by Equation (18).

$$y(t) = f \left( x_1(t-1), x_1(t-3), x_2(t-1), x_2(t-2), x_3(t-1), x_3(t-2), x_4(t-1), x_4(t-3), y(t-1), y(t-2), y(t-3) \right). \quad (18)$$

The learning of 20,000 epochs on 327 daily means of solar radiation intensity, which is organized into 280 data for training (in-sample) and 47 for testing (out-sample) gives results as shown in Figure 2.
Figure 2. Comparison of graph patterns for solar radiation intensity based on meteorological variables without rainfall between WNN model output (blue) and actual data (red)

Based on Figure 2, it can be seen that the output of the WNN model has shown a similar pattern to the original data, except on the first data. The pattern of data was generated by the WNN model provides the data of WNN model is in the interval of original data; that is:

\[
\min y_m < \min y_o < \max y_m < \max y_o
\]

Where \( y_m \) and \( y_o \) represent the data of the WNN model and the original data, respectively.

Statistically, the accuracy of WNN for modeling the intensity of solar radiation based on the influence of meteorological factors is given by some indicators, which are minimum and maximum value and root of mean square error (RMSE) as an indicator of model performance (Table 1). Regarding the average value, Table 1 provides information that the difference between the mean value of the data generated by the WNN model and the original data is quite small compared to the distribution of the original data, which is 3.6%.

Table 1. The performance of the WNN model for modeling solar radiation intensity based on meteorological variables without rainfall

<table>
<thead>
<tr>
<th>Data</th>
<th>Statistical Indicators</th>
<th>Performance (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimal Data</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td>Actual</td>
</tr>
<tr>
<td>Training</td>
<td>18.7547</td>
<td>0</td>
</tr>
<tr>
<td>Testing</td>
<td>30.6461</td>
<td>2.3077</td>
</tr>
</tbody>
</table>

Based on Table 1, the range of data generated by the WNN model is 95.7088 on training data (in-sample) and 135.375 on testing data (out-sample). Furthermore, the average follows the distribution pattern, where those produced by the WNN model is slightly higher (61.8790) than the actual (57.0018). This implies that the intensity of the solar radiation generated by the WNN model is in a safe position in predictions and its impacts.

4.3. Solar Radiation Model based on Meteorological Variables with Rainfall

Based on Equation (17), the relationship between the intensity of solar radiation (as a response variable) and meteorological variables with rainfall (as a predictor variable) is given by Equation (19).
\[
y(t) = f \left( x_1(t-1), x_1(t-3), x_2(t-1), x_2(t-3), x_3(t-1), x_3(t-3), x_4(t-1), x_4(t-3), x_5(t-2), x_5(t-3), y(t-1), y(t-3) \right)
\]

The learning of 20,000 epochs on 327 daily mean on the intensity of solar radiation based on meteorological variables with rainfall, which are organized into 280 data for training data (in-sample) and 47 for testing (out-sample) gives results as shown in Figure 3.

![Comparison of graph patterns for solar radiation intensity based on meteorological variables with rainfall between WNN model output (blue) and actual data (red)](image)

Figure 3. Comparison of graph patterns for solar radiation intensity based on meteorological variables with rainfall between WNN model output (blue) and actual data (red).

Statistically, the accuracy of WNN for modeling the intensity of solar radiation based on the influence of meteorological factors is given by some indicators, which are minimum and maximum value and root of mean square error (RMSE) as an indicator of model performance (Table 2).

Based on Table 1 and Table 2, the influence of the rainfall variable provides increased accuracy based on the RMSE indicator on the WNN model in predicting the intensity of solar radiation, which is 2.7513 in the training data and 0.0177 in the test data. This conclusion is also reinforced by the difference in the average value of the WNN model output data for the meteorological variable, and the rainfall variable with the actual data is 2.4247. This value is smaller than the case for the meteorological variable without the rainfall variable, which is 4.8772.

Based on the graph in Figure 2, it can be seen that the data on the intensity of solar radiation in the City of Mataram contains approximately ten outlier's data, which is represented by a significant increase in the intensity of solar radiation for 3 days and a sharp decrease in the intensity of solar irradiation for 7 days. This fact contributes to a significant error rate.

Table 2. The performance of the WNN model for modeling solar radiation intensity based on meteorological variables with rainfall

<table>
<thead>
<tr>
<th>Data</th>
<th>Statistical Indicators</th>
<th>Performance (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimal Data</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td>Actual</td>
</tr>
<tr>
<td>Training</td>
<td>10.0886</td>
<td>0</td>
</tr>
<tr>
<td>Testing</td>
<td>30.4592</td>
<td>2.3077</td>
</tr>
</tbody>
</table>

From Table 2, the range of data generated by the WNN model is 66.7444 on the training (in-sample) and 135.375 on the testing (out-sample). Furthermore, the mean follows the distribution pattern, where those for the data produced by the WNN model is slightly higher (59.4265) than the actual (57.0018). This means that the solar radiation intensity generated by the WNN model is in a safe position in prediction and its impacts.

5. Conclusion

In this study, the modeling of the intensity of solar radiation is based on the hybrid feed-forward Neural Network using the wavelet method. Based on the RMSE indicator, the application of the
FF-WNN model gave good results, as shown in the graphical pattern of the output compared to the target data, especially in the in-sample section. Furthermore, it is seen from the two observed cases that the WNN model for the case of solar radiation intensity based on the influence of meteorological factors with rainfall is statistically and relatively better than the cases without rainfall as a predictor. Moreover, the distribution of data generated by the WNN model for predictor factors with rainfall has a smaller (compressed) range than without.

Acknowledgment

The authors are grateful to the West Nusa Tenggara Environment and Forestry Office and the Nusa Tenggara I River Basin Office, Directorate General of Water Resources, Ministry of Public Works and Public Housing (PUPR) for providing the data used in this study. The authors are also grateful to all parties who have provided input for the improvement of this study.

References

2014.


