TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Prediction of power fluctuation classes for photovoltaic installations and potential benefits of dynamic reserve allocation

Citation for published version (APA):
Nijhuis, M., Rawn, B. G., & Gibescu, M. (2014). Prediction of power fluctuation classes for photovoltaic installations and potential benefits of dynamic reserve allocation. *IET Renewable Power Generation*, *8*(3), 314-323. https://doi.org/10.1049/iet-rpg.2013.0098

DOI:
10.1049/iet-rpg.2013.0098

Document status and date:
Published: 01/01/2014

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 07. Dec. 2021

www.ietdl.org

Published in IET Renewable Power Generation
Received on 1st April 2013
Revised on 17th June 2013
Accepted on 2nd July 2013
doi: 10.1049/iet-rpg.2013.0098

ISSN 1752-1416

Prediction of power fluctuation classes for photovoltaic installations and potential benefits of dynamic reserve allocation

Michiel Nijhuis[1], *Barry Rawn*[2], *Madeleine Gibescu*[2]

[1]*Alliander N.V.P.O. Box 50, 6920 AB Duiven, Netherlands*
[2]*Electrical Sustainable Energy, Delft University of Technology, Delft, Netherlands*
E-mail: mnijhuis@aol.nl

Abstract: During partly cloudy conditions, the power delivered by a photovoltaic array can easily fluctuate by three quarters of its rated power in 10 s. Fluctuations from photovoltaics of this size and on this time scale may necessitate adding an additional component to power system secondary and primary reserves to regulate frequency. This study quantifies the benefit of dynamically sizing a reserve component to cover photovoltaic fluctuations so that the additional reserves are different for each hour. The concept of categorising an hour as belonging to one of three possible fluctuation classes is presented. Based on historical array data and weather forecast information, several methods of forecasting these classes are evaluated, including persistence, Markov chains and a neural network. A practical tool based on class forecasting is proposed to aid in estimating a photovoltaic reserve requirement ahead of time for horizons ranging from 1 to 24 h. Results indicate that of a 10% possible reduction in total reserves held, most of this benefit (8%) can be obtained for hour-ahead scheduling with persistence forecasting and that a similar benefit may be possible for four-hour ahead scheduling if neural networks based on weather forecast information are introduced.

1 Introduction

Solar irradiance can contain cloud induced fluctuations, especially in oceanic climates [1, 2]. As a result, from one hour to another, the electric power delivered by a photovoltaic array can change from being very smooth to undergoing large variations. This is illustrated in Fig. 1 which shows the power time series of a 12.6 kW photovoltaic (PV) array for two consecutive days. Variability of the produced power is evident on multiple time scales. The first day starts off cloudy, but from 1PM the clouds are all gone and a smooth profile can be seen, whereas on the second day there is also a cloudy start of the day, as indicated by the low level of irradiance, but just after noon the sky becomes partly cloudy and there are fluctuations of up to 0.75 [pu/10 s] present in the solar irradiance signal. These different sky conditions have been described in [3].

Active power fluctuations can pose problems to the operation of the power system, on both the local and regional scale. Locally, voltage rise and regulation are a matter that requires attention, but can be managed [4]. On a regional scale these power fluctuations can still be significant after aggregation, [5–8] and may therefore require additional reserves to be allocated.

Additional reserves to balance photovoltaic fluctuations may not be required at all times, if periods of high fluctuation can be anticipated additional reserve can be allocated only when necessary. Based on an analysis of its irradiance, a given hour can be classified as partly cloudy, sunny or cloudy, [9], and these classes have been shown in previous work to result in significantly different reserve requirements [10]. If the class can be forecasted in advance for multiple locations, then significant reductions in the reserve allocated could be achieved by scheduling less reserve for the hours when most locations experience sunny or cloudy classes, resulting in a more efficient operation of the power system.

This paper evaluates several methods of forecasting the hourly fluctuation classification for a single location. This differs from energy forecasting [11] or cloudiness forecasting [12]. The hourly classes identified by forecasting do not match exactly those identified with detailed analysis. However, the reserves required for each type of forecasted class are significantly different. This means reductions are possible through dynamic reserve scheduling whereas still maintaining acceptable risk levels of frequency excursions and triggering of other reserves. The main contribution of this paper is a complete methodology to convert weather forecast information to classifications of future PV fluctuation states and allocate necessary reserves. To support the view that fluctuation forecasting based on classification of irradiance is worth further consideration, the paper begins with a demonstration of the potential gains of applying dynamic reserves on a regional level based on the same type of classification. The

© The Institution of Engineering and Technology 2014

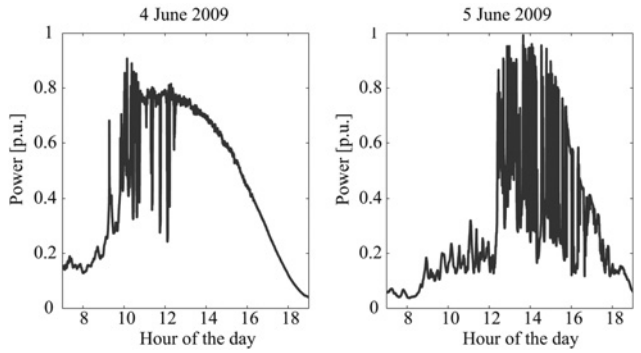


Fig. 1 *Solar power measured on 4th June and 5th June from TU Delft 12.6 kW photovoltaic array*

sections of the paper are as follows. Section 2 outlines the classification method and explains its relevance to dynamic reserves. Section 3 evaluates the methods and data which could be used for forecasting. In Section 4, forecasting results are discussed. In Section 5, the application to dynamic reserve allocation is evaluated.

2 Dynamic reserve and classification of irradiance fluctuation

The power system supplies a demand for power that is always varying throughout the day. Imbalances must be dealt with immediately and the system is continually subjected to random events, such as sudden disconnections of generators or loads, line outages, changes in the output of renewable generators and the normal fluctuations of electrical load because of human activities. Deviations from the forecasted load can always be expected. On a time frame of seconds, the kinetic energy of rotating machines connected to the system can deliver or absorb power imbalances. However generators must be capable of altering their power up or down as needed, to correct for the random types of events and processes just mentioned and changes in schedule. This capability is referred to as power system reserve and available margins of adjusting power upwards are distinguished from adjustments downwards, the former being most important for power system reliability in the face of failures. There are several categories of reserve, generally distinguished by the rate of response and whether the compensated imbalance is an event, a random variation, or a slow change [13]. Based on past experience of typical forecast errors, credible contingencies and net load variability, different quantities of each type of reserve are allocated. These quantities have augmented with an additional component in some systems because of the integration of generators based on renewable sources such as wind power [14]. The concept of adjusting this extra component over time based on the expected amount of renewable production is referred to as dynamic reserve allocation.

When examining solar irradiance data it becomes apparent that it can be divided into three different classes which correspond to different levels of cloud coverage [9, 10, 15]. The characteristics of these three classes as defined in [10] are given in Table 1 and Fig. 2 shows the power output of a typical hour for each class. This classification method has been defined for areas with fast changing weather conditions, like those with oceanic climates and is therefore performed on a hourly basis. The fluctuations shown in

Table 1 Summary of the three fluctuation classes for irradiation

Class	Mean irradiance	Fluctuation
overcast	low	low
partly cloudy	high/low	high
sunny	high	low

Fig. 2 can add to the need for regulating reserve (used to counter random variations on the time frame of seconds) as well as for ramping reserve (used to address forecast error imbalance). The next two sections describe the method of classification and its implications for reserve requirements.

2.1 Classification method

The classification is performed in two steps. The first step uses the 'clearness index', which detrends irradiance data by dividing it with the ideal values from a clear sky model [16]. The clearness index $k(t)$ is defined as

$$k(t) = \frac{G(t)}{I_0 E_0(t) \cos(\theta_i(t))} \tag{1}$$

where $G(t)$ is the solar irradiance, I_0 is the solar constant (1367 W/m^2), $E_0(t)$ is the eccentricity correction factor and θ_i is the angle of incidence the sun has with respect to the PV-panel. $E_0(t)$ and θ_i can both be calculated from astronomical relationships and depend only on time and the geographical location.

When conditions are consistently clear or overcast, the clearness index over that time interval has a probability distribution of exponential form, where the decay constant or 'rate parameter' is negative (overcast) or positive (clear) [16]. The most clear and overcast hours can therefore be distinguished by computing the rate parameter and checking its sign. This constitutes the first step in the classification process.

In the second step of processing, further characteristics are extracted from those hours in the dataset that were most clearly identified. These characteristics are used to group those hours that were not easily identified as clear or overcast. Wavelet transformations are applied to compute the energy in selected frequency bands [17] and the exponential rate parameters of the clearness index probability distribution are computed for hours in the sunny

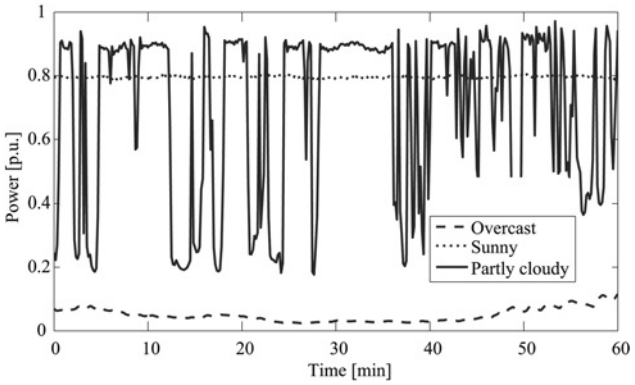


Fig. 2 *Examples of produced solar power by TU Delft 12.6 kW array for the three irradiance classes in June 2009*

© The Institution of Engineering and Technology 2014

Table 2 Single site: occurrence and characteristics of the three classes for the TU Delft dataset

class	Percentage of daylight hours, %	Mean power, pu	ΔP^{xx}, pu, 10 min		
			ΔP^{95}	ΔP^{99}	max
overcast	41	0.17	0.17	0.31	0.63
partly cloudy	33	0.46	0.31	0.44	0.69
sunny	26	0.51	0.18	0.31	0.54
all classes	100	0.35	0.23	0.37	0.69

and overcast categories, [10]. These two metrics become the defining characteristic of an hour. If an hour has similar characteristics to either the sunny or overcast category it is classified as such, otherwise it is classified as partly cloudy.

Table 2 offers more insight into the classes. Data with a 10 s time resolution measured in 2009 from a roof-mounted PV array were processed using the classification method to obtain subsets. Within each subset, changes in the 10 min rolling average power between a sample and that 10 min later were calculated as a measure ΔP of the fluctuations. The characteristics of these distributions (mean power and the 95th, 99th percentile and the maximum change in power) and the probability of a given hour belonging to each subset are presented in the table. It can be seen that the most fluctuating content is present in the partly cloudy hours.

2.2 Implications for power system reserves

The relation of the irradiance classes to reserve requirements on a regional scale will be demonstrated by studying the aggregation of eight measurement sites distributed over the Netherlands, by using 10 min maximum, minimum and mean data from the Royal Dutch meterological institute (KNMI). The available information is from irradiance sensors, which in general will result in larger fluctuations than the actual output of a whole district of rooftop (or field) solar arrays [2]. However, for a 10 min sampling interval, the difference between the two is small [18, 19]. Equal capacity at the eight sites is assumed, representing an assumption of locally uniform PV distribution and the sites fit with in a 250 km diameter. Fig. 3*a* shows that the 95th percentile of aggregate fluctuations increases in direct proportion to the number of sites concurrently having a partly-cloudy classification. The magnitude of fluctuations associated with other classifications is greater or equal to those experienced under partly-cloudy classification until the concurrence of sites with partly cloudy conditions is more than one third. Fig. 3*b* demonstrates how the probability of concurrence involving one third of sites or more is only 25% (22% + 3%) of all daylight hours.

A conservative measure of upward reserve requirements $R(n)$ for a given interval is the difference between average production and the lowest production level

$$R(n) = P_{\text{mean}}(n) - P_{\text{min}}(n+1) \quad (2)$$

where P_{mean} is the 10 min rolling average of the aggregated solar output and P_{min} the minimum power in the subsequent 10 min interval. Fig. 4 shows a sorted curve (dash-dot) of $R(n)$ required for each daylight hour of the year. When subsets of the year are chosen, namely the three ranges of

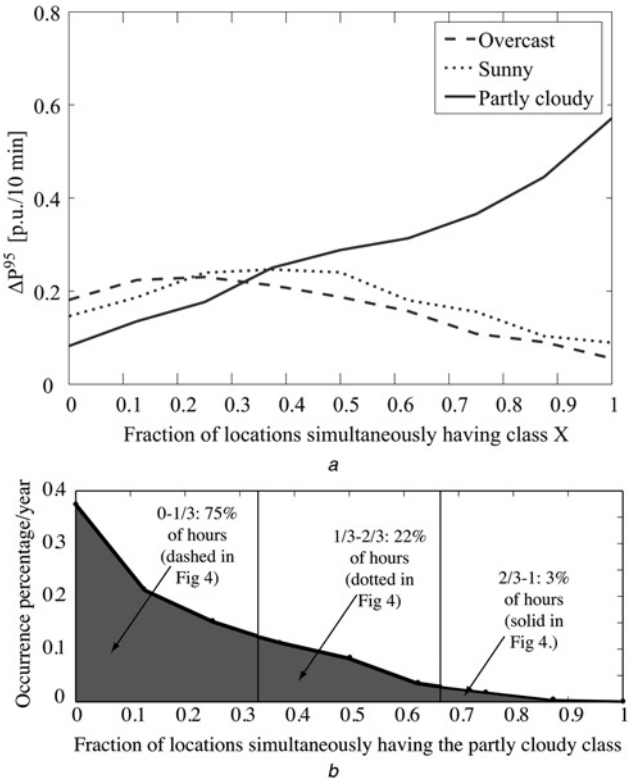


Fig. 3 *Analysis of concurrence of sites with the same classification on a 10 min. base, eight sites in Netherlands*

a Ten minute aggregated fluctuations, ΔP^{95} against concurrence
b Probability of concurrence, partly cloudy classification

concurrence marked in Fig. 3*b*, it is clear that the distribution of required reserves changes significantly with the subset.

Rather than holding sufficient reserves to cover the relatively rare maximum value of the distribution, a percentile R^{xx} of the reserve distribution may be chosen to cover xx percent of cases. Further, because the largest reserve levels do not have an equal probability of occurence at all times, holding a fixed level of reserve over the year is inefficient, as presented in Fig. 4 and Table 3. For example, a fixed reserve of 0.16 pu could be chosen based on the duration curve of all hours of the year (dash-dot curve) to cover 95% of all possible fluctuations (the R^{95} value of the complete distribution). This corresponds to a certain risk of insufficient reserves, which we define simply by reading of the x-value associated (in the example the the risk of insufficient reserves is 5%). However, the actual risk of insufficient reserves would change over time, as shown by the three subsets. Seventy five percent of the time at the

Table 3 Dependence of upward reserve requirements on different extents of partly cloudiness for the aggregated solar power of eight sites in the Netherlands

Subset of daylight hours where partly cloudy in:	Probability p_k, $k = \{a, b, c, d\}$	Reserves percentile R^{xx}		
		R^{95}	R^{99}	max
a: from 0 to 1/3 of sites	0.75	0.05	0.09	0.23
b: from 1/3 to 2/3 of sites	0.22	0.13	0.19	0.31
c: from 2/3 to all of sites	0.03	0.27	0.37	0.57
d: from 0 to all of sites	1.0	0.16	0.29	0.57

© The Institution of Engineering and Technology 2014

doi: 10.1049/iet-rpg.2013.0098

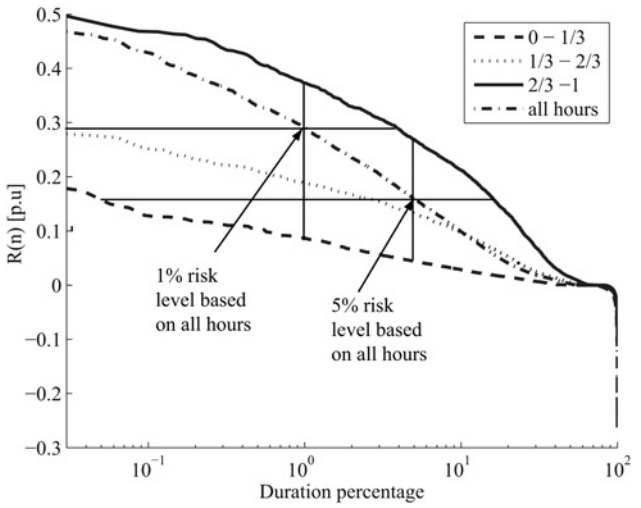


Fig. 4 *Sorted curve of upward reserves R(n) based on aggregated fluctuations of eight sites in the Netherlands, for all hours (dash-dot) and for periods having increasing fractions of partly cloudy sites (dotted, dashed, solid). Cross hairs indicate variable risk of fixed reserve (horizontal hair) and variable reserve necessary for fixed risk (vertical hair).*

regional level, the risk would be 0.05% or a hundred times less than intended. During the worst periods, a risk of 14% would be experienced instead of 5%. Unlike the approach just discussed for a single site, we propose to select three reserve levels dynamically based on the three ranges of concurrency defined in Fig. 3*b*, such that a more constant risk level of 5% can be maintained in each subset of hours. The resulting average level of reserves over a year would be less than half, as shown in Table 4. Similarly, a 1% level of risk could be obtained using dynamic reserves with an average value of 0.11 pu, instead the fixed level reserve of 0.29 pu suggested by the dash-dot line in Fig. 4 that aggregates all hours into one reserve class.

We conclude from analysis of this dataset that, in practice reduction of additional reserve levels by over half might be possible with knowledge about the fluctuation classification of different sites. The calculations presented identify reserve levels based on the duration curve of hours categorised using historical yearly information to identify the category of each site each hour. Since, the reserve level would have to be predicted using only past information and current weather forecasts, based on understanding historical duration curve of hours that were categorised. The higher the accuracy of forecasting, the greater the distinction between the categories and the more to be gained by dynamically scheduling reserves.

In further sections of this paper, we demonstrate the feasibility of forecasting categories at a single site. The gains that are apparently possible at a regional scale make further investigation of forecasting attractive. The specific reductions in reserves available for the two cases should not be compared, as the appropriate way to implement a dynamic reserve philsophy is different (i.e. using three classes against using fraction of partly cloudy sites).

Table 4 Yearly average reserve allocation based on aggregated solar power from eight sites in the Netherlands as fraction of array power

Allocation approach	Average reserve		
	Formula	$\overline{R}^{95}$	$\overline{R}^{99}$
fixed, pu	$\overline{R}^{xx} = p_d R_d^{xx}$	0.16	0.29
dynamic, pu	$\overline{R}^{xx} = \sum_{k\in a,b,c} p_k R_k^{xx}$	0.07	0.12
savings from dynamic reserves	absolute difference	0.09	0.17
	% reduction	56%	58%

3 Forecasting methodology and results

For power system operation purposes it is necessary to know the irradiance class of a given hour at least one hour in advance to coincide with the clearing of regulation and reserve markets. The class is a nominal categorical variable; it takes on three discrete values and there is no meaningful order to these values. The forecasting of this type of variable therefore requires the use of forecasting methods different from those used to quantitatively forecast solar irradiance which is a continuous variable. Inputs that may be relevant to forecasting the classification include the current and past values of the irradiance class and weather forecast information.

In the following development, we consider four models that show promise for forecasting a variable like the irradiance class. The first two are simple models based on the probabilities of the class changing between the three discrete values. These models predict the class in the next hour only based on the current class. The issue of predicting the class based on several of its previous values could be more deeply examined. However, the irradiance class can be expected to have a strong relation to weather conditions, for which detailed forecasts are available. Therefore the other two models evaluated include one or more weather variable forecasts.

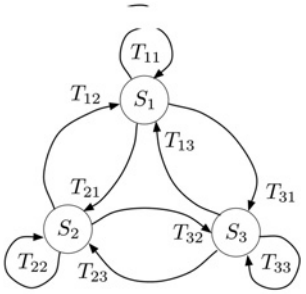
Some of these models require training before use. Training is done using the 10 s data from the year 2008, which consisted of 3000 valid daylight hours. Data from the year 2009 were used for evaluation of forecast accuracy.

3.1 Class-based models: Markov and persistence

Although the phenomena described by the irradiance class are complex rather than random, it can still be helpful to view the time series of the hourly irradiance class as a random variable X_t that takes on three states, S_1, S_2 and S_3. The table in Fig. 5*a* indicates the total probability of being in each state based on historial data obtained from the TU Delft array. Fig. 5 depicts

State	$p(S_i)$
Overcast S_1:	0.41
Partly Cloudy S_2:	0.33
Sunny S_3:	0.26

a



b

Fig. 5 *Probabilities associated with being in and transitioning between irradiance class states S_i*

a Total probability
b State diagram based on conditional probability T_{ij}

© The Institution of Engineering and Technology 2014

Table 5 Conditional probabilities of transition T_{ij}

$T_{i,j}$		X_{t-1}		
		S_1	S_2	S_3
X_t	S_1	0.759	0.204	0.139
	S_2	0.127	0.591	0.180
	S_3	0.114	0.204	0.681

the conditional probabilities of state transitions between the irradiance classs cloudy, partly cloudy and sunny.

Conditional probabilities are represented with a matrix $\boldsymbol{T}$ with entries T_{ij}

$$T_{i,j} = p\left(X_t = S_i | X_{t-1} = S_j\right) \tag{3}$$

For example, if S_1 is the overcast irradiance class and S_2 is the partly cloudy irradiance class, then $T_{1,2}$ is the probability that an overcast hour follows a partly cloudy hour. The matrix $\boldsymbol{T}$ can be derived from historical data and its values (shown in table of Fig. 5*a*) enable some observations about possible models.

First, the probabilities of moving between the states differ significantly from what might be expected from the total probabilities. For example, whereas Table 5 indicates there is a roughly 1/3 chance of being in the partly cloudy state S_2 in general, all of the probabilities associated with staying in or transitioning to S_2 shown in the middle row of Table 4 are different from 1/3. Thus taking account of these conditional probabilities may be important. Further, for each state, the probability of staying in that state (the diagonal of Table 4) is signifcantly more than the total probabilitity of being in that state (i.e. $p(X_t = S_i/X_{t-1} = S_i) > p(X_t = S_i)$). Thus, past values will matter to the prediction of the irradiance class time series.

Two simple models can be used to take advantage of these properties. A persistence model of the irradiance class makes the simplest assumption possible about how past values of the irradiance class affect present values. The class of the next hour is simply forecast as being equal to the class of the current hour. Access to the photovoltaic array measurements from the previous hour and knowledge of the irradiance class scheme are sufficient to determine this forecast. The persistence model does not require statistical information about past values.

Slightly more sophisticated than persistence, Markov chains are based on the assumption that a state transition has a probability that depends only on the current state and not on past values [20]. Future states are predicted based on the probabilities associated with transitioning from one state to the other that were observed in the past. The state diagram depicted in Fig. 5*b* completely specifies the Markov chain. The forecasted class of the next hour is determined using a random number generator based on the transition probabilities estimated from past data and the class experienced in the current hour.

A random selection from the three states defines the minimum possible accuracy of forecasts. If the irradiance class time series were truly random, it would not be possible to improve upon the forecast accuracy of such a selection. Table 6 compares the requirements and forecasting accuracy of the class-only models. Both persistence and Markov models forecast with an accuracy significantly better than random for the next hour, but do little better than a random guess after 4 h. Therefore the irradiance class series is not purely random. The fact that persistence performs better than the Markov chain may indicate that an irradiance class often lasts longer than an hour. Thus for an hourly irradiance class method, accounting for past values could yield more accurate predictions.

3.2 Forecast-based models: neural net and hidden Markov

The forecasting of class would most likely be closely related to the forecasting of the weather. Therefore it makes sense to consider models that incorporate those weather variables which are already forecasted and available. For the Netherlands many weather variables are forecasted by the KNMI (Royal Dutch Meteorological Institute). In this section, two methods for forecasting the irradiance class time from weather information are considered. Both methods involve a trained weighting and summation of forecasting inputs. The first method, a neural network, is used to determine which weather variables contribute the most to the accuracy of a forecast and is used to assess the value of incorporating a large number of inputs. The second method, a hidden Markov model, is used to evaluate the use of a single weather input.

The neural net is a highly adaptive model which can be trained to predict time series or nominal categorical variables [21]. The neural net consists of interconnected 'neurons'. For the forecasting of the irradiance class a three-layer feed-forward network, with sigmoid hidden and output neurons is chosen. The inputs of the neural net are passed on to the hidden layer together with a bias where each neuron computes a weighted average of the inputs plus this bias. The weighted averages of the hidden layer are inserted in the activation function. To set the weighted averages of each neuron and the biases, a scaled conjugate gradient backpropagation algorithm is used [21], as imlplemented in the Matlab neural net toolbox. The neural net is modelled by using the pattern recognition network from the Matlab neural net toolbox, with an input layer consisting of seven neurons, an output layer which consists of three neurons and one hidden layer consisting of six neurons. Weather forecasts and the current value of the irradiance class time series are used as inputs, whereas the

Table 6 Characteristics and forecast accuracy of models based only on class.

Method	Inputs	State variables	Parameters	% Correct, by horizon	
				(1 h ahead)	(4 h ahead)
average random	0	0	0	33.3	33.3
persistence	1	1	0	68.1	38.0
Markov chain	1	1	9	53	35.0

© The Institution of Engineering and Technology 2014

doi: 10.1049/iet-rpg.2013.0098

next value of the irradiance class time series is used as the desired output for training the neural net.

For the neural net model a large number of variables can be used; however, this makes the model more complicated and increases computation time. Not all the weather variables are therefore used as input for these two models. To be able to assess which variables are the most important, a sensitivity analysis has been performed. The sensitivity analysis is performed by using the neural net and omitting a single variable from the inputs. The difference in percentage between the forecast with all the variables and with a single variable less are calculated. The variables with the largest difference in forecast accuracy between the cases when they are omitted and included are seen as the most important variables for the prediction of the irradiance class. An overview of the sensitivity analyses is shown in Table 7.

The top six weather variables which are selected for inputs to the neural net are radiation, wind direction, cloud cover, precipitation, temperature and wind speed. Two neural net configurations will be considered, one where the current irradiance class is also included and one where it is not.

The top six weather variables were considered one by one for use as states in a hidden Markov model. The hidden Markov model is characterised by a set of states, a set of outputs, a transition probability matrix and an output probability matrix [22]. The transition probability matrix is the chance of switching between states and the output probability matrix is the chance of emitting an output from a state. A single weather forecast variable can be used to define model states; discrete variables can be taken directly as states, whereas equally-spaced ranges can be defined for continuous variables. The three states of the irradiance class are taken as outputs. The training dataset is used to determine appropriate weights mapping the states to the outputs. Forecasting results for each of the top six weather variables were extremely similar in their 1, 4 and 24 h forecast accuracies. The cloud cover variable produced the highest accuracy forecasts of the irradiance class and also has the advantage of already having discrete states and of having a logical relationship to the cloud conditions that produce fluctuations.

Table 8 shows forecasting results for models that incorporate weather forecast information. The complexity and data requirements of these are evident in the number of states and parameters. However, they also have a higher average forecast accuracy than the methods in Table 6 that used current class information alone. Two key observations can be made. Incorporating the current irradiance class improves the forecast in a way that adding additional weather variables cannot (1 h ahead results, comparing rows 1 and 2 against row 3) – but this advantage fades for longer periods (rows 2 and 3, 4 h ahead results). Adding additional weather variables is important for longer term forecasts (row 1 against row 3).

Table 7 Results from the sensitivity analysis for the weather variables and current irradiance class

Origin of variable	Name of variable	Reduction in forecast accuracy when omitted, %
measurement	current irradiance class	10.4
weather forecast (chosen for analysis)	global radiation	8.62
	mean wind direction	7.66
	cloud cover	7.48
	temperature	7.22
	precipitation	7.04
	hourly mean wind speed	6.95
weather forecast (discarded)	sunshine duration	6.42
	present weather code	5.63
	mean wind speed	4.93
	fog	4.14
	minimum temperature	3.87
	maximum wind gust	3.61
	hourly precipitation amount	3.43
	rainfall	3.43
	horizontal visibility	3.26
	thunder	2.99
	air pressure	2.73
	relative atmospheric humidity	2.02
	dew point temperature	1.85
	ice formation	1.58
	snow	1.14

3.3 Summary of forecast results

From Tables 6 and 8 it can be seen that the neural net incorporating both weather forecast and current class information is the best in predicting the future class. The very simple persistence model is almost the worst, but it also employs the least information. The average forecast accuracy for the persistance, neural net and hidden Markov model, is plotted against forecast horizon in Fig. 6. The baseline accuracy of 33% from simply choosing the class at random is also given in the figure. From this it becomes apparent that, no method can accurately predict a partly cloudy class a day ahead. However, these models can be judged to provide increasing predictive power in line with their complexity. The simple Markov chain was discarded when judged against this criteria, since it required more parameters than the persistence method without significantly improving forecast accuracy.

For prediction of the next hour's class, the persistence model is only 15% worse than the most advanced neural net. The persistence model could thus be simply implemented without the need for any other data sources than the photovoltaic array itself. Therefore the persistence

Table 8 Characteristics and average forecast accuracy of weather-forecast based models

Method	Inputs	States	Parameters	% Correct, by horizon	
				(1 h)	(4 h)
hidden Markov	2	9	27	72.4	51.3
neural net, only weather forecast)	6	100	403	72.5	68.4
neural net, weather forecast and current measured class	7	101	407	82.1	73.9

© The Institution of Engineering and Technology 2014

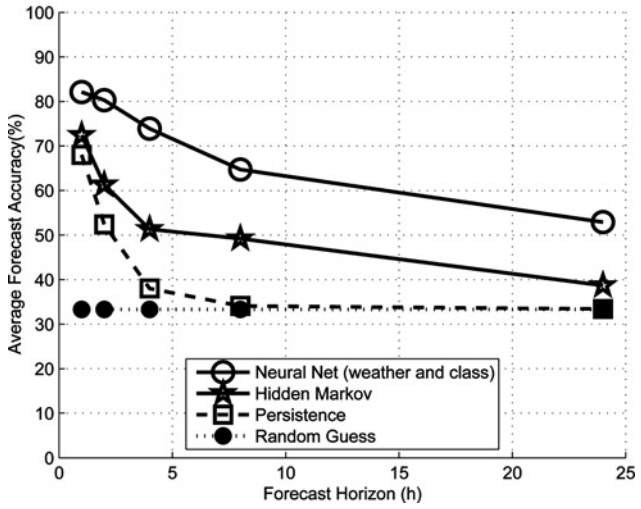


Fig. 6 *Average forecast accuracy against horizon for three forecasting methods.*

is an interesting option. However, when the forecast horizon is extended, the percentage correct predictions begins to decrease rapidly, whereas for the neural net the percentage decreases far more gradually (see Fig. 6). Also, a comparison of Tables 6 and 8 reveals that by using central forecast information, after its training the neural net can provide irradiance class forecasting as good as the persistence method without local measurements. Therefore these two models structures are most interesting and merit further investigation. In particular, it is important to consider not just average forecast accuracy, but accuracy broken down by class.

To gain more insight into the forecasting accuracy, the confusion matrix is generated. A confusion matrix is a conditional probability table that indicates the probability of an actual class occurring, given a predicted class. Confusion matrices for the persistence forecasts and for the neural net are given in Tables 9 and 10.

From these confusion matrices a number of interesting points can be observed. For both methods the prediction of the partly clouded class seems to be the hardest and the overcast hours are easiest. It can also be seen that the neural net uniformly improves the forecast because the percentage of wrong forecasts is reduced for each possibility. The reduction in confusion is 48% for the sunny hours, 48% for overcast hours and 36% improvement for partly cloudy hours.

The partly cloudy conditions produce the largest need for reserves. Allocating reserves efficiently will therefore be affected most by a method's ability to accurately distinguish partly cloudy conditions from the other two categories. To the extent the method fails, the 99th percentile of sunny and overcast fluctuation distributions may rise, because of the inclusion of erroneously categorised partly cloudy hours. There will thus be less difference in fluctuations between the categories and therefore less advantage will be

Table 9 Confusion matrix showing normalised results from the forecast of the persistence

		Predicted class		
		Overcast, %	Partly cloudy, %	Sunny, %
Real class	overcast	75.9	20.4	13.9
	partly cloudy	12.7	59.1	18.0
	sunny	11.4	20.4	68.1
		$\sum = 100\%$	$\sum = 100\%$	$\sum = 100\%$

Table 10 Confusion matrix showing normalised results from the testing of the neural net

		Predicted class		
		Overcast, %	Partly cloudy, %	Sunny, %
Real class	overcast	87.4	10.2	4.9
	partly cloudy	7.4	73.7	11.6
	sunny	5.3	16.1	83.5
		$\sum = 100\%$	$\sum = 100\%$	$\sum = 100\%$

conferred by using these forecasted categories to dynamically allocate reserve.

Fig. 7 shows the forecasting accuracy of the neural net and persistence methods for only the partly cloudy category. The accuracies of the two methods are closer together and poorer in general than the average accuracy over all classes, as shown in Fig. 6. The accuracy of a neural net using only weather forecast information is also shown.

4 Effects on dynamic reserve allocation at a single site

Whether on a regional scale or at a single site, dynamic reserve allocation by irradiance class makes a reduction in average reserves possible. Reserve levels are based on the 99th percentile of the set of hours comprising the current class of weather. Achieving the reduction requires two things: knowing the future class and knowing the distribution of fluctuations for that class.

In a practical implementation of dynamic reserve allocation, current weather and/or site information is used to determine the future class. After sufficient running time, a collection of hours can be obtained for each forecasted class. The distribution of fluctuations in this collection yields a reserve level necessary to achieve a desired risk level. Assuming the aspects of the irradiance class are stationary and adequately expressed in both the training and the running data, the desired level of risk can be achieved in practice by setting reserves based on the forecasted class. For a single site, the three different levels of dynamic

Table 11 Yearly average allocation $\overline{R}^{99}$ as fraction of array power for a single location (Delft, over the year 2009)

		Fixed reserves	Persistence (1 h)	Neural net (1 h)	Perfect forecast
$\overline{R}^{99}$ (per-unit)		0.580	0.532	0.521	0.519
Savings from dynamic reserves	absolute	—	0.048	0.059	0.061
	% reduction	—	8.2	10.2	10.5

© The Institution of Engineering and Technology 2014

doi: 10.1049/iet-rpg.2013.0098

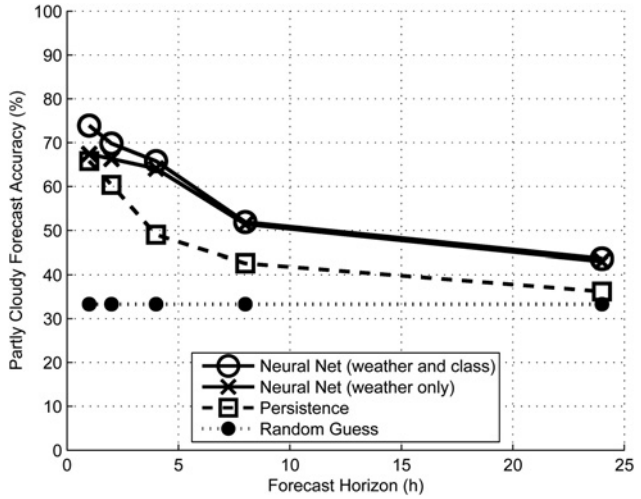


Fig. 7 *Forecast accuracy for partly cloudy conditions against horizon for three forecasting methods*

reserve are defined for each class, instead of being based on the fraction of cloudy sites as in the example of Fig. 3*b*, where a multi-site set-up with perfect forecast was illustrated.

The reduction of reserves is lessened when the forecast is inaccurate. However, the distribution of forecasted classes with normally low fluctuations may contain erroneously classified hours with high fluctuations, the difference between these two types of classes will be smaller. To illustrate the effect and compare, a base case of calculating reserves based on the whole year (no dynamic reserve) is calculated. If the irradiance class data is available, then the reserves can be set based on the 99th percentile of each class's subset of hours. The reserve allocation based on the forecasted class demonstrates the practically achievable gains, whereas the reserve allocation based on the actual irradiance class (equivalent to a perfect forecast) demonstrates the upper limit on the effectiveness of applying dynamic reserves.

For the case of fixed reserves, reserves are calculated from the duration curve of all daylight hours of the year. For the perfect forecasted case the duration curves for the three classes for hours determined by the actual irradiance class method are used. For the two prediction methods of neural nets and persistence, a duration curve needs to be computed for each possibility of forecasted and real class. These duration curves can then be used in combination with the forecast error to calculate the reserves required.

The duration curves for the two prediction methods and for the perfectly forecasted case are shown in Fig. 8. The level of fluctuation is higher than in Figs. 3*a* and 4, because of the lack of smoothing from multiple sites. From Fig. 8*a*, it can be seen that the collection of hours actually classified as partly cloudy has a distribution of fluctuations that is slightly more severe than the collection of hours identified by the two chosen forecasting methods. The consequence of forecasting error is more easily seen in the distribution of overcast and sunny hours. The hours identified by forecast methods as belonging to these two categories have distributions that include larger fluctuations. This is likely because the collections include many hours that are probably actually partly cloudy. Future reserve levels would be generated based on the characteristics of the forecasted collections shown in Fig. 8, which are but an estimate of the real characteristics.

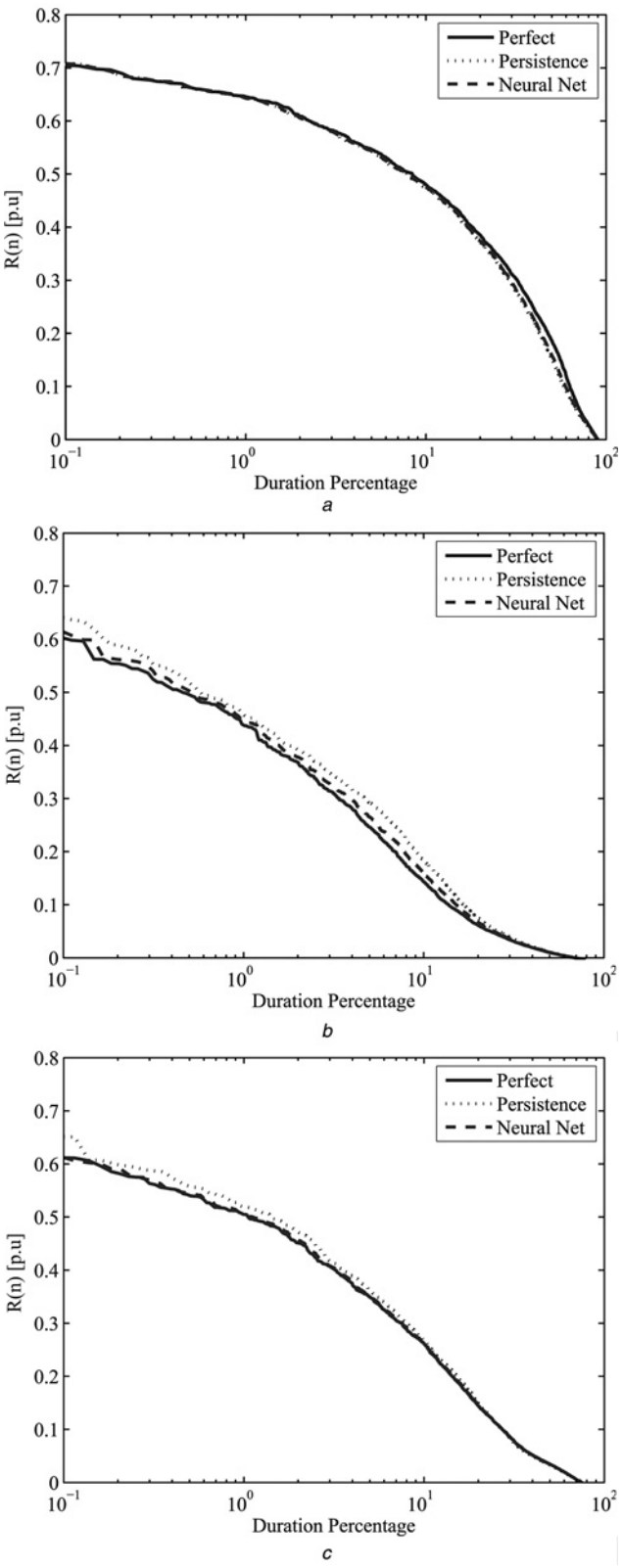


Fig. 8 *Reserve required for desired level of risks, ranging from 0.1 to 100%. Duration curves are for forecasted class types at a single site, compared with reserve required if knowledge of the class type was perfect*

a Hours classified as partly cloudy
b Hours classified as overcast
c Hours classified as sunny

Table 11 shows how an allocation of reserves based on the properties of forecastable classes that can achieve a reduction

© The Institution of Engineering and Technology 2014

in average reserves used. Most of the reduction potential of dynamic scheduling is achieved when using either of the forecasting methods. Hour ahead adjustments have been assumed and on this time frame the neural net forecast is not significantly superior to the persistence forecast. If the forecast accuracy for partly cloudy conditions can be taken as a proxy for reserve savings, then the results of Fig. 7 suggest two interesting conclusions. First, for hour ahead adjustments, either persistence forecasts based on local array measurements or neural net forecasts based on centralised weather forecasts based on local array measurements or neural net forecasts based on centralised weather forecasts could be used to achieve an 8% reduction in reserves. Second, if adjustments must be made further than one hour in advance, then neural net forecasts based on centralised weather forecasts could still be an option to achieve similar reductions.

5 Conclusions

From previous work by Nijhuis *et al.*, it was known that the reserves required to counteract power fluctuations from photovoltaic arrays are not the same from hour to hour, on both a local and regional aggregated scale. A classification of each hour into one of three irradiance classes based on fluctuations had been found to clarify this effect. The three irradiance classes were proposed as useful for implementing a dynamic reserve methodology, either by direct sorting into three levels at a single site or by allowing a determination of partly cloudy fraction over a region. This paper examined the prospect of forecasting the fluctuation classes in advance based on array and weather forecast information. Combined with knowledge of the size of fluctuations that characterise hours belonging to each forecasted class, the potential to reduce the average amount of reserves through dynamic reserve allocation throughout the year was evaluated. At a single site, the results indicated that dynamic allocation based on forecasts could achieve most of possible reduction available from using three reserve levels, which was 10% or 0.06 pu For aggregated fluctuations on a regional scale, the potential reduction might be as much as 50%. The two potential reductions cannot be directly compared because the regional and single site require different methods of implementing dynamic reserves. This paper could not provide proof of the benefit of forecasting for the regional aggregated case, as data were not available to do so. The paper did demonstrate that at both the single site and the regional level, successful dynamic allocation depends on distinguishing partly cloudy conditions from other conditions.

It was found that the hourly classes can be predicted best by using a neural net, but the use of a persistence model also shows promise as it is a simple method with reasonable forecast accuracy. Average forecast accuracies of 65–80% were achievable for 1 and 4 h horizons, compared to the baseline accuracy of 26–41% achievable based on total probabilities for each class shown in Fig. 5*a*. Information about the present hourly class was found to be as important as weather forecast information for hour ahead forecasts and combining these two sources of information improved forecasts. Present class information lost relevance for longer time frames and use of multiple weather variables improved forecasts on a 4 h horizon.

The persistence with its simplicity could therefore be seen as the most suitable model to use for the forecasting of the irradiance class. However, when the forecasting horizon is expanded the accuracy drops quickly and the persistence model can no longer be improved whereas the neural net could still be improved by for instance using more accurate weather forecast. The forecasting performance for partly cloudy conditions suggests that the reserve reductions of persistence on the 1 h horizon could still be available on the 4 h horizon if a neural net based on weather forecast variables is applied. Since the neural net has capacity for improvement, if longer time horizons are required it can be seen as the best option for forecasting.

6 Future work

The possible reserve reductions observed through this work justify further investigation of forecasting methods. A valuable next step would be to obtain weather information for all eight sites in the Netherlands and apply the forecasting methods to examine whether the fraction of partly cloudy sites on a regional scale, which was shown to be a meaningful proxy for additional reserve requirements, can be be accurately predicted. As data becomes available from jurisdictions with large amounts of solar power, it would be worth isolating periods of high photovoltaic production, determining an accompanying series of hourly irradiance class and testing whether activation of reserves is indeed higher during partly cloudy conditions.

7 References

1 Soubdhan, T., Emilion, R., Calif, R.: 'Classification of daily solar radiation distributions using a mixture of Dirichlet distributions', *Sol. Energy*, 2009, **83**, (7), pp. 1056–1063
2 Lave, M., Kleissl, J., Arias-Castro, E.: 'High-frequency irradiance fluctuations and geographic smoothing', *Sol. Energy*, 2012, **86**, (8), pp. 2190–2199
3 Jewell, W., Ramakumar, R.: 'Effects of moving clouds on electric utilities with dispersed photovoltaic generation', *IEEE Trans. Energy Convers.*, 1987, **EC-2**, (4), pp. 570–576
4 Thomson, M., Infield, D.: 'Impact of widespread photovoltaics generation on distribution systems', *IET Renew. Power Gener.*, 2007, **1**, (1), pp. 33–40
5 Otani, K., Murata, A., Sskuta, K., Minowa, J., Kurokawa, K.: 'Statistical smoothing of power delivered to utilities by distributed PV systems'. Proc. Second World Conf. and Exhibition on Photovoltaic Solar Energy Conversion, 1998, 6–10 July 1998
6 Curtright, A.E., Apt, J.: 'The character of power output from utility-scale photovoltaic systems', *Prog. Photovol., Res. Appl.*, 2008, **16**, (3), pp. 241–247
7 Wiemken, E., Beyer, H.G., Heydenreich, W., Kiefer, K.: 'Power characteristics of PV ensembles: experiences from the combined power production of 100 grid connected PV systems distributed over the area of Germany', *Sol. Energy*, 2001, **70**, (6), pp. 513–518
8 van Haaren, R., Morjaria, M., Fthenakis, V.: 'Empirical assessment of short-term variability from utility-scale solar PV plants'. *Prog. Photovol., Res. Appl.*, 2012
9 Harrouni, S., Guessoum, A., Maafi, A.: 'Classification of daily solar irradiation by fractional analysis of 10-min-means of solar irradiance', *Theory Appl. Climatol.*, 2005, **80**, (1), pp. 27–36
10 Nijhuis, M., Rawn, B.G., Gibescu, M.: 'Classification technique to quantify the significance of partly cloudy conditions for reserve requirements due to photovoltaic plants'. Proc. 2011 IEEE Trondheim PowerTech., 2011, pp. 1–7
11 Yang, D., Jirutitijaroen, P., Walsh, W.M.: 'Hourly solar irradiance time series forecasting using cloud cover index', *Sol. Energy*, 2012, **86**, (12), pp. 3531–3543
12 Mantzari, V.H., Mantzaris, D.H.: 'Solar radiation: cloudiness forecasting using a soft computing approach', *Artif. Intell. Res.*, 2012, **2**, (1), pp. 69
13 Holttinen, H., Milligan, M., Ela, E., *et al.*: 'Methodologies to determine operating reserves due to increased wind power', *IEEE Trans. Sustain. Energy*, 2012, **3**, (4), pp. 713–723

© The Institution of Engineering and Technology 2014

doi: 10.1049/iet-rpg.2013.0098

14 De Vos, K., Morbee, J., Driesen, J., Belmans, R.: 'Impact of wind power on sizing and allocation of reserve requirements'. IET Renew. Power Gener., 2013
15 Perpinan, O., Marcos, J., Lorenzo, E.: 'Electrical power fluctuations in a network of DC/AC inverters in a large PV plant: relationship between correlation, distance and time scale', *Sol. Energy*, 2013, **88**, (0), pp. 227–241
16 Jurado, M., Caridad, J., Ruiz, V.: 'Statistical distribution of the clearness index with radiation data integrated over five minute intervals', *Sol. Energy*, 1995, **55**, (6), pp. 469–473
17 Woyte, A., Belmans, R., Nijs, J.: 'Analysing short-time irradiance fluctuations by their characteristic time scales', *Proc. of the 3rd World conf. on Photovolatic Energy Conversion*, 2003, pp. 2290–2293
18 Mills, A.: 'Implications of wide-area geographic diversity for short-term variability of solar power', 2010
19 Marcos, J., Marroyo, L., Lorenzo, E., García, M.: 'Smoothing of PV power fluctuations by geographical dispersion', *Prog. Photovol., Res. Appl.*, 2012, **20**, (2), pp. 226–237
20 Papoulis, A., Pillai, S.U.: 'Probability, random variables, and stochastic processes. McGraw-Hill electrical and electronic engineering series' (McGraw-Hill, 2002)
21 Haykin, S.: 'Neural networks: a comprehensive foundation upper saddle river'. IEEE ASSP Magazine, 1998, pp. 1–842
22 Rabiner, L., Juang, B.: 'An introduction to hidden Markov models', *IEEE ASSP Mag.*, 1986, **3**, (1), pp. 4–16

© The Institution of Engineering and Technology 2014