Current **A**pplied **S**cience and **T**echnology e-ISSN: 2586-9396

Curr. Appl. Sci. Technol. 2025, Vol. 25 (No. 6), e0262956

Research article

Estimation of Diffuse Fraction from Clearness Index and Atmospheric Parameters in Thailand

Danuch Phaisathit, Korntip Tohsing* and Serm Janjai

Solar Energy Research Laboratory, Department of Physics, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand

Received: 16 April 2024, Revised: 16 November 2024, Accepted: 3 December 2024, Published: 22 April 2025

Abstract

Diffuse solar radiation is a component of solar radiation that can be applied in modern and innovative solar energy technologies. However, there are a few ground monitoring stations in Thailand, leading to a lack of diffuse radiation data. Therefore, a monthly average model of diffuse solar radiation based on the relation between the diffuse fraction and the clearness index was developed using mathematical and machine learning methods: multiple linear regression (MLR), cross-validation (CV), and artificial neural networks (ANN). The ground-based data was measured at four stations in Thailand, namely the Nakhon Pathom station at Silpakorn University (13.82°N, 100.04°E), the Chiang Mai station in the north (18.78°N, 98.98°E), the Ubon Ratchathani station in the northeast (15.25°N, 104.87°E), and the Songkhla station in the south (7.20°N, 100.60°E), during 2016-2023. In addition, some atmospheric parameters such as the aerosol optical depth, precipitable water, total ozone column, cloud cover, air temperature, relative humidity, and wind speed provided by the Aerosol Robotic Network (AERONET) and the Thai Meteorological Department (TMD) were considered to increase the performance of the model. Statistical indicators were applied to investigate the performance of the models with measured data from the four stations. In the comparisons between the monthly average diffuse fraction estimated from the purposed model and the data obtained from the measurement, the ANN model showed outstanding performance, with a mean bias difference (MBD) of -0.224%, a root mean square difference (RMSD) of 7.33%, and an index of agreement (IOA) of 0.93. These results show that the ANN model demonstrated the best performance and can be applied for calculating diffuse solar radiation, which can be utilized for research and applications in solar energy, agriculture and building efficiency.

Keywords: solar radiation; diffuse fraction; clearness index; atmospheric parameters; machine learning

*Corresponding author: E-mail: tohsing_k@silpakorn.edu
https://doi.org/10.55003/cast.2025.262956

1. Introduction

The consumption of energy, specifically fossil fuels, is increasing by approximately 1.4% every single year and especially in the last few decades, the world has witnessed an exponential rise in its demand for energy (Salcedo-Sanz et al., 2015). On the contrary, the sources of petroleum, natural gas, and coal products, which are the main sources of fossil fuels, are decreasing, and furthermore they are non-renewable resources. Moreover, the combustion of fossil fuels and biomass causes greenhouse gas emissions that lead to global warming, air and marine pollution, climate abnormalities, and extreme weather events. The depletion of non-renewable energy sources and increasing energy demand are increasingly attracting worldwide attention to renewable energy. One of these is solar radiation, which can provide sustainable energy and reduce environmental pollution (Wang et al., 2018).

Global solar radiation is the sum of diffuse and direct solar radiation. Direct solar radiation is the portion of solar radiation that directly comes to the Earth's surface from the sun without any obstacles. On the other hand, diffuse solar radiation is the solar radiation scattered by atmospheric parameters. In this work, we focus on diffuse solar radiation, which can be applied in many solar radiation technologies, such as bifacial solar panels (Nussbaumer et al., 2019) and solar heat gains in buildings (Iqbal, 1981). Diffuse solar radiation can be measured using a pyranometer equipped with a shader for obstructing direct solar radiation. Previously, most of the work used a shadow band or shadow ring as a shader. However, there is a need to adjust the band or the ring almost every day. Moreover, the diffuse radiation from the pyranometer with the shadow ring must be corrected by using a ring correction algorithm (Drummond, 1956). In addition, this shadow ring correction factor depends on the sky conditions, e.g., clear sky, cloudy sky, and overcast sky (Siren, 1987). In this work, a shading ball and suntracker were used to solve these problems. This involved the incorporation of a fully automated machine tracking the sun. In Thailand, there are a few recent ground monitoring stations that can measure diffuse solar radiation, and there is still insufficient diffuse radiation data.

Since the 1970s, many researchers in the solar radiation field have found a specific relation between diffuse fraction and clearness index, which is expressed in the form of a logistic function and used as a main model for estimating diffuse radiation (Orgill & Hollands, 1977; Collares-Pereira & Rabl, 1979; Erbs et al., 1982; Reindl et al., 1990; Hove & Göttsche, 1999; Scarpa et al. 2017). Nevertheless, the model with only one predictor showed non-satisfactory performance. Subsequently, researchers included more atmospheric parameters, indicating that the performance of the model could be improved (Iqbal, 1979; Ridley et al., 2009; Abal et al., 2016; Zhou et al, 2019; Zhu et al., 2021). Furthermore, the relationship between the diffuse fraction and clearness index in many sites around the world has been evaluated using a mathematical approach (Kambezidis et al., 2021). As an increasing of photovoltaic (PV) system, the sky hemisphere was simulated to estimate the diffuse fraction for PV technology (Hofmann & Seckmeyer, 2017). A comparison of models delivered daily diffuse fraction in India was presented (Pandey & Katiyar, 2009). The NCEP/NCAR satellite data was also used for investigating the relation between the diffuse fraction and the clearness index (Babatunde & Vincent, 2021).

In general, the intensity of diffuse solar radiation depends on geography, latitude, longitude, and local climatology, as well as atmospheric parameters, which mostly consist of opaque objects that collide in the air, such as clouds, aerosols, and precipitable water. Empirical models were mostly used for estimating the diffuse fraction (all works mentioned above) because the mathematical models were simplified. However, the empirical models

were highly dependent on local data, which affected their performance when applied to different locations. Furthermore, machine learning is increasingly being used to improve the performance of models (Alam et al., 2008; Husain & Khan, 2021).

Therefore, the objectives of this study were to develop a monthly average diffuse fraction model using the relationship with clearness index and to try to include more atmospheric parameters in the initial equation to estimate diffuse fraction in the main regions of Thailand. The diffuse fraction can be converted back to the diffuse solar radiation, which is an important portion of the solar radiation and is used in many applications. For instance, diffuse radiation data is needed when designing solar panels for producing electricity under cloudy sky conditions or conditions where buildings obstruct direct solar radiation. In addition, architects and engineers integrate diffuse solar radiation data in building design to optimize natural light for reducing energy consumption. Furthermore, diffuse solar radiation plays an important role in agriculture for promoting crop growth, especially in areas with high cloud cover (Liu et al., 2022).

2. Materials and Methods

2.1 Solar radiation and atmospheric data

The initial model formulation in this work was focused on the relation between the diffuse fraction and clearness index. The necessary components included diffuse solar radiation, global solar radiation, and extraterrestrial solar radiation. The results of a previous study (Orgill & Hollands, 1977; Collares-Pereira & Rabl, 1979 ; Erbs et al., 1982) showed that only one predictor model delivered insufficiently accurate diffuse solar radiation. Then, we gathered more atmospheric parameters as predictors to improve the model.

In this work, global and diffuse solar radiation data was collected at four main regions of Thailand, namely the Silpakorn University, Nakhon Pathom station in the central region (13.82°N, 100.04°E), the Chiang Mai station in the northern part (18.78°N, 98.98°E), the Ubon Ratchathani station in the northeastern part (15.25°N, 104.87°E), and the Songkhla station in the southern part (7.20°N, 100.60°E), over the period of 1 January, 2016 to 31 December, 2023. The data was collected from various regions of Thailand to present the different climatic conditions and geographical features. The global solar radiation was measured by a CM21 pyranometer from Kipp & Zonen. Suntracker two-axis positioner (2AP) with a CM11 pyranometer from Kipp & Zonen equipped with a shading ball was used for measuring diffuse solar radiation. The four solar radiation monitoring stations and instruments are shown in Figure 1. The data was recorded and cumulated to hourly data using a DX2000 datalogger from Yokogawa. After that, monthly average solar radiation data were prepared for obtaining the diffuse fraction related to the clearness index and other parameters.

The spectral distribution of radiation arriving at the top of the atmosphere of the Earth (TOA) is very important in solar radiation applications such as satellites and spaceships. The extraterrestrial solar radiation deviates from the solar constant value for two reasons. The first is the variation in the radiation emitted by the sun itself. The second is the variation in distance between the Sun and the Earth due to the Earth's elliptical orbit. The daily extraterrestrial solar radiation can be calculated by the following equation (Iqbal, 1983):

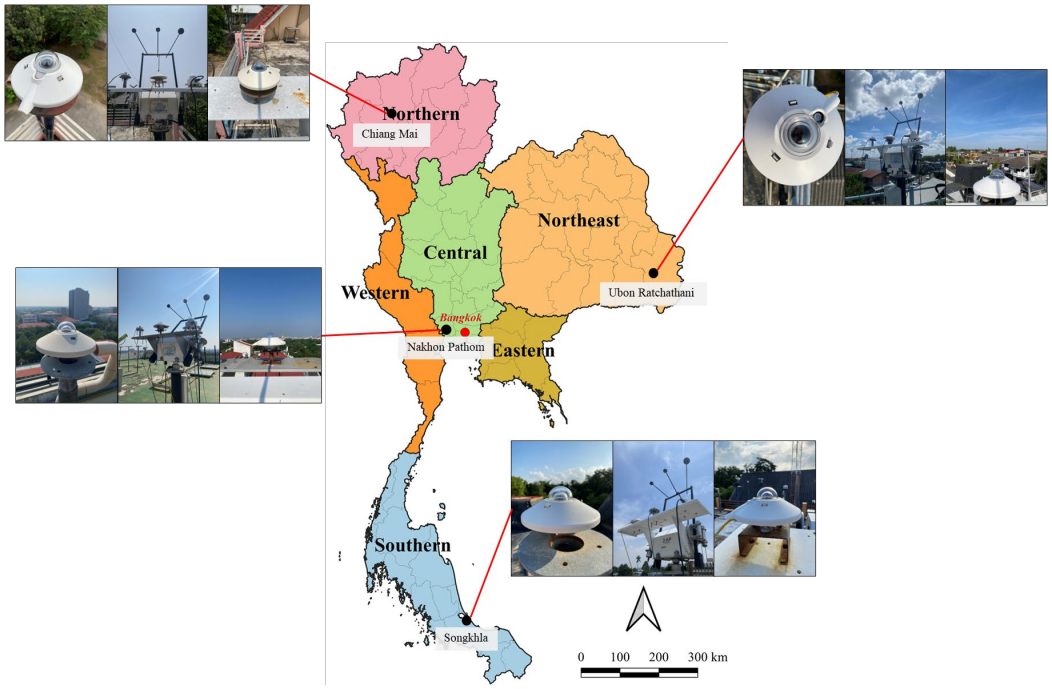

Figure 1. Global solar radiation (left), suntracker (middle), and diffuse solar radiation (right) instruments at four main regions of Thailand.

$$H_{0h} = \frac{24}{\pi} I_{sc} E_0 \int_0^{\omega_s} (\sin\delta \sin\phi) + (\cos\delta \cos\phi \sin\omega)\, d\omega$$

Or

$$H_{0h} = \frac{24}{\pi} I_{sc} E_0 [(\pi/180)\omega_s (\sin\delta \sin\phi) + (\cos\delta \cos\phi \sin\omega_s)] \quad (1)$$

where H_{0h} is the daily extraterrestrial solar radiation on a horizontal plane (W/m^2), I_{sc} is the solar constant (1367 W/m^2) (Iqbal, 1983), E_0 is the eccentricity correction factor (dimensionless), δ is the solar declination angle (°), ϕ is the latitude of the station (°), and ω_s is the sunset hour angle (°).

The monthly average extraterrestrial solar radiation is a summation of daily values and can be calculated using:

$$\bar{H}_{0h} = \frac{\sum_{i=1}^{N} H_{0h}}{N} \quad (2)$$

The eccentricity correction factor of the Earth is an adjustive function due to the variation of distance between the Sun and the Earth at different seasons. The distance is 1.014 AU in the summer solstice on June 22, 0.984 AU in the winter solstice on December 22, and 1 AU in the vernal and autumnal equinoxes on March 21 and September 22, respectively. The eccentricity correction factor is expressed by the equation as follows (Duffie & Beckman, 1980):

$$E_0 = 1 + 0.033 \cos\left[\frac{360 d_n}{365}\right] \quad (3)$$

where d_n is the day number; on January 1, $d_n = 1$ and February 1, $d_n = 32$.

The solar declination angle is an angle between the solar position and the celestial equator measured along the great circle containing the Sun and the celestial poles, calculated by the following equation (Iqbal, 1983):

$$\delta = 23.45° \sin\left[\frac{360}{365}(d_n + 284)\right] \quad (4)$$

The hour angle is the dihedral angle measured at the celestial pole between the observer's meridian (the path from north to south on the hemisphere or local meridian) and the solar meridian (the Sun's position). The sunset hour angle is the hour angle from midday to the evening. The sunset hour angle (ω_s) is calculated using the equation below (Iqbal, 1983):

$$\omega_s = \cos^{-1}(-\tan\phi \tan\delta) \quad (5)$$

Note that the sunrise hour angle is equal to the sunset hour angle except for an opposite sign.

The diffuse fraction is defined as the ratio of diffuse solar radiation to global solar radiation. The diffuse fraction varies between 0 and 1; the maximum shows that atmospheric constituents can reduce the intensity of global solar radiation to that of the diffuse. Whereas, the lowest value represents a clear sky condition. The diffuse fraction can be computed as follows:

$$\bar{F}_D = \frac{\bar{H}_{dh}}{\bar{H}_h} \quad (6)$$

where $\bar{F}_d$ is the monthly average diffuse fraction (dimensionless), $\bar{H}_{dh}$ is the monthly average diffuse horizontal solar radiation (W/m^2) and $\bar{H}_h$ is the monthly average global horizontal solar radiation (W/m^2).

The clearness index (K_T) is a parameter for indicating the atmosphere's clearness and is defined as the ratio of global solar radiation to extraterrestrial solar radiation and is expressed in equation (7). Theoretically, the clearness index varies in the range 0 to 1, which represents a high value under a clear sky and a low value under cloudy or overcast conditions.

$$\bar{K}_T = \frac{\bar{H}_h}{\bar{H}_{0h}} \quad (7)$$

The solar radiation reaching the Earth's surface is attenuated depending on the atmospheric or meteorological parameters. In this study, these parameters were also considered for the model improvement. The aerosol optical depth (AOD), precipitable water (w), and total ozone column (O_3) were obtained at the same stations from the Aerosol Robotic Network (AERONET) operated by the National Aeronautics and Space

Administration (NASA). Other meteorological data, such as air temperature (T_a), relative humidity (RH), cloud cover (cc), and wind speed (win) were provided by the Thai Meteorological Department (TMD).

The solar radiation data was maintained by expert meteorologists and researchers. However, it was necessary to control the quality of the solar radiation data to eliminate inaccurate measurements caused by the cosine response error of the sensors or other technical reasons. The hourly irradiance data was checked using the steps adapted from (Abal, 2019) as follows:

1) Global and diffuse solar irradiance was greater than 0 W/m².

2) The cosine of the solar zenith angle was more than 0.1219 (solar altitude greater than 7°).

3) Global horizontal irradiance was greater than 5 W/m².

4) Global horizontal irradiance was greater than diffuse horizontal irradiance.

5) Global horizontal irradiance was less than extraterrestrial solar irradiance on the horizontal plane.

6) Diffuse horizontal irradiance was less than or equal to half of extraterrestrial solar irradiance on the horizontal plane.

There were some short periods of missing data for various reasons, including malfunctioning instrumentation and power failure. Then, the month with daily data less than 15 days was rejected. After performing the quality check of the data with these steps, reliable solar radiation data was used to formulate the monthly average diffuse fraction model.

2.2 Modelling

In this study, three different methods were applied to obtain a model to estimate the diffuse fraction. These models were multiple linear regression as a mathematical method, cross-validation as a combined mathematical and machine learning method, and artificial neural networks as a full machine learning method. The data from the years 2016-2021 was normalized to eliminate redundant data and minimize data modification errors before being used to train the model. The other dataset of the years 2022-2023 was used for testing the models.

Multiple linear regression (MLR) is a statistical method used in many applications that attempts to determine the strength and character of the relationship between the diffuse fraction and the clearness index in the initial model. It is called simple regression or ordinary least squares (OLS). Linear regression is the most common form of this technique and is usually used to create empirical models for the diffuse fraction, as previously described. However, in this work, multiple predictors, such as the atmospheric parameters, were delivered into the model. Therefore, multiple linear regression was selected for mathematical modelling. MLR was used to formulate an equation relating the diffuse fraction and all other parameters, with coefficients and intercepts corresponding to each parameter. The coefficients in the MRL model were obtained by using the OLS method (Kutner et al., 2005). This delivers an empirical model for estimating the diffuse fraction.

The cross-validation (CV) method is a combination of mathematical and machine learning methods, meaning that the algorithms used to process the model are machine learning approaches that are then computed using standard regression techniques. The algoritm separates the data into the 3 datasets for training, validating and testing. The testing data is applied for testing the model with unseen diffuse fractions from the same year against other models. The cross-validation procedure started with separating the data

from 8 years during 2016-2023 into a modelling dataset for 75% of all data (the data from 6 years during 2016-2021) for training and validating. Then, the modelling dataset was divided into five subdatasets (20% for each subdataset): 80% for training and 20% for validating. During the application of the cross-validation technique, the dataset was randomly split 80-20 into sets of training and regression. Five iterations were applied, with each iteration using a different validating subdataset. This validating subset was used for fine-tuning the model. After that, the machine learning algorithm found hyperparameters to tune the model in each iterature and summarized the results from five iterations into single cross-validation model. The final step, the model was tested by using a testing dataset for 25% of all data (the data from 2 years during 2022-2023) from the first separation. After that the performance of the monthly average diffuse fraction model was evaluated by the cross-validation method (Irizarry, 2019). A schematic diagram of the cross-validation is shown in Figure 2.

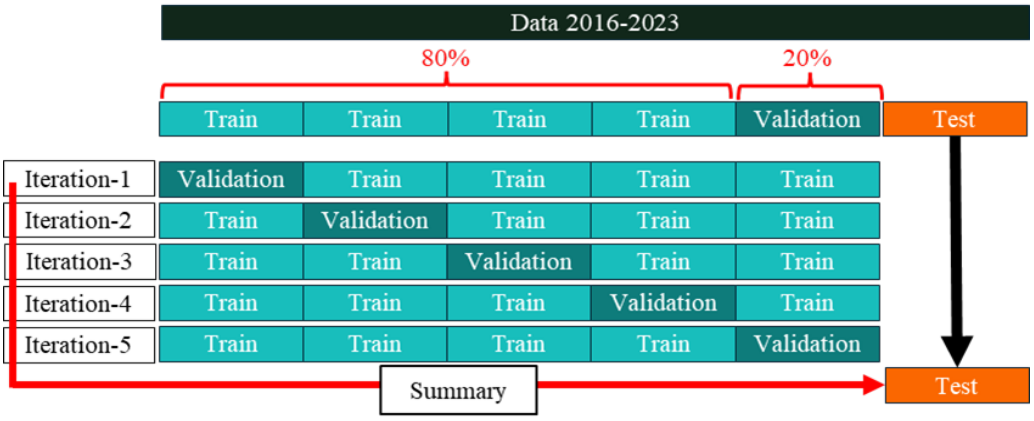

Figure 2. The structure of cross-validation model performed in this work.

In machine learning model, an artificial neural network (ANN) is a model inspired by the neuronal organs found in the biological neural networks in animal brains. The neural network model mainly includes an input layer, a hidden layer, and an output layer, as shown in Figure 3.

The input layers are responsible for receiving seven input parameters in this work. The hidden layers were operated using a non-linear function, which commonly uses a sigmoid function. Each node in the hidden layers was used for computing and delivering the weight and the bias. The weight adjusted the strength of the input parameter, while the bias was used for adjusting the output values to be closer to the actual value. The equation in each node of the hidden layers can be expressed as equation (8). The output layers produce and present the final network outputs resulting from the process performed by the neurons in the previous layers. The algorithm of the ANN starts with a feedforward process, where the input data is fed through the input layer and passed through multiple hidden layers. Each node in these layers calculates the output, weight, and bias. For the final stage, the output is sent to the output layer, where the result is presented as a desire value. Importantly, the weights and biases are fed back into the hidden layer, which is called a backpropagation. This mechanism recalculates the output based on the weight, bias, and the loss function. The multiple calculations involved in the backpropagation facilitate continuous improvement of the model (Dangeti, 2017).

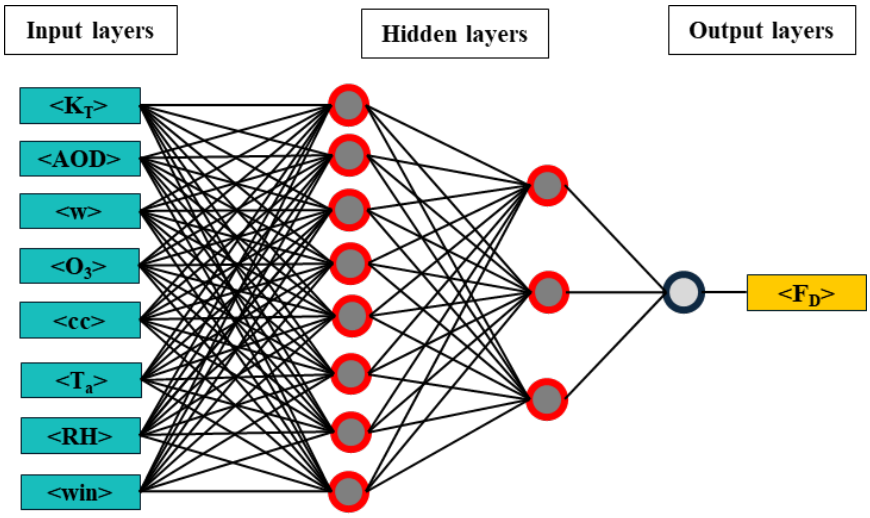

Figure 3. Schematic illustration of an artifitial neural network model.

$$y = F\left[\sum_{i=1}^{N} w_i x_i + b\right] \tag{8}$$

where y is the output (monthly average diffuse fraction), x_i is the input (all predictors), w_i is the weight for each predictor, b is the bias, and F is the sigmoid function.

The accuracy and performance of the presented models for estimating the monthly average diffuse fraction were evaluated and compared using three statistical indicators, namely the mean bias difference (MBD), the root mean square difference (RMSD) and the index of agreement (IOA). The MBD shows the bias of the diffuse fraction calculated from the model to the measured diffuse fraction. A positive value shows an overestimation of the model and a negative value presents an underestimation of the model. The MBD can be calculated from the formula:

$$MBD = \frac{\left[\frac{1}{N}\sum_{i=1}^{N} \bar{F}_{Dmea} - \bar{F}_{Dest}\right]}{\bar{F}_{Dest,ave}} \times 100\% \tag{9}$$

where $\bar{F}_{Dmea}$ is measured monthly average diffuse fraction, $\bar{F}_{Dest}$ is the estimated monthly average diffuse fraction, $\bar{F}_{Dest,ave}$ is the estimated monthly average diffuse fraction and N is the amount of data.

The RMSD shows the deviation of the measured diffuse fraction from the model and can be calculated by:

$$RMSD = \frac{\left[\frac{1}{N}\sum_{i=1}^{N} (\bar{F}_{Dmea} - \bar{F}_{Dest})^2\right]^{\frac{1}{2}}}{\bar{F}_{Dest,ave}} \times 100\% \tag{10}$$

The IOA is used for quantifying the degree of prediction error in models ranging between 0 and 1, where a value of 1 signifies a perfect match between model predictions

and observed data, while a value of 0 indicates no agreement. The IOA is calculated using the equation below (Willmott, 1981):

$$IOA = 1.0 - \frac{\sum_{i=1}^{N}(\bar{F}_{Dmea} - \bar{F}_{Dest})^2}{\sum_{i=1}^{N}(\,|\bar{F}_{Dest} - \langle\bar{F}_{Dmea}\rangle| + |\bar{F}_{Dmea} - \langle\bar{F}_{Dmea}\rangle|\,)^2} \quad (11)$$

3. Results and Discussion

The clearness index and seven atmospheric parameters from four stations in the main regions of Thailand during 2016-2021 were used as input for estimating the monthly average diffuse fraction for Thailand using the MLR, CV, and ANN models. The results from each model were then compared with those obtained from the measurements during the years 2022-2023.

3.1 MLR model

The mathematical method used to find the monthly average diffuse fraction, which is a multiple linear equation, is expressed in equation (12) and shown in Figure 4. The results show a reasonable agreement between both datasets, with a coefficient of determination (R^2) of 0.97, an MBD of 0.442%, an RMSD of 17.67%, and an IOA of 0.80. The scatter plot shows widespread data points. This might be because the relation between the diffuse fraction and clearness index is a logistic function and hardly fits the linear function.

$$\begin{aligned}\bar{F}_D = \; & 0.47 - 1.25\bar{K}_T - 0.07\frac{AOD}{AOD_{max}} - 0.03\frac{w}{w_{max}} + 0.53\frac{O_3}{O_{3,max}} - 0.08\frac{cc}{cc_{max}} \\ & + 0.8\frac{T_a}{T_{a,max}} + 0.26\frac{RH}{RH_{max}} + 0.10\frac{win}{win_{max}}\end{aligned} \quad (12)$$

3.2 CV model

The comparison of calculated and measured monthly average diffuse fractions for the cross-validation model (Figure 5) shows good statistical indicators. The coefficient of determination is equal to 0.99, the MBD is -0.051%, the RMSD is 9.77%, and the IOA is 0.82. The scatter plot demonstrates a slight deviation from the expected result because the CV model still uses regression to generate the model, leading to the same problem. However, the machine learning algorithm used in CV may solve the problem and discard trivial errors.

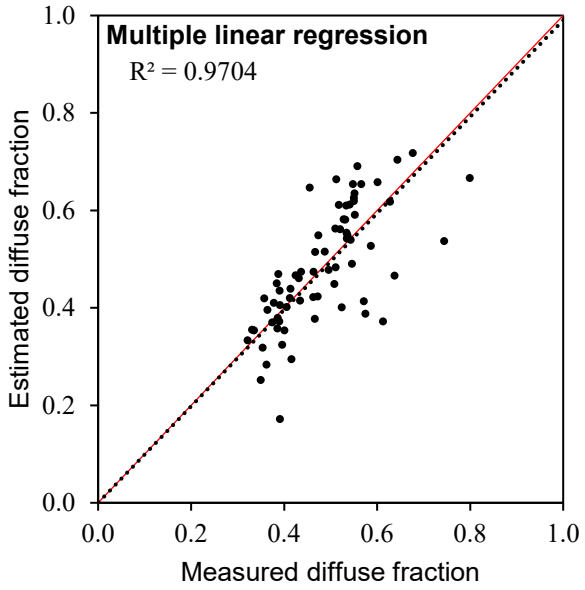

Figure 4. Comparison between estimates from the MLR model and measured monthly average diffuse fractions. The dashed line is the trend line, and the red line is the 1:1 line.

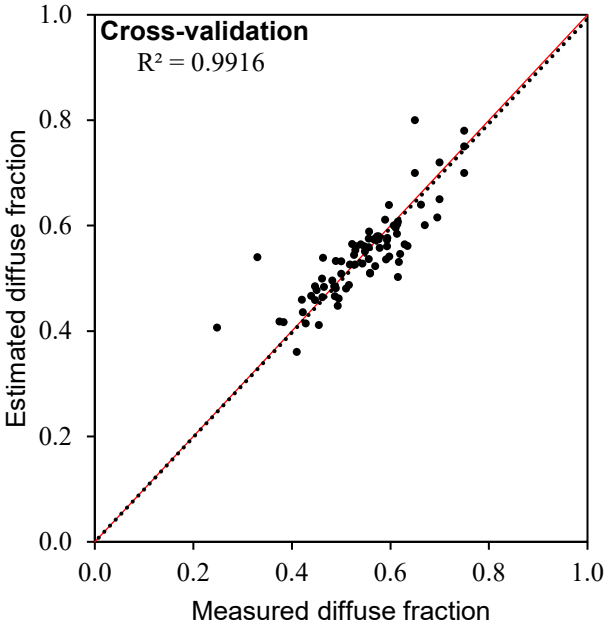

Figure 5. Comparison between estimates from the CV model and measured monthly average diffuse fractions. The dashed line is the trend line, and the red line is the 1:1 line.

3.3 ANN model

The resulting comparison from the artificial neural network model (Figure 6) presents a superfine agreement between the calculated and measured monthly average diffuse fractions with an excellent coefficient of determination equal to 0.99, the MBD of -0.224%, the RMSD of 7.33%, and the IOA of 0.93. The scatter plot depicts that the data plotted had a slight deviation. Because the ANN model uses a sigmoid function in each node to compute the diffuse fraction, the model is more suitable for the relationship between diffuse fraction and clearness index, including atmospheric parameters.

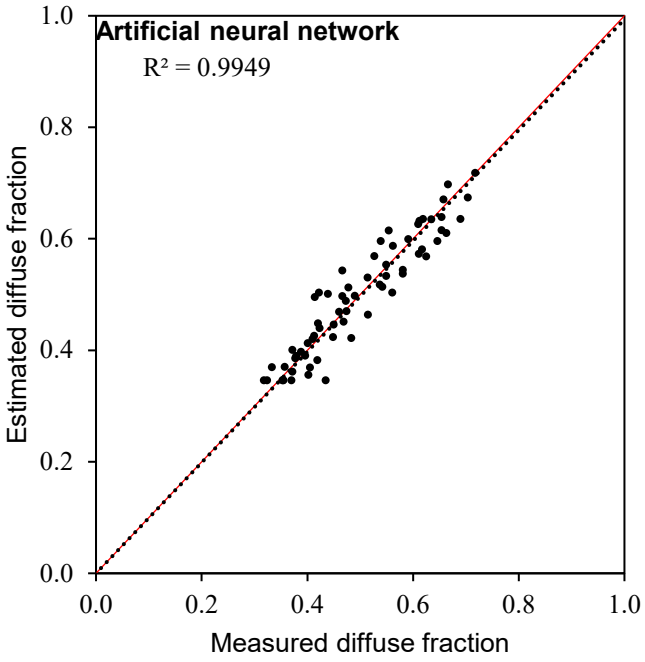

Figure 6. Comparison between estimates from the ANN model and measured monthly average diffuse fractions. The dashed line is the trend line, and the red line is the 1:1 line.

4. Conclusions

The aim of this work was to develop a model that could accurately estimate the monthly average diffuse fraction from the clearness index and other atmospheric parameters. Three models, MLR, CV, and ANN, were selected for this purpose. After that, the performance of the diffuse fraction model was investigated with measured data from four stations in the main regions of Thailand during 2022-2023 using two statistical indicators.

The comparison between calculated and measured monthly average diffuse fractions showed that all models could be applied to estimate the monthly average diffuse fraction. The diffuse fraction from multiple linear regressions had the biggest RMSD. The cross-validation presented a fine RMSD, but there were some deviations in the scatter plot. The artificial neural network delivered excellent results; an RMSD of 7.33% and an IOA of

0.93. It can be concluded that the ANN model demonstrated the best performance for obtaining the diffuse fraction.

The monthly average diffuse fraction from ANN model can be applied to estimate the monthly average diffuse solar radiation through the monthly average diffuse fraction in any area of Thailand. Futhermore, the model can fill in the monthly data gap and obtain the data for an area, in which no data measuring takes place. This can enhance the continuity and reliability of diffuse solar radiation data that can be utilized in research and applied in a range of applications.

5. Acknowledgements

We wish to thank the Department of Physics at Silpakorn University and Thai Meteorological Department (TMD) for providing instruments and data. We would also like to thank everyone in the Solar Lab, Silpakorn University, for their kind support.

6. Conflicts of Interest

The authors declare that they have no conflicts of interest.

ORCID

Danuch Phaisathit https://orcid.org/0009-0009-9795-7522

Korntip Tohsing https://orcid.org/0000-0002-7085-9238

References

Abal, G., Aicardi, D., Suárez, R. A., & Laguarda, A. (2016). Performance of empirical models for diffuse fraction in Uruguay. *Solar Energy*, 141, 166-181. https://doi.org/10.1016/j.solener.2016.11.030

Alam, S., Kaushik, S., & Garg, S. (2008). Assessment of diffuse solar energy under general sky condition using artificial neural network. *Applied Energy*, 86(4), 554-564. https://doi.org/10.1016/j.apenergy.2008.09.004

Babatunde, A., & Vincent, O. O. (2021). Estimation of hourly clearness index and diffuse fraction over coastal and sahel regions of Nigeria using NCEP/NCAR satellite data. *Journal of Energy Research and Reviews*, 7(4), 1-18. https://doi.org/10.9734/jenrr/2021/v7i430195

Collares-Pereira, M., & Rabl, A. (1979). Derivation of method for predicting long term average energy delivery of solar collectors. *Solar Energy*, 23(3), 223-233. https://doi.org/10.1016/0038-092x(79)90162-2

Dangeti, P. (2017). *Statistics for machine learning*. Packt Publishing Ltd.

Drummond, A. J. (1956). On the measurement of sky radiation. *Archiv Für Meteorologie Geophysik Und Bioklimatologie Serie B*, 7(3-4), 413-436. https://doi.org/10.1007/bf02242969

Duffle, J. A., & Beckman, W. A. (1980). *Solar engineering of thermal processes.* John Wiley & Sons, Inc.

Erbs, D., Klein, S., & Duffie, J. (1982). Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Solar Energy*, 28(4), 293-302. https://doi.org/10.1016/0038-092x(82)90302-4

Hofmann, M., & Seckmeyer, G. (2017). A new model for estimating the diffuse fraction of solar irradiance for photovoltaic system simulations. *Energies*, 10(2), Article 248. https://doi.org/10.3390/en10020248

Hove, T., & Göttsche, J. (1999). Mapping global, diffuse and beam solar radiation over Zimbabwe. *Renewable Energy*, 18(4), 535-556. https://doi.org/10.1016/s0960-1481(98)00782-4

Husain, S., & Khan, U. A. (2021). Machine learning models to predict diffuse solar radiation based on diffuse fraction and diffusion coefficient models for humid-subtropical climatic zone of India. *Cleaner Engineering and Technology*, 5, Article 100262. https://doi.org/10.1016/j.clet.2021.100262

Iqbal, M. (1979). Correlation of average diffuse and beam radiation with hours of bright sunshine. *Solar Energy*, 23(2), 169-173. https://doi.org/10.1016/0038-092x(79)90118-x

Iqbal, M. (1981). The influence of collector azimuth on solar heating of residential buildings and the effect of anisotropic sky-diffuse radiation. *Solar Energy*, 26(3), 249-257. https://doi.org/10.1016/0038-092x(81)90210-3

Iqbal, M. (1983). *An introduction to solar radiation*. Elsevier. https://doi.org/10.1016/b978-0-12-373750-2.x5001-0

Irizarry, R. A. (2019). *Introduction to data science: Data analysis and prediction algorithms with R*. Chapman and Hall/CRC.

Kambezidis, H. D., Kampezidou, S. I., & Kampezidou, D. (2021). Mathematical determination of the upper and lower limits of the diffuse fraction at any site. *Applied Sciences*, 11(18), Article 8654. https://doi.org/10.3390/app11188654

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. McGraw-hill.

Liu, P., Tong, X., Zhang, J., Meng, P., Li, J., Zhang, J., & Zhou, Y. (2022). Effect of diffuse fraction on gross primary productivity and light use efficiency in a warm-temperate mixed plantation. *Frontiers in Plant Science*, 13, Article 966125. https://doi.org/10.3389/fpls.2022.966125

Nussbaumer, H., Janssen, G., Berrian, D., Wittmer, B., Klenk, M., Baumann, T., Baumgartner, F., Morf, M., Burgers, A., Libal, J., & Mermoud, A. (2019). Accuracy of simulated data for bifacial systems with varying tilt angles and share of diffuse radiation. *Solar Energy*, 197, 6-21. https://doi.org/10.1016/j.solener.2019.12.071

Orgill, J., & Hollands, K. (1977). Correlation equation for hourly diffuse radiation on a horizontal surface. *Solar Energy*, 19(4), 357-359. https://doi.org/10.1016/0038092x(77)90006-8

Pandey, C. K., & Katiyar, A. (2009). A comparative study to estimate daily diffuse solar radiation over India. *Energy*, 34(11), 1792-1796. https://doi.org/10.1016/j.energy.2009.07.026

Reindl, D. T., Beckman, W. A., & Duffie, J. A. (1990). Diffuse fraction correlations. *Solar Energy*, *45*(1), 1-7. https://doi.org/10.1016/0038-092x(90)90060-p

Ridley, B., Boland, J., & Lauret, P. (2009). Modelling of diffuse solar fraction with multiple predictors. *Renewable Energy*, 35(2), 478-483. https://doi.org/10.1016/j.renene.2009.07.018

Salcedo-Sanz, S., Muñoz-Bulnes, J., Portilla-Figueras, J., & Del Ser, J. (2015). One-year-ahead energy demand estimation from macroeconomic variables using computational intelligence algorithms. *Energy Conversion and Management*, 99, 62-71. https://doi.org/10.1016/j.enconman.2015.03.109

Scarpa, F., Marchitto, A., & Tagliafico, L. (2017). Splitting the solar radiation in direct and diffuse components; insights and constrains on the clearness-diffuse fraction representation. *International Journal of Heat and Technology*, 35(2), 325-329. https://doi.org/10.18280/ijht.350213

Siren, K. (1987). The shadow band correction for diffuse irradiation based on a two-component sky radiance model. *Solar Energy*, 39(5), 433-438. https://doi.org/10.1016/s0038-092x(87)80062-2

Wang, L., Lu, Y., Zou, L., Feng, L., Wei, J., Qin, W., & Niu, Z. (2018). Prediction of diffuse solar radiation based on multiple variables in China. *Renewable and Sustainable Energy Reviews*, 103, 151-216. https://doi.org/10.1016/j.rser.2018.12.029

Willmott, C. J. (1981). On the validation of models. *Physical Geography*, 2(2), 184-194. https://doi.org/10.1080/02723646.1981.10642213

Zhou, Y., Liu, Y., Chen, Y., & Wang, D. (2019). General models for estimating daily diffuse solar radiation in China: Diffuse fraction and diffuse coefficient models. *Energy Procedia*, 158, 351-356. https://doi.org/10.1016/j.egypro.2019.01.101

Zhu, T., Li, J., He, L., Wu, D., Tong, X., Mu, Q., & Yu, Q. (2021). The improvement and comparison of diffuse radiation models in different climatic zones of China. *Atmospheric Research*, 254, Article 105505. https://doi.org/10.1016/j.atmosres.2021.105505