Storage sizing for grid connected hybrid wind and storage power plants taking into account forecast errors autocorrelation

Andrea Michiorri, Jesus Lugaro, Nils Siebert, Robin Girard, Georges Kariniotakis

To cite this version:

HAL Id: hal-01626067
https://hal-mines-paristech.archives-ouvertes.fr/hal-01626067
Submitted on 30 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ATTENTION, this is a preprint of:

Available at:

Storage sizing for grid connected hybrid wind and storage power plants taking into account forecast errors autocorrelation

Andrea Michiorri², Jesus Lugaro², Nils Siebert², Robin Girard², George Kariniotakis²
² MINES ParisTech, PSL - Research University, PERSEE - Centre for processes, Renewable Energies and Energy Systems, 1, Rue Claude Daunesse, 06904 Sophia-Antipolis, France
Corresponding author: andrea.michiorri@mines-paristech.fr

Abstract
This paper describes a research on the influence of wind power prediction error autocorrelation on the sizing of storage coupled with a wind farm.

The stochastic nature of renewable energies resources such as wind speed or solar radiation represents a challenge for the grid integration of renewable energy plants. The imbalances between renewable power predictions and realised production are generally penalised by system operators since additional reserves are required to maintain the stability of the grid. The coupling of storage devices with renewable energy plants is one of the solutions studied to reduce those imbalances. In this work, a methodology to manage imbalances and to size storage in order to achieve a determined level of controllability is proposed. It is applied to a specific use case: the integration of a combined wind-storage plant in French Guyana. The influence of the autocorrelations of errors on the battery size is investigated in detail and a methodology for producing wind prediction errors time series is presented.

Keywords
Wind, forecast, storage, sizing
1 Introduction

This paper describes the methodology and the findings of a research on the influence of wind power prediction error autocorrelation on the sizing of storage coupled with a wind farm.

The integration of wind power into power systems contributes to the reduction of CO$_2$ emissions and the security of supply through the use of endogenous resources. However, the increased share of electricity produced by fluctuating and not completely predictable renewable generation is altering the traditional operations of power systems. Its effects can be seen in an increased unpredictability of electricity prices patterns, in an increase to the use of intraday adjustments and on the increased request for reserves. Furthermore, the reduced utilisation factor of renewable generators (in the region of 30% for wind turbines and 15% for photovoltaic plants, although with strong differences between different regions and plants) reduce the profitability of the farms, because of additional network reinforcements. However, the limited predictability of wind power production represents a constraint for large-scale wind power integration. The difference between the power injected in the system and the power previously scheduled (due to errors in the production prediction) must be compensated by other generators, to maintain the balance among generation and demand [1].

A solution for part of these issues is the development of network storage capacity both associated and not associated with renewable producers. In particular, Storage associated with renewable power plants represents the most straightforward solution to mitigate renewables variability. It is also the only possible one if other solutions are not available, economically viable or they are not allowed. On the other hand storage non-associated with the generators presents several advantages: it can be mutualised with other users, it can exploit site-specific features (of primary importance in the case of pumped hydro or compressed air energy storage) [2], or it can be replaced by an interface with other energy networks such as gas or hydrogen [3]. Storage can be used to absorb excess electricity in hours of high renewable production and low demand and to re-inject it into the network later. The operating performance of such storage can be highly improved by considering simple charge-discharge plans based on short-term predictions of the renewable production. These predictions, generated by physical or statistical models using weather predictions and measurements as input are nowadays widely used by system operators in several countries. This increased performance can be translated directly into a reduced size of the storage.

A particular negative aspect of predictions errors for renewable energy production is represented by their autocorrelation. For example, if for a specific day an average wind production was predicted but in reality the day turned out to be very wind or very calm, the error for the wind farm production prediction will be constantly over or under estimated. Therefore, if an energy storage system is used to counterbalance this prediction error, it should be able to absorb or produce all or part of the renewable production for a full day or for a considerable number of hours. This simple example illustrates that prediction error autocorrelation should be considered when sizing storage devices for such applications whereas the use of a simple white noise would not result in an acceptable result. A solution to overcome this problem is the use of historical production and prediction data for a particular application. But if historical data is not available, prediction errors should be estimated through simulations.

In order to present clearly this study, the manuscript is structured in the following way: Firstly the state of the art is presented in Section 1.1, the methodology proposed is then described in Section 2, with the method proposed for the generation of forecast errors detailed in Section 2.2 and the methodology used for the sizing detailed in Section 2.3. The case study is the presented in Section 3, describing the specific needs for wind farm connection in French Guyana. Then the results of the

---

study on storage sizing are described in Section 4, with special attention given to the discussion of the importance and the limitations of the study. Finally conclusions are drawn in Section 5.

1.1 State of the art

Several solutions have been envisaged for the coordination between wind farms and storage plants. An example is the participation in electricity markets using pumped storage. Castronuovo and Lopes proposed in [4] the coordination between a wind farm and a pumped storage facility for increasing the controllability of the wind farm and maximize profits when participating in the Portuguese market. In [5] the same authors described new considerations about the optimal size of the pumped storage station. In [6] various methodologies are presented for coupling wind generation and storage for coordinated market participation. Korpås et al. [7] proposed a method for sizing and operating an energy storage system coupled with a wind power plant under the Norwegian market conditions. Koeppel and Korpås [8] analysed the utilization of a generic energy storage device for balancing the differences between predicted and real productions in a wind farm located in Norway when acting in a market environment. Garcia-Gonzalez et al. [9] analysed the combined operation of wind farms and a pumped storage facility for participating in the Spanish electricity market considering the uncertainties of both wind power generation and market prices.

All the above mentioned works showed that the optimal management of pumped storage coupled to wind farms results in an economic benefit and increase the controllability of the wind farm. However, for the specific case of isolated systems without hydropower potential, the use of DES is necessary. Several DES technologies such as battery storage [10], [11], ultra-capacitors [12] or flywheels [13], [14] are considered as efficient way, together with accurate prediction models, to increase renewable energy penetration in islands without installing additional reserves (i.e. based on thermal generation), with the additional advantage of increasing energetic independence of these areas. DES can also be used to overcome network congestion problems or to allow wind farms to respect grid codes. The work in [15] showed that pumped storage can be also very useful in isolated systems, improving both the dynamic security and the economic operation of the grid. Another example is presented in [16] where the sizing of a battery for a grid connected wind farm in order to provide frequency support is described.

Research on storage sizing for renewable energy integration is driven in part by the high capital cost of storage. In general storage sizing is studied as a minimisation problem of the fixed and variable costs of the storage and its application. An example can be seen in [17] where the initial capital cost of a DES is considered along with the operating cost and in a microgrid. In a similar case [18], a complete analysis of the cash flow of a DES used for integrating renewable power is studied and used to optimise the sizing of the storage. In this paper the return on investment (ROI) of the system is calculated taking into account the already mentioned cash flow and the expected length of the battery life, calculated considering the known lifetime in cycles and the number of expected cycles of the storage. The problem of aging in the sizing of electrochemical DES is also studied in [19], where it is found a linear relationship between battery lifetime and battery size was found as well as an inverse relationship between available export and lifetime. Furthermore, the desired lifetime is chosen for determining the optimal utilisation of the battery in the conditions chosen. In [20] the sizing of a DES for an isolated microgrid is calculated considering a risk based trade off method where the probability of not supplying the load is weighted against the cost of the necessary incremental capacity of the storage.

The use of prediction and prediction errors for storage sizing is not present or is not a fundamental part of the works described above. A first example of this problem can be seen in [21], where punctual prediction of load, wind and solar production are used for developing the optimal schedule for a DES in a microgrid for reducing its operating cost. The same approach is followed in [22], where attention is given to the modelling of a fuzzy Energy Management System. In general prediction errors are modelled as white noise following a normal distribution, but in [23] a more accurate definition of the probability distribution of prediction errors is studied and is modelled with a beta
distribution. The impact of the use of a more complex error distribution on the sizing of a DES devise is also described in the text. The prediction errors autocorrelation problem is studied in [24], where wind power trajectories are derived from predictions in order to study the optimal dynamic storage sizing for grid scale applications. The development of trajectories, according to the methodology described in [25], implicitly incorporates the existence of errors autocorrelation in predictions. Another study on the impact of day-ahead errors autocorrelation can be found in [26], where an autoregressive model is used for developing the necessary auto correlated time series. A particular solution is proposed in [27], where a sizing methodology is developed for a two component battery with better performance for slow and fast dynamics.

As shown in this section, the use of DES for renewable energy integration is an active research field centred mainly on the development of scheduling strategies and on the definition of optimal sizing for each application. This is shown also by the presence of several commercial packages such as HOMER, HOGA, TRNSYS and RETScreen and others as summarised in [28] and [29]. Anyway, historical wind speed measurement must be used in sizing and when these are not available, inputs are calculated from Weibull distributions as verified in [30], losing the autocorrelation of wind speed time series. This problem is addressed in [31] where artificial wind speed time series able to replicate the standard deviation and the autocorrelation of the measured ones are calculated with an ARMA model in order to be used in a storage sizing study for a hybrid system. Anyway in most of these studies, the storage management strategy applied does not take into account forecasts and therefore the topic of forecast errors autocorrelation is seldom faced.

1.2 Conclusions and contributions to knowledge

In summary, it can be said that previous work on storage sizing for wind power support: 1) are particularly interested on off grid applications, 2) make use of historical wind profiles or use a Weibull distribution in the case of Monte Carlo simulations, 3) predictive management, where forecasts for the wind power productions are used to schedule the charge and discharge of the storage, is seldom used. Therefore the impact of the autocorrelation of forecast errors on the sizing of the storage is not taken into account, with the exception of [26]. This paper aims at filling this gap presenting a methodology for producing auto-correlated prediction errors is presented, integrated with an optimal Monte Carlo based sizing procedure for the energy storage.

2 Methodology

The sizing of the battery is an optimisation problem where the characteristics of the storage, as well as its operation are taken into account. The storage system considered is composed by an inverter and by an electrochemical battery. The inverter is characterised by its own rating \( R_p \) [kW] and its own cost \( C_p \) [€/kW]. The battery is also characterised by its rating \( R_e \) [kWh] and its cost \( C_e \) [€/kWh]. Depending on the technology of the battery, a relation could exist between the energy rating \( R_e \) and the power rating \( R_p \) and the maximum charge rate can be different than the maximum discharge rate. In this work a value of 1 for the ratio between the charge and discharge power of the battery is chosen. This results in a charge rating equal to the discharge rating. The battery is used to filter the prediction error \( E \), a time series of \( m \) elements representing the errors for each hour \( dt \). The sizing of the battery is an optimisation problem, described by Equations (3-6) aimed at identifying the optimal power and energy rating of the storage \( R_{p\text{opt}}, R_{e\text{opt}} \) that maximise a selected objective function \( f_{\text{obj}} \) under a series of constraints \( \text{cons} \).

The objective function is considered to be equal to ROI, calculated as in Equation (1). The terms \( V_b \) [€] and \( C_b \) [€] represent the value of the action of the battery and the cost of the battery respectively. The cost of the battery \( C_b \) is calculated as in Equation (2) and the value of the battery is calculated as in Equation (3). The terms \( j \) and \( N_y \) are respectively the discount rate, considered in this work equal to 10% and the expected lifetime of the battery in years, calculated as in Equation (4), where \( N_{\text{max}} \) is the maximum number of cycles of the battery.
\[ f_{\text{obj}} = \text{ROI} = \frac{(V_b - C_b)}{C_b} \quad (1) \]

\[ C_b = (R_p + C_p + R_e + C_e) \cdot (1 + i)^N y \quad (2) \]

\[ V_p = \sum_{i=0}^{N y} \frac{aPen_i}{(1+i)^i} \quad (3) \]

\[ N y = \frac{\sum_{i=0}^{N y} \text{abs}(e_i)}{8760 \cdot \text{max}_0} \quad (4) \]

The constraint of the optimisation problem is represented by the necessity for the battery to avoid at least a certain percentage \( a\text{PenTarget} \) of the penalties that would be paid if the battery was not connected. This can be expressed analytically as in Equation (5). The terms \( a\text{Pen} \) (the avoided penalties), \( p\text{Pen} \) (the penalties paid) and \( \text{Pen} \) (the penalties paid without the battery) found in Equations (5-7-8) can be calculated as in Equation (8), where \( x \) [kWh] is the value of a prediction error and \( c\text{Pen} \) [€/kWh] is the cost paid for each penalty.

\[ \text{Constraints:} \frac{a\text{Pen}}{p\text{Pen}} > a\text{PenTarget} \quad (5) \]

\[ a\text{Pen} = \sum_j \left( \text{Pen}(E_j) \right) - p\text{Pen} \quad (6) \]

\[ p\text{Pen} = \sum_j \left( \text{Pen}(E_f) \right) \quad (7) \]

\[ \text{Pen}(x) = c\text{Pen} \cdot \text{abs}(x) \quad (8) \]

The terms \( E \) and \( E_f \) in Equation (8-9) represent the time series of prediction errors before and after the filter applied by the battery. In order to describe clearly the calculation of the filtered errors it must be considered that the battery in this study is used as a filter of the prediction errors \( E \), absorbing the errors \( E_a \) and letting to pass the errors \( E_f \). This can be represented as shown in Equation (9). This means also that the absorbed errors \( E_a \) represents the power effectively charged and discharged by the battery, whose \( SOC \), can be calculated as the integral of the absorbed errors as shown in Equation (10). The term \( P_{aux} \) represents the power absorbed by the auxiliaries of the battery, for example for heating or cooling the device. The absorbed errors \( E_a \) are also limited by the saturation of the power capacity and the energy capacity of the battery, as shown in Equation (11) and (12). The filtered errors can then be calculated considering Equation (11-12-14). Here the term \( \eta^2 \) takes into account the battery charge and discharge efficiency: \( \eta \) is the full cycle charge / discharge efficiency and \( z \) is equal to \( \frac{1}{2} \) or \( \frac{1}{2} \) if the battery is charging or discharging respectively.

\[ E = E_a + P_{aux} + E_f \quad (9) \]

\[ SOC_i = SOC_{i-1} + \frac{E_a}{R_e} \frac{dt}{\eta^2} \quad (10) \]

\[ \text{abs}(E_{a,i}) < \text{abs}(R_p) \quad (11) \]

\[ 0 < SOC_{i-1} \cdot R_e + E_{a,i} \cdot dt \cdot \eta^2 < R_e \quad (12) \]

### 2.1 Discussion on the sizing methodology

In a real-world situation, as seen in the use cases described in Section 3, the sizing of the storage will be subject to a larger number of constraints, such as a minimum ROI of the project, a minimum value for the filtered errors, a more complicated penalties function, etc. The objective of this study is
however, to clarify the relation between the statistical properties of the prediction errors and the sizing of the battery. For this reason, the number of hypothesis and assumptions was kept to the minimum, and simple constraints or linear relations have been used. The four strongest simplifications or assumptions done are: 1) the use and the shape of the objective function and of the constraints, 2) the use a shape of the penalties function, 3) the battery utilisation strategy and 4) the eventual constraints given by particular storage technology. More in detail:

1) Optimisation function. A simple objective function equal to the ROI has been used during the optimisation. The only constraint considered is the objective of reducing the amount of penalties paid by more than a pre-set value (in this case arbitrarily set at 70%). Other possible objectives or constraints such as the reduction of the number of penalties have not been considered.

2) Penalties function. The penalty function considered is linear and symmetrical. Positive and negative errors have the same cost and this is equal to the value of the electricity produced by the wind farm, in this case equal to the feed in tariff. In other situations, small errors can be penalised less than larger ones, and errors going in the direction of helping network stability can have a lower cost.

3) Battery utilisation. In this study, the battery is used simply to filter prediction errors. It is assumed that all the information available has been used to provide the most accurate predictions, therefore the resulting errors are unavoidable. It is also chosen to not consider other services potentially offered by the battery, such as price arbitrage or reserve for stability. Finally, incorporating a more complex battery management strategy would have made the results dependent on the particular strategy chosen and not only on the input errors. Furthermore complex control techniques do not always perform better with respect to simpler ones, as shown in [1].

4) Battery technology. This study is technology agnostic concerning the type of battery used. Different battery technologies present multiple constraints, such as specific ratios between power and energy rating, non-linear aging functions, etc. The objective is to let the user select the most appropriate technology according to the use of the battery determined by the prediction errors properties. In any case, it is necessary to fix a value for the cost of the storage, and values of 100 €/kW and 400 €/kWh have been chosen for the inverter and the battery of the storage unit.

2.2 Generation of the prediction errors time series

The error time series are characterised by a standard deviation $\sigma$ and a Kriging scale parameter $\alpha$. This last represents the level of autocorrelation in the time series. With this method an autocorrelation plot (as the one shown later in Figure 5) is approximated with a kernel function, in this case an exponential with a negative exponent. This function is then utilised to calculate the elements of a transformation matrix used in turn to obtain the desired time series from a random one. An autocorrelation matrix $A[a_{ij}]$ is created as in Equation (13). The errors array $E$ is then calculated as in Equation (14), where $R$ is an array of random Gaussian elements. The resulting time series is scaled as shown in Equation (15) in order to obtain the desired standard deviation $\sigma$. The time series is then filtered in order to remove the elements larger than 1 in absolute value as shown in Equation (16).

\[
a_{ij} = e^{-\frac{\text{abs}(i-j)}{\alpha}} \tag{13}
\]

\[
E = A^{0.5} \cdot R \tag{14}
\]

\[
E = \frac{\sigma}{\text{stdev}(E)} \cdot E \tag{15}
\]
\[ E = \max[-1; \min[1; E]] \cdot P_{\text{max}} \]  

(16)

The resulting time series, generated with this method are represented in Figure 1, where, for each combination of the standard deviation and the Kriging scale parameter, the resulting time series is depicted. The standard deviation ranges from 0.05 and 0.5 whilst the scale parameter ranges from 1 to 10. Each series in the plots includes 500 hourly values. This method, which is not intended to reproduce bias, was tested with the same standard deviation and kriging parameter of the measured forecast errors, producing a new errors time series with the same autocorrelation.

Figure 1: Generated errors time series

2.3 Optimal sizing of the storage device

An example of the results obtained from the application of this procedure is shown in Figure 2, where the different parameters used in the multi objective optimisation are plotted for different combinations of the power and energy rating of the storage. The results are obtained for an error’s time series with a standard deviation of 0.2 and a kriging scale parameter of 5 and with a wind farm rating of 1 MW. In Figure 2 (a), the ratio of penalties larger than 15% observed with and without the battery is shown. Clearly the ratio decreases as the storage size increases, both in power and in energy. In Figure 2 (b) the ROI of the battery is plotted. This parameter is larger for small batteries, although small values of the energy or power rating have a negative effect, since the amount of avoided penalties is reduced. The area of values with positive ROI is bounded by the red iso-curve of null ROI.

The optimal design, corresponding to the higher value of the objective function is highlighted with a diamond in the two subplots. Its value is obtained by selecting the combination of \( R_p \) and \( R_e \) which maximise the ROI whilst maintaining the value of avoided penalties higher than a defined value. This process is repeated iteratively with a Monte Carlo approach and the values of the optimal rating for power and energy are stored in two arrays \( Y_1 \) and \( Y_2 \). After each iteration \( N_{\text{it}} \) the inequality described in Equation (17) is tested, and if true for both the arrays the iteration is stopped. The average value of \( Y_1 \) and \( Y_2 \) is taken as the result of the simulation.
\[ N_{it} < \left( \frac{3 \cdot \text{std}(Y)}{\text{average}(Y)} \right)^2 \]  

(17)

Figure 2: Map of avoided penalties ratio (a) and ROI (b) for different values of storage power and energy rating. The diamond shows the optimal battery configuration.

3 Case study

The present study is based on a wind-storage plant projected for installation in French Guyana. This project was developed following a tender launched in 2010 by the French Energy Regulation Commission for the installation of new wind farms. The technical requirements of the tender included both the installation of local storage to increase the controllability of the wind farm output and the use of short term wind power predictions. In this context, the authors conducted a research study to better size the storage and to develop performing management strategies. The wind farm nominal power is in the order of 9 MW.

In the specifications of the above mentioned tender, each wind farm operator must inform daily the system operator of the predicted production. The wind farm operator will then be paid and penalized according to his ability to produce the expected power. The predictions are composed of 48 power values, one for each 30-minutes period of the day. The obligation is to use storage to smooth-out the
power output of the combined plant and more precisely to provide a constant level of power for each 30 minute period. Upper and lower bounds are imposed that define penalties for deviations. In Figure 3, the predictions are shown in red for three periods of 30 minutes. The real production is shown in black and the admissible range of production in blue. However, it is possible to see that this constraint is met in periods one and three but it is not met in period 2. The producer would be penalized by the operator if this situation occurred in reality.

![Figure 3](image)

**Figure 3**: Example (3 periods) of day-ahead wind power prediction (red), real production (black) and admissible band (blue)

### 3.1 Wind farm production estimation

The present study is based on a project of building a wind-storage plant. A campaign of measurements was required to simulate the production of the projected wind farm. The measurements were realized during 2 years (2007 and 2008) with two anemometers placed at the future location of the wind farm. Each anemometer was placed at a different height: 49.5 m and 72 m respectively. Since the nacelle of the wind turbines is located at a height of 95 m, the wind speed power law shown in Equation (18) was used to extrapolate wind speeds. This relationship is commonly used in wind power engineering and gives a good approximation of the speed [32].

\[ v = v_0 \left( \frac{h}{h_0} \right)^{\beta} \]  

(18)

Then, the manufacturer’s power curve was used to estimate the power output of the wind farm, composed by five wind turbines of a total capacity of 9 MW. In this work the simulations cover a 6 months period, from November 2007 to April 2008. Due to the particular wind regimes in French Guyana, this period corresponds to about the 70% of the annual expected energy production of the wind farm.

### 3.2 Penalties structure

The technical requirements of the tender are summarized in this section. The goal is to describe precisely the conditions that will allow the formulation of the wind-storage plant management algorithm. Firstly, wind farm developers must install storage devices at the plant location in order to increase the controllability of the power plant. The four controllability conditions to be met are:

1. The difference between the injected power and the production plan (production forecast) should not exceed ± 25% of the wind farm nominal power \( P_{\text{max}} \) during the first year of operation, ± 20% of \( P_{\text{max}} \) during the second year or ± 15% of \( P_{\text{max}} \) for the following years. The
The rationale behind the reduction of the tolerance threshold in time is that an improvement in the accuracy of wind power predictions could be expected in time.

2. During the transitions between two power levels defined in the day-ahead production plan, the plant must complete the passage from from 0 to $P_{\text{max}}$ in a period between 30 seconds and 5 minutes, and the change between $P_{\text{max}}$ to zero in a period from 1 and 10 minutes.

3. While injecting a power $P_{\text{inj}}$, the plant must have a power reserve of:

   \[
   \begin{cases}
   10\% \cdot P_{\text{max}} & 0 < P_{\text{inj}} < 90\% \cdot P_{\text{max}} \\
   P_{\text{max}} - P_{\text{inj}} & P_{\text{inj}} > 90\% \cdot P_{\text{max}}
   \end{cases}
   \]

   This reserve can be requested by the system operator for at least 15 minutes and should be available in less than 0.5 seconds.

4. The wind power plant should be able to regulate the voltage at the point of grid connection. In particular, the voltage should vary between 95% and 105% of the nominal voltage at the point of grid connection.

The wind farm developer is subject to financial penalties if one or several of the conditions defined above are not met. The penalties imposed by the system operator are either financial or operational. An operational penalty is represented by the disconnection from the grid. In both cases, they result into financial losses reducing the total revenue of the plant. The penalties are assigned according to the following rules:

1. If the instantaneous measurement of the injected power (on a one minute average) does not respect controllability conditions 1 and 2, the production during the ten minutes interval comprising the event will be penalized at 50% of the feed-in tariff.

2. Each event of no respect of controllability conditions 3 or 4 leads to a penalty equivalent to the revenue obtained by the independent power producer if producing 6 hours at nominal power.

3. If more than 100 events of no respect of conditions 1, 2, 3 or 4 occur during the last 30 consecutive days, the installation might be disconnected from the grid.

Finally, the independent power producer can be exempted from penalties in the following cases: i) if the storage is requested by the grid operator during the 12 hours preceding the event, ii) if the energy produced by the wind farm is smaller than 20% of $P_{\text{max}}$ and iii) if the wind farm is disconnected due to other issues in the power systems (congestions, outages, etc).

### 3.3 Wind power prediction and errors

Wind power production prediction is an active area of research but will not be treated in detail in this work. A detailed overview of the state-of-the-art can be found in [33]. The description of the state-of-the-art wind power prediction model used in this paper, based on the Kernel Density Estimators (KDE) method can be found in [34] and [35]. The method takes into account Numerical Weather Predictions and historical production records to produce a probability distribution of the expected wind power production. Recent developments in wind power forecasting, show that the use of time-adaptive kernel density estimation, by making use of the latest measured sample of wind farm production, is able to improve the accuracy of the prediction. An example of such system is shown in [36], where its performances are compared with an off-line model such as the one used in this article.

In Figure 4 (a) the forecasts for the period 02-08/10/2007 are shown, along with the measured value of the wind power production. In this study the predictions are calculated each day at 12:00, in order to let the producer propose a production schedule to the operator. This is at the origin of the discontinuities that is possible to observe in the diagram. The difference between the median of the distribution and the observed wind farm production is the error that the battery is called to filter. The error for the same period is represented Figure 4(b). The main characteristics of the time series are summarised in Table 1.
The analysis of the Kriging scale parameter of the KDE model highlights also the higher correlation of the predicted power with respect to the observed wind power production, at the origin of the error correlation. The error variation at each time step on the contrary presents a very low correlation. The Kriging parameter, described in detail in Section 2.2 can be considered as a measure of the autocorrelation of the time series.

Central to this work is the autocorrelation of the forecast error, shown in Figure 5 along with the correlogram of the predicted and measured production. Anyway it is important to remember that the predicted power represents the median of a distribution of possible outputs, and not a single scenario, explaining in part its higher autocorrelation respect to the measured value of the production. Observing the autocorrelation plot for the errors time series, it is possible to see how its value is higher than 0.5 for the first 6 lags. This means that a significant autocorrelation is observable for a period of about three hours and a weaker autocorrelation is observed for a period of up to 10 hours.

![Figure 4](image_url)

**Figure 4:** (a) Prediction of wind power production showing the quantiles of the prediction probability distribution, its median and the observed production in the same period. (b) Dimensionless error between observed wind power production and its day-ahead prediction. The diagrams are relative to the period 02-08/10/07
Table 1: Statistical description of the time series for the Observed Power, Predicted power and Error

<table>
<thead>
<tr>
<th></th>
<th>Observed Power</th>
<th>Predicted Power</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min [MW]</td>
<td>0</td>
<td>0.469</td>
<td>-7.014</td>
</tr>
<tr>
<td>Average [MW]</td>
<td>3.11</td>
<td>2.656</td>
<td>-0.454</td>
</tr>
<tr>
<td>Max [MW]</td>
<td>9</td>
<td>5.904</td>
<td>4.458</td>
</tr>
<tr>
<td>StDev [MW]</td>
<td>2.231</td>
<td>1.090</td>
<td>1.802</td>
</tr>
<tr>
<td>Kriging Scale Parameter</td>
<td>13.9</td>
<td>36</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Figure 5: Autocorrelation plot of the time series for the Observed Power, Predicted Power and Error

### 3.4 Storage management solutions

The sizing methodology proposed in this work is dependent on the storage management strategy. In this work, the algorithm proposed in [37], able to maximise a renewable producer revenue by bidding in an energy market with the help of a storage, is used. This is a typical problem of optimisation under uncertainty as explained in [38]. Much work has been devoted to the problem of decision-making under uncertainty, which is often modelled as an optimization problem [39], [40], [41].

The algorithm used in this study is divided in two stages: the definition of the day-ahead production plan and the operation of the wind-storage plant. The inputs are the wind power predictions, real time measurements of the plant such as wind production and the storage state of charge (SOC), the technical parameters of the wind farm, of the storage and the tender controllability conditions. The production plan is made of a series of 30 minutes constant power intervals and the power gap between two different intervals must respect the maximum variation rates defined in Section 3.2.

For the first stage, a simple production plan is used, consisting in submitting to the operator a production plan equal to the latest available wind farm production prediction. Also, the objective is restricted to satisfying condition 1 and condition 2 of the tender, since conditions 3 and 4 depend on external conditions and requests from the grid operator for which no information was available.

For the second stage the wind-storage hybrid plant operation algorithm is based on a simple filter approach, which aims at reducing the real time difference between the wind power and the day-ahead production plan in the periods where the controllability condition 1 is not met, while respecting condition 2. If the difference between the production plan and the effective wind farm
production is less than 15%, the batteries are not used. If the difference is higher, the batteries are charged or discharged, if possible, to mitigate the difference. The penalties will not be avoided when the battery power rating is smaller than the error to be absorbed and the battery SOC is close to its limits.

More advanced planning and control approaches, taking into account the penalties structure, the real time SOC of the battery and the possibility of wind farm down regulation have been considered, but are not reported in this paper centred on storage sizing.

3.5 Case study summary table

In this section, simulations of a large energy storage system (5 MW / 36 MWh) are presented for the management of wind power imbalances. This sizing results in 4 hours storage time for the studied wind farm (9 MW) and a power rating equivalent to the 55% of the wind farm nominal power. Four simulations were realised using simple approaches to define the production plan (deterministic prediction) and to operate the hybrid plant (simple filter without wind turbines regulation). The four simulations are characterised by different types of wind power prediction. This will demonstrate the impact of predictions properties on the storage operation performance, before proceeding to the sizing optimisation. In other words, the goal of the results presented here is not to find the optimal sizing, which will be presented in Section 4, but to show the impact of the predictions errors magnitude, bias and temporal correlation on the number of penalties and disconnection time received by the wind-storage plant and on the storage lifetime.

Table 2 summarises the characteristics of the 4 types of predictions (A-D). The four cases are characterised by the presence of different error time series with different normalised bias (NBIAS) and normalised mean absolute error (NMAE). Normalisation is done using the rated power of the wind farm. Simulations B, C and D are based on hypothesis on the performance of forecasts and cannot be obtained in an operation setting. The calculated life spans take into account only the aging for cycling for a NAS battery and not the calendar aging.

<table>
<thead>
<tr>
<th>Scenario parameters</th>
<th>Simulation Cases</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage size</td>
<td></td>
<td>5 MW</td>
<td>5 MW</td>
<td>5 MW</td>
<td>5 MW</td>
</tr>
<tr>
<td>Wind power prediction</td>
<td>KDE</td>
<td>KDE unbiased &amp; reduced NMAE</td>
<td>Measurement +white noise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBIAS (% Pn)</td>
<td></td>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NMAE (% Pn)</td>
<td></td>
<td>17.3</td>
<td>17.5</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Grid disconnection times (% of days)</td>
<td>167 (90.7%)</td>
<td>150 (81.5%)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Number of penalties</td>
<td></td>
<td>3214</td>
<td>2532</td>
<td>59</td>
<td>38</td>
</tr>
<tr>
<td>Number of full discharge battery cycles</td>
<td>62.9</td>
<td>51</td>
<td>18.5</td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>Lifetime estimated (years)</td>
<td>35</td>
<td>44</td>
<td>121</td>
<td>114</td>
<td></td>
</tr>
</tbody>
</table>

First, it is clear from simulations A and B that with high wind prediction errors (NMAE around 17%) and the considered storage size and operation strategy, the number of penalties is excessive leading
to unacceptable disconnection time of the plant (above 80% of the days). This result justifies the necessity to find advanced control strategies in order to avoid the need of very large storage system to meet the grid requirements.

Second, it is observed that most of the penalties (about 90%) in simulation A were positive penalties because of two reasons: i) in occurrences of power below 20% of nominal power the wind farm is exempted of penalties and ii) positive bias of wind power prediction considered in simulation A (NBIAS of 2.6%).

From a comparison between results of simulations A and B, it is observed that the studied storage unit performs better in managing wind imbalances when the wind prediction is unbiased. In particular, if the bias from the wind prediction used in simulation A is removed, the number of penalties is reduced by more than 20% in the present case study. This result can be explained by the fact that the bias of a prediction means that there is either a domination of positive over negative errors or the inverse and this disequilibrium increases the number of times that the battery is respectively fully charged or fully discharged (when using a simple control strategy). These states of the battery (SOC=0% or SOC=100%) should be considered as critical states since the risk for storage to be unable to compensate predictions error is higher. In conclusion, the presence of bias makes it more difficult for storage to manage imbalances and increases the chance of penalization. In simulations C and B, by reducing the NMAE from 17% to 6% in the same prediction series, the number of penalties is drastically reduced to 59, allowing the plant to be always connected to the grid for injecting the wind production.

When comparing simulations C and D, it is possible to understand the impact of the temporal correlation of errors on the penalties. To isolate the impact of the temporal correlation from the bias or the error magnitude, an artificial prediction was created for simulation D with a white noise in the top of the wind production measurements. It is important to remind that white noise does not present any correlation by definition. The number of penalties for simulation D is 38 in comparison to 59 in simulation C. This reduction around 30% demonstrates that the temporal correlation of errors makes it more difficult for storage to manage imbalances and increases the chances of penalization. The reason behind this fact is that high temporal correlation of errors means that the same type of error (positive or negative) is repeated over a large period of time. The storage may be able to compensate the imbalances for a certain time but saturation states will be reached more frequently.

Concerning battery lifetime, it was observed that the number of cycles was low, resulting in high lifetime (above 20 years) for all simulations, probably due to the large size of the storage considered.

In conclusion, the ability of the storage to manage wind power variability does not only depend on the error magnitude that needs to be mitigated but also on the bias and temporal correlation of errors. The reduction of the bias, of the error magnitude and the error autocorrelation can be achieved by better predictions. This is particularly true for the bias and the error magnitude which could present better performance also thanks to the geographical characteristics of a site. Anyway, the latter is believed to be more difficult to reduce, because it is more dependent on mesoscale meteorological effects which are difficult to control and measure at the scale of a wind farm.

4 Results and discussion

4.1 Effect of errors standard deviation and autocorrelation on battery sizing

In general it is reasonable to think that an increase of the standard deviation of the errors would result in an increase of the power rating of the battery. A battery with a larger power rating will be able to absorb larger errors, reducing the penalties. At the same time it is possible to think that an increase in the autocorrelation of the errors results in an increase of the energy rating of the battery. A battery with a larger energy rating will be saturated less easily which will result in a lower number of penalties. This can be seen in the charts reported from Figure 6 to Figure 9, where the main results
of the study are presented. The value of the parameters used in this simulations are summarised in Table 3.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter [Unit]</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{\text{max}}$</td>
<td>Wind farm rating [kW]</td>
<td>1000</td>
</tr>
<tr>
<td>$i$</td>
<td>Discount rate [%]</td>
<td>10</td>
</tr>
<tr>
<td>$N_{\text{max}}$</td>
<td>Maximum number of cycles</td>
<td>5000</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Cost of power rating [€/kW]</td>
<td>100</td>
</tr>
<tr>
<td>$C_e$</td>
<td>Cost of energy rating [€/kWh]</td>
<td>900</td>
</tr>
<tr>
<td>$c_{\text{Pen}}$</td>
<td>Cost of penalty &gt;10% [€/kW]</td>
<td>0.09</td>
</tr>
</tbody>
</table>

These figures show the optimal value of the power rating and of the energy rating of the storage for different couples of values of the errors standard deviation and autocorrelation, represented by the kriging scale parameter. As in the previous section, the standard deviation ranges from 0.05 to 0.5 whilst the kriging scale parameter ranges from 1 to 10. For clarity, the results are reported for a wind farm with a nominal rating of 1MW, producing a maximum error of 1 MW. The figures report also the resulting expected ROI in Figure 10, the lifetime for the optimised battery in Figure 9 along with the initial capital cost of the battery in Figure 8.

Concerning the power rating, Figure 6 shows that this design parameter is mainly influenced by the error standard deviation, and only marginally by the error autocorrelation. The results can be interpolated with a linear function as shown in Equation (19), with the parameters $a$ and $b$ calculated with an interpolation from the results. In this case they were observed values equivalent to $a = 6.8$ and $b = 1091$. The small value of $a$ suggests that its influence is negligible, at least in the conditions of this study, and that the effect of error autocorrelation on the optimal power rating should be captured by a more complex model.

$$R_{P,\text{opt}} = a \cdot \sigma + b \cdot \sigma$$  \hspace{1cm} (19)

Concerning the optimal energy rating, Figure 7 shows that this parameter grows both with the error standard deviation and with the error autocorrelation. The results can be interpolated with the function shown in Equation (20), in which the value of the parameter $c$ is estimated to $c = 766$.

$$R_{E,\text{opt}} = c \cdot \alpha \cdot \sigma$$  \hspace{1cm} (20)

The reason of this difference in the response of the optimisation algorithm can be explained by the higher cost of the battery with respect to the power rating. When possible, the algorithm tries to reduce the penalties, first thanks to a larger power rating, and only after, through an increase of the storage energy rating.

The behaviour of the optimal energy rating can be found qualitatively in the behaviour of the resulting initial capital cost for the optimised battery shown in Figure 8.

Concerning the expected lifetime of the battery, Figure 9 shows that its behaviour is roughly linearly dependent on the error autocorrelation and ranges from roughly 2 to 17 years for low and high values of error autocorrelation respectively. This means that independently from the error standard deviation, when the error autocorrelation increases the storage is subject to a smaller number of charge and discharge cycles during a time unit, resulting in a longer lifetime. This, in turn, influences the ROI, as seen in Figure 10, which mimics the behaviour of the battery lifetime and falling roughly linearly with the increase of error autocorrelation from a minimum value of -30% to a maximum of -60%.

It is important also to remember that these numeric values are relative to the specific costs considered in this work for battery capital cost and penalties and to the battery management.
strategy. The limitations of this study and their impact on the results are described in detail in Section 4.3.

Finally, according to the methodology the optimal size of the battery for the conditions of the case studies described in Section 3.5 (standard deviation $\sigma = 17\%$ and autocorrelation $\alpha = 4.7$) would be of about 200 kW and 500 kWh for each MW of error. Considering a rating of 9 MW for the wind farm, this is translated in a battery of about 1.8 MW and 4.5 MWh. This would result into a reduction in the rating values of 3.2 MW and 0.5 MWh with respect to the battery considered. The optimised storage is also characterized by a shorter lifespan (roughly 8 years) with respect to the storage simulated in the use case (more than 35 years) as a direct consequence of the smaller energy rating. The shortest lifespan of the battery means that every 8 years the battery must be replaced with a similar one and the proposed methodology takes this into account. Anyway, it is reasonable to think that with time, battery costs can change, either downwards thanks to the development of the sector or upwards because of raw materials scarcity; these long term cost evolutions have not been taken into account into this study. The return of investment of the proposed solution is negative (roughly -20%), this means that with the current costs for energy storage considered, the use of the battery represent an additional cost necessary to respect the connection requirements.

![Figure 6: Optimal storage power rating for different values of prediction errors standard deviation and autocorrelation](image-url)
Figure 7: Optimal storage energy rating for different values of prediction errors standard deviation and autocorrelation

Figure 8: Optimal storage capital cost for different values of prediction errors standard deviation and autocorrelation
4.2 Importance of the findings for storage and renewable energy integration

The main contribution of this work is the demonstration of the importance of the autocorrelation in renewable energy production prediction errors on the sizing of storage devices used for renewable energy integration. The paper presents also a methodology for storage sizing taking into account prediction errors and their autocorrelation that can be used in further works on battery sizing.

The methodology proposed has the twofold advantage for storage manufacturers and renewable energy developers. For storage manufacturers it provides the possibility to precisely calculate the behaviour of their products, thus it allows increasing their added value for customised solutions. For renewable developers wishing or constrained to smooth out the variability of their renewable energy production.
generation thanks to the use of a storage system, a more precise sizing would reduce the initial capital cost whilst maintaining the performance for the system. Concerning the methodology proposed, based on error time series generation, optimisation and Monte Carlo estimation, it is believed that its importance lies in the reliability of the results and in the flexibility in adapting to sites changes. The possibility to apply the proposed methodology to a new installation site for which there is not available an historical production is also an advantage, provided a previous study on typical prediction errors autocorrelation for different climates and geographical areas.

4.3 Limitations of the study and further research perspectives

In this work, a balance has been found between the need to describe a realistic use case and an abstraction necessary to extract the fundamental behaviour of the phenomenon. For this reason, the limitations in the various hypotheses, which are mainly related to the numerical value of the results more than to methodology described, are also listed and described in detail in this section. The main limitations identified are due to the:
- generic battery technology representation
- wind profiles and predictions which are relative to a specific climate
- penalties structure
- battery management strategy
- optimisation objective function and constraints

Although the study was conceived to be technology agnostic concerning the type of storage considered, information about specific batteries technology are still present in the results. This is because of the relative values of the cost used for the power and the energy rating of the battery compared to the penalties. Different battery technologies would have different relative values of these parameters. The simplified model used can also be improved by modifying Equation 4 and if necessary adding new relations able to describe the behaviour of the specific battery technology. Also the results obtained are relative to a simplified storage with perfect efficiency and no losses from auxiliaries.

In this work, wind speed measurements and predictions from the French Guyana have been used, influencing the results. Different wind regimes can be characterised by a different variability, a higher or lower average value and a different predictability, influencing the result of the sizing exercise. The penalty structure used in this work is taken from the use case described in Section 3.2, but other structures could be considered. The penalties threshold could be higher, lower or even not present, or more thresholds could be used. Penalties could be nonlinear, or asymmetrical, and their value depends on current grid or market conditions.

The battery management strategy is expected to considerably affect the optimal sizing and it is likely that in real-world applications a more complex strategy might be applied than the simple filter strategy considered here. Now, the power announced in day ahead is simply equal to the median of the probabilistic forecast, but an advanced bidding strategy based on the use of probabilistic predictions and a detailed description of the penalty structure would reduce the amount of errors to be filtered. The use of intraday rescheduling based on updated predictions can also have an impact on the sizing as shown in [42].

Finally, the optimisation objective function and constraints have also an impact on the results of the work: the choice of maximising the ROI is reasonable but could be coupled with the optimisation of other parameters, and the choice of reaching a specific reduction of the amount of penalties paid is also arbitrary. Other possible criteria could be the minimisation of the break-even time of the investment or the reduction of the number of penalties independently from their amplitude.

In light of these considerations, it is possible to see that the objective of the work, the study of the influence of prediction errors correlation on storage sizing for wind power integration has been achieved. The methodology proposed can be applied with the necessary adaptations to a large range
of use cases. Possible further area of research on this field can be the study of the impact of different storage management strategies, the study of typical forecast errors autocorrelation values in different geographical areas, the use of a non-Gaussian probability distribution function for prediction errors and the validation of the time series generation technique for other types of renewable energy such as solar or marine power.

5 Conclusion

The influence of wind prediction errors autocorrelation on the sizing of energy storage for wind power network integration has been studied in this paper. This problem has arisen during a series of studies for the integration of wind power in French Guyana and can be generalised for other cases of where storage is used for facilitating the integration of wind power into the electric network. The increased penetration of weather-dependent and not completely dispatchable renewable resources is increasing the interest on this problem, as can be seen also from the analysis of previous work. The paper has presented the original case study which is briefly summarised here: in order to be connected to the network, a wind farm must provide predictions for its day-ahead production and when predictions are not met within a certain tolerance band, penalties must be paid to the network operator. In this scenario, a storage system is used for reducing the number and the amount of penalties. In the case study the wind power and storage system is modelled and its behaviour simulated over a period of about six months. Wind power production predictions have been created using a Kernel Density Estimation based algorithm. The case study and its data have been taken as example for developing a storage sizing methodology taking into account the observed phenomenon of prediction errors autocorrelation. The methodology proposed is divided in two parts: firstly auto-correlated errors time series characterised by a defined variance and autocorrelation are generated with a Kriging approach. The behaviour of the storage is then simulated according to the specific storage management approach used. The simulations are repeated for different combinations of storage power and energy ratings until the optimal combination of the two parameters is found. The sizing methodology was then applied to the data used in the case study. As expected, it was proven that an increase of errors variance is translated directly into an increase of the optimal storage power rating, whilst an increase in errors autocorrelation is linked with an increase of the storage energy rating. This relation is not linear, probably because of the higher per unit cost of energy rating respect to the power rating (the higher cost of the electrochemical battery respect to the inverter). Furthermore, thanks to the study, it was possible to calculate the optimal size of the storage that would have been necessary for the wind farm of the use case: for the conditions measured at the wind farm (prediction errors standard deviation $\sigma = 17\%$ and autocorrelation $\alpha = 4.7$), a 1.8 MW and 4.5 MWh storage would have been necessary for reducing the amount of penalties paid by 70%.

Although the problem can be particularly complex, the work presented in this paper used simple models wherever it was possible, describing in detail the simplifications made and, when stronger hypothesis were done, the reason behind the choice. This has been necessary in order to maintain the study general enough for describing its objective, the relation between storage sizing and prediction error autocorrelation, instead of a particular use case. In particular, the storage management approach is expected to have a considerable impact on storage sizing and the study of the impact of different storage management algorithms on the storage optimal sizing can be considered a new promising field of research.

A final observation concerns the impact on grid efficiency of storage used to minimise renewables volatility. Storage introduces additional losses due to its charge discharge efficiency in the order of 95% and 85% for Li-ion and NaS battery respectively [43], values that would represent problems if applied to the whole renewable energy production. Anyway when they are used only for filtering large forecast errors, as in this paper, the additional losses must be applied only to the fraction of
energy represented by the errors. In the case study presented, forecast errors processed by the battery represent roughly the 73% of the total energy produced by the wind farm, therefore losses would account for roughly the 3.65% of the total energy produced. This value can be reduced further with improved wind power forecast and possible with improved battery management strategies where allowed by the grid code.

**Acknowledgement**

The authors would like to acknowledge Mr. Sebastien Bourgeois for his collaboration in carrying out this study as well as the company Cr’Eole for supporting this work and for providing the data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Nomenclature**

- **A**: autocorrelation matrix
- **a**: [kW] ratio between power rating and errors kriging factor
- **a_pen**: [€] avoided penalties thanks to the battery
- **b**: [kW] ratio between power rating and errors standard deviation
- **c**: [kWh] ratio between energy rating and product of errors kriging factor and standard deviation
- **C_b**: [€] cost of the battery €
- **C_e**: [€/kWh] cost per unit of energy storage
- **C_p**: [€/kW] cost per unit of inverter’s power rating
- **c_pen**: [€/kW] cost of penalties
- **E**: [kW] array of the prediction errors
- **E_a**: [kW] array of errors absorbed by the battery
- **E_f**: [kW] array of the prediction errors filtered by the battery
- **E_max**: [kW] maximum value of the error
- **E_t**: [kW] prediction error at the time t
- **f_obj**: optimisation objective function
- **h**: [m] height at what the wind speed is calculated
- **h_0**: [m] height at what the wind speed is measured
- **I**: [%] discount rate for the project
- **N_max0**: [years] maximum number of cycles of the battery
- **N_y**: [years] lifetime of the battery in years
- **P_aux**: [kW] power absorbed by the auxiliaries of the battery
- **P_en**: [€] penalties paid by the wind farm without battery
- **P_inj**: [kW] power injected into the network by the wind farm
- **P_max**: [kW] wind farm nominal power
- **p_pen**: [€] paid penalties when the battery is connected
- **R**: array of random Gaussian elements
- **R_e**: [kWh] energy rating
- **R_{eopt}**: [kWh] optimal energy rating
- **R_p**: [kW] power rating
- **R_{popt}**: [kW] optimal power rating
- **v**: [m/s] wind speed
\( v_0 \) [m/s] wind speed at reference height

\( V_b \) [€] value of the battery €

\( \alpha \) kriging scale parameter

\( \beta \) wind power law coefficient

\( \sigma \) standard deviation

\( \eta \) charge / discharge battery efficiency

References


