Enhancements in Day-Ahead Forecasts of Solar Irradiation with Machine Learning: A Novel Analysis with the Japanese Mesoscale Model

JOAO GARI DA SILVA FONSECA JR.

Ogimoto Laboratory, Institute of Industrial Science, The University of Tokyo, Meguro, Japan

FUMICHIKA UNO, HIDEAKI OHTAKE, AND TAKASHI OOZEKI

Systems Team, Research Center of Photovoltaics, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

KAZUHIKO OGIMOTO

Ogimoto Laboratory, Institute of Industrial Science, The University of Tokyo, Meguro, Japan

(Manuscript received 29 September 2019, in final form 13 April 2020)

ABSTRACT

The objective of this study is to propose and evaluate a set of modifications to enhance a machine-learning-based method for forecasting day-ahead solar irradiation. To assess the proposed modifications, they were implemented in an initial forecast method, and their effectiveness was analyzed using two years of data on a national scale in Japan. In addition, the accuracy of the modified method was compared with one of the forecast methods for solar irradiation used by the Japan Meteorological Agency (JMA), namely, the mesoscale model (MSM). Such forecasts were made publicly available only recently, which makes this study one of the first ones to compare them with machine-learning-based forecasts. The annual root-mean-square error (RMSE) of local forecasts of the JMA-MSM varied from 0.1 to 0.14 kW h m^{-2}; the regional equivalent varied from 0.062 to 0.091 kW h m^{-2}. In comparison with these results, the modified model achieved an average RMSE reduction of 7.5% on the local scale and 16% on the regional scale. The modified model also had a skill score that was 23% higher than that of the JMA model. Furthermore, the performance of the JMA model had strong spatial and seasonal dependencies, which were reduced in the machine-learning-based forecasts. The results show that the proposed modifications are effective in reducing large forecasts errors, but they cannot compensate for situations in which the input data used to make the forecasts are highly inaccurate.

1. Introduction

The worldwide dissemination of photovoltaic (PV) power systems in the current decade has been remarkable. Just between 2010 and 2017, the worldwide installed capacity of PV power increased from 40 to 403 GW (International Energy Agency 2018). Such growth, associated with the intrinsic weather-dependent variability typical of PV power generation, has caused a strong demand for improvements in day-ahead forecasting of solar irradiation.

Day-ahead forecasts of solar irradiation are essential to support the stability of power grids in scenarios of high penetration of PV power and in countries with day-ahead markets. In Japan, for example, such forecasts are employed to schedule the curtailment of residential and nonresidential PV power in regions with high penetration of PV and insufficient grid flexibility and power demand in order to consume all generation (Ministry of Economy, Trade and Industry 2014). Due to the characteristics of the Japanese PV power curtailment system, errors in the day-ahead forecasts of solar irradiation can cause economic losses to the PV systems' owners as well as power supply–demand issues to the power grid operators.

One of the following three approaches is usually adopted to forecast solar irradiation one day ahead of time. In the first approach, the forecast is performed directly from the numerical weather prediction (NWP) models. Examples of such models include the ones developed by the European Centre for Medium-Range

Corresponding author: Joao Gari da Silva Fonseca Jr., jfonseca@iis.u-tokyo.ac.jp

Weather Forecasts (ECMWF), the Weather Research and Forecasting (WRF) Model of the U.S. National Center for Atmospheric Research (Powers et al. 2017), and the nonhydrostatic model (NHM) of the Japan Meteorological Agency (JMA; Saito et al. 2007). Such models are continuously developed, the main purpose of which is to achieve better weather predictions in general. Recently, however, with the growth of PV systems, NWP solar irradiation models have received considerable attention. For example, a version of the WRF focused on solar resources (WRF-Solar) has been developed (Jiménez et al. 2016a,b), and it is already being updated (Golnas 2018). Another example comes from the JMA, which in February 2017 replaced its original mesoscale model called NHM (Saito et al. 2007) with a new model called a system based on a unified concept for atmosphere (ASUCA; Japan Meteorological Agency 2014). This new model is expected to increase the accuracy of low-level cloud forecasts by implementing horizontal discretization of the atmosphere with a finite-difference method (the NHM used a spectral approach).

The second approach often adopted to forecast solar irradiation one day ahead of time is to use NWP data as predictors and machine learning or statistical models to obtain the solar irradiation as the output, as in Antonanzas et al. (2016) and Voyant et al. (2017). The following are examples of other methods employing this approach. Pereira et al. (2019) developed an artificial neural network (ANN) to postprocess up to 72 h-ahead forecasts of the global horizontal irradiation obtained using the ECMWF NWP model. The authors trained their ANN to correct the forecasts according to several weather-related variables. For 103 days of data and four locations in Portugal, they showed that their corrective algorithm increased the forecast skill by 23% when compared with forecasts without postprocessing. Qing and Niu (2018) proposed a method based on long short-term memory (LSTM) networks to forecast day-ahead solar irradiance. They compared the performance of their method with that of persistence, linear regression, and ANN-based models with over six months of data from the island of Cape Verde. The authors found that the LSTM method outperformed the ANN, achieving a root-mean-square error (RMSE) of 122.7 against 150.2 W m^{-2} for the ANN. In the same study, using a dataset with 11 years of data from the National Renewable Energy Laboratory (NREL), the authors showed that their method yielded an RMSE 42.7% lower than the methods based on ANNs. Srivastava and Lessmann (2018) also tested the performance of LSTM networks in day-ahead forecasts of global horizontal irradiance (GHI) but using satellite data. They compared the performance of LSTM against persistence, ANN, and

gradient boosting regression methods over 21 locations in Europe and the United States and one year of test data. The authors found that the LSTM yielded a median reduction of the forecasts' RMSE of 8.6% when compared with the ANN and gradient boosting regression-related forecasts (maximum RMSE reduction of 18% in one location and maximum RMSE increase of 7.1% in another).

Currently, the potential benefits of machine-learning techniques that use random projections of the inputs into high-dimensional space, such as extreme learning machines (ELM) and reservoir computing, are also being explored in the problem of PV power forecasts. Such techniques can also be extended to the problem of solar irradiation forecasts. The main advantage of these techniques is that they significantly simplify the training process, as they require only the updating of the weights of the output layer. For example, Le Cadre et al. (2015) proposed the use of ELM and an information-based rule to generate day-ahead probabilistic PV power generation forecasts. Li et al. (2015) used ELM to forecast the hourly PV power generation of a system in Shanghai, China, for 6 months. They created three models based on the weather conditions and employed the most suitable one according to the predicted weather of each target day. Their method outperformed a reference ANN-based model by approximately 5% in terms of a normalized RMSE. Hossain et al. (2017) used ELM to forecast the PV power of three residential systems in Malaysia. They tested their model with three months of data and found that the ELM-based model outperformed both support vector regression (SVR) and ANN models. For the solar irradiance forecast itself, Basterrech et al. (2013) developed an echo state network and differential polynomial neural networks to forecast the day-ahead solar irradiance in the Czech Republic. They found that both methods achieved similar performance for a test set of seven days. Alomar et al. (2016) used reservoir computing to predict daily global solar irradiation from air temperature data. They used four years of data for training and one year for testing and found that their model outperformed regression-based empirical models by approximately 6% in terms of a normalized RMSE.

The third approach that is often used for solar irradiation forecasts is to increase the complexity of the forecast models, using meta aggregators to generate ensembles of forecasts done with the first or second approaches. An example of this kind of approach is the Watt-sun forecasting system based on a multiexpert learning approach for multihorizon forecasts of solar irradiation and PV power (Hamann 2017). This system

combines the North American Mesoscale Forecast System, the National Centers for Environmental Prediction, and the Global Forecast System data with random forests, support vector machines, and linear models, in a 12-member ensemble forecast. According to Hamann (2017), the Watt-sun system can yield forecasts of solar power with a mean absolute error of 20% lower than forecasts based on the ECMWF (evaluations based on two-day-ahead forecasts for three locations in the United States, and more than six months of data). Uno et al. (2018) used a grand ensemble technique based on four leading NWP models from around the world to evaluate the potential error reduction in the daily forecasts of solar irradiation in the Kanto region in Japan, with a horizon from 1 to 6 days ahead of time. They showed that an ensemble of models with coarse grid spacings $(40, 140,$ and $110 \, \text{km}^2)$ can yield forecast errors as low as those provided by a 25-km^2 model. Perez et al. (2018) developed a technique based on historical satellite data and a multimodel blending of four advanced NWP models to improve the day-ahead forecast of solar irradiation. They showed that their technique, over a period of nine months and seven locations in the United States, was able to yield forecasts with an RMSE 1.7% lower than their reference model.

The literature available on this topic shows that although advances have been made, forecast error characterization and error reduction in the day-ahead horizon remains an open problem. Because of the eventual occurrence of large errors, their reduction is a constant objective. Nonetheless, even the quantification of advances achieved by new methods is a difficult task in some regions of the world, due to the lack of publicly available forecasts of solar irradiation that can be used as a common baseline. This was the case in Japan until December 2017, when the JMA started to make publicly available its mesoscale model forecasts of solar irradiation (Japan Meteorological Agency 2017).

In consonance with the need to improve day-ahead forecasts of solar irradiation, and to characterize such improvements using a publicly available standard, this study has two objectives. The first one aims to evaluate the impact of the three modifications on improving a forecast model based on the application of machine learning with NWP data. To verify the potential of these modifications, we implemented them on a forecast model developed previously (Fonseca et al. 2015) and verified how the forecast error changed after the implementation. The second objective aims to provide a countrywide characterization of the performance of day-ahead forecasts of solar irradiation in Japan, using the initial model, the modified model, and forecasts from the JMA mesoscale model (MSM) (which are

bound to become the standard baseline in Japan after becoming publicly available). To do that, we compared the performance of the three forecast models in 2016 and 2017 at 41 locations distributed across all the main regions of Japan. Furthermore, we evaluated the forecast accuracy on local as well as on regional scales. Although assessments of the observed solar irradiation in Japan as well as the stand-alone seasonal evaluations of the JMA NWP models have been studied before (Watanabe et al. 2016; Ohtake et al. 2015, 2018), to the best of our knowledge, this is the first comprehensive countrywide evaluation of machine-learning-based forecasts using the JMA-MSM forecasts as the baseline. Thus, aside from showing the potential of the proposed enhancement techniques, the analysis can also be a useful reference to other researchers when comparing the performance of day-ahead forecasts of solar irradiation in different parts of the world.

This paper is organized as follows: in section 2, day-ahead solar irradiation forecast models used are introduced; in section 3, forecast targets (regional and local) are described; in section 4, the performance metrics employed are presented; and in section 5, the results are presented and discussed. The conclusions of the study are presented in section 6.

2. Forecast models

The forecast models used in this study are the persistence model, the JMA-MSM, the initial machine-learning model developed previously, and a new model modified with enhancement techniques. The persistence model was used to provide an easily reproducible reference to which the skill of the other three models could be compared.

a. Day-ahead persistence forecasts

Forecasts with persistence are done by using the values regarded as the observed ones for a day as the predictions for the next day. In persistence-based forecasts, it is assumed that the weather conditions of a day will persist or be the same for the next day. In the day-ahead horizon, any other nonnaive forecast method should generally outperform persistence to justify its application. Thus, persistence was used only to evaluate the skill of the other three models.

b. NWP forecasts: The JMA-MSM

The mesoscale model of the JMA entered into operation in May 2005 with the development of the NHM (Saito et al. 2006). A detailed description of it can be found in Saito et al. (2007). The spatial domain of the model covers the region between 22.4° and 47.6°N

latitudes and 120° and 150°E longitudes, with an effective grid spacing of 25 km^2 (a total area of approximately 10 900 km^2 including and surrounding Japan). The approximate model top is 22 km with 50 vertical levels. A data assimilation technique using a coarser spatial resolution model, the JMA global spectral model (GSM) is used to initialize the MSM. In its current version, 39-h-ahead forecasts are provided 6 times per day. For operational reasons, the forecasts are released with a latency of approximately 2.5 h from each initial time (Japan Meteorological Business Support Center 2019).

The JMA releases MSM-based predictions of several weather-related variables in 16 vertical layers of pressure (from 100 to 1000 hPa) and at the surface level. Until December 2017, the published surface-level data included only mean sea level atmospheric pressure, surface atmospheric pressure, relative humidity, accumulated precipitation, and cloudiness. All the predictions have a temporal resolution of 1 h. These data are made available through a gridpoint value mesoscale model prediction (GPV-MSM) dataset. Solar irradiation predictions from the MSM started being released to the public through the GPV-MSM dataset only from December 2017 onward.

Solar irradiation forecasts of the MSM are calculated using a radiative transfer scheme based on a 22-band model, including the parameterizations of several elements affecting solar irradiation. The difference in the optical characteristics of ice and water particles, as well as the monthly profiles of aerosols in the targeted areas, is also considered in the scheme. The effect of clouds with a size smaller than the model resolution is estimated with a partial condensation model proposed by Sommeria and Deardorff (1977).

Last, it should be mentioned that, although solar irradiation forecasts of the JMA were made publicly available only after 2017, since a part of this study was carried out within a Japan Science and Technology Agency–sponsored research project in which the JMA participated, access to its day-ahead solar irradiation forecasts from before 2017 was granted.

c. Machine-learning-based forecasts 1: The initial model

Until December 2017, the JMA did not publicly release solar irradiation forecasts; hence to forecast solar irradiation, we previously proposed a machine-learning-based forecast method that used the available NWP data (Fonseca et al. 2011). The machine-learning algorithm used was the ν-SVR proposed by Schölkopf et al. (1998).

The objective of SVR is to fit as many instances of the data as possible within the margins set with a fixed distance between them while controlling the number of outliers that can be left outside the margins. A hyperparameter ε controls the distance between the margins (which is also called the ε tube); in the ν-SVR formulation, this parameter is automatically set during the training procedure. Mathematically, for each instance of the data i, the ν-SVR maps input variables $\mathbf{x}$ into a high-dimensional space with a map function ϕ and a bias $\mathbf{b}$, as shown in Eq. (1):

$$y_i \approx \langle \mathbf{w} \cdot \phi(x_i) \rangle + \mathbf{b}. \tag{1}$$

By expanding the weight vector $\mathbf{w}$ in Eq. (1) as a combination of the instances of $\mathbf{x}$ after the mapping with the Lagrange multipliers $\alpha_i \alpha_i^*$, and replacing the resulting inner product by a kernel function, the optimization problem becomes one of determining the values of $\alpha_i \alpha_i^*$ that maximizes the following objective function for n instances of data:

$$\text{maximize} \left[-\frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*) k(x_i, x_j)(\alpha_j - \alpha_j^*) - \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) y_i \right], \tag{2}$$

subject to

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0, \tag{3}$$

$$\sum_{i=1}^{n} (\alpha_i + \alpha_i^*) \leq C\nu, \quad \text{and} \tag{4}$$

$$0 \leq \alpha_i^*; \quad \alpha_i \leq C/n. \tag{5}$$

In Eqs. (4) and (5), C is the cost hyperparameter that controls how the errors are handled by the model, and ν controls the fraction of instances that can become support vectors (used to build the margins containing the instances) during training. Besides these two hyperparameters, a kernel function also has to be chosen. In this forecast model, the Gaussian radial basis function was selected as the kernel. To set the hyperparameters, a model ensemble like that proposed in Fonseca et al. (2015) was used. One forecast model was built per targeted day using data of the 60 preceding days as the training dataset (this value was based on evaluations with regard to the training data size). Thus, the forecast for each day was done using a model trained with the most recent data available.

A training dataset consists of NWP data, which are used as input data, and observed values of solar irradiation, which are used in the calculation of the forecast errors during training. The input–output variables of the initial model are given in Table 1. All the input variables

TABLE 1. Input and output variables of the initial machine-learning-based forecast model.

	Variables	Source
Input[a] (predictors)	3-level cloudiness, temperature[b], and humidity[b]	JMA-MSM surface level predictions done at 1200 JST of the day preceding the targeted days
	Extraterrestrial solar irradiance[b]	Calculated with a theoretical model
Output (target variable)	1-h accumulated solar irradiation (kW h m^{-2})[c]	—

[a] All predictors were normalized.
[b] Values for the targeted period and the hour preceding the targeted one are used as input.
[c] Output values are compared with observed values at JMA weather stations and used to calculate forecast errors.

were normalized by their maximum historical values in Japan. Solar irradiation is the only variable in the study that had forecast and observed values. Observed values are used only in the calculation of forecast errors during the training stage and the performance evaluations. Last, NWP forecasts of solar irradiation are not used at all in the approach.

This model was regarded as the initial model. It was implemented in the Python programming language (version 3.6) using the port of a library for support vector machines (LIBSVM; Chang and Lin 2001) for scikit-learn (Pedregosa et al. 2011).

As the initial model uses NWP data as the input, it has a role similar to that of a radiative transfer scheme in a weather forecast system. NWP solar irradiation is not used at all. Thus, it is suitable for situations where numerical predictions of solar irradiation are not available. Last, with regard to the use of the SVR itself, it was chosen because it showed good results against an ANN (Fonseca et al. 2011). The suitability of SVR in solar irradiation forecasting has also been shown by other researchers (Voyant et al. 2017; Lauret et al. 2015; Cheng et al. 2014; Zendehboudi et al. 2018).

d. Machine-learning-based forecasts 2: The modified model

Although the initial model presented in section 2c performed well (Fonseca et al. 2011), further improvements are desired. To achieve better performance, we considered three modifications to improve the accuracy of the forecasts of the initial model. The first modification is the use of three additional numerically predicted weather-related variables as the input variables. The variables added were the atmospheric pressure, wind velocity in the north–south direction, and wind velocity in the east–west direction. These three variables contain day-ahead predictions done with the JMA-MSM, and they were retrieved from the GPV-MSM dataset following the same approach adopted for the other NWP input data described in section 2c.

The second modification is aimed at customizing the learning of the model by selecting target-dependent training patterns. The use of 60 days of data preceding a targeted day, as proposed in the initial model, has two main advantages. First, it imposes a small burden with regard to the storage of data, as only two months are needed to make forecasts for a whole day. Second, the continuous use of the most recent past data to train a forecast model ensures that the most recent weather trends are always available to the machine-learning algorithm. However, this approach also has the disadvantage of forcing the forecast model to learn all the weather patterns in the set of 60 days of training data, regardless of their similarity to the ones targeted. The result is that a forecast model trained to forecast a day expected to be sunny might also end up learning patterns from cloudy and rainy days.

To address this problem, we propose to increase the initial database available to train a model to one year of past input–output data and to implement a training pattern selection procedure. In this procedure, for each hour of a targeted day, the hours most similar to it, in terms of the input data, are searched and a specific forecast model for the targeted hour is built. The Euclidean distance is used as a similarity metric, and 720 h of training data are selected from the initial database. The following input variables were used in the Euclidean distance calculations of similarity between any 2 h: predicted cloudiness (three levels), predicted humidity, predicted temperature, and predicted extraterrestrial solar irradiance (top-of-atmosphere GHI). The number of hours and variables used were set based on evaluations with spare data.

With this procedure, we expect that the forecast models will be able to characterize better the solar irradiation of distinctive weather patterns. However, there are two main limitations when using this or any other training pattern selection technique for a solar irradiation forecast. The first limitation comes from the assumption of similarity itself. It is assumed that the past year of data will contain enough hours that are sufficiently similar to the targeted hour, in terms of the input data of the forecasts, so that the targeted hour weather pattern will be properly learned. If the input data of the targeted hour are atypical or rare, the training data selection will not be effective. The second limitation is that the training
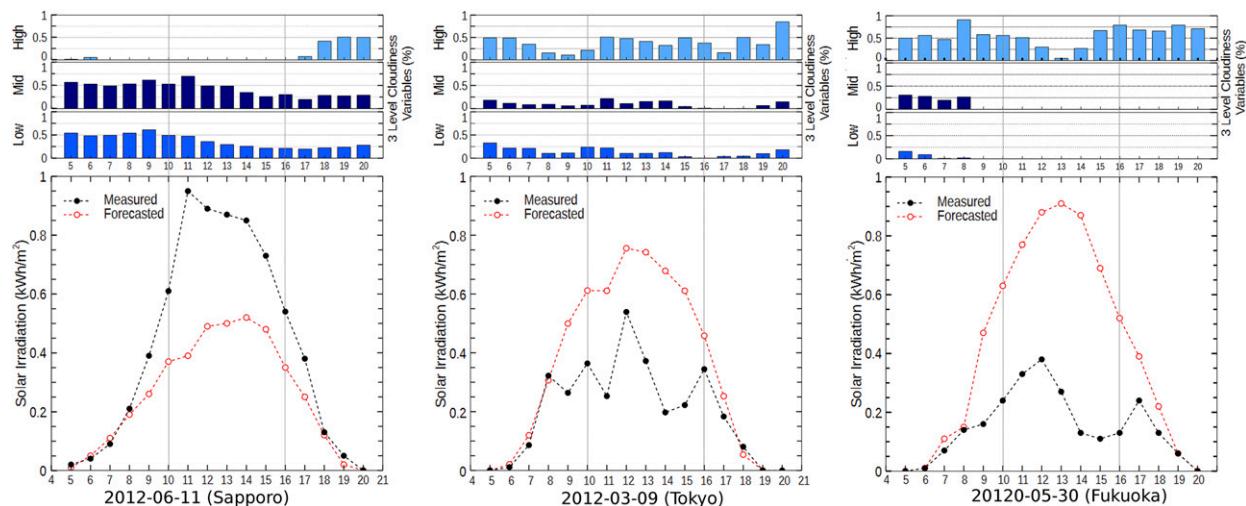
FIG. 1. Inaccurate day-ahead predictions of solar irradiation forecasts and the corresponding cloudiness used as predictors [three locations in (left) northern (Sapporo), (center) central (Tokyo), and (right) southwestern (Fukuoka) regions in Japan].

pattern selection should not improve a forecast if the input data of a targeted hour is completely inaccurate.

The motivation for the third modification is based on the analysis of examples of large forecast errors obtained with the initial model for different locations and time. By analyzing such errors, we observed among many patterns two distinct ones. In some cases, although one of the cloudiness variables had a high value (indicating cloudy weather), the model was forecasting high levels of solar irradiation. In other cases, one of the three cloudiness variables was simply wrong, causing the model to forecast low solar irradiation values when they should be high. Examples of these patterns in three locations in Japan are in Fig. 1. In the left panel of Fig. 1, the realized solar irradiation values suggest that the predicted midlevel cloudiness values were incorrect around noontime, causing a strong underestimation of the realized value. The overestimation in the center and right panels of Fig. 1 suggests that the model ignored the high-level cloudiness predictions in favor of low and midlevel ones, resulting again in large forecast errors.

Because of the limitations of the JMA-MSM (and any other current NWP model), the accuracy of the three cloudiness predictors in the day-ahead horizon is highly variable. The result is that the day-ahead predicted cloudiness can be highly inaccurate. In spite of that, when a machine-learning model is trained with data of these variables, it will learn their correlation with solar irradiation and will give weight to one variable more than another accordingly. For example, Table 2 has the correlation between cloudiness and solar irradiation normalized by extraterrestrial irradiance for the same three locations of Fig. 1, using 3 years of data. For

these three locations the machine-learning model, when trained with each location's data, should give a high weight to midlevel cloudiness and a low weight to high-level cloudiness information.

Although, this is the expected characterization of the effect of the three MSM cloudiness variables, the predicted cloudiness that receives the highest weight is not always the most accurate variable of the three cloudiness variables. Thus, in cases where it is inaccurate, we hypothesize that the initial model will likely yield solar irradiation forecast with a large error, as was shown in the case for midlevel cloudiness in Sapporo, Japan (Fig. 1, left panel). On the other hand, although the cloudiness variable that receives the lowest weighting is not always inaccurate, we hypothesize that its information will be generally ignored by the initial model, as what happened in Tokyo, Japan (Fig. 1, center panel), and Fukuoka, Japan (Fig. 1, right panel). Since the three cloudiness variables were not measured, direct evidence of this problem cannot be provided. However, if this problem is happening, devising a method to prevent too

TABLE 2. Correlation coefficient among each of the three day-ahead cloudiness predictors and the solar irradiation normalized by extraterrestrial solar irradiance during three years (2012–14) and three cities in Japan (only for hours with extraterrestrial irradiance greater than 0 in each location).

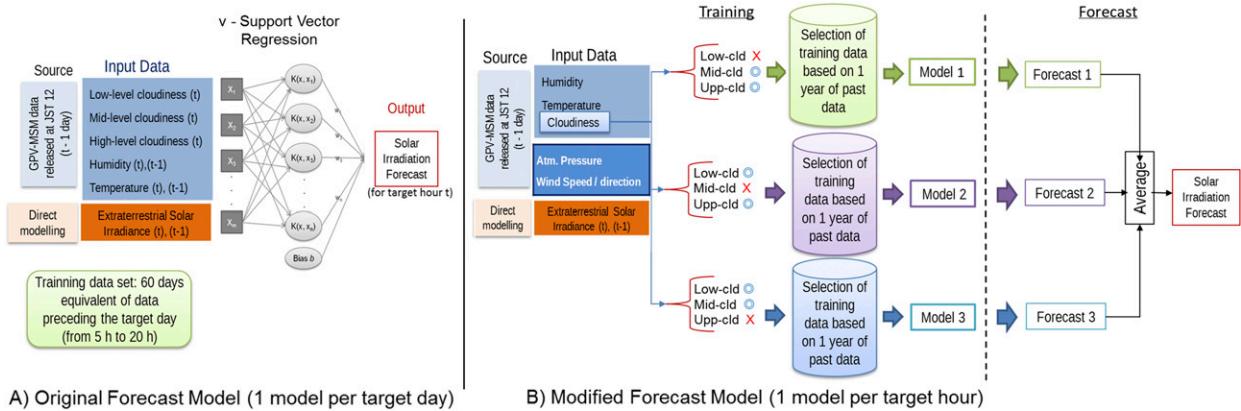|	Sapporo (northern Japan)	Tokyo	Fukuoka (western Japan)
Low-level cloudiness	−0.33	−0.49	−0.41
Mid-level cloudiness	−0.37	−0.52	−0.52
High-level cloudiness	−0.29	−0.38	−0.28

FIG. 2. (a) Initial and (b) modified machine-learning-based models used to forecast day-ahead solar irradiation.

much or too little weighting of such variables by the initial model should cause a reduction of the forecast error. That is the purpose of the third modification.

The third modification proposed is based on a three-member ensemble approach. Instead of training a single model with the three predicted cloudiness values as input variables, three models are trained, each with only two cloudiness variables. After that, the three models are used to make a forecast, and the three results are averaged.

The differences between the modified and initial models can be visualized by comparing Figs. 2a and 2b. In this study, the combined effect of all three modifications together is presented. A preliminary analysis of the individual effects of the modifications was presented in a workshop (Fonseca et al. 2018).

3. Target locations' description

Forecasts were done for 41 locations in Japan from the northernmost area, Hokkaido, to the southernmost one, Okinawa. They were done for every day, from 0500 to 2000 local time (UTC + 9 h), of the target period, from 1 January 2016 to 31 December 2017. Days with missing or faulty input or output data were excluded. The median of faulty days was 4 days per location with a minimum of 3 days and a maximum of 16 days over the 2 years of data studied.

To train the forecast models, the observed solar irradiation data from the JMA weather stations at each location were used. The data from the GPV-MSM closest grid point to the corresponding JMA weather stations were used as the input data.

Regional forecasts were also done to characterize the accuracy of methods at the regional level and to estimate the regional smoothing effect of the forecast error in Japan. They were calculated by averaging the corresponding values of individual locations within each targeted region. Both the machine-learning forecasts and the JMA-MSM regional forecasts were calculated in the same way. The target locations and regions are shown in Fig. 3.

Regional forecasts of solar irradiation are of particular interest to power utilities and stakeholders in the PV power market in Japan. Thus, the regional division adopted the geopolitical division of regions in Japan with their division according to the original operation area of the power utilities, and the availability of data. For example, since the power services of Niigata Prefecture (index 12 in Fig. 3) are managed by the same utility company that manages the power services of the entire Tohoku region, Niigata was added to the Tohoku region in the regional error calculations. The same was done for Kofu (index 22 in Fig. 3), which is in the Chubu region but is served by the power utility that provides services for all of Kanto region. Finally, the regions of Kansai, Chugoku, and Shikoku Island were regarded as a single region (the yellow area called Ks + Ch + Sh between Chubu and Kyushu in Fig. 3), because there are only a few weather stations measuring the solar irradiation in these regions individually.

4. Performance metrics

The performance of the forecast models was analyzed using the skill score (SS) defined in Eq. (6) and the RMSE in kilowatt hours per meter squared, given by Eq. (7):

$$\text{SS} = 1 - \frac{1/n\sum_{i=1}^{n}(H_{\text{fct},i} - H_{\text{msd},i})^2}{1/n\sum_{i=1}^{n}(H_{\text{ref},i} - H_{\text{msd},i})^2} \quad \text{and} \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_{\text{fct},i} - H_{\text{msd},i})^2}. \quad (7)$$
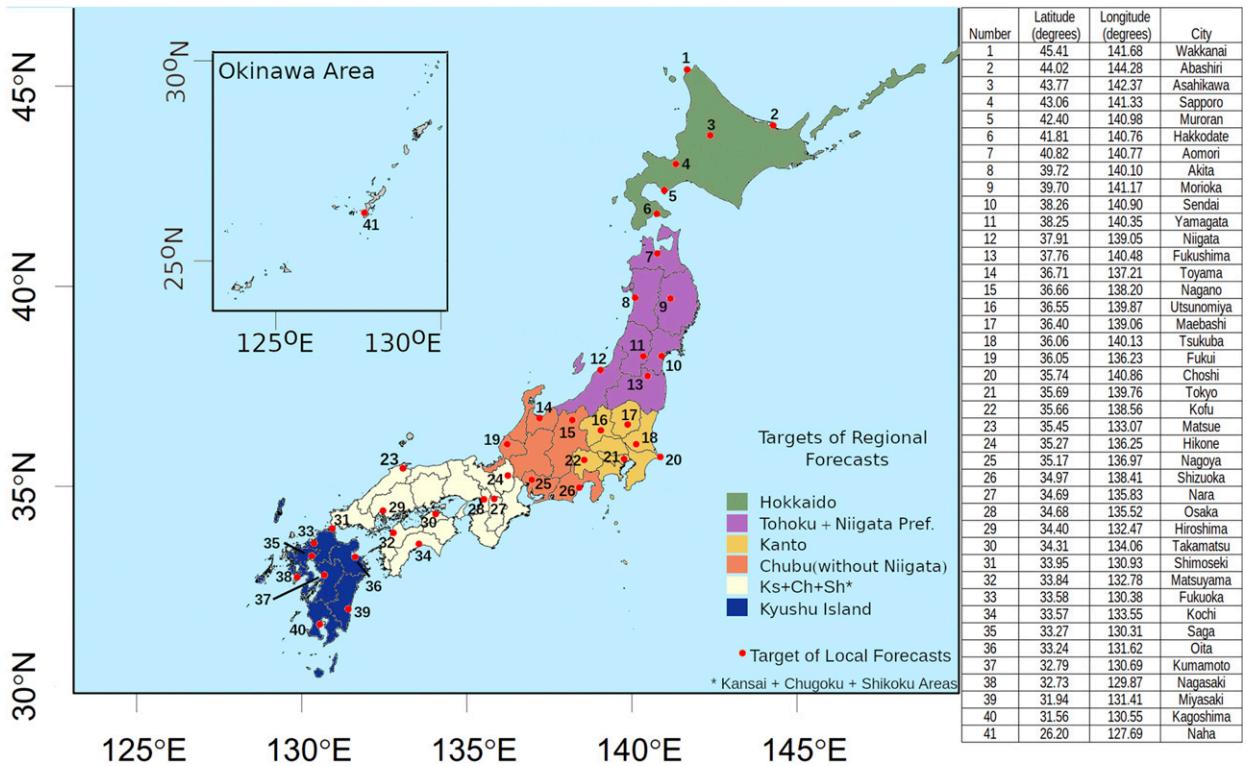
FIG. 3. (left) Location of the day-ahead forecast targets on local and regional scales in Japan. (right) Table with the respective location name, latitude (°N), and longitude (°E).

In Eqs. (6) and (7), $H_{fct,i}$ is the forecast of the global horizontal solar irradiation at hour i in kilowatt hours per meter squared, and $H_{msd,i}$ is the observed value of the same quantity at hour i, also in kilowatt hours per meter squared. In Eq. (6), $H_{ref,i}$ is the forecast of the global horizontal solar irradiation provided by a reference model. In these equations, n is the total number of hours for which the calculation was performed. As mentioned in section 2a, the day-ahead persistence of solar irradiation was used as an initial reference. However, the skill of the machine-learning models was also directly compared with the JMA-MSM forecasts. Hence, the reference model used in the SS calculations is always indicated in the related results.

For the RMSE, two normalized versions of it are also calculated. The first one, called $RMSE_n$, is the RMSE of a series of forecasts divided by the standard deviation of the observed values of the same location and period. As any mean error is a measure of deviation or dispersion (from observed values), it stands to reason to use as a normalizing factor, a typical measure of dispersion. However, as many similar studies use the mean of observed values as a normalizing factor, to facilitate comparisons with other studies available in the technical literature, we also present the RMSE normalized by such mean, defined here as the $RMSE_M$.

Forecast error is defined as the difference between the forecast value and what was regarded as the observed value. A positive value indicates an overestimation, and a negative value indicates an underestimation of the observed value. The Pearson correlation coefficient was used to evaluate the linear correlation between the forecast and observed values. Similar to the case of the $RMSE_n$, the standard deviations of the forecast values, when presented, were also normalized by the standard deviation of the observed values of solar irradiation for each location in the target period. Thus, a standard deviation of one for the forecasts of a given location means that they have the same standard deviation as that of the observed values. Finally, the concept of outliers was also used to identify and evaluate large forecast errors. Their definition follows the one used in the construction of a boxplot. An outlier was regarded as any forecast error outside the whiskers of the boxplot. The whiskers extend to the last point within 1.5 times the interquartile range of a forecast error distribution counted from the first and third quantiles, respectively.

5. Results

In this section, the results are divided into two sections, namely, local and regional forecasts. Local forecasts are
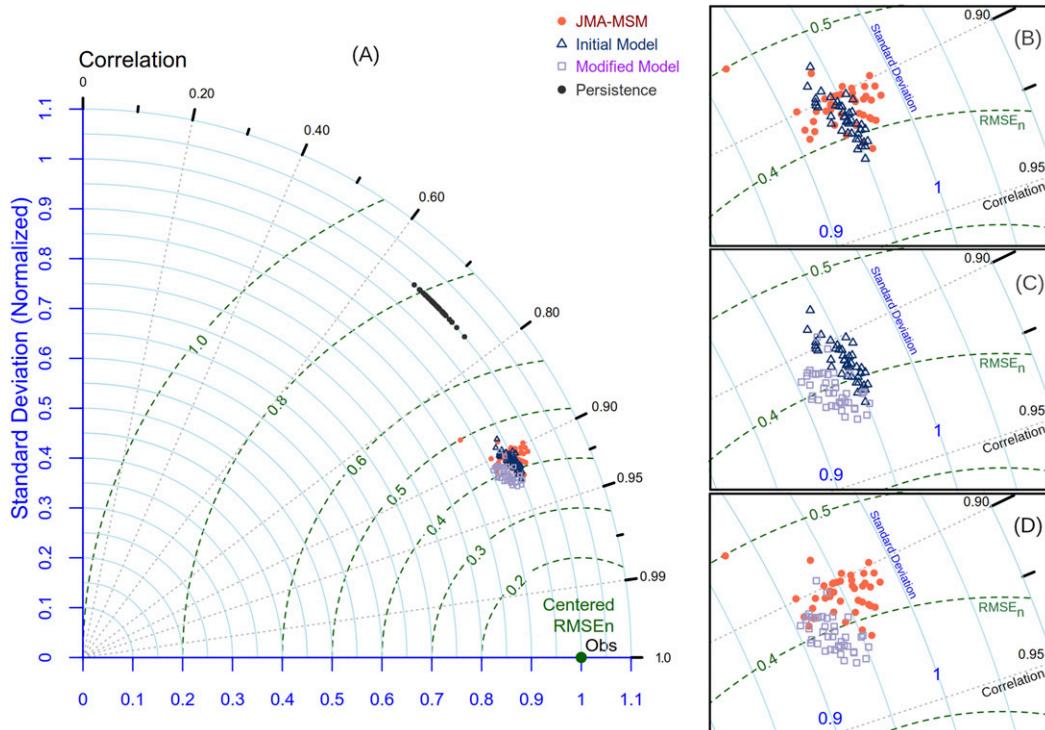
FIG. 4. (a) Full Taylor diagram of local day-ahead solar irradiation forecasts for 41 locations in Japan done with the three models and with persistence (2 years). (b)–(d) Magnification of the area, comparing the performance of two models at a time.

those made for a single location point (the red dots in Fig. 3). Regional forecasts are those obtained by up-scaling the local forecasts within the five regions, as shown in Fig. 3.

a. Local forecasts

To show the performance of the forecast models at a local scale throughout the country, a Taylor diagram of each model's results for each location during the 2-yr period studied is presented in Fig. 4a. In Figs. 4b–d, the area of the diagram containing the results is magnified, and the results of pairs of models are plotted in them to visualize their differences.

Figure 4a shows that on the local scale, the persistence-based forecast in Japan has a correlation coefficient with observed values varying from approximately 0.68 to 0.77. Moreover, the $RMSE_n$ of the persistence model varied from 0.65 to 0.82 standard deviations of the respective observed values.

Focusing on the three forecast models studied, Fig. 4a shows that they had an $RMSE_n$ ranging from 0.35 to 0.45, normalized standard deviations higher than 0.90, and a correlation coefficient between forecasts and observed values higher than 0.85. With regard to the exact variation of the $RMSE_n$ of the day-ahead solar irradiation

forecasts, it varied from 0.505 to 0.385 (23.7%) with the JMA-MSM, from 0.470 to 0.378 (19.6%) with the initial model, and from 0.446 to 0.366 (17.9%) with the modified model.

Furthermore, in Fig. 4a, it can be seen that the JMA-MSM showed poor performance in Naha (index 41 in Fig. 3), reaching an $RMSE_n$ of 0.5 standard deviations of the observed values. One reason for this poor perfor-mance is that its 5 km by 5 km grid spacing is not high enough to characterize the weather of small islands. Another reason is the tendency that the JMA-MSM has of not properly reproducing clouds typical of subtropical climate. For example, Ohtake et al. (2015) showed that the JMA-MSM forecasts have a bias in the direction of overestimation of solar irradiation, whenever the observed weather in the area of Okinawa had cumulonimbus clouds.

Machine-learning models are trained with the past data of their target location. Thus, they can learn (even if partially) the characteristics of their target location and are less affected by the limitations of physical models. Figure 4a shows that the machine-learning models' forecasts for Naha have $RMSE_n$ values consistent with the values found for other locations with the same models.
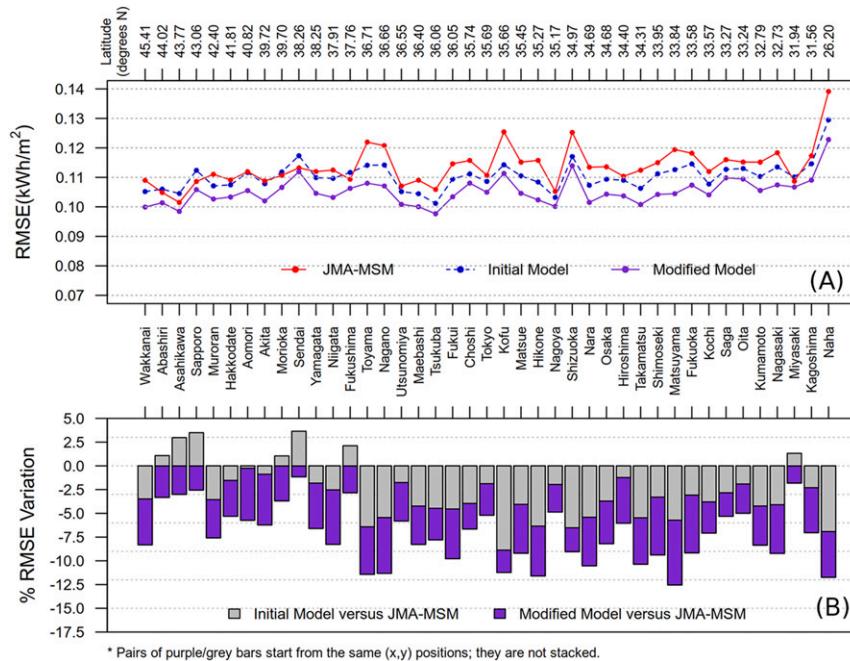
FIG. 5. (a) RMSE for 2 years of day-ahead forecasts of solar irradiation for 41 locations in Japan with three methods, and (b) the respective percent variation based on JMA-MSM forecast results.

Figure 4b shows that the JMA-MSM model had a performance similar to the initial machine-learning model only in a few locations. Moreover, the JMA-MSM forecasts were spread in the direction of the standard deviation axis with an $RMSE_n$ higher than 0.400 and a correlation above 0.90. In contrast, the initial model forecasts had less variation in the former and more variation in their $RMSE_n$ and correlation coefficient. Still, Fig. 4b shows that the $RMSE_n$ of the initial model forecasts was generally lower than that of the JMA-MSM model. The median of the $RMSE_n$ of the JMA-MSM forecasts throughout the country was 0.430 compared to 0.417 of the initial model. In comparison, Figs. 4c and 4d show that the modified model provided a further reduction of the $RMSE_n$ and further improvement of the correlation coefficient. With the modified model, the forecasts of 24 locations had an $RMSE_n$ lower than 0.400 and a correlation coefficient higher than 0.925. Such performance was achieved only for eight locations with the JMA-MSM model. Even for locations with $RMSE_n$ values greater than 0.400, Figs. 4b–d show that the modified model improves the correlation coefficients and the $RMSE_n$ of the forecasts (the median of the $RMSE_n$ of the forecasts of the modified model was 0.396).

On the other hand, the modified model caused a slight reduction in the standard deviation of the forecasts compared with the initial model's results. The median of

the normalized standard deviation of the forecasts with the modified model was 0.928. With the JMA-MSM, it was 0.953, and with the initial model, it was 0.947. We hypothesize that this reduction is related to the ensemble procedure described in section 2d.

In Fig. 5a, the RMSE (kW h m^{-2}) for 41 locations are provided. The locations are ordered according to the latitude from high to low (the same ordering used in Fig. 3). The corresponding percentage variations of the RMSE values of the machine-learning models based on the JMA-MSM results are plotted in Fig. 5b.

A general increase in the RMSE with the decrease in latitude is observed in Fig. 5a. The RMSE of JMA-MSM ranges from around 0.105 kW h m^{-2} at Hokkaido to around 0.115 kW h m^{-2} at Oita (index 36 in Fig. 3): a variation of 9.52% from Hokkaido to Kyushu Island. Including Naha in Okinawa in the evaluation makes the total RMSE variation range from 0.105 to 0.139 kW h m^{-2} (32.4%). The same variation was from 0.101 to 0.129 kW h m^{-2} (27.7%) with the initial model and from 0.098 to 0.123 kW h m^{-2} (25.5%) with the modified model.

In addition, Fig. 5b shows that the initial model yields better forecasts than the JMA-MSM model in all but 7 of the 41 locations studied. The results in Fig. 5 also show that the initial model's forecasts RMSE varied from −3% to +3% of the JMA-MSM forecasts RMSE (error reduction in 7 of 13 locations) in Hokkaido and Tohoku. However, for the rest of the country, the initial

TABLE 3. Five-number summary for the RMSE$_M$ (normalized by the mean of the observed values) of each forecast model over all 41 locations studied.

	Min	1st quantile	Median	3rd quantile	Max
JMA-MSM	0.415	0.467	0.490	0.506	0.552
Initial model	0.407	0.447	0.479	0.500	0.548
Modified model	0.395	0.425	0.456	0.477	0.520

model always had lower RMSE values than those for the JMA-MSM.

When compared with the results of the initial model, the results of the modified model showed reductions in the RMSE of the forecasts for all locations studied in Japan. When the baseline was the JMA-MSM model, even in the seven locations in which it outperformed the initial model, such trends were reversed with the modified model. Still using the JMA-MSM as the baseline, the RMSE was reduced with the modified model by almost 12.5% in the best case and 1.15% in the worst case.

To facilitate comparisons with other studies, Table 3 contains a summary of the RMSE values presented in Fig. 5, normalized by the mean of their respective observed values RMSE$_M$.

An important characteristic of a forecast model is its ability to consistently provide forecasts with low error. One way to assess such ability is to study the large forecast errors associated with a given model. In this study, we used the statistical definition presented in section 4. We focused on two characteristics of the outliers, namely, their average value and frequency. In Figs. 6 and 7, both characteristics are shown per location and per model.

Figure 6a shows that the average value of the outliers varied from 0.175 to 0.275 kW h m^{-2}, with frequent values around 0.20 kW h m^{-2}. In other words, the average values of large day-ahead forecast errors were around 20% of the maximum value of solar irradiation in Japan (where the peak value is around 1 kW h m^{-2}).
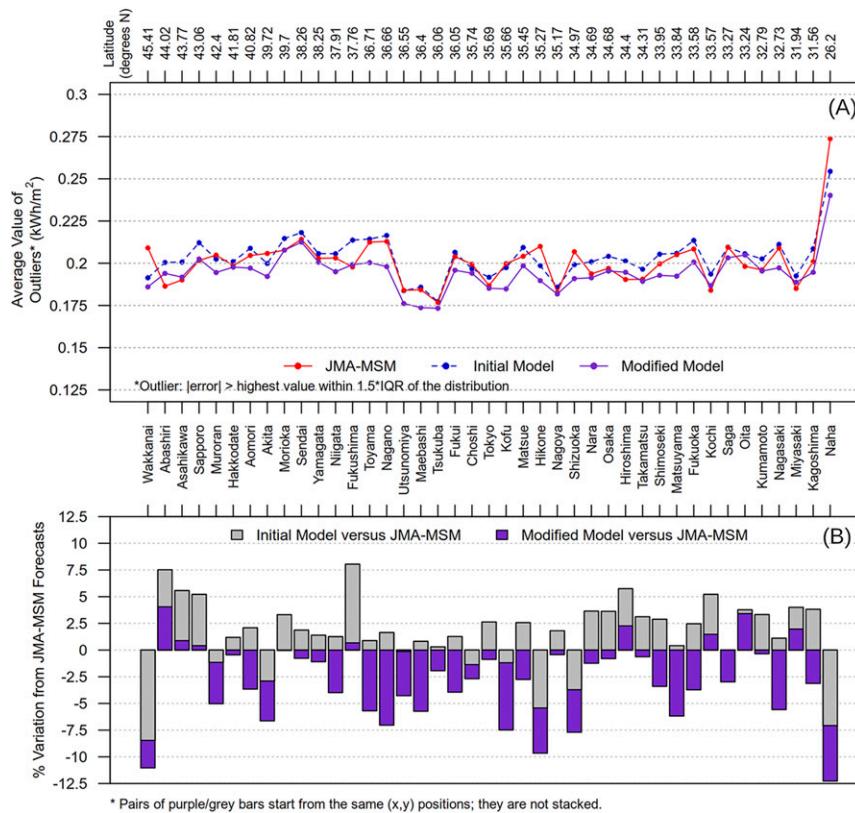

FIG. 6. (a) Average value of the outliers of forecast errors per location and model; (b) percent variation of the outliers' average value based on JMA-MSM results.
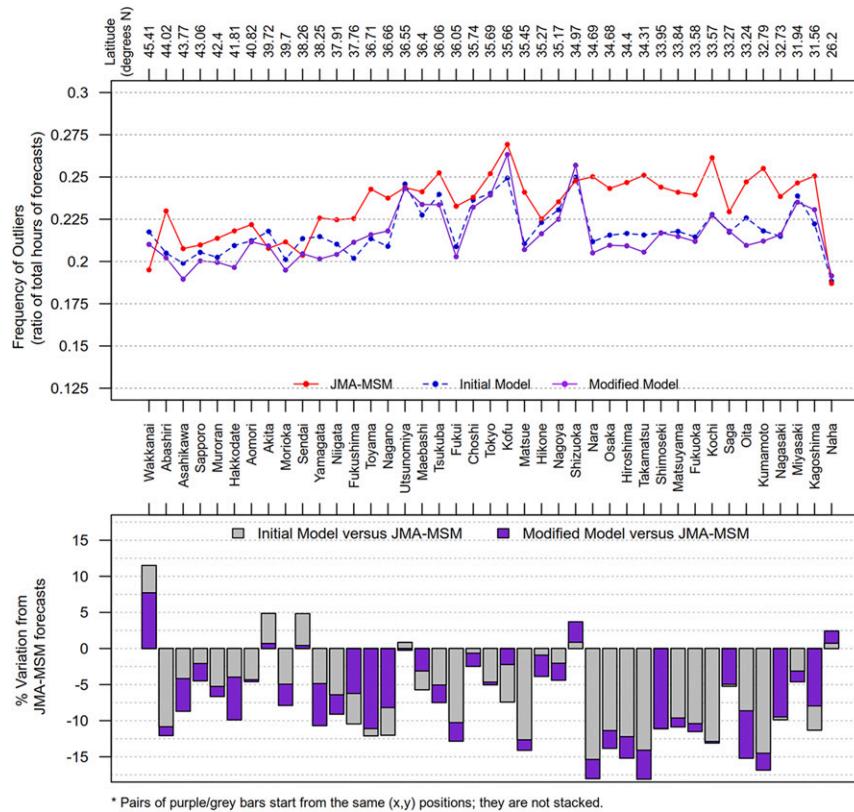
FIG. 7. (top) Outliers' frequency per location and model; (bottom) percent variation of outliers' frequency based on JMA-MSM results.

The forecast errors of the modified model had a lower average value of outliers in all the locations when compared with the initial model. Compared with the results of the JMA-MSM model, the former had lower average outlier values in 33 of the 41 locations targeted. In terms of the frequency of outliers, Fig. 7 (top panel) shows that they make up approximately 19%–27% of all the forecast errors in Japan. Moreover, the top panel of Fig. 7 also shows a clear trend of increasing frequency of outliers going southward with the JMA-MSM forecasts. With both initial and modified models, this trend is less accentuated. Figure 7 (bottom panel) shows that the modified model caused a reduction in the frequency of outliers of the forecast errors in all but 5 locations. With respect to the same criteria, the modified model also outperformed the initial model in 30 locations. The mean reduction in the frequency of outliers over the JMA-MSM results was also better than the one achieved with the initial model: 7.59% against 6.26%, respectively (not shown in Fig. 7).

Despite the overwhelmingly positive results, the modified model had a poor performance in a few locations. As mentioned in section 2d, the modified model should result in the reduction of forecast errors, when there is a

reasonable level of accuracy in the information provided by the predictors (so that the training pattern selection is effective), and when the cloudiness variables have some level of accuracy (so that the three-member ensemble can have a positive effect). When these assumptions are not satisfied, the forecast errors might become larger (as a model might end up being trained to forecast the solar irradiation of a weather pattern unrelated to the realized one). We hypothesized that in the few locations where the modified model had poor performance with regard to outliers, the high level of inaccuracy of predictors was the main cause.

In Fig. 8, we present the distribution per location of the monthly forecast skill scores during the 2-yr period. The notches in the boxplots follow the definition of McGill et al. (1978). According to McGill et al. (1978), nonoverlapping notches when comparing two or more series of data provide statistical evidence that the respective medians are different.

In Fig. 8a, it can be seen that the JMA-MSM forecasts throughout Japan show a noticeable low skill in July, August, and January. In the months of April, May, June, and October, its skill is at the same level as the other two machine-learning models. As shown in
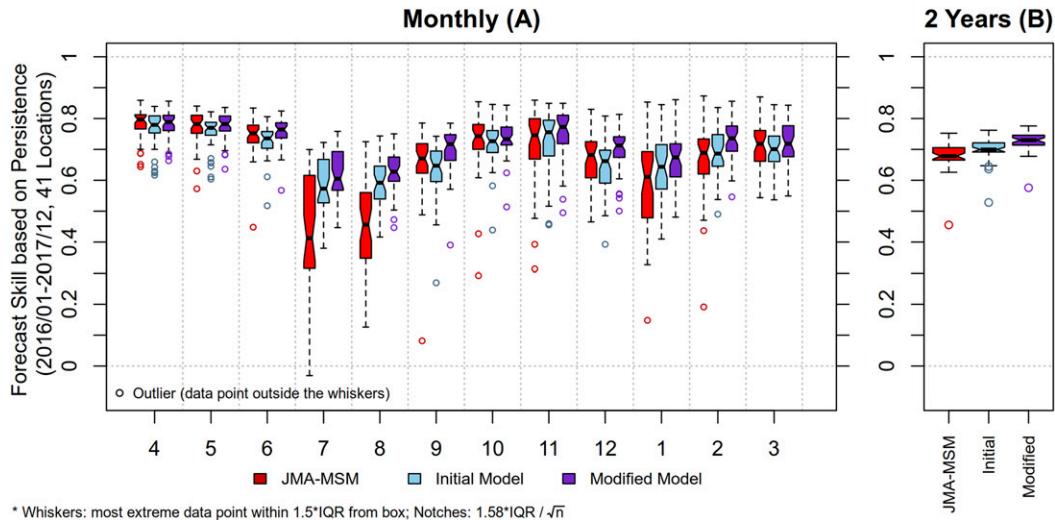
FIG. 8. (a) Monthly and (b) annual distribution of the forecast skill score of 41 locations in Japan, per forecast model, over 2 years.

Ohtake et al. (2015), the data from 2008 to 2012 suggests a strong bias of the JMA-MSM in summer and winter. This characteristic is clearly still present in the data of 2016 and 2017. The results in Fig. 8a also show that in 9 of the 12 months of the year, the JMA-MSM had outliers indicative of its lowest skill among the three models. Although not shown in Fig. 8a, the location with the worst forecast skill scores was usually in Hokkaido (5 months) and Naha (4 months). These results indicate a location dependency with regard to the skill of the JMA-MSM forecasts, in addition to the seasonal one.

The modified model showed the highest national forecast skill median for 9 months. However, in Fig. 8a, the notches suggest that statistically such difference from the results of the other two models, might be significant only in September and February. However, considering only the locations with the worst monthly forecast skills, the modified model yielded the best result in all months. We conclude from these that the proposed modifications were particularly effective in locations with originally poor forecast skill or high forecast errors, and also in summer and winter when the JMA-MSM more frequently exhibited poor performance.

Both the initial and modified models also show a reduction of skill in summer and winter, but the skill reductions are clearly less accentuated than that shown by the JMA-MSM model. Figure 8b shows the cumulative effects of the less accentuated drops in skill in summer and winter, and of the improvements achieved by the modified model in the locations with the worst monthly skill. They were important enough to make it yield an outlier with the highest skill among the three models,

and also the highest national median forecast skill (this time, with a clear indication of statistical significance).

b. Regional forecasts

A Taylor diagram with the regional results obtained using each model is shown in Fig. 9. Comparing the regional results with the local ones in Fig. 4, there was a perceptible reduction in the forecast error regardless of the model used. The $RMSE_n$ of the persistence model ranged from approximately 0.55 to 0.70. For local forecasts, the corresponding range was 0.65–0.82. The regional $RMSE_n$ for the three forecast models now varies around the 0.25 line as compared with 0.40 in the local case: a local to regional variation of 37.5%. The correlation coefficient between the forecasts and observed values was also better on the regional scale than on the local one. For the three forecast models, the correlation coefficient varied from approximately 0.95 to 0.975 on the regional scale and from 0.90 to 0.95 on the local one. There was also an improvement in the standard deviation of the forecasts. On the local scale, the forecasts had a standard deviation of around 92.5% of the standard deviation of the observed data. On the regional scale, this value was around 95%.

With regard to the performances of the three models, there was a similar trend to the one noted on the local scale. The JMA-MSM had the highest $RMSE_n$ and the lowest correlation coefficient in all the regions. Moreover, the JMA-MSM regional forecasts showed the highest $RMSE_n$ variation, a correlation variation compared with the other two models. These results indicate that on the regional scale also, the JMA-MSM performance has a strong spatial and geographical
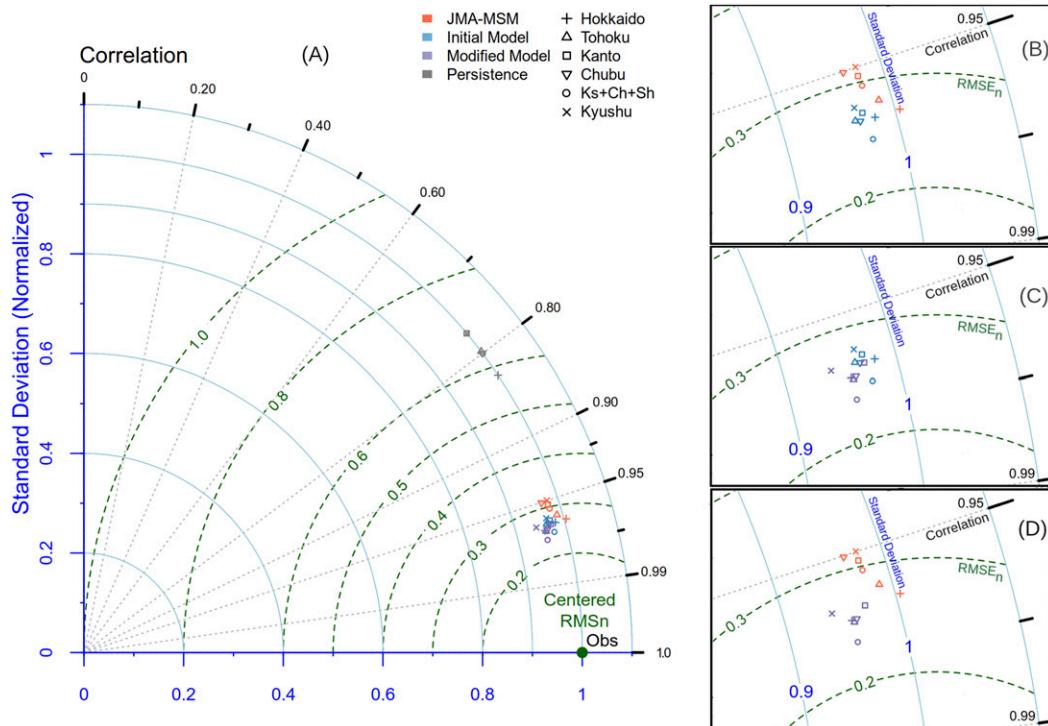
FIG. 9. (a) Full Taylor diagram of regional day-ahead solar irradiation forecasts for all regions in Japan done with the three models, and persistence (2 years). (b)–(d) Magnification of the area, comparing the performance of two models at a time.

dependency. The results in Fig. 9c also show that on the regional scale, the modified model improved the $RMSE_n$ and correlation coefficient in all regions, when compared with the initial model results.

In Fig. 10, we present the RMSE for the 2-yr period and the three models studied. For reference, in Table 4 the same RMSE is presented normalized by each region's respective mean of observed values. Figure 10 and Table 4 indicate that the modified model yielded meaningful reductions of the RMSE in all regions. When compared with the JMA-MSM-related results, Fig. 10 shows that RMSE reductions of up to 21.1%
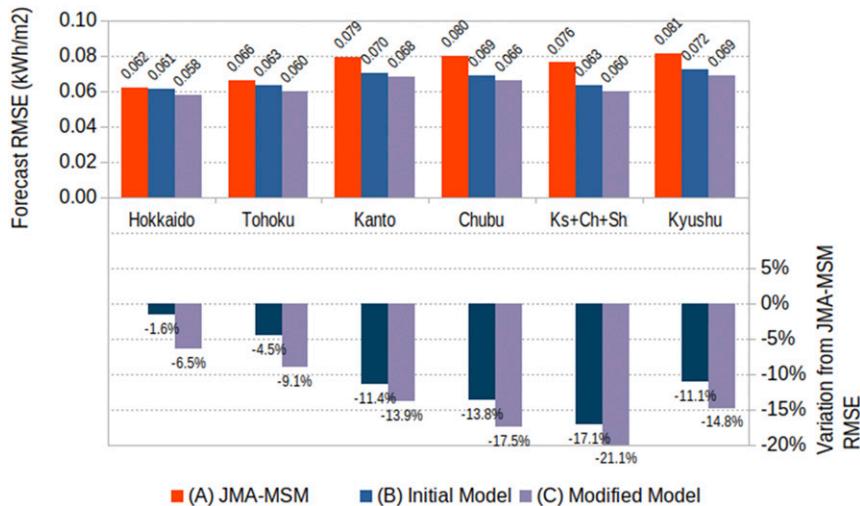

FIG. 10. Two-year RMSE of regional day-ahead forecasts of solar irradiation for all regions in Japan.

TABLE 4. Regional RMSE normalized by the mean of the observed values per region, $RMSE_M$.

	Hokkaido	Tohoku	Kanto	Ks + Sh + Ch	Kyushu
JMA-MSM	0.294	0.304	0.323	0.326	0.341
Initial Model	0.288	0.289	0.289	0.278	0.304
Modified Model	0.276	0.274	0.281	0.267	0.291

were achieved on the regional scale in the best case and 6.45% in the worst one.

The RMSE is sensitive to large errors. Thus, a stronger reduction of the RMSE is an indication that the modified method is working well during hours in which the other two methods had large errors. Further evidence of this is given in Table 5, which presents the average values of the outliers of the regional forecast errors. The initial machine-learning model already yields lower average values than the JMA-MSM. However, by using the modified model, further reductions in the average values of the outliers were achieved in all the regions.

The skill scores [Eq. (6)] of the regional forecasts, using persistence as a reference, is shown in Fig. 11a. The smoothing effect that occurred on the regional scale, resulted in a higher forecast skill and lower variation when compared with the local forecasts. The regional skill score varied from 0.75 to 0.86, according to the region and the forecast model, as compared with 0.40–0.80 in the local case.

The improvement in skill provided by the modified model over the initial model was small, varying from 1% to 2%. Still, in all the regions, the former yielded forecasts with the highest skill. As day-ahead persistence is regarded as the minimum threshold, which any forecast model is expected to outperform, the forecast skill scores based on persistence are expected to be high, and the difference between skill scores might look deceivingly small. To provide a better assessment, in Fig. 11b we present the skill of the forecasts using the JMA-MSM forecasts as a reference.

On average, the initial model had a skill 18.6% higher than that of the JMA-MSM, reaching its maximum value in the region comprising Kansai, Chugoku, and Shikoku. This region showed the best results in general

because it had the largest area among the regions studied. Consequently, it also had the strongest smoothing effect on errors. The modified model had a skill 25.3% higher (on average) than that of the JMA-MSM. This result demonstrates (on the regional scale as well) the effectiveness of the modifications proposed in section 2d in improving day-ahead forecasts of solar irradiation.

6. Conclusions

This study had two objectives: to propose a modified model to improve the accuracy of day-ahead forecasts of solar irradiation, and to perform a national evaluation of the accuracy of an NWP model and two machine-learning models in forecasting solar irradiation in the day-ahead horizon.

For the national evaluation, the results showed that on the local scale, the day-ahead forecasts of solar irradiation in Japan have an annual RMSE varying from 0.100 to 0.140 kW h m^{-2}. On the regional scale, values ranging from 0.060 to 0.070 kW h m^{-2} were found for the five regions studied. The skill score of the forecasts varied around 0.70 on the local scale and around 0.80 on the regional scale. The monthly skill scores of the forecasts showed that the machine-learning-based approaches were less sensitive to seasonal and spatial variations than the JMA-MSM.

With regard to the proposed modifications, the results show that although the initial machine-learning model was already slightly better than the JMA-MSM model in general, the proposed modifications caused a further reduction of the RMSE and improvement of the skill score of the forecasts. The modified method also yielded the lowest RMSE values on the local and regional scales. Compared with the JMA-MSM results, RMSE reductions varied from 1.15% to 12.5% on the local scale and

TABLE 5. Average value (kW h m^{-2}) of the outliers of forecast errors per region and per forecast model.

	Hokkaido	Tohoku	Kanto	Chubu	Ks + Ch + Sk	Kyushu
JMA-MSM	0.121	0.125	0.138	0.141	0.135	0.142
Initial model (variation over JMA-MSM)[a]	0.115 (−4.9%)	0.124 (−0.08%)	0.131 (−5.1%)	0.134 (−4.9%)	0.122 (−9.6%)	0.137 (−3.5%)
Modified model (variation over JMA-MSM)[a]	0.11 (−9.1%)	0.115 (−8%)	0.128 (−7.2%)	0.125 (−11.3%)	0.117 (−13.3%)	0.131 (−7.7%)

[a] Negative values indicate reduction of the average value of the outliers.
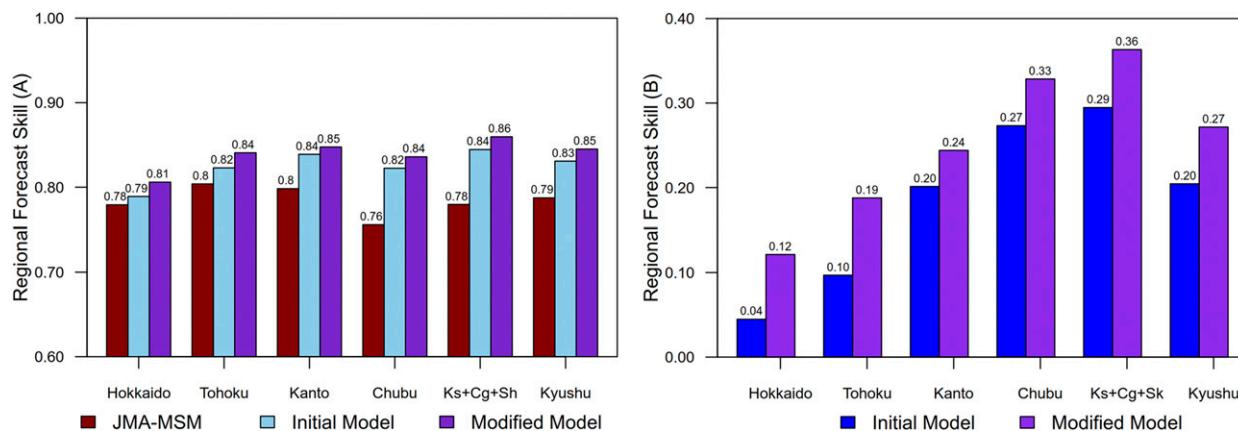
FIG. 11. Regional forecast skill scores of the three models using (a) using persistence as a reference and (b) of the two machine-learning models using the JMA-MSM as a reference.

from 6.45% to 21.1% on the regional scale. These values are within the same level of improvements as that achieved by other authors with more complex approaches in the problem of PV forecasting (Pierro et al. 2016) and solar irradiation forecasting (Perez et al. 2018). For the forecast skill score, it was improved over the JMA-MSM by at least 10%. The analyses of outliers indicated that the modified model reduced the average values of large errors as well as their frequencies and that it was effective in locations with forecasts that originally had low skill.

Two venues for further improvements in forecasts are anticipated based on the results of this study. One is about the further reduction of large errors. The proposed modifications were not effective when the predictors were highly inaccurate. Possible solutions to this problem include improving the predictors' accuracy or using different NWP models. The second venue for improvement comes from the findings of the comparisons with the JMA-MSM. In spring and autumn months, the JMA-MSM forecasts had comparable skills to the machine-learning models. Therefore, further improvements in day-ahead forecasts of solar irradiation may be achieved by combining JMA-MSM forecasts with those of the modified model. Both venues for improvement of the forecast will be explored in further studies.

Data availability statement: NWP solar irradiation data used in this study were allowed by the JMA only to members of the CREST project mentioned in the acknowledgments and to project-related activities. Other NWP data and the observed solar irradiation data used are available at the Japan Meteorological Business Support Center. Any other data used in the study are only available in the forms (graphs and tables) presented in the study to comply with the regulations of both funding projects mentioned in the acknowledgments and because of the restrictions with regard to their use by third parties.

REFERENCES

Alomar, M. L., V. Canals, J. L. Rosselló, and V. Martínez-Moll, 2016: Estimation of global solar radiation from air temperature using artificial neural networks based on reservoir computing. *Proc. EuroSun2016*, Palma de Mallorca, Spain, International Solar Energy Society, 1–8.

Antonanzas, J., N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, 2016: Review of photovoltaic power forecasting. *Sol. Energy*, **136**, 78–111, https://doi.org/10.1016/j.solener.2016.06.069.

Basterrech, S., L. Zjavka, L. Prokop, and S. Mišák, 2013: Irradiance prediction using echo state queueing networks and differential polynomial neural networks. *2013 13th Int. Conf. on Intelligent Systems Design and Applications*, Bangi, Malaysia, IEEE, 271–276.

Chang, C.-C., and C.-J. Lin, 2001: LIBSVM : A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.*, **2** (3), 1–27, https://doi.org/10.1145/1961189.1961199.

Cheng, H.-Y., C.-C. Yu, and S.-J. Lin, 2014: Bi-model short-term solar irradiance prediction using support vector regressors. *Energy*, **70**, 121–127, https://doi.org/10.1016/j.energy.2014.03.096.

Fonseca, J. G. S., Jr., T. Oozeki, H. Ohtake, K. Shimose, T. Takashima, and K. Ogimoto, 2011: Analysis of the use of support vector regression and neural networks to forecast insolation for 25 locations in Japan. *Proc. ISES Solar World Congress 2011*, Kassel, Germany, ISES, 4128–4135.

——, ——, ——, T. Takashima, and K. Ogimoto, 2015: Regional forecasts of photovoltaic power generation according to different data availability scenarios: A study of four methods.

Prog. Photovolt. Res. Appl., **23**, 1203–1218, https://doi.org/10.1002/pip.2528.

——, F. Uno, T. Oozeki, and K. Ogimoto, 2018: On the improvement of day-ahead forecasts of solar irradiance with simple ensembles and training data selection in Japan: A countrywide assessment. *Proc. Eighth Int. Workshop on the Integration of Solar Power into Power Systems*, Stockholm, Sweden, Energynautics GmbH, 4C_SIW18-273, [ISBN: 978-3-9820080-4-2].

Golnas, T., 2018: Overview of DOE solar forecasting II. *ESIG Forecast Workshop 2018*, Saint Paul, MN, ESIG, https://www.esig.energy/download/session-2-overview-of-doe-solar-forecasting-ii-foa-tassos-golnas/?wpdmdl=4203&refresh=5d93944952df21569952841.

Hamann, H. F., 2017: A multi-scale, multi-model, machine-learning solar forecasting technology. DOE Tech. Rep. DE-EE-0006017, 50 pp., https://doi.org/10.2172/1395344.

Hossain, M., S. Mekhilef, M. Danesh, L. Olatomiwa, and S. Shamshirband, 2017: Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. *J. Clean. Prod.*, **167**, 395–405, https://doi.org/10.1016/j.jclepro.2017.08.081.

International Energy Agency, 2018: Trends 2018 in photovoltaic applications. IEA Photovoltaic Power Systems Programme Rep. T1-34:2018, 88 pp., https://iea-pvps.org/wp-content/uploads/2020/01/2018_iea-pvps_report_2018.pdf.

Japan Meteorological Agency, 2014: The next generation non-hydrostatic model Asuca (in Japanese). NWP Section Tech. Supplement, 156 pp., https://www.jma.go.jp/jma/kishou/books/nwpreport/60/No60_all.pdf.

——, 2017: On the provisioning of solar irradiation forecast data (in Japanese). Accessed 10 April 2019, https://www.jma.go.jp/jma/press/1712/05a/20171205_nissha.html.

Japan Meteorological Business Support Center, 2019: Specification of the mesoscale model GPV (MSM) (in Japanese). JMBSC, accessed 25 June 2019, http://www.jmbsc.or.jp/jp/online/file/f-online10200.html.

Jiménez, P. A., and Coauthors, 2016a: WRF-Solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction. *Bull. Amer. Meteor. Soc.*, **97**, 1249–1264, https://doi.org/10.1175/BAMS-D-14-00279.1.

——, S. Alessandrini, S. E. Haupt, A. Deng, B. Kosovic, J. A. Lee, and L. Delle Monache, 2016b: The role of unresolved clouds on short-range global horizontal irradiance predictability. *Mon. Wea. Rev.*, **144**, 3099–3107, https://doi.org/10.1175/MWR-D-16-0104.1.

Lauret, P., C. Voyant, T. Soubdhan, M. David, and P. Poggi, 2015: A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. *Sol. Energy*, **112**, 446–457, https://doi.org/10.1016/j.solener.2014.12.014.

Le Cadre, H., I. Aravena, and A. Papavasiliou, 2015: Solar PV power forecasting using extreme learning machine and information fusion. *European Symp. on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, Belgium, HAL-01145680, https://www.archives-ouvertes.fr/hal-01145680/document.

Li, Z., C. Zang, P. Zeng, H. Yu, and H. Li, 2015: Day-ahead hourly photovoltaic generation forecasting using extreme learning machine. *2015 IEEE Int. Conf. on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, Shenyang, China, IEEE, 779–783.

McGill, R., J. W. Tukey, and W. A. Larsen, 1978: Variations of box plots. *Amer. Stat.*, **32**, 12–16, https://doi.org/10.1080/00031305.1978.10479236.

Ministry of Economy, Trade and Industry, 2014: On the revision of the operation of the feed-in tariff program aiming large-scale introduction of renewable energy (in Japanese). METI Japan, accessed 17 August 2018, https://warp.da.ndl.go.jp/info:ndljp/pid/9450762/www.meti.go.jp/press/2014/12/20141218001/20141218001.html.

Ohtake, H., J. G. da Silva Fonseca, T. Takashima, T. Oozeki, K. Shimose, and Y. Yamada, 2015: Regional and seasonal characteristics of global horizontal irradiance forecasts obtained from the Japan Meteorological Agency meso-scale model. *Sol. Energy*, **116**, 83–99, https://doi.org/10.1016/j.solener.2015.03.020.

——, F. Uno, T. Oozeki, Y. Yamada, H. Takenaka, and T. Y. Nakajima, 2018: Outlier events of solar forecasts for regional power grid in Japan using JMA mesoscale model. *Energies*, **11**, 2714, https://doi.org/10.3390/en11102714.

Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.

Pereira, S., P. Canhoto, R. Salgado, and M. J. Costa, 2019: Development of an ANN based corrective algorithm of the operational ECMWF global horizontal irradiation forecasts. *Sol. Energy*, **185**, 387–405, https://doi.org/10.1016/j.solener.2019.04.070.

Perez, R., J. Schlemmer, S. Kivalov, J. Dise, P. Keelin, M. Grammatico, T. E. Hoff, and A. Tuohy, 2018: A new version of the SUNY Solar Forecast Model: A scalable approach to site-specific model training. *Proc. Seventh World Conf. on Photovoltaic Energy Conversion*, Waikoloa, HI, IEEE, http://asrc.albany.edu/people/faculty/perez/2018/IEEE-2018.pdf.

Pierro, M., F. Bucci, M. De Felice, E. Maggioni, D. Moser, A. Perotto, F. Spada, and C. Cornaro, 2016: Multi-model ensemble for day ahead prediction of photovoltaic power generation. *Sol. Energy*, **134**, 132–146, https://doi.org/10.1016/j.solener.2016.04.040.

Powers, J. G., and Coauthors, 2017: The weather research and forecasting model: Overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc.*, **98**, 1717–1737, https://doi.org/10.1175/BAMS-D-15-00308.1.

Qing, X., and Y. Niu, 2018: Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, **148**, 461–468, https://doi.org/10.1016/j.energy.2018.01.177.

Saito, K., and Coauthors, 2006: The operational JMA non-hydrostatic mesoscale model. *Mon. Wea. Rev.*, **134**, 1266–1298, https://doi.org/10.1175/MWR3120.1.

——, J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita, and Y. Honda, 2007: Nonhydrostatic atmospheric models and operational development at JMA. *J. Meteor. Soc. Japan*, **85B**, 271–304, https://doi.org/10.2151/jmsj.85B.271.

Schölkopf, B., P. Bartlett, A. Smola, and R. Williamson, 1998: Support vector regression with automatic accuracy control. *ICANN 1998: Perspectives in Neural Computing*, L. Niklasson, M. Bodén, and T. Ziemke, Eds., Springer, 111–116.

Sommeria, G., and J. W. Deardorff, 1977: Subgrid-scale condensation in models of nonprecipitating clouds. *J. Atmos. Sci.*, **34**, 344–355, https://doi.org/10.1175/1520-0469(1977)034<0344:SSCIMO>2.0.CO;2.

Srivastava, S., and S. Lessmann, 2018: A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Sol. Energy*, **162**, 232–247, https://doi.org/10.1016/j.solener.2018.01.005.

Uno, F., H. Ohtake, M. Matsueda, and Y. Yamada, 2018: A diagnostic for advance detection of forecast busts of regional surface solar radiation using multi-center grand

ensemble forecasts. *Sol. Energy*, **162**, 196–204, https://doi.org/10.1016/j.solener.2017.12.060.

Voyant, C., G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, 2017: Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, **105**, 569–582, https://doi.org/10.1016/j.renene.2016.12.095.

Watanabe, T., T. Takamatsu, and T. Y. Nakajima, 2016: Evaluation of variation in surface solar irradiance and clustering of observation stations in Japan. *J. Appl. Meteor. Climatol.*, **55**, 2165–2180, https://doi.org/10.1175/JAMC-D-15-0227.1.

Zendehboudi, A., M. A. Baseer, and R. Saidur, 2018: Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.*, **199**, 272–285, https://doi.org/10.1016/j.jclepro.2018.07.164.