

Article

Mitigation of Short-Term Wind Power Ramps through Forecast-Based Curtailment

Oliver Probst ^{1,*}  and Luis I. Minchala ² 

¹ School of Engineering and Sciences, Campus Monterrey, Tecnológico de Monterrey, Eugenio Garza Sada 2501 Sur, Monterrey CP 64849, Mexico

² School of Engineering and Sciences, Campus Guadalajara, Tecnológico de Monterrey, General Ramon Corona 2514, Zapopan CP 45138, Mexico; ismael.minchala@tec.mx

* Correspondence: oprobst@tec.mx

Featured Application: The proposed innovation allows for a significant reduction of large drops in wind power by using only the self-regulating capabilities of a wind turbine or wind power plant, as well as upstream wind measurement instrumentation, which may already be present at the wind farm.

Abstract: As the penetration of renewable energy generation in electric grids becomes more substantial, its contribution to the variability of the net load becomes more noticeable. Particularly in small or weak grids, the rate at which the output power of a wind farm decreases may become a concern to grid operators. In the present work, a novel approach, called forecast-based curtailment (FBC), is shown to be able to self-mitigate downward ramps on short time scales at a very small energy penalty, compared to conventional mitigation schemes, such as flat curtailment or up-ramp limitations. FBC allows to achieve compliance with ramp limits imposed by system operators at a very small energy cost and modest additional upfront investments.

Keywords: wind power variability; short-term forecast; curtailment



Citation: Probst, O.; Minchala, L.I. Mitigation of Short-Term Wind Power Ramps through Forecast-Based Curtailment. *Appl. Sci.* **2021**, *11*, 4371. <https://doi.org/10.3390/app11104371>

Academic Editors: S. M. Muyeen, Junji Tamura, Masaki Yagami, Kenta Koiwa and Rifat Hazari

Received: 2 April 2021
Accepted: 6 May 2021
Published: 12 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

Renewable energy (RE) technologies are becoming increasingly part of the mainstream in the electricity sectors worldwide. As opposed to a few decades ago, large-scale wind and solar photovoltaic power plants are now the cheapest options for generating electricity [1]. This major achievement is finally paving the way towards the deep decarbonization of the power sector. The low cost of wind and solar power also enables the transition of the transport and fuel sectors, both by vehicle electrification [2,3] and the production of green hydrogen [4]. A natural concern regarding variable renewable energy (VRE) technologies relates to the variability of their power output. Though the impact of this variability on power systems operations is often misunderstood or overestimated, there is no doubt that the *increase in variability* of the net system load has to be addressed in power sectors with high and very high VRE penetration. Das et al. [5], in their assessment of the impact of wind energy on the Indian power system, concluded that at a 10% penetration level (calculated on an energy basis) the increase in variability of the net load was hardly noticeable; for penetration levels of up to 30% the increase in variability was of the order of the penetration level or less, an increase which can be comfortably handled with traditional grid management tools. In small systems, however, e.g., island grids [6], or grids tending to become fragmented because of insufficient transmission capacities interconnecting high-resource regions and major load centers, the variability of wind or PV plants may become an issue even at moderate penetrations. This is where grid operators may often require the installation of an energy storage system (EES) [7], typically battery-based (BESS), to provide grid support services, such as frequency response and primary frequency

regulation, and sometimes Automatic Generation Control (AGC). These requirements are often justified on the grounds of the variability of the VRE plant, confusing services *provided by the plant to the grid* with actions designed to mitigate the variability of the VRE plant output itself. This distinction is important, since though the VRE plant contributes to the variability of the net load (generally by increasing it), the marginal increase of one additional plant is often small. In either case, given the discretionary power often invested in grid operators, some restriction as to the tolerable level of variability of the power output is likely to be imposed on individual power plants seeking an interconnection permit. Such a requirement may often be phrased as a maximal negative slope in power output tolerated by the grid.

The main ideas for managing and mitigating wind power variability in general and (negative) ramps in particular can be summarized as follows. Not unexpectedly, most authors consider some kind of storage device (generally battery-based) as the main tool for smoothing wind power output, often proposing specific algorithms for the optimized use of such devices [8–11]. Another line of thought is the overall reduction of wind power variability by an appropriate siting of wind farms, thereby tapping into the complementarity of wind generation at different locations [12,13]. This geographic diversification goes a long way towards smoothing the overall wind power generation in the system; its impact is, however, generally limited to the planning stage, at least if all wind farms are expected to always produce their maximum possible power output. An obvious option for the management of an already built-out fleet of wind farms is to selectively restrict the output of certain farms at given times, in order to “tailor” the output according to pre-established criteria for ramps or variability. This can be done by flat-curtailing a given wind farm [14] (i.e., restricting the output to a maximal value or fraction of the rated output), or restricting the rate of increase (or positive ramp rate) to a certain value. The rationale behind this latter idea is that a positive ramp will be generally followed by a negative one, so restricting the former will on the average reduce the latter. This approach was studied in an empirical manner by Martín-Martínez [15] for the case of the Spanish grid, demonstrating that the general ideas works, albeit at a considerable energy penalty.

A different approach consists in accepting the wind power output as it is, but managing its impact on the grid. This can be done by considering wind power ramps as an explicit variable in the dispatch operations, allowing for sufficient operating reserves to be allocated to account for the possibility of large wind power ramps. Pinto et al. [16] took such an approach by considering ramps as a risk variable in their stochastic unit commitment problem; ramps were not predicted but used as an external variable in their optimization problem. Being able to predict the general ramping characteristics is of course useful for any optimization approach, so the development of suitable wind farm models is important too. Sorensen et al. [17] presented an approach for wind farm modeling based on coherence functions and tested their model on the Danish offshore wind farm Horns Rev. The authors were able to show that their model provided a fair prediction of the spectral characteristics and the 10-min ramp distributions.

While the papers discussed above deal with the mitigation or management of wind power fluctuations, other authors have discussed approaches through which wind farms and turbines can provide frequency support to the grid through adjustments of their active power in response to variations in grid frequency. The main mechanisms include turbine de-loading or de-rating by operating wind turbines above their MPP shaft speed, de-loading through pitching, and energy storage using the DC link of turbines with a full-scale converter [18]; also see Reference [19] for an interesting combination of these mechanisms. As mentioned above, frequency support deals with the ways wind turbines and farms respond to grid fluctuations, rather than with the mitigation of their fluctuations, but the two topics obviously bear a natural relationship.

As discussed above, the variability in wind power output, and particularly negative ramps, can be minimized at the planning stage (by appropriate siting), managed during grid operations (through risk-oriented probabilistic unit commitment), modeled and fore-

cast, mitigated by storage devices and optimal strategies, and reduced by curtailment measures, such as flat curtailment or ramp-rate restrictions. While storage-based solutions require substantial add-on investments (and also imply a certain loss of energy because of finite round-trip efficiencies), curtailment-based options lead to quite substantial energy losses and should be considered as the last resort, restricted to systems with substantial lags in grid upgrades and expansion. An overview of the recent investigation into the management of wind power variability and grid support is shown in Table 1.

A way out of the dilemma described above was proposed by Probst [20] by introducing and developing a new strategy called *Forecast-Based Curtailment* (FBC). The basic idea is to slightly curtail (if possible) the power output of a wind turbine or farm before a large expected fall in power, i.e., a large negative ramp, based on an accurate forecast of the power output for two consecutive time intervals after the last measurement. Analytical expressions for the increased compliance with a given critical ramp rate and the mitigation efficiency were provided for typical power change distributions. The analytical framework included both perfect and non-ideal forecasts, and the results were validated against numerical simulations using both simulated power change data and data from an operating wind farm. The actual capabilities of realistic approaches to short-term forecasting were, however, not subject of the work mentioned. Exploring the suitability of state-of-the-art forecasting methods for the FBC method is the main objective of the present work.

Table 1. Overview of the recent literature on the study and mitigation of wind power ramps.

Authors	Year	Title	Main Objective	Main Contribution	Subject Classification
Probst [20]	2020	A new strategy for short-term ramp rate control in wind farms	Develop a no-storage proposal for negative ramp reduction of wind farms	Analytical framework assessing compliance, considering a finite forecast error	Negative ramp reduction by wind farm self-regulation
Simla et al. [21]	2020	Reducing the impact of wind farms on the electric power system by the use of energy storage	Assess the cycling cost of coal-fired plants due to wind power intermittency	Evaluation of different alternatives with varying degree of storage	Conventional plant cycling/storage
Kazari et al. [22]	2019	Assessing the Effect of Wind Farm Layout on Energy Storage Requirement for Power Fluctuation Mitigation	Simultaneous optimization of wind farm output and battery size	Development and use of stochastic wind farm model to devise optimal wind farm LO	Turbulence/WF modeling/storage
Lyu et al. [19]	2019	Coordinated Control Strategies of PMSG-Based Wind Turbine for Smoothing Power Fluctuations	Smooth power output by de-loading, pitching, and DC-link charging/discharging	Evaluation of the benefit of a hierarchical use of the three strategies	Turbine control
Jin et al. [8]	2019	Dynamic modeling and design of a hybrid CAES & WT system for wind power fluctuation reduction	Model the capability of a CAES system to reduce short-term fluctuations	Thermodynamic analysis of a wind turbine operating with a CAESS	Storage modeling
Lamsal et al. [9]	2019	Output power smoothing control approaches for wind and photovoltaic generation systems A review	Review techniques for reduction of power output fluctuations: wind and PV	Review techniques with and without storage systems	Storage technologies (review)
Musselman et al. [12]	2019	Optimizing wind farm siting to reduce power system impacts of wind variability	Bi-objective optimization to reduce residual demand and wind power variability	Algorithms for siting of new wind farms	Fluctuation mitigation through wind farm siting
Takayama et al. [23]	2018	Study on the ramp fluctuation suppression control wind power generation output using optimization method	Bi-objective optimization to reduce wind power variability and storage size	New grid control method	Grid control with storage and power forecast
Ren et al. [10]	2017	Overview of wind power intermittency: Impacts, measurements, and mitigation solutions	In-depth review of wind power intermittency and mitigating measures	Quantitative measures of intermittency, grid integration costs etc.	Storage technologies (review)
Gong et al. [24]	2016	Ramp Event Forecast Based Wind Power Ramp Control With Energy Storage System	Devise a wind power ramp control method using storage	Optimal control method for ramp control with storage	Ramp control/storage

Table 1. Cont.

Authors	Year	Title	Main Objective	Main Contribution	Subject Classification
Bai et al. [25]	2015	A stochastic power curve for wind turbines with reduced variability using conditional copula	Construct a stochastic on-site wind power curve using copula	Methodology for power construction with decreased variability and obeying BIC	Stochastic power curve construction
Islam et al. [26]	2013	Smoothing of Wind Farm Output by Prediction and Supervisory-Control-Unit-Based FESS	Smooth wind power output using flywheel storage	Supervisory control of a flywheel storage unit	Control of storage devices
Jiang et al. [27]	2013	A Battery Energy Storage System Dual-Layer Control Strategy for Mitigating Wind Farm Fluctuations	Smooth combined wind power/storage output, while optimizing power allocation	Two-layer optimization strategy	Control of storage devices
Martín-Martínez et al. [15]	2013	Analysis of positive ramp limitation control strategies for reducing wind power fluctuations	Reduction of aggregated negative wind power ramps	Several ramp reduction strategies are compared	Fluctuation mitigation strategies
Rahimi et al. [11]	2013	On the management of wind power intermittency	Description of wind power intermittency and reduction approaches (storage)	Descriptive review of hybrid approaches to reduction wind power fluctuations	Storage technologies (review)
Wang et al. [28]	2011	Reduction of Power Fluctuations of a Large-Scale Grid-Connected Offshore Wind Farm	Reduce power output fluctuations after a severe wind speed ramp	PID damping controller for DFIG generators	Turbine control
Tarroja et al. [13]	2011	Spatial and temporal analysis of electric wind generation intermittency and dynamics	Spectral analysis of the effect of aggregation of wind farms on intermittency	Quantitative data-driven analysis based on operating data from S. California	Fluctuation mitigation through aggregation
Hori et al. [29]	2010	Disconnection Control of Wind Power Generators for the Purpose of Reducing Frequency Fluctuation	Disconnect wind farms based on turbulence level	New control scheme, validated with model system	Grid control
Sorensen et al. [17]	2007	Power Fluctuations From Large Wind Farms	Simulate ramps and compare to observational data	Coherence models for wind farm simulation	Turbulence modeling for wind farms

Current methods for forecasting wind speed or wind power can be grouped into the following general approaches [30]: (1) Physical methods with parametric models of the atmosphere [31], (2) statistical methods such times series models [32,33] and artificial neural networks [34], (3) machine-learning approaches with either supervised or interactive learning [35], and (4) hybrid methods combining different techniques in order to perform short-term and medium-term predictions [36].

Among the different approaches, the Kalman filter (KF) method stands out in its ability to provide accurate wind speed forecasts [37]. For instance, Poncela et al. [38] report a recursive wind power forecasting system based on KF, where the parameters of the filter are tuned with the use of an expectation-maximization algorithm. Liu et al. [39] demonstrate the combination of an autoregressive integrated moving average (ARIMA) model of the wind generation process with the KF method for wind speed prediction. An experimental comparison of three forecasting methods for the one-step ahead wind speed based on robust Kalman filtering is reported by Zuluaga et al. [40]. The robust KF methods studied by the authors use the definition of irrational data (wind speed data not physically possible), as well as unnatural data (wind speed data that produce low wind output power), as outliers. These data are not processed in the KF algorithm and are replaced by weighted previous states. Khosravi et al. [34] present a comparison of machine-learning algorithms for predicting wind speed time series with different time scales (5-min, 10-min, 15-min, and 30-min). Although the results are fair for one-step ahead prediction purposes, the computational cost of the algorithms is not discussed. Judging by the reported architecture of the predictors, involving the implementation of multi-layer neural networks with up to 15 neurons per each layer for 10 layers, this approach implies

the use of important computational resources, making it potentially difficult to implement in real-time applications.

In the present work, the first practical demonstration of the FBC method, introduced in Reference [20], will be presented, using KF-based forecasting approaches. In the methodology section, Section 2, Forecast-Based Curtailment will be reviewed, and a simple probabilistic model for the detection efficiency of negative ramp events exceeding a certain threshold is presented. A case will be made for the importance of analyzing the correlation between the changes in power output and the corresponding forecast error. It will be argued that naive forecast methods based on autoregressive techniques and similar approaches typically produce a negative correlation, leading to unsatisfactory results in the prediction of (negative) ramps. In the following subsection, the short-term forecast methods used in this work will be discussed. All methods are based on the Kalman Filter (KF) approach; the main points of distinction lie with the assimilation of exogenous data, mostly stemming from an additional met tower lying upstream of the target tower. In the results Section 3, the effect of correlation on the ability of a forecast algorithm to detect in-compliance events will be discussed. Realistic forecasting results for the site of interest will be shown next. In the final results subsection, practical results of the FBC method, based on site data and the forecasting methods implemented in this work will be presented. The main findings of the work will be wrapped up and discussed in Section 4. Conclusions and suggestions for future work will be provided in Section 5.

2. Methods and Data

2.1. Forecast-Based Curtailment (FBC)

In the following, a brief overview of FBC method will be provided, and some new results, relevant to the present context, will be reported. For a detailed description of the analytical framework of the FBC method, the reader is referred to Reference [20]. We start our discussion by introducing the random variables

$$X = v_n - v_{n-1}, Y = v_{n+1} - v_n, \quad (1)$$

$$X_p = P_n - P_{n-1}, Y_p = P_{n+1} - P_n, \quad (2)$$

which denote the changes in wind speed v and output power P between pairs of time intervals t_{n-1}, t_n , and t_{n+1} , respectively. The FBC method focuses on short-term changes, so typically the time step $\Delta t = t_n - t_{n-1}$ will be of the order of 1 to 10 min. t_{n-1} is the last time step with a measured wind speed value; time steps t_n and t_{n+1} lie in the future, so v_n and v_{n+1} have to be predicted. X and Y have an associated joint or bivariate probability density function (pdf) $f_{XY}(x, y)$. Similarly, a bi-pdf can be defined for the changes in output power: $f_{X_p, Y_p}(x_p, y_p)$. As argued in Reference [20], for short time intervals (1–10 min), the correlation between consecutive changes is generally negligible, and the bivariate pdf therefore, factorizes: $f_{XY}(x, y) = f_X(x)f_Y(y)$. For the 10-min observational data studied in this work, this assumption continues to hold to a good approximation.

The wind speed values are converted to corresponding turbine power values by applying the turbine power curve $P(v)$, i.e., $P_n = P(v_n)$. For the purposes of this work, it will be assumed that $P(v)$ is a monotonically increasing function for the wind speed range $[v_{\text{cut-in}}, v_{\text{rated}}]$ and constant elsewhere ($= 0$ for $v < v_{\text{cut-in}}$, $= P_{\text{rated}}$, for $v > v_{\text{rated}}$). The considerations in this paper can be generalized to stochastic power curves [25]; alternatively, power, rather than wind speed, can be used for the forecasting procedures. However, for the sake of brevity, only wind speed forecasting and monotonic power curves will be considered.

In the following, it will be assumed that negative changes in power can be tolerated by the electricity grid only up to a critical slope value $-\alpha = (dP/dt)_{\text{crit}}$. Events with $dP/dt > -\alpha$ will be referred to as (slope) *compliant*. For the purposes of the discrete formulation, it is convenient to define the positive quantity $a_p = \alpha\Delta t > 0$. The variable of primary interest is the power change in the second time interval, i.e., y_p . If $y_p < -a_p$,

the wind farm or turbine is said to be in compliance; however, often, this in compliance can be avoided by a slight reduction (curtailment) of the output power at t_n , resulting in a less negative power change y_p in the second time interval. Figure 1 shows the general idea of the FBC method.

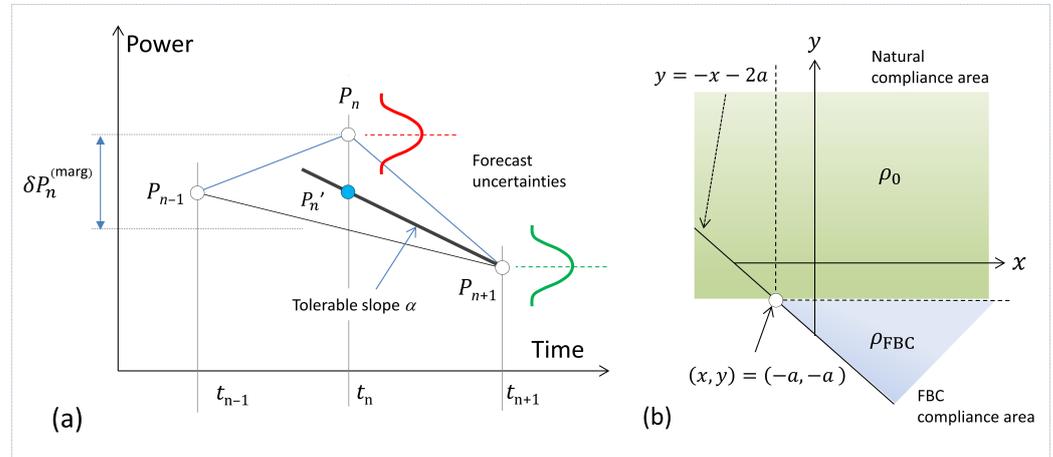


Figure 1. Basic idea of forecast-based curtailment: (a) Three events on the output power time line; P_{n-1} is the last measured event, P_n and P_{n+1} lie in the future. (b) Compliance areas in the x - y -plane, where the $x_p = P_n - P_{n-1}$ and $y_p = P_{n+1} - P_n$.

2.1.1. Ideal Forecasts

In order to illustrate the idea we will first derive an expression for the FBC correction in the case of ideal forecasts, i.e., $\hat{P}_n = P_n$ and $\hat{P}_{n+1} = P_{n+1}$. In the case of an *incompliance event*, i.e., $y_p < -a_p$, the criterion for curtailing power at t_n is for the *average true slope* $\frac{1}{2}(x_p + y_p)$ to be larger than the critical slope $-a_p$. Then the expression for the adjusted power P'_n at t_n becomes

$$P'_n = \begin{cases} P_{n+1} + a_p & \text{if } P_m \geq 0 \\ P_n & \text{if } P_m < 0, \end{cases} \quad (3)$$

where the *curtailment margin* $P_m = \frac{1}{2}(x_p + y_p) + a_p$ has been introduced. Equation (3), applicable to the case $y_p < -a_p$, can be restated in a compact form in terms of the power changes x_p and y_p :

$$x'_p = x_p + (y_p + a)\theta(x_p + y_p + 2a_p), \quad (4)$$

where $\theta(z)$ is Heaviside's theta function. This expression can be easily extended to the case of arbitrary y -values by applying the factor $1 - \theta(y_p + a_p)$. Taking into account that a decrease in x_p translates into a corresponding increase in y_p , we have

$$\begin{aligned} x'_p &= x_p - \Delta x_p = x_p + [(y_p + a_p)\theta(x_p + y_p + 2a_p)](1 - \theta(y_p + a_p)) \\ y'_p &= y_p + \Delta x_p = y_p - [(y_p + a_p)\theta(x_p + y_p + 2a_p)](1 - \theta(y_p + a_p)) \end{aligned} \quad (5)$$

where the positive quantity $\Delta x_p = -[(y_p + a_p)\theta(x_p + y_p + 2a_p)](1 - \theta(y_p + a_p))$ has been introduced. Equation (5) allows for a straightforward generation of a FBC-processed time series for the case of ideal forecasts ($\hat{x}_p = x_p$ and $\hat{y}_p = y_p$). Analytical expressions for the FBC compliance fraction ρ_{FBC}

$$\rho_{FBC} = \frac{N(y_p < -a_p \wedge y'_p > -a_p)}{N_{\text{all}}} \quad (6)$$

and the *natural compliance fraction*

$$\rho_0 = \frac{N(y_p \geq -a_p)}{N_{\text{all}}} \quad (7)$$

were derived in Reference [20] for simple and generalized Laplace-type, as well as Cauchy-Lorentz power change distributions, under the assumption of vanishing correlation between subsequent power changes, and their validity was demonstrated with numerical simulations. $N(\cdot)$ is the number of events fulfilling the logical condition in brackets, and N_{all} is the total number of registered events (power changes). Laplace distributions were found to be a good fit to empirical 1-min power change distributions of an operating wind farm. It is also convenient to introduce the *FBC efficiency* η_{FBC} by

$$\eta_{\text{FBC}} = \frac{\rho_{\text{FBC}}}{1 - \rho_0}, \quad (8)$$

which identifies the degree to which incompliance events are mitigated by the FBC method. In Reference [20], analytical expression for the *FBC curtailment cost*, i.e., the energy not served, were also derived. The energy cost of curtailing wind power based on forecasts for a given critical or *tolerable* slope a was found to be very small, i.e., around two orders of magnitude smaller than the cost for *flat curtailment*, i.e., the permanent reduction of output power to a certain fraction of the potential output power at any given time step.

2.1.2. Considerations for Finite Forecast Accuracy and Correlation

As shown in Reference [20], a finite forecast accuracy can be mitigated quite effectively for the idealized case of vanishing correlation between the wind speed or power changes and the corresponding forecast errors. Then, the true critical slope a can be replaced in the algorithm by an *operational* slope limit $a' = a - \delta a$, resulting in the detection of less negative changes and, consequently, in potentially higher curtailment losses. If, for simplicity, it is assumed that the predicted wind speed changes \hat{x} and \hat{y} are contained in the intervals $[x - \delta a, x + \delta a]$ and $[y - \delta a, y + \delta a]$, then it is easy to show that the true curtailment margin P_m is always positive for predicted ($\hat{y} < -a'$) incompliance events with a positive predicted curtailment margin \hat{P}_m . For simplicity, the subscript “p” has been dropped, as all considerations apply equally to wind speed and power.

One key insight of the present work is, however, the observation that *practical* forecasting techniques for short time horizons, which are often based on some version of time series analysis, generally exhibit some degree of negative correlation between the changes of the variable of interest and the forecast errors. This is particularly notorious for the (trivial) case of persistence (i.e., for the model with $\hat{x}_{n+1} = x_n$), which leads to a perfect negative correlation ($\hat{x}_{n+1} - x_{n+1} = -(x_{n+1} - x_n)$). Less trivial methods obviously improve on this behavior, but the improvement is often less than spectacular.

As we will show below (Section 2.2), negative correlation has a profound impact on the ability of an algorithm to forecast a given event, in the current context one where the wind speed or power change two-steps ahead is more negative than a given threshold value. The practical implication of a negative correlation between the wind speed changes and the forecast error is the fact that the forecast variable does not vary symmetrically around the true value anymore. For negative correlations, we then have $\hat{y} \in [y - \delta a + y_1, y + \delta a + y_1]$, where y_1 is a positive quantity if the true event y is negative. If y_1 is large, then the necessary head room $\delta a + y_1$ for ensuring a positive curtailment margin at incompliance events may become prohibitively large. A practical workaround is to take a somewhat more conservative position by reducing the power change x in the first interval in such that way that the resulting change is equal to the critical slope, i.e., by setting $x' = -a'$. This strategy was used for all results obtained with the FBC algorithm (see Section 3.6).

2.2. Detection Efficiency for Incompliance Events

The probability for detecting an (*incompliance*) event, i.e., a specific instance with $y^{(0)} < -a$, can be written as the *conditional* probability for the predicted value to be smaller than $-a$ as well ($\hat{y} < -a$). Writing $\hat{y} = y + \delta y$, we have

$$p(\hat{y} < -a | y^{(0)}) = \int_{-\infty}^{-a-y^{(0)}} f_{Y, \delta Y}(\delta y | y^{(0)}) d(\delta y), \tag{9}$$

where $f_{Y, \delta Y}(\delta y | y^{(0)})$ is the conditional probability density for observing a forecast error δy , given a true change $y^{(0)}$:

$$f_{Y, \delta Y}(\delta y | y^{(0)}) = \frac{f_{Y, \delta Y}(y^{(0)}, \delta y)}{f_Y(y^{(0)})}. \tag{10}$$

$f_{Y, \delta Y}(\delta y, y^{(0)})$ is the bivariate probability density function (bi-pdf) for variables y and δy , and $f_Y(y^{(0)})$ is the marginal pdf for $y^{(0)}$. In the special case of vanishing correlation between the variables (i.e., $\text{corr}(\delta y, y^{(0)}) = 0$), the conditional probability density simply reduces to the marginal density for δy . In this work, however, we will remain focused on the general case of finite correlation.

We can easily state the *total* probability for detecting an incompliance event by multiplying the conditional probability $p(\hat{y} < -a | y^{(0)})$ (Equation (9)) with the marginal density for $y^{(0)}$ and integrating over $dy^{(0)}$:

$$P(\hat{y} < -a | y < -a) = \int_{-\infty}^{-a} f_Y(y^{(0)}) \int_{-\infty}^{-a-y^{(0)}} f_{Y, \delta Y}(\delta y | y^{(0)}) d(\delta y) dy^{(0)}. \tag{11}$$

Using Equation (10), we can simplify the former expression to read:

$$P(\hat{y} < -a | y < -a) = \int_{-\infty}^{-a} \int_{-\infty}^{-a-y} f_{Y, \delta Y}(\delta y, y) d(\delta y) dy, \tag{12}$$

where the superscript (0) has been dropped for brevity. Therefore, the total probability for detecting an incompliance event is simply obtained from a bi-dimensional integral over the joint or bivariate pdf $f_{Y, \delta Y}(\delta y, y)$ for the variables y and δy . We can also calculate the fraction of *true positives* from

$$r_{TP} = \frac{P(\hat{y} < -a | y < -a)}{P(y < -a)}, \tag{13}$$

where

$$P(y < -a) = \int_{-\infty}^{-a} f_Y(y) dy \tag{14}$$

is the true probability for the occurrence of an incompliance event. Equation (12) allows for a geometric interpretation, which is highly useful for the interpretation of the efficacy of a given forecasting method. For this interpretation, we turn to Figure 2, where the detection region $\mathbb{D}_{TP} = \{(y, \delta y) | -\infty < y \leq -a \wedge -\infty < \delta y \leq -a - y\}$ for true positives has been indicated, together with pictorial representations of bivariate pdfs for uncorrelated (Figure 2b) and negatively correlated variables (Figure 2c), respectively.

It can be seen (Figure 2a) that only events $(y, \delta y)$ contained in \mathbb{D}_{TP} lead to the correct identification of an incompliance event, whereas events in the “lost” area go undetected. It then becomes clear that designing a forecast method is about avoiding a large fraction of the joint pdf for y and δy to fall within this lost area. In Figure 2b, representative contours, say, for 95% of cumulative probability, of the joint pdf have been drawn for the case of comparable variations in the forecast error and the wind speed changes (b.1) and significantly smaller forecast errors (b.2). The events unaccounted for by the forecasting system in case (b.1) are represented by the dark pie slice-shaped area; note that the fraction of undetected events is proportional to the ratio of the dark area and the total area of the circle. It can be seen that even in this relatively stringent case, where the variability of both y and δy is of the same order as a , the fraction of undetected events is relatively small. The situation changes, however, quite drastically if a negative correlation exists between

the two variables, as shown in Figure 2c. In this case, a significant fraction of all events overlaps with the lost area; a large fraction of non-compliance events then goes undetected.

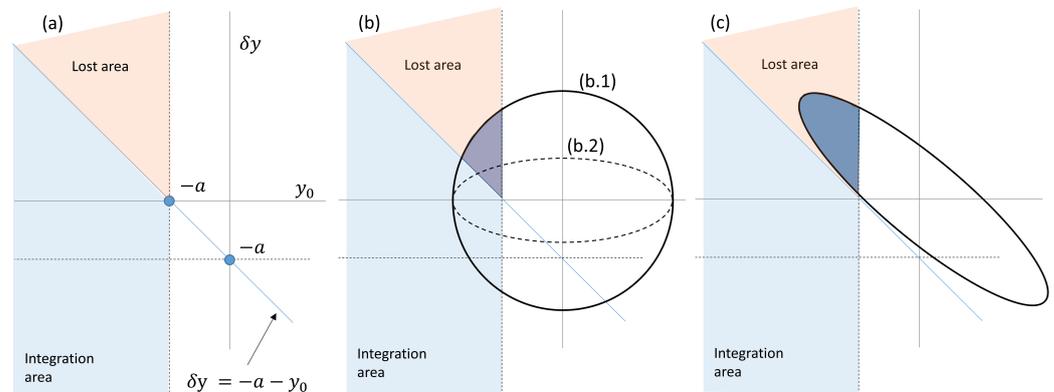


Figure 2. Geometric interpretation of the condition for the detection of non-compliance events ($y < -a$). (a) Detection region \mathbb{D}_{TP} and lost area in the coordinate space spanned by the true wind speed changes (horizontal axis) and the forecast error (vertical axis). (b) Non-correlated joint probability distributions. (b.1) Similar variability of changes and forecast errors. (b.2) Forecast errors about half the size of the variations. (c) Negatively correlated forecast errors and wind speed changes.

In order to conduct a systematic assessment of the combined effect of forecast uncertainty δy and the correlation between y and δy , bi-normal pdfs have been explored first. Though empirical joint pdfs will be shown to be generally more complex, studying bi-normal distributions provides a general feel for the situation and prepares the reader for the discussion of the empirical results. Introducing normalized variables $\tilde{y} = y/\sigma_y$ and $\tilde{\delta y} = \delta y/\sigma_{\delta y}$, the variability ratio $\kappa = \sigma_y/\sigma_{\delta y}$, and the correlation coefficient $\rho = \text{corr}(y, \delta y) = \text{cov}(\tilde{y}, \tilde{\delta y})\kappa$, the joint bi-normal pdf can be written as:

$$f_{\tilde{y}, \tilde{\delta y}}(\tilde{y}, \tilde{\delta y}) = \frac{\kappa}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{\tilde{y}^2 + \kappa^2\tilde{\delta y}^2 - 2\rho\kappa\tilde{y}\tilde{\delta y}}{2(1-\rho^2)}\right) \tag{15}$$

2.3. Forecasting Methods

The prediction methods previously mentioned use the standard formulation of Kalman Filters (KF), which implies that wind speed observations are assumed to have Gaussian noise. KF is a recursive estimation algorithm capable of predicting future states of a system based on a previous state. The standard KF formulation is as follows [41]:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k), \end{aligned} \tag{16}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, and $\mathbf{y} \in \mathbb{R}^p$ are the state vector, the input vector, and the vector of measured output signals, respectively; \mathbf{w} is the vector of the state noise; \mathbf{v} denotes the measurement noise. Note that the letters \mathbf{x} and \mathbf{y} used for the state and the output vectors are not to be confused with the variables x and y (identifying wind speed or power changes for the first and second forecast interval, respectively) of the previous section. Noise vectors are independent Gaussian processes with the following properties:

$$\begin{aligned} E\left[\begin{pmatrix} \mathbf{w}(k) \\ \mathbf{v}(k) \end{pmatrix}\right] &= \mathbf{0} \\ E\left[\begin{pmatrix} \mathbf{w}(k) \\ \mathbf{v}(k) \end{pmatrix} \begin{pmatrix} \mathbf{w}'(l) & \mathbf{v}'(l) \end{pmatrix}\right] &= \begin{pmatrix} \mathbf{Q}_w & \mathbf{Q}_{wv} \\ \mathbf{Q}'_{wv} & \mathbf{Q}_v \end{pmatrix} \delta_{kl}. \end{aligned} \tag{17}$$

The problem to be solved consists in estimating the system’s state from the previously observed values:

$$\begin{bmatrix} y(k - k_0) & y(k - k_0 + 1) & \dots & y(k - 1) & y(k) \\ u(k - k_0) & u(k - k_0 + 1) & \dots & u(k - 1) & u(k) \end{bmatrix}.$$

It is convenient to define *a priori* state estimation $\hat{\mathbf{x}}^- \in \mathbb{R}^n$ at step k , given knowledge prior to step k of the observed values of \mathbf{y} and \mathbf{u} , and a posteriori state estimate at step k , $\hat{\mathbf{x}} \in \mathbb{R}^n$ given the measurement of $\mathbf{y}(k)$. *A priori* and a posteriori errors are defined as follows:

$$\mathbf{e}^-(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)^- \tag{18}$$

$$\mathbf{e}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k). \tag{19}$$

A priori estimate error covariance and a posteriori estimate error covariance are calculated as follows:

$$\mathbf{P}^-(k) = E[\mathbf{e}(k)^- \mathbf{e}^T(k)^-], \tag{20}$$

$$\mathbf{P}(k) = E[\mathbf{e}(k) \mathbf{e}^T(k)]. \tag{21}$$

A formulation for the posteriori state estimation can then be written as:

$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k)^- + \mathbf{K}(k)(\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k)^-), \tag{22}$$

which resembles a state observer. The gain $\mathbf{K}(k)$ is also called blending factor, and its purpose is to minimize the a posteriori error covariance, Equation (21). This optimization procedure, the details of which are not presented here for brevity, can be conducted by substituting Equation (22) into the definition of the error, Equation (19), performing the corresponding expectation operations, taking the derivative of the trace of the result with respect to \mathbf{K} , setting the result equal to zero and finally solving the equation for \mathbf{K} . One particular form of \mathbf{K} that minimizes Equation (21) is given by:

$$\mathbf{K}(k) = \mathbf{P}(k)\mathbf{C}^T (\mathbf{C}\mathbf{P}(k)^-\mathbf{C}^T + \mathbf{Q}_v)^{-1}. \tag{23}$$

Finally, the KF algorithm implies solving Equations (20) to (23) recursively. In this research, we use the standard KF formulation described above to formulate six forecasting approaches, five of which (methods KF1, 2, 3, 5, and 6) assimilate exogenous variables, i.e., wind speed measurements from towers (T9 & T3) other than the target tower. The remaining method (KF4) only uses data from the tower at which future readings are to be predicted (tower T2). The time step of all time series is 10-min. See Section 2.4.

The state matrix, as well as the input matrix of each proposed model, is obtained by data fitting of an autoregressive integrated moving average model with exogenous variables (ARIMAX) as a state space model of the wind generation process. The Akaike Information Criterion (AIC) [42] is used as the evaluation method for data fitting of each ARIMAX model of the six approaches. The KF parameters are tuned by considering a prediction horizon of two samples, $N_p = 2$. The optimization method for selecting these parameters uses the mean-square error (MSE) of each approach as the cost function. The search algorithm for the optimization is the interior-point method.

Table 2 shows a summary of the setup and the nominal performance of each model as measured by its root mean squared (RMS) error value. RMS values are typically used as figures of merit in the forecasting literature; however, as will become apparent below (see Section 2.2), the RMS error is actually not a very good predictor of the capability of a method to forecast wind speed changes one or two steps into the future, as required by the FBC method.

Table 2. KF model details and performance summary ($N_p = 2$). T2 is the target tower, and T3 and T9 are additional towers providing exogenous wind speed data.

Model Details		RMS [m/s]
KF1	$\mathbf{x}_1(k+1) = \mathbf{A}_1\mathbf{x}_1(k) + \mathbf{B}_1 \begin{bmatrix} T9(k) \\ T3(k) \end{bmatrix} + \mathbf{K}_1[T2(k) - \mathbf{C}\mathbf{x}_1(k)]$	0.5534
KF2	$\mathbf{x}_2(k+1) = \mathbf{A}_2\mathbf{x}_2(k) + \mathbf{B}_2T9(k) + \mathbf{K}_2[T2(k) - \mathbf{C}\mathbf{x}_2(k)]$	0.5699
KF3	$\mathbf{x}_3(k+1) = \mathbf{A}_3\mathbf{x}_3(k) + \mathbf{B}_3T3(k) + \mathbf{K}_3[T2(k) - \mathbf{C}\mathbf{x}_3(k)]$	0.5426
KF4	$\mathbf{x}_4(k+1) = \mathbf{A}_4\mathbf{x}_4(k) + \mathbf{K}_4[T2(k) - \mathbf{C}\mathbf{x}_4(k)]$	0.7081
KF5	$\mathbf{x}_5(k+1) = \mathbf{A}_1\mathbf{x}_5(k) + \mathbf{B}_{51} \begin{bmatrix} T9(k) \\ T3(k) \end{bmatrix} + \mathbf{B}_{52}T9(k-1) + \mathbf{K}_5[T2(k) - \mathbf{C}\mathbf{x}_5(k)]$	0.6733
KF6	$\mathbf{x}_6(k+1) = \mathbf{A}_1\mathbf{x}_6(k) + \mathbf{B}_{61}T9(k) + \mathbf{B}_{62}T9(k-1) + \mathbf{K}_1[T2(k) - \mathbf{C}\mathbf{x}_6(k)]$	0.7145

As described in previous sections, the variables of interest to the FBC method are the wind speed changes x and y . The predicted values (\hat{x} and \hat{y}) of these variables were calculated as follows:

$$\begin{aligned} \hat{x} &= \hat{v}_{n,1} - v_{n-1,0} \\ \hat{y} &= \hat{v}_{n+1,2} - v_{n,1} \end{aligned} \tag{24}$$

where $v_{n,1}$ is the wind speed prediction for time step t_n obtained with either of the forecast methods tuned for predictions one time step ahead; similarly, $v_{n+1,2}$ is the prediction for time step t_{n+1} forecast two time steps ahead. $v_{n-1,0}$ is the last measured wind speed.

In addition to the KF methods summarized in Table 2, three *ensemble* methods, combining the predictions of all methods assimilating exogenous data, have been constructed. In those cases, no wind speed time series were constructed. Rather, the indices of detected events $\text{ind}_i = \{\text{index}(\hat{y}) | \hat{y} < -a\}$ by each of the methods ($i \in [1, 2, 3, 5, 6]$) were calculated, and then the union of all detected events was formed: $\text{ind} = \cup \text{ind}_i$.

All methods introduced in this work were coded by the authors, based on the equations exhibited, using MATLAB.

2.4. Site Description and Wind Data

On-site tall tower wind resource data from the development phase of a commercial wind farm in Mexico were used for model construction and validation. All towers at the site were equipped with three pairs of redundant cup anemometers (class I for primary sensors, standard for redundant) placed at 80, 60, and 40 m above ground level. The wind direction was measured at two levels. Data were recorded in 10-min intervals; for each variable the mean, maximum, minimum, and standard deviation were recorded. A subset of three towers was selected for this study. One full year of concurrent information with only minor data gaps was selected to avoid seasonal biases. There were few gaps in general, and those only extend to a few hours in most cases. Apart from missing data, a few stretches of data were eliminated because of poor data quality. Initial quality assurance was conducted in a semi-automatic way using Windographer. Overall data recovery after quality assurance was 99.9 %. Only 80-m wind speed data were used for the purposes of this study.

The relative location of the three towers used in this study (T2, the target tower, T9 the “upstream tower”, and the additional tower T3) are shown in Figure 3a; the wind roses are exhibited in Figure 3b.

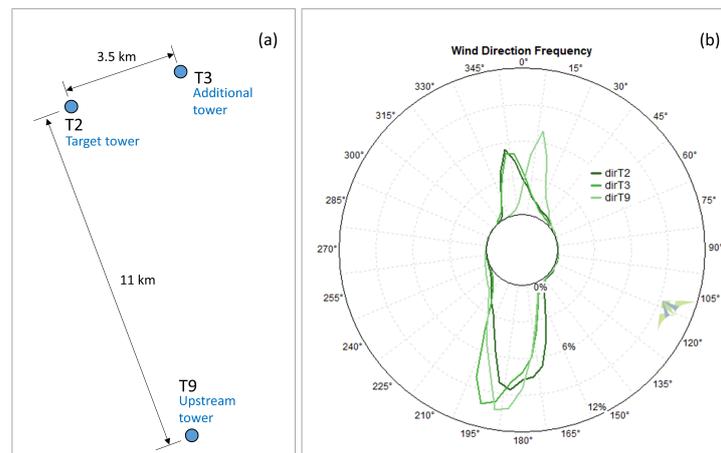


Figure 3. Site characteristics. (a) Relative locations of meteorological towers used in this work. (b) Wind roses from all towers.

3. Results & Discussion

3.1. Detection of Incompliance Events for Bi-Normal Joint Pdfs

As described above, some intuition can be built with regards to the impact of (a) forecast accuracy and (b) correlation between changes in the primary variable (wind speed) and the forecast error by studying bi-normal pdfs (Equation (15)). In Section 2.2, a case was made for the importance of minimizing these correlations in the forecasting process, in addition to the obvious requirement of minimizing the rms forecast error. Using Equation (15), the absolute fractions of *true positives* and their rates relative to the true number of events (the detection rate for true positives) can be calculated as a function of the correlation coefficient ρ and the ratio $\kappa = \sigma_y / \sigma_{\delta y}$ of the standard deviations of the wind speed changes and the forecast error, respectively. Some representative results are shown in Figure 4. The results of the numerical simulations obtained with two normally distributed time series with $T = 5 \times 10^6$ time steps and a given correlation coefficient are shown in open circles, whereas the results obtained from the integration of Equation (12) using Equation (15) are shown as continuous lines. It can be seen that the theory and the simulation produce identical results.

As expected intuitively, the probability of forecasting a large negative change (an incompliance event) decreases with the size of the event (i.e., the tolerable negative slope), and the overall behavior is similar for joint pdfs with negative (Figure 4a) or zero (Figure 4b) correlation, albeit with a varying sensitivity with respect to the standard deviation ratio. Given that the probability of occurrence of such events decreases, as well (not shown for brevity), it is more relevant for the present context to inspect the detection efficiency, again as a function of the size of the event (the tolerable slope). From the comparison of Figure 4c,d, it can be seen that the case with negative correlation (Figure 4c) not only has an overall smaller detection efficiency for a given standard deviation ratio, but also decays much faster as a function of the tolerable slope, i.e., incompliance events with large negative wind speed changes are more difficult to detect, compared to the uncorrelated case. This is a fully consistent with the geometric interpretation offered in Figure 2. The practical implication is that, for detecting small wind speed changes, it is important to reduce the overall forecast error, whereas, for larger incompliance events, it is more important to have a small correlation coefficient. Evidently, this is the more relevant case for practical applications.

3.2. Realistic Forecasts: Exploratory Results

After these conceptual discussions we now turn to actual forecasts obtained for the target site (T2). Figure 5 shows the true wind speed time series, as well as the forecast wind speed obtained with one method assimilating exogenous data (KF1), and the one method in this study using only on-site (local) data (KF4), both for a prediction horizon of two steps into the future. As evidenced by the figure, the local method shows the usual characteristics

associated with autoregressive techniques, a lagging forecast signal, emulating the true wind speed time series, but always trying to catch up. Clearly, such a forecast is of very limited use for detecting future changes, and for acting on them. The non-local method (KF1), on the other hand, can be seen to actually predict the ramps, both upwards and downwards.

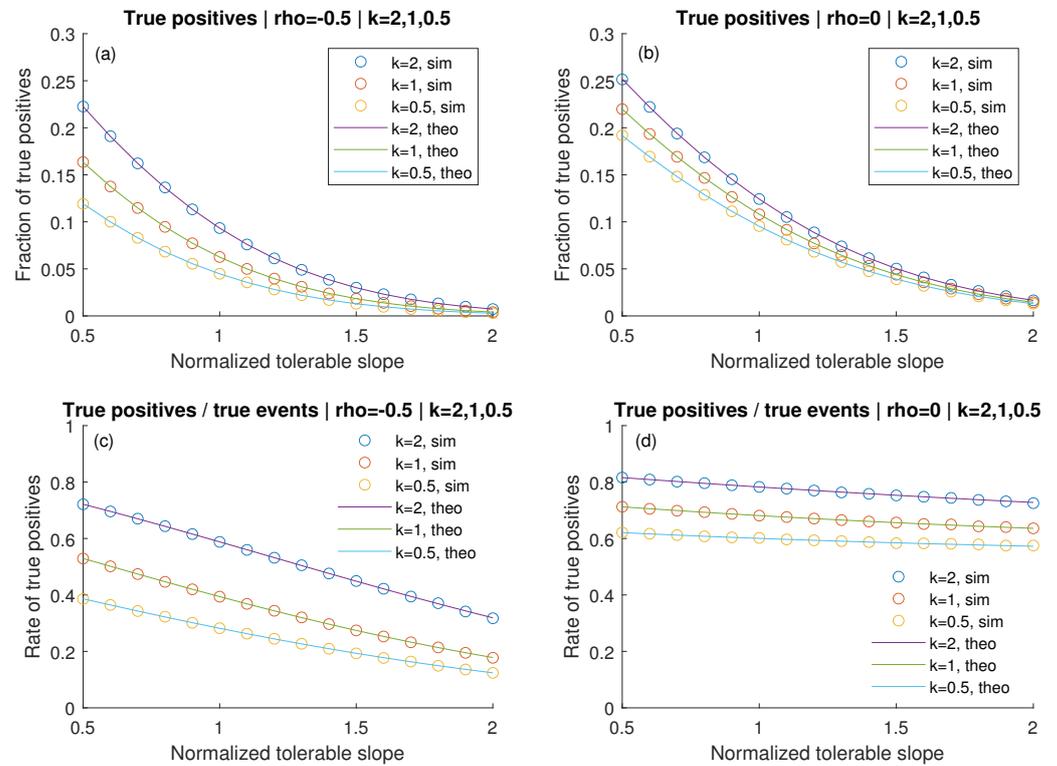


Figure 4. Fraction of true positives observed (a,b) and the respective detection efficiencies (c,d), for different combinations of the correlation coefficient ρ and the standard deviation ratio κ . Left column (a,c): Negative correlation coefficient. Right column (b,d): Zero correlation. Circles: Numerical simulation. Solid lines: Theory.

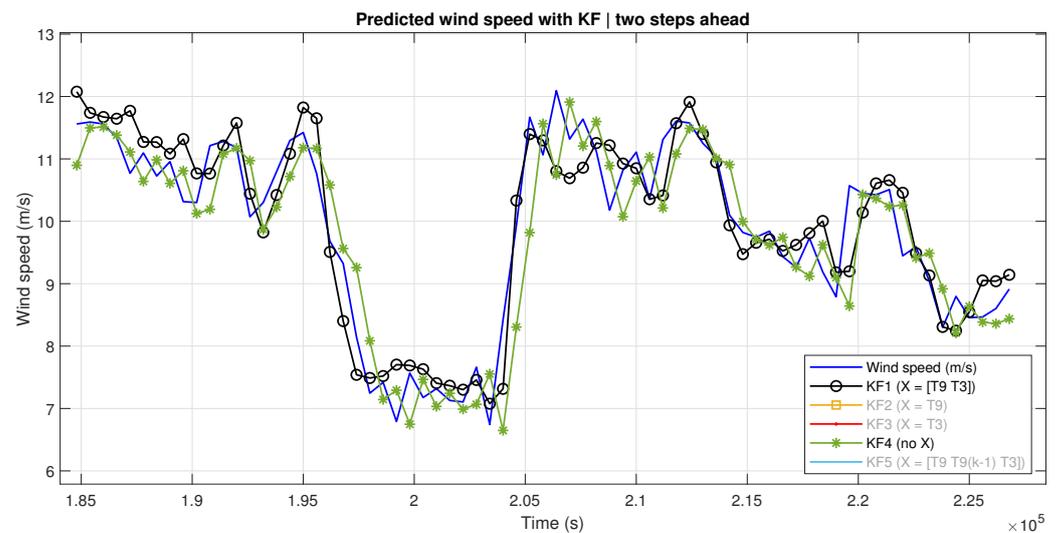


Figure 5. Example time series of the true wind speed, and the forecast wind speed, using one method with exogenous data (KF1) and the one without (KF4); see Table 2.

In order to illustrate how these findings translate into the relevant variables of the FBC method, the two-step ahead wind speed changes (y) and the curtailment margin (P_m), we refer to Figure 6, where results obtained with method KF1 are shown as an example. It can be seen that three out of five incompliance events (taken to be events with $y < -a = -1$ m/s) are correctly identified by the forecast. The curtailment margin ($P_m = \frac{1}{2}(x + y) + a$) can be seen to be positive at all times, and the forecast margin \hat{P}_m is close to the true value at most times, particularly at the detected incompliance events. Therefore, this qualitative inspection suggests that incompliance events can indeed be mitigated, using the FBC method, based on a realistic forecast method.

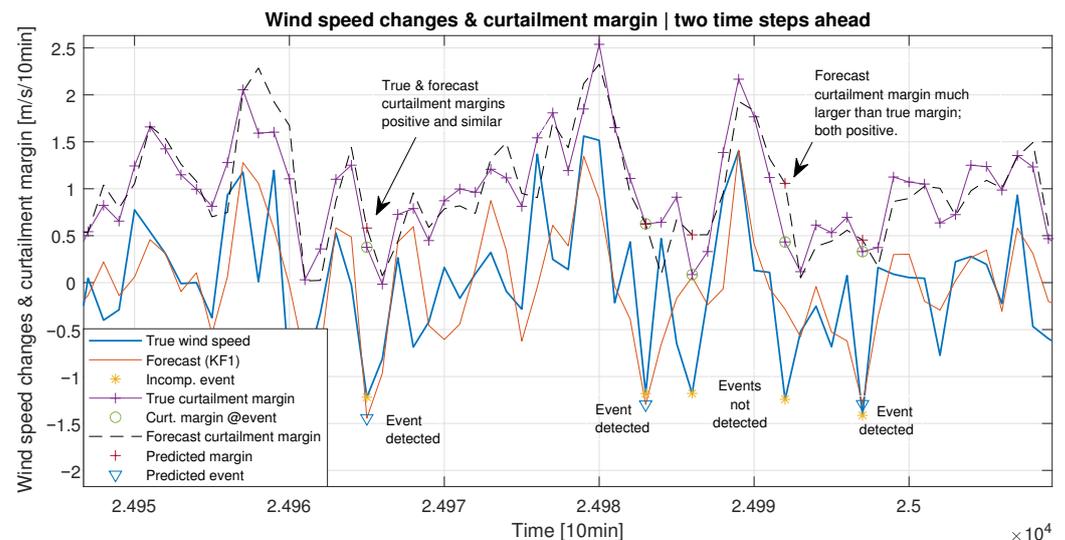


Figure 6. Example time series of the true wind speed changes, the forecast wind speed changes, as well as the true and the forecast curtailment margins, all for two time steps into the future, shown for methods KF1 and KF4; see Table 2. The critical slope was set to $a = 1$ m/s in this example.

3.3. Correlations and Empirical Kernel Density Estimates for Different Forecast Methods

As discussed in Section 2.3, six versions of Kalman Filter-based forecast methods (KF1–KF6) have been implemented in this work. Method KF1, 2, 3, 5, and 6 assimilate exogenous data, whereas method KF4 only uses data from the target site. From Table 2, it can be seen that the methods using exogenous data have similar standard errors (with an rms error of the order of 0.55 m/s), whereas method KF4 has a somewhat higher error (rms = 0.71 m/s), i.e., some 30% higher. While this does point to an inferior performance by method KF4, the rms error does not fully capture the impact on the capability of the method to predict future incompliance events. Based on the considerations described in Section 2.2, it is convenient to inspect the empirical joint probability density function for the wind speed changes y and the prediction error δy for those changes, both for two time steps into the future.

The results for method KF4 and a representative method with exogenous data (KF2) are shown in Figure 7. It can be seen that not only is the forecast error (vertical axis) dramatically smaller for method KF2, but also is there a significant difference in the correlation structure between the two methods. Whereas the bi-pdf of method KF4 is reminiscent of a bi-normal pdf, the probability density for KF2 shows two distinctive lobes or modes, and also a larger concentration of the density near the origin. Both methods show a negative correlation between the wind speed changes and the forecast errors, but the effect is much less pronounced in the case of method KF2. From the comparison of Figure 7 and Figure 2, it can be concluded that a much larger fraction of the events will fall into the “lost” area in Figure 2, i.e., correspond to incompliance events that go undetected. Method KF2, on the other hand, while also showing some overlap with the lost area, mainly because of the lobe with a high negative slope (of around -1), an important fraction of the data falls

onto the lower lobe, with its much less negative slope, and thereby increases the probability for the detection of incompliance events.

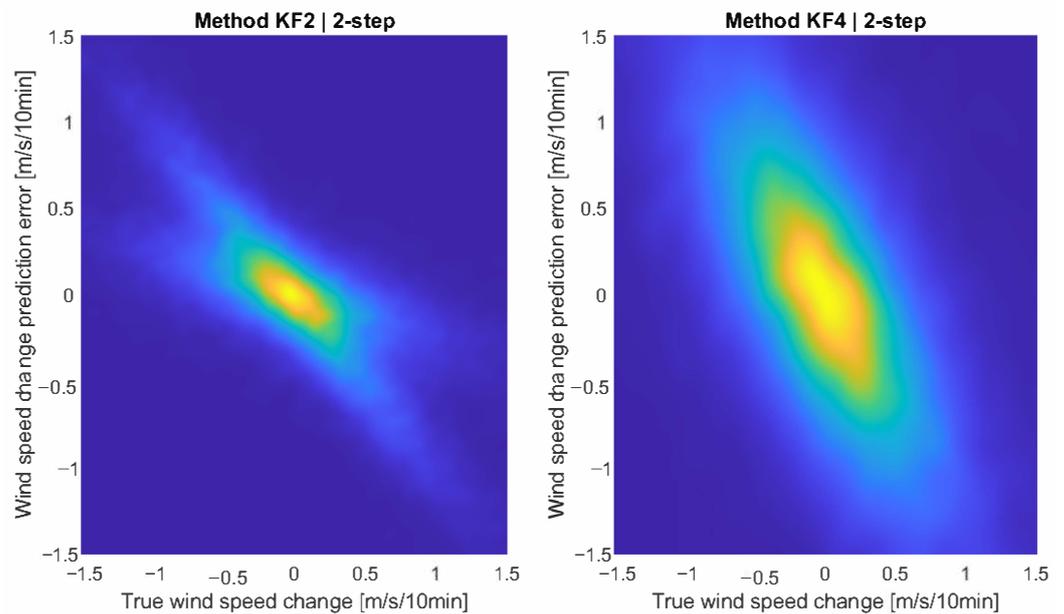


Figure 7. Estimates of the joint probability density functions for variables y (wind speed changes) and δy (forecast error of wind speed changes) for a forecast horizon of two steps into the future. (left) Method KF2, (right) KF4.

3.4. Detection Efficiency and Bivariate Pdfs for Realistic Forecasts

The qualitative findings obtained from the inspection of the joint pdf for y and δy can be translated into quantitative results by integrating the probability density according to Equation (12). The results are shown, again for methods KF2 and KF4, in Figure 8, both of the simulation (dots) and the numerical integration (solid lines) of the empirical joint pdf for y and δy . The results of the simulation and the integration can be seen to be consistent, with some discrepancies occurring at the high a -values, where the number of both the true events and the forecast true positives are small numbers.

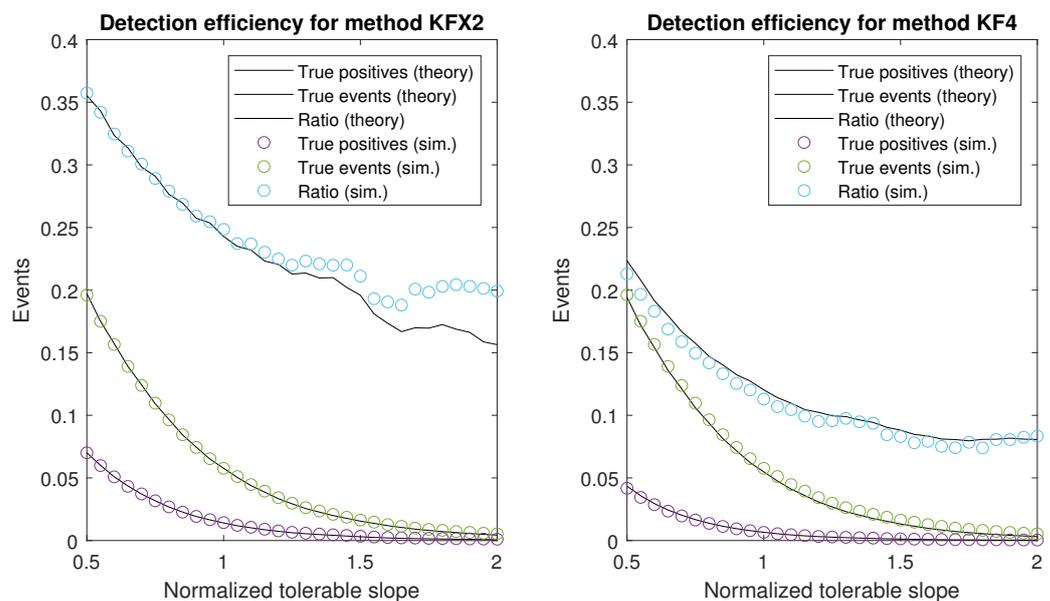


Figure 8. Rates of true events and true positives, as well as their ratio (=detection efficiency) for methods KFX2 (left) and KF4 (right).

As evidenced by Figure 8, method KF2 clearly outperforms the local method KF4, detecting about 20% of all true noncompliance events at large a -values, as opposed to KF4, which only detects about 10%.

3.5. Detection Efficiency and Curtailment Margin Prediction for Different Methods

After discussing the fundamental differences between a forecasting method (KF4) based exclusively on data from the target site and a representative method using exogenous data (KF2), it is interesting to inspect the performance of the different methods introduced in this work. Apart from the true-positive rate (or *detection efficiency*), already introduced in Section 2.2, it is convenient to also keep track of the false-positive rate, as the latter is an indicative measure of the energy cost associated with forecast-based curtailment (FBC).

Figure 9 shows the true-positive (TP) and false-positive (FP) rates for all six KF-based methods introduced in Table 2, as well as the three ensemble methods (see Section 2.3) combining the predicted noncompliance events from certain combinations of exogenous methods. It can be seen (Figure 9a) that all exogenous methods outperform the local method KF4 by far, with KF5 and KF6 showing both the highest detection efficiency for all individual methods and the slowest decay with a , exhibiting an almost constant value of nearly 40%. Interestingly, KF6, the method assimilating exogenous data from only the upstream tower T9 but for two previous time steps, has a much lower false-positive rate (~ 0.1 for KF6 vs. ~ 0.3 for KF5, Figure 9b), with KF6 exhibiting the highest performance ratio (=true/false positives) of all methods at large a -values, being only second to KF2 for small a -values. (Not shown for brevity.)

Combining all exogenous methods into an ensemble approach, a further boost to the true-positive rate is obtained. All three ensemble methods achieve a detection rate of around 50%, with KF12356 being the top-ranking method in term of its efficiency for detecting noncompliance events. In other words, nearly half of all noncompliance events can be detected two time steps ahead, as required by the FBC method. This is quite encouraging. In terms of the cost-benefit ratio, the highest-ranking method is KF1236, given its relatively low false-positive rate (see Figure 9b). It should be noted that, in the context of the FBC method, given the quite low false-positive rates, the overall energy penalty is quite small to begin with, and variations in the true/false positive ratio, therefore, have a small impact on the overall performance of the FBC method.

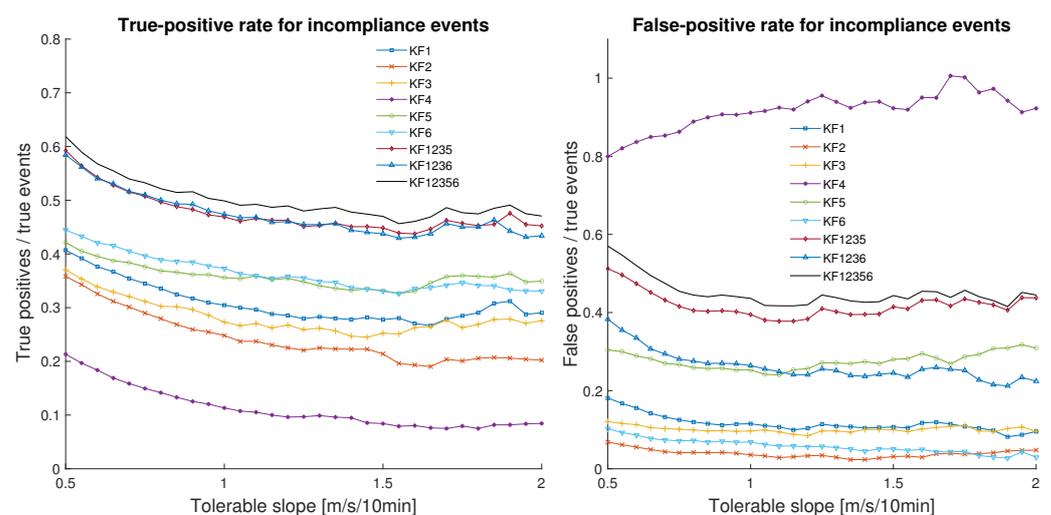


Figure 9. Performance of all forecast methods for the detection of noncompliance events two time steps (2×10 min) into the future. **Left:** True-positive rate vs. the critical slope a . **Right:** False-positive rate vs. the critical slope a .

3.6. Forecast-Based Curtailment

Now that the capability of the different forecast approaches to predict wind speed changes two time steps into the future has been documented at some detail, it can be explored how these results translate into the mitigation of incompliance events through forecast-based curtailment (FBC). Two metrics will be discussed: (1) The FBC efficiency (Equation (8)), and (2) the relative energy cost $\Delta E/E$, where ΔE is the curtailed amount of energy, and E is the uncurtailed energy production for the observation period ($\sim 1a$). Both quantities depend on the critical or tolerable slope a introduced before.

The results can be inspected in Figure 10, together with those of the ideal case, i.e., for perfect forecasts. The ideal FBC efficiency curve can be seen to be nearly flat, with an average value of around 80% and a slight nearly linear increase towards higher critical slope values. While not all incompliance events can be mitigated with the FBC method, it is certainly encouraging to see that a method relying solely on the self-regulation capabilities of a wind turbine should be able to mitigate most of the events that might be of concern to the operator of the electric grid. The associated energy penalty for the ideal case is very small, about 0.2% for a critical slope a of 1 m/s and practically zero for slopes of 2 m/s or higher. (It should be remembered that a is always taken as a positive quantity, and all incompliance events correspond to $y < -a$.) This is particularly noteworthy in the light of the common practice of utilities or independent system operators to require a *flat* curtailment of wind power plants during periods of low regulation reserves (i.e., a restriction to a maximum power output below rated power, or reduction by a constant factor), which may result in very substantial energy penalties. (See Reference [14] for an introduction to the subject and Reference [20] for a systematic comparison between the energy cost of FBC and flat curtailment.)

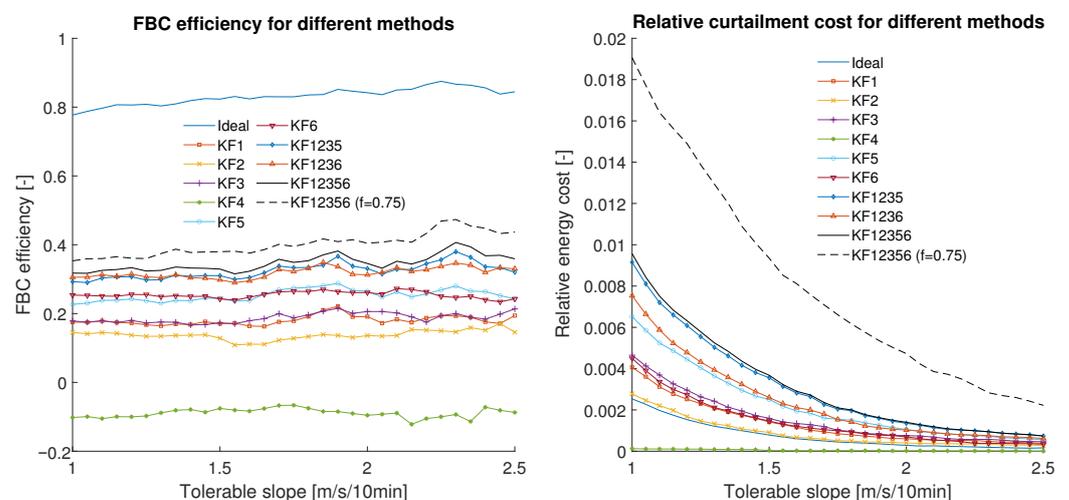


Figure 10. FBC results for all methods. Left: FBC efficiency, right: curtailment cost.

Not unexpectedly, the practical forecast methods implemented in this work do not quite live up to the expectations of the ideal method, but their contributions are still fairly substantial. As before, the ensemble methods constructed from the methods using exogenous data fare best, with an FBC efficiency of around 35%, as shown in Figure 10. The FBC efficiency is nearly independent of the critical slope a , with only a slight increase towards larger a -values. Interestingly, the reference method (KF4), based on measurement data from the target tower alone, not only fails to improve compliance but even decreases it, as demonstrated by its negative FBC efficiency in Figure 10. This is readily explained in terms of the unfavorable ratio between false and true positives of the method; see Figures 9 and 11. Given that the rate for false positives is about one order of magnitude higher than the rate of true positives (which is only about 10%), there are many events that lead to overcorrection. This stresses the need for exogenous data, as successfully demonstrated in this work.

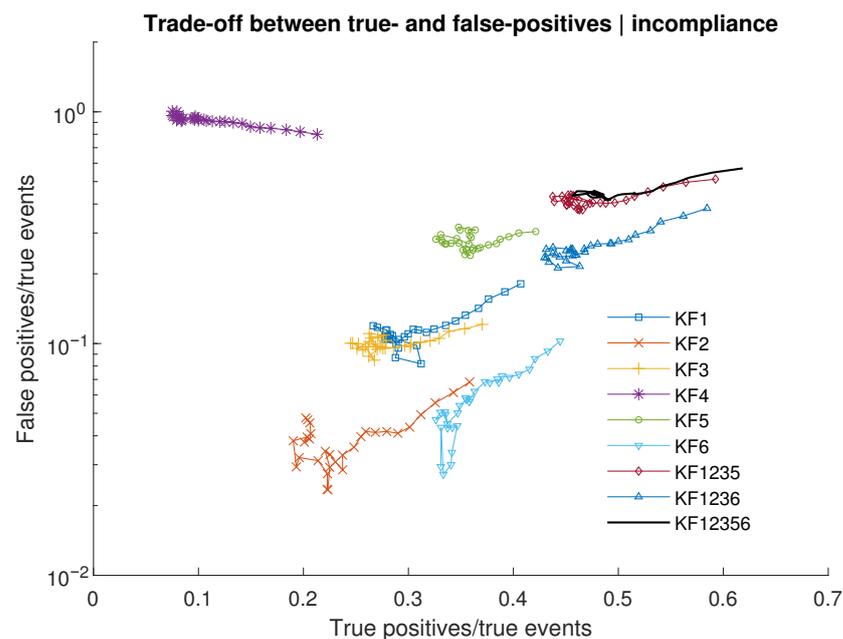


Figure 11. Trajectories of all methods in the true-positive rate—false-positive plane. Note the semi-logarithmic ordinate.

4. Discussion and Practical Implications

As shown in the previous section, the forecasting methods used in this work assimilate exogenous data from two meteorological towers in addition to the target tower to accurately forecast the wind speed one and two time steps ahead; one of the extra towers is located at an upstream position with respect to the target tower, considering the predominant lobe of the bi-modal wind direction distribution at the site. An additional method, using only on-location records from the target tower, was used for reference purposes. The assimilation of exogenous data dramatically boosts the capability of predicting wind speed changes, particularly in the critically important case of two-time-steps-ahead forecasts, compared to the reference method. While the reference method detects only about 10% of large ($\delta ws < -1$ m/s) negative wind speed (ws) changes, the highest-ranking methods with exogenous data detect up to 40% of such events, while producing only about 20% false positives. The detection efficiency for incomppliance events can be further boosted by building *ensemble* methods, which combine the predicted incomppliance events of several exogenous methods, leading to a detection efficiency of about 50%.

The application of these forecasts to the FBC method translates into similar results for the *FBC efficiency* η_{FBC} , defined as the additional compliance achieved by the FBC method, divided by the natural incomppliance fraction, at any given critical slope value. η_{FBC} measures the capability of mitigating naturally occurring incomppliance events. For the case of the target tower studied in this work, the ideal FBC efficiency is around 80%, where it is assumed that the wind speed changes one and two time steps ahead can be predicted perfectly. For the highest-ranking method implemented in this work, based on an ensemble approach, the realistic FBC efficiency is about 35%, i.e., a little below half the ideal value.

The key advantage of the FBC method, as opposed to other curtailment strategies, such as *flat* curtailment (e.g., methods reducing power output by a constant factor or limiting power to a constant maximum) or fixed positive ramp rate limitations, is its very low energy penalty. Particularly at high critical slopes, where compliance matters the most, the energy penalty is of the order of 0.1% for all methods; even for very small tolerable slopes, the curtailment cost is well below 1%. These numbers have to be compared with the tens of percents of energy which are lost by conventional curtailment approaches under similar conditions. Battery energy storage systems (BESS), finally, are well suited to

mitigate negative ramps but represent a significant upfront investment. A brief summary of the methods for ramp control is given in Table 3.

Table 3. Comparison of the FBC method vs. other strategies for short-term ramp control.

Method	Advantages	Disadvantages
FBC (this work)	Very low energy penalty	Requires upstream measurements
Flat curtailment	No instrumentation/investment required	High energy penalty
Positive ramp limit	No instrumentation/investment required	Moderate to high energy penalty
BESS	Low to moderate energy penalty	High upfront investment

The FBC efficiency can be further boosted by activating the FBC algorithm at *operational* critical slope values that are less negative than the true system limit. In the indicative case of triggering curtailment at 75% of the true critical value, the FBC efficiency is increased to over 40% at high slopes, but the energy penalty now increases more drastically (compared to the variations among forecast methods), while still remaining well below 1% for relevant slope values.

As pointed out before, the FBC method critically relies on the capability of forecasting future wind speed changes one, and more importantly, two time steps ahead. In this work, it was argued that this capability is typically limited by the negative correlation between the wind speed changes and the corresponding forecast errors. A simple probabilistic model was developed to demonstrate and explore this effect for idealized joint probability density functions (bi-pdfs) for the changes and forecast errors, and a geometrical explanation was offered. The combined effect of correlation and forecast error was explored with bi-normal pdfs, and the numerical simulations were found to be fully consistent with the theory. Practical bi-pdfs are not bi-normal but show a qualitatively similar pattern for the case of the local forecast method (i.e., the one without exogenous data). Not only is the forecast error distribution much wider for the local method than for the methods with exogenous data, but also is there a strong negative correlation. The exogenous methods, on the other hand, were found to display a bi-modal structure in their joint probability density functions for the changes and forecast errors, with one of the lobes showing a more narrower error distribution. This translates into a much higher detection efficiency for non-compliance events, as was demonstrated by applying the probabilistic model to the empirical bi-pdf.

5. Conclusions and Outlook to Future Work

The present work represents the first practical demonstration of a novel method for mitigating negative ramps in wind turbine power, previously introduced by one of the authors. This method, termed Forecast-Based Curtailment (FBC), relies on the accurate forecast of wind speed or power changes one and two steps into the future; wind speed forecasting was used in the present work. FBC only uses the self-regulating capacity of wind turbines and does not require external storage devices, such as Battery Energy Storage (BES) systems, although a combination with BES systems is possible. The FBC focuses on the detection of *incompliance* events, where the wind speed or power change (or tolerable slope) is more negative than a preset value; this value will typically depend on the requirements of the system operator of the electric grid.

It was shown in this paper that about 40% of all non-compliance events can be mitigated, at an energy cost of only about 0.1%. This mitigation efficiency increases somewhat towards large tolerable slopes (where avoiding non-compliance is more critical), but the increase is small. While a mitigation efficiency for non-compliance events of about 40% as demonstrated in this work, at a negligible energy cost, can be considered a substantial improvement over traditional curtailment schemes, it will depend on the specific case if these mitigation figures are enough or not. An obvious extension of the method is the combination with an energy storage system, such as a BESS, which can be downsized compared to the base case with no ramp mitigation method, leading to a more competitive economic proposal. As an alternative, the BESS could be left at its original capacity, but now being allowed to

provide additional services and increase its value stack. Such options will be explored in follow-up work.

A further improvement of the FBC method may be possible even without building FBC-BESS hybrid systems, by assimilating additional upstream data. In the current work, exogenous data were limited to one upstream tower located at a distance of about 11 km, which is equivalent to a naive (i.e., ballistic time-of-flight) lead time of about 10 min (i.e., one time step) for a wind speed of 8 m/s and two time steps for 4 m/s. This somewhat limits the capability of predicting two time steps ahead. Furthermore, upstream measurements were available only for the (predominant) southern lobe of the wind rose, not for the (minor) northern part. Future improvements of the method will explore forecasting approaches that fully exploit the predictive power of upstream measurements in more generalized wind direction distributions.

Author Contributions: Conceptualization, O.P.; methodology, O.P. and L.I.M.; software, O.P. and L.I.M.; validation, O.P.; formal analysis, O.P. and L.I.M.; investigation, O.P. and L.I.M.; resources, O.P.; data curation, O.P.; writing—original draft preparation, O.P. and L.I.M.; writing—review and editing, O.P.; visualization, O.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge the support from the Energy & Climate Change and the Mechatronics research groups, as well as the open access publishing fund at the School of Engineering and Sciences at Tecnológico de Monterrey.

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

References

1. Lazard's Levelized Cost of Energy and Storage, v14. 2020. Available online: <https://www.lazard.com/perspective/lcoe2020> (accessed on 14 March 2021).
2. Aragon-Aviles, S.; Trivedi, A.; Williamson, S.S. Smart Power Electronics Based Solutions to Interface Solar-Photovoltaics (PV), Smart Grid, and Electrified Transportation: State-of-the-Art and Future Prospects. *Appl. Sci.* **2020**, *10*, 4988. [\[CrossRef\]](#)
3. Marmioli, B.; Messagie, M.; Dotelli, G.; Van Mierlo, J. Electricity Generation in LCA of Electric Vehicles: A Review. *Appl. Sci.* **2018**, *8*, 1384. [\[CrossRef\]](#)
4. Noussan, M.; Raimondi, P.P.; Scita, R.; Hafner, M. The Role of Green and Blue Hydrogen in the Energy Transition A Technological and Geopolitical Perspective. *Sustainability* **2021**, *13*, 298. [\[CrossRef\]](#)
5. Das, P.; Mathuria, P.; Bhakar, R.; Mathur, J.; Kanudia, A.; Singh, A. Flexibility requirement for large-scale renewable energy integration in Indian power system: Technology, policy and modeling options. *Energy Strategy Rev.* **2020**, *29*, 100482. [\[CrossRef\]](#)
6. Miller, N.; Manz, D.; Johal, H.; Achilles, S.; Roose, L.; Griffin, J.P. Integrating high levels of wind in island systems: Lessons from Hawaii. In Proceedings of the 2011 16th International Conference on Intelligent System Applications to Power Systems, Hersonissos, Greece, 25–28 September 2011; pp. 1–8. [\[CrossRef\]](#)
7. Fernandes, D.; de Azevedo, R.; Probst, O. Storage regulations and technologies. In *Transforming the Grid Towards Fully Renewable Energy*; Probst, O., Castellanos, S., Palacios, R., Eds.; Institution of Engineering and Technology: Stevenage, UK, 2020; pp. 179–218. [\[CrossRef\]](#)
8. Jin, H.; Liu, P.; Li, Z. Dynamic modeling and design of a hybrid compressed air energy storage and wind turbine system for wind power fluctuation reduction. *Comput. Chem. Eng.* **2019**, *122*, 59–65. [\[CrossRef\]](#)
9. Lamsal, D.; Sreeram, V.; Mishra, Y.; Kumar, D. Output power smoothing control approaches for wind and photovoltaic generation systems: A review. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109245. [\[CrossRef\]](#)
10. Ren, G.; Liu, J.; Wan, J.; Guo, Y.; Yu, D. Overview of wind power intermittency: Impacts, measurements, and mitigation solutions. *Appl. Energy* **2017**, *204*, 47–65. [\[CrossRef\]](#)
11. Rahimi, E.; Rabiee, A.; Aghaei, J.; Muttaqi, K.M.; Esmael Nezhad, A. On the management of wind power intermittency. *Renew. Sustain. Energy Rev.* **2013**, *28*, 643–653. [\[CrossRef\]](#)
12. Musselman, A.; Thomas, V.M.; Boland, N.; Nazzal, D. Optimizing wind farm siting to reduce power system impacts of wind variability. *Wind Energy* **2019**, *22*, 894–907. [\[CrossRef\]](#)
13. Tarroja, B.; Mueller, F.; Eichman, J.D.; Brouwer, J.; Samuelsen, S. Spatial and temporal analysis of electric wind generation intermittency and dynamics. *Renew. Energy* **2011**, *36*, 3424–3432. [\[CrossRef\]](#)
14. Rose, S.; Apt, J. The cost of curtailing wind turbines for secondary frequency regulation capacity. *Energy Syst.* **2014**, *5*, 407–422. [\[CrossRef\]](#)
15. Martín Martínez, S.; Gómez Lázaro, E.; Viguera Rodríguez, A.; Fuentes Moreno, J.A.; Molina García, A. Analysis of positive ramp limitation control strategies for reducing wind power fluctuations. *IET Renew. Power Gener.* **2013**, *7*, 593–602. [\[CrossRef\]](#)

16. Pinto, M.; Miranda, V.; Saavedra, O.; Carvalho, L.; Sumaili, J. Mitigation in the Very Short-term of Risk from Wind Ramps with Unforeseen Severity. *J. Control. Autom. Electr. Syst.* **2017**, *28*, 247–258. [[CrossRef](#)]
17. Sorensen, P.; Cutululis, N.A.; Viguera-Rodriguez, A.; Jensen, L.E.; Hjerrild, J.; Donovan, M.H.; Madsen, H. Power Fluctuations From Large Wind Farms. *IEEE Trans. Power Syst.* **2007**, *22*, 958–965. [[CrossRef](#)]
18. Attya, A.B.; Dominguez-Garcia, J.L.; Anaya-Lara, O. A review on frequency support provision by wind power plants: Current and future challenges. *Renew. Sustain. Energy Rev.* **2018**, *81*, 2071–2087. [[CrossRef](#)]
19. Lyu, X.; Zhao, J.; Jia, Y.; Xu, Z.; Po Wong, K. Coordinated Control Strategies of PMSG-Based Wind Turbine for Smoothing Power Fluctuations. *IEEE Trans. Power Syst.* **2019**, *34*, 391–401. [[CrossRef](#)]
20. Probst, O. A new strategy for short-term ramp rate control in wind farms. *Int. J. Electr. Power Energy Syst.* **2020**, *120*, 105969. [[CrossRef](#)]
21. Simla, T.; Stanek, W. Reducing the impact of wind farms on the electric power system by the use of energy storage. *Renew. Energy* **2020**, *145*, 772–782. [[CrossRef](#)]
22. Kazari, H.; Oraee, H.; Pal, B.C. Assessing the Effect of Wind Farm Layout on Energy Storage Requirement for Power Fluctuation Mitigation. *IEEE Trans. Sustain. Energy* **2019**, *10*, 558–568. [[CrossRef](#)]
23. Takayama, S.; Ishigame, A. Study on the ramp fluctuation suppression control of wind power generation output using optimization method. *Electr. Eng. Jpn.* **2018**, *205*, 8–15. [[CrossRef](#)]
24. Gong, Y.; Jiang, Q.; Baldick, R. Ramp Event Forecast Based Wind Power Ramp Control with Energy Storage System. *IEEE Trans. Power Syst.* **2016**, *31*, 1831–1844. [[CrossRef](#)]
25. Bai, G.; Fleck, B.; Zuo, M.J. A stochastic power curve for wind turbines with reduced variability using conditional copula. *Wind Energy* **2016**, *19*, 1519–1534. [[CrossRef](#)]
26. Islam, F.; Al-Durra, A.; Mueen, S.M. Smoothing of Wind Farm Output by Prediction and Supervisory-Control-Unit-Based FESS. *IEEE Trans. Sustain. Energy* **2013**, *4*, 925–933. [[CrossRef](#)]
27. Jiang, Q.; Gong, Y.; Wang, H. A Battery Energy Storage System Dual-Layer Control Strategy for Mitigating Wind Farm Fluctuations. *IEEE Trans. Power Syst.* **2013**, *28*, 3263–3273. [[CrossRef](#)]
28. Wang, L.; Chen, L. Reduction of Power Fluctuations of a Large-Scale Grid-Connected Offshore Wind Farm Using a Variable Frequency Transformer. *IEEE Trans. Sustain. Energy* **2011**, *2*, 226–234. [[CrossRef](#)]
29. Hori, Y.; Saitoh, H. Disconnection control of wind power generators for the purpose of reducing frequency fluctuation. *IEEE Trans. Power Energy* **2008**, *128*, 721–727. [[CrossRef](#)]
30. Soman, S.S.; Zareipour, H.; Malik, O.; Mandal, P. A review of wind power and wind speed forecasting methods with different time horizons. In Proceedings of the North American Power Symposium 2010 (NAPS 2010), Arlington, TX, USA, 26–28 September 2010. [[CrossRef](#)]
31. Candy, B.; English, S.J.; Keogh, S.J. A Comparison of the Impact of QuikScat and WindSat Wind Vector Products on Met Office Analyses and Forecasts. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1632–1640. [[CrossRef](#)]
32. Kolumbán, S.; Kapodistria, S.; Nooraee, N. Short and long-term wind turbine power output prediction. Technical Report. *arXiv* **2017**, arXiv:1707.06497.
33. Tian, Z.; Li, S.; Wang, Y. A prediction approach using ensemble empirical mode decomposition? permutation entropy and regularized extreme learning machine for short-term wind speed. *Wind Energy* **2020**, *23*, 177–206. [[CrossRef](#)]
34. Khosravi, A.; Machado, L.; Nunes, R.O. Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil. *Appl. Energy* **2018**, *224*, 550–566. [[CrossRef](#)]
35. Lim, B.; Zohren, S. Time series forecasting with deep learning: A survey. *arXiv* **2020**, arXiv:2004.13408.
36. Shi, J.; Guo, J.; Zheng, S. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3471–3480. [[CrossRef](#)]
37. Hur, S.H. Short-term wind speed prediction using Extended Kalman filter and machine learning. *Energy Rep.* **2020**, *7*, 1046–1054. [[CrossRef](#)]
38. Poncela, M.; Poncela, P.; Perán, J.R. Automatic tuning of Kalman filters by maximum likelihood methods for wind energy forecasting. *Appl. Energy* **2013**, *108*, 349–362. [[CrossRef](#)]
39. Liu, H.; Tian, H.Q.; Li, Y.F. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Appl. Energy* **2012**, *98*, 415–424. [[CrossRef](#)]
40. Zuluaga, C.D.; Álvarez, M.A.; Giraldo, E. Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison. *Appl. Energy* **2015**, *156*, 321–330. [[CrossRef](#)]
41. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME Basic Eng.* **1960**, *82*, 35–45. [[CrossRef](#)]
42. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]