

Article

A Learning-Based Methodology to Optimally Fit Short-Term Wind-Energy Bands

Claudio Risso ^{1,*}  and Gustavo Guerberoff ² ¹ Computer Science Institute, Engineering Faculty, University of the Republic, Montevideo 11200, Uruguay² IMERL, Engineering Faculty, University of the Republic, Montevideo 11200, Uruguay; gguerber@fing.edu.uy

* Correspondence: crisso@fing.edu.uy

Abstract: The increasing rate of penetration of non-conventional renewable energies is affecting the traditional assumption of controllability over energy sources. Power dispatch scheduling methods need to integrate the intrinsic randomness of some new sources, among which, wind energy is particularly difficult to treat. This work aims at the optimal construction of energy bands around wind energy forecasts. Complementarily, a remarkable fact of the proposed technique is that it can be extended to integrate multiple forecasts into a single one, whose band width is narrower at the same level of confidence. The work is based upon a real-world application case, developed for the Uruguayan Electricity Market, a world leader in the penetration of renewable energies.

Keywords: wind power; non-conventional renewable energy; forecasting; energy bands; combinatorial optimization



Citation: Risso, C.; Guerberoff, G. A Learning-Based Methodology to Optimally Fit Short-Term Wind-Energy Bands. *Appl. Sci.* **2021**, *11*, 5137. <https://doi.org/10.3390/app11115137>

Academic Editor: Luis Hernández-Callejo

Received: 21 April 2021
Accepted: 18 May 2021
Published: 31 May 2021

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



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

1. Introduction

Whether due to economic pressure or environmental concerns, the rate of penetration of non-conventional renewable energies has been increasing rapidly over recent years, and it is expected to grow even faster in the years to come. Short-term operation and maintenance of electrical systems relies on optimal power dispatch scheduling methods.

Either renewable or not, conventional energy sources are dispatchable on request, i.e., authorities can control when and how much power will be provided from each source. Conversely, non-conventional renewable energies are not controllable, are intermittent and uncertain, even within a few hours period ahead. The intrinsic stochastic nature of the new energy sources turns out the short-term dispatch of the grid into a much harder challenge, which necessarily must coexist with randomness coming from significant portions of the installed power plant. This work regards with the optimal crafting of wind-energy bands (in a sense precisely defined in Section 3). It is based on an application case of Uruguay, a worldwide leader in the usage of renewable energies. The Uruguayan case was chosen as reference because: (i) Uruguay is the country of these authors, so the case is very proximate to our research group; (ii) the country counts an immense relative penetration of renewable generation of diverse sources; and (iii) there is plenty of open access information available. Although this document is guided by that reference application case, the methodology is general and its results extendable, so it can be ported to another system or country.

1.1. Literature Overview

In the context of forecast of daily scenarios there exist a vast literature and plenty of methods proposed to accurately predict a likely behavior for the stochastic process involved. Several of them use purely statistics techniques—parametric, semi-parametric and nonparametric—to infer forecasts. Particularly, there are standard and well-known techniques coming from the analysis of time series for the daily forecasting electricity prices and electricity demand (see [1]) that can be adapted to wind power inference. Statistical methods like those, perform well to forecast time series without very strong fluctuations,

like: electricity prices, electrical demand time series, hydraulic contributions to water reservoirs, or even to forecast mid and long-term availability of wind power. When the application regards with short-term wind power generation, the accuracy of such statistical methods degrades notably, even over short periods of time that range from a few hours till two or three days ahead. This is because wind power is much more volatile than electricity prices and electrical demand time series (and the other examples mentioned before).

Complementarily, there are approaches for short-term wind power forecasting based on numerical simulations of atmosphere's wind flows (see [2,3]). For a couple of days ahead period, or even larger time windows, numerical simulations are usually more accurate than purely statistical models. Such models are deterministic, while the underlying physical phenomena is chaotic by nature. So, they perform better than purely statistical methods to follow the process whereabouts at early stages but are far from being trustworthy in what respects to the construction of likely scenarios at larger times. Summarizing: the scheduling of short-term dispatch must coexist with randomness, so, even though wind simulations provide valuable information, they must be enriched in order to account the intrinsic stochasticity of the process.

In the last decade, the concept of prediction interval associated to probabilistic forecasts was introduced (see [4]). The construction is based on a nonparametric approach to estimate—at once, for all instant of time within a selected grid over a forecast horizon—the interdependent quantiles for the (unknown) distribution probability of the wind power process. For each time, the corresponding prediction interval provides an estimate of the expected accuracy of predictions with respect to what the actual value of the wind power will be. Therefore, this technique crafts bands inside which the wind power process is expected to stay with a given probability (the nominal coverage rate [5,6]). Complementing the previous line of work, in [7] other methods that also use deterministic forecasts as input are introduced, and through a subtle analysis that involves historical data to estimate nonparametric forecast error densities, for some relevant times chosen appropriately, the authors have succeeded in generating wind power scenarios with their respective probabilities.

Referred to the construction of confidence bands, several papers have been issued. In other works (see [8]), parametric processes guided by stochastic differential equations (SDEs) are studied. These processes involve a drift term, which acts as a force that tends to attract the trajectories towards the forecast (which is known as an input), and also involve a diffusive term modulated by a Wiener process factor as usual. Different parametric models are studied by considering specific forms for the drift and diffusion terms. The basic idea of the mentioned work consists in approximate these parametric non-Gaussian stochastic processes by Gaussian processes with the same mean, variance and covariance structure. Such approximations allow the estimation of the parameters through maximum likelihood techniques. Following the ideas of these authors and using the same data set for wind power in Uruguay as in the present work, the article [9] synthesizes multiple realizations of the calibrated process in order to build confidence bands. As a subsequent and tightly coupled step towards tackling the crafting of stochastic optimal short-term dispatches schedules of the grid, the previous research team uses those SDEs as supply for a Continuous-Time Stochastic Optimal Control (CTSOC) [10], with very promising results. However, scalable numerical techniques to solve optimal control problems derive from dynamical programming [11], what limits the number of state variables to integrate to the problem. The reference [10] is a pretty good example of what can be done when the number of states is manageable, but many times, the intrinsic structure of the problem makes it untreatable through such approaches. Some practical applications inevitably require crafting scenarios to use other optimization techniques, where the availability of wind-energy bands is essential.

Energy bands crafted in this article were used as a supply for an optimal short-term dispatch problem for the reference application case, which combines generation units with complex commitments, temporal dependencies among them, and other intrinsic

characteristics that makes the problem too hard to be tackled when approached with dynamical programming or related techniques. We suggest [12] as an illustrative reference for practical applications of this work.

In recent years a considerable effort was put into controlling strong variability of weather conditions through the incorporation of what is called in the literature Ensemble Weather Forecasting (see for instance [13], focused on wind-power predictions in Japan); a strategy significantly different from that present in Section 4 of this paper. Methodologies aside, objectives are in fact quite similar: to control rare events, in that case through numerical studies over an area-averaged, added to a rigorous probabilistic analysis. Of course, the particularities of the Uruguayan case are notably different to the Japanese case due to the frequent existence in the eastern Pacific region of strong climatic anomalies as the passing of extratropical cyclones, that the authors called wind ramp events. Their probabilistic wind-power prediction achieved a good statistical reliability through confidence interval for the wind-power variability. Numerical Weather Predictions and Ensemble Weather Forecasting are also used in [14].

Finally, researchers from South Korea introduce an interesting ensemble [15] through different machine learning techniques, combining multilayer perceptron (MLP), support vector regression (SVR), and CatBoost to improve power forecasting of renewable sources. As we see later on, the article here presented focuses upon wind energy rather than power forecasts, and in fact, uses power forecasts as a supply. Furthermore, instead of using existing learning techniques as in [15], this work introduces a novel one, conceptually simpler, and yet promising according on its results.

1.2. Particulars of the Reference Application Case

As we previously mentioned, the methodology elaborated in this work is general and it can be applied to other cases. However, practical relevance of these results relies upon the magnitude and diversity of renewable energy sources in the particular application. Uruguay is one the countries of the world with highest penetration of renewable energies. Nowadays, over 98% of the annual energy consumed by the country or exported to neighbors (i.e., Argentina and Brazil) comes from renewable sources. Table 1 presents the installed power plant and annual generation by type of energy source (Information regarding capacity and energy is available at: <https://portal.ute.com.uy/composicion-energetica-y-potencias> while geographic distribution is in: <https://www.ute.com.uy/institucional/infraestructura/fuentes-de-generacion> Other historical or real-time data is available at: <https://www.adme.com.uy> and <https://adme.com.uy/mmee/infannual.php>. A few years ago, when data of this work was acquired, that fraction was 96%, slightly lower but still remarkable high.

Table 1. Details of the installed power plant and energy by type of source [ADME: 2019, UTE: 2019].

Energy by Type of Source	Installed Power Plant (MW)	Relative Subtotal	Produced Energy Total 2019 (GWh)	Relative Subtotal
Hydroelectric	1534	31.5%	6134	55.6%
Wind-power	1506	30.9%	3690	33.5%
Solar-Photovoltaic	254	5.2%	314	2.8%
Biomass (thermal)	413	8.5%	660	6.0%
Fossile (thermal)	1170	24.0%	225	2.0%

Figure 1 on the other hand shows the geographical distribution of renewable units.

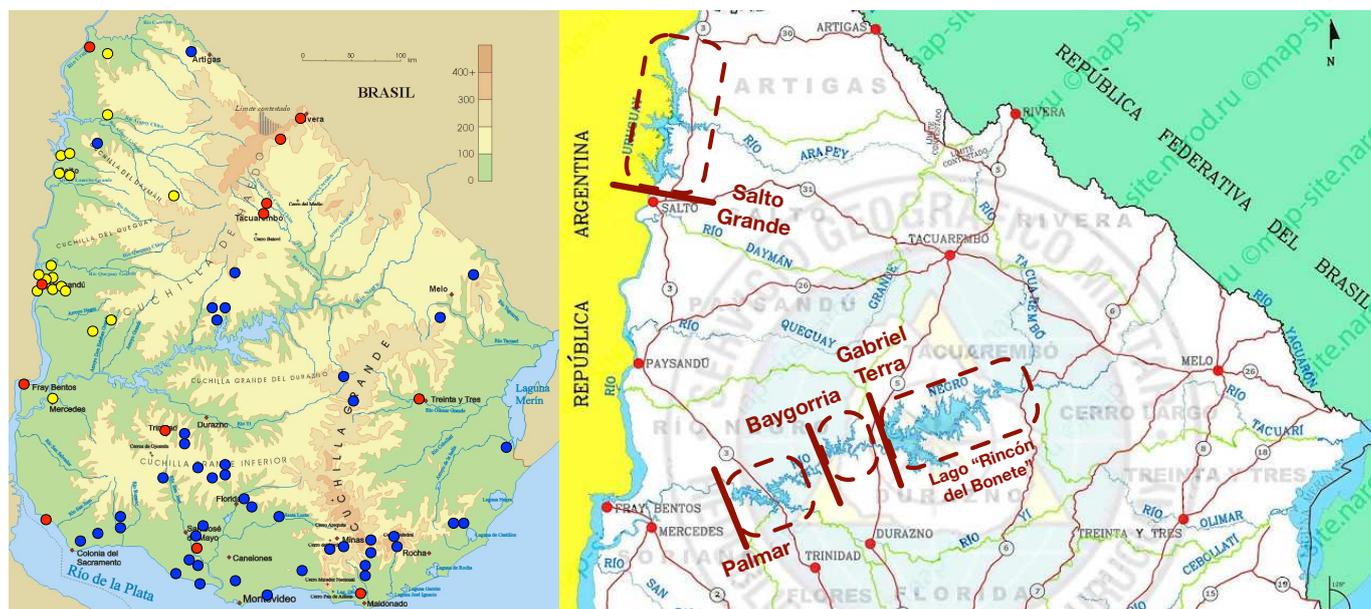


Figure 1. Geographical distribution of renewable sources. Leftmost: wind-power (blue), solar-photovoltaic (yellow) and biomass thermal units (red). Rightmost: hydroelectric dams. [source UTE: 2019 and [12]].

The information previously presented is public and accessible through the provided URLs. Detailed historic information about actual wind-power and related forecasts was also public by the time this work was realized, but unfortunately is no longer so.

1.3. Main Goals of This Work

Regardless of the technique used to narrow uncertainty with bands, those works mentioned in Section 1.1 share a common characteristic: the electric-power is the magnitude to be captured. For some applications and/or contexts, the *energy* coming from wind sources—which is a derivative of the wind's power anyway—is an important magnitude itself, and since it is the outcome of integrating the previous, results less noisy and easy to capture. This work aims on crafting bands representative of the energy to be produced along some periods, rather than focusing on accurate measures of the instant power. It is worth mentioning that since both magnitudes are strictly dependent, a fitted power-process cannot be outside energy-bands for too long. Thus, though this kind of bands could resemble the other in their shape, strictly speaking, they are different.

Regarding the process to craft such wind-energy bands, the idea goes by designing a learning model that is fed from: wind-power forecasts as a mainstream of what to expect for the days to come, and actual wind-power data to incorporate information related to the historical deviations of the process. The area inside the band is a measure of the quality of such calibration. The smaller the area, the better the quality of the fitting. So, crafting energy-bands of minimal area is our objective. Assuming persistent behavior, an optimal fitting over a training set is expected to replicate reasonably well over other independent instance, so a historical calibration could be used to estimate energy-bands in the near future.

Summarizing, our interest is not focused on particular power trajectories and their probabilities but aims on crafting optimal area energy-bands around wind-power forecasts, in such a way that the energy outside those bands be bounded. At this respect, our article presents a novel approach. The method is purely nonparametric, since it makes no assumptions on the physical phenomena, nor any hypothesis about the involved random processes (conditions on homogeneity in time, seasonal behavior of the temporal series or any kind of markovian hypotheses, are not necessary for this work). Our proposal is based

upon a mixed integer optimization problem, which aims on getting the narrowest average band around a set of forecasts, or a combination of forecasts, that keeps the off-band aggregated energy below a given threshold. As an innovation, to allow the optimization to go as far as possible, the model enables to discard up to a given percentage of training samples that are treated as *atypical profiles*.

At first glance this last feature might look risky, in the sense that, in advance, one cannot tell whether the day to come will match or not an atypical profile. However, as we see later, when the method is trained with more than one forecast, whenever they are conditionally independent or weakly dependent, a combination of those forecasts and their bands allows to regain the lost confidence; the result is a more accurate band than those of constructions computed by separate. This is another remarkable point of the present work, since it helps to improve band's quality by taking the best of more than one forecasts provider.

1.4. Structure of This Document

This work uses data coming from two independent wind-power forecast providers for the Uruguayan Electricity Market: Garrad Hassan and Meteorológica, which was available during the period. Complementarily, a third and purely probabilistic forecast was constructed up from historical wind-power realizations, by closely following other documented ideas (see [16]). Regarding actual wind-process realizations, we also have used power records measured over the Uruguayan grid. Therefore, the experimental evaluation here presented for training and test sets is based upon real-world data.

The structure of this article matches the stages of the novel technique. In Section 2 we describe the main characteristics of wind-power in Uruguay, together with the forecasts used as supplies for computations and analysis. Section 3 presents the optimization model to create likely bands of minimum width as well as results from experimental evaluation; while Section 4 shows how after filtering atypical days, a combination of bands calibrated up from independent forecasts performs better than any of them by separate. Finally, Section 5 summarizes the main results and possible applications.

2. Wind-Power Uncertainty and the Use of Forecasts

This section shows how variable wind-power is -when described as a stochastic process—and it presents the forecasts that are used to anticipate power realizations. The historical of wind-power data in Uruguay has a few years and along this period the installed power plant has been firmly growing. So, instead of expressing power in term of MW, we use the Plant Load Factor (PLF), which corresponds to the actual power generated at each time divided by the sum of the installed power capacity of each wind turbine in the system at that moment (i.e., the wind-power plant). Therefore, PLF is a dimensionless quantity that takes values between 0 and 1. Over an hour time-slot basis we consider the average PLF, i.e., the average power along each hour divided by the power plant; this is the main variable we use along this work. In this way the information is normalized, and we can disregard of changes in the installed capacity during the period of analysis. In what follows we present a summary of the behavior of wind-power in Uruguay. The data sample used involves around 730 days of the years 2014 to 2016. We decided to use this period as reference because of the homogeneity of forecast providers and the availability of open data sources.

2.1. Seasonal Regularity

Figure 2 shows the average energy and relative deviations of the daily PLF with respect to the historical mean. Daily PLF is the cumulated value of hourly PLFs over a day, so it ranges from 0 to 24. The temporal horizon varies from 1 day to 730 days ahead.

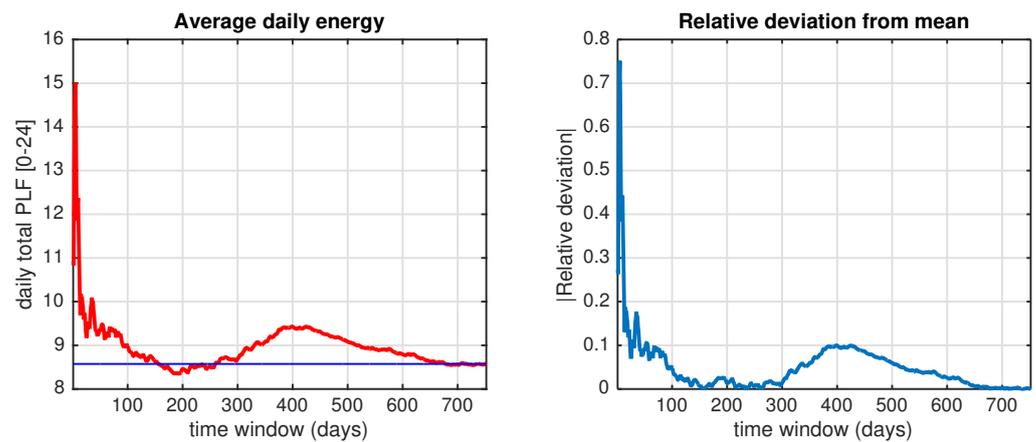


Figure 2. Total and relative deviations of actual PLF with respect to the historical average (leftmost horizontal blue line) as a function of the horizon (730 days).

The intermediate rebound in the deviation is due to a seasonal phenomenon; this effect decreases considerably if the records are limited to a single season. For instance, Figure 3 shows the equivalent plot when only summer days of the same period are considered.

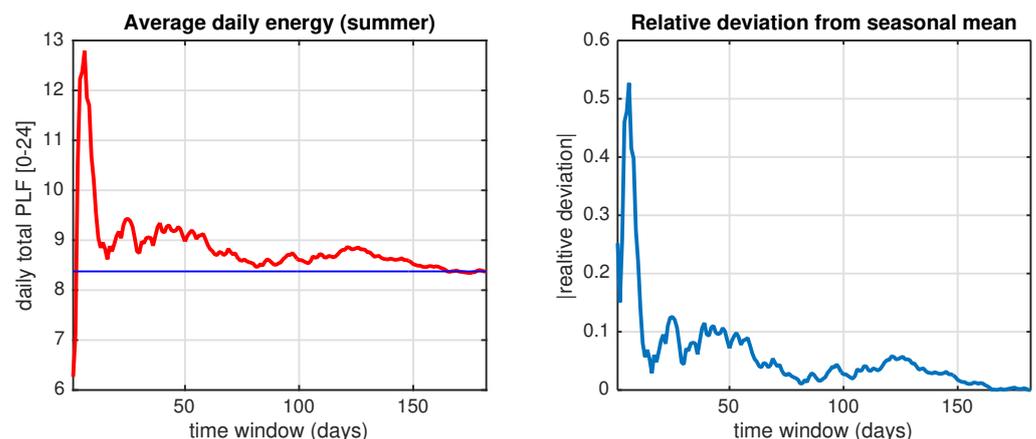


Figure 3. Total and relative deviations of actual PLF with respect to the historical average for summer season (leftmost horizontal blue line).

The previous figure shows how after a week or two, the process goes inside the 10% error band, respect to the average value for that season. We must conclude that wind-power is fairly regular when used in medium-term planning. For shorter periods of time, the situation is quite the opposite.

2.2. Changing Daily Behavior

Managing the electric grid of a country is a challenging task that must be carried out carefully and optimally. To accomplish that, multiple problems are to be solved, spanning different scales of time and components. Medium-term planning usually refers to the valuation of intangible resources, such as the height of the lake in an electric dam accounted as an economic asset. Seasonal regularity as that observed in Section 2.1 is an advisable characteristic to develop mid-term optimization planning models. Short-term planning consists in crafting optimal dispatch schedules for some days ahead, and its aim is upon efficiently coordinate the usage of available resources. The object of this work is on suppling energy-bands for the short-term power dispatch of the Uruguayan grid, whose outcome sets the prices of energy in the electricity market. Due to its short scale of time (a few days ahead) and time-step (of an hour each) short-term planning requires accurate PLF estimations hour-by-hour over some days to come.

A histogram of daily cumulated PLFs along the available two years of samples is shown over the leftmost of Figure 4. Observe that many days within the period have daily cumulated PLFs below 5 (which is approximately 20% of the power plant). The number of samples whose cumulated PLFs are above 19 (80% of the power plant) are lesser, and yet there are days where the average wind-power was pretty close to the power plant. To get an insight about hourly behavior, the rightmost of Figure 4 shows the hour-by-hour mean marked with asterisks. The mean PLF is around 20% higher in the night hours compared with the hours of sun. Complementarily, the rightmost image plots actual wind-power samples, concretely those 30% with higher distance ($\|\cdot\|_2$) respect to the average over the first year. Observe how divergent are these samples when compared with the mean trajectory. We are not going further in the direction of standard statistical descriptive, since it exposes the predictions of wind-power to important errors and thus is seldom used.

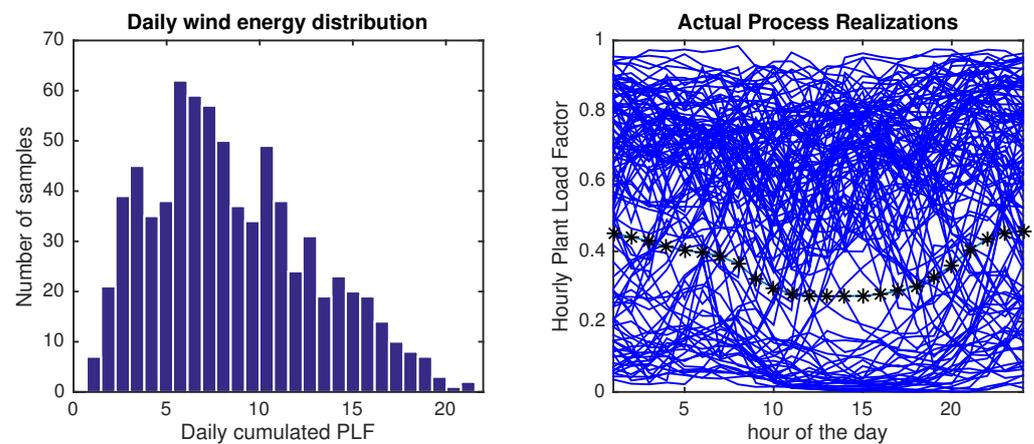


Figure 4. Histogram of daily wind energy samples [daily PLF upon the leftmost] and 30% most atypical realizations for Uruguayan wind-power over a year [rightmost].

2.3. Additional Accuracy Coming from Numerical Forecasts

This section experimentally analyzes the benefits of using short-term wind-power forecastings based on numerical simulations of atmosphere's wind flows. As a measure of the error incurred we use the $\|\cdot\|_1$ distance between forecasts and actual realizations. Hence, the total energy error for the period starting on day d (denoted err^d) is computed as: $err^d = \sum_{t=1}^T |w_t^d - p_t^d|$, being w_t^d and p_t^d , respectively, the actual power (PLF) for the day d as seen t hours ahead and its corresponding forecasted value. On the other hand, T is the time horizon of the forecasts: 72 h in our examples. In Section 3 we explain the convenience of using such an error measure.

This work used information for the Uruguayan grid, which was of public domain by the time computations were realized. Two independent forecast providers are considered: Garrad Hassan and Meteológica. Their common samples span around 300 days, starting in early 2016. A third forecast—about which we elaborate later on—that uses a purely statistics analysis (following PSF ideas [16]) is built to benchmark statistical and numerical forecast. In spite of its lower performance as an isolated technique, we see in Section 4 that this final forecast (referred to as PSF44) increases the overall quality of a convex combinations of filtered forecasts.

Regarding statistical moments of the series, the mean value of the PLF error samples is 6.80 for Garrad Hassan and 5.99 for Meteológica, with respective variances of values 2.48 and 2.21. On the other hand, those figures for the PSF44 are: 13.99 and 4.52; notoriously worse than numerical forecasts. Complementarily, histograms in Figure 5 represent error distributions for each case, reinforcing the idea that Meteológica's forecast slightly outperforms Garrad Hassan's, while both are much better than PSF44.

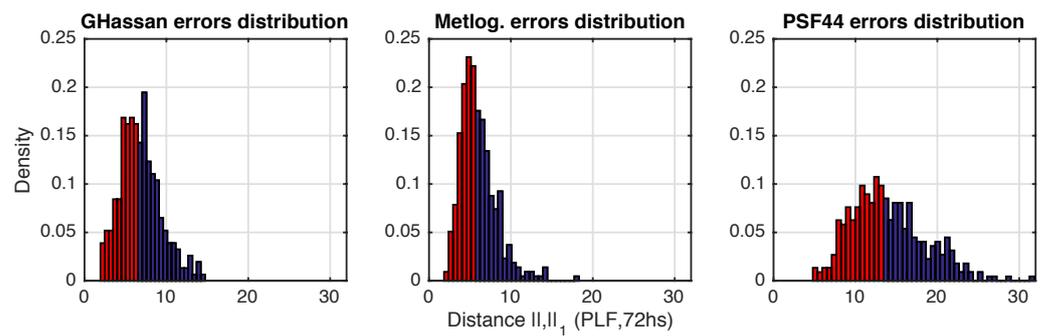


Figure 5. Histograms for total deviations of forecasts within a horizon of 72 h ahead. Red and blue areas concentrate 50% of probability in all cases [leftmost: Garrad Hassan, center: Meteorológica, rightmost: PSF44].

2.4. PSF-Like Forecast

Pattern Sequence-based Forecasting algorithm (PSF) is a novel nonparametric approach to infer forecasts. This method has provided promising results when applied to an assortment of time series forecasting in several international markets, at a horizon of one or a few days ahead. The main idea of the PSF algorithm -and more recent variants- involves three parts working sequentially:

1. The historical of time series is clustered into groups that have similar temporal daily profiles;
2. The time series of daily profiles is converted into a discrete sequence of labels for the cluster each day belongs to;
3. Given the label corresponding to the current day, a window of days of fixed length is used to seek along the whole past sequence of labels for those that match the window ending at present time. The profile for the day to be predicted is constructed averaging profiles that follow to each one of those identical windows in the past.

A remarkable advantage of the PSF method is its reduced number of parameters. There are only two main parameters to adjust: the number of clusters K and the historical time window W .

Figure 6 shows the centroids computed over the actual wind-power data-set when using different numbers of clusters. For example purposes, assume $K = 3$ and $W = 4$; so, a sequence of D days translates into a sequence $\{s_d\}$ ($d \in D$) of digits, with $s_d \in \{1, 2, 3\}$. Here, s_d is the index of the closest centroid to the realization of the day d .

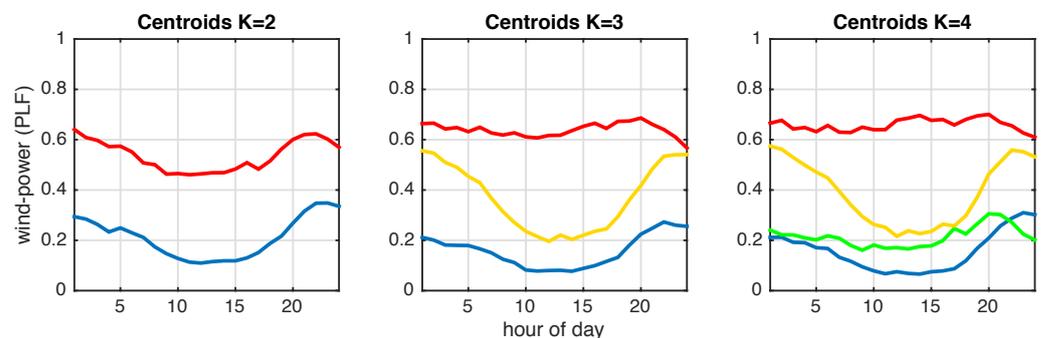


Figure 6. Centroids of actual wind-power samples for different number of clusters.

Within such sequence, we aim now on finding subsequences with $W = 4$ symbols, for instance, the sequence 1231 (we are assuming that current day belongs to cluster 1, given by the last symbol in this sequence). Figure 7 shows examples where that subsequence could be found. The outcome of this predictor (one day ahead) is the average of the actual PLF realizations, among those samples immediately following the subsequences registered. To extend the construction to a two-days ahead forecast, one could repeat the

process seeking for the subsequence composed of the previous $W - 1$ days plus the new forecasted one, and so on.

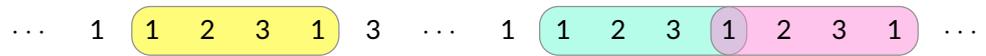


Figure 7. Example of matching subsequences within a historical of symbols.

The purpose behind the development of this forecast is not devaluating statistical methods. On the contrary, this work shows an example of how such a simple method, may contribute to the overall quality when combined with forecasts coming from complementary techniques.

3. Optimization of Wind-Energy Bands

Optimization of wind-energy bands is in the core of this framework. We provide an expression to compute a band around any forecast, and, for that concrete formula, we seek for the narrowest band that satisfies a set of constraints, which imposes limits to the actual process in its deviations in accordance with the historical behavior. A traditional approach would go the way of setting constraints to keep the power deviation under certain boundaries. Conversely, this work aims on minimizing the expected off-band energy.

The previous is explained by the particulars of the Uruguayan electricity installed plant, but it is also justified by trends of new technologies. Around 98% of the electricity annually consumed in this country comes from renewable sources (see [17]). In average, 50% is from hydroelectric sources, while 35% is from wind-power. All of the hydroelectric dams in Uruguayan have water reservoirs; two of them (Bonete and Salto Grande, see Figure 1) are particularly huge. Almost 40% of the hydroelectric capacity is located after the greater lake (Bonete’s), which would take 5 months to empty at full-power. Therefore, in fact, the hydroelectric plant also constitutes an accumulator, i.e., a kind of battery that can plenty compensate short-term fluctuations of the power coming from wind turbines. Hence, regarding Uruguayan short-term planning concerns, an accurate prediction of energy boundaries is more convenient than a power forecast of limited punctual quality. In a complementary manner, smart-grids capabilities are rapidly advancing towards active applications, capable of dynamically adjusting portions of the demand to adapt them to fit system needs (see [18]), while electricity storage units based on batteries are just around the corner (read [19] and also see the “Neoen & Tesla Motors” project in Australia). Therefore, in the near future, this work could be a useful experience for other countries.

The information required to determine an instance of our optimization problem comprises the following data sets. At first place, we need a historical of wind-power forecasts. We consider a collection \mathcal{P} of deterministic registers that involves short-term point forecasts over a horizon of a few days ahead; i.e., a family of vectors $p^d \in [0, 1]^T$, with fixed T , which is set by the number of samples along the time horizon. Here, d is the index for each day on which the construction of a band begins; $d \in D$, being D the set of indices for days with historical observations. Wind-power forecasts usually span from one up to three days, i.e., from 24 to 72 h, and time is discretized at a rate of one sample per hour. Let $T - 1$ be the limit of hours ahead available for each forecast. We assume that all forecasts share the same time horizon, and that in $t = 0$ the current power is the only data known for sure. As we mentioned earlier, for simplicity the wind-power is expressed as the PLF, which corresponds to the actual power generated divided by the sum of the installed power capacity of wind turbines in the system at each moment. Thus $p_t^d \in [0, 1]$ is the normalized point forecast of the wind-power t hours ahead, within the vector associated to the forecast issued on the day d .

The second part of the input data set comprises the actual historical wind-power time series samples, grouped into a collection: \mathcal{W} , whose elements $w^d \in [0, 1]^T$ are also assumed normalized. Hence, $w_t^d \in [0, 1]$ is the actual PLF measured t hours after the beginning of the day d . For consistency, since the current state can be measured rather than forecasted, $p_0^d = w_0^d$ for each day d . Observe that the set \mathcal{W} usually has duplicated records, for instance:

$w_{24}^d = w_0^{d+1}$. Despite that, we have chosen this format to simplify those expressions that link with forecast information. Regarding forecasts, however, the previous equality doesn't hold. In fact, p_{24}^d (a sample, forecasted 24 h ahead) is different from p_0^{d+1} (the actual value measured a day later).

It is clear that, given any two bands containing the real process inside of them at the same instants, the narrower band is of better quality. Wind-power generation is a process hard to anticipate, and violations to computed bands is a fact we must coexist with. However, not every violation has the same severity in terms of its impact to the power grid. In the context of the short-term energy dispatch, how much cumulated energy falls down outside the band is a convenient metric to assess the confidence of the pair: forecast plus computed band. In this work, we define the following expression as a metric for the reliability (The expression on the right hand side corresponds precisely to the anti-reliability, which of course is the complement of the reliability; hence the notation for the left hand side) \mathcal{R}^d , of a band around a given forecast p^d :

$$1 - \mathcal{R}^d(w, lb, ub) = \frac{1}{T} \sum_{t=0}^{T-1} \max[w(t) - ub(p^d, t), 0] + \frac{1}{T} \sum_{t=0}^{T-1} \max[lb(p^d, t) - w(t), 0] \quad (1)$$

where ub and lb respectively are the functions that determine upper and lower limits for the bands along the forecasted period, and w is the actual generation, unknown until the near future where reality is revealed. Functions lb and ub take a forecast (p^d) and an instant (t) as their inputs, while their outputs are the respective bounds to expect.

As mentioned, the feasible region of the optimization model imposes limits to the severity of violations to the band. Besides, in order to improve the quality as further as possible, the model allows to discard up to a limit of elements in the training set, which are *atypical*, specially bad forecasts that whether included would either: deteriorate the accuracy of the result, or force us to use too broad bands. So, to complete an instance we must set values to those quantities. The parameter $\theta \in [0, 1]$ limits the amount of energy allowed to fall down outside the band along the optimization horizon. The parameter $\lambda \in [0, 1]$ sets a minimum fraction of *regular* (i.e., not atypical) forecasts to be used in the effective training set or, in other terms, $(1 - \lambda)$ is the maximum fraction of atypical days allowed to be discarded. It is worth mentioning that the limit for off-band energy only accounts over regular forecasts.

3.1. Minimal Relative Width of Bands

This work considers those bands defined by relative deviations with respect to forecasted values, which are simple to calculate and optimize, and yet lead to accurate results. Let $\{x_t \geq 0\}$ be a set of coefficients associated to the time series analyzed, which delimits the width of the band. That is, for any instant t within the time horizon of the forecast issued on day d , we take p_t^d and compute the lower and upper limits of the band using the expressions $lb_t^d = \max[0, (1 - x_t)p_t^d]$ and $ub_t^d = \min[1, (1 + x_t)p_t^d]$ respectively. Hence, $\{x_t : 0 \leq t \leq T - 1\}$ comprises the first set of control variables that modulates the relative width of the band for a given forecast $p^d \in \mathcal{P}$. Figure 8 sketches about how these variables and derivatives are related, through a hypothetical forecast (centroid of the band, highlighted in blue), its correspondent energy-band (shaded in grey), and the actual power-process registered afterwards (red curve).

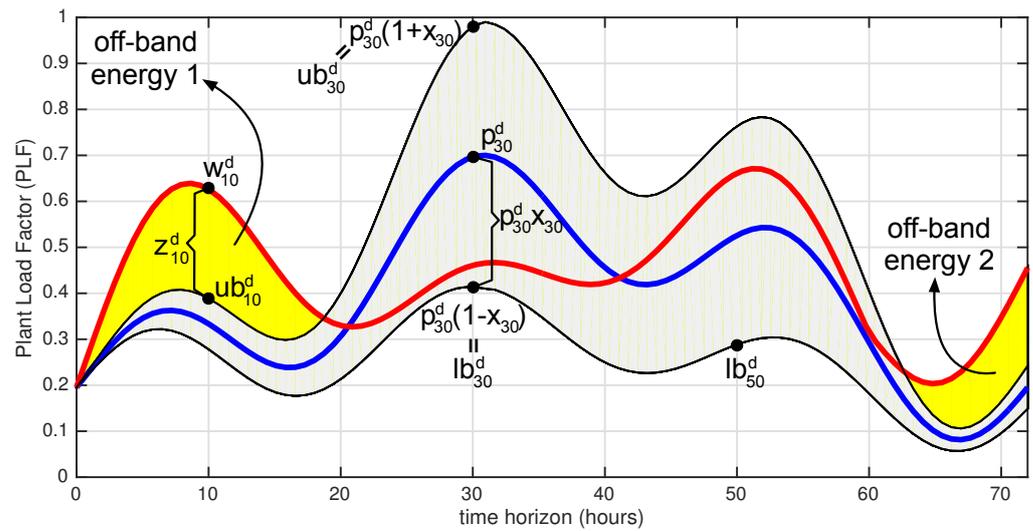


Figure 8. A wind-power band (grey) crafted after a forecast (blue), for some day d , and the actual process (red).

The objective function of this optimization is $\sum_{t=0}^{T-1} \hat{w}_t x_t$, where $\hat{w}_t = (\sum_{d \in \bar{D}} w_t^d) / |\bar{D}|$ is the average PLF at time t over a historical record of observations \bar{D} , eventually different from that of the training set. In other words, \hat{w}_t corresponds to the sequence of asterisks in the rightmost of Figure 4, while $\sum_{t=0}^{T-1} \hat{w}_t x_t$ matches the average grey area in Figure 8. Whenever forecasts are statistically reliable, the objective function corresponds with the expected absolute PLF area of the band along the period T .

Defined so, the optimization is not instant-to-instant greedy, in the sense that it could deteriorate the performance at some points in order to surpass the overall performance by gaining more in others. That differentiates this work from related ones (like [7]), whose intention is to track power rather than energy. In fact, this model doesn't need conventional hypotheses about stochastic processes, such as homogeneity or markovianity.

The second group of control variables is composed by those who determine which are the regular forecasts. The variable $y_d \in \{0, 1\}$ indicates whether the forecast issued on day d should be considered regular ($y_d = 1$), or atypical ($y_d = 0$). Unlike the $\{x_t\}$ variables, these new ones are boolean. We denote D to the set of days for which the optimization problem is implemented (i.e., the training-set). The complete combinatorial optimization model is that in (2).

$$\left\{ \begin{array}{l} \min \sum_{t=0}^{T-1} \hat{w}_t x_t \\ p_t^d x_t - y_d + z_t^d \geq |w_t^d - p_t^d| - 1, 0 \leq t \leq T-1, d \in D, \quad (i) \\ \sum_{t=0}^{T-1} z_t^d \leq T(\theta + 1 - y_d), d \in D, \quad (ii) \\ \sum_{d \in D} y_d \geq \lambda D, \quad (iii) \\ y_d \in \{0, 1\}, 0 \leq x_t, 0 \leq z_t^d \leq 1. \end{array} \right. \quad (2)$$

The auxiliary variables (z_t^d) account by how much power the process (w_t^d) violates the band around the forecast (p_t^d), either at the top or the bottom, for those days classified as regular (i.e., when $y_d = 1$). For instance, if $y_d = 1$ and $w_t^d \geq p_t^d(1 + x_t)$, then it must be held $z_t^d \geq w_t^d - p_t^d(1 + x_t) \geq 0$ to satisfy equation (i) in (2) for that day d at time t . When $y_d = 1$ and $w_t^d \leq p_t^d(1 - x_t)$, then z_t^d should verify $z_t^d \geq p_t^d(1 - x_t) - w_t^d \geq 0$ to satisfy equation (i). For a graphical reference about both situations, please see Figure 8. The optimization process pushes down the z_t^d values, which ultimately are to be set to

$\max [0, w_t^d - p_t^d(1 + x_t), p_t^d(1 - x_t) - w_t^d]$, the anti-reliability of (1). That equation is always satisfied when $y_d = 0$ simply by choosing $z_t^d = 0$ for every t ; therefore, atypical days are disregarded for violations.

Given any day d , when $y_d = 1$ (an effective day of the training set), the second equation guarantees that the time-normalized cumulated off-band energy along the forecasted period T is below θ . That is, in terms of the reliability: $1 - \mathcal{R}^d = (\sum_{t=0}^{T-1} z_t^d) / T \leq \theta$; so θ bounds the energy that lies outside the band to a fraction of the installed power plant. As it happens with (i), equation (ii) is automatically satisfied when $y_d = 0$. Coming back to Figure 8 as a reference instance, by combining equations (i) and (ii) inside an optimization process, we are forcing the total off-band energy (the result of adding up both yellow areas) to be under a desired threshold. Finally, equation (iii) forces the problem to select at least λD days to be regular, which combined with the persistence hypothesis conveys likelihood to the result.

3.2. Experimental Evaluation

The experimental evaluation of this work is based upon a later open data from the Uruguayan Electricity Market. From that past repository, we chose two independent forecast sources: Garrad Hassan and Meteológica. The data were pre-processed using a power assimilation methodology, which fits forecasts along the first 6 h in order to match the starting state (w_0^d). The exact process is described in paper [3]. The used forecasts from Garrad Hassan were those issued at 1AM between 5 April 2016, and 10 March 2017. Within this period there are 302 days where both, forecast and actual data, are complete. Regarding the other provider (Meteológica), the number of complete records is 394, with dates of issue ranging from 1 January 2016, to 10 March 2017. Regarding our own forecast (PSF44), synthesized up from a series of actual power registers, we used the same 730 days of between years 2014 to 2016 that were used upon the first part of Section 2.1. Best performance was found by using $K = 4$ and $W = 4$ (acronym PSF44 refers to those parameters).

Throughout this work, we relied upon IBM(R) ILOG(R) CPLEX(R) Interactive Optimizer 12.6.3 as the optimization solver. The server was an HP ProLiant DL385 G7, with 24 AMD Opteron(tm) Processor 6172 with 64 GB of RAM. After running model (2) over a training set comprising around 30% of Meteológica's days, we find bands like those sketched in Figure 9.

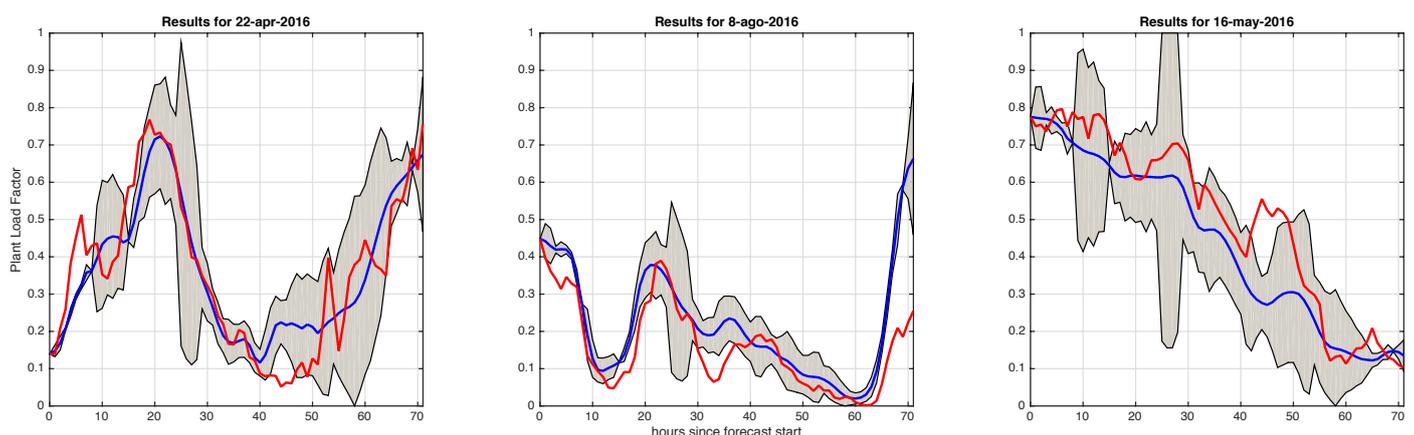


Figure 9. Wind-energy bands for three random days within Meteológica's training set [$\lambda = 1$, $\theta = 0.05$].

The x-axis represents the number of hours ahead for each forecast, while the y-axis corresponds to the PLF. Blue curves are associated with power forecasts while red ones are the actual values. Finally, the grey area represents the wind-power band for $\theta = 0.05$ and $\lambda = 1$. Since λ equals 1, every day within the training set must be effectively included; that is, $y_d = 1$ for each $d \in D$, so all days are treated as regular. Furthermore, when $\lambda = 1$ then (2) turns out to be a pure linear programming problem, and running times are within

the second. Fixing λ to 1, it is of interest to explore how θ modifies the bands. Figure 10 shows the result over the same training set when $\theta = 0.01$.

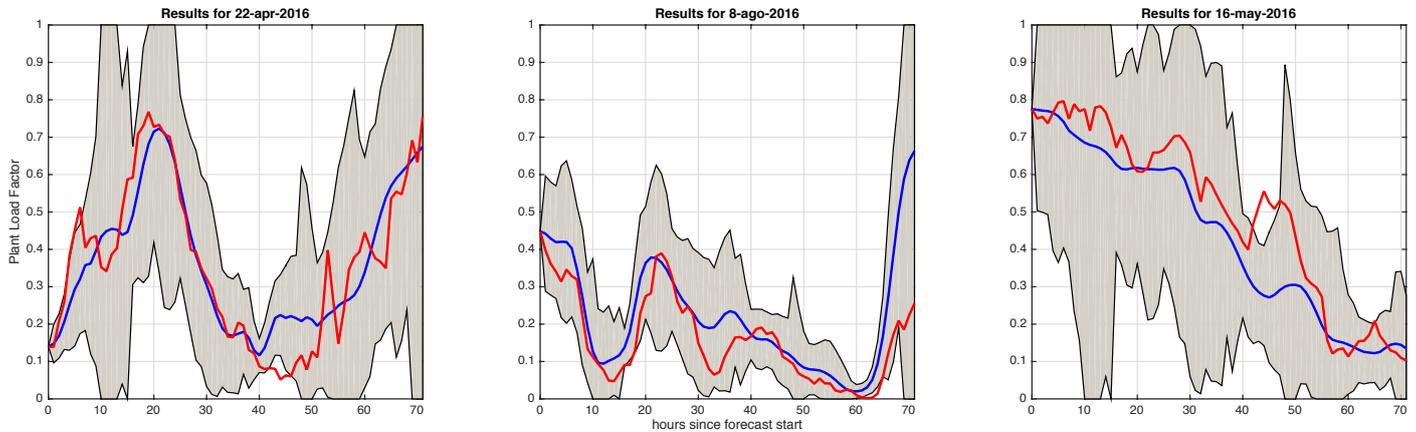


Figure 10. Wind-energy bands for the same days when $\theta = 0.01$ instead of $\theta = 0.05$ [$\lambda = 1$].

Observe that bands in Figure 10 are wider than in Figure 9, which is expected since we are less tolerant respect to how much energy lies outside those bands. In order to balance reliability and thickness, it is of interest to compute how much area do bands cover as we change θ while keeping $\lambda = 1$.

The training in all of the previous cases was performed over a set D of 120 randomly selected days out of a set of 300 days in common for all providers. The common complement, i.e., the set of (180) days shared by these three forecasts and not being in the *training-set*, is used as the *test-set*. The calibration of PSF44 was crafted using the set of 430 contiguous days previous to those of training and test sets. Experimental evaluation (see leftmost of Figure 11) verifies that the average width of the bands, when trained over the entire training set of forecasts ($\lambda = 1$), falls down rapidly to 0, which is reached upon both companies when θ is close to 0.2. Although similar, Meteológica’s bands (blue) are always better than Garrad Hassan’s (red). PSF44 (green) requires much wider bands to achieve the same grades of reliability. The middle plot shows the relative difference between widths of original bands (those of the training set), and widths computed over the test-set using the corresponding x vector found for each θ . It is worth mentioning that widths are always similar (divergence is low), so the objective function in (2) when evaluated over the training set is representative of what happens outside it. This is sustained even for relatively higher θ values right below 0.2, where widths tend to zero and the relative deviation makes no sense to be accounted.

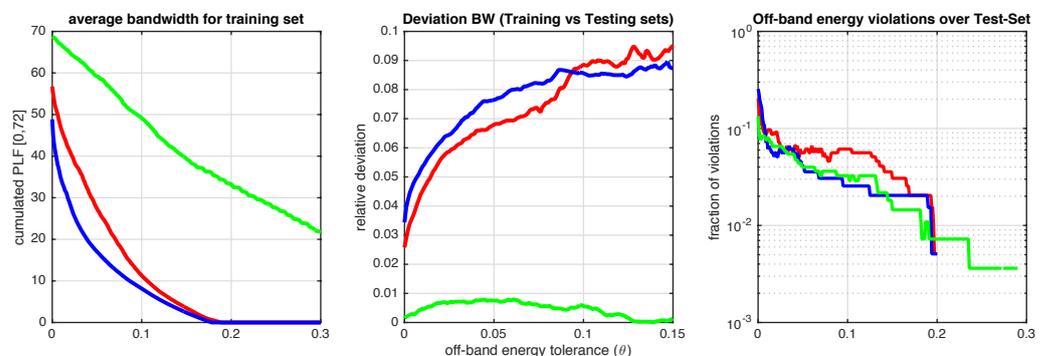


Figure 11. Average width of bands found for the training set as θ changes while λ is fixed in 1 [leftmost], relative deviation between average widths register for training and testing sets [middle], and violations of off-band energy limits over testing set. [red samples correspond to Garrad Hassan, blue ones to Meteológica and green to PSF44].

Regarding off-band energy violations to the limit θ when computed over the test-set (they do not happen in the training-set because of (2)(ii)), the rightmost of Figure 11 shows the fraction of those violations, i.e., the fraction of samples where the off-band energy surpasses $T\theta$. They are also low for all forecasts and are particularly lower as values of θ get apart from zero. The previous exercise experimentally justifies the persistence hypothesis this technique is based on.

The goal of this work is providing stochastic short-term optimal power dispatch schedulers, with accurate wind-energy bands, in the context of the Uruguayan Electricity Market. In particular, our interest is keeping off-band energy below 10% of the average PLF, which is around 0.35; so we consider $\theta = 0.035$. In Uruguay, 35% of electricity comes from wind-power sources, thus that value of θ corresponds to 1.23% of the average energy consumed, what is ambitious. That value is used as reference during this work.

The other parameter to consider is λ which attends to the fact that, whatever accurate a family of forecasts may be, there will always be samples that degrade the overall quality of the whole. Table 2 shows how some attributes of the bands change as λ decreases from 1 to 0.6, while θ remains fixed in 0.035 (our target off-band violation).

Table 2. Experimentally estimated attributes for confidence $\theta = 0.035$ bands as λ decreases.

λ	% Anomalous	\overline{BW}	% \overline{BW}	t (s)	% ANOMALOUS	\overline{BW}	% \overline{BW}	t (s)	% Anomalous	\overline{BW}	% \overline{BW}	t (s)
1.00	6.6%	21.3	29.6%	<1	5.6%	34.3	47.6%	<1	5.4%	62.3	86.5%	<1
0.95	14.8%	16.0	22.2%	117	12.2%	26.8	37.2%	129	8.0%	56.6	78.6%	77
0.90	20.4%	13.8	19.2%	138	16.3%	24.8	34.4%	428	11.2%	51.6	71.7%	66
0.85	27.0%	12.5	17.4%	265	22.5%	22.1	30.7%	272	13.4%	48.4	67.2%	292
0.80	32.1%	11.7	16.3%	664	25.5%	19.7	27.4%	1518	16.3%	45.6	63.3%	421
0.75	36.7%	10.8	15.0%	1296	37.2%	17.4	24.2%	1097	20.7%	41.8	58.1%	657
0.70	45.4%	9.9	13.7%	2020	45.4%	15.6	21.7%	1950	28.3%	38.5	53.5%	842
0.65	49.0%	8.9	12.3%	1225	49.0%	13.1	18.2%	1480	28.6%	35.9	49.9%	1468
0.60	54.6%	8.3	11.5%	1390	52.6%	11.9	16.5%	1560	37.0%	34.2	47.5%	2425
Meteorológica				Garrad Hassan				PSF44				

The first three four columns correspond to Meteorológica forecasts, the second part does to Garrad Hassan's and the last one to PSF44 forecasts. These metrics were computed over the test-set by using optimal x coefficients for each λ over the samples in the training-set. Columns labeled as %*anomalous* indicate the percentage of the samples, in each case, whose off-band energies surpasses the 0.035 of the total plant factor ($T\theta$). We decided to use different adjectives to distinguish between *atypical days*: samples intentionally excluded from the training-set, and *anomalous days*: samples in the testing-set that by chance surpass the off-band energy limit. The columns \overline{BW} and % \overline{BW} respectively show the average absolute and relative areas of the wind-power bands over the test set, using 72 as the full plant factor for the time horizon. Finally, the number of seconds spent by the solver to find the optimal solution for each case appears in the column labeled as $t(s)$. Observe that as λ decreases so it does the expected width for energy bands, because the solver is allowed to select down to λD days during the training, and the optimization ends up by crafting bands for the best subset with a λ fraction of the original number of days. Computation times ascend, because (2) becomes combinatorial for $\lambda < 1$. Conversely, the percentage of *anomalous days* (i.e., those whose off-band energy falls outside the limit) increases, since the calibration performed over a partial/elite training-set is no longer representative over the complement (i.e., the test-set). A second goal of this work is keeping the percentage of *anomalous days* below 10%, which translates into attaining the target θ at least 90% of the times. The final goal is over the allowed variance for wind-power. Until now, we have focused upon energy rather than power. Keeping the process within narrower bands

is equivalent to expect lower power variations. According to official sources, the total electricity produced by Uruguay during 2017 (to meet internal demand plus energy exports to Argentina and Brazil) was of 12,600 GWh. The equivalent hourly average power is 1438 MW. The total wind-power plant by late 2017 was of 1437 MW (the fact these final figures match is just a coincidence). Hence, aiming on having energy bands whose relative width is below 20% is equivalent to expect average power fluctuations (either upwards or downwards the centroid) below 10% of the installed wind-power plant, which matches the average power consumption. In summary, our targets are: $\theta \leq 0.035$, $\% \overline{BW} \leq 20\%$ and $\% \text{anomalous} \leq 10\%$. Observe that no record in Table 2 fulfills all these goals simultaneously. Along Section 4 we see how to deal with that issue.

4. Combining Forecasts

At first sight, we might think that a convex combination of forecasts and their bands would inherit the width of each one, and that we cannot improve bands quality by means of combining them. The only mechanism we have seen that can get narrower bands goes by reducing λ . As a drawback, this also increases the percentage of anomalous days. However, we might regain confidence if anomalous days -for the different forecasts- were somehow independent, since a combination of anomalous situations in all bands would be rarer than in any of them by separate. That’s the idea behind this section. To check the consistency of this idea we analyze how independent anomalous days are, by its correlation matrices. Table 3 recapitulates figures of anomalous days for different values of λ with $\theta = 0.035$.

Table 3. Correlation matrices for anomalous days for different λ 's and $\theta = 0.035$.

	MT	GH	PS		MT	GH	PS		MT	GH	PS
MT	1.0000	0.2621	−0.0178	MT	1.0000	0.2854	0.0147	MT	1.0000	0.2797	−0.0093
GH	0.2621	1.0000	−0.0648	GH	0.2854	1.0000	−0.0720	GH	0.2797	1.0000	−0.0308
PS	−0.0178	−0.0648	1.0000	PS	0.0147	−0.0720	1.0000	PS	−0.0093	−0.0308	1.0000
	$\lambda = 0.60$				$\lambda = 0.65$				$\lambda = 0.70$		
	MT	GH	PS		MT	GH	PS		MT	GH	PS
MT	1.0000	0.2450	0.1127	MT	1.0000	0.2488	0.0581	MT	1.0000	0.2507	0.0109
GH	0.2450	1.0000	0.0817	GH	0.2488	1.0000	0.0810	GH	0.2507	1.0000	0.0929
PS	0.1127	0.0817	1.0000	PS	0.0581	0.0810	1.0000	PS	0.0109	0.0929	1.0000
	$\lambda = 0.75$				$\lambda = 0.80$				$\lambda = 0.85$		
	MT	GH	PS		MT	GH	PS		MT	GH	PS
MT	1.0000	0.3246	0.0262	MT	1.0000	0.2830	−0.0926	MT	1.0000	0.2921	−0.0707
GH	0.3246	1.0000	0.1119	GH	0.2830	1.0000	0.0281	GH	0.2921	1.0000	0.0243
PS	0.0262	0.1119	1.0000	PS	−0.0926	0.0281	1.0000	PS	−0.0707	0.0243	1.0000
	$\lambda = 0.9$				$\lambda = 0.95$				$\lambda = 1.00$		

These numbers were computed over the test-set for Bernoulli’s random variables, $Mt(d)$, $Gt(d)$ and $Pt(d)$, indicator of the event of finding an anomalous day: they evaluate to 1 (respectively 0) if and only if the forecast for day d of the respective corresponding -Metológica, Garrad Hassan and PSF44- classifies as anomalous (resp. regular). From these correlation values, we infer that anomalous days of Metológica and Garrad Hassan are positively but weakly correlated. By running simple simulations with two sets of dependent Bernoulli’s random variables with the same expected value, we observe that the correlations values of the table appeared when 1 out of between 3 to 4 samples of one set copy the value

of the other. PSF44 is basically independent of the others providers, and in fact can either be positively or negatively correlated with them.

Given the three sets of forecasts: \mathcal{P}_{mt} , \mathcal{P}_{gh} and \mathcal{P}_{ps} , and their corresponding functions to compute bands (lower and upper bounds): $bdMT_{\lambda}(p) \rightarrow [0,1]^{T \times 2}$, $bdGH_{\lambda}(p) \rightarrow [0,1]^{T \times 2}$ and $bdPS_{\lambda}(p) \rightarrow [0,1]^{T \times 2}$, we explore convex combinations of them: $bdMX = \alpha \cdot bdMT_{\lambda_1} + \beta \cdot bdGH_{\lambda_2} + (1 - \alpha - \beta) \cdot bdPS_{\lambda_3}$, with $0 \leq \alpha, \beta \leq \alpha + \beta \leq 1$, for different combinations of λ 's. The goal is on finding the combination that is closest to satisfy the targets: $\theta \leq 0.035$, $\%BW \leq 20\%$ and $\%anomalous \leq 10\%$. This second stage of training was performed over the half of the test-set (90 days). The other half remains as the definite test-set to check results. The most convenient combination over the new training-set was found for values: $\alpha = 0.66$, $\beta = 0.25$, $\gamma = 0.09$, $\lambda_1 = 0.85$, $\lambda_2 = 0.70$ and $\lambda_3 = 0.65$. After checking over the now reduced test-set we verify that for $\theta = 0.035$ as limit for off-band energy, 8.9% of the days fall into the anomalous condition, while the value for the average bandwidth is $\%BW = 21.7\%$. This final figure does not attain our original goal (i.e., 20%), but it is pretty close to it.

Performance of Optimally Combined Bands

The lecture of the previos figures indicates that the most performant family of forecasts (Metológica) contributes with 66% of the weight when is calibrated using 85% of its better forecast samples. Despite having similar performance (recall Figures 5 and 11), Garrad Hassan's forecasts only contributes with 25% of the weight, and that is after filtering 30% of its samples. Unexpectedly, being the worst by far, PSF44 contributes with almost 10% to the final result, although after purging 35% of its samples. Probably, the higher weight of PSF44 comes from its almost independence (small correlation) with respect to the other forecasts, rather than its quality.

To analyze the performance of the combined band we present qualitative and quantitative evidence. Figure 12 sketches random bands, its centroid and the corresponding actual power over six days within the test-set. The last two figures (middle and rightmost plots in the bottom row) correspond to two of the eight anomalous days found. Although the off-band energy surpasses the $T\theta$ limit, overall, the performance of those bands doesn't look that bad either.

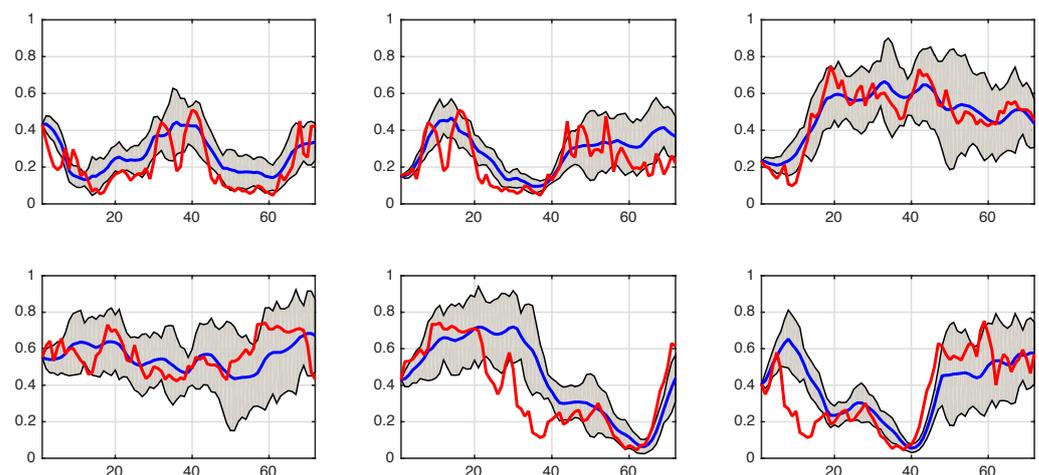


Figure 12. Hybrid bands for six random days in the test-set [blue is the centroid, the red one is the actual power].

Figure 13 shows other group of six random bands. This case does not include anomalous samples, but there are a couple of samples where the area of the band is above the target value. The most notorious case is that on the leftmost of the bottom row.

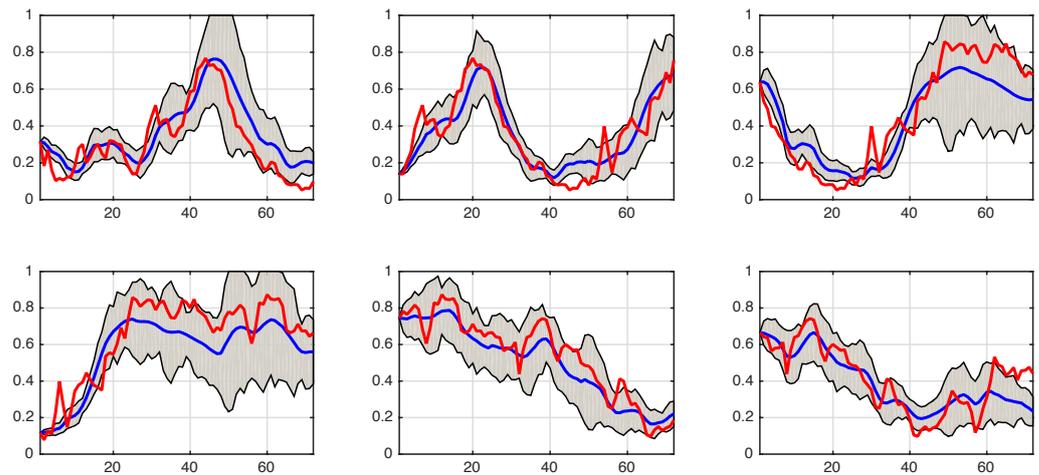


Figure 13. Hybrid bands for six random days in the test-set [blue is the centroid, the red one is the actual power].

It is worth wondering how much energy lies outside the band when violations happen, and how narrow confidence bands are. The residual test-set is so small (90 samples) that, although biased, we decided to use the old one to craft histograms. Figure 14 shows histograms computed up from the original (180 samples) test-set. The leftmost corresponds to the distribution of the off-band energy normalized by the total PLF along the period (72). It is observed that no sample disagrees in more than 7.3%, while in 50% of the samples (those colored with red) that percentage is lower than 1.6%. The rightmost represents the distribution of normalized widths (% \overline{BW}). As in the previous case, samples colored in red add up to 50% and all of them are lower than 14.6%.

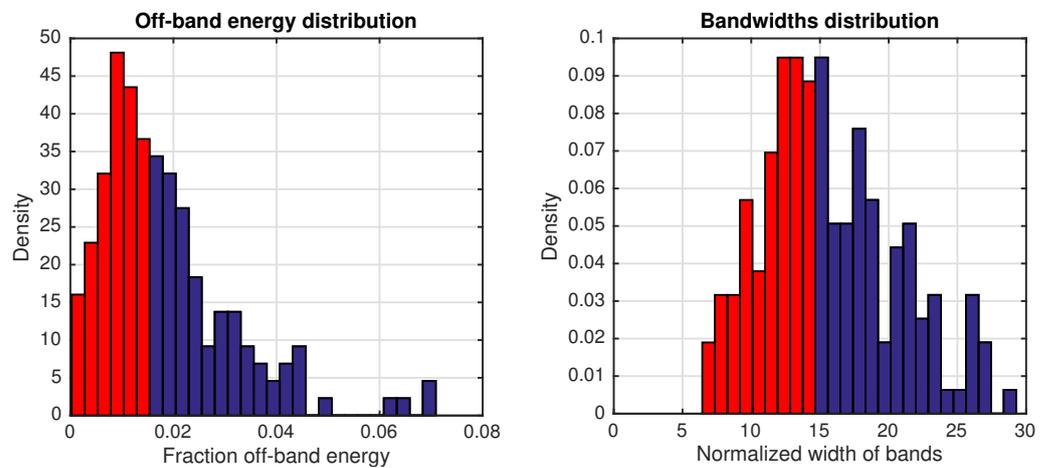


Figure 14. Histograms for relative off-band energy and widths of the bands.

Complementing the previous figure, Figure 15 marks with green the quantiles where the values of either: off-band energy [leftmost] or normalized widths [rightmost] satisfy original targets. The cumulated probability of samples in the first totalizes 90.17%, while those over the rightmost add up to 79.8%. These results reflect the quality of the forecasts and bands computed by this method. We conclude then, that the result is not only satisfactory regarding our initial average performance goals ($\theta \leq 0.035$, { but it is pretty good in terms of the overall quality of the bands and specially in terms of the energy confidence of them.

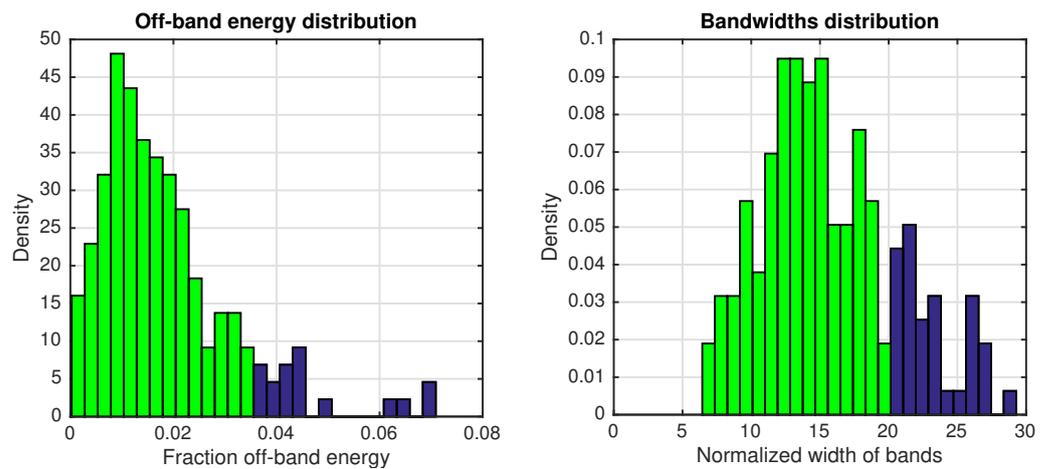


Figure 15. Histograms for relative off-band energy and widths of the bands.

5. Conclusions and Future Work

In this work we introduce a novel learning technique for crafting wind energy bands around forecasts of wind-power generation over a horizon of 72 h ahead. The analysis is based on a historical data set provided by the Uruguayan Electricity Market. The technique allows to discard a portion of atypical days in the training-set, while controls the average cumulated energy that lies outside bands. With an appropriate choice of the parameters involved in the analysis, the model has succeeded in providing bands satisfying natural requirements on confidence and width.

A remarkable conclusion of this work is that the use of an optimal convex combination of conditionally independent (or weakly dependent) forecasts and its corresponding bands improves significantly the performance of the model. For instance, the experimental evaluation of Section 3 suggests that Meteorológica forecasts performs, in average, better than Garrad Hassan's and PSF44. However, an appropriate convex combination of all of them (even when the performance of PSF44 is rather bad) provides better results. While most of the weight of the combination goes to Meteorológica, the inclusion of Garrad Hassan and PSF44 forecasts conveys stability to the result, compensating the fact that some anomalous days for one forecast are regular according to the others. This idea could be extended of course including more forecasts providers.

A drawback of the analysis that we mention here is that the available data set at the moment this work was developed was not too large (around 300 days). Regarding the quality of bands, we expect the performance of the technique will work even better with a more extensive data-set, perhaps spanning a few years. Nevertheless, this size introduces a challenge: increasing the training-set significantly increases computation times. Notice that after adding up computation times reported in Table 2, the total time is above 6 h, which is pretty good for the purposes of these experiments. However, those times are expected to be much higher as the training data-set increases in size, so, in such situations is necessary the introduction of specific algorithms to solve the optimization problem in (2), i.e., not to rely upon standard solvers. A line of future work precisely goes the way of experimenting with other exact methods or derivatives thereof and the exploration of Metaheuristics, in order to find more efficient algorithms to tackle the problem.

Complementarily, the current model uses a single set of x 's variables to delimit bands around forecasts, which results in symmetric widths either upwards or downwards. It is worth testing this hypothesis by including two sets of x 's, on per each direction, and letting the optimization to find solutions over a larger search space. A previous clusterization of forecasts might also improve the performance. Since the training-set indistinguishably comprises both: samples for windy and not-windy days, the relative deviation at a time t necessary to reposition the process within a band shall be greater for forecasts of low prospected energy than the necessary for high energy ones. The previous over-penalizes widths of bands in forecasts with higher expected energy. Training different bands for

different seasons might also improve the quality. Most of these ideas however, require historical data sets much larger than the one currently available.

Regarding the application of bands as those developed in this work. They may be particularly important to craft scenarios in stochastic optimization problems where the complexity of state variables does not allow using other techniques, such as dynamical programming. Examples of such situations arise from a combination of: generation units with complex commitments (limit to minimum power, a slow starting/stopping process, a minimum uptime operation once started); temporal correlation between generation units (e.g., dam water reservoirs where water influxes come from another hydroelectric dam); control deferrable consumption (e.g., electrical residential water heating that must be accounted within certain time windows); large scale energy storage to be later returned to the grid; among others. That results in a wide spectrum of potential application cases.

Author Contributions: Both authors have contribute equally. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: None.

Informed Consent Statement: None.

Data Availability Statement: None.

Acknowledgments: This work was partially supported by PEDECIBA-Informática and PEDECIBA-Matemática (Uruguay), by the STIC-AMSUD project 15STIC-07 DAT (joint project Chile-France-Uruguay), and by ANII (Agencia Nacional de Investigación e Innovación, Uruguay)-Fondo Sectorial de Energía 2015, ANII-FSE_110454.

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

References

1. Weron, R.; Misiorek, A. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *Int. J. Forecast.* **2008**, *24*, 744–763. [[CrossRef](#)]
2. Cazes, G.; Ortelli, S. Minimum-Cost Numerical Prediction System for Wind Power in Uruguay, with an Assessment of the Diurnal and Seasonal Cycles of its Quality. *Ciência Nat.* **2018**, *40*, 205–210.
3. De Mello, S.; Cazes, G.; Gutiérrez, A. Operational wind energy forecast with power assimilation. In Proceedings of the 14th International Conference on Wind Engineering, Porto Alegre, Brazil, 21–26 June 2015.
4. Pinson, P. Estimation of the Uncertainty in Wind Power Forecasting. Ph.D. Thesis, École des Mines de Paris, Paris, France, 2006.
5. Pinson, P.; Nielsen, H.A.; Møller, J.K.; Madsen, H.; Kariniotakis, G.N. Non-parametric probabilistic forecasts of wind power: Required properties and evaluation. *Wind Energy* **2007**, *10*, 497–516. [[CrossRef](#)]
6. Pinson, P.; McSharry, P.; Madsen, H. Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation. *Q. J. R. Meteorol. Soc.* **2010**, *136*, 77–90. [[CrossRef](#)]
7. Staid, A.; Watson, J.P.; Wets, R.J.B.; Woodruff, D.L. Generating short-term probabilistic wind power scenarios via nonparametric forecast error density estimators. *Wind Energy* **2017**, *20*, 1911–1925. [[CrossRef](#)]
8. Møller, J.K.; Zugno, M.; Madsen, H. Probabilistic Forecasts of Wind Power Generation by Stochastic Differential Equation Models. *J. Forecast.* **2016**, *35*, 189–205. [[CrossRef](#)]
9. Elkantassi, S.; Kalligiannaki, E.; Tempone, R. *Inference and Sensitivity in Stochastic Wind Power Forecast Models*; UNCECOMP: Rhodes Island, Greece, 2017.
10. Caballero, R.M. Stochastic Optimal Control of Renewable Energy. Master's Thesis, King Abdullah University of Science and Technology, Tuwa, Saudi Arabia, 2019. [[CrossRef](#)]
11. Bertsekas, D.P. *Dynamic Programming and Optimal Control*; Athena Scientific: Belmont, MA, USA, 1995; Volume 1.
12. Risso, C. Benefits of demands control in a smart-grid to compensate the volatility of non-conventional energies. *Rev. Fac. Ingeniería Univ. Antioq.* **2019**, 19–31. [[CrossRef](#)]
13. Nohara, D.; Ohba, M.; Watanabe, T.; Kadokura, S. Probabilistic Wind Power Prediction Based on Ensemble Weather Forecasting. *IFAC Pap.* **2020**, *53*, 12151–12156.
14. Foley, A.M.; Leahy, P.G.; Marvuglia, A.; McKeogh, E.J. Current methods and advances in forecasting of wind power generation. *Renew. Energy* **2012**, *37*, 1–8.
15. Khan, P.W.; Byun, Y.C.; Lee, S.J.; Kang, D.H.; Kang, J.Y.; Park, H.S. Machine Learning-Based Approach to Predict Energy Consumption of Renewable and Nonrenewable Power Sources. *Energies* **2020**, *13*, 4870.

16. Álvarez, F.M.; Troncoso, A.; Riquelme, J.C.; Ruiz, J.S.A. Energy Time Series Forecasting Based on Pattern Sequence Similarity. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 1230–1243 [[CrossRef](#)]
17. REN21 Secretariat. *Renewables 2018 Global Status Report*; REN21 Secretariat: Paris, France, 2018. [[CrossRef](#)]
18. Jeon, W.; Lamadrid, A.J.; Mo, J.Y.; Mount, T.D. Using deferrable demand in a smart grid to reduce the cost of electricity for customers. *J. Regul. Econ.* **2015**, *47*, 239–272. [[CrossRef](#)]
19. Mo, J.Y.; Jeon, W. How Does Energy Storage Increase the Efficiency of an Electricity Market with Integrated Wind and Solar Power Generation? A Case Study of Korea. *Sustainability* **2017**, *9*, 1797. [[CrossRef](#)]