Progress in PHOTOVOLTAICS

PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS
Prog. Photovolt: Res. Appl. 2012; **20**:874–882
Published online 22 July 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/pip.1152

RESEARCH ARTICLE

Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in Kitakyushu, Japan

Joao Gari da Silva Fonseca Jr.[1]*, Takashi Oozeki[1], Takumi Takashima[1], Gentarou Koshimizu[2], Yoshihisa Uchida[2] and Kazuhiko Ogimoto[3]

[1] National Institute of Advanced Industrial Science and Technology, Research Center for Photovoltaics Technologies, AIST Tsukuba Central 2, 305-8568 Tsukuba, Ibaraki, Japan
[2] Electric Power Development Co., Ltd. (JPOWER), 104-8165 Ginza 6-15-1, Chuo-ku, Tokyo, Japan
[3] Institute of Industrial Science (IIS), Collaborative Research Center for Energy Engineering (CEE), Tokyo University, 153-8505 Komaba 4-6-1, Meguro-ku, Tokyo, Japan

ABSTRACT

The development of a methodology to forecast accurately the power produced by photovoltaic systems can be an important tool for the dissemination and integration of such systems on the public electricity grids. Thus, the objective of this study was to forecast the power production of a 1-MW photovoltaic power plant in Kitakyushu, Japan, using a new methodology based on support vector machines and on the use of several numerically predicted weather variables, including cloudiness. Hourly forecasts of the power produced for 1 year were carried out. Moreover, the effect of the use of numerically predicted cloudiness on the quality of the forecasts was also investigated. The forecasts of power production obtained with the proposed methodology had a root mean square error of 0.0948 MW h and a mean absolute error of 0.058 MW h. It was also found that the forecast and measured values of power production had a good level of correlation varying from 0.8 to 0.88 according to the season of the year. Finally, the use of numerically predicted cloudiness had an important role in the accuracy of the forecasts, and when cloudiness was not used, the root mean square error of the forecasts increased more than 32%, and the mean absolute error increased more than 42%. Copyright © 2011 John Wiley & Sons, Ltd.

KEYWORDS
photovoltaic systems; power production forecast; support vector regression; numerically predicted cloudiness

***Correspondence**
Joao Gari da Silva Fonseca Jr., National Institute of Advanced Industrial Science and Technology, Research Center for Photovoltaics, AIST Tsukuba Central 2, 305-8568, Tsukuba, Ibaraki, Japan.
E-mail: jon-fonseca@aist.go.jp

Received 5 January 2011; Revised 5 April 2011

1. INTRODUCTION

One of the main problems regarding the use of photovoltaic (PV) systems in large scale is their characteristic intermittent power production. Such unstable power production can cause control and operation issues for users and administrators of public electricity grids, as they would have to deal frequently with sudden surpluses or drops in the power production. One way to address this problem is to forecast the amount of power production of PV systems. In this way, eventual drops or surpluses in the power production can be predicted and compensated, for example, balancing the use of PV with the other kinds of energy systems.

The power production of PV systems can be directly or indirectly forecasted. When indirectly forecasted, a variable related to PV systems' power production, such as the insolation, is first forecasted and then, using appropriate models, the power production itself is calculated. Insolation forecast has application in several fields besides solar energy systems. For example, it can be used for crop modeling [1], building design [2], meteorological studies, and other applications. Furthermore, several approaches have been employed in the problem of insolation forecasts [3,4]. Regardless of the approach and application, state-of-the-art insolation forecast methods developed until now are still unable to provide high accuracy insolation forecasts for a large range of weather patterns. As a consequence, the forecast of the power production of PV systems based on insolation forecast is directly affected, as it incorporates the errors of the latter.

The main difference between direct and indirect methods to forecast power production of PV systems is that the former does not require insolation forecasts as an intermediate step in the calculations. The elimination of this step can produce results with better accuracy, as the models used to convert insolation to power output in a PV system are not used. Direct methods of forecasting the power production of PV systems can be based on empirical equations or on computer techniques. If computer techniques as artificial intelligence models are used, the input variables used to forecast insolation can also be applied to directly forecast the power production of PV systems. This is explained by the strong dependence of the power production of a PV system on insolation.

Artificial intelligence models have been used in PV applications [5]. Typical applications involve the use of feed-forward or radial basis functions artificial neural networks [6,7]. Moreover, a combined use of artificial neural networks with some other techniques, such as wavelet theory and fuzzy logic, has also been investigated [8]. The main advantages of artificial intelligence models are related with their ability to handle noisy or incomplete data, to generalize and to adapt to different problem configurations, and to solve non-linear problems. Furthermore, these methods can be used with parallel processing enabling high speed calculations. In spite of these advantages, specifically on the problem of forecasts of PV systems power production, further improvements of methods and techniques are still required to achieve high accuracy forecasts, for instance, on an hourly basis and for a year of weather patterns.

One of the most advanced artificial intelligence techniques currently available is known as support vector machines (SVMs). According to Cristianini and Shawe-Taylor [9], SVMs are based on recent advances in statistical learning theory and can deliver state-of-the-art performance in categorization and classification problems. Furthermore, SVM can also be used for regression and prediction problems [10,11]. In this way, it should be investigated if SVM can also deliver improved performance in the problem of the forecast of PV system power production.

Another way to improve the accuracy of the forecasts of PV systems' power production is to improve the quality of the information available that can be used as input data of the forecast methods. Several studies available in the technical literature apply insolation forecast models with previously measured meteorological data, such as temperature, geographical data, and with auxiliary variables as the clearness index [12,13].With the constant development of better weather forecast models, forecasted meteorological data have their accuracy improved. The use of such data can help forecast models to provide better results. In Japan, one of the forecasted meteorological data that has its accuracy improved since 2007 is the numerically predicted cloudiness, provided by the grid point value–meso-scale model (GPV–MSM) weather forecast system of the Japan Meteorology Agency. Cloudiness is an important variable that affects the power production of PV systems. Moreover, it was already proven that insolation forecast methods using predicted cloudiness and other input variables can improve the accuracy of the forecasts provided by models, based on empirical equations or based on artificial neural networks [14,15]. In the same way, better input data used with a state-of-the-art forecast method may yield forecasts of power production of PV systems with improved accuracy. Therefore, it should be verified whether such benefits can be obtained using SVM and numerically predicted cloudiness.

The objective of this study is to directly forecast the power production of a PV power plant in Japan using SVMs and numerically predicted cloudiness. Moreover, the effect of the use of numerically predicted cloudiness on the forecasts carried out with SVMs was also studied, comparing forecasts using and not using such data.

The power production data used in this study belong to a PV power plant located in the city of Kitakyushu. The PV power plant has a maximum output of 1 MW h. The forecasts of power production were hourly, and they were carried out for a period of 1 year.

2. FORECAST METHODS

To verify whether the use of SVM is an interesting method to forecast power production of PV power plants, we compared SVM's forecasts with the forecasts carried out using a persistence method. In this study, persistence at a given hour was regarded as the measured power production of the PV power plant at the same hour of the day prior to the forecast day.

The SVM used was the ν-support vector regression (ν-SVR). The implementation of the ν-SVR is available in the LIBSVM software library [16,17]. The port of this software library for MATLAB was used to make the forecasts of power production of the PV power plant.

2.1. Support vector machines

Support vector machines are a class of learning algorithms originally developed for pattern recognition. Later, the method was extended to deal with regression problems with the devising of an ε-insensitive loss function by Vapnik [18], where ε is a prescribed parameter indicating the deviation a forecast can have from the target answer. This method was further developed by Schölkopf *et al.* [19], who added modifications that automatically minimize ε, yielding estimates as accurate as possible. The modified method is called ν-SVR, and it was the one used in this study to make the forecasts of power production of the PV power plant.

The SVR finds a function that can relate a set of input patterns $(x_1,\ldots, x_n)$ with a set of targets $(y_1,\ldots, y_n)$, mapping the input patterns to a higher dimension space, with a map function ϕ, and performing a linear fit with maximum deviation of ε from the actual targets. If the input patterns x_i are mapped to a higher dimension space with a map function ϕ, their relation with the targets y_i can be linearly described with a weight vector w_i and a bias b as shown in Equation (1).

$$y_i \approx \sum_{j=1}^{m} w_j \phi\left(x_{ij}\right) + b \tag{1}$$

The problem can be rewritten using a Lagrange function and upon doing implicit mapping. In this way, the weight vector can be expanded using the Lagrange multipliers, α_i and α_i^*, as shown in Equation (2).

$$w = \sum_{j=1}^{n} \left(\alpha_i - \alpha_i^*\right) \phi(x_t) \tag{2}$$

Substituting Equation (2) with Equation (1) and replacing the resulting dot product with a kernel function belonging to the ϕ mapping, Equation (1) can be rewritten as

$$y_i \approx \sum_{j=1}^{m} \left(\alpha_i - \alpha_i^*\right) k\left(x_i, x_j\right) + b \tag{3}$$

To solve Equation (3), we have to obtain the Lagrange multipliers α_i and α_i^* and to choose a kernel function. Therefore, the v-SVR problem is for $C > 0$ and $v \geq 0$ to maximize the following quadratic problem:

$$-\frac{1}{2} \sum_{i,j=1}^{n} \left(\alpha_i - \alpha_i^*\right) k\left(x_i, x_j\right) \left(\alpha_j - \alpha_j^*\right) - \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) y_i \tag{4}$$

subject to

$$\sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) = 0 \tag{5}$$

$$0 \leq \alpha_i^* \leq \frac{C}{n} \tag{6}$$

$$\sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) \leq C \cdot v \tag{7}$$

The kernel chosen to be used with v-SVR in this study was the radial basis function described in Equation (8). This kernel was chosen because it yielded the best results in the tests carried out using several kinds of kernels.

$$k\left(x_i, x_j\right) = e^{-\gamma |x_i - x_j^2|} \tag{8}$$

To use the v-SVR, we have to choose beforehand the configuration parameters such as γ in Equation (8), the cost C in Equations (6) and (7), and v. The parameter C controls how errors are regarded by the algorithm. High values of C usually imply lower errors during the training stage, but they also may lead to an SVR with lower generalization ability. To find suitable values for these parameters, we used 1 year of measured power production data. These data are regarded as the validation data set and were different from the data used for the forecasts. Series of annual forecasts were carried out with v-SVR, using different values for the configuration parameters. The parameter values that yield the best annual forecasts, within the range of values tested, were selected.

Further information about the development of SVM can be found in Smola and Schölkopf [20] and Haykin [21].

3. PROBLEM DESCRIPTION

The PV power plant analyzed in this study is located at Kitakyushu, Japan (latitude 33°55′, longitude 130°44′). The power plant maximum output is 1 MW divided in 40 static arrays of 25 kW. The arrays were installed facing south and have a tilt angle of 20°. The plant covers an area of 16 107 m^2 and started operation on 2 February 2008. The power production data used in the study were from May of 2008 to July of 2010. The data were composed of hourly measured power production. One year of forecasts were carried out from 3 July 2009 to 2 July 2010. The remaining data were used to make the parameter configuration of the SVR.

To forecast the power production of the PV power plant, six variables were used as input data in the forecast method. The six variables are described in Table I.

The variables presented in Table I are divided according to their sources. Weather forecast data and calculated data were used. The forecasted temperature, relative humidity in per cent, and cloudiness, also in per cent, in three levels regarding altitude, showed in Table I, were provided by the weather forecast system GPV–MSM of the Japan Meteorology Agency. This system uses a grid covering Japan with 505 × 481 points (latitude precision of 0.05° and longitude precision of 0.0625°). The GPV–MSM system forecasts several weather-related variables every 3 h, in Coordinated Universal Time, providing forecasts for the next 15 or 33 h, depending on the forecast time.

Table I. Variables used as input of the support vector regression (SVR) forecast method.

Source	Variable used as input
GPV–MSM data (values for the current[a] and the previous hour)	Normalized temperature, relative humidity, low level cloudiness, mid-level cloudiness, upper-level cloudiness
Calculated data (values for the current[a] and previous hour)	Extraterrestrial insolation

[a]Current hour is regarded as the hour for which the power production will be forecasted.

 Prog. Photovolt: Res. Appl. 2012; **20**:874–882 © 2011 John Wiley & Sons, Ltd.
DOI: 10.1002/pip

To have the best possible forecasts, we used 3 h of forecasts from every forecasted time. For example, from the forecasts carried out at 12 AM, the forecast values of cloudiness for 12 AM, 1 AM, and 2 AM were used; from forecasts carried out at 3 AM, the forecast values for 3 AM, 4 AM, and 5 AM were used and so on until a day of forecasts is completed. The last variable shown in Table I is the extraterrestrial insolation. This quantity was calculated in kilowatt hour per square meter for every hour of the forecast period using the method available in [22] and the simplified equation of time available in [23].

To use the SVR as a forecast method, we have to execute a training procedure. The input variables used in the training stage are the same as that described in Table I. In addition to those variables, during the training stage, the power production, in megawatt hour, was used to calculate the forecast errors and to apply the learning algorithm. The measured power production data were provided by the operator of the PV power plant. The training of the SVR for every day of forecasts was carried out using 60 days of data prior to the day being forecasted.

4. PERFORMANCE PARAMETERS

To evaluate the accuracy of the power production forecasts carried out with the SVR, three kinds of errors were considered: the root mean square error (*RMSE*), the mean absolute error (*MAE*), and the mean absolute percent error (*MAPE*). The *RMSE* and the *MAE* were evaluated on an hourly basis. As a zero power production value in the first and last hours of some days during the year prevents the *MAPE* calculation on an hourly basis, the *MAPE* was calculated on a daily basis. To calculate the *MAPE* on a daily basis, the total forecast of power production in a whole day was compared with the total measured power production in the same day. The *RMSE*, *MAE*, and daily *MAPE* were calculated using Equations (9)–(11), respectively.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_{\mathrm{fcs},i} - P_{\mathrm{msd},i}\right)^2} \tag{9}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|P_{\mathrm{fcs},i} - P_{\mathrm{msd},i}\right| \tag{10}$$

$$MAPE_{\mathrm{d}} = \frac{\left|\sum_{i=1}^{H} P_{\mathrm{fcs},i} - \sum_{i=1}^{H} P_{\mathrm{msd},i}\right|}{\sum_{i=1}^{H} P_{\mathrm{msd},i}} \times 100 \tag{11}$$

In Equations (9)–(11), $P_{\mathrm{fcs},i}$ is the power production forecast at hour i in megawatt hour, and $P_{\mathrm{msd},i}$ is the measured power production at hour i in megawatt hour. The *RMSE* and *MAE* were calculated, in megawatt hour, for all the N hours evaluated in a period of 1 year. In the case of the daily error calculation in Equation (11), the amount of hours of a single day, H, was used to calculate the daily *MAPE*, $MAPE_{\mathrm{d}}$, in percentage of the daily measured power production.

5. RESULTS

5.1. Forecast of power production with SVR and persistence

In Figure 1(a)–(d) are the forecast versus measured values of power production of the PV power plant in every season during the period evaluated (3 July 2009 to 2 July 2010). In the same figures are the coefficient of determination, indicating the correlation between the forecasts carried out with SVR and the corresponding measured values of power production. Also, in Figure 1 are the equations of the linear regression between forecast and measured value according to the season.

The results in Figure 1 show that the coefficient of determination varied between 0.80 and 0.88, indicating a good agreement between forecasts and measured values especially in spring. In principle, the best forecasts are expected in summer. However, the results can be better understood if it is considered that during summer in Japan, there is a rainy season when the weather is cloudy and highly unstable. Such weather affects the power production of the PV power plant, and also, it makes the forecasts of weather-related variables difficult. In this way, the forecast of power production is also affected. In spite of that, the determination coefficient obtained in summer shows that the forecast values had a good level of correlation with the measured values of power production.

Spring had the best coefficient of determination during the period evaluated as illustrated in Figure 1(d). This can be attributed to the fact that there is naturally a higher level of insolation in spring than winter and autumn. Moreover, spring has a more stable weather than summer (as there was no rainy season). The weather conditions in spring generated several sunny days with a high level of insolation, indicated by the highest values of measured power production in Figure 1(d). Such sunny days are easier to forecast, and they help to explain the better coefficient of determination reached in the period.

Table II shows the annual values of the errors of the forecasts obtained with SVR and with the persistence method. The table also shows the peak and average values of measured power production during the period of the forecasts.

As it can be noted in Table II, the annual *RMSE* of the forecast carried out with SVR was more than 45% lower than the one obtained with the persistence method. Moreover, both the annual *MAE* and annual $MAPE_{\mathrm{d}}$ of the forecasts of power production carried out with SVR were also significantly lower than the respective errors of the forecasts carried out with the persistence method.

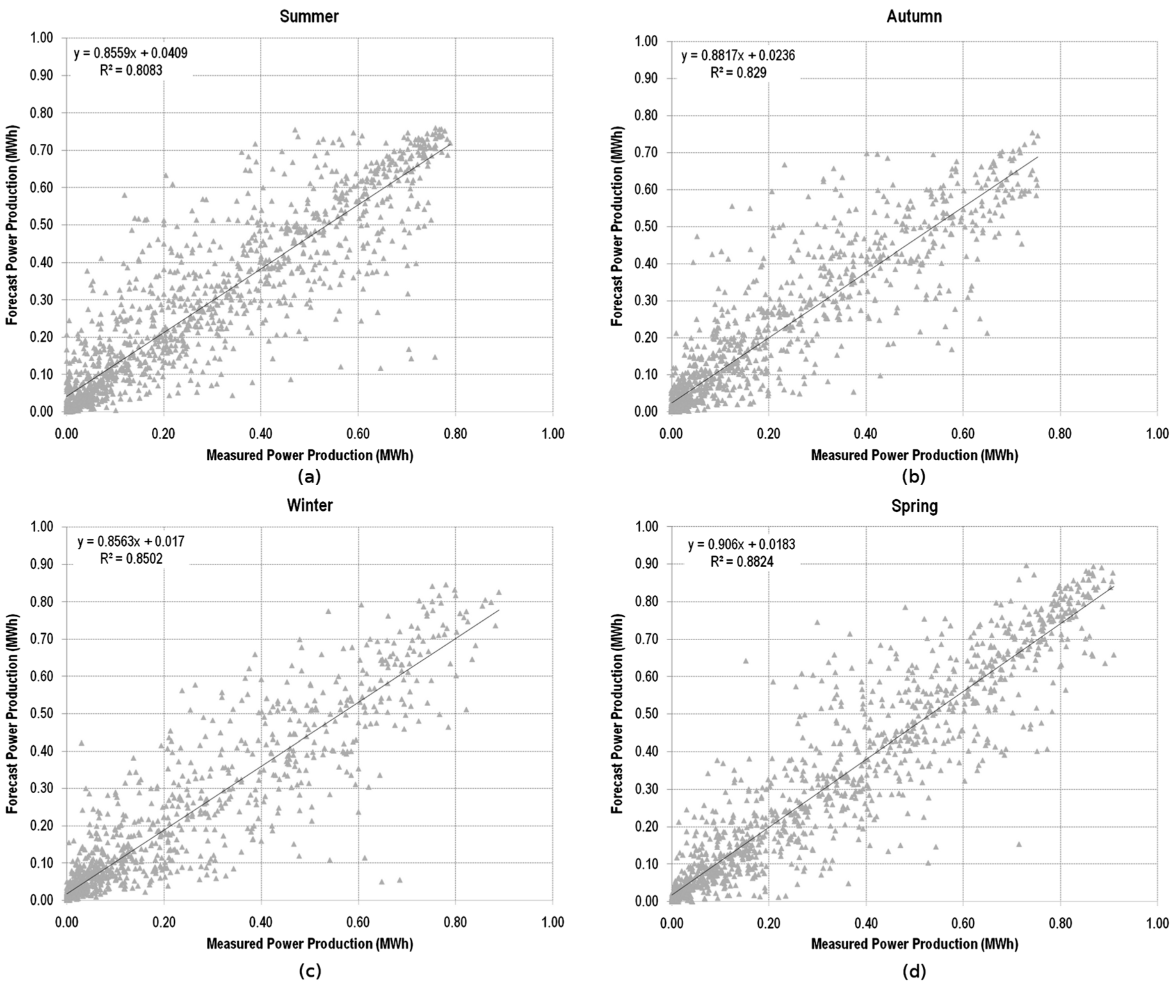

Figure 1. Forecast versus measured power production with support vector regression in summer (a), autumn (b), winter (c), and spring (d).

Table II. Annual measured power production values and average errors of the forecasts.

	Forecasts using SVR	Forecasts using persistence
RMSE[a] (MW h)	0.0948	0.2020
MAE[a] (MW h)	0.0580	0.1264
$MAPE_d$ (%)	29.53	118
Peak power production[a] (MW h)	0.909	
Average power production[a] (MW h)	0.227	

SVR, support vector regression.
[a]Hourly values

To verify the forecast accuracy under different conditions, three consecutive days of forecast and measured power production in every season were plotted in Figure 2. The specific days were selected as representatives of the conditions of mid-summer, Figure 2(a), mid-autumn, Figure 2(b), mid-winter, Figure 2(c), and mid-spring, Figure 2(d).

As seen in Figure 2, the SVR forecasts had a good level of accuracy on sunny days (1 and 2 May on Figure 2(d)) and on completely cloudy days (1 November in Figure 2(b) and 1 February on Figure 2(c)). On the other hand, in some days partially cloudy as the ones represented by 2 and 3 November, the accuracy of the forecasts of power production decreased, and the power production of the PV power plant in the peak hours is underestimated or overestimated. Furthermore, in days as 1 August in Figure 2(a) or 3 May in Figure 2(d), sudden drops or increases in the power production were not properly forecasted. In spite of these problems, the forecast versus measured values in Figure 2 and the forecast errors in Table II show that, in general, the daily trend of the power production was forecasted with good accuracy.

5.2. Effect of numerically predicted cloudiness

To investigate the effect of the use of numerically predicted cloudiness in the forecasts of the power production carried out with SVR, we removed the three variables representing cloudiness from the input data set. The forecasted initial conditions were kept the same, but a new selection of the configuration parameters values of the SVR method was carried out. The procedure to select the

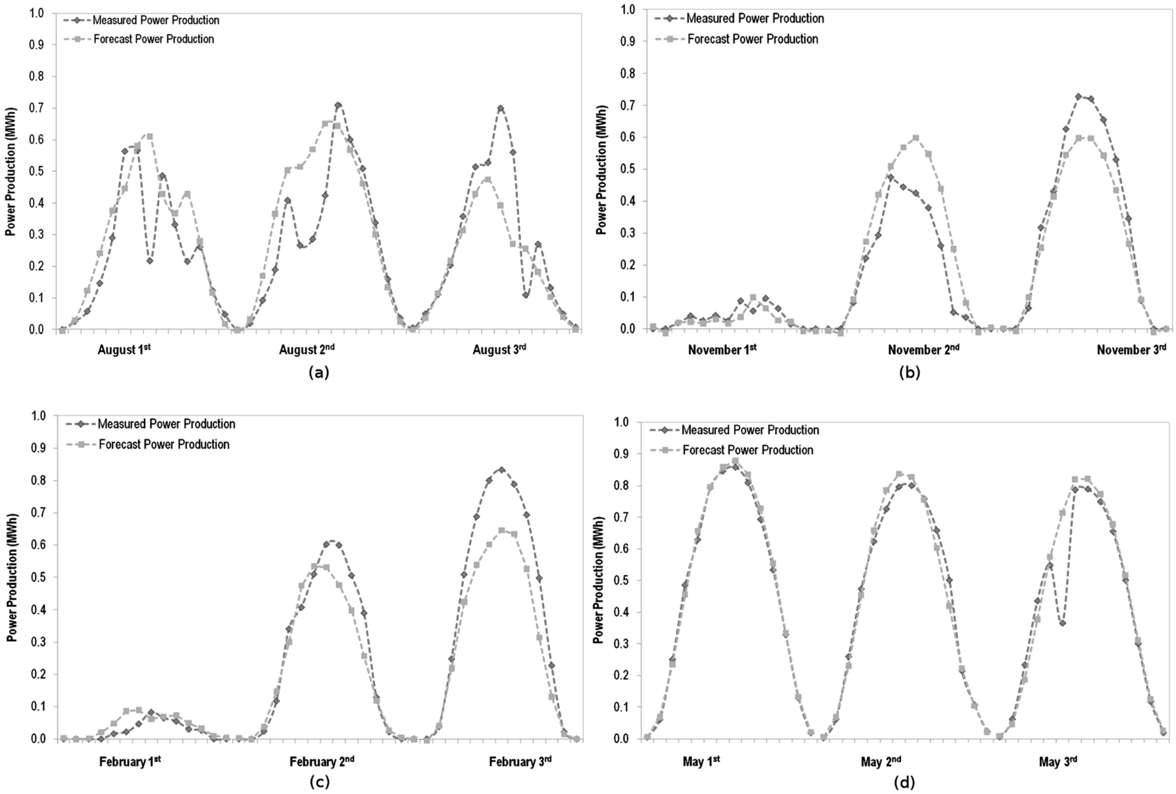

Figure 2. Power production measured and forecasted during three consecutive days (from 6 to 19 h) of mid-summer of 2009 (a), mid-autumn of 2009 (b), mid-winter of 2010 (c), and mid-spring of 2010 (d).

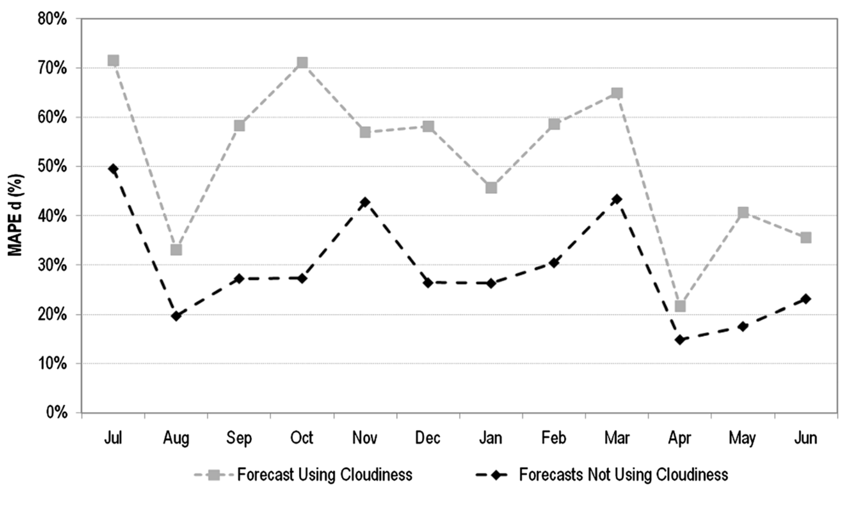

Figure 3. Monthly $MAPE_d$ of power production forecast using support vector regression with and without numerically predicted cloudiness (from July 2009 to June 2010).

values of the configuration parameters of the SVR for the case where cloudiness is not used was the same as that described in Section 2.1.

In Figure 3, the monthly $MAPE_d$ of the forecasts carried out with and without numerically predicted cloudiness are shown. In every month of the period studied, the use of cloudiness improved the accuracy of the forecasts of power production. Moreover, the increase of the $MAPE_d$ of the forecasts carried out without the use of numerically predicted cloudiness was not constant throughout the period studied, being more significant in autumn and winter than spring and summer.

The $MAPE_d$ is a daily measure of error, and it is highly affected by the strong variations that happen in the days of low power production. In Equation (11), it can be seen that in the days of low power production, a relatively small difference between measured and forecast values will be overemphasized. To verify the difference between the forecasts carried out with and without numerically predicted cloudiness without such effect, the hourly *MAE* in every month is presented in Figure 4.

As it can be noted in Figure 4, the use of numerically predicted cloudiness with SVR resulted in forecasts of power production with the *MAE* significantly lower than the *MAE* of the forecasts carried out without cloudiness. Moreover, the use of cloudiness resulted in better forecasts all over the year, confirming the trend showed by the $MAPE_d$ in Figure 3. These results indicate that the improvements caused by the use of numerically predicted cloudiness are not dependent on seasonal conditions.

Table III has the annual values of the errors calculated for the forecasts carried out with and without numerically predicted cloudiness. The same table also shows the percent differences between the errors of both forecasts of power production.

Table III shows that the *RMSE* of the forecasts carried out without cloudiness increased by almost 33% and the $MAPE_d$ of the same kind of forecasts increased by almost 80%. From these results, it can be concluded that when using SVR to forecast power production of a PV power plant, the use of numerically predicted cloudiness as input data has an important role, significantly improving the accuracy of the forecasts.

6. CONCLUSIONS

The objective of this work was to forecast the power production of a 1-MW PV power plant located in Kitakyushu, Japan, using SVMs and numerically predicted cloudiness. The SVM applied was the ν-SVR, and three variables describing cloudiness in different altitudes were used. One year of forecasts were carried out from July of 2009 to July of 2010.

The results showed that the use of the ν-SVR and numerically predicted cloudiness yield forecasts of power production with a good level of accuracy. The annual *RMSE* was around 10% of the peak value of power production, and the *MAE* was even lower reaching nearly 6% of the same peak value. Also, a good level of correlation was found between forecast and measured power production values.

Table III. Annual average errors of the forecasts using support vector regression with and without numerically predicted cloudiness.

	Forecasts using numerically predicted cloudiness (A)	Forecasts *not* using numerically predicted cloudiness (B)	% Variation (using (A) as reference)
RMSE[a] (MW h)	0.0948	0.1257	+32.60
MAE[a] (MW h)	0.0580	0.0826	+42.40
$MAPE_d$ (%)	29.53	52.95	+79.30

[a]Hourly values

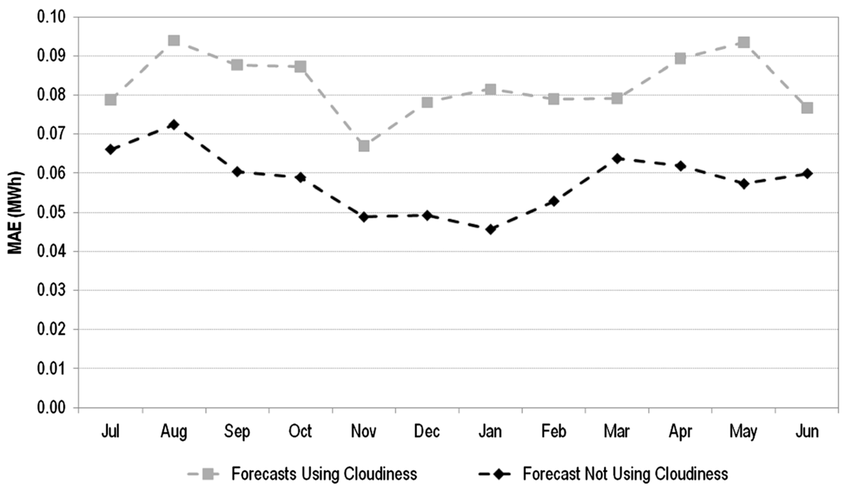

Figure 4. Monthly *MAE* of power production forecast using support vector regression with and without numerically predicted cloudiness (from July 2009 to June 2010).

DOI: 10.1002/pip

In spite of the good results, the proposed forecast method was not able to provide highly accurate forecast of power production in partially clouded days and days with sudden changes in the amount of insolation reaching the power plant. In such days, the numerically predicted cloudiness still has a low accuracy. When such cloudiness is used as input of the forecast method, its low accuracy in days with unstable weather conditions affected the forecast of power production of the PV power plant. Still, the results presented clearly showed that even with such problems, the use of numerically predicted cloudiness as input data of the proposed forecast method has an important role in improving the accuracy of the forecasts of power production with SVR throughout the period studied. Such conclusion is corroborated by the results of the three error parameters evaluated regarding the forecasts carried out with and without the use of numerically predicted cloudiness. Further improvements regarding the accuracy of the forecasts could be achieved with better selection of the training patterns or with the use of other forecast weather data such as wind speed.

The forecast methodology presented in this study had several merits. One of them, for instance, is the use of forecast and calculated variables only as input of the forecast method. In this way, local measurements of weather variables as air temperature or relative humidity, for example, are not necessary. This fact makes it possible to use the forecast method in locations where such databases are not available. Moreover, the small amount of training patterns, 60 days, used to forecast the power production of the PV power plant in each day also dispenses the use of large databases, making the forecasting methodology easily applicable.

ACKNOWLEDGEMENT

This research was supported by the Environment Research and Technology Development Fund (E-0903) of the Ministry of the Environment, Japan.

REFERENCES

1. Hunt LA, Kuchar L, Swanton CJ. Estimation of solar radiation for use in crop modelling. *Agricultural and Forest Meteorology* 1998; **91**(3): 293–300.
2. Li DHW, Lam JC. Solar heat gain factors and the implications to building designs in subtropical regions. *Energy and Buildings* 2000; **32**(1): 47–55.
3. Mohandes M, Rehman S, Halawani TO. Estimation of global solar radiation using artificial neural networks. *Renewable Energy* 1998; **14**(1): 179–184.
4. Remund J, Perez R, Lorenz E. Comparison of solar radiation forecasts for the USA. *Proceedings of the 23rd European PV Conference*, 2008; 1–9.
5. Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: a review. *Progress in Energy and Combustion Science* 2008; **34**(5): 574–632.
6. Yona A, Senjyu T, Saber A, Funabashi T, Sekine H, Chul-Hwan K. Application of neural network to 24-hour-ahead generating power forecasting for PV system. *Proceedings of the IEEE Power and Energy Society General Meeting—Conversion and Delivery of Electrical Energy in the 21st Century*, 2008; 1–6.
7. Mellit A, Pavan AM. A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy* 2010; **84**(5): 807–821.
8. Cao J, Lin X. Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique, *Engineering Applications of Artificial Intelligence* 2008; **21**(8): 1255–1263.
9. Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press: Cambridge, 2000.
10. Müller K, Smola A, Rätsch G, Schölkopf B, Kohlmorgen J, Vapnik V. Predicting time series with support vector machines. *Proceedings of the Artificial Neural Networks—ICANN'97*, 1997; **1327**: 999–1004.
11. Juhos I, Makra L, Tóth B. Forecasting of traffic origin NO and NO_2 concentrations by support vector machines and neural networks using principal component analysis. *Simulation Modelling Practice and Theory* 2008; **16**(9): 1488–1502.
12. López G, Batlles FJ, Tovar-Pescador J. Selection of input parameters to model direct solar irradiance by using artificial neural networks. *Energy* 2005; **30**(9): 1675–1684.
13. Dorvlo AS, Jervase JA, Al-Lawati A. Solar radiation estimation using artificial neural networks. *Applied Energy* 2002; **71**(4): 307–319.
14. Yujiro K, Fujiwara K, Ishihara Y, Funabashi T, Okuno Y, Nakashima H. Study on forecasting irradiation by using cloudiness of numerical prediction data. *Proceedings of JSES/JWEA Joint Conference*, 2009; 127–130.
15. Fonseca Jr. JGS, Oozeki T, Takashima T, Ogimoto K. Solar irradiation forecasts with neural networks and numerically predicted cloudiness data. *Proceedings of the 5th World Conference on Photovoltaic Energy Conversion*, 2010; 5007–5010.
16. Chang C, Lin C. LIBSVM: a library for support vector machines, 2001.
17. Chang C, Lin C. Training nu-support vector regression: theory and algorithms. *Neural Computation* 2002; **14**(8): 1959–1977.

18. Vapnik V. *The Nature of Statistical Learning Theory*. Springer: New York; 1999.
19. Schölkopf B, Bartlett P, Smola A, Williamson R. Support vector regression with automatic accuracy control. *Proceedings of ICANN'98, Perspectives in Neural Computing*, 1998; 111–116.
20. Smola AJ. Schölkopf B. A tutorial on support vector regression. *Statistics and Computing* 2004; **14**(3): 199–222.
21. Haykin S. *Neural Networks and Learning Machines*. Pearson Education: Upper Saddle River, NJ; 2008.
22. Iqbal M. *Introduction to Solar Radiation*. Academic Press, Inc.: New York, 1984.
23. Watt Engineering Ltd. On the nature and distribution of solar radiation. Prepared for the Dept. of Energy, Assistant Secretary for Energy Technology, Division of Solar Technology, Environmental and Resource Assessments Branch, 1978.