Estimation of Atmospheric Precipitable Water in Thailand using an Artificial Neural Network

Sumaman Buntoung[1]*, Jindarat Pariyothon[1] and Patsakorn Detkhon[2]

[1]Department of Physics, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand

[2]Program of Physics, Department of Science and Technology, Faculty of Liberal Arts and Science, Roi Et Rajabhat University, Roi Et 45120, Thailand

* Corresponding author. E-mail address: s.buntoung@gmail.com

Received: 18 June 2020; Revised: 11 August 2020; Accepted: 17 August 2020; Available online: 10 September 2020

Abstract

In this work, an Artificial Neural Network (ANN) was proposed to estimate monthly average precipitable water (PW) using ambient air relative humidity, ambient air temperature, saturated water vapour pressure and the order of the month as input data. The PW data measured from ground-based sunphotometer were used as output data. The multilayer perceptron ANN using the back propagation training algorithm with two hidden layers was employed for deriving the PW. A five-year period (2009-2013) of the data collected from four meteorological stations, namely Chiang Mai (18.98°N, 98.98°E), Ubon Ratchathani (15.25°N, 104.87°E), Bangkok (13.67°N, 100.60°E) and Songkhla (7.20°N, 100.60°E) were used to train the ANN. An independent two-year period (2014-2015) of the data from the same stations were used to evaluate the performance of the trained ANN model. The result shows that PW at the four stations derived from the ANN agrees well with those obtained from the measurement, with the discrepancy in terms of root mean square error (RMSE) and mean bias error (MBE) of 7.5% and -0.1%, respectively.

Keywords: precipitable water, artificial neural network, back propagation algorithm

Introduction

Water vapour is one of the atmospheric compositions that is highly varies with height, region and time. The amount of water vapour can count from zero at the dry zone up to about 4% of the atmospheric compositions at humid zone (Kämpfer, 2013). The atmospheric water vapour contains mostly in the tropospheric layer and some in the stratospheric layer. It plays an important role in the physical and chemical processes of the Earth's atmosphere. The tropospheric water vapour has influences on optical properties of aerosols resulting in the formation of clouds and rain (Iqbal, 1983). It also changes visibility of the atmosphere. Moreover, the water vapour is well-known as an active greenhouse gas in the atmosphere as it can absorb atmospheric radiation. More than 60% of the natural greenhouse effect is contributed by water vapour (Iqbal, 1983; Nunez, 1993; Taylor, 2005). For the upper atmosphere, there is very small amount of water vapour but it is also important for circulation and stability of the atmosphere (Kämpfer, 2013). Therefore, a number of researchers have been interested in studying the atmospheric water vapour (Maghrabi & Dajani, 2013; Abbasy, Abbasi, Asgari, & Ghods, 2017; Yue & Ye, 2019). For example, Yue and Ye (2019) aimed to predict water vapour at Antarctic Zhongshan station (69.37°S, 76.37°E) using global positioning system (GPS) zenith tropospheric delay. Akatsuka, Oyoshi, & Takeuchi (2010) used the brightness temperature data from the MTSAT satellite to calculate water vapour in East and Southeast Asia and the western Pacific.

The amount of water vapour can be quantified by precipitable water (PW) which measures the depth of the water that would be condensed from the top of the atmosphere to the surface (Senkal, 2015). In general, PW can be measured by ground-based instruments using microwave radiometer and sunphotometer (Pérez-Ramírez et al., 2014), or estimated using GPS technique (Yue & Ye, 2019). However, such instruments have high prices and thus ground-based stations are very sparse. Therefore, several models have been proposed to estimate PW (Hay, 1971; Reitan, 1960). However, such models are complicated and need a number of input data which are sometime not available. Recently, artificial neural network (ANN) has been widely used in many fields including atmospheric sciences (Hung, Babel, Weesakul, & Tripathi, 2009; Kurtgoz, Karagoz, & Deniz, 2017; Yari, Ayoobi, & Ghassemi, 2014).

In this work, the ANN is proposed to estimate PW as it does not require the establishment of equations to calculate PW. In addition, the study focuses on estimating PW in the regions of Thailand.

Methods and Materials

Collection of Data

Several meteorological surface data e.g. ambient air relative humidity (RH) and ambient air temperature (T) in three-hours basis were collected during 2009-2015 from the meteorological stations in the four main regions of Thailand, namely Chiang Mai (CM; 18.98°N, 98.98°E) in the northern region, Ubon Ratchathani (UB; 15.25°N, 104.87°E) in the northeastern region, Bangkok (BK; 13.67°N, 100.60°E) in the central region and Songkhla (SK; 7.20°N, 100.60°E) in the southern region. The ambient air temperature was used to calculate saturated water vapour pressure (p_{vs}) using a formula as follows (Iqbal, 1983):

$$p_{vs} = \exp\left(26.23 - \frac{5416}{T}\right) \quad (1)$$

where p_{vs} is saturated water vapour pressure (hPa) and T is ambient air temperature (K).

Afterward, the three datasets including ambient relative humidity, ambient air temperature and saturated water vapour pressure were normalized with their maximum values and then processed to obtain the monthly average data and used in training and testing ANN.

Apart from the meteorological data, the PW data used in this work were measured by sunphotometers (Cimel Electronique, model CE318) of the Tropical Atmospheric Physics Laboratory of Silpakorn University, Thailand which is a member of Aerosol Robotic Network (AERONET) of NASA (https://aeronet.gsfc.nasa.gov). The sunphotometer is an instrument measured the direct sun spectral irradiance and the sky radiance at the wavelength of 340, 380, 440, 500, 675, 870, 940 and 1020 nm (Holben et al., 1998). The retrieval of the PW is based on measurements taken at 940 nm (water absorption peak) and 1020 nm (no water absorption) (Reagan, Thome, Herman, & Gall, 1987; Bruegge et al., 1992). There are several products such as aerosol optical depth, aerosol size distribution and PW data in instantaneous and daily averages formats from the AERONET website. However, in this work, only the daily averages of PW data at Level 1.5 (cloud screened and quality controlled) were selected. The PW data measured from the sunphotometers were collected during 2009-2015 at the ground-based stations, namely CM, UB, SK and Nakhon Pathom (NP; 13.82°N, 100.04°E). NP is the nearest station to Bangkok. The daily average data were normalized with its maximum value and then calculated to obtain monthly average normalized PW data.

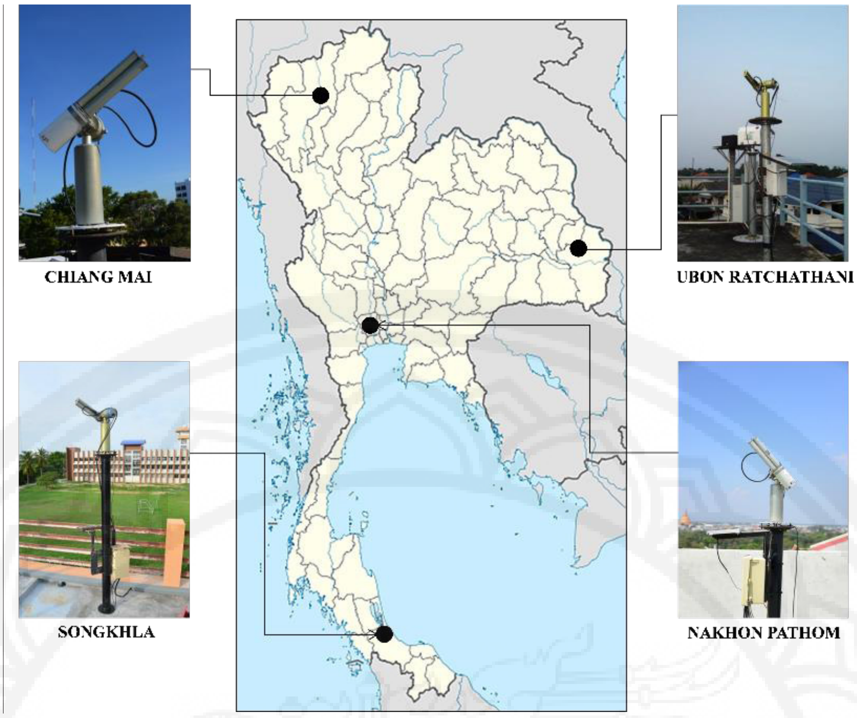

Figure 1 Locations and pictorial views of the sunphotometer stations

ANN model

Artificial Neural Network (ANN) is a mathematical model simulating biological neural system in the human brain. Generally, feedforward neural network called multilayer perceptron is mostly popular. It consists one input layer, multiple hidden layers and one output layer (Figure 2).

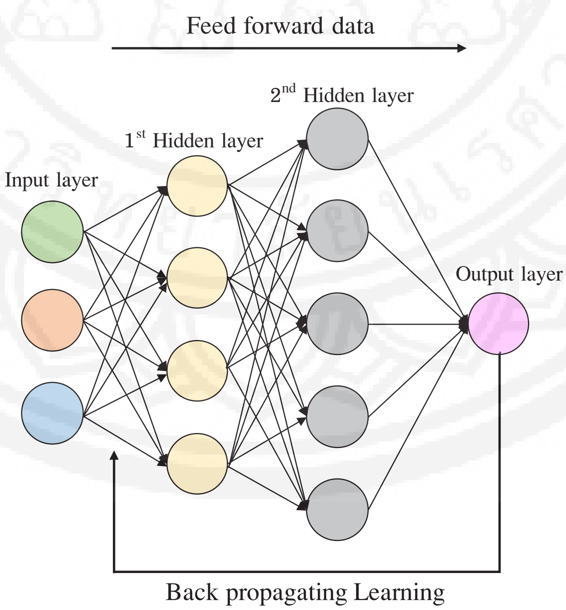

Figure 2 Diagram of the feedforward neural network

In Figure 2, it shows an example of the feedforward neural network including one input layer, two hidden layers and one output layer. All nodes of the input data are connected to each node of the first hidden layer via weighted connections. Similarly, all nodes of the first hidden layers are also linked to each node of the second hidden layer. At the end, all nodes of the second hidden layer are crossed to the output layer.

At each node, a weighted sum of those input data is calculated and this weighted sum is then sent through activation function (Figure 3). The output of each node is then transferred to nodes of the next hidden layers.

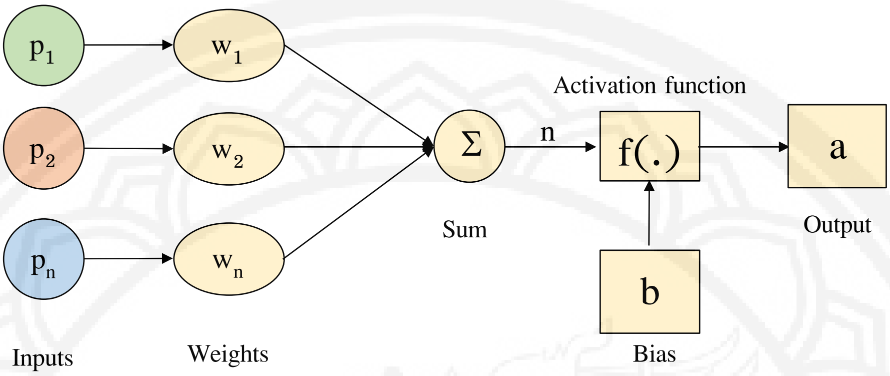

Figure 3 Diagram of the feedforward neural network

The corresponding mathematical model (Yari et al., 2014; Lv et al., 2018) calculated at each node is as follows:

$$n = \sum_{i=1}^{n} w_i p_i + b \tag{2}$$

where n is the net input of the hidden layer, w_i is the weight between the i^{th} input and the hidden layer, p_i is the i^{th} input, b is the bias of the hidden layer, and n is the total number of input.

The net input (n) passes through an activation function which can generate the neuron output (a).

$$a = f(n) \tag{3}$$

The activation function selected in this work is the sigmoid function which can written as follows:

$$f(x) = \frac{1}{1+e^{-x}} \tag{4}$$

where x is the input of function.

To train the network, the back propagation algorithm is commonly used to minimize an error between the actual and predicted outputs (Gardner & Dorling, 1998). The error is calculated and sent back to the system to adjust the weight for the next input (Figure 4). The network learns by adjusting the interconnections between the layers. When the learning or training procedure is completed, a suitable output is produced at the output layer.

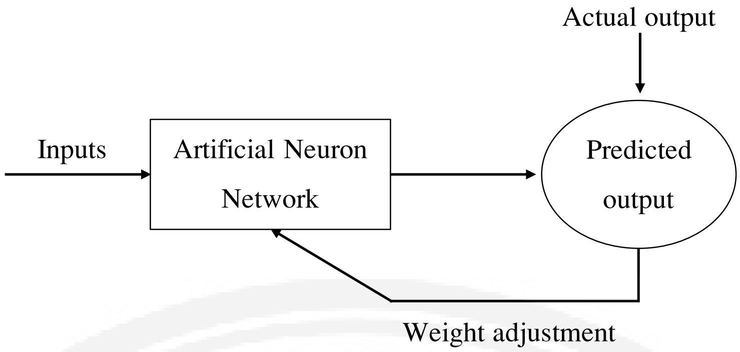

Figure 4 The process of the back propagation ANN

In this study, the multilayer perceptron ANN model written in the form of a software called WEKA (Frank, Hall, & Witten, 2016) was used to estimate PW on monthly basis. This software was developed by the University of Waikato, New Zealand. There are two steps in this study. The first step is to train the model and the second step is to test the performance of the model. For the first step, the normalized values of monthly average RH, T and p_{vs} data at CM, UB, BK and SK, and the order of the month (m) were used as input data as they influents the PW, while the normalized values of monthly average PW data at CM, UB, NP and SK were used as expected output of the model. The data during 2009-2013 were used for training the ANN model. Initially, number and node of hidden layers were specified to suit the model. The less numbers and nodes may cause difficulties in learning process of the model while more numbers and nodes may take more processing time (Hung et al., 2009). The data were trained and the weight values were adjusted until the ANN model can provide the best result. In the second step, the optimal ANN model was employed to estimate the PW data. To achieve this issue, the normalized values of RH, T and p_{vs} during 2014-2015 were used as input data in the trained ANN model. The estimated PW values obtained from the ANN were compared with the ground-based measured PW values. It is observed that the estimated PW data at BK station were tested with those at NP station as there is no such ground-based instrument at BK. So that the available data from the nearest site e.g. NP were used.

The comparison result is presented in terms of root mean square error (RMSE, in % and cm), mean bias error (MBE, in %) and mean absolute error (MAE, in cm) which can be expressed as follows:

$$\mathrm{RMSE}\ (\%) = \frac{\sqrt{\frac{\sum_{i=1}^{N}(PW_{model}-PW_{meas})^2}{N}}}{\frac{\sum_{i=1}^{N}PW_{meas}}{N}} \times 100\% \quad (4)$$

$$\mathrm{MBE}\ (\%) = \frac{\frac{\sum_{i=1}^{N}(PW_{model}-PW_{meas})}{N}}{\frac{\sum_{i=1}^{N}PW_{meas}}{N}} \times 100\% \quad (5)$$

$$\mathrm{RMSE}\ (\mathrm{cm}) = \sqrt{\frac{\sum_{i=1}^{N}(PW_{model}-PW_{meas})^2}{N}} \quad (6)$$

$$\mathrm{MAE}\ (\mathrm{cm}) = \frac{\sum_{i=1}^{N}(PW_{model}-PW_{meas})}{N} \quad (7)$$

where PW_{model} is monthly average precipitable water obtained from the ANN (cm), PW_{meas} is monthly average precipitable water measured from the sunphotometer (cm) and N is number of data.

Results and Discussions

Using the corresponding data during 2009-2013, the ANN model has been designed with two hidden layers (Figure 5). The first hidden layer consists of four neurons and the second hidden layer has three neurons. This architecture gave the best result.

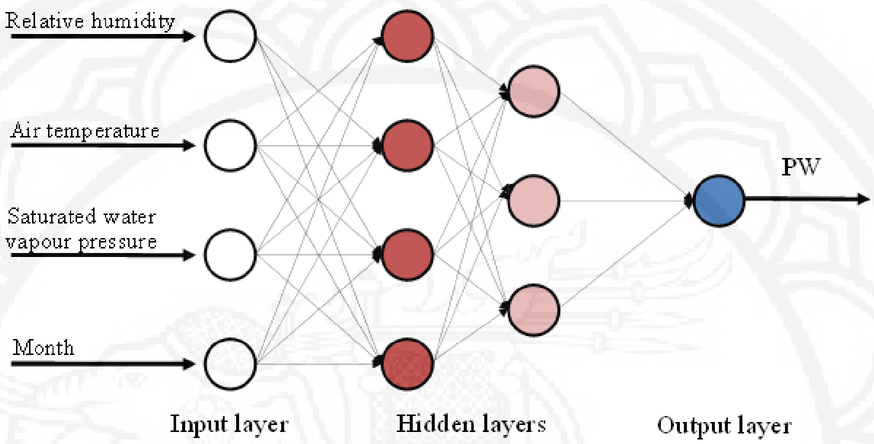

Figure 5 Structure of ANN for PW estimation

The comparison result between the PW data estimated from the optimized ANN and measured from the sunphotometers is shown in Figure 6 for each station and Figure 7 for all stations. The summary of the differences is presented in Table 1. From the results, these two datasets were in good agreement with RMSE and MBE of 7.5% and -0.1%, respectively.

Table 1 The differences between the PW data from the ANN model and from the measurement in terms of root mean square error (RMSE), mean bias error (MBE), mean absolute error (MAE) and coefficient of determination (R^2). N is number of data

Stations	RMSE (%)	MBE (%)	RMSE (cm)	MAE (cm)	R^2	N
Chiang Mai	7.75	1.98	0.29	0.074	0.96	24
Ubon Ratchathani	7.01	-3.26	0.31	-0.146	0.92	19
Bangkok/Nakhon Pathom	7.97	0.81	0.37	0.038	0.90	22
Songkhla	7.11	-0.06	0.34	-0.003	0.57	19
All stations	7.50	-0.1	0.33	-0.003	0.91	84

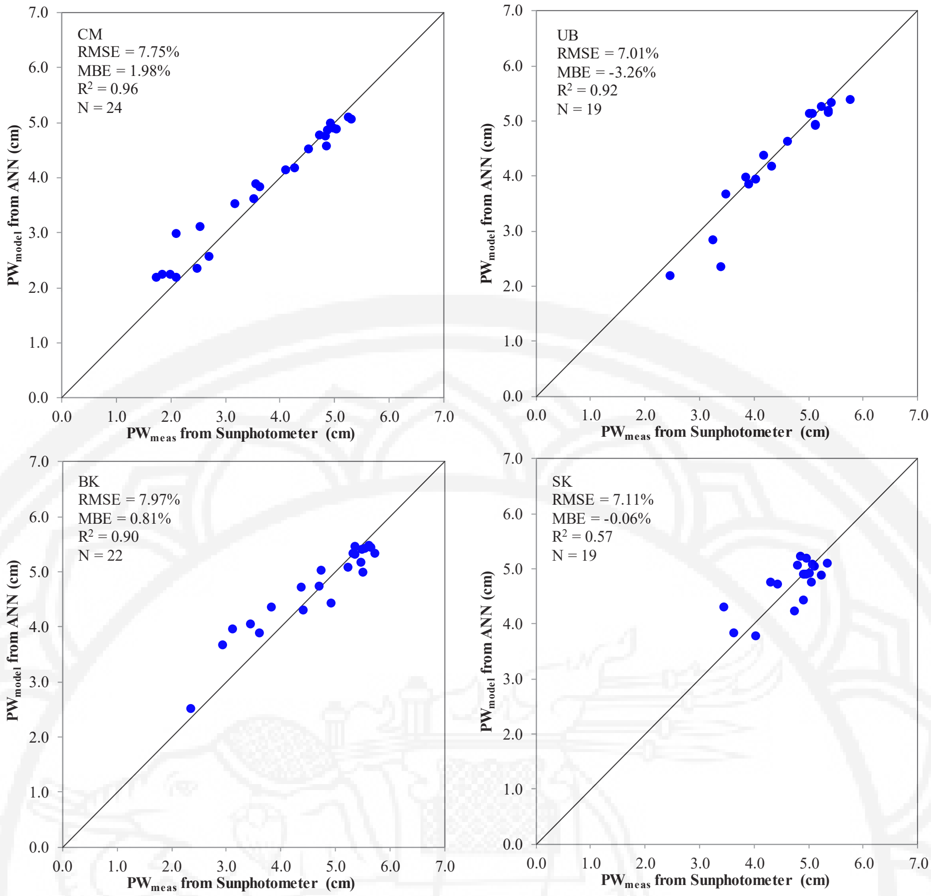

Figure 6 Comparison between the monthly average precipitable water (PW) from the ANN model and the sunphotometers for the four stations

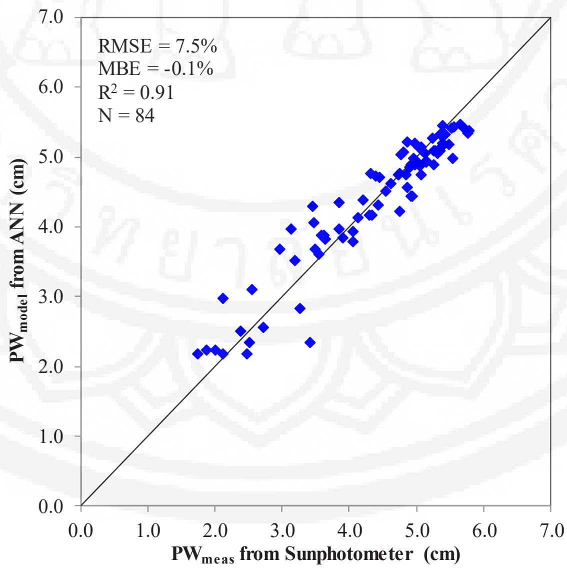

Figure 7 Comparison between the monthly average precipitable water (PW) from the ANN model and the sunphotometers for all stations

From literature reviews, there are several studies used ANN model to estimate PW data in different countries. For example, Shastri and Pathak (2018) predicted PW at Hyderabad station (13.02°N, 77.57°E), India during

2011-2015 using three methods: ANN model, support vector machine and multiple linear regression. The accuracies of the estimated PW data were analyzed by comparing the PW retrieved from the models with the actual values. The results show that the data from the ANN model performs better than the others two models, with RMSE and MAE to be less than 0.5 cm in most cases. In Europe, Basili et al. (2008) also used the feedforward ANN model to estimate PW during April-May, 2005-2007 over Mediterranean area using brightness temperature as an input data. The performance of the model was presented in terms of RMSE and MAE. The RMSE under clear sky condition is 0.158 cm over sea and 0.336 cm over land, with negligible bias. In the case of Thailand, no ANN work for predicting PW has been reported. However, Phokate and Atyotha (2018) studied the PW in Thailand using a statistical model. Their results have RMSE and coefficient of determination (R^2) of 0.528 cm and 0.85, respectively. Concerning our present study, the performance of our proposed ANN is comparable to those from the literatures and our model performs better than that of Phokate and Atyotha (2018). It is observed that R^2 of Songkhla station is lower than that of the other three stations. This may be due to the fact that the location of Songkhla station is near the sea, causing high atmospheric water vapour all year round. In addition, this station is situated in the southern region with highly different climate, as compared to the climate of the other regions. However, the error of the prediction in terms of RMSE and MBE are still relatively low. This indicates that the proposed ANN model can be used to estimate PW in Thailand with good accuracy.

In order to improve the accuracy of the result, it is suggested that other related data such as altitude of the station be additionally used as input of the model.

Conclusion

In this work, the ANN was used to estimate monthly average precipitable water. The input data of ANN are ambient relative humidity, ambient temperature, saturated water vapour pressure and the order of the month. The ANN was trained using the back propagation algorithm and the data from the four stations in Thailand during the period of 2009-2013. The precipitable water from the ground-based sunphotometers during a two-year period (2014-2015) were used to demonstrate the performance of the ANN. The result shows that the precipitable water estimated from the ANN agree well with those obtained from the measurement with RMSE and MBE of 7.5% and -0.1%, respectively, and correlation coefficient (R^2) of 0.91.

Acknowledgments

The authors are grateful to the Thai Meteorological Department for providing ambient relative humidity and temperature data at the ground-based meteorological stations used in this work. The authors also thank AERONET for processing and providing the precipitable water data.

References

Abbasy, S., Abbasi, M., Asgari, J., & Ghods, A. (2017). Precipitable water vapour estimation using the permanent single GPS station in Zanjan, Iran. *Meteorological Applications*, *24,* 415-422.

Akatsuka, S., Oyoshi, K., & Takeuchi, W. (2010). Mapping of precipitable water using MTSAT data. In *Proceedings of 31st Asian Conference on Remote Sensing (ACRS),* Vietnam: Hanoi.

Basili, P., Bonafoni, S., Mattioli, V., Pelliccia, F., Serpolla, A, Bocci, E., & Ciotti, P. (2008). Development of a neural network for precipitable water vapor retrieval over ocean and land. In *2008 Microwave Radiometry and Remote Sensing of the Environment, 11-14 March 2008* (pp. 1-4). Firenze, Italy: Institute of Electrical and Electronics Engineers.

Bruegge, C. J., Conel, J. E., Green, R. O., Margolis, J. S., Holm, R. G., & Toon, G. (1992). Water vapor column abundance retrievals during FIFE. *Journal of Geophysical Research: Atmospheres, 97*(D17), 18759-18768.

Frank, E., Hall, M. A., & Witten, I. H. (2016). The WEKA Workbench. USA: Morgan Kaufmann.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences, *Atmospheric Environment*, *32*(14-15), 2627-2636.

Hay, J. E. (1971). Precipitable water over Canada: II distribution. *Atmosphere*, *9*(4), 101-111.

Holben, B. N., Eck, T. F., Slutsker, I. a., Tanre, D., Buis, J., Setzer, A., & Nakajima, T. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, *66*(1), 1-16.

Hung, N. Q., Babel, M. S., Weesakul, S., & Tripathi, N. K. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Science*, *13,* 1413-1425.

Iqbal, M. (1983). *An Introduction to Solar Radiation.* New York, USA: Academic Press.

Kämpfer, N. (2013). *Monitoring atmospheric water vapour, Ground-Based Remote Sensing and In-situ Methods.* New York, USA: Springer.

Kurtgoz, Y., Karagoz, M., & Deniz, E. (2017). Biogas engine performance estimation using ANN. *Engineering Science and Technology, an International Journal*, *20,* 1563-1570.

Lv, C., Xing, Y., Zhang, J., Na, X., Li, Y., Liu, T., … Wang, F.Y. (2018). Levenberg-Marquardt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System. *IEEE Transactions on Industrial Informatics*, *14*(8), 3436-3446.

Maghrabi, A., & Dajani, H. M. (2013). Estimation of precipitable water vapour using vapour pressure and air temperature in an arid region in central Saudi Arabia. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, *14*(1), 1-8.

Nunez, M. (1993). The development of a satellite-based insolation model for the tropical western Pacific Ocean. *International Journal of Climatology*, *13*(6), 607-627.

Pérez-Ramírez, D., Whiteman, D. N., Smirnov, A., Lyamani, H., Holben, B. N., Pinker, R., … Alados-Arboledas, L. (2014). Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites. *Journal of Geophysical Research: Atmospheres*, *119,* 9596-9613.

Phokate, S., & Atyotha, V. (2018). Determination of the aount of water vapor in the troposphere over Thailand using surface data. *Kasem Bundit Engineering Journal*, *8,* 364-372.

Reagan, J. A., Thome, K., Herman, B., & Gall, R. (1987). Water vapor measurements in the 0.94 micron absorption band: Calibration, measurements and data applications. *Electrical and Computer Engineering, 5,* 63-67.

Reitan, C. H. (1960). Distribution of precipitable water vapor over the continental United States. *Bulletin of the American Meteorological Society*, *41*(2), 79-87.

Senkal, O. (2015). Solar radiation and precipitable water modeling for Turkey using artificial neural networks. *Meteorology and Atmospheric Physics, 127,* 481-488.

Shastri, N., & Pathak, K. (2018). Estimation of perceptible water vapor of atmosphere using artificial neural network, support vector machine and multiple linear regression algorithm and their comparative study. *AIP Conferernce Proceeding, 1953*(1), 140114-1-140114-4.

Taylor, F. W. (2005). *Elementary climate physics*. England: Oxford University Press.

Yari, E., Ayoobi, A., & Ghassemi, H. (2014). Applying the Artificial Neural Network to Estimate the Drag Force for an Autonomous Underwater Vehicle. *Open Journal of Fluid Dynamics*, *4,* 334-346.

Yue, Y., & Ye, T. (2019). Predicting precipitable water vapor by using ANN from GPS ZTD data at Antarctic Zhongshan Station. *Journal of Atmospheric and Solar-Terrestrial Physics*, *191*, 1-9.