




## Article

# Regional Solar Irradiance Forecast for Kanto Region by Support Vector Regression Using Forecast of Meso-Ensemble Prediction System

Takahiro Takamatsu <sup>1,\*</sup>, Hideaki Ohtake <sup>1,2</sup>, Takashi Oozeki <sup>1</sup>, Toshiyuki Nakaegawa <sup>2</sup>, Yuki Honda <sup>3</sup> and Masahiro Kazumori <sup>3</sup>

<sup>1</sup> National Institute of Advanced Industrial Science and Technology, Fukushima 963-0298, Japan; hideaki-ootake@aist.go.jp (H.O.); takashi.oozeki@aist.go.jp (T.O.)

<sup>2</sup> Meteorological Research Institute, Ibaraki 305-0052, Japan; tnakaega@mri-jma.go.jp

<sup>3</sup> Japan Meteorological Agency, Tokyo 105-8431, Japan; honda.yuuki@met.kishou.go.jp (Y.H.); kazumori@met.kishou.go.jp (M.K.)

\* Correspondence: takamatsu.1849@aist.go.jp; Tel.: +81-29-861-0521

**Abstract:** From the perspective of stable operation of the power transmission system, the transmission system operators (TSO) needs to procure reserve adjustment power at the stage of the previous day based on solar power forecast information from global horizontal irradiance (GHI). Because the reserve adjustment power is determined based on information on major outliers in past forecasts, reducing the maximum forecast error in addition to improving the average forecast accuracy is extremely important from the perspective of grid operation. In the past, researchers have proposed various methods combining the numerical weather prediction (NWP) and machine learning techniques for the one day-ahead solar power forecasting, but the accuracy of NWP has been a bottleneck issue. In recent years, the development of the ensemble prediction system (EPS) forecasts based on probabilistic approaches has been promoted to improve the accuracy of NWP, and in Japan, EPS forecasts in the mesoscale domain, called mesoscale ensemble prediction system (MEPS), have been distributed by the Japan Meteorological Agency (JMA). The use of EPS as a machine learning model is expected to improve the maximum forecast error, as well as the accuracy, since the predictor can utilize various weather scenarios as information. The purpose of this study is to examine the effect of EPS on the GHI prediction and the structure of the machine learning model that can effectively use EPS. In this study, we constructed the support vector regression (SVR)-based predictors with multiple network configurations using MEPS as input and evaluated the forecast error of the Kanto region GHI by each model. Through the comparison of the prediction results, it was shown that the machine learning model can achieve average accuracy improvement while reducing the maximum prediction error by MEPS, and knowledge was obtained on how to effectively provide EPS information to the predictor. In addition, machine learning was found to be useful in improving the systematic error of MEPS.

**Keywords:** meso-ensemble prediction system; support vector regression; solar irradiance forecast; ensemble learning



**Citation:** Takamatsu, T.; Ohtake, H.; Oozeki, T.; Nakaegawa, T.; Honda, Y.; Kazumori, M. Regional Solar Irradiance Forecast for Kanto Region by Support Vector Regression Using Forecast of Meso-Ensemble Prediction System. *Energies* **2021**, *14*, 3245. <https://doi.org/10.3390/en14113245>

Academic Editor: Marco Pasetti

Received: 14 April 2021

Accepted: 24 May 2021

Published: 2 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the basic structure of the electricity market where the power generation sector and the transmission sector are separated, the transmission system operator (TSO) bears responsibility for coordinating the gap from the balancing group's (BG's) plan of the supply and demand in real time. Therefore, TSO is bound to compensate for the imbalance in real time due to the error of the day-ahead forecast of the photovoltaic (PV) power supply by the power resource. Hence, the cost of securing a power source that is not used in the actual supply and demand is always incurred as a countermeasure against the case

of excessive solar radiation, and this cost will ultimately be borne by the consumer. For this reason, in addition to improving the average accuracy of solar radiation forecasts, it is important to reduce excessive forecasts or the huge error in forecasting as of the day before the actual power supply and demand for efficient operation of the electric power system; that is, minimizing reserve power on standby and maintaining optimal operating conditions of the plant [1].

Weather forecast based on the Numerical Weather Prediction (NWP) is advantageous in the 1–3 day-ahead forecast task, which is related to the transmission schedule and/or day-ahead electricity market [2]: in Japan, Grid Point Value (GPV) of the Mesoscale Model's (MSM's) prediction presented by Japan Meteorological Agency (JMA) is widely used. However, NWP has the limitation derived from its insufficient spatial resolution to resolve the detail of the cloud, which causes the error of the local effect, and systematic error is contained from the pre-process operation for the stable acquisition of the numerical solution. To refine the NWP, post-processing methods are frequently used: machine learning methods are applied to NWP. Such as artificial neural-network (ANN) [3], hybrid model of auto-regressive moving average model (ARMA) and ANN [4], and support-vector machine (SVM) [5].

By constructing a predictor based on the NWP, more accurate global horizontal irradiance (GHI) forecast becomes possible, while there is the limitation on the accuracy of NWP. Although NWP is calculated from a deterministic approach, its forecast errors may occur due to the boundary condition and/or initial values, and the error can rapidly grow by the chaotic behavior of the atmosphere [6]. For the above intrinsic unpredictability, machine learning approach based on NWP has adopted various ensemble techniques: multiple outputs of different machine learning modes [7], shifting the forecast horizon of the NWP [8], picking up a historical data similar to patterns of predictions by analog ensemble [9] or k-nearest neighbor (k-NN) [10]. Furthermore, recent researchers applied probabilistic forecasting approach to solar power forecasting from NWP. As the non-parametric approach, the quantile regression (QR) is combined with machine learning technique. The QR makes possible to estimate  $\tau$ -th quantile without assuming the probability distribution [11], and this approach is useful for the probabilistic solar power forecasting because of the heterogeneous dispersion of solar radiation. M. P. Almeida et al. applies NWP data to the QR based on the random forests and estimate the quantiles for predicting the distribution of PV generation [12]. R. Judan et al. applied the multiple QR with mapping to the radial basis function (RBF) kernel for the solar power forecasting in the 2014 Global Energy Forecasting Competition (GEFCom2014) [13]. Y. He et al. utilize the support vector machine for the QR with the fuzzy information granulation [14]. On the other hand, as the parametric approach, J. Fonseca Jr. et al. estimated the confidence interval with assuming probability distribution function by comparing the input data with the historical data of the previous 60 days [15]. However, the accuracy of the NWP presumably become the bottleneck of the deterministic solar power forecasting and the probabilistic one.

To take measures for the problem of the NWP based on the deterministic approach, alternatively, Ensemble Prediction System (EPS) has been studied: stochastic approach combined with the deterministic NWP model [16,17]. Ensemble mean of the multiple forecasts, which generated by adding the appropriate perturbation to the base NWP model, gives better performance. Besides, in 2018, F. Uno et al. reported the correlation between the huge GHI forecast errors and the variances of the EPS members through the analysis of the four leading NWP centers' EPSs [18]. From the above, it is possible to improve the accuracy of weather forecasting and analyze the reliability information by using EPS.

Operational forecasting of EPS was first introduced to the global scale model by National Meteorological Center (NMC) in 1992 [19]. With the development of the EPS for the global scale model, operational centers have expanded the EPS to the regional scale model [20]. Today, forecasting systems that apply ensemble prediction to the mesoscale region model have been put into the operational forecast in various countries: The United States National Centers for Environmental Prediction (NCEP) started operating short

range ensemble forecasts (SREF) in 2001 [21]; United Kingdom Met Office (UKMO) started Met Office Global and Regional Ensemble Prediction System (MOGREPS) in 2005 [22]; Meteorological Service of Canada (MSC) started Regional Ensemble Prediction System (REPS) in 2011 [23]; Deutscher Wetterdienst (DWD) started ensemble prediction system based on the Consortium for Small-scale Modeling for German region (COSMO-DE-EPS) in 2012 [24].

For the regional scale forecast in Japan region, JMA started operational forecast of the Meso-Ensemble Prediction System (MEPS): regional EPS based on the MSM. MEPS generates 21 member forecasts by singular vector (SV) method, and it provides the up to 39 hour-ahead forecast. MEPS is adept at representing a variety of meteorological scenarios. Hence, by combining MEPS data with machine learning technique, more accurate and versatile day-ahead solar irradiance forecast system for Japan region can be expected. In other words, by the machine learning model referring to MEPS data, the one day-ahead solar power forecasts' precision can assumedly be improved by correcting the bias in MEPS with suppressing the huge forecasting error of concern to TSO and BGs.

There are several works applying EPS data to the machine learning models for the short term prediction of the renewable resource or the weather data by using EPS data from European Centre for Medium-Range Weather Forecasts (ECMWF). C. Junk et al. applied ensemble model output statistics (EMOS) and analog-based EMOS to ECMWF's EPS data for wind power forecasting [25], S. Sperati et al. applied neural-network (NN) to ECMWF's EPS data for solar power forecasting [26], and S. Rasp and S. Lerch used NN with ECMWF's EPS data for estimating air temperature at 2 m [27]. L. Massidda and M. Marrocu applied QR to each ECMWF's EPS member and combined Integrated Forecasting System (IFS) data with their QR outputs [28]. However, to our best knowledge, there are few research on the GHI forecast based on the machine learning approach using EPS data despite the works in References [26,28].

Note that, in previous research, ECMWF's EPS forecasts are only used for the machine learning-based solar power forecasting, and few studies have focused on the maximum prediction error, regardless of its practical importance for the power transmission systems. In addition, the affection of the difference of the machine learning models' network which process the EPS's multiple members is not sufficiently investigated. It is necessary to study the configuration that enables machine learning models to utilize EPS forecasts effectively.

In order to improve the average accuracy and reduce the maximum prediction error by the machine learning model using EPS forecasts, and to verify the effective way of constructing the machine learning model using EPS predictions, this study constructed several network configurations of SVR-based predictors for regional GHI prediction in the Kanto region using MEPS forecasts as input, and verified the accuracy of each predictor by the ground observation data at meteorological stations.

By comparing the results of regional GHI prediction of each predictor, we discussed how to effectively construct a machine learning model using EPS. Besides, by comparing the prediction results with those of MSM and MEPS, we confirmed that the machine learning model is effective in improving the systematic error of GHI prediction by MEPS. Note that the MEPS forecast used in this paper incorporates a new model from the JMA for the calculation of physical processes, and mentions its bias error in the discussion. Furthermore, in addition to improving the average accuracy of the predictor, we showed that the maximum prediction error, which is of practical importance, can be reduced by the machine learning model using EPS.

The paper is organized as follows. In Section 2, we denote the data description. Section 3 explains the fundamentals of our study. Support vector regression and evaluation methods. Section 4 shows our prediction models, and their precision are discussed in Section 5. In Section 6, conclusions are denoted.

## 2. Data Description

### 2.1. Meso-Scale Ensemble Prediction System

Forecasts that are calculated from accurately estimated initial and boundary values are called deterministic forecasts, and higher accuracy has been promoted by improving observation systems and data assimilation technology, improving NWP models, and introduction of high-performance computers. As such deterministic forecasts, in Japan region, MSM data is distributed from the JMA.

JMA works diligently for disaster prevention, such as an intensive rain storm, through the improvement of the MSM. However, because the error of deterministic forecasts may increase depending on the initial variables and boundary condition by the chaotic behavior of the atmosphere, its intrinsic predictability is limited. Even though the practical predictability can be extended through the improvement of the NWP models, it is necessary to objectively grasp their reliability through the probabilistic approach.

Ensemble prediction enables to obtain the stochastic characteristics of the weather forecasts through the multiple predictions of the members created by perturbing the numerical conditions. By utilizing the information from ensemble members' behavior, efficient risk management becomes possible. Besides, their expected value can improve the precision of the forecast.

For the examination of the regional EPS, JMA started the development of mesoscale SV in 2005. After a number of comparative tests, MEPS had developed for the implementation from 2012. Along with the operation of the 10th Generation Numerical Analysis Prediction System (NAPS10), the operational forecast of MEPS started in 2019 [29].

MEPS is the system generating 21 ensemble members by perturbing the initial and lateral boundary conditions of the control member, which is equal to MSM-GPV: their perturbations are given by the meso SV and the global SV with ensuring their consistency. Forecasts of MEPS are operated at 00, 06, 12, 18 UTC, and their forecasts can be made up to 39 h in advance. The spatial resolution of MEPS is 5 km. GHI forecasts from MEPS are averaged over the previous hour, and the ground measurements used for comparison are converted to average value of the previous hour according to the MEPS's GHI. Note that MEPS forecasts used in our work adopt the asuca model, which is the improved model of dynamical processes from conventional JMA Non-Hydrostatic Model (JMA-NHM) (see more details in Reference [30]).

GHI forecasts provided from MEPS can be useful for solar power forecasts; however, MEPS is coordinated for windstorm disaster prevention, and the control-run of the MEPS has systematic errors due to inconsistencies in cloud cover calculations between cloud microphysical processes and radiative and boundary layer processes.

For such biases caused by sub-grid scale physics, it is expected that correction by machine learning technique is useful. By combining machine learning techniques with MEPS, it is expected that it will be possible to build an one day-ahead solar power forecasting model that can respond to various situations while correcting for systematic errors in MEPS.

### 2.2. Input Data

In our work, we aim to verify the improvement of the global horizontal irradiation (GHI) forecast from MSM or MEPS through the prediction for the regional GHI averaged at five points in the Kanto region (Maebashi, Tsukuba, Utsunomiya, Tokyo, Choshi) at 1-h intervals from 6 June 2018 to 6 October 2018.

Our prediction models' explanatory variables consist of the weather forecasts and the theoretical total solar irradiation intensity: MSM-GPV's forecast data (temperature and relative humidity on a surface, high/middle/low level cloud covers), MEPS's GHI forecast obtained from 21 ensemble members, the calculated values of the solar insolation at the top of atmosphere on a horizontal surface [31]. For one day-ahead forecast, we adopt the forecast data delivered at JST 15:00 of the previous day of the machine learning prediction

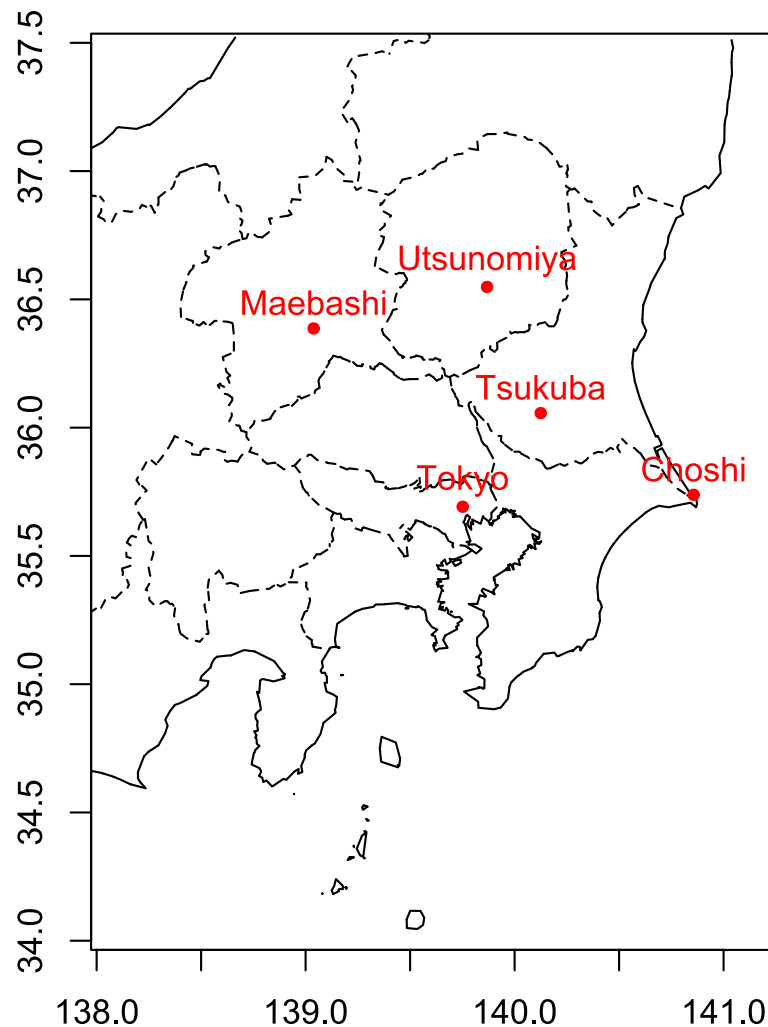
target date. We describe the MSM datasets excluded GHI forecast data and GHI from MEPS of each site as  $MSM_s(h)$ ,  $\widehat{GHI}_s(h)$ , and we denote them as follows.

$$MSM_s(h) = [\widehat{T}_s(h) \quad \widehat{RH}_s(h) \quad \widehat{HC}_s(h) \quad \widehat{MC}_s(h) \quad \widehat{LC}_s(h)], \quad (1)$$

$$\widehat{GHI}_s(h) = [\widehat{GHI}_s^0(h) \quad \widehat{GHI}_s^1(h) \quad \dots \quad \widehat{GHI}_s^m(h) \quad \dots \quad \widehat{GHI}_s^{N_m}(h)], \quad (2)$$

where  $s = 1, 2, \dots, N_s$  and  $m = 0, 1, \dots, N_m$  indicate the index number of the sites and the members of MEPS,  $h$  is the hour of the forecast date,  $\widehat{T}_s(h)$  and  $\widehat{RH}_s(h)$  are the forecast values of the temperature and the relative humidity on a surface, and  $\widehat{HC}_s(h)$ ,  $\widehat{MC}_s(h)$  and  $\widehat{LC}_s(h)$  are the forecasts of the high / middle / low level cloud covers. In particular, MSM's GHI forecast data correspond to the control member of the MEPS members, that is tagged by the 0-th number ( $m = 0$ ). Each site's extra-atmospheric solar radiation on a horizontal surface is denoted by  $E_s(h)$ .

For the training and evaluation of the predictor, we use the ground-based solar radiation data observed by the JMA station at the target site in Figure 1 as the objective variable.



**Figure 1.** Location of the JMA Stations. JMA records ground-based observation data for GHI at 5 station in Kanto region. Maebashi, Tsukuba, Utsunomiya, Tokyo, Choshi.

### 3. Fundamentals

#### 3.1. Support Vector Regression

SVR is the machine learning technique that applies the support vector networks [32] to the regression problem. K. R. Müller et al. and H. Drucker et al. showed the SVR obtained the excellent performance in the time series forecast [33,34]. In case time series data is given by  $(x_i, y_i), i = 1, 2, \dots, n$ , where  $x_i$  is the explanatory variable, and  $y_i$  the objective variable, the basic idea of SVR is the optimization technique for estimating the regression formula denoted by

$$\hat{f}(x) = \sum_{i=1}^n \alpha_i K(x_i, x) + \beta, \quad (3)$$

where  $K(x_p, x_q)$  is the kernel function,  $\alpha_i$  is the weight parameter, and  $\beta$  is the bias. SVR firstly projects the datasets into the feature space given by the non-linear map function linking to the kernel function, and secondly estimates the line minimizing the loss function's penalty in the feature space through the optimization of the dual problem. Standard SVR formula adopts the RBF kernel and the  $\epsilon$ -insensitive loss function:

$$K_{RBF}(x_p, x_q) = e^{-\frac{\|x_p - x_q\|_2^2}{2\sigma^2}}, \quad (4)$$

$$l_\epsilon(y, f(x)) = \max(|y - f(x)| - \epsilon, 0). \quad (5)$$

$\epsilon$  corresponds to the size of insensitive tube, and its optimal value depends on the noise level in the datasets.  $\sigma^2$  is the variance of the Gaussian function for the data fitting. For constructing support vector machine tuning  $\epsilon$ -insensitive tube size automatically, Scholköpf et al. developed the  $\nu$ -SVR [35]. Through using  $\nu$ -SVR,  $\epsilon$  parameter can be automatically adjusted without assuming the noise level. The implementation method and kernel functions of SVR are detailed in Reference [36].

In our study, we applied  $\nu$ -SVR method to the machine learners of the system operating regional solar irradiance forecast task. The LibSVM wrapper for R ("e1071" package) is used for the implementation [37,38].

#### 3.2. Index of Error

There are several ways to evaluate prediction results. Mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination ( $R^2$ ) are frequently used for the comparisons of the multiple forecasts' precision:

$$MBE = \frac{1}{n} \sum_{j=1}^n (\hat{y}_j - y_j), \quad (6)$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |\hat{y}_j - y_j|, \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (\hat{y}_j - y_j)^2}, \quad (8)$$

$$R^2 = 1 - \frac{\sum_{j=1}^n (\hat{y}_j - y_j)^2}{\sum_{j=1}^n (\bar{y} - y_j)^2}, \quad (9)$$

$$\bar{y} = \frac{1}{n} \sum_{j=1}^n y_j, \quad (10)$$

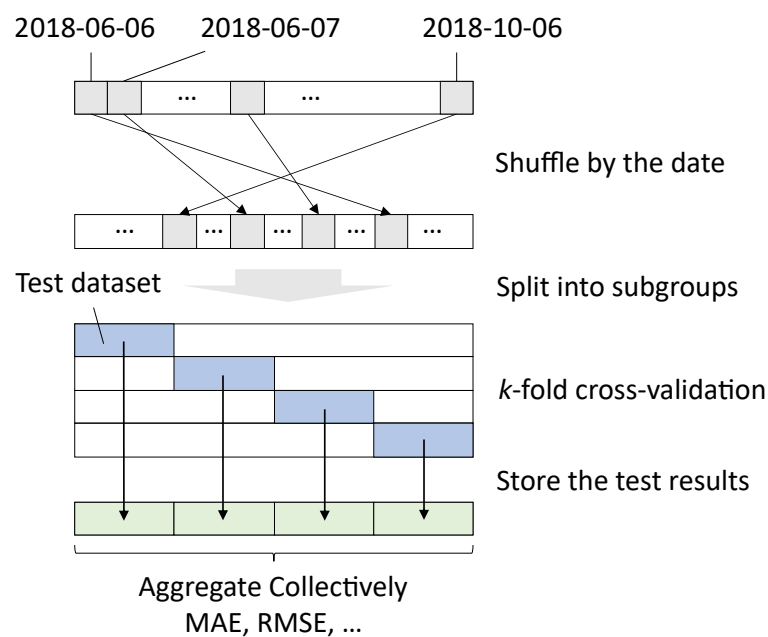
where  $n$  is the number of the samples,  $y_j$  is the real value, and  $\hat{y}_j$  is the prediction value. MAE evaluates the prediction error by the Manhattan distance, while RMSE evaluates the prediction error by the Euclidean distance: In case prediction values are given by the constant, median minimizes MAE, and mean minimizes the RMSE.  $R^2$  gives the

performance index compared with the case that mean is given as the prediction value. The characteristics of these indices are detailed in Reference [39].

### 3.3. Cross-Validation

In addition to measuring the performance score of the prediction result, it is needed to adequately train and predict the machine learning models for the evaluation. Ideally, it is desirable that the training dataset and the test dataset should be strictly divided for verifying the performance of time-series forecast of the machine learning models. However, if enough long dataset is not available,  $k$ -fold cross-validation (CV) is adopted for the verification process: firstly, the dataset is split into  $k$  subgroups; secondly, the machine learner trains the dataset excepting  $k$ -th subgroups' data; thirdly,  $k$ -th data is used for the test of the machine learning model's prediction; those process is repeated until all subgroups' test results are acquired. In particular, if the number of folds  $k$  is equal to the number of samples  $n$ , that is called leave-one-out CV.

In our work, 4-fold CV is adopted for the evaluation of the prediction result, and in order to eliminate the influence of the time trend, random shuffle is performed for each date before the 4-fold CV. Besides, we evaluate the error data obtained from CV collectively;  $k$ -fold CV generally aggregates the score by groups and obtains the mean of their score, but we totally treat the prediction results of each sample without separating groups. The configuration of our evaluation process is shown in Figure 2.



**Figure 2.** Configuration of the evaluation process. Dataset is shuffled by the date before the  $k$ -fold CV, and prediction results are aggregated collectively.

## 4. Prediction Model

Accuracy of the regional solar power forecast based on machine learning models may depend on the strategies for obtaining its prediction value: configuration of the prediction model. In Reference [40], J. Fonseca, Jr., et al. shows the strategy obtaining the regional PV power forecast from local SVR forecasts (Strategy1) gives better performance than the strategy that SVR directly obtains the regional PV power forecast (Strategy2). For this reason, it is important to check the prediction results of the prediction model using MEPS in multiple configurations.

MEPS has 21 forecast members, that are calculated individually for each ensemble member. Hence, the constitution of the predictor could be designed variously: one machine learner could treat all 21 forecast data for the prediction collectively; multiple weak

learners could treat each MEPS member's forecast for make estimation, and their output are integrated at ensemble node (Integrator: INT) for the final decision of the prediction value. In addition, there are two strategies to integrate regional solar radiation at target sites: machine learner directly trains and predicts only the averaged GHI value. Each site's GHI is individually predicted by machine learners, and then they are averaged.

In case adopting the weak learner for generating the ensemble prediction of the GHI for each site, disregarding the combination,  $(N_m + 1) \times N_s$  prediction are obtained, and they have the two axes: index of the MEPS members  $m$ , index of the sites  $s$ . To project them on the  $s$  axis, there are two patterns for the locus of the ensemble nodes. One arranges the ensemble nodes at each site while another arranges the ensemble node at last, that is, integrators output the local GHI estimation or regional GHI estimation. Besides, there are options for how to join in the ensemble node.

In this study, we consider the prediction models based on the SVR machine learners, which are composed as Figures 3–6 and Table 1. Each SVR model (with the exception of the integrator) uses weather forecast data, extra-atmospheric solar radiation on a horizontal surface, and their value of the time of one hour delay. Besides, standardization is applied to the pre-processing of the training. Each case has a different approach to predicting regional GHI. Case1 is an approach to obtain the regional GHI directly without splitting the data, while Case2 is a method to accumulate the predicted GHI for each region. Cases3–4 use the EPS forecast of each location to construct a weak learner for each member, and the integration method of the weak learner is changed in each case. Case3 takes an approach to forecast the regional GHI by passing the results of the weak learner to the integrator and accumulating the local GHI forecasts for each location. In contrast, Case4 takes the approach of calculating regional GHI forecasts for each ensemble member from the results of the weak learner, and finally integrating and forecasting them in the integrator. The mathematical representation of each case and the different labels within the cases are described below.

**Table 1.** Classification of prediction models.

Label	Use of GHI	Output of SVRs	Locus of INT	INT Function
Case1a	MSM Only	Regional GHI	-	-
Case1b	MEPS All Members	Regional GHI	-	-
Case2a	MSM Only	Local GHI	-	-
Case2b	MEPS All Members	Local GHI	-	-
Case3a	MEPS All Members	Local GHI	At Each Site	Mean
Case3b	MEPS All Members	Local GHI	At Each Site	SVR
Case4	MEPS All Members	Local GHI	At Last Layer	SVR

In Case1 shown in Figure 3, only one machine learner is used for the prediction; all sites' data are collectively inputted into it, and regional average GHI is predicted directly.



In Case1a, the predictor uses MSM for GHI forecast input; in Case1b, the predictor uses MEPS 21 member data collectively. Their predictions are given by:

$$\hat{y}_{avg}(h)|_{Case1a} = SVR^0(\mathbf{r}^0(h); \mathbf{u}(h)), \tag{11}$$

$$\hat{y}_{avg}(h)|_{Case1b} = SVR(\mathbf{r}(h); \mathbf{u}(h)), \tag{12}$$

$$\mathbf{r}_s^m(h) = [\widehat{GHI}_s^m(h) \quad \widehat{GHI}_s^m(h-1)], \tag{13}$$

$$\mathbf{r}_s(h) = [\widehat{GHI}_s(h) \quad \widehat{GHI}_s(h-1)], \tag{14}$$

$$\mathbf{r}^m(h) = [\mathbf{r}_1^m(h) \quad \mathbf{r}_2^m(h) \quad \dots \quad \mathbf{r}_{N_s}^m(h)], \tag{15}$$

$$\mathbf{r}(h) = [\mathbf{r}_1(h) \quad \mathbf{r}_2(h) \quad \dots \quad \mathbf{r}_{N_s}(h)], \tag{16}$$

$$\mathbf{u}_s(h) = [MSM_s(h) \quad MSM_s(h-1) \quad E_s(h) \quad E_s(h-1)], \tag{17}$$

$$\mathbf{u}(h) = [\mathbf{u}_1(h) \quad \mathbf{u}_2(h) \quad \dots \quad \mathbf{u}_{N_s}(h)]. \tag{18}$$

In Case2 shown in Figure 4, multiple machine learners predict GHI of each site, and their outputs are summed up to get the regional average GHI forecast. In Case2a, predictors use only the control member data ( $m = 0$ ). In Case2b, each site’s MEPS 21 members’ GHI are put together into the machine learners, which predict each site’s local GHI, respectively, and their outputs are averaged to obtain the regional GHI prediction.

$$\hat{y}_{avg}(h)|_{Case2a} = \frac{1}{N_s} \sum_{s=1}^{N_s} \hat{y}_s^0(h), \tag{19}$$

$$\hat{y}_{avg}(h)|_{Case2b} = \frac{1}{N_s} \sum_{s=1}^{N_s} \hat{y}_s(h), \tag{20}$$

$$\hat{y}_s^m(h) = SVR_s^m(\mathbf{r}_s^m(h); \mathbf{u}_s(h)), \tag{21}$$

$$\hat{y}_s(h) = SVR_s(\mathbf{r}_s(h); \mathbf{u}_s(h)), \tag{22}$$

where  $\hat{y}_s^m(h)$  is the local GHI prediction based on the  $m$ -th member’s GHI forecast data from MEPS.

In Case3 shown in Figure 5, the prediction model consists of weak learners and integrators in each subsystem; weak learners try to improve GHI prediction from each MEPS member’s GHI forecast individually; integrators are, respectively, output each site’s GHI prediction, and regional average GHI is obtained from them:

$$\hat{y}_{avg}(h)|_{Case3} = \frac{1}{N_s} \sum_{s=1}^{N_s} INT_s(\hat{y}_s^0(h), \hat{y}_s^1(h), \dots, \hat{y}_s^{N_m}(h)), \tag{23}$$

where  $INT_s(\cdot)$  is the function corresponding to the integrator of each site. They predict the local GHI for each site. They use only the weak learners’ output if SVR models are used in the ensemble nodes.

In Case4 shown in Figure 6, this model consists of weak learners and an integrator as with the Case3; weak learners predict each site’s GHI from each MEPS member data, and their outputs are summed up for estimating regional average GHI, respectively, for each member’s group; the integrator finally produces the regional average GHI prediction:

$$\hat{y}_{avg}(h)|_{Case4} = INT(\hat{y}^0(h), \hat{y}^1(h), \dots, \hat{y}^{N_m}(h)), \tag{24}$$

$$\hat{y}^m(h) = \frac{1}{N_s} \sum_{s=1}^{N_s} \hat{y}_s^m(h), \tag{25}$$

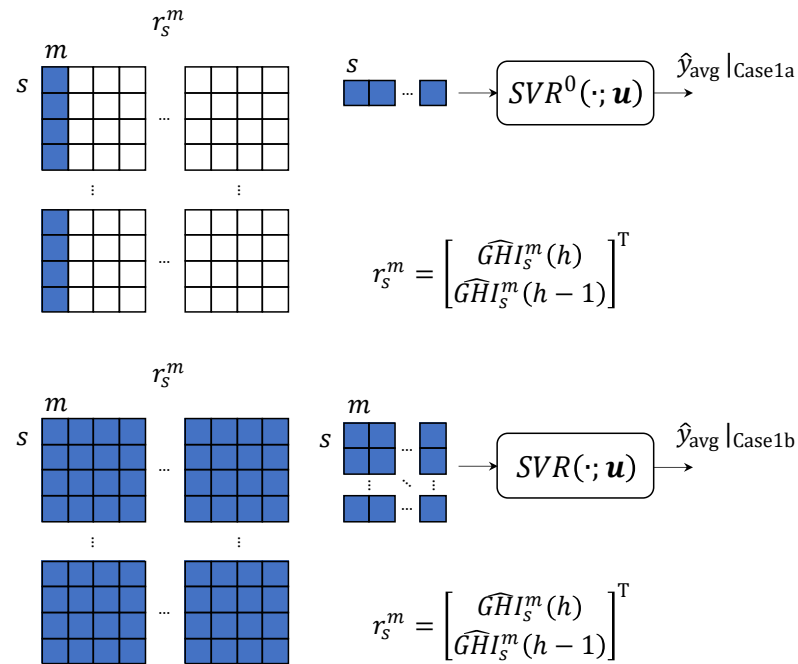
where  $INT(\bullet)$  is the integrator placed in the last layer of the prediction model. It uses each MEPS member’s regional GHI prediction values as the explanatory variables.

In Case3, we adopt mean operation and SVR as integration process. However, in Case4, we only use SVR for the function of the integrator because the mean operation's result is equal to the result of Case3's mean operation case.

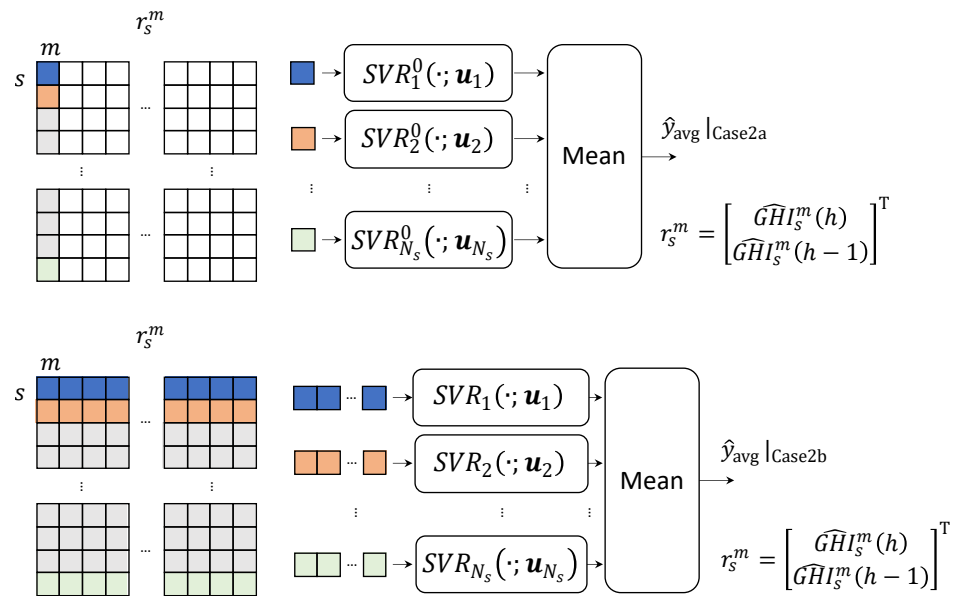
For the comparison with the persistence model, which is often treated as the benchmark, we give the prediction value denoted by:

$$\hat{y}_{\text{avg}}(h)|_{\text{Persistence}} = \frac{1}{N_s} \sum_{s=1}^{N_s} y_s(h - 24). \tag{26}$$

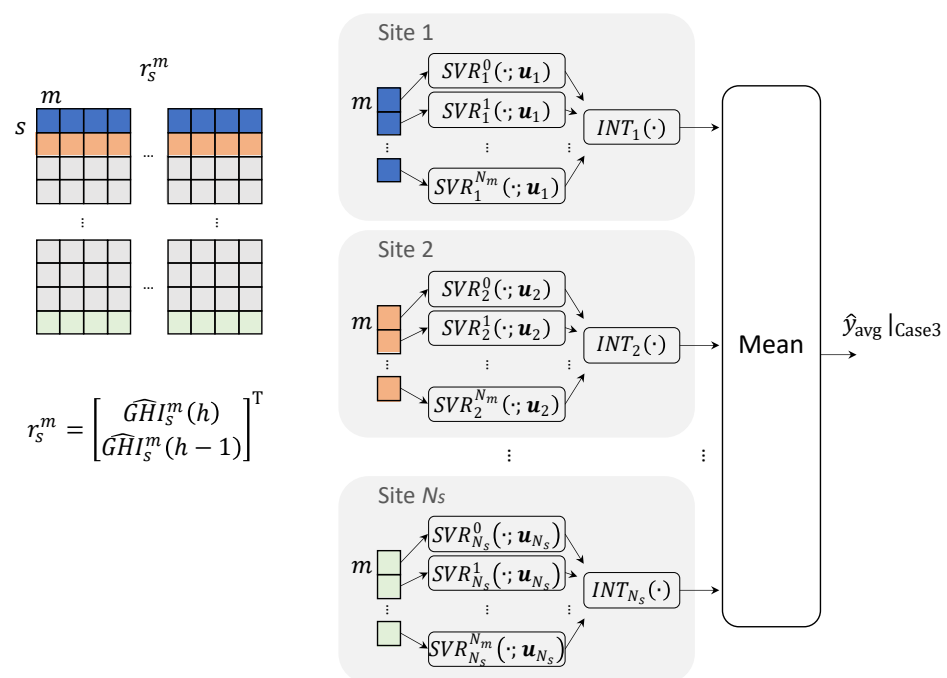
The persistence model simply outputs the one day-ahead observed GHI as the forecast.



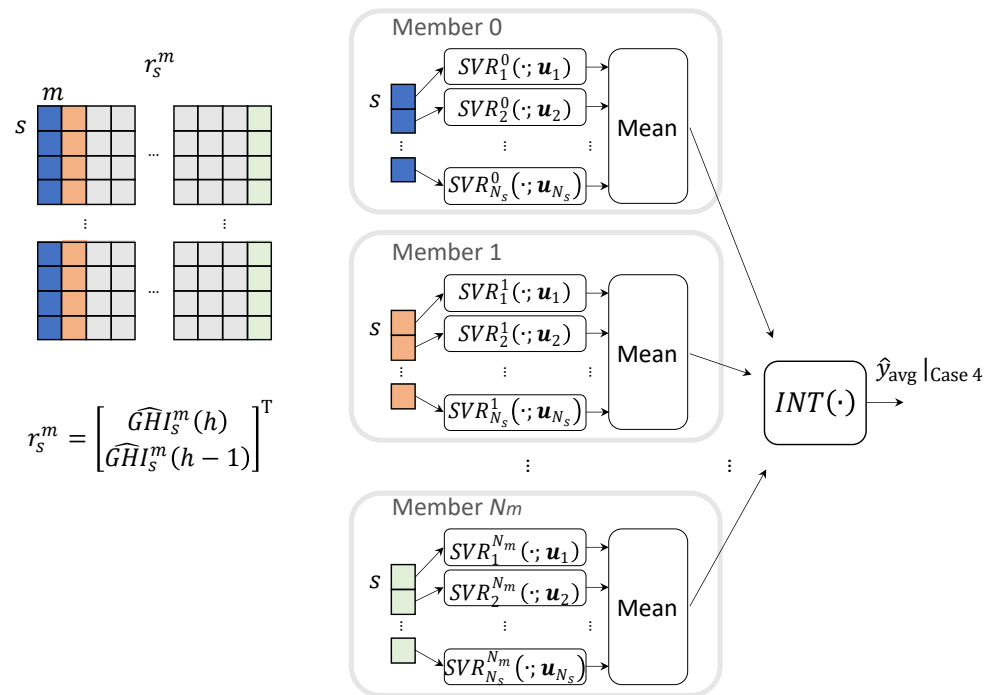
**Figure 3.** Schematic of the access of the MEPS's GHI forecasts in Case1. Case1a uses the MSM's GHI ( $m = 0$ ) only, and Case1b uses MEPS all member ( $m = 0, 1, \dots, 20$ ). They directly predict regional GHI with one SVR model.



**Figure 4.** Schematic of the access of the MEPS's GHI forecasts in Case2. Case2a-b predict local GHI at each site, and their predictions are averaged to obtain the regional GHI prediction. Case2a uses MSM's GHI only, and Case2b accesses the all MEPS member's GHI.



**Figure 5.** Schematic of the access of the MEPS's GHI forecasts in Case3. This prediction model places the integrators (INT<sub>s</sub>(•), s = 1, 2, ..., 5) at each site. Weak learners (SVR<sub>s</sub><sup>m</sup>(•, u<sub>s</sub>)) output the prediction, and then they are gathered into the integrators at ensemble nodes. Each ensemble node forecasts the local GHI. Then, regional GHI is obtained from the mean function. Integrators use mean function or SVR in Case3a and Case3b.



**Figure 6.** Schematic of the access of the MEPS’s GHI forecasts in Case4. This prediction model places the integrator ( $INT(\bullet)$ ) at the last layer. Weak learners ( $SVR_s^m(\bullet, \mathbf{u}_s)$ ) output are gathered into the mean function for each MEPS ensemble member to obtain the regional GHI ensemble prediction, and then their forecasts enter the ensemble node.

**5. Results and Discussion**

Evaluation scores for Case1–4, Persistence Model, MSM’s forecast (MSM-GPV), and MEPS ensemble mean are in Table 2 and Figure 7. In comparison with the Persistence, MSM and MEPS forecasts, all Cases refine the accuracy of the regional GHI forecast for the application of the machine learning model. In particular, Case2b presents the best performance, and the prediction values including MEPS have higher performance than the Case1a, MSM-GPV, and MEPS ensemble mean. The results of Case1b and Case2b show that the forecast model using MEPS forecast data scores better than the other model within the same approach (Case1a, Case2a).

**Table 2.** Evaluation score of the prediction models. MBE, MAE, and RMSE are in  $W \cdot m^{-2}$ . R2 is in percentage.

Prediction Model	MBE	MAE	RMSE	R2
Persistence	−1.1	128.0	192.8	47.8
MSM	−37.6	72.1	110.7	80.1
MEPS	−37.6	69.6	105.0	80.7
Case1a	−1.0	70.0	100.3	84.4
Case1b	−2.2	63.3	91.4	86.9
Case2a	4.3	62.3	90.6	87.4
Case2b	1.2	60.2	87.2	88.4
Case3a	4.5	61.9	89.6	87.6
Case3b	5.4	61.5	89.5	88.0
Case4	2.1	62.4	91.2	87.3

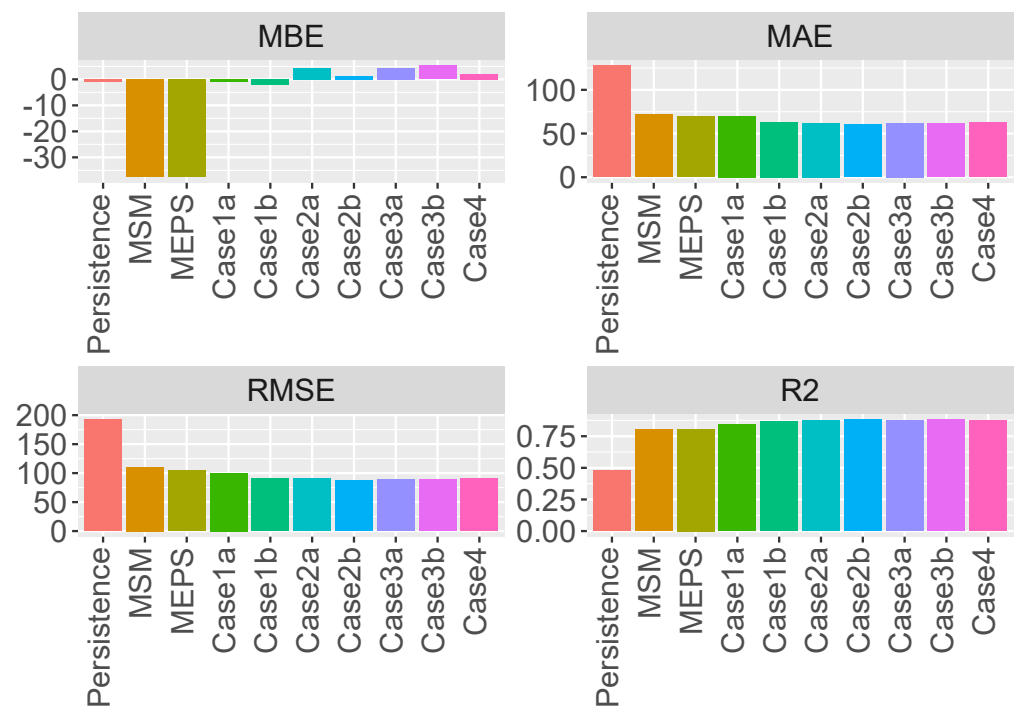


Figure 7. Index of the precision of the forecasts.

In addition, it is confirmed that the strategy averaging multiple local forecasts gives better performance by comparison Case1 with Case2. Among the prediction models using MEPS forecasts, the approach of Case3 is the best after Case2b. Case3 is also based on the accumulation of regional GHI forecasts. As with regional PV forecast in Reference [40], regional GHI forecast that adopts MEPS ensemble members shows the similar trends. The systematic error of the NWP is not uniform among points, and it is thought that the systematic error caused by the topography of each point can be corrected by machine learning models. Therefore, it is inferred that it is effective to divide the model by sub-region in the forecasting model using EPS forecasts.

However, in contrast to the improvement of the precision by the strategy of the local GHI forecast integration, the score of Case3–4, which renovated the MEPS ensemble members through weak learners, gives worse performance than Case2b, while MEPS ensemble members improve the forecast in Case1–2 and MSM without weak learners. Score of the Case3–4 are almost equal to the score of the Case2a. Each weak learner in Case3–4 uses each MEPS member's GHI forecast individually, and their configuration is equal to the predictor in Case2a. By constructing multiple SVR by changing the choice of the MEPS member simply, multi-model ensemble approach does not contribute to the accuracy improvement. The ensemble mean approach assumes that the ensemble members contain the appropriate perturbation. By applying MEPS GHI forecast to the SVR directly, weak learners output loses the effectiveness, that is, the spread of probability distribution is lost. To refine the precision of the models in Case3–4, it is needed to maintain the ensemble members' diversity through the post-process in Reference [26], such as EMOS and variance deficit (VD). Hence, in case no post-processing is used, it is better to input the EPS forecasts together in one model to construct an effective predictor.

Focusing on the MBE, MSM, and MEPS has the negative bias (about  $-37.5$  [ $W/m^2$ ]), while the other predictors configured by the machine learning technique correct the MBE and median properly (from  $-2.2$  [ $W/m^2$ ] to  $5.4$  [ $W/m^2$ ]). Through the training process, certainly, the predictors compensate the effect of the micro-scale physic unexpressed correctly in the MSM and MEPS.

MSM, which is the base model of the MEPS, has the tendency to underestimate the regional GHI. In previous research, Ohtake et al. (2015) analyzed the seasonal variations

in the GHI forecast errors of the MSM-GPV during the year from 2008 to 2012, and MSM had the tendency to underestimate GHI from  $-29.5$  to  $50.6$  in summer (from June to November), regardless of the initialization times [41]. In 2008, JMA also reported MSM underestimated GHI in summer during 2004 to 2007 [42]. Over only the Kanto region, Ohtake et al. (2013) investigated the relationship between the relatively large GHI forecast error and the frequency of the cloud type observed at JMA station, and they conjectured cirrus appeared in the underestimation case [43]. Referring to the Köppen-Geiger climate classification [44], the Kanto region of Japan is classified as a humid subtropical climate (Cfa). In East Asia, the monsoon tends to carry humid air from low latitudes to high latitudes during the summer, resulting in high humidity in the Kanto region. In relation to upper clouds (e.g., cirrus) produced by humid air, the MSM does not consider ice supersaturation in the calculation of cloud cover diagnostics in radiative processes. This suggests that overestimation of upper clouds is one of the causes of underestimation in summer. MEPS ensemble members are generated from the control run which corresponds to the MSM. Hence, it can be inferred the ensemble members of the MEPS similarly has the tendency to underestimate the GHI, while the control-run of the members replaced the model of dynamical process from JMA-NHM to asuca. Because the machine learning model includes relative humidity and temperature as explanatory variables, the effect of overestimation of upper cloud cover is considered to be corrected by learning in Case1–4 regardless of the differences in networks.

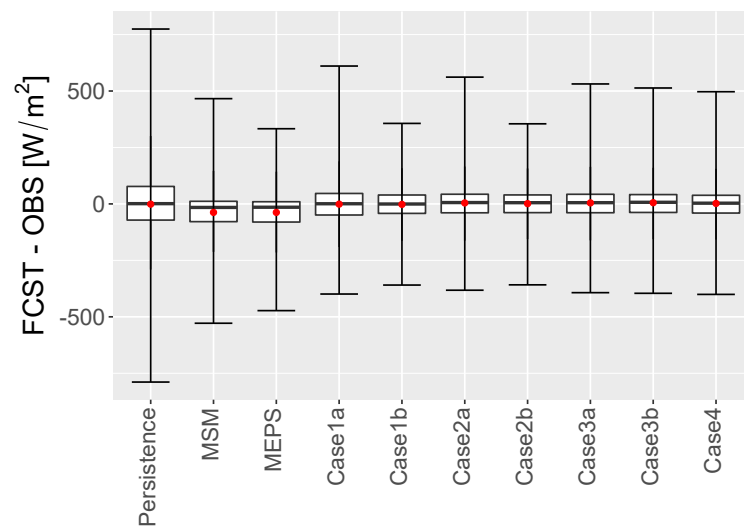
From the above, the following can be inferred about the average accuracy of each predictor in this study.

- If the predictors have the same network structure, the average accuracy of the predictors can be improved by adding EPS predictions to the explanatory variables.
- An approach that divides the predictor by sub-region and combines local GHI is effective for predicting regional GHI.
- When post-processing is not used for weak learners, it is more effective to use a single predictor to handle EPS forecasts rather than to construct a multistage predictor for each ensemble member.
- Regardless of the network structure, each predictor compensates for the systematic errors of the EPS.

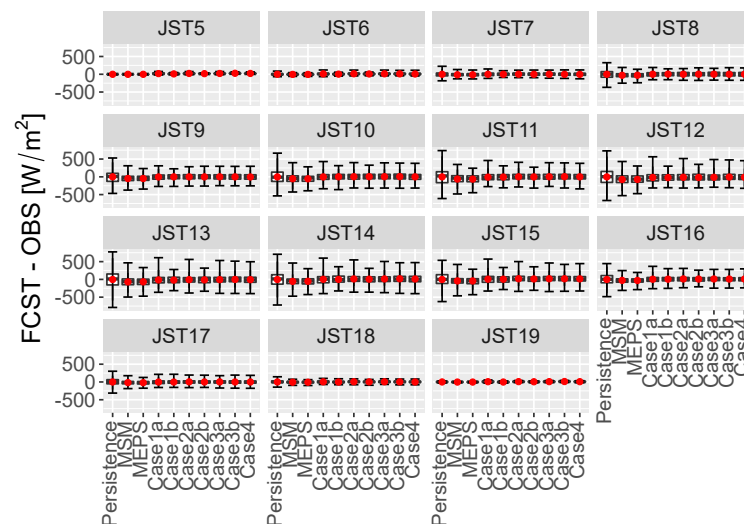
In addition to the average accuracy of the forecast, confirmation of the maximum forecast error and error distribution is particularly important from the practical standpoint of day-ahead solar power forecasting. To confirm the distribution of the forecast errors ( $\hat{y}_k - y_k$ ), quantile of the error is in Table 3, and their box plot are shown in Figures 8 and 9. Focusing on the maximum error (100% quantile), persistence model, Case1a, Case2a, and Case3–4 give larger maximum error than the MSM, that is, their prediction models hardly decrease the amount of the regulating power that TSO purvey on the day before the operation. MEPS without machine learning models, Case1b and Case2b, however, show smaller maximum error (100% quantile) in whole period, and they suppress the overestimation of the regional GHI on JST10:00–JST15:00. This means that it is possible to reduce the overestimation of PV power generation based on the GHI forecast during the sunny hours when the power demand is high in the summer. Notably, when compared with the predictors only using MSM, the maximum error decreases significantly by the utilization of MEPS: Case1b reduces 100% quantile error by 41.6% from the result of Case1a; Case2b reduces 100% quantile error by 36.8% from the result of Case2a. Case1b and Case2b coordinate MEPS ensemble member forecasts collectively; thus, machine learning models assumably obtain the performance in the various situation by referring to the MEPS ensemble members' multiple behaviors. From above, by collectively introducing EPS forecasts into the machine learning model's explanatory variables, the prediction model can improve the maximum forecast error, and the bias in the NWP models can also be corrected.

**Table 3.** Quantile of the forecast error. Quantile values are in  $W \cdot m^{-2}$ .

Prediction Model	Q0%	Q25%	Q50%	Q75%	Q100%
Persistence	-788.89	-71.67	1.11	77.36	774.44
MSM	-528.42	-78.53	-15.61	11.55	466.15
MEPS	-472.56	-80.26	-14.69	9.64	332.98
Case1a	-399.11	-49.13	0.75	46.27	610.60
Case1b	-359.35	-41.94	-0.37	39.44	356.49
Case2a	-382.21	-39.38	6.02	42.85	561.54
Case2b	-358.36	-38.68	5.43	39.77	354.77
Case3a	-393.03	-39.00	5.48	42.62	531.29
Case3b	-395.88	-37.81	7.23	41.29	513.43
Case4	-400.72	-40.52	3.25	38.21	496.86



**Figure 8.** Box plot of the forecast error in total. Note that red point and black line indicate MBE and median of the error.



**Figure 9.** Box plot of the forecast error for each time cross section. Prediction results in the whole period are aggregated by time.

## 6. Conclusions

In this study, we applied several configurations of the machine learning model to the MEPS data, which is a regional EPS data of Japan, and evaluated the prediction errors in order to verify the improvement effect of the GHI forecast and the effective configuration method of the predictor using EPS. The main contributions of our paper are the following:

- We investigated how to construct a regional GHI predictor using regional EPS and how to give an effective network configuration.
- It was confirmed that systematic errors in MEPS forecasts can be improved by machine learning.
- By applying machine learning methods to the prediction of MEPS, we showed that it is possible to construct a predictor that reduces the maximum prediction error while correcting for systematic errors in MEPS.

Multiple configurations were considered for design the predictors, and machine learning-based predictors utilizing MEPS ensemble members forecasts improve the MBE, MAE, RMSE,  $R^2$  in almost all cases when compared with the predictor using only MSM. Furthermore, maximum error of the machine learning model decreased in case MEPS forecasts are collectively input into the predictor; in other words, machine learning-based predictors' overestimation can be suppressed by the MEPS data utilization. However, duration of our available MEPS data is limited, and they can have the seasonal deviation (e.g., Ohtake et al., 2013, 2015). For the more accurate verification, it is needed to confirm their effect in longer term: in spring, in autumn, in winter, in annual. In addition, while the comparison of the prediction results has provided us with knowledge on how to construct an effective network for machine learning models using EPS, it is necessary to conduct future comparisons of predictors with multi-stage configurations using weak learners with appropriate post-processing for predicting weak learners. Besides, machine learners used only GHI from the MEPS forecasts as the first analysis. Hence, it is desirable that various parameters of MEPS forecasts are utilized in the machine learning models. In order to investigate even deeper, we will extend the period of the available data and utilize the multiple parameters of the MEPS ensemble members' forecasts.

**Author Contributions:** Conceptualization, T.O.; methodology, T.T.; software, T.T.; validation, T.T.; formal analysis, T.T.; investigation, T.T. and H.O.; resources, T.N., Y.H. and M.K.; data curation, T.T. and H.O.; writing—original draft preparation, T.T., H.O., T.N.; writing—review and editing, T.T.; visualization, T.T.; supervision, H.O. and T.O.; project administration, T.O.; funding acquisition, T.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

<i>MBE</i>	mean bias error
<i>MAE</i>	mean absolute error
<i>RMSE</i>	root mean square error
$R^2$	coefficient of determination
<i>SVR</i>	function constructed by support vector regression
<i>INT</i>	function that means integrator
<i>GHI</i>	global horizontal irradiation
<i>T</i>	temperature
<i>RH</i>	relative humidity
<i>HC</i>	high level cloud cover



MC	middle level cloud cover
LC	low level cloud cover
$N_s$	number of sites
$N_m$	number of ensemble members
$s$	index of site number
$m$	index of ensemble member number

## References

- Rashid, K.; Safdarnejad, S.M.; Powell, K.M. Process intensification of solar thermal power using hybridization, flexible heat integration, and real-time optimization. *Chem. Eng. Process. Process. Intensif.* **2019**, *139*, 155–171. [CrossRef]
- Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N.. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew. Sustain. Energy Rev.* **2013**, *27*, 65–76. [CrossRef]
- Marquez, R.; Coimbra, C.F. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. *Sol. Energy* **2011**, *85*, 746–756. [CrossRef]
- Voyant, C.; Muselli, M.; Paoli, C.; Nivet, M.L. Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. *Energy* **2012**, *39*, 341–355. [CrossRef]
- Fonseca, J.G.D.S., Jr.; Takashi, O.; Hideaki, O.; Ken-ichi, S.; Takumi, T.; Kazuhiko, O. Analysis of different techniques to set support vector regression to forecast insolation in Tsukuba, Japan. *J. Int. Counc. Electr. Eng.* **2013**, *3*, 121–128. [CrossRef]
- Lorenz, E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
- Chaouachi, A.; Kamel, R.M.; Ichikawa, R.; Hayashi, H.; Nagasaka, K. Neural network ensemble-based solar power generation short-term forecasting. *World Acad. Sci. Eng. Technol.* **2009**, *54*, 54–59. [CrossRef]
- Liu, Y.; Shimada, S.; Yoshino, J.; Kobayashi, T.; Miwa, Y.; Furuta, K. Ensemble forecasting of solar irradiance by applying a mesoscale meteorological model. *Sol. Energy* **2016**, *136*, 597–605. [CrossRef]
- Alessandrini, S.; Delle Monache, L.; Sperati, S.; Cervone, G. An analog ensemble for short-term probabilistic solar power forecast. *Appl. Energy* **2015**, *157*, 95–110. [CrossRef]
- Yamazaki, T.; Homma, H.; Wakao, S.; Fujimoto, Y.; Hayashi, Y. Estimation prediction interval of solar irradiance based on just-in-time modeling for photovoltaic output prediction. *Electr. Eng. Jpn.* **2016**, *195*, 1–10. [CrossRef]
- Koenker, R.; Bassett, G., Jr. Regression quantiles. *Econom. J. Econom. Soc.* **1975**, *46*, 33–50.
- Almeida, M.P.; Perpiñán, O.; Narvarte, L. PV power forecast using a nonparametric PV model. *Sol. Energy* **2015**, *115*, 354–368. [CrossRef]
- Juban, R.; Ohlsson, H.; Maasoumy, M.; Poirier, L.; Kolter, J.Z. A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014. *Int. J. Forecast.* **2016**, *32*, 1094–1102. [CrossRef]
- He, Y.; Yan, Y.; Xu, Q. Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression. *Int. J. Electr. Power Energy Syst.* **2019**, *113*, 515–527. [CrossRef]
- Fonseca, J.G.D.S., Jr.; Oozeki, T.; Ohtake, H.; Takashima, T.; Kazuhiko, O. On the use of maximum likelihood and input data similarity to obtain prediction intervals for forecasts of photovoltaic power generation. *J. Electr. Eng. Technol.* **2015**, *10*, 1342–1348. [CrossRef]
- Epstein, E.S. Stochastic dynamic prediction. *Tellus* **1969**, *21*, 739–759. [CrossRef]
- Leith, C.E. Theoretical skill of Monte Carlo forecasts. *Mon. Weather. Rev.* **1974**, *102*, 409–418. [CrossRef]
- Uno, F.; Ohtake, H.; Matsueda, M.; Yamada, Y. A diagnostic for advance detection of forecast busts of regional surface solar radiation using multi-center grand ensemble forecasts. *Sol. Energy* **2018**, *162*, 196–204. [CrossRef]
- Toth, Z.; Kalnay, E. Ensemble forecasting at NMC: The generation of perturbations. *Bull. Am. Meteorol. Soc.* **1993**, *74*, 2317–2330. [CrossRef]
- Hou, D.; Kalnay, E.; Droegemeier, K.K. Objective verification of the SAMEX'98 ensemble forecasts. *Mon. Weather. Rev.* **2001**, *129*, 73–91. [CrossRef]
- Du, J.; Tracton, M.S. Implementation of a Real-Time Short Range Ensemble Forecasting System at NCEP: An update. In Proceedings of the 9th Conference on Mesoscale Processes, Fort Lauderdale, FL, USA, 29 July–2 August 2011.
- Bowler, N.E.; Arribas, A.; Mylne, K.; Robertson, K.B.R.; Beare, S.E. The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc. J. Atmos. Sci. Appl. Meteorol. Phys. Oceanogr.* **2008**, *134*, 703–722. [CrossRef]
- Erfani, A.; Frenette R.; Gagnon, N.; Charron, M.; Beaugaurd, S.; Giguère, A.; Parent, A. The New Regional Ensemble Prediction System at 15 km Horizontal Grid Spacing (REPS 2.0.1), Canadian Meteorological Centre Technical Note. Available online: [https://collaboration.cmc.ec.gc.ca/cmc/cmci/product\\_guide/docs/lib/technote\\_reps201\\_20131204\\_e.pdf](https://collaboration.cmc.ec.gc.ca/cmc/cmci/product_guide/docs/lib/technote_reps201_20131204_e.pdf) (accessed on 13 April 2021).
- Gebhardt, C.; Theis, S.E.; Paulat, M.; Bouallègue, Z.B. Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmos. Res.* **2011**, *100*, 168–177. [CrossRef]
- Junk, C.; Delle Monache, L.; Alessandrini, S. Analog-based ensemble model output statistics. *Mon. Weather. Rev.* **2015**, *143*, 2909–2917. [CrossRef]
- Sperati, S.; Alessandrini, S.; Delle Monache, L. An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting. *Sol. Energy* **2016**, *133*, 437–450. [CrossRef]

27. Rasp, S.; Lerch, S. Neural networks for postprocessing ensemble weather forecasts. *Mon. Weather. Rev.* **2018**, *146*, 3885–3900. [[CrossRef](#)]
28. Massidda, L.; Marrocu, M. Quantile regression post-processing of weather forecast for short-term solar power probabilistic forecasting. *Energies* **2018**, *11*, 1763. [[CrossRef](#)]
29. Ono, K.; Kunii, M.; Honda, Y. The regional model-based Mesoscale Ensemble Prediction System, MEPS, at the Japan Meteorological Agency. *Q. J. R. Meteorol. Soc.* **2021**, *147*, 465–484. [[CrossRef](#)]
30. Japan Meteorological Agency | Numerical Weather Prediction Activities. Available online: <https://www.jma.go.jp/jma/en/Activities/nwp.html> (accessed on 14 April 2021).
31. Liou, K.N. *An Introduction to Atmospheric Radiation*, 2nd ed.; Academic Press: London, UK, 2002.
32. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
33. Müller, K.R.; Smola, A.J.; Rätsch, G.; Schölkopf, B.; Kohlmorgen, J.; Vapnik, V. Predicting time series with support vector machines. In Proceedings of the 7th International Conference, Lausanne, Switzerland, 8–10 October 1997.
34. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **1997**, *9*, 155–161.
35. Schölkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. New support vector algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [[CrossRef](#)] [[PubMed](#)]
36. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
37. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *Acm Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 1–27. [[CrossRef](#)]
38. Meyer D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Friedrich Leisch. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, R Package Version 1.7-4. 2020. Available online: <https://CRAN.R-project.org/package=e1071> (accessed on 14 April 2021).
39. Van der Meer, D.W.; Widén, J.; Munkhammar, J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1484–1512. [[CrossRef](#)]
40. Fonseca, J.G.D.S., Jr.; Oozeki, T.; Ohtake, H.; Shimose, K.I.; Takashima, T.; Ogimoto, K. Forecasting regional photovoltaic power generation—a comparison of strategies to obtain one-day-ahead data. *Energy Procedia* **2014**, *57*, 1337–1345.
41. Ohtake, H.; Fonseca, J.G.D.S., Jr.; Takashima, T.; Oozeki, T.; Shimose, K.I.; Yamada, Y. Regional and seasonal characteristics of global horizontal irradiance forecasts obtained from the Japan Meteorological Agency mesoscale model. *Sol. Energy* **2015**, *116*, 83–99. [[CrossRef](#)]
42. Nagasawa, R. Report of Numerical Prediction. *Radiat. Process. Div.* **2008**, *54*, 149–165. (In Japanese, Suuchiyohouka Bessatsu Houkoku)
43. Ohtake, H.; Shimose, K.I.; da Silva Fonseca, J.G., Jr.; Takashima, T.; Oozeki, T.; Yamada, Y. Accuracy of the solar irradiance forecasts of the Japan Meteorological Agency mesoscale model for the Kanto region, Japan. *Sol. Energy* **2013**, *98*, 138–152. [[CrossRef](#)]
44. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [[CrossRef](#)]