Stat
The ISI's Journal for the Rapid
Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/sta.0000

Very short-term solar power forecasting with hidden regimes

Alexis Gerossier[1] **and Pierre Pinson**[2*]

Received 20 December 2015; Accepted ??

Generating electricity from renewable energy sources brings new challenges related to its modelling and forecasting, with their nonstationary and nonlinear dynamics. In the case of solar power, high-resolution forecasting is made especially difficult owing to cloud passages. Our objective is to explore how regime-switching approaches, with a hidden Markov chain, hence permitting to model switching dynamics without explanatory variables such as sky images, can allow to improve short-term forecasts. In practice, a novel Markov-Switching AutoRegressive (MSAR) model-based approach using a generalized logit transformation is applied to generate predictive densities for a 1-minute lead-time. An extension of an existing recursive and adaptive estimation scheme allows tracking parameter changes over time. The proposed model-based prediction approach is evaluated in a probabilistic forecasting framework using data from Queensland, Australia, while benchmarked against classical time-series-based approaches. Our results show that significant improvements can be obtained without using explanatory information from sky images.

Keywords: solar power, nonlinear time-series, regime-switching, probabilistic forecasting, generalized logit transform

1. Introduction

The share of renewable energy sources in the overall energy mix is to increase dramatically in the near future. The European Union plans to reach 20% of its energy consumption from green energies only before 2020. Similarly, several US states would like to attain 40% before 2030. Electricity operators have to face new challenges with this new type of electricity generation. Electricity may now be produced by small distributed installations, strongly subject to meteorological phenomena, e.g., wind turbines and solar panels. Schleicher-Tappeser (2012) explains that renewable energies are imposing a new bottom-up logic requiring the electricity market to adapt over time. Currently, although solar power is clearly the most abundant energy, water and wind power are predominant. And yet, photovoltaic (PV) panels have considerable advantages: they can be massively produced, they do not require a lot of maintenance and, installed on a rooftop, they can almost cover the consumption of an entire family. The leftover needs can be supplied

[1] **Mines ParisTech, PERSEE Centre, Sophia Antipolis, France** [2] **Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark**
[*] **Email: ppin@dtu.dk**

Prepared using staauth.cls [Version: 2012/05/12 v1.00]

by the electricity market. Current practice in electricity markets and power system operation require using forecasts as input to a wealth of decision-making problems, ideally in a probabilistic framework.

Solar production time-series have many specific characteristics. First of all, owing to daily cycles, power generation is null at night, then increasing and decreasing throughout the day. To normalize such time series and remove such seasonality, Bird and Hulstrom (1981) proposed to construct clear-sky models. By using local characteristics, such as latitude, panels angle and air humidity, it gives a theoretical power production if no cloud-related perturbation. After normalization, the time series exhibit non-linear patterns induced by cloud movements. Since physical clear-sky models require precise knowledge of the solar panel characteristics, it is often impractical. That is why Bacher *et al.* (2009) suggested to smooth past data instead to obtain such clear sky models for normalization.

While some studies made use of exogenous inputs for their models, e.g. satellite images for Heinemann *et al.* (2006), numerical weather predictions for Marquez and Coimbra (2011) and Chen *et al.* (2011) or sky images for Chow *et al.* (2011), others tried to generate forecasts with endogenous models, i.e., by using past values of power generation time series only. Martín *et al.* (2010) and Pedro and Coimbra (2012) made extensive comparisons of several models including ARIMA and neural networks. They both concluded that neural networks give the most efficient results for a few hours horizon. For short-term prediction, Bouzerdoum *et al.* (2013) took a hybrid approach: a SVM framework classifies the current sky regime, then an ARIMA model is employed to predict power output.

Our work places emphasis on very short-term forecasting (1 minute ahead) with an endogenous model. Our overall objective is to see how much forecast quality can be improved by using advanced regime-switching approaches to account for cloud-related perturbations. A clear-sky model, established by smoothing previous data, normalizes the solar time series between 0 and 1. A generalized logit transformation, whose asymmetry is controlled by positive unknown parameter, once applied to this normalized time series, yields a real-valued time series. We then assume that the resulting time series is generated by a Markov-Switching AutoRegressive (MSAR) model. A discrete Markov chain indexes weather regimes, which could stand for no perturbation, a few clouds, storms and so on. Switching from one regime to another at each instant is governed by unknown probabilities to be estimated. Each regime has its own set of AR parameters and its own noise variance. A recursive and adaptive scheme is proposed to track AR parameters, noise variances, regime-switching probabilities and the shape parameter of the generalized logit transformation.

The proposal of Pinson and Madsen (2012) for application to wind power forecasting is used as a starting point. The interested reader is referred to it for more details since similar notations are adopted here. Our specific contribution on the modelling and estimation side is to generalize this framework by also tracking the parameter of the generalized logit transformation. Besides, even though that transformation was originally proposed in Pinson (2012) the bounds for solar power production are not static but varying throughout the day depending upon a clear-sky model that defines the upper bound for power production.

This paper is divided in three parts. The two transformations applied to originally observed solar power production time series, first the clear-sky normalization, then the generalized logit transformation, are detailed in Section 2. The MSAR model along with our estimation scheme are presented in Section 3. In Section 4, the MSAR model is used to give predictive densities for a 1-minute lead time. Evaluation is performed in a probabilistic framework based on the Continuous Ranked Probability Score (CRPS).

2. Raw data and relevant data transformation

2.1. Data

Data used in this paper is obtained from the University of Queensland, Australia, as available through its website, `http://solar.uq.edu.au/user/reportPower.php`, for various panels on campus. We focus on a specific location for a rooftop at the UQ Centre on St Lucia Campus in Brisbane.

Solar power production at night is null and hence such period are discarded in both modelling and forecasting parts of our exercise. Moreover, if a single observation were missing, we simply repeated the previous one, as it occurred extremely rarely, only 6 times during the year 2014 (for a 1-minute resolution time series). Power production at a given time t is denoted by w_t in the following.

2.2. Clear-sky model

A clear-sky model gives the maximal power that can be generated at a specific time when no cloud is present. It is crucial in solar power forecasting since it defines the upper bound of our double-bounded time series. By taking the ratio between the actual output and the clear-sky value at that same time, the time-series is normalized between 0 and 1. The outcome series is equal to 1 when the sky is indeed clear and is less than 1 when clouds are obstructing the PV panels. Bird and Hulstrom (1981) and Gueymard (1989) proposed using physical models for that purpose, though these often become impractical to use since requiring substantial knowledge on location and panel characteristics. For instance, Ineichen (2006) reviewed different clear-sky models to highlight crucial characteristics such as turbidity. A thorough review of different models was carried out by Inman *et al.* (2013, Section 2). Besides, these clear-sky models are for solar irradiance, and not solar power generation, which the our variable of interest. Finally, since this clear-sky model is just used as a first transformation for our solar power time series, we do not need it to be extremely precise: an upper bound which takes roughly the same form as the production is enough to remove general trends.

Following the method of Bacher *et al.* (2009), our clear-sky model is constructed with a quantile regression based approach. The smoothing kernel is a two-dimensional Gaussian density: one dimension for the surrounding minutes and one dimension for the surrounding days. By choosing standard deviations equal to 25 minutes and 5 days, respectively, we truncate densities by keeping a temporal window of respectively 100 minutes and 20 days around the current instant. A quantile with nominal proportion 0.99 is chosen to almost always obtain a superior bound. This regression problem is solved using data from 2014 (and inevitably the 20 last days of 2013 and 20 first days of 2015). It yields a theoretical clear-sky production for every minute of every day of an entire year. Meteorologists would refer to that as a local climatology.

The clear-sky time series is denoted by $\{\hat{w}_t^{\text{cs}}\}$. The normalized time series — the ratio between w_t and $\hat{w}_t^{\text{cs}}$ at every time t — is $\{y_t\}$. It sometimes happens that the actual production is greater than the theoretical upper clear-sky bound, owing to the reflection of clouds beginning or finishing to obstruct panels in extraordinary layouts. In these rare cases, a prediction error is inevitable. The value of y_t is there rounded to 1.

2.3. Generalized logit transformation

Predicting the series $\{y_t\}$ constrained between 0 and 1 prevents from employing classical Gaussian noise assumptions. A second transformation is proposed to obtain a real-valued series $\{x_t\}$ that would allow to work in Gaussian framework. Originally suggested by Mead (1965) for carrot growths and used by Pinson (2012) for wind power forecasting, the generalized logit function stands as a bijection between the open segment $(0, 1)$ and the real-axis. Additionally to the

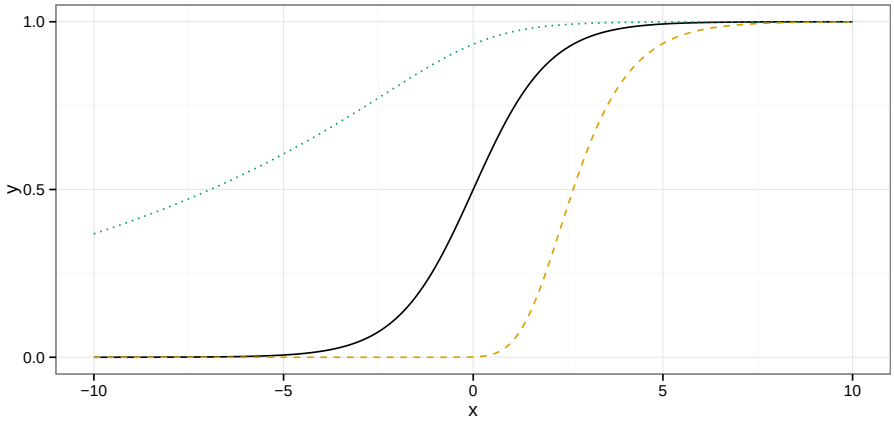


Figure 1. Examples of the $\text{lotto}^{-1}(\cdot;\kappa)$ function for three values of κ: 0.1 (red dashed line), 1 (black solid line) and 10 (green dotted line).

usual logit transform, a parameter κ changes the position of the mode relatively to the upper bound and creates an asymmetry between values close to 0 and 1. For all real x, $y \in (0,1)$ and $\kappa > 0$,

$$\text{logit}(y;\kappa) = \log\left(\frac{y^\kappa}{1-y^\kappa}\right), \qquad \text{logit}^{-1}(x;\kappa) = \left(\frac{1}{1+\exp(-x)}\right)^{1/\kappa}. \tag{1}$$

Three examples of the logit^{-1} function are depicted in Figure 1. For $\kappa = 10$ (green dotted line) the values close to 1 are more spread than for lower κ.

Although it is possible to choose a fixed κ and work with the transformed time-series $\{x_t\}$, a different approach is followed here. κ is considered as an unknown parameter to be estimated and tracked along with the dynamic model parameters themselves. Since our proposed estimation scheme is adaptive, κ can smoothly evolve through time to capture potential seasonal changes in the clear-sky impact on nonlinear dynamics of solar power generation.

In Figure 2, the time-series for 16 March 2015 is plotted alongside its two successive transformations: after normalization in the middle and after generalized logit transformation with $\kappa = 10$ on the right. Both transformations are reducing the series seasonality during a major part of the day. Series behavior looks the same from approximately 8:00 to 16:00. However the ratio just after sunrise and just before sunset clearly follows a trend. The clear-sky model would need to be refined for these periods.

3. Markov-Switching AutoRegressive (MSAR) model

3.1. Definition

The Markov-Switching AutoRegressive model was introduced in the context of wind forecasting by Ailliot *et al.* (2006) with the idea that an unobservable chain describing different regimes is more efficient than an observable one, based on Smooth Transition AutoRegressive (STAR) model for instance. We extend here the theoretical framework presented in Pinson and Madsen (2012) to add the generalized logit parameter κ. Instead of predicting the twice transformed

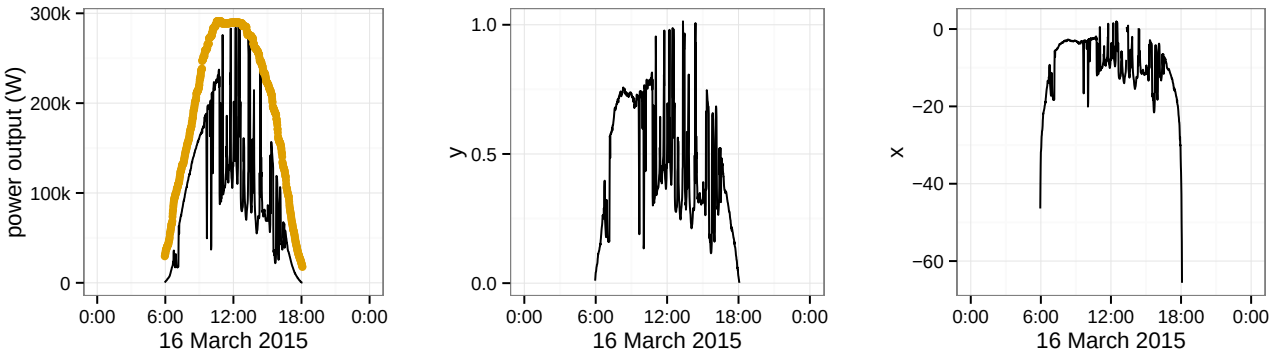


Figure 2. The power output and its two successive transformations. On the left, the real time series $\{w_t\}$ (in black) and the clear-sky one $\{\hat{w}_t^{cs}\}$ (in orange). In the middle, the normalized series $\{y_t\}$ between 0 and 1 and on the left, the generalized logit transformation $\{x_t\}$ for a fixed $\kappa = 10$.

series $\{x_t\}$, we incorporate the generalized logit transformation into the model,

$$\text{logit}(y_t;\kappa) = \theta_0^{(z_t)} + \theta_1^{(z_t)}\text{logit}(y_{t-1};\kappa) + \theta_2^{(z_t)}\text{logit}(y_{t-2};\kappa) + \varepsilon_t^{(z_t)}. \tag{2}$$

Here, 2 lags are used only (as it was found to be the most relevant number of lags through a cross-validation exercise) but the following is valid for any number of lags. The regime sequence $\{z_t\}$ takes values in the discrete set $\{1,\dots,r\}$ for a fixed integer r. It gives the regime at instant t and indexes the AR parameters as well as the standard deviation $\sigma^{(j)}$ of the noise $\varepsilon_t^{(j)}$ assumed to be i.i.d. Gaussian. For every $j \in \{1,\dots,r\}$, this noise series is independent from the $\{z_t\}$ sequence. This hidden regime chain is to capture non-linearities in power production time-series. It is expected that one regime would correspond to clear-sky conditions, while one or more would be related various cloudy conditions, with different noise variance. At each instant, there is a probability p^{ij} of switching from regime i to j. These probabilities entirely define the $\{z_t\}$ series assumed to a first-order Markov-chain.

Parameters are also transformed for them to be real-valued and easier to estimate. For all $(i,j) \in \{1,\dots,r\}^2$, we define the logarithm of the deviations $\tilde{\sigma}^{(j)} = \log\sigma^{(j)}$, the logit of the square root of probabilities $\tilde{s}^{ij} = \text{logit}\left(\sqrt{p^{ij}}\right)$ and the logarithm of the shape parameter $\tilde{\kappa} = \log\kappa$. They are gathered in our parameter vector, i.e.,

$$\Theta = \left[\theta_0^{(1)}, \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_2^{(r)}, \tilde{\sigma}^{(1)}, \dots, \tilde{\sigma}^{(r)}, \tilde{s}^{11}, \dots, \tilde{s}^{1r}, \dots, \tilde{s}^{rr}, \tilde{\kappa}\right]^\intercal. \tag{3}$$

The probability parameters are constrained since, for all $(i,j) \in \{1,\dots,r\}^2$, $0 < p^{ij} < 1$ and $\sum_j p^{ij} = 1$. We note $\mathcal{S}$, the space where Θ satisfies these conditions (which corresponds to a r-dimensional sphere, see Pinson and Madsen (2012)).

3.2. Adaptive and recursive estimation scheme

Holst *et al.* (1994) and Rydén (1997) first described how to recursively estimate MSAR parameters. A recursive framework offers several advantages: easier computation and lower memory usage make it more suitable for online application and adaptive parameter estimation. In the solar forecasting case, weather dynamics are expected to vary throughout the year, and hence its impact on solar power generation. The parameter vector is therefore then indexed with a time variable (Θ_t). A forgetting factor in the recursive estimation scheme allows for smooth variations in parameter values to be tracked..

Our two-step scheme, based on Newton-Raphson step, can be summarized as

$$\hat{\mathbf{R}}_t = \lambda \hat{\mathbf{R}}_{t-1} + (1-\lambda)(\nu \mathbf{I} + \hat{h}_t \hat{h}_t^{\intercal}), \qquad \hat{\Theta}_t = \pi_{\mathcal{S}} \left(\hat{\Theta}_{t-1} + (1-\lambda) \hat{\mathbf{R}}_t^{-1} \hat{h}_t \right), \tag{4}$$

where λ is the so-called forgetting parameter (commonly set to a value slightly less than 1), ν is a regularization parameter, $\mathbf{I}$ is an identity matrix of the same size as Θ, $\hat{h}_t$ is the gradient of the log-likelihood at time t estimated with $\hat{\mathbf{R}}_{t-1}$, while finally $\pi_{\mathcal{S}}$ is the projection onto the constrained probability sphere $\mathcal{S}$ (so that elements on each and every line of the transition probability matrix naturally sum to 1).

In order to write $\hat{h}_t$, the likelihood must be differentiated with respect to the parameter vector. Details for the general MSAR case were already given in Pinson and Madsen (2012, Appendix). Our extension to accounting for the shape parameter of the generalized logit transform reflects solely on the noise density. It is now a generalized logit-normal (GL-normal) density. Omitting the estimation hat and the time variable for clarity, it writes for regime $j \in \{1, \ldots, r\}$, for all $y \in (0,1)$,

$$\eta^{(j)}(y; \Theta_{t-1}) = \frac{1}{\sigma^{(j)}\sqrt{2\pi}} \frac{\kappa}{y\,(1-y^{\kappa})} \exp\left(-\frac{1}{2} \left(\frac{\mathrm{logit}(y;\kappa) - \theta_0^{(j)} - \theta_1^{(j)} \mathrm{logit}(y_{t-1};\kappa) - \theta_2^{(j)} \mathrm{logit}(y_{t-2};\kappa)}{\sigma^{(j)}} \right)^2 \right). \tag{5}$$

Necessary derivatives of that function are null for many of the parameters in our parameter vector. For the others, one first has, for the AR parameter in regime j,

$$\partial_{\theta_0^{(j)}} \eta^{(j)}(y; \Theta_{t-1}) = \frac{\varepsilon_t^{(j)}}{\sigma^{(j)2}} \eta^{(j)}(y; \Theta_{t-1}), \tag{6}$$

$$\partial_{\theta_i^{(j)}} \eta^{(j)}(y; \Theta_{t-1}) = \frac{\varepsilon_t^{(j)}}{\sigma^{(j)2}} \mathrm{logit}(y_{t-i};\kappa) \eta^{(j)}(y; \Theta_{t-1}), \qquad \forall i \in \{1,2\}, \tag{7}$$

while for the standard deviation parameter,

$$\partial_{\tilde{\sigma}^{(j)}} \eta^{(j)}(y; \Theta_{t-1}) = \left(\left(\frac{\varepsilon_t^{(j)}}{\sigma^{(j)}} \right)^2 - 1 \right) \eta^{(j)}(y; \Theta_{t-1}), \tag{8}$$

and finally for the shape parameter of the GL-normal density,

$$\partial_{\tilde{\kappa}} \eta^{(j)}(y; \Theta_{t-1}) = \left(\frac{1}{\kappa} + \frac{y^{\kappa} \log y}{1-y^{\kappa}} - \frac{\varepsilon_t^{(j)}}{\sigma^{(j)2}} \left(\frac{\log y}{1-y^{\kappa}} - \sum_{i=1}^{2} \theta_i^{(j)} \frac{\log y_{t-i}}{1-y_{t-i}^{\kappa}} \right) \right) \eta^{(j)}(y; \Theta_{t-1}). \tag{9}$$

4. Results

4.1. Predictive density

While some users might be interested in single-valued forecasts only, emphasis is placed here on predictive densities. Forecasts ought to be probabilistic in nature, then leaving the possibility to derive optimal single-valued forecasts in a decision-theoretic framework (Gneiting, 2011).

The predictive density issued at time t for time $t+1$ is a mixture of GL-normal densities as in (5). For each regime j, this GL-normal density is weighted by the predicted probability $\hat{\xi}_{t+1|t}^{(j)}$ of being in such regime at $t+1$. $\hat{\xi}_{t+1|t}^{(j)}$ is obtained by multiplying the switching probability $\hat{p}_t^{ij}$ by the probability of being in the ith regime at instant t. This

probability of being in the ith regime at instant t is computed recursively with (23) in Pinson and Madsen (2012). Then, by summing for every $i \in \{1, \dots, r\}$, one obtains $\hat{\xi}^{(j)}_{t+1|t}$. This eventually yields, for all $w \in (0, \hat{w}^{\text{cs}}_{t+1})$,

$$\hat{f}_{t+1|t}(w) = \sum_{j=1}^{r} \frac{\hat{\xi}^{(j)}_{t+1|t}}{\hat{\sigma}^{(j)}\sqrt{2\pi}} \frac{\hat{\kappa}}{w\left(1 - \left(\frac{w}{\hat{w}^{\text{cs}}_{t+1}}\right)^{\hat{\kappa}}\right)} \exp\left(-\frac{\left(\text{logit}\left(\frac{w}{\hat{w}^{\text{cs}}_{t+1}}; \hat{\kappa}\right) - \hat{\theta}^{(j)}_0 - \hat{\theta}^{(j)}_1 \text{logit}(y_t; \hat{\kappa}) - \hat{\theta}^{(j)}_2 \text{logit}(y_{t-1}; \hat{\kappa})\right)^2}{2\hat{\sigma}^{(j)}}\right), \tag{10}$$

4.2. Cross-validation with CRPS

Some of the model parameters not included in Θ have to be selected once and for all: number of regimes r, forgetting parameter λ and regularization parameter ν. In our case, those parameters are chosen based on a grid search in a cross-validation framework. The predictive quality of the various model setups is evaluated with the Continuous Ranked Probability Score (CRPS). This score describes how close a predictive density (say, expressed as a cumulative distributive function $\hat{F}_{t+1|t}$) is to a realization w_{t+1}, as

$$\text{CRPS}(\hat{F}_{t+1|t}, w_{t+1}) = \int_w \left(\hat{F}_{t+1|t}(w) - \mathbf{1}\{w_{t+1} < w\}\right)^2 dw, \tag{11}$$

where $\mathbf{1}\{w_{t+1} < w\}$ represents the cumulative distribution function for a Dirac placed on w_{t+1}. More details about this score can be found in Gneiting and Raftery (2007).

Here, in a one-fold cross-validation framework, we choose the set of parameters that produces the lowest CRPS in January 2015, which is $(r, \lambda, \nu) = (4, 0.995, 0.5)$. The Θ vector was first initialized at the beginning of 2014 and hence forgotten when reaching January 2015, leaving the whole year of 2014 as a burn-in period for our time-varying parameter estimation.

4.3. Comparison

Table 1 gives all the model parameters in Θ at two instants, on 1 February and on 1 July 2015. We see that two regimes, the first and fourth ones, have greater variances. Probabilities of switching from one regime to the other are important: around 20% in February and 10% in July. Those two regimes describe approximately the same state when there are clouds and perturbations, i.e. when production is uncertain. The two others have lower variances and have an extremely high probability of staying in the same regime (around 95%). They describe clear-sky state; no clouds are obstructing the panels and the production is easily predicted. One can notice that between the two instants, parameters evolve quantitatively but stay qualitatively the same. The extra parameter $\tilde{\kappa}$ slowly changes over time.

Figure 3 depicts predictive densities at two instants, i.e., on 1 July 2015. The top panel shows how the solar production varies on this very day, with two dots showing the two instants detailed below. On the left panel, the prediction obtained for 12:00 (knowing data up to 11:59), on the right for 12:10 (knowing data up to 12:09). The vertical lines represent the actual productions (hence productions to be predicted), the black curves are the predictive densities with our model and the orange dashed curve is the predictive density given by a simple AR model. Since the predictive densities for our models are mixtures of GL-normal densities, the details of the weights and deviations are gathered in the bottom tables.

By looking at the top panel, one can see that the 12:00 value can be easily predicted. Over that preceding period, the weather conditions were most likely clear. Therefore, through the iterative filtering performed by the MSAR, it then

1 Feb. 2015, 12:00	θ_0	θ_1	θ_2	σ	$p^{\cdot 1}$	$p^{\cdot 2}$	$p^{\cdot 3}$	$p^{\cdot 4}$	κ
regime 1	2.45	0.15	-0.01	0.95	75	.2	.3	25	0.10
regime 2	0.04	1.12	-0.13	0.01	.2	95	4	.5	
regime 3	0.02	1.43	-0.44	0.04	.5	5	89	5	
regime 4	-0.06	1.03	-0.01	0.12	18	.1	13	68	
1 July 2015, 12:00	θ_0	θ_1	θ_2	σ	$p^{\cdot 1}$	$p^{\cdot 2}$	$p^{\cdot 3}$	$p^{\cdot 4}$	κ
regime 1	0.96	0.73	0.03	0.67	86	.1	.2	14	0.04
regime 2	-0.01	1.47	-0.47	0.03	.1	94	2	3	
regime 3	0.01	1.72	-0.72	0.01	.3	2	95	3	
regime 4	0.07	1.01	-0.03	0.11	8	7	3	83	

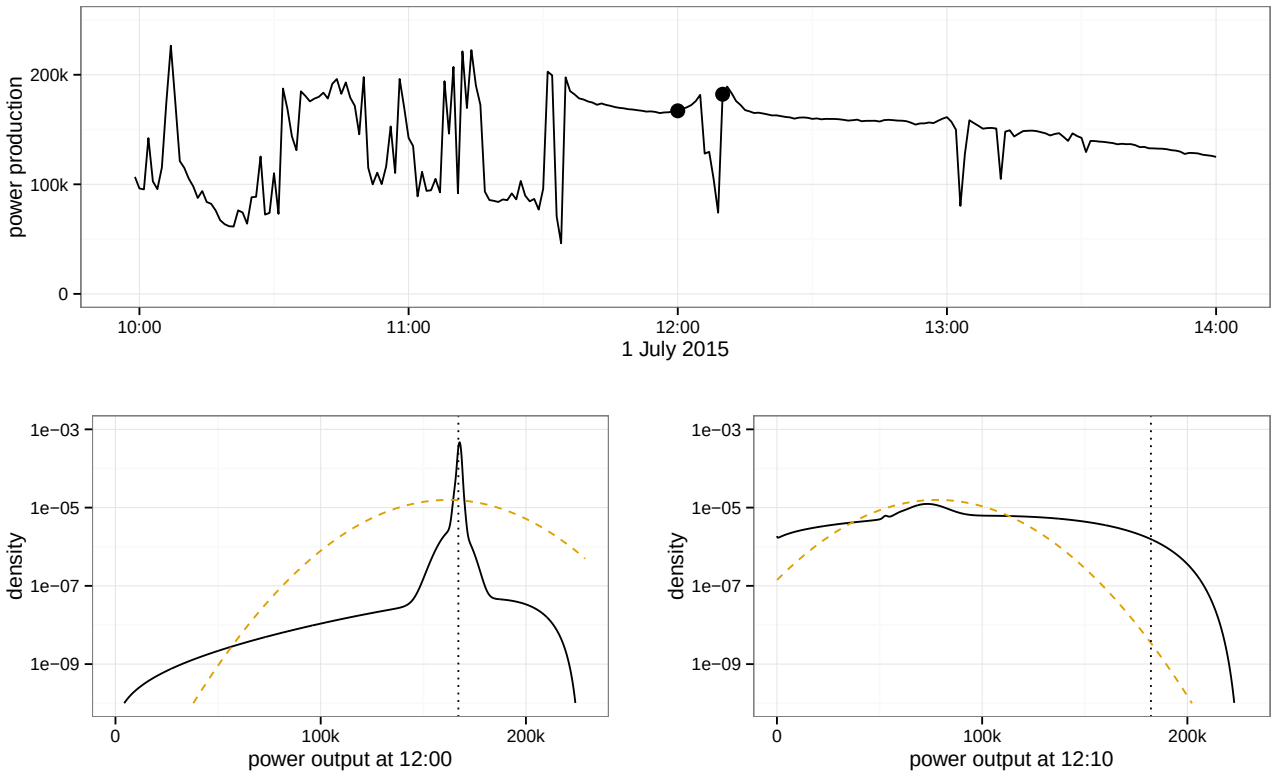
Table 1. Value of the parameters vector Θ for the 4 regimes at two instants: on 1st February 2015 and 1st July 2015 at noon. The probability to be in regime $i \in \{1, \dots, 4\}$ at next instant is denoted by $p^{\cdot i}$ and is multiplied by 100.

tend to give more weight to the clear-sky regimes (2nd and 3rd ones), hence with lower nose variance. It generates a high peak on the actual value. On the contrary, the AR model has an unconditional noise variance and does not make use of the stable period we are in. The second point, on the right, is taken at the end of a period with more power fluctuations. During a few minutes, the solar production was quite low meaning that clouds were probably obstructing the panels. Our model favors the two perturbation regimes (1st and 4th ones) with high variances because it knows that the next realization of the process will most likely be under the same regimes. The density gains a 10^3 factor on its density compared to the naive AR model.

The CRPS averaged over February to July 2015, excluding the night periods, is equal to 6926W for our MSAR model and 11187W for the AR model. The improvement is almost 40%. It is to be noted that, for a process like solar power generation, one should carefully appraise these score values. This is since such a unique score value does not reflect the fact that there is a natural tendency for the CRPS, if calculated as a function of the time of day, to increase from sunrise to mid-day and then to decrease from mid-day to sunset. This does not change the superiority of the MSAR-based approach over simpler AR models, whatever the time of day.

5. Conclusions

The MSAR-based approach presented here strives to give reliable solar power forecasts based on past value of the power time-series only. However, it allows to account for, and accommodate, some of the nonlinearities present in these time-series, mainly due to regime-switching dynamics linked to cloud passages. It was shown based on a real-world test case with high-resolution solar power generation data that issuing probabilistic forecasts based on MSAR models would yield significant benefits in comparison to using AR models (here, a 40% improvement in terms of CRPS). However, naturally, the unobservable regime framework used in this paper suffers from an unavoidable issue. At the very instant when the regime is switching (producing an important gap of production with the previous instant), the model is doomed to make a substantial forecast error. The only solution to prevent such errors would be to additionally use exogenous information as input, for instance provided by sky cameras. We would still expect that, in terms of average CRPS, MSAR-based approaches will be difficult to outperform, especially considering their relative simplicity compared to the idea of extracting and using information from sky images.



regime	1	2	3	4
probability $\hat{\xi}$ (%)	0.4	28	68	4
deviation $\hat{\sigma}$	0.67	0.03	0.01	0.11

regime	1	2	3	4
probability $\hat{\xi}$ (%)	85	0.2	0.2	14
deviation $\hat{\sigma}$	0.67	0.03	0.01	0.11

Figure 3. The MSAR model probabilistic forecasts for 1 July 2015. The top panel shows the power production from 10:00 to 14:00. At 11:59 and 12:09, forecasts are issued for 12:00 and 12:10, respectively. The black curves in the middle panel are the predicted densities with our model. The vertical dotted lines represent the realized values while the dashed orange curves are the densities given by the AR model. The tables below gather the probabilities of being in the various regime along with the corresponding standard deviation of the noise in these regimes.

References

Ailliot, P, Monbet, V, Prevosto, M (2006), An autoregressive model with time-varying coefficients for wind fields, *Environmetrics*, **17**, 107–117.

Bacher, P, Madsen, H and Nielsen, H. A. (2009), Online short-term solar power forecasting, *Solar Energy*, Elsevier **83**(10), 1772–1783.

Bird, R. E. and Hulstrom, R. L. (1981), Simplified clear sky model for direct and diffuse insolation on horizontal surfaces, *Solar Energy Research Inst.*, Golden, CO.

Bouzerdoum, M, Mellit, A and Massi Pavan, A (2013), A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant, *Solar Energy*, Elsevier, **98**, 226–235.

Chen, C, Duan, S, Cai, T and Liu, B (2011), Online 24-h solar power forecasting based on weather type classification using artificial neural network, *Solar Energy*, Elsevier, **85**(11), 2856–2870.

Chow, C. W., Urquhart, B, Lave, M, Dominguez, A, Kleissl, J Shields, J and Washom, B (2011), Intra-hour forecasting with a total sky image at the UC San Diego solar energy testbed, *Solar Energy*, Elsevier, textbf85(11), 2881–2893.

Gneiting, T (2011), Quantiles as optimal point forecasts, *International of Forecasting*, Elsevier, **27**, 197–207.

Gneiting, T and Raftery, A. E. (2007), Sctrictly proper scoring rules, prediction, and estimation, *Journal of the American Statistical Association*, **102**(477), 359–378.

Gueymard, C (1989), A two-band model for the calculation of clear sky solar irradiance, illuminance, and photosynthetically active radiation at the earth's surface, *Solar Energy*, Elsevier, **43**(5), 253–265.

Heinemann, D, Lorenz, E and Girodo, M (2006), Forecasting of solar radiation, *Solar energy resource management for electricity generation from local level to global scale*, Nova Science Publishers, New York.

Holst, U, Lindgren, G, Holst, J and Thuvesholmen, M (1994), Recursive estimation in switching auroregressions with a Markov regime, *Journal of time series analysis*, **15**(5), 489–506.

Ineichen, P (2006), Comparison of eight clear sky broadband models against 16 independent data banks, *Solar Energy*, Elsevier, **80**(4), 468–478.

Inman, P, Pedro H. T. C. and Coimbra, C. F. M (2013), Solar forecasting methods for renewable energy integration, *Progress in energy and combustion science*, **39**(6), 535–576.

Marquez, R and Coimbra, C. F. M. (2011), Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database, *Solar Energy*, Elsevier, **85**(5), 746–756.

Martín, L, Zarzalejo, L, Jesú s, P, Navarro, A, Marchante, R and Cony, M (2010), Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning, *Solar Energy*, Elsevier, **84**(10), 1772–1781.

Mead, R (1965), A generalised logit-normal distribution, *Biometrics*, **21**(3), 721–732.

Pedro, H. T. C. and Coimbra, C. F. M. (2012), Assessment of forecasting techniques for solar power production with no exogenous inputs, *Solar Energy*, Elsevier, **86**(7), 2017–2028.

Pinson, P (2012), Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions, *Journal of the Royal Statistical Society: Series C (Applied Statistics*, **61**(4), 555–576.

Pinson, P and Madsen, H (2012), Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-Switching AutoRegressive models, *Journal of Forecasting*, **31**(4), 281–313.

Rydén, T (1997), On recursive estimation for hidden Markov models, *Stochastic Processes and their Applications*, **66**(1), 281–313.

Schleicher-Tappeser, R (2012), How renewables will change electricity markets in the next five years, *Energy policy*, Elsevier, **48**, 64–75.