



Universidad
Carlos III de Madrid



This is a postprint version of the following published document:

Usaola, J. (2014). Synthesis of hourly wind power series using the Moving Block Bootstrap method. En: Institute of Electrical and Electronics Engineers. *2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. (7-10 July 2014. Durham, United Kingdom).
DOI: 10.1109/PMAPS.2014.6960602

© 2014. IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Synthesis of hourly wind power series using the Moving Block Bootstrap method

Julio Usaola

Universidad Carlos III de Madrid. Department of Electrical Engineering.
Av. Universidad 30 28911 Leganés. Spain
jusaola@ing.uc3m.es

Abstract—Reliability studies of power systems will have to include different renewable resources, because of their increasing importance in the generation mix. Wind is now the most developed intermittent renewable source of energy, and wind power series are needed to include it in Monte Carlo based studies. However, to generate correct wind power series for a given power system using ARIMA/GARCH methods is complex because of their nonlinear character and seasonality. In this paper a simpler approach based on the Moving Block Bootstrap method is proposed. This method is easy to implement, and effective if there is enough representative data. The paper describes the features of wind power production in peninsular Spain, proposes a bootstrap method to generate wind power production scenarios and shows the applicability of the method by means of an adequacy study of the IEEE-RTS system adding wind and concentrating solar power.

Index Terms—Reliability, moving block bootstrap, wind generation.

I. INTRODUCTION

Future reliability and adequacy studies will have to include an increasing amount of renewable sources of energy. Unlike the dispatchable conventional power plants, renewable resources are intermittent and with limited or no dispatchability. For this reason, wind and solar plants must be modelled in a specific way.

Monte Carlo methods present many advantages for these adequacy studies, since they allow great flexibility in modelling different generation resources. The main disadvantage of these methods is the computation time they require, but this time can be decreased using mathematical techniques, and also the computing capabilities are steadily increasing.

Monte Carlo methods in reliability require the generation of consistent scenarios of generation and demand, including production of solar and wind plants, which reproduce usually a complete year. Then, generation availability and demand is calculated for every hour in a year for the system, and this analysis is performed thousands of years until an accurate value of a given reliability index is found. These scenarios must consider the seasonality of demand, solar and wind resources, and for this reason a chronological/sequential Monte Carlo method is needed, which increases the complexity of this method.

In the case of wind energy, there are many studies that integrates this resource into the adequacy studies, many of them aiming at calculating its Capacity Credit. One early study is reference [1], where the effect of wind energy on reliability and system adequacy is assessed. References [2], [3] and [4] are studies on the capacity credit on wind energy, where [2] gives results for Portugal and [3] presents the effects of combining wind with hydro. Reference [4] is a survey of methods for obtaining the capacity credit of wind energy. The capacity credit is obtained by analytical methods in [6] and [8], and closed expressions for it have been proposed in [7].

Most of these studies did not model sequentially the wind power series, disregarding the possible dependence of the wind production with demand. More recently, however, methods for creating time series for wind production have been proposed [9], [10]. Reference [9] finds an ARIMA time series model able to reproduce the production of a wind farm. Reference [10] is more general, and an ARIMA based method is proposed to create time series of wind production for a whole area, in this case, the UK. These studies use as input wind speed series, which should be reproduced by these ARIMA time series models. These studies require a complex fitting process, since the involved series are difficult to characterize. Besides, they provide only the production of a single location (through an adequate model of wind farm), and then the results must be joined to those of other areas, taking into account the time dependence between the different areas considered.

A different approach is followed in this paper. The input data are the power production series of the whole studied system, and then a Moving Block Bootstrap (MBB) [11] technique is applied, putting together random blocks and creating as many series as needed. This method has been already used in hydrological studies [12] with different variants, as explained below. This method, as explained later, is only applicable to power systems with an already important penetration of wind energy and long enough data series to be statistically meaningful. If this is not the case, it could be used to generate time series of wind speed and proceed then to convert these wind series into power series. MBB has also the disadvantage that no extreme values beyond those already recorded can be considered. A long time series is therefore needed to capture a large enough range of events. Its simplicity and good statistic properties, however, make it worth of consideration.

The Spanish Ministry of Education supported this work through the project IREMEL (Ref ENE2010-16074).

This paper describes the proposed method to synthesize any number of wind power series from the Spanish peninsular wind production between 2007 and 2010. To do this it is necessary to characterize and adapt the data, which is a study that has interest in itself. The method is applied to the reliability assessment of a power system with different levels of wind power. To the wind power series, a series of Concentrating Solar Power (CSP) production has been added, to assess the differences between these two renewable resources, as well as the differences between sequential and non-sequential Monte Carlo methods.

The paper is structured as follows. Section II analyzes the input data, describing the main features and seasonality of peninsular Spain wind power production. In section III the Moving Block Bootstrap method is described. Section IV describes how the CSP production series have been created. In section V the example is described, and the results of the simulation study are given. The paper ends with the conclusion of section VI.

II. ANALYSIS OF WIND POWER PRODUCTION OF SPAIN.

The available data are the average hourly productions in peninsular Spain for the years 2007-2010. The series is represented in Figure 1 and data are given in Table 1.

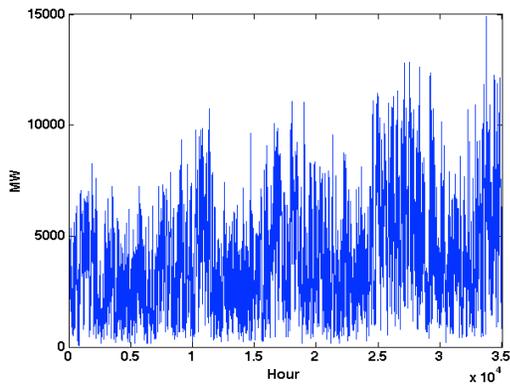


Figure 1. Hourly average production in peninsular Spain between 2007-2010.

These data include the installed wind power at the end of each year and the maximum, average and minimum values of each year, normalized to the installed power. It has been considered that the installed power changes linearly within the year. At the end of 2006 there were 11470 MW of wind power installed.

TABLE I. DATA OF SPANISH PENINSULAR WIND PRODUCTION. (P.U. OVER THE INSTALLED POWER).

	Minimum (p.u.)	Average (p.u.)	Maximum (p.u.)	Installed power (MW)
2007	0.0037	0.2319	0.6884	14107
2008	0.0157	0.2384	0.7291	16148
2009	0.0095	0.2354	0.6787	18961
2010	0.0102	0.2506	0.7498	20057

The statistical distribution of the production along this time, normalized to the installed power, is shown in Figure 2, together with the Weibull approximation of the

distribution, which are very close to each other. It must be remarked that the Weibull distribution is what is normally used for modelling the distribution of wind speed in the long term, while the data shown here are power outputs of the whole peninsular Spain.

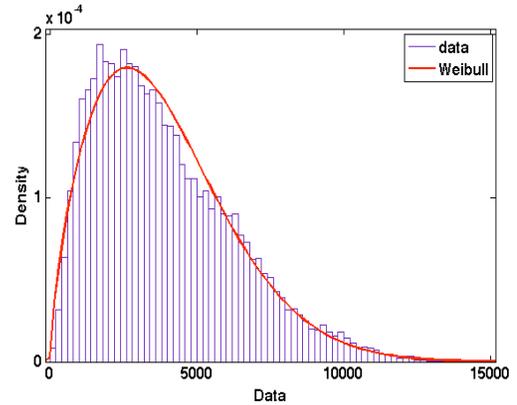


Figure 2. Distribution of wind production in peninsular Spain between 2007 and 2010.

The frequency components of the series are shown in Figure 3, where two main periodic components may be seen, corresponding to a yearly and a daily components. The yearly periodicity becomes apparent, if the moving average (2 months) of the series and the periodic component coming from the FFT analysis are represented, as in Figure 4.

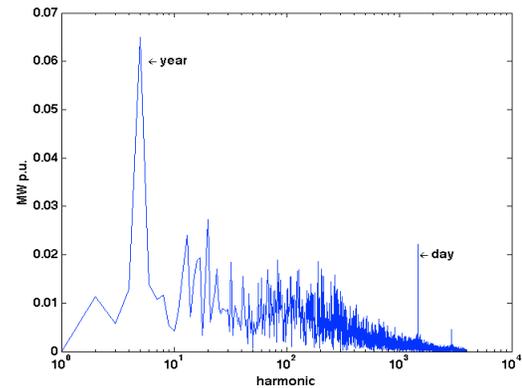


Figure 3. Spectrum of the production in peninsular Spain between 2007 and 2010.

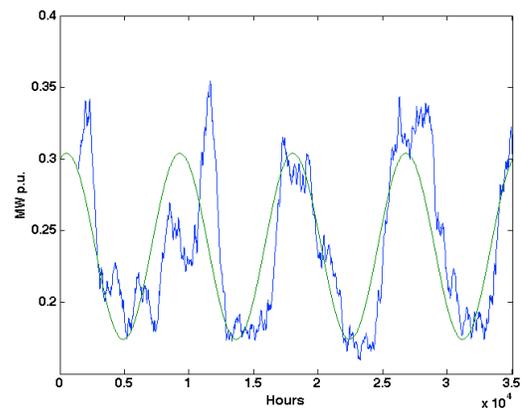


Figure 4. Moving average (two months) and yearly harmonic of the wind power production in Spain between 2007 and 2010.

To see the daily component of the series, it is useful to look at the average hourly value along the four examined years, shown in Figure 5. In this figure, the average daily production per season has been also included. Winter includes the months January-March; spring goes from April to June, summer from July to September, and autumn from October to December. The differences between seasons cannot be easily seen in the frequency plot. It can be concluded that the wind production reaches its peak at 20 hours, and that the production is higher in winter and lower in summer. This pattern has points in common with the demand in Spain, where the maximum is at 20 hours and the highest consumption takes place in winter. The dispersion of this production is however very large, as may be seen in Figure 6, where the boxplot of the distribution of hourly productions along 4 years are represented.

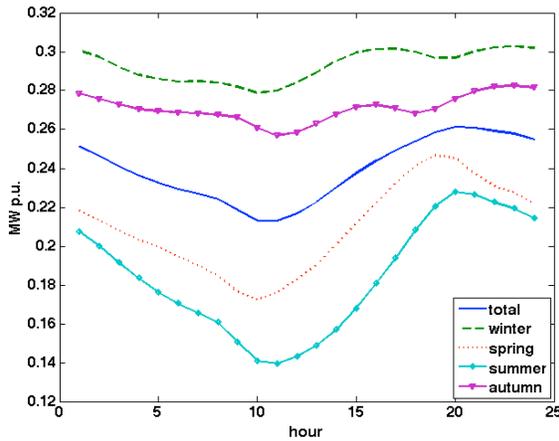


Figure 5. Daily average production total and per season. Values in p.u. over the installed power.

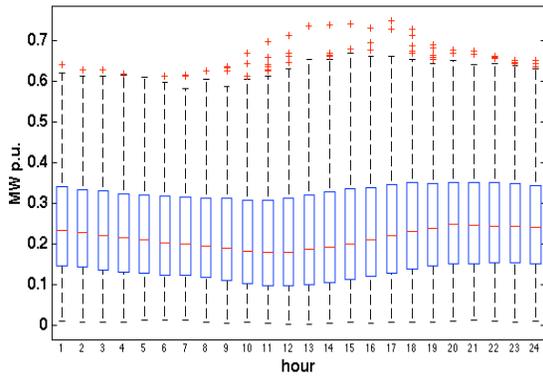


Figure 6. Boxplot of the daily wind production in peninsular Spain between 2007 and 2010.

In this series the curtailments (reductions of wind energy production) imposed by the TSO due to security reasons are also included. Although they suppose only a small part of the energy produced (0,18% in 2011[13]), they do alter the time series sequence, introducing a non-predictable, random component.

The recorded wind power series has important features, which make it difficult to reproduce or estimate. Figure 7 shows the autocorrelation and the partial correlation of the normalized series, showing its complex structure, difficult to

model by ARIMA series. The residuals after an adjustment to a linear series shows at least heterokedasticity and different trials for fitting have shown that the residuals are not Gaussian. Even if more in-depth analysis could yield better results, the fitting of the series is very problematic and there are serious doubts that new synthesized series could give general results. All these difficulties lead to look for an alternative and simpler method to generate scenarios.

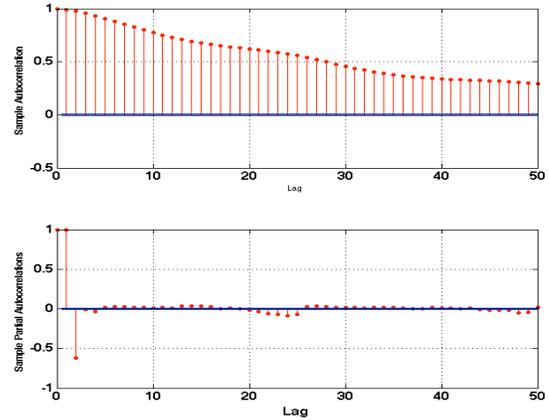


Figure 7. Autocorrelation and partial correlation of series .

III. WIND POWER SERIES SYNTHESIS USING THE MOVING BLOCK BOOTSTRAP METHOD.

The method proposed here to produce wind power series is a variant of the Moving Block Bootstrap method [11]. Bootstrapping is a widely used technique consisting in resampling a record, with replacement, to generate B bootstrap samples, from which can simulate B estimates of a given statistic, leading to an empirical probability distribution of the statistic [12]. It is a nonparametric method, simple to implement, although it presents some limitations, as will be shown later.

These nonparametric techniques have been used for synthesizing meteorological series [12]. Among them it might be mentioned the k -nearest neighborhood resampling [13] [15] or the hybrid approach of [16]. The best and simplest one for the problem considered here is the moving block bootstrap used in [12]. The other proposed techniques were developed mainly for hydrological series, which have different features from the hourly wind production series considered here. Besides, the k -nearest neighborhood raises considerable problems for the selection of the optimal dimensioning of the kernel functions needed in it [16], and the residual resampling schemes issued in the hybrid approach does not have great advantages over the simpler moving block bootstrap method [12]. In the method proposed here, a variant of the k -nearest neighborhood will be applied for improving the selection of the subsequent block, aiming at improving the similarity between the original and synthesized series. One of the disadvantages of bootstrap methods is that no extreme values can be reached beyond the limits of the available data, from which the series are generated. To use it, a significant number of samples must be available.

The outline of the method used here is given below. To preserve the statistical properties of the original series, some criteria have been used to select an appropriate block among all the possible ones.

- The series from which the consecutive days are taken is normalized (divided by the installed power at each hour) and the periodic yearly component is removed.
- A new yearly series is synthesized by taking random blocks of consecutive values of the so transformed series and putting them sequentially. Each sample block must include a few days, in order to preserve the daily periodicity of the series.
- The transition between consecutive blocks should be smooth, so that the difference between the last value of one block and the first value of the next block should be smaller than a given value.
- To preserve the seasonal properties of the series, the values of one season (spring, summer, autumn or winter) are taken from days of the same season of the original series.
- Finally, the periodic component is added and the series is scaled to the installed wind power.

From the given data, new yearly wind power series have been synthesized and the comparison between them and the original ones is given below. For the sake of comparison, 4 years have been synthesized and compared with the original data in Figure 8. It can be seen that the properties of the synthesized series are similar to the original one: the average daily value per season is shown in Figure 9 and compared to the average values of years 2007-2010. The autocorrelation and partial autocorrelation (Figure 10), are also similar to the original series (Figure 7). Other analysis yield also similar results: the synthesized series are a good replica of actual production series, with the same daily and seasonal pattern and preserving the time series properties.

IV. SERIES OF SOLAR THERMAL PLANTS PRODUCTION.

The adequacy of a power system with wind and solar thermal plants will be assessed as an example of the application of the proposed method of synthesis of hourly wind power series. For this reason, a short explanation of the synthesis of Concentrating Solar Power (CSP) production series is given in this section.

Concentrating Solar Power plants use the direct component of solar radiation to heat a fluid and produce electricity by means of a thermodynamical cycle. A special feature of this technology is its ability to store thermal energy and use it after sunset. These capacity makes them dispatchable within certain limits.

Including CSP in power systems adequacy studies requires the generation of random series of irradiation from which the production series is obtained. Since in planning studies the location of the future plants is not known with accuracy, and the required data might not be available, the approach followed here is to generate these series from public available data [17] following the method described in [21] and [22], which is summarized below.

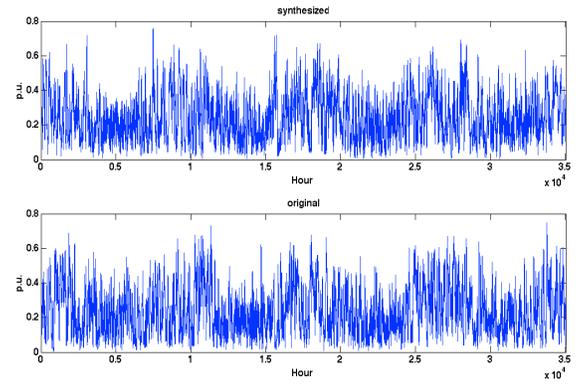


Figure 8. Original and synthesized normalised wind power series.

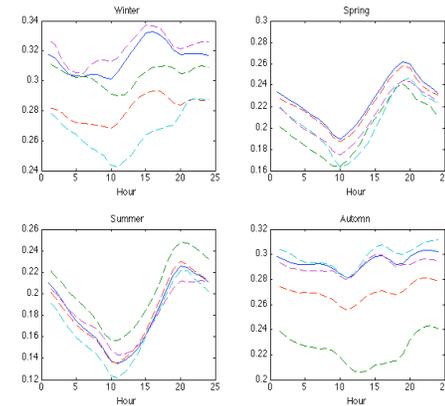


Figure 9. Average daily profile of synthesised series (solid line) and the four data years (dotted).

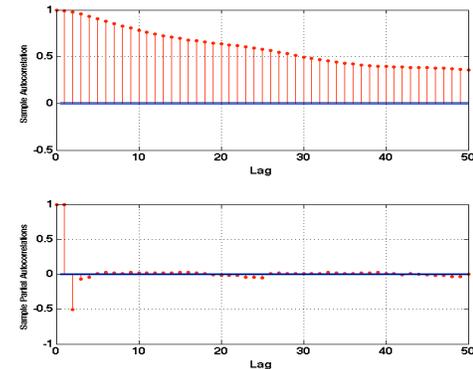


Figure 10. Autocorrelation and partial autocorrelation of synthesized series.

From the deterministic value of extraterrestrial irradiance value the irradiance on a horizontal plan on the surface of Earth is obtained multiplying it by the hourly clearness index. A series of this clearness index is created from average monthly values available in [17]. First, random daily values are obtained using Markov matrices [18] and from them, hourly values are obtained using the technique described in [19]. Then, the direct (beam) component of the irradiance must be separated. This is made using the method of Erbs [20]. From the value of beam

radiation on a horizontal surface the incident radiation on the parabolic mirrors is obtained.

The irradiance values obtained in this way are converted into power series by means of a simplified mathematical model of the CSP plant [22]. This model calculates the output power of the solar field taking into account its average efficiency, and the plant is managed in a simple way: the incoming solar energy is used to produce electricity up to the rated power of the generator; the excess of energy is stored and used after sunset. The excess energy when the storage is full is dumped. The warm up energy needed to start the thermal process, the startup energy of the plant and the self consumption of the plant have been also considered. The daily average production of a plant with 7.5 hours of storage is shown in Figure 12.

V. EXAMPLE.

To illustrate the application this method, the generated wind power series have been used for reliability calculations in the IEEE-RTS test system [23] with a peak demand of 2850 MW and 3405 MW of installed thermal power plants. For these values the value of LOLE is 9.338 hours/year. New values of Loss of Load Probability (LOLE) have been obtained for different penetration levels of wind energy with sequential and non-sequential Monte Carlo simulation methods. CSP generation has been also added to assess the effect of both renewable resources on reliability

The CSP generation consists in four plants of the same rated power and technology placed in four sites in Spain given in [22]. Their production has been considered. The plants may store 7.5 hours of the rated power.

The non-sequential adequacy calculations are made in two steps. First, a sequential list of randomly generated values of generation margin (conventional generation minus demand) is created. To this margin, values of wind/CSP generated power are added, coming from the generation scenarios created for both renewable resources. Although both assumptions and method are arguable, these calculations have been performed in this paper just to assess the interest of the wind power scenario generation, and the results are not intended to reproduce any real situation.

The limits of this approach must be also clearly said. Although the sequential nature of the solar production and the wind energy is considered in this study, the climatic dependence of windy and sunny years have been not modelled. Also, in the case of CSP, the demand profile of countries with enough resource is different from that considered here, and this has impact on the reliability indices. Although the obtained values may be an approximation to the real ones, specific studies should be made in a real system, to obtain accurate values of the contribution of these renewable resources to the overall system reliability. The main purpose of this example is just to illustrate the application of sequentially generated wind power series.

The values of the Lost of Load Expectation (LOLE) are presented in Figure 11. for different penetration levels for wind energy alone and CSP alone. The penetration level is the percentage of demand covered by the renewable resource.

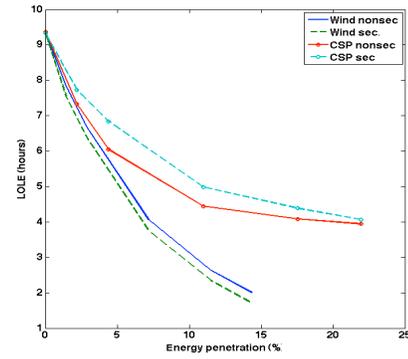


Figure 11. LOLE values for increasing penetration levels of wind energy and CSP.

In this figure it can be observed that the increasing penetration of wind energy leads to lower LOLE values, i.e. more reliable systems, than the CSP, both for sequential and non sequential methods. This is due to the largest time when the CSP technology does not produce because there is no sun and the storage is empty. The production is also lower in the winter months, where the demand chosen here is higher. A comparison of the daily average patterns of load and CSP production, in p.u. over their respective maximum values, is given in Figure 12.

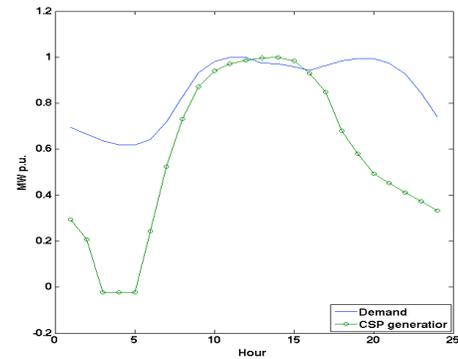


Figure 12. Daily average pattern (in p.u.) of demand and CSP generation.

TABLE II. SIMULATION RESULTS.

P (MW)	E(%)	Wind		CSP		
		LOLE (h/year)		E(%)	LOLE (h/year)	
		seq	n/seq		seq	n/seq
100	1.43	7.56	7.85	2.19	7.73	7.32
200	2.87	6.39	6.67	4.39	6.84	6.05
500	7.17	3.76	4.07	10.97	4.99	4.45
800	11.47	2.35	2.64	17.56	4.38	4.08
1000	14.34	1.71	2.00	21.95	4.07	3.95

Results are also given numerically in Table II, where ‘P’ is the installed power of wind or CSP, ‘E’ is the penetration level in % of the demand, and ‘seq’ and ‘n/seq’ the values of LOLE with the sequential and non sequential method, respectively. For the same installed power, the energy given by the CSP plants is higher due to the storage capability. The table and the figure give the values when only wind or CSP plants are added.

The author gratefully acknowledges the contribution of Prof. Ismael Sánchez for his useful hints about the bootstrap method. The data have been friendly provided by Red Eléctrica de España, the Spanish TSO.

REFERENCES

- [1] R. Billinton, A.A. Chowdhury "Incorporation of wind energy conversion systems in conventional generating capacity adequacy assessment" *IEE Proceedings-C* vol 139 pp. 47-56 (1992).
- [2] R.M.G. Castro, L.A.F.M. Ferreira "A comparison between chronological and probabilistic methods to estimate wind power capacity credit". *IEEE Transactions on Power Systems* vol. 16 pp. 904-909 (2001).
- [3] M.J. Castro, A.A. Shakoob, D. Pudjianto, G. Strbac "Evaluating the capacity value of wind generation in systems with hydro generation" *Proc. of 16th PSCC*. Glasgow (2008).
- [4] C. Ensslin, M. Milligan, H. Holttinen, M. O'Malley, A. Keane "Current Methods to Calculate Capacity Credit of Wind Power, IEA Collaboration" *Proceedings of 2008 IEEE Power and Energy Society General Meeting*. Pittsburgh (2008).
- [5] R. Billinton, H. Chen, R. Ghajar "Time-series models for reliability evaluation of power systems including wind energy," *Microelectron. Reliab.* vol 36 (9), pp. 1253-1261, 1996.
- [6] M. Amelin "Comparison of capacity credit calculation methods for conventional power plants and wind power". *IEEE Transactions on Power Systems* vol. 24, pp. 685-691 (2009).
- [7] K.R. Voorspools, W.D. D'haeseleer "An analytical formula for the capacity credit of wind power" *Renewable Energy* 31 pp. 45-54 (2006).
- [8] N. Zhang, C. Kang, D.S. Kirschen, Q. Xia "Rigorous model for evaluating wind power capacity credit". *IET Renew. Power Gener.* vol 7, pp. 504-513 (2013).
- [9] P. Chen, T. Pedersen, B. Bak-Jensen, Z. Chen "ARIMA-Based Time Series Model of Stochastic Wind Power Generation," *IEEE Trans. on Power Systems*, vol. 25, pp. 667-675, May 2010.
- [10] A. Sturt, G. Strbac, "Time-series modelling for the aggregate Great Britain wind output circa 2030," *IET Renew. Power Gener.*, vol. 7 (1), pp. 36-44, 2013.
- [11] M. R. Chernick, *Bootstrap Methods. A Practitioner's Guide*. New York: Wiley, 1999.
- [12] R. M. Vogel, A. L. Shallcross, "The moving blockstrap versus parametric time series," *Water Resources Research*, vol. 32, pp. 1875-1882, June 1996.
- [13] B. Rajagopalan, U. Lall, "A k-nearest neighbor simulator for daily precipitation and other weather variables," *Water Resources Research*, vol. 35, pp. 3089-3101, Oct. 1999.
- [14] AEE *Eolica '13* (in Spanish). Available at <http://www.aeeolica.org>.
- [15] T. Lee, T. B. M. J. Ouarda "Identification of model order and number of neighbors for k-neares neighbor sampling," *Journal of Hydrology*, vol. 404, pp. 136-145, 2011.
- [16] V.V. Srinivas, K. Srinivasan, "Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows," *Journal of Hydrology* vol. 302, pp. 307-330, 2005.
- [17] NASA Atmospheric Science Data Center. *Surface meteorology and Solar Energy*. http://eosweb.larc.nasa.gov/HPDOCS/proj_sup.html
- [18] R.J. Aguiar, M. Collares-Pereira, J.P. Conde "Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices" *Solar Energy* 40, pp. 269-279, 1988.
- [19] R. Aguiar, M. Collares-Pereira "TAG: a time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation". *Solar Energy* vol. 49, pp. 167-174, 1992.
- [20] D.G. Erbs, S.A. Klein, J.A. Duffie "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation." *Solar Energy*, vol. 28, pp. 293 - 302, 1982.
- [21] Usaola, J.: "Stochastic production of CSP solar plants for generation planning". In *Proc. Second Int. Workshop on Integration of Solar Power into Power Systems*, Lisbon, November 2012.
- [22] J. Usaola, "Capacity credit of concentrating solar power," *IET Renewable Power Generation*, in press.
- [23] Reliability Test System Task Force "The IEEE reliability test system 1996". *IEEE Transactions on Power Systems*. Vol. 14 pp. 1010-1020 (1999).

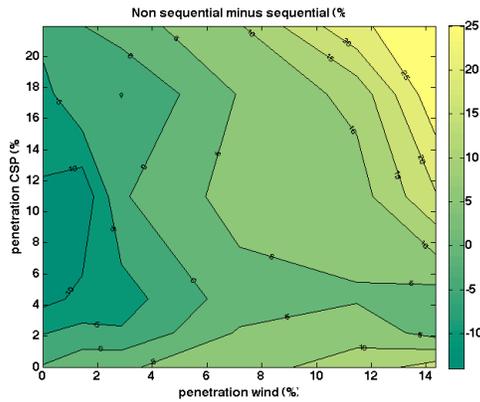


Figure 13. Relative differences between non-sequential and sequential methods (%) for different penetration levels of wind and CSP.

The differences between the LOLE values obtained with a non-sequential method and a sequential method, in percentage relative to the non-sequential method results (taken as reference) are shown in Figure 13. It can be noticed that with a higher proportion of wind the reliability given by the sequential method is higher.

The difference between the non-sequential and the sequential methods can be explained for this reason: CSP generation is, in this case, negatively correlated to demand, and this correlation can be considered only if a sequential method is used. A non-sequential method does not consider the fact that the CSP generation tends to be lower when the demand (in this example) is higher. On the other hand, the correlation between the wind and the demand in this case is positive (see Figure 5. and Figure 12.), so the results with a sequential method yield higher results.

VI. CONCLUSION.

The method of Moving Block Bootstrap generates yearly series of hourly wind power production when there are enough and representative production data for a given area. This method allows to perform sequential Monte Carlo based adequacy studies. The method is simple to understand and implement.

The examined wind power series in peninsular Spain has two main seasonal patterns, daily and yearly. The difference between seasons is also important, but more difficult to model.

After a simulation example assessing the reliability of the IEEE-RTS test power system including wind generation and concentrating solar power, some preliminary conclusions could be extracted.

The non-sequential Monte Carlo method yields more reliability for the CSP generation than the sequential one, but less reliability for the wind generation.

This is due to the temporal correlation between these resources and the demand. The correlation is greater for wind power since it produces at night and mostly in winter, when the demand of the test case is higher.

These conclusions might be different for other load profiles, particularly for load profiles more likely to exist in countries where CSP generation is present.