

Accuracy Assessment of Typical Meteorological Year Data for a Photovoltaic System using a Bootstrap Method

Uwineza Laetitia¹ · Kim Hyun-Goo^{2†} · Kim Chang Ki² · Kim Boyoung³ · Kim Jin-Young³

¹Doctoral Student, Department of Renewable Energy Engineering, University of Science & Technology

²Principal Researcher, New & Renewable Energy Resource Map Laboratory, Korea Institute of Energy Research

³Senior Researcher, New & Renewable Energy Resource Map Laboratory, Korea Institute of Energy Research

[†]Corresponding author: hyungoo@kier.re.kr

Abstract

The typical meteorological year (TMY) is a key element in the long-term performance predictions of photovoltaic systems. However, there is a loss of some hourly measured weather data during the formation of the TMY dataset owing to a lack of observations. Therefore, various statistical techniques or satellite observations can be utilized to address this limitation, and these interpolated data may represent some degree of uncertainty that can influence the accuracy of the TMY data. This may in turn affect the long-term planning reliability of the photovoltaic systems. Therefore, a bootstrap method was used in this study to evaluate the accuracy of TMY datasets in the prediction of long-term photovoltaic yields. The electricity production uncertainties obtained through a simulation conducted using this TMY were calculated with a confidence level of 95% for each resulting value as validated based on the long-term average electricity yield. The results show that the bootstrap method provides valid and more useful information than the deterministic approach. Therefore, it is the best method for quantitatively analyzing uncertainty in TMY datasets. The results described herein can serve as a guide for risk management strategies and other business decisions related to solar energy projects. In addition, the results can help planners achieve greater confidence when applying TMY data to feasibility studies.

Keywords: Photovoltaic system, Typical meteorological year, Bootstrap method

1. Introduction

The design of a photovoltaic generation plant system strongly depends on the long-term weather dataset used to predict the long-term performance of the system. To achieve this goal, it is necessary to have a typical meteorological year or long-term weather dataset of the most frequent conditions at the relevant location, which allows a simulation of the solar power production supplied by the system over the long term. However, using a long-term weather dataset is time-consuming, inconvenient, and expensive compared to typical meteorological year¹⁾. To mitigate this problem, typical

 OPEN ACCESS



Journal of the Korean Solar Energy Society
Vol.41, No.4, pp.115-129, August 2021
<https://doi.org/10.7836/kses.2021.41.4.115>

pISSN: 1598-6411

eISSN: 2508-3562

Received: 7 April 2021

Revised: 1 June 2021

Accepted: 7 June 2021

Copyright © Korean Solar Energy Society

This is an Open-Access article distributed under the terms of the Creative Commons Attribution NonCommercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

meteorological year datasets have been used with simulation tools such as PVSyst software to obtain photovoltaic output results quickly²⁾.

Over the past few decades, many statistical methods have been used to calculate the time-series of the meteorological parameters of typical meteorological dataset. Lund's method is widely used in Europe³⁾, while the Finkelstein-Schafer statistical method is well known in American publications. These statistical methods require a long-term measured weather dataset to generate TMY data for a given location⁴⁾. Typical meteorological year datasets consist of 12 typical months determined by considering various meteorological elements, including global horizontal radiation and direct normal radiation as well as dry bulb temperatures, dew point temperatures, and wind speeds. However, some hourly measured weather datasets are found to be missing during the formation of the TMY due to a lack of observations; this gap can then be filled by statistical methods, satellite observations, and simulation algorithms⁵⁾. Therefore, those interpolated datasets can reduce the accuracy of TMY datasets⁶⁾. For this reason, the validation of TMY datasets is required, with an acceptable level of accuracy. In addition, stakeholders need TMY datasets with high accuracy to use them for solar project planning, as they translate directly into forecasts of electricity generation and revenues.

Previous studies⁷⁾ have used deterministic methods to validate TMY data, though this method is limited and ignores the influence of uncertainty on the estimation of the TMY dataset. In such studies, error metrics and deterministic predictions were used to calculate error, simplified mean, variance, and standard deviation values, which are limited and cannot reflect the inherent integrated uncertainty⁸⁾. Although these studies used deterministic methods, the corresponding statistical analyses generated unreliable means of values, suggesting that the TMY is not an appropriate tool for assessing individual sites to secure investments in long-term planning models and that risk levels associated with solar energy projects can be very high⁹⁾.

Table 1 summarizes different studies of validation of TMY by using traditional statistics. However, for the practical design of renewable energy projects, it seems more appropriate to use probabilistic models than the deterministic approach, which is good for only one point in time¹⁰⁾. Therefore, we adopted the bootstrap method as a probabilistic method to model the uncertainty of TMY data. As a result, the electrical energy produced by PV power generation system was estimated after the simulation of each TMY with prediction interval. In addition, this information is useful for long-term performance of the PV system in absence of the measurement of hourly weather data series, and further will enable the planners to make more informed decision¹¹⁾. The bootstrap method is frequently applied to make a probabilistic forecast in

medicine, water assessment, and sustainable agriculture¹²⁾. The main advantage of this method is that it is simple to use and generates samples from the given data. However, when the sample size n is very small, bootstrapping may fail.

Table 1. Summary of previous studies

Period	Location	Methods	Validation	Findings	References
1975-2008	Nigeria	Filkenstein Schafer method	Deterministic approach	TMY showed a good fit with long-term weather data with a mean percentage error of less than 2%.	Ohunakin et al. (2013) ¹³⁾
1999-2015	United State	Filkenstein Schafer method	Kolmogorov Smirnov Integral test (KSI)	TMY datasets were less representative to the long-term energy production for Concentrated-PV systems.	Realpe et al. (2016) ¹⁴⁾
1971-2000	Poland	Filkenstein Schafer method	Multi-criteria analysis method	Finding showed that TMY were different from the long term average solar radiation. Author critiqued that TMY should not be considered as an appropriate alternative for the long-term measured data.	Nelken et al. (2015) ¹⁵⁾
2003-2014	Portugal	Filkenstein Schafer method	Global Performance Index (GPI)	TMY showed the closest results to the long-term weather datasets.	Abreu et al. (2018) ¹⁶⁾

Despite several studies validating TMY data at different locations, no study has focused on the bootstrap method to quantify the reliability or uncertainty of TMY datasets. In addition, the deterministic predictions used in the aforementioned studies are based on laws of physics that are too optimistic and provide an unreliable mean value that can be useless or even misleading. Furthermore, the impact of uncertainty in the estimates of the means and variance of TMY datasets was not taken into account despite the importance of providing information about uncertainty to those making long-term investment decisions related to solar energy systems. Therefore, this study assesses the bootstrap method for a quantitative analysis of the uncertainty of TMY datasets. The results will likely be widely used in feasibility studies and in policy decision-making activities in the solar power business in the future¹⁷⁾.

There are three main objectives of this paper. First, we seek to estimate long-term PV electricity production levels from TMY and long-term measurement datasets. Second, the confidence interval for long-term electricity production and TMY electricity production will be established using the bootstrap method. Finally, the accuracy of TMY in estimating long-term PV electricity production levels will be

evaluated. The validation method proposed in this study plays an important role in quantifying the prediction interval of the TMY in feasibility studies of solar energy projects. In addition, such information can help planners and investors to measure and manage risks and can assist with many decisions using TMY datasets in renewable projects.

2. Meteorological Data

To validate the TMY datasets, 31 sites were selected as study areas (Fig. 1) to analyze the effects of uncertainty in the TMY data. The TMY data (Finkelstein-Schafer statistical method) and historical year weather data (from 2008 to 2017) with the meteorological variables of the solar radiation, dry bulb temperatures and wind speeds are obtained from the New & Renewable Energy Resource Map database and from the Korea Meteorological Administration. The weather data is used as input for the simulation tool to study the performance of the photovoltaic system. Fig. 1 shows the geographical characteristics of the weather stations with the corresponding code, while Table 2 summarizes the selected locations.

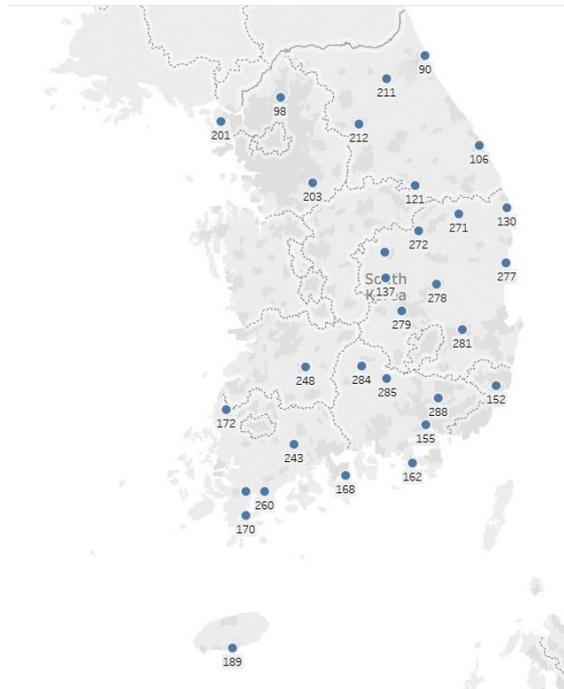


Fig. 1 Sites of the typical meteorological year datasets

Table 2. Summary of selection sites

Code	Province	Station name	Latitude	Longitude
155	Gyeonnam province	Changwon	35°1'	128°5'
152	Gyeonnam province	Ulsan	35°5'	129°3'
162	Gyeonnam province	Tongyeong	34°8'	128°4'
284	Gyeonnam province	Geochang	35°6'	127°9'
285	Gyeonnam province	Hapcheon	35°5'	128°1'
288	Gyeonnam province	Milyang	35°4'	128°7'
260	Jeonam province	Haenamgun	34°9'	125°2'
262	Jeonam province	Yeosu	34°7'	127°7'
168	Jeonam province	Wando	34°3'	126°7'
170	Jeonam province	Jangheongun	34°6'	126°9'
201	Incheon province	Ganghwa	37°7'	126°4'
172	Jeonbuk province	Gochang	35°3'	126°5'
243	Jeonbuk province	Namwon	35°0'	127°2'
248	Jeonbuk province	Sangju	35°6'	127°3'
98	Gyeonggi province	Dongducheon	37°9'	127°0'
203	Gyeonggi province	Icheon	37°2'	127°4'
189	Jeju province	Seoguiipo	33°2'	126.5'
271	Gyeongbuk province	Bonghwa	36°9'	128°9'
273	Gyeongbuk province	Moonkyeong	36°6'	128°1'
277	Gyeongbuk province	Yeongduk	36°5'	129°4'
272	Gyeongbuk province	Yeongju	36°8'	128°5'
278	Gyeongbuk province	Uiseong	36°3'	128°6'
279	Gyeongbuk province	Gumi	34°1'	128°3'
281	Gyeongbuk province	Yeongcheon	35°9'	128°9'
130	Gyeongbuk province	Uljin	36°9'	129°4'
137	Gyeongbuk province	Sangju	36°4'	128°1'
90	Gangwon province	Sokcho	38°2'	128°5'
106	Gangwon province	Donghae	37°5'	129°1'
121	Gangwon province	Yeongwol	37°1'	128°4'
212	Gangwon province	Hongcheon	37°6'	127°8'
211	Gangwon province	Inje	38°0'	128°1'

3. Methodology

In order to assess the accuracy of the TMY dataset, a comparison was made between the TMY energy yield results and the multi-year average values after the simulation of the weather dataset in PVSyst software. The adequacy of the simulation tool was tested by means of the simulated electric energy production and measured electric energy outcome. The findings show that the error between PV energy

production and simulated PV energy was close to 5%¹⁸⁾. Long-term average weather datasets containing 8760 hourly irradiance, air temperature, wind direction, wind speed values and the TMY datasets were fed into the simulation tool to generate the long-term electrical power of the PV power generation system. The behavior of the TMY datasets was evaluated using a 100-kW PV system, where the modules of the system are arranged on a fixed plan array with an optimal tilt angle of 30 degrees and oriented on 0-degree azimuths facing south. After the simulation process, all annual hours of the energy yield results were transferred to a CSV file for further analysis. The long-term average electric energy production was used as a reference value to validate the TMY electric energy production using a probabilistic and deterministic model. The indicator F1, obtained from Equation 1, is the relative error percentage of the TMY annual average of the electrical energy produced by the PV power generation system and the long-term average; these values were used to underscore the errors. Fig. 2 shows the flowchart that was used to evaluate the accuracy of the TMY datasets.

The relative error percentage (REP):

$$F1 = \frac{EP_{LT} - EP_{TMY}}{EP_{LT}} \times 100 \tag{1}$$

where EP_{LT} is an annual average electricity production of the long-term average, while EP_{TMY} is an annual average of electricity production of TMY.

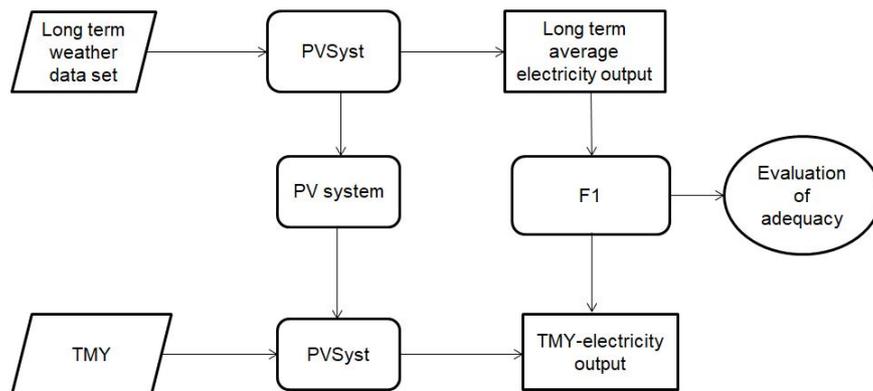


Fig. 2 Schematic representation of the analysis procedure

3.1 Traditional Statistics

The use of traditional statistics is very common for the validation of TMY data. The most common

descriptive statistics are the mean and standard deviation, which are used to describe the variability around the mean of a sample of a population. According to this approach, deterministic predictions cannot reflect uncertainty, and the risks can be very high in the economics of solar energy projects. However, in terms of statistics, uncertainty can be expressed as a 95% confidence interval from probabilistic modelling instead of a deterministic approach¹⁹⁾.

3.2 Bootstrap Method

The bootstrap method, introduced and named by Efron in 1979, is used as a resampling technique to obtain information about the uncertainty of the estimates of means and variances in any confidence interval of any distribution²⁰⁾. The idea behind bootstrapping is to generate multiple datasets from an actual dataset without making any assumptions. This method can be used to generate a sample of the same size as the original dataset by randomly selecting actual observations one by one. In addition, it is a probabilistic simulation for decision making in uncertain situations. In this study, the bootstrap method was utilized to measure quantitatively the uncertainty of the estimation in terms of the mean values and confidence intervals of the annual average electrical energy production generated by the TMY and the historical year average. From a solar resource risk assessment perspective, it is critical to quantify the uncertainty in resource data. If the solar resource uncertainty is low, the PV system performance can then be relied upon with a higher level of confidence. Therefore, the uncertainty of probabilistic estimates represents crucial information in solar energy resource assessments, as it serves to mitigate risks in renewable energy projects.

The analysis process is presented below:

- (1) Draw a new sample of size n with replacement from the original sample.
- (2) Calculate the mean or variance of the new sample and refer to it as m_i .
- (3) Repeat steps 1 and 2, 10,000 times, calling the i -th new sample mean or variance m_i .
- (4) Plot the distribution of these 10,000 sample means or variances.
- (5) Calculate the lower and upper bound of the 95% CI of m_i as the 2.5% and 97.5% percentiles (P2.5 and P97.5) of the bootstrapped distribution derived from the histogram.

3.3 Description of the Proposed Method

The bootstrap method greatly contributes to the decision-making process of TMY data in solar project planning. In addition, it helps in evaluations of the uncertainty of TMY datasets in solar system feasibility studies²¹⁾. The application of this method greatly increases the objectivity of the mean, variance and

uncertainty analyses of TMY annual electrical energy production levels and long-term averages. The energy performance of a PV system was assessed using PVSyst software. Therefore, taking full advantage of the bootstrap method, an effective decision making method was used to determine the accuracy of TMY datasets to support the long-term planning of utility-scale PV generation at any site.

The main steps are given below:

Step 1: Simulation tool. Several types of data, including TMY and long-term weather data, as well as component specifications are entered into the PVSyst software. After the simulation process, the hourly electrical energy production values were exported as a CSV file, with a deterministic method then used to evaluate the accuracy of the TMY data

Step 2: Uncertainty analysis. The bootstrap method is utilized for different sites. In addition, the probability distribution for the annual average TMY electricity production and the long-term average was constructed with estimation uncertainty in the form of means and confidence intervals. Finally, the relative error was applied to analyze the difference between the long-term average electric power production and the TMY electric power production.

4. Results and Discussion

In this study, in order to assess the adequacy of the TMY datasets, the long-term average data and TMY of each weather station were simulated using PVSyst software. After the simulation, the annual average power output of the TMY was compared with the long-term average using two different methods, probabilistic and deterministic, to evaluate the uncertainty of the TMY data. Therefore, we undertook the calculation of the mean and standard deviation with the deterministic method (without considering sampling uncertainty) and the bootstrap method (considering sampling uncertainty) in the TMY power output and the long-term average power output.

To quantify the performance capabilities of the methods in terms of uncertainty, we initially analyzed the mean and standard deviation of deterministic statistics and probabilistic statistics to obtain a more reliable measure of the uncertainty. The uncertainty associated with TMY datasets is directly related to the long-term expected yield of solar PV systems²²⁾. Therefore, project investors must understand the accuracy of their TMY datasets. If the TMY uncertainty is lower than the long-term average electricity yield, predictions of long-term solar PV yields can then be based on TMY datasets with a high level of confidence.

Fig. 3 shows the calculation results of the standard deviation of the TMY-electricity output and the long-term average electricity for all sites when using the two aforementioned methods. As shown in Fig. 3,

for most sites, a higher value is visualized for the mean and standard deviation as determined by traditional statistics compared to the 95% confidence intervals from the bootstrap method. As a result, the deterministic approaches did not provide reliable uncertainty estimates compared to the probabilistic approaches.

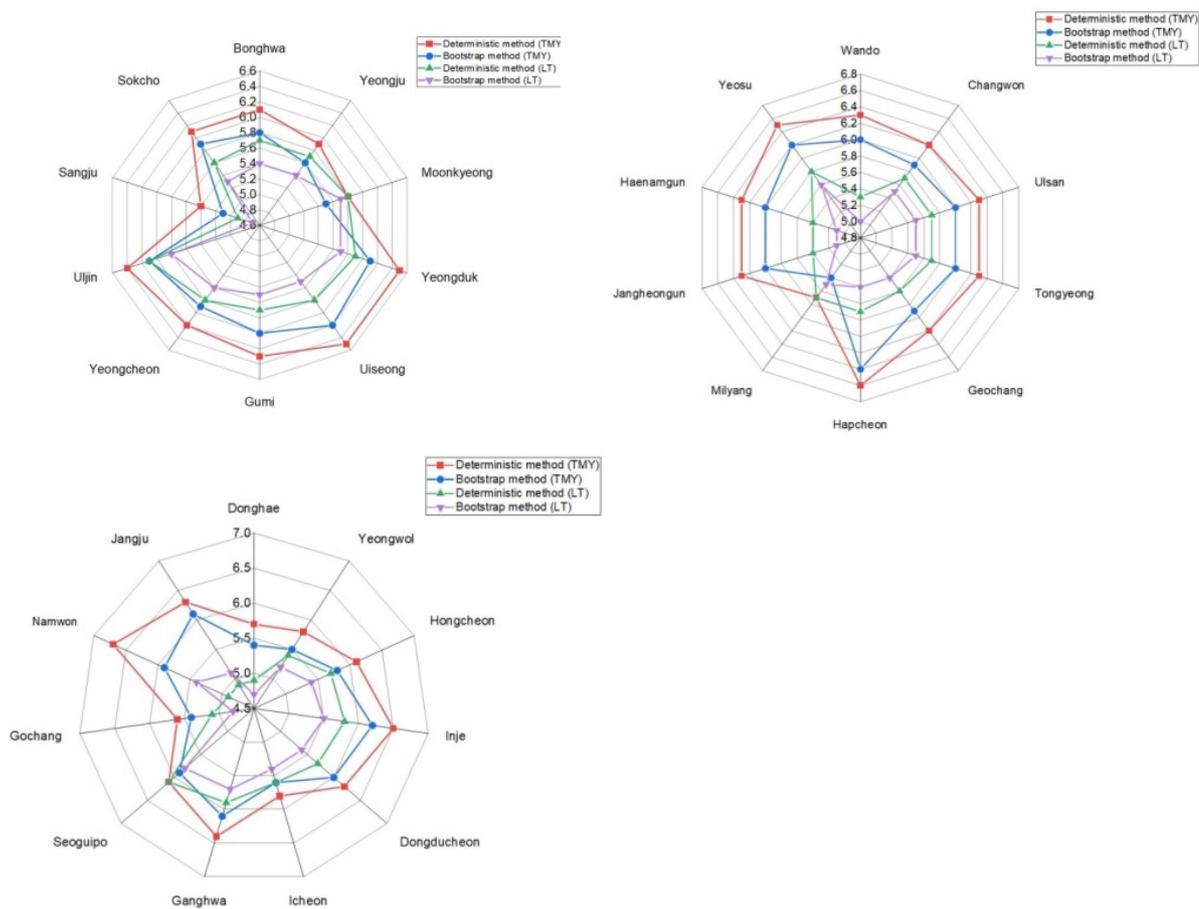


Fig. 3 The standard deviation scores of 31 weather stations using deterministic and bootstrap methods

According to the deterministic values indicated in Fig. 3, there is deviation from the bootstrap approach based on the probabilistic approaches. The planning of a solar project with a deterministic model can lead to over- or under-investment in the financial viability of the solar energy project. However, by using the 95% distribution, PV developers can quantify the prediction interval of the TMY-electricity output and uncertainty in feasibility studies. Therefore, various interpretations by investors are possible with the probabilistic approaches; in addition, the results can clarify the expansion capacity of PV systems more precisely and effectively than a deterministic approach. The classification results of the standard deviation scores of the 31 weather stations in Fig. 3 were randomly selected.

4.1 Probabilistic Prediction

The uncertainty in the estimates of the average values of the TMY electricity output and the long-term average electricity were calculated in order to validate the TMY datasets using the bootstrap method (as described in Section 3). The bootstrap method shows a 95% confidence interval. First, the mean values (P50) and standard deviation of the probabilistic prediction of the TMY electricity output are compared with the mean values of the deterministic approaches. Therefore, the equality of means and the standard deviation between the TMY electricity output and the long-term average electricity output were utilized. All P50 standard deviations of the TMY annual average electricity are slightly smaller than those from the deterministic statistics for almost all sites because the TMY estimate from the bootstrap method in each case is already corrected by the iterative resampling method using the 95% bootstrap confidence interval. In particular, the bootstrap method reduces the errors by 2% to 4% in terms of the P50 standard deviation for all stations, where P50 indicates values are almost equal to those for the deterministic statistics for all sites. With lower uncertainty due to the bootstrap method, a higher level of performance can be assumed at all sites. For example, if a solar farm has an annual average estimated yield production of 10 kW with uncertainty $\pm 10\%$ at a prediction interval of 95%, the solar farm would then produce only 9 kW at the low end of the uncertainty. If the uncertainty is $\pm 5\%$, the solar farm would produce 9.5 kW. The 9 or 9.5 kW would then be used to estimate the debt fraction. With lower uncertainty, the estimate of the return on equity would increase.

In the TMY electricity output case when using the bootstrap method, Fig. 3 shows that this approach indicates variability closer to the long-term average, meaning that the TMY data offers good performance in predictions of the electricity output. This information could be useful to measure and manage risk and will help those who must make many decisions using TMY datasets in renewable projects. In addition, the mean value and confidence interval (P2.5 and P97.5) of the TMY electricity output have values close to the long-term average with the least uncertainty.

These results reveal that the deterministic prediction can underestimate or overestimate the true mean value of a TMY in estimations of the electrical energy produced by a PV power generation system, mainly because this approach neglects the sampling uncertainty²¹⁾. As a result, future revenue values would decrease with high risk. The comparison of the two methods with regard to the TMY and the long-term average PV power output shows that the bootstrap method provides a more reliable credibility range for the mean value. Furthermore, the method proposed in this work can be used to analyze the uncertainty of TMY datasets quantitatively, and it is believed that it can be widely used for mitigating the risks of the solar power projects in the future.

4.2 Relative Error Analysis

The impact of uncertainty in the TMY data was evaluated using two different methods (deterministic and probabilistic prediction) with a comparison of the relative error of the annual average value of the TMY and the long-term electric energy production outcome. Fig. 4 presents the relative error percentage for each method. The relative error was used to validate the TMY electricity output calculated by the two methods. Fig. 4 illustrates the calculation results of the relative error percentage (REP) at different locations. It can be seen that the REP outcomes are wholly different for two methods, except for the Uljin and Icheon sites. The difference is mainly due to deterministic method itself, as it neglects the effect of sampling uncertainty on the estimations of the average and standard deviation. Furthermore, it gives a single outcome, while the bootstrap method incorporates the uncertainty explicitly in the mean value and variance. Therefore, it offers a confidence interval of the possible outcomes of the model.

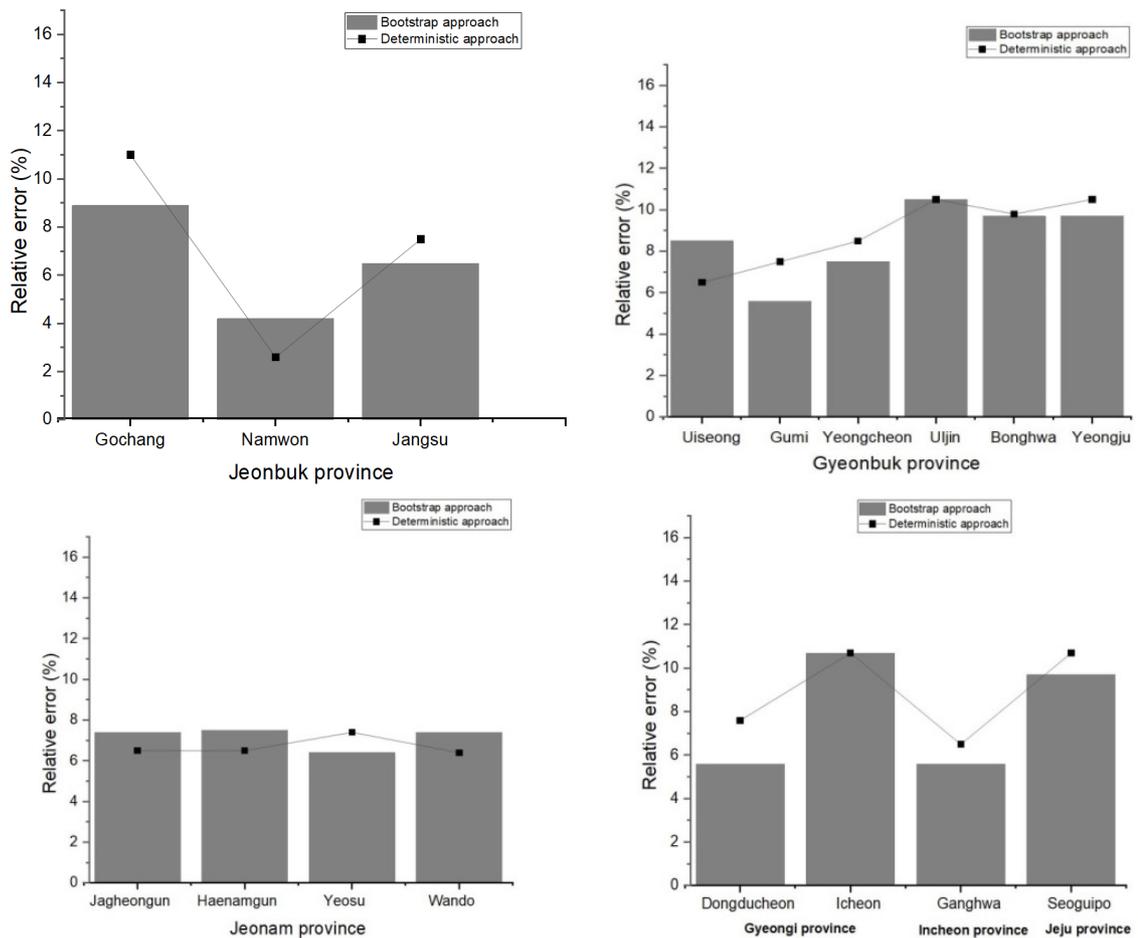


Fig. 4 Comparison of relative error percentage between deterministic and probabilistic prediction

The error percentages according to the deterministic method (Fig. 4) at Yeongju, Gumi, Yeongcheon, Dongducheon, Ganghwa, Seoguipo, Gochang, Jongsu, and Yeosu site appeared to be higher than the probabilistic estimates. In addition, this approach had limitations that may lead to underestimations or overestimations of the mean and standard deviation of the TMY electricity output because it does not consider the uncertainties caused by missing data. However, knowledge of the accuracy and uncertainties of TMY data is very important for debt financing of a solar energy project. Thus, higher uncertainty levels would affect the estimated performance of the electricity output as well as financial investments of the solar energy project. In contrast, the probabilistic method was found to be very useful as it provided a valid and informative estimate of the long-term feasibility of the photovoltaic system.

It is significant that the error metrics calculated in other studies using deterministic models tend to have the same values according to the data, as shown in Table 3.

Table 3. Summary of validation of TMY datasets in other studies using the deterministic method

Location	The difference between TMY and long term average.	References
Beijing	4%	Jiang et al. (2010) ²³⁾
Hongkong	4.14%	Chan et al. (2016) ²⁴⁾
Thailand	5%	Pattarapanitchai et al. (2014) ²⁵⁾

The bootstrap method can provide not only the expected estimate but also the prediction interval, which can be used to analyze the uncertainty of the TMY for estimating the electric energy produced by a PV power generation system in the long term. The deterministic approach can increase the expectation of the profitability of energy output in a solar power project. This may affect the renewable energy tax credit that is applied based on the amount of energy produced. From the above discussions, it is clear that uncertainty analyses of TMY datasets should be based on the bootstrap method (probabilistic prediction) rather than on deterministic predictions.

5. Conclusion

Evaluations of solar projects planning require high accuracy of the TMY datasets. Therefore, we designed a photovoltaic system in the PVSyst simulation tool to test the adequacy of TMY datasets. Furthermore, two statistical methods, in this case deterministic and probabilistic methods, were compared by evaluating the uncertainties in TMY datasets. In general, the probabilistic method provides more accurate results than the

deterministic method. The uncertainty was evaluated by comparing the annual means and variances of TMY electricity production and the long-term average electricity production. The bootstrap method was utilized to analyze the uncertainties of TMY datasets in a feasibility study of a solar energy system using 95% confidence intervals, which provided both valid and more informative estimates than the deterministic prediction.

The main important findings are summarized below.

- (1) The uncertainty in the estimates of the variances was quantified by calculating the 95% confidence intervals of the TMY and the long-term average electricity output at all stations. The results showed that the probabilistic prediction of the TMY was close to the long-term average electricity output at all sites. Furthermore, compared to the deterministic prediction, the bootstrap method provides the uncertainty value at the 95% confidence level, which can be used to estimate the debt fraction of a solar energy project. A proper understanding of uncertainty in TMY datasets is a fundamental element for ensuring competitive financing of solar energy projects.
- (2) The bootstrap method presented in this paper considers a proper assessment of the uncertainty in the TMY dataset used for a solar energy system to obtain reliable results, and it is advisable to express the uncertainty as a prediction interval in statistics. In addition, the annual average TMY electricity output estimate with the confidence interval will be useful in a solar energy resources assessment and can help planners and investors to use TMY datasets more confidently in renewable energy deployment decision, also likely playing a crucial role in mitigating the risk of renewable energy projects.

The policy implications of this research are several. First, the results of this research can serve as advice for project planners as they conduct studies of the long-term feasibility of solar energy projects with TMY datasets. Second, the uncertainty analysis of TMY datasets developed by the bootstrap method can help in the decision-making process when weather datasets are missing. Thirdly, Korean policymakers should encourage the use of TMY datasets in PV system design projects because these datasets can reduce the computational burden and save time. In addition, in the future, the bootstrap method with the confidence interval of the TMY electricity output will help those involved in evaluations of the uncertainty in the economic performance of solar energy systems in terms of the levelized cost of energy and the net present cost.

ACKNOWLEDGEMENT

This work was supported by the Technology Innovation Program (20016225) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

REFERENCES

1. De Miguel, A. and Bilbao, J., Test Reference Year Generation from Meteorological and Simulated Solar Radiation Data, *Solar Energy*, Vol. 78, No. 6, pp. 695-703, 2005.
2. Crawley, D. B., Which Weather Data Should You Use for Energy Simulations of Commercial Buildings?, *Transactions-American Society of Heating Refrigerating and Air Conditioning Engineers*, 104, pp. 498-515, 1998.
3. Lund, H. and Eidorff, S., Selection Methods for Production of Test Reference Years, Report no EUR 7306 EN. - Tech. Univ. Denmark, Dept. of Buildings and Energy, 1985.
4. Hall, I., Prairie, R., Anderson, H., and Boes, E., Generation of Typical Meteorological Years for 26 SOLMET Stations, Rapport technique SAND78-1601, - Sandia National Lab., Albuquerque, 1978.
5. Moazami, A., Nik, V. M., Carlucci, S., and Geving, S., Impacts of Future Weather Data Typology on Building Energy Performance-Investigating Long-term Patterns of Climate Change and Extreme Weather Conditions, *Applied Energy*, Vol. 238, pp. 696-720, 2019.
6. Cui, Y., Yan, D., Hong, T., Xiao, C., Luo, X., and Zhang, Q., Comparison of Typical Year and Multiyear Building Simulations Using a 55-year Actual Weather Dataset from China, *Applied Energy*, Vol. 195, pp. 890-904, 2017.
7. Bre, F. and Fachinotti, V. D., Generation of Typical Meteorological Years for the Argentine Littoral Region, *Energy and Buildings*, Vol. 129, pp. 432-444, 2016.
8. Ho, C. K., Khalsa, S. S., and Kolb, G. J., Methods for Probabilistic Modeling of Concentrating Solar Power Plants, *Solar Energy*, Vol. 85, No. 4, pp. 669-675, 2011.
9. Nelken, K. and Zmudzka, E., TMY Versus Multi-year Time Series of Meteorological Conditions for the Characterization of Central Poland's Suitability for Photovoltaics, *Meteorologische Zeitschrift*, Vol. 26, No. 1, pp. 21-31, 2016.
10. Khosravi, A., Nahavandi, S., and Creighton, D., Prediction Intervals for Short-term Wind Farm Power Generation Forecasts, *IEEE Transactions on Sustainable Energy*, Vol. 4, No. 3, pp. 602-610, 2013.
11. Chu, Y. and Coimbra, C.F., Short-term Probabilistic Forecasts for Direct Normal Irradiance, *Renewable Energy*, 101, pp. 526-536, 2017.
12. Hu, Y. M., Liang, Z. M., Liu, Y. W., Wang, J., Yao, L., and Ning, Y., Uncertainty Analysis of SPI Calculation and Drought Assessment Based on the Application of Bootstrap, *International Journal of Climatology*, Vol. 35, No. 8, pp. 1847-1857, 2015.
13. Ohunakin, O. S., Adaramola, M. S., Oyewola, O. M. and Fagbenle, R. O., Generation of a Typical Meteorological Year for North-east, Nigeria, *Applied Energy*, Vol. 112, pp. 152-159, 2013.
14. Realpe, A., Vernay, C., Pitaval, S., Lenoir, C., and Blanc, P., Benchmarking of Five Typical Meteorological Year Datasets Dedicated to Concentrated-PV Systems, *Energy Procedia*, Vol. 97, pp.108-115, 2016.
15. Nelken, K. and Zmudzka, E., TMY Versus Multi-year Time Series of Meteorological Conditions for the Characterization of Central Poland's Suitability for Photovoltaics, *Meteorologische Zeitschrift*, Vol. 26, No. 1, pp. 21-31, 2016.
16. Abreu, E. F., Canhoto, P., Prior, V., and Melicio, R., Solar Resource Assessment through Long-term Statistical Analysis and Typical Data Generation with Different Time Resolutions Using GHI Measurements, *Renewable Energy*, Vol. 127, pp. 398-411, 2018.
17. Moreno-Tejera, S., Silva-Perez, M. A., Lillo-Bravo, I., and Ramirez-Santigosa, L., Solar Resource Assessment in Seville, Spain, Statistical Characterisation of Solar Radiation at Different Time Resolutions, *Sol. Energy*,

Vol. 132, 430e441, 2016.

18. Carigiet, F., Baumgartner, F., Sutterlueti, J., Allet, N., Pezzotti, M., and Haller, J., Verification of Measured PV Energy Yield Versus Forecast and Loss Analysis, In 28th European PV Solar Energy Conference, October 2013 .
19. Rostron, P. D., Fearn, T., and Ramsey, M. H., Confidence Intervals for Robust Estimates of Measurement Uncertainty, Accreditation and Quality Assurance, pp. 1-13, 2020.
20. Tsamardinos, I., Greasidou, E., and Borboudakis, G., Bootstrapping the Out-of-sample Predictions for Efficient and Accurate Cross-validation, Machine Learning, Vol. 107, No. 12, pp. 1895-1922, 2018.
21. Ahn, H., Rim, D., Pavlak, G. S., and Freihaut, J. D., Uncertainty Analysis of Energy and Economic Performances of Hybrid Solar Photovoltaic and Combined Cooling, Heating, and Power (CCHP+ PV) Systems Using a Monte-Carlo Method, Applied Energy, Vol. 255, 113753, 2019.
22. Khosravi, A., Nahavandi, S., and Creighton, D., Prediction Intervals for Short-term Wind Farm Power Generation Forecasts, IEEE Transactions on Sustainable Energy, Vol. 4, No. 3, pp. 602-610, 2013.
23. Jiang, Y., Generation of Typical Meteorological Year for Different Climates of China, Energy, Vol. 35, No. 5, pp. 1946-1953, 2010.
24. Chan, A. L. S., Generation of Typical Meteorological Years Using Genetic Algorithm for Different Energy Systems, Renewable Energy, Vol. 90, pp. 1-13.
25. Pattarapanitchai, S., Tohsing, K., Pankaew, P., and Janjai, S., Generation of Typical Meteorological Year Datasets for 20 Stations in Thailand, In 2014 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE) (pp. 1-6), IEEE, March 2014.