EFFECTIVENESS OF USING PREDICTED SOLAR RADIATION DATA IN BUILDING PERFORMANCE SIMULATION

Abstract

In real-time building performance simulation, real-time weather data is required. Solar radiation information is one of the most important weather parameters; however, it is not readily available. This paper presents an Artificial Neural Network algorithm that predicts global solar radiation based on easily accessible weather data; i.e., temperature and humidity. Diffuse and direct normal solar radiations are generated from predicted global solar radiation using the EnergyPlus™ weather converter program, which is also used as a weather packing tool to create the EPW weather file for EnergyPlus simulation. An office building is used as a case study for analysis. Three simulation scenarios are developed using: (1) complete Typical Meteorological Year (TMY3) file, (2) Limited_TMY3 file, and (3) Predicted _TMY3 file. This study analyzes the feasibility of using the predicted solar radiation data in the building performance simulation. The simulation results of three different scenarios are compared and analyzed.

Authors

Hany Gaballa and Soolyeon Cho
North Carolina State University

Keywords

Artificial Neural Network, heating and cooling loads, solar radiation prediction

Introduction

Predicting solar radiation has become a very important topic. The predicted solar radiation can be used to size PV power systems (Demirtas et al., 2012) and to analyze building energy performance. It can help increase agricultural productions (Torshizi & Mighani, 2017). Solar radiation could be predicted through various approaches (Zhang et al., 2017). Artificial Neural Network (ANN) is considered one of the most effective methods used in prediction (Qazi et al., 2015). It can be used for a computational simulation with complex, non-linear systems (Gani et al., 2016). The ANN model consists of three layers: input, hidden, and output layers.

Previous studies using the ANN model as a prediction method show that researchers used various parameters for the input layer, while the aim in this paper is to focus on easily measurable data from the Building Automation System (BAS) such as temperature and relative humidity. The hidden layer contains the calculations that happened in the hidden box to generate outputs, which is, in our case, hourly global solar radiation. An ANN model for predicting solar radiation was developed (Gaballa & Cho, 2019) within an acceptable error range according to ASHRAE Guideline 14 (ASHRAE, 2014). The Coefficient of Variance of the Root Mean Square Error CV(RMSE), Normalized Mean Bias Error (NMBE), and Coefficient of determination (R-squared) are used to calculate the uncertainties between the measured and predicted solar radiations.

Before running the building simulation, there is a need to make sure that weather data with limited number of parameters, including solar radiation, temperature, and humidity, will give a reasonable result compared to using the original weather data files. The EnergyPlus weather converter program was used to generate two different EnergyPlus weather (EPW) data files. The first is the limited TMY3 weather data files (Limited_TMY3) by using only five parameters: dry-bulb temperature, relative humidity, direct normal radiation, and global and diffuse horizontal radiation. The second is the predicted TMY3 weather data files (Predicted_TMY3) using the output from the ANN model, which is the hourly global solar radiation. The three weather data files were used to run the simulation using the EnergyPlus simulation program for an office building to show the differences between using original, limited, and predicted TMY3 weather files.

Literature Review

By looking at the previous studies that have a mutual interest area, it can be divided into two main categories. First, papers using the ANN model to predict solar radiation and the kind of similar articles discussed before and that can be viewed here (Gaballa & Cho, 2019). Second, papers using different methods to run a real-time building energy simulation. Through the literature review, there was a lack of articles using the ANN model as a method to develop a real-time weather file and then use it in real-time energy performance prediction. Otherwise, they use onsite instruments or sensors to develop weather files. In some other papers, the author used different models like the Seo model, which requires more than six parameters to generate the solar radiation data; some of these parameters are not readily available. In the following, some of these papers are presented.

Pang et al. (2011) used a Building Control Virtual Test Bed (BCVTB) as a software platform to provide data linkage between the EnergyPlus model and a database. Also, the authors used building automation and control network (BACnet) to make measurements accessible to EnergyPlus. The main part of this article includes that the author installed sensors to calculate the needed data for simulation like dry-bulb temperature, relative humidity, wind speed, wind direction, direct normal solar radiation, and diffuse solar radiation. The authors updated their work and published it later in three different articles (Pang et al., 2012; O'Neill et al., 2014; Pang et al., 2016). The new version of their work is mainly about developing a new platform but still using the same way of getting weather parameters.

Kwak and Huh (2016) also aimed to develop a method of real-time building energy simulation. First, the author used the Korea Meteorological Administration (KMA), which offers data on dry-bulb temperature, relative humidity, cloud cover, wind direction, and wind speed every three hours. After that, the authors used an equation to provide data every hour. About solar radiation data, the Seo model is used based on the hourly weather elements mentioned before. Then direct and diffuse solar radiation are calculated based on extraterrestrial horizontal irradiance, clearness index, and some other coefficients. Finally, the weather data should be ready for simulation. A similar method was presented by Lee et al. (2017), using KMA first and, after that, using the Seo model for solar radiation prediction.

Xu et al. (2017) used the real-time meteorological data observed by the Institute of Geographic Sciences and Natural Resources Research to run the real-time energy consumption simulation. Asadi et al. (2019) aimed to develop an algorithm to calibrate the real-time energy simulation model by looking at occupants, equipment, and lighting schedule. The authors used the Actual Meteorological Year (AMY) file to generate the real-time weather files needed.

SUMMARY OF LITERATURE REVIEW

It was found that no study uses the ANN model to predict solar radiation using only readily available weather data such as temperature and relative humidity, or even uses this prediction method to run real-time building performance simulation. The main goal of this paper is to fill this gap; in the following section, the methodology used will be discussed in detail to reach this goal.

Methodology

The weather file used in this paper is Typical Meteorological Year (TMY) weather data files for Raleigh-Durham International Airport weather station (Raleigh-Durham.Intl. AP_TMY3).

EnergyPlus uses EPW weather file format to run the simulation, which can be developed using a weather packing tool like EnergyPlus weather converter program. The EnergyPlus weather converter program accepts a Comma Separated Value (CSV) file format as an input. If some data is missing, a specific number should be written according to each parameter, so the program understands it is missing. Some of these parameters are dependent on others. For example, if we have dry-bulb temperature and relative humidity, the dew point temperature will be calculated automatically. By looking at solar radiation data, there are three components: global, direct, and diffuse solar radiations. Through the Perez model (EERE, 2018), which is run automatically during the EPW file packing process in the EnergyPlus weather converter program, direct and diffuse solar radiation can be calculated from global solar radiation. This process is written with more detail in the EnergyPlus auxiliary programs document. From here, the ANN algorithm takes place to predict global solar radiation using hourly input data; temperature, relative humidity, solar zenith angle, and time. After the prediction process, the EnergyPlus weather converter program is used to calculate direct and diffuse solar radiation from the predicted global solar radiation. Then, a new EPW file is packed using the same program, called (Predicted_TMY3).

EnergyPlus weather converter program is also used to generate another EPW file to see how a weather file with limited data can perform in simulation compared to the original weather file with full data. This time the EPW file is generated by only using temperature, relative humidity, and solar radiation data from TMY3, in another way by deleting any other data that cannot be measured or predicted easily like wind speed, wind direction, cloud cover, etc. The file is named as (Limited_TMY3). For more clarity, Table 1 shows the contents of the three weather files used in this paper. An abbreviation will be given for each case: C_1, C_2, and C_3 for original, limited, and predicted TMY3, respectively.

Case	**Dry-bulb temp.**	**Relative humidity**	**Global horizontal solar radiation**	**Direct normal solar radiation**	**Diffuse horizontal solar radiation**	**Dew point temp.**	**Horizont. Infrared Radiation Intensity**	**Wind direction**	**Wind speed**	**Extraterrestrial radiation data**	**Cloud cover**	**Others**
Case 1 [C_1] Original TMY3	√	√	√	√	√	√	√	√	√	√	√	√
Case 2 [C_2] Limited TMY3	√	√	√	√	√	C	C	C				
Case 3 [C_3] Predicted TMY3	√	√	P	C	C	C	C	C				

C: Calculated through the EnergyPlus weather converter program
P: Predicted from the ANN model

Table 1: The three weather cases contents.

ANN ALGORITHM

An ANN model is developed to predict hourly global solar radiation using Python 3.6, which is considered one of the commonly used programming languages for machine learning (Guo, 2014). The TMY3 data was divided into four seasons. As an initial step, only summer data was used in this paper as training (80% of summer data) and testing (20% of summer data) in the ANN model. The process of the ANN algorithm goes through three layers. ANN needs inputs to predict the output through calculations happening in the hidden layer. This study used only the readily available meteorological data for the input layer; i.e., temperature, relative humidity, solar zenith angle, and time every hour. The output layer is the hourly global solar radiation, as shown in Figure 1. The hidden layer consists of hidden neurons. It varies depending on each case. Figure 1 shows the architecture of the ANN model and how it works, starting from the initial randomized weights assigned to each parameter in the input layer to be connected to a hidden neuron in the hidden layer.

After this process, the sigmoid function should be applied to get a new number. The same process is repeated but now between the hidden and output layer with different weights. The whole process is called Feed-Forward Neural Network (FFNN), starting from the input layer to reach the output layer. After that, the error rate is calculated, and then another process happens to update all weights, which is called Back-Propagation (BP). The number of the whole loop, including FFNN and BP that repeat until reaching the minimum error rate, is called the number of epochs. There is an essential parameter in the BP process called learning rate, which is used to settle the changes in the weights at the end of each epoch (Rezrazi et al., 2016).

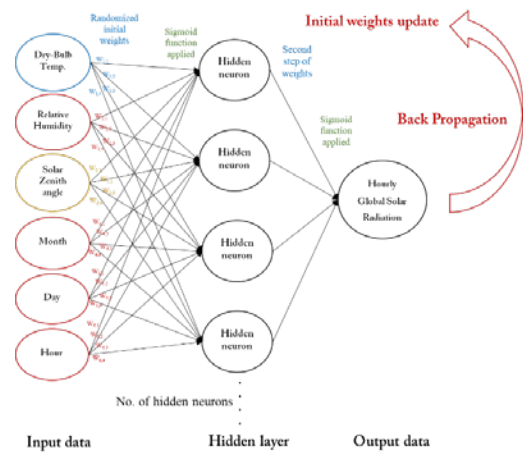

Figure 1: ANN architecture for global solar radiation prediction.

In this paper, ANN hidden process has three different variables: the number of hidden neurons, the number of epochs, and the learning rate. An optimization process is conducted to give a specific value for each one of the three variables depending on which season the simulation is done to provide the most accurate output results. The accuracy of the ANN model is verified according to ASHRAE Guideline 14 by calculating CV(RMSE), NMBE, and R-squared to find the error differences between the measured and predicted global solar radiation data. The results showed that error differences are within ASHRAE Guideline 14 recommendations (ASHRAE, 2014); the NMBE value less than 10% and CV(RMSE) value less than 30% according to hourly calibration data, more details about this ANN model can be found here (Gaballa & Cho, 2019).

By calculating the error differences for the predicted global horizontal radiation, CV(RMSE) and NMBE between $C_{1,2}$ and C_3 found to be 23% and -0.5%, respectively. Figures 2 and 3 show a comparison between the two cases and the difference between them.

Figure 2 shows the relation between global solar radiation and hours from June 21st at 4 pm until July 9th at 9 am on the horizontal axis, while Figure 3 shows the solar radiation and its relation to the dry-bulb temperature.

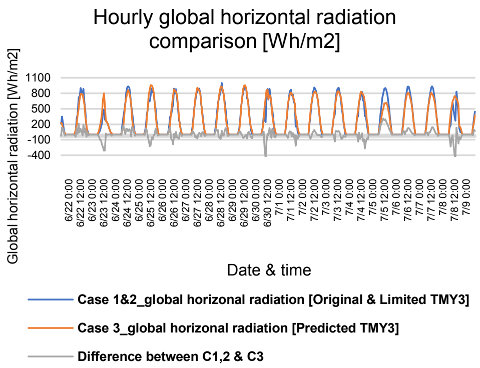

Figure 2: Hourly global horizontal solar radiation comparison between $C_{1,2}$ & C_3.

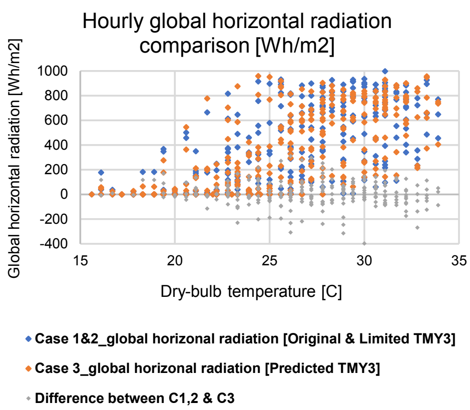

Figure 3: The relation between hourly global horizontal solar radiation and temperature.

CASE STUDY BUILDING SIMULATION

A case study office building is selected in the Research Triangle Park (RTP), NC, located five miles away from Raleigh-Durham (RDU) International Airport (Figure 4). This location is climate zone 4A. The building consists of three floors, as shown in Figure 4, with a total area of 43,265 ft^2, including the conditioned space of 42,332 ft^2. There are 29 thermal zones on three floors: lake level, 1st floor, and 2nd floor with 11, 9, and 9 thermal zones, respectively. The south elevation is assigned to be adiabatic as it is attached to another building. An 80-foot tall building is located 30 feet away from the east elevation, so a shading surface group was drawn with the same dimensions and distance in the model. The window to wall ratio is 42%, with a total glazing area of 18,245 ft^2.

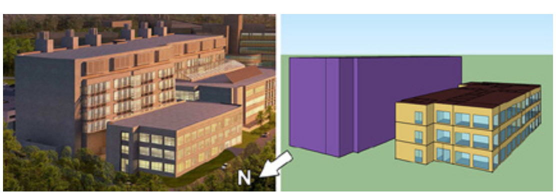

Figure 4: Case study building, Durham, NC.

The EnergyPlus program V8.9.0 is used for simulation modeling. Only load calculations are performed. The only difference is changing the weather data files used in simulation; C_1, C_2, and C_3. A comparison for the whole year could be made between original and limited weather data files. On the other hand, we can compare only 20% of summer data when using the predicted weather data files (C_3). The 20% represents the testing data used in the ANN model, which includes the predicted hourly global solar radiation.

Results

Heating and cooling loads are selected as indicators to make a comparison between the three cases. Figure 5 shows a comparison of the entire year for heating and cooling loads between C_1 and C_2. The total cooling loads found to be as follows; C_1 is 1,042,497 kBtu/yr while C_2 is 1,048,852 kBtu/yr. The total heating loads for C_1 is 259,272 kBtu/yr, while for C_2 is 268,120 kBtu/yr. The results show only a 0.6% difference for the total cooling loads and a 3.3% difference for the total heating loads.

Next step, a comparison was made between C_1, C_2, and C_3 for a specific period. As mentioned before, the ANN results show just 20% of the summer season data, which reflects the predicted data from the ANN model. So the comparison took place from June 21st at 4 pm until July 9th at 9 am, which means 426 hours. Figure 6 shows the total cooling loads comparison between the three cases for the 426 hours of the summer season.

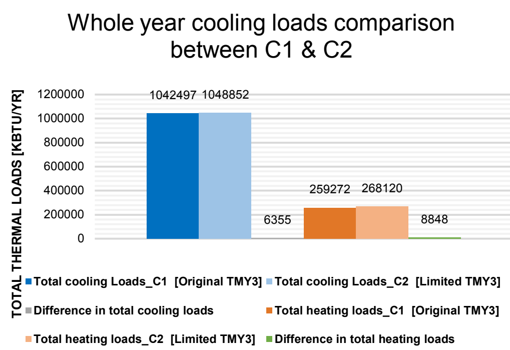

Figure 5: Heating & Cooling loads comparison between C_1 and C_2 for the whole year.

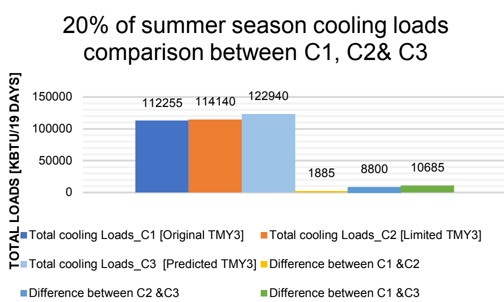

Figure 6: Cooling loads comparison between C_1, C_2 and C_3 for 20% of the summer season.

The cooling loads are 112,255 kBtu for C_1, 114,140 kBtu for C_2, and 122,940 kBtu for C_3. The result shows a 1.65% difference between C_1 and C_2, a 7.15% difference between C_2 and C_3, and an 8.7% difference between C_1 and C_3 for the total cooling loads for this specific period.

Period	Cases	CV(RMSE)	NMBE
Whole year	C_1 vs. C_2 Heating Loads	13.9%	3.3%
	C_1 vs. C_2 Cooling Loads	4.9%	0.6%
20% of summer season	C_1 vs. C_2 Cooling Loads	3.2%	1.7%
	C_1 vs. C_3 Cooling loads	14.9%	8.7%
	C_2 vs. C_3 Cooling loads	14.3%	7.2%

Table 2: CV(RMSE) and NMBE results between the three weather cases.

Figure 7 presents the cooling loads' comparison graph using the three weather files for the same period. By looking at the graph, it shows a good correlation between the three cases. To measure the accuracy of using predicted weather files in simulation, CV(RMSE) and NMBE were calculated. Table 2 shows these results, which meet ASHRAE Guideline 14 (ASHRAE, 2014) requirements according to hourly data comparison.

Analysis and Discussion

The simulation between C_1 and C_2 shows small differences, and this is logically related to the missing data in the weather data files. Missing extraterrestrial solar data for the whole year affected the simulation results in winter, which made the heating loads increase by 3.3%. In summer, the cooling loads increased by only 0.6%, and this might be related to the missing wind data.

The previous results show that using limited data will not severely affect the simulation results, which verifies that using ANN models to predict global solar radiation might also give good results compared to measured data.

By running the simulation using the predicted data (C_3), it shows a little bit more differences than using the limited data (C_2). These differences come to two main reasons. First, the prediction accuracy of the ANN model output, which is the hourly global solar radiation. Second, the accuracy of the Perez model used to calculate the direct and diffuse solar radiation data from the predicted global solar radiation data. The difference shows an 8.7% increase in the total cooling loads for 426 hours. After calculating CV(RMSE) and NMBE, it proves that using the predicted weather files gives an acceptable error range according to ASHRAE Guideline 14.

Now it is time for building operators to give their opinion about the accuracy of the predicted weather files used to run the simulation and if this error percentage is acceptable or not for the real-life operation situation. By proving the accuracy of this process, it will help designers to have a good vision looking at the real-time building performance prediction and make a comparison with the actual case.

Conclusion and Future Studies

ANN technique is considered one of the most reliable methods used in prediction. ANN model was developed using the easily measured data in buildings to predict hourly global solar radiation. Only summer data from TMY3 was used; 80% for training and 20% for testing. Using the EnergyPlus weather converter program, a new weather data file was generated using the predicted global solar radiation from the ANN model to be called Predicted_TMY3 (C_3).

To see the effectiveness of using only data that can be easily measured or predicted, new weather data file developed using only limited data called Limited_TMY3 (C_2). This weather file also is developed using the EnergyPlus weather converter program.

The simulation shows some differences between the three cases—these differences related to the prediction error rate and missing parameters in the weather data files as well.

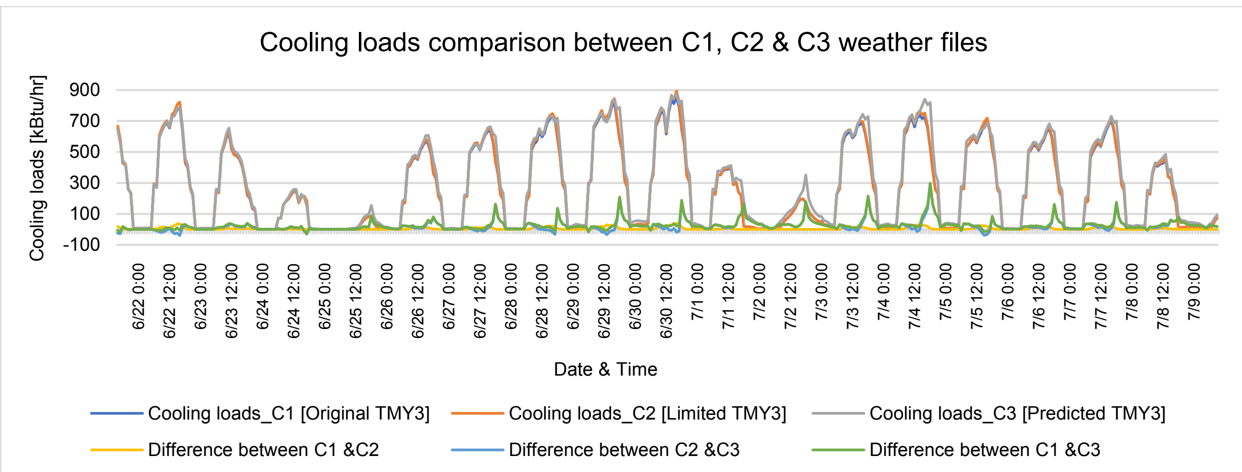

Figure 7: Hourly Cooling Loads comparison between C_1, C_2 and C_3.

In this study, TMY3, which is used to run simulations, represents historical data. For future studies, real-time weather files will be developed using 2017 data to get real-time building performance predictions. By doing so, it helps to keep an eye on the building performance and see if the building has a low efficiency at a specific time than the way it should operate. Also, it will help the building operators find where the building problem is in real-time and fix it immediately, consequently reducing the building energy loss.

References

Asadi, S., Mostavi, E., Boussaa, D., & Indaganti, M. (2019). A building energy model calibration using automated optimization-based algorithm. *Journal of Energy and Buildings, 198,* 106–114.

ASHRAE Guideline 14. (2014). American Society of Heating, Refrigerating and Air-Conditioning Engineers. *Measurement of energy, demand, and water savings.*

Demirtas, M., Yesilbudak, M., Sagiroglu, S., & Colak, I. (2012). Prediction of solar radiation using meteorological data. *2012 International Conference on Renewable Energy Research and Applications (ICRERA),* 1–4.

EERE (Department of Energy Efficiency and Renewable Energy). (2018). Ground heat transfer in EnergyPlus. Documentation, EnergyPlus™ Version 8.9 - Auxiliary Programs, 36.

Gaballa, H., & Cho, S. (2019). Prediction of Hourly Solar Radiation using Temperature and Humidity for Real-Time Building Energy Simulation. *Journal of Physics: Conference Series (JPCS).*

Gani, A., Mohammadi, K., Shamshirband, S., Khorasanizadeh, H., Danesh, A. S., Piri, J., & Zamani, M. (2016). Day of the year-based prediction of horizontal global solar radiation by a neural network auto-regressive model. *Journal of Theoretical and Applied Climatology, 125,* 679–689.

Guo, P. (2014). Python is now the most popular introductory teaching language at top US universities. *BLOG@ CACM, July, 47.*

Kwak, Y., & Huh, J. (2016). Development of a method of real-time building energy simulation for efficient predictive control. *Journal of Energy Conversion and Management, 113,* 220–229.

Lee, J., Lee, S., Kim, J., Song, D., & Jeong, H. (2017). A middleware platform for the validation and utilization of short-term weather forecast data for office. *Journal of Energy and Buildings, 149,* 192–203.

O'Neill, Z., Pang, X., Shashanka, M., Haves, F., & Bailey, T. (2013). Model-based real-time whole building energy performance monitoring and diagnostics. *Journal of Building Performance Simulation, 7,* 83–99.

Pang, X., Bhattacharya, P., O'Neill, Z., Haves, P., Wetter, M., & Bailey, T. (2011). Real-time Building Energy Simulation using EnergyPlus and the Building Controls Virtual Test Bed. *The 12th International Building Performance Simulation Association Conference (IBPSA),* 12, 2890–2896.

Pang, X., Nouidui, T., Wetter, M., Fuller, D., Liao, A., & Haves, P. (2016). Building Energy Simulation in Real Time through an Open Standard Interface. *Journal of Energy and Buildings, 117,* 282–289.

Pang, X., Wetter, M., Bhattacharya, P., & Haves, P. (2012). A framework for simulation-based real-time whole building performance assessment. *Journal of Building and Environment, 54,* 100–108.

Qazi, A., Fayez, H., Wadi, A., Raj, R. G., Rahim, N. A., & Khan, W. A. (2015). The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. *Journal of Cleaner Production, 104,* 1–12.

Rezrazi, A., Hanini, S., & Laidi, M. (2016). An optimization methodology of artificial neural network models for predicting solar radiation: a case study. *Journal of Theoretical and Applied Climatology, 123,* 769–783.

Torshizi, M.V., & Mighani, A.H. (2017). The Application of Solar Energy in Agricultural Systems. *Journal of Renewable Energy and Sustainable Development, 3,* 234–240.

Xu, C., Li, Y., Jin, X., Yuan, L., & Cheng, H. (2017). A Real-Time Energy Consumption Simulation and Comparison of Buildings in Different Construction Years in the Olympic Central Area in Beijing. *Journal of Sustainability, 9.*

Zhang, J., Zhao, L., Deng, S., Xu, W., & Zhang, Y. (2017). A critical review of the models used to estimate solar radiation. *Journal of Renewable and Sustainable Energy Reviews, 70,* 314–329.