Electric Power Systems Research 211 (2022) 108599

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Adequacy assessment using data-driven models to account for aerodynamic losses in offshore wind generation

Thuy-hai Nguyen [a,*], Jean-François Toubeau [a], Emmanuel De Jaeger [b], François Vallée [a]

[a] Power Systems and Market Research Group, University of Mons, Mons, Belgium
[b] Department of Mechatronic, Electrical Energy, and Dynamic Systems, Université Catholique de Louvain, Louvain, Belgium

ARTICLE INFO

Keywords:
Adequacy
Machine Learning
Offshore wind generation
Wake effects

ABSTRACT

Offshore wind generation has developed rapidly in the past few years, leading to an increasing importance in power systems. Therefore, it becomes essential to properly account for aerodynamic effects that affect the power extracted from the wind, and to assess their impact on the power system adequacy. In adequacy studies, due to computational constraints, the power output of offshore wind farms is currently modelled in a simple and approximate way, neglecting important factors such as turbulence and wake effects. This may lead to erroneous, and thus misleading adequacy estimations. Hence, the focus of this paper is to develop data-driven proxy models able to learn the complex relation between free flow wind information and the aggregated power of wind farms. Those Machine Learning-based models are used as fast and reliable surrogates of numerical simulations based on computational fluid dynamics. The developed models are then included in an adequacy study built upon sequential Monte-Carlo simulations. The obtained outcomes are compared with traditional modelling approaches, which allows to quantify the value of the proposed procedure.

1. Introduction

Offshore wind energy has become an essential part for a large-scale energy transition. However, wind energy is by nature intermittent and uncertain. There are thus growing concerns regarding the reliability of future power systems. One major aspect of the power system reliability assessment lies in its adequacy computation, i.e., the evaluation of the long-term ability to cover the load in steady-state conditions. Currently, probabilistic iterative methods relying on (sequential) Monte-Carlo simulations are used to perform reliable adequacy calculations. However the iterative nature of such simulations may prevent the use of advanced wind farm models to ensure the tractability of the analysis. Aerodynamic effects such as wind shear, turbulence and wake effects, clearly affect the expected power output, and they cannot be disregarded. Wake effects mean that a wind turbine placed behind another one will experience a lower wind speed and thus will produce less power. In modern wind farms, the interaction of wind turbine wakes can cause annual energy losses of 10 to 20% [1]. In the current literature, this aspect is either neglected using a traditional power curve approach or it is modelled in a highly simplified fashion through an efficiency coefficient (typically assumed to be equal to 90%–95% of the total wind farm power [2] or computed using the approximated Jensen wake model [3]). Wake losses are a major issue for offshore wind farms, where the low ambient turbulence accentuates the wake effects. This motivates this work of incorporating them into power system adequacy studies. Specifically, this paper focuses on the inclusion of intra-park aerodynamic effects in offshore wind generation models, which are integrated in Monte-Carlo-based adequacy computations. The challenge is to avoid the high computational cost associated with numerical simulations. This is achieved by using the recent developments in Machine Learning (ML), which is able to capture the complex characteristics of wind generation in a fast and reliable way. Overall, this paper has two main contributions.

Firstly, data-driven proxy models are developed based on ML techniques to improve the way offshore wind parks are represented within adequacy tools. However, instead of using measurements, which are often difficult to obtain and with a limited history (that does not fully capture all possible states), the models are trained using the output of aerodynamic computations. This presents the additional advantage to allow the evaluation of the impact of different projects involving new wind parks to be installed in the power system. For each wind farm, a database is firstly created based on the output of Computational Fluid Dynamics (CFD) simulations, which perform very accurate power predictions but are time-consuming and computationally expensive. This database is then leveraged to train ML models, with the goal of

* Corresponding author.
E-mail addresses: thuy-hai.nguyen@umons.ac.be (T.-h. Nguyen), jean-francois.toubeau@umons.ac.be (J.-F. Toubeau), emmanuel.dejaeger@uclouvain.be (E. De Jaeger), francois.vallee@umons.ac.be (F. Vallée).

https://doi.org/10.1016/j.epsr.2022.108599
Received 3 October 2021; Received in revised form 17 April 2022; Accepted 2 July 2022
Available online 13 July 2022

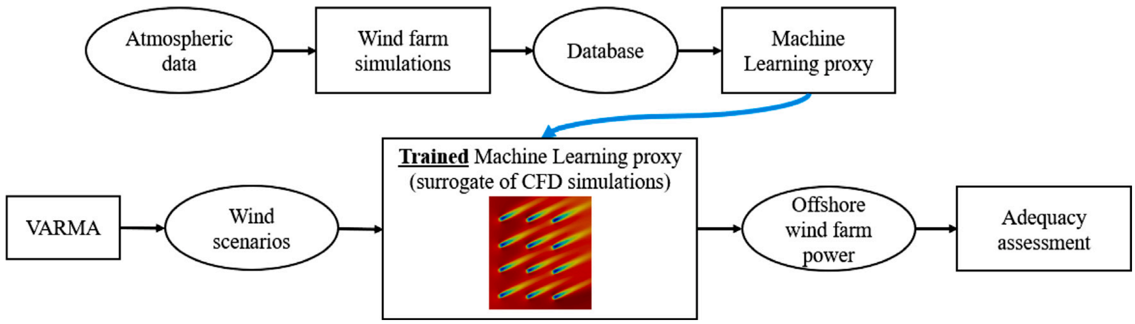

Fig. 1. Methodology scheme.

predicting the output power of offshore wind farms based on free-flow wind information (wind speed and wind direction). Several ML algorithms are trained with the database, and their performance in terms of accuracy, computational time and complexity are compared. In order to validate the ML surrogates, a benchmark consisting of SCADA data from the Alpha Ventus offshore wind farm is set up.

Secondly, the data-driven surrogates are integrated in Monte-Carlo simulations for adequacy assessment. Representative yearly time series of correlated wind speeds and wind directions are generated with a procedure based on vector auto-regressive moving average (VARMA) models. The scenarios are tested on a reliability test system in order to appreciate the impact of an improved offshore wind generation modelling on adequacy results. The current simplified modelling approach for offshore generation is used to benchmark the enhanced models.

The methodology presented in this paper to produce accurate and fast power outputs for a wind farm is summarized in Fig. 1.

The remainder of this paper is organized as follows. Section 2 shows how databases are created from wind farm simulations run with computational fluid dynamics. The methodology is verified in a benchmark using real measurements in Section 2.5. In Section 3, several Machine Learning algorithms are presented and compared based on their prediction accuracy, computational time and complexity. Section 4 shows how to integrate those wind farm models within adequacy studies. A case study is then presented in Section 5 and results are analysed. Finally, main findings and perspectives are summarized in the last section.

2. Wind farm simulations based on CFD

Wind farm numerical simulations cannot be directly integrated in Monte-Carlo simulations. Indeed, in that context, millions of simulated power system states are needed to converge towards reliable outcomes. However, a database can be generated from wind farm simulations, then be used to train fast Machine Learning models (to be embedded in the Monte-Carlo framework) [4].

2.1. Method

There are several ways to run wind simulations. The simplest and less expensive method is to use analytical wake engineering models. In a general form, wake models apply aerodynamic simulations considering mass and momentum conservation principles. However, the equations governing the models rely on many assumptions on aerodynamics, model parameters often need to be tuned (either with measurements or more advanced techniques like CFD) and power predictions deep inside wind farms usually have bias. Even though their speed and simplicity make them attractive to use in the context of this work, their lack of accuracy motivates the use of more advanced techniques.

Computational Fluid Dynamics rely on a set of partial differential equations to solve with initial and boundary conditions, a discrete

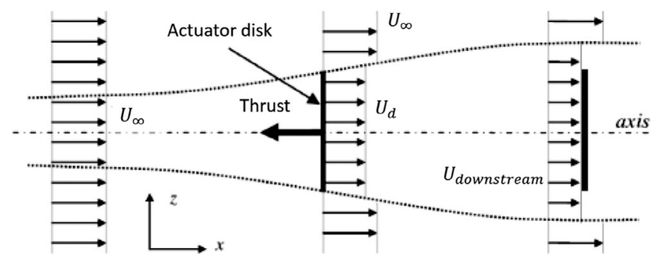

Fig. 2. Actuator disk concept and downstream wind turbine, edited from [7].

representation of the geometry and flow domain (the mesh) and on a numerical procedure (spatial and temporal discretization schemes). For turbulent flow simulations (as it is the case in offshore wind farms), there are different computational strategies with respect to computed length scales. However, to build a relevant database to later train the ML surrogates, thousands of simulations need to be run, for different sets of input parameters. In practice, it is not tractable to run such a number of simulations with very advanced CFD models, as it would take too long to build the database. For example, Large Eddy Simulations [5,6] are able to capture transitory effects and large turbulent scales are resolved. Nonetheless, in the context of adequacy studies, the mean power is needed on an hourly basis. The small fluctuations around the mean value are not necessary (which would not be the case if the power was needed, e.g., to assess the ability of wind turbines to participate to the frequency balancing). This is why Reynolds-Averaged Navier–Stokes (RANS) are used in this paper to build a database linking free-flow wind information with power output. A RANS solver models the averaged turbulent quantities completely so that only mean flow and statistical moments are obtained, in a steady-state simulation. Even though the accuracy is lower than for more advanced CFD solvers (that remain too expensive in time and resources when they are used on large wind farms), it is still more accurate than analytical wake models. Moreover, it is tractable to run hundreds of RANS simulations in a reasonable time, if parallel computing is used.

2.2. Wind turbine representation

The wind turbine geometry is not physically modelled; instead, the rotor forces are represented by an Actuator Disk (AD), shown in Fig. 2.

The wind turbine is then seen as a device that extracts momentum and energy from the wind, uniformly over the rotor area. It acts as a momentum source term in the Navier–Stokes equations. Authors in [8] showed that as long as the AD is subjected to ambient atmospheric turbulence, the averaged velocity deficit calculated by the AD is similar to that from a CFD simulation in which the full rotor geometry is represented. The thrust, i.e., the force exercised by an AD on the external flow, determines the amount of momentum that is extracted from the wind and is therefore very important in wind turbine wake simulations that are modelled with ADs. The determination of the actuator disk forces in multiple wake configurations is not straightforward. In the cases where downstream ADs experience the velocity deficit of upstream ADs, the downstream ADs that are positioned in the full wake of others should experience lower normal and tangential forces compared to those that are subject to the undisturbed flow, as shown in Fig. 2. Therefore, they will produce less power than the upstream ADs. Moreover, determining the power output of waked wind turbines is not trivial because the power extracted by an AD from the wind can be written as:

$$P_{wind\ turbine} = 0.5 * \rho_{air} * \pi * R^2 * C_p(U_\infty) * U_\infty^3 \quad (1)$$

where R is the radius of the wind turbine rotor (and thus the radius of the actuator disk), C_p is the power coefficient and U_∞ is the upstream undisturbed flow velocity at hub height. Eq. (1) assumes that the wind turbine is actively controlled in order to optimize the extracted power without exceeding the maximum allowed power and maximum rotor speed and that the nacelle is aligned with the main wind direction.

However, for a waked wind turbine, the reference velocity U_∞ is not readily known and would require arbitrary decisions about which upstream distance to use when specifying the velocity. Instead, for wind farms, it is useful to base the relations for power on the prevailing axial velocity at the rotor disk position, U_d, such that:

$$P_{wind\ turbine} = 0.5 * \rho_{air} * \pi * R^2 * C_p^*(U_d) * U_d^3 \quad (2)$$

The relationship between C_p and U_∞ for most manufactured wind turbines is readily available in the form of power curves or C_p curves. However, the calibrated coefficient C_p^* as a function of the averaged AD velocity U_d is not directly provided. Therefore, a calibration procedure is carried out in order to determine the value of C_p^* with regard to U_d: single standalone wind turbine simulations are run for $2 \leq U_\infty \leq 30$ m/s with equidistant intervals of 0.5 m/s. From these simulations, it is possible to extract the corresponding U_d and its associated C_p^* [9].

2.3. Modelled wind farm

The modelled offshore wind farm characteristics are based on Alpha Ventus, Germany's first offshore wind farm located in the North Sea and built in 2009 [10]. It consists of 12 wind turbines equally spaced: 6 Adwen AD 5–116 of 5 MW with a diameter of 116 m and 6 REpower 5 M of 5 MW with a diameter of 126 m (the layout can be seen in Fig. 3).

2.4. Numerical setup

In CFD simulations, the mesh should be carefully chosen, as a mesh too coarse leads to approximate results and a mesh too refined is computationally too heavy. In this paper, the computed domain dimensions are 5500 m x 5000 m x 500 m, divided into 182 x 167 x 17 cells respectively in the x, y and z direction. This means that the meshing is roughly equivalent to 4 points per wind turbine diameter. In order to capture the most relevant phenomena around the wind turbines rotors, the mesh is refined in the actuator disks areas: the mesh is then equivalent to 16 points per diameter. This refinement is kept for 10 diameters behind each wind turbine so that wake effects are correctly modelled. Simulations are carried out for wind speeds between 2 m/s and 32 m/s with intervals of 1 m/s, and for wind directions ranging from 0° to 360° by step of 5°. This amounts to a total of 2232 simulations, where each simulation computes the mean hourly power for a given combination of hourly wind speed and direction. This allows to get the power output of each wind turbine for many combinations of wind speed and wind direction.

The RANS simulations are carried out using OpenFOAM, a free open source CFD software [11]. Each simulation is run in parallel on 25 CPUs for approximately 4 min, which amounts to 1.67 CPU hour. Simulations for a wind speed of 8 m/s and wind directions of 0°, 5° and 25° can be seen in Fig. 3. Comparatively with more advanced CFD models, a simulation of one single wind turbine using Large Eddy simulations needs a computation time of 8000 CPU hours [12].

2.5. Verification and validation

Measurements have been collected in the scope of the "Research at Alpha Ventus" project (RAVE) since 2009 by a multitude of sensors on 4 of the 12 turbines in the wind farm, on the substation and within the area of the wind farm itself. The 100-m-high measuring mast Fino 1 is located directly alongside the wind farm, allowing to record meteorological data. Times series from 2011 to 2014 of wind speeds and wind directions from Fino 1 as well as SCADA measurements of electrical power output for one wind turbine in the centre of the wind farm were used as a reference for benchmarking. The measurements were pre-processed and filtered before being used to assess the accuracy of the CFD simulations.

The benchmark focuses on a waked situation, when the second turbine of the row is in the wake of the first turbine of the row. Given the layout of the Alpha Ventus wind farm, this situation arises notably when the wind comes from the west. Therefore, the normalized power of the second turbine is plotted against the wind direction for the wind sector [257°–287°], with a wind speed of 9.5 m/s. It can be seen in Fig. 4 that the RANS simulations show a good agreement with the measurements: they are able to predict the width and depth of the power deficit. To compare with an analytical wake model, the power outputs were also computed with the gaussian model [13]. Results show that the power deficit is overestimated by the model, and that the RANS curve follows the SCADA curve more accurately. The power curve approach does not consider wake effects and the power output of the wind turbine is constant for a given wind speed, independently of the wind direction. Considering that the maximum power loss is approximately ≈50% in full-wake conditions, this clearly emphasizes the limitations of such an approach.

3. Machine learning proxies

Using the database built on CFD simulations, a data-driven model based on Machine Learning algorithms is developed. The dataset consists of 2232 data samples, each representing an hourly value of wind speed and its associated wind direction, with the corresponding hourly power as output. The data are divided into a training set (1256 samples, to fit the model), a validation set (418 samples, to select the best hyperparameters), and a test set (558 samples, to estimate the model performance).

The relationship between the wind power output and raw wind information (wind speed and wind direction) is highly non-linear, which justifies the use of the four following supervised ML algorithms: Decision tree, Random forest, Gradient Boosting Regression Tree (GBRT) and Neural network (more specifically, Multilayer Perceptron).

Decisions trees can be easily interpreted if their size is reasonable. Moreover, they are able to work with features of different scale without a burdensome data pre-processing. However they often overfit and offer thus poor generalization performance. To alleviate this problem, multiple decision trees can be combined in order to decrease the variability of the resulting model.

Random forests are a collection of independent decision trees. Each tree is built on a random subset of features, using a random subsample of the training data set. By averaging the results of all trees, the overall overfitting is highly reduced. One downside is that random forests performance generally increases with the number of trees in the forest, which inherently increases the computational time and memory requirement.

The gradient boosting regression trees algorithm is another ensemble method combining multiple trees. Unlike random forests, gradient

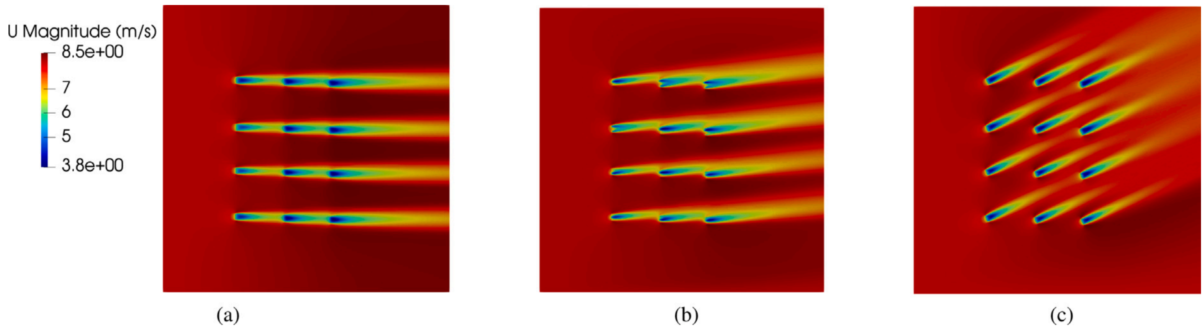

Fig. 3. RANS simulations of the Alpha Ventus wind farm, for a wind speed of 8 m/s and a wind direction of (a) 0°, (b) 5° and (c) 25°.

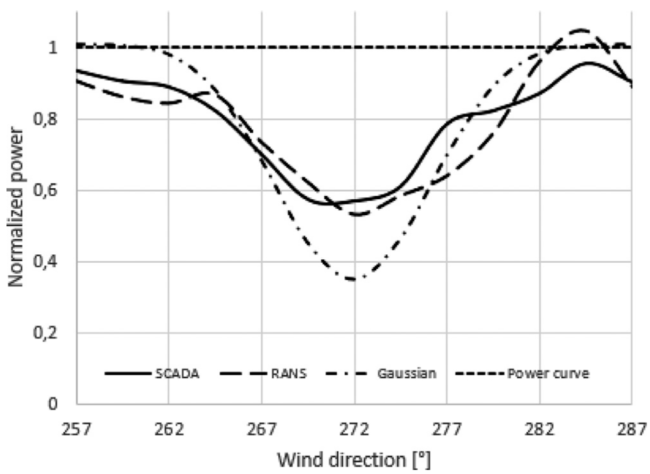

Fig. 4. Comparison of generator power in wake conditions.

boosting aims at constructing trees in a sequential way, and each new tree attempts to rectify the errors made by the previous one. The model is faster in terms of prediction time and needs less memory than random forests because the trees used in this method are not deep. The main drawback of the GBRT method is its high sensitivity to the calibration of hyperparameters, which therefore leads to a complicated fine tuning.

The fourth algorithm used in this paper involves the use of multi-layer perceptrons (MLP) or feed-forward neural networks. A neural network is composed of several processing units (or neurons), connected to each other by learnable weighted connections with the goal of mathematically representing any relationship between inputs and outputs.

The hyperparameters of each of the four considered algorithms are selected based on a compromise between model complexity, prediction accuracy on the validation test and computational time [14]. In this paper, the model accuracy is assessed using the root mean squared error (RMSE), and the mean absolute error (MAE). These metrics are expressed as follows:

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{3}$$

$$MAE = \frac{1}{n} * \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{4}$$

where y_i is the true value of the ith sample, $\hat{y}_i$ is the corresponding predicted value and n is the number of samples in the validation set. It should be noted that lower RMSE and MAE values are associated with more accurate models.

4. Adequacy studies

In the current literature, the more accurate adequacy calculations rely on sequential Monte-Carlo simulations. In practice, Monte Carlo simulations can be used to estimate reliability indices by simulating the actual process and random behaviour of the considered electrical system [15]. The Monte-Carlo sampling process is sequential, i.e., it models all contingencies and operating characteristics inherent to the power system in a chronological way. Sequential simulations are ideally suited to the analysis of intermittent generating sources such as offshore wind generation because they allow the use of detailed hourly generation and load models. The idea is to sample successive system states while keeping the time correlation between consecutive steps. In this paper, scenarios of wind speed and direction are generated, along with scenarios of load and possible failures of conventional generation units.

4.1. Load

Because the simulations are sequential, an hourly load profile describing the evolution of load during the entire year is needed. This profile integrates seasonal trends, diurnal cycle as well as weekday/weekend patterns.

4.2. Conventional generation

Conventional generation units are represented using a two-state model that can be either in up state or down state. Using a random sampling technique from the corresponding state residence time probability distributions, the up-down-up cycle for a yearly sequence can be generated. In this paper, the time to failure (TTF) and time to repair (TTR) are assumed to be exponentially distributed and can be expressed as follows:

$$TTF = -MTTF \ * \ ln\, U \tag{5}$$

$$TTR = -MTTR \ * \ ln\, U' \tag{6}$$

where U and U' are two uniformly distributed random number sequences between 0 and 1, $MTTF$ is the mean time to failure and $MTTR$ is the mean time to repair.

4.3. Offshore wind generation

Two parts compose the offshore generation model: the wind model and the wind turbine generator model.

4.3.1. Wind model

In traditional adequacy computations, only a wind speed model is needed. However, when taking the wake effects into account, the wind direction also plays a significant influence on the power output of wind turbines. Indeed, depending on the orientation of the wind, some wind turbines of the offshore farm can be completely waked (Fig. 3(a)) or partially waked (Fig. 3(b)), while for some directions, the wake effects can be negligible (Fig. 3(c)). When generating wind data, it is important to keep, at a given location, the correlation between the generated wind speeds and wind directions. To that end, we use a Vector Auto-Regressive Moving Average (VARMA) model, which augments the ability of ARMA models with a representation of cross-variables correlations. ARMA models accurately represent time dependencies: each value in the simulated time series depends on its own lagged values (AR part) but also on current and various past values of a stochastic term (AM part).

The application of ARMA models requires the time series to be stationary, i.e., the input should have a constant mean, variance and autocovariance that does not change with time. Given the non-stationary nature of the hourly wind speed distribution due to the daily cycle and seasonality, it is necessary to carry out a transformation of the input time series.

$$y_t = \frac{OW_t - \mu_t}{\sigma_t} \tag{7}$$

where OW_t is the observed wind speed at hour t, μ_t is the mean observed wind speed at hour t, and σ_t is the standard deviation of observed wind speed at hour t. The data series y_t can then be used to build an ARMA time series model.

VARMA(p,q) models used to simulate correlated wind speeds and wind directions are just a generalization of ARMA(p,q) models. The wind speed not only depends on its own lagged values, but also on the lagged values of wind directions. The same goes for the wind direction time series and the wind vector is written as:

$$\begin{aligned} y_t &= \Phi_1 * y_{t-1} + \Phi_2 * y_{t-2} + \cdots + \Phi_p * y_{t-p} \\ &\quad + \alpha_t - \Theta_1 * \alpha_{t-1} - \cdots - \Theta_q * \alpha_{t-q} \end{aligned} \tag{8}$$

where $y_t = [y_s, y_d]^t$ contains the data series corresponding to wind speed and wind direction, and Φ_i and Θ_j are matrices of auto-regressive and moving average parameters of the model respectively. The methodology for estimating the parameters and for choosing the (p,q) order is detailed in [16].

Once the time series model is established, the simulated wind should be computed as follows:

$$SW_t = \mu_t + \sigma_t * y_t \tag{9}$$

4.3.2. Wind turbine generator output

A wind turbine is also subject to outages. In order to consider this aspect, the operating cycle of a wind turbine is simulated using a similar procedure as for conventional generation (see Section 4.2). In particular, the failures are modelled using a probabilistic model, which provides time steps during which wind turbines are unavailable.

4.4. Sequential Monte-Carlo simulations

The simulation procedure for adequacy assessment using Monte-Carlo simulations can be briefly presented as follows:

(1) Model the availability of conventional generating units using chronological simulations
(2) Build a model for the wind power output of the offshore wind farm using the time-series VARMA models and the Machine Learning proxies
(3) Sum the conventional and wind power levels and compute the total generation capacity of the system. Then compare it with the total load
(4) Compute the reliability indices, which are averaged over all generated scenarios

This process is carried out on a yearly basis (8736 h), and repeated until convergence is achieved, i.e., when a specified degree of confidence has been reached. The simulation can then be terminated and the results collected. The stopping criterion used in this paper is:

$$\frac{\sigma(X)}{\sqrt{N} * E(X)} < \epsilon \tag{10}$$

where X is a reliability index, $\sigma(X)$ is the standard deviation, $E(X)$ is the expected value, N is the number of sampling years and ϵ is a convergence threshold (chosen to be 0.01 in this paper). The reliability indices used in this work are the Loss Of Load Expectation (LOLE) [h/year] and the Loss Of Energy Expectation (LOEE) [MWh/year]. The LOLE is the expected number of hours in the year during which the total generation is not enough to cover the load. The LOEE is the expected amount of energy that will not be served during the year [17].

5. Case study

The methodology presented in this paper was applied on a test case in order to quantify the evolution of reliability indices when considering aerodynamic losses in offshore wind farms. An adequacy study is thus conducted using a reliability test system. To validate our methodology, we use the "copper plate" approach: the network constraints are not considered.

5.1. Description of the test case

The Roy Billinton Test System (RBTS) is used as base for the adequacy assessment. The detailed data of the RBTS are presented in [14]. Some offshore wind generation is added to the base system.

5.1.1. Offshore wind generation

The RBTS system is modified with the addition of the modelled offshore wind farm based on Alpha Ventus. The total installed capacity is 60 MW.

The wind database used to construct the VARMA model is a dataset provided by the Royal Meteorological Institute of the Netherlands (KNMI). The KNMI North Sea Wind atlas contains climatological and hourly undisturbed wind data (without wind turbine wake effects) and is based on the ERA-Interim reanalyses dataset, which captures more than 40 years of meteorological measurements and generates 3D wind fields consistent with these measurements and the laws of physics [18]. The wind speeds in the atlas have a climatological accuracy comparable to that of wind measurements (better than 0.2 m/s at wind turbine heights offshore). The hourly time interval is sufficient for this application as the adequacy study is carried out in an hourly way. Using wind data at the location of Alpha Ventus (offshore conditions at sea), a VARMA(5, 4) model is built to generate correlated wind speed and wind direction time series.

5.1.2. Load profile

The RBTS chronological load profile consists of 8736 load points, with an annual peak load of 185 MW. The profile is assembled using data of weekly peak load in percent of annual peak, daily load in percent of weekly peak and hourly peak load in percent of daily peak.

5.1.3. Conventional generation

The RBTS consists of 11 conventional generating units of hydro and thermal type. The size ranges from 5 MW to 40 MW, for a total capacity of 240 MW. The time to failure and time to repair distributions of conventional units are assumed to be exponential.

Table 1
Machine learning models performance.

	Tree	Random forest	GBRT	NN-MLP
RMSE on test set [MW]	0.199	0.176	0.148	0.203
MAE on test set [MW]	0.086	0.067	0.07	0.105
Training time [s]	0.05	0.2	0.1	4.7

5.2. Performance of machine learning proxies

For the modelled wind farm, the four ML models are trained using the database produced with RANS simulations. After a dedicated sensitivity analysis, the following hyperparameters were selected for each algorithm:

- Decision tree: maximum tree depth of 5
- Random forest: 100 trees, a maximum tree depth of 8
- Gradient Boosting Regression Tree: 60 estimators, a maximum depth of 4 and a learning rate of 0.4
- Multi-layer perceptron: 2 hidden layers composed of 10 nodes each, with the use of hyperbolic tangent as activation function

The performances of each ML model are given in Table 1. The computation times needed by the trained models to predict a yearly power output (8736 samples, single CPU) are:

- 0.005 s for decision tree
- 0.057 s for random forest
- 0.015 s for GBRT
- 0.03 s for neural network

The time that would be necessary to produce such an output using RANS simulations is approximately 14 589 CPU hours. The huge difference in computation time between CFD simulations and the ML models clearly justifies the relevance of using such ML models as surrogates of wind farm simulations in Monte-Carlo runs, where hundreds of years are usually needed to reach convergence.

The four ML algorithm show strong performance in terms of training time and prediction accuracy. The decision tree has the fastest training and inference time (used in the Monte Carlo iterations), but its accuracy is slightly lower than the other models. The MLP needs more time to be trained and is less accurate. Both GBRT and random forests exhibit a promising accuracy and a reasonable training time as well, but the inference time is lower for GBRT. Given those observations, the GBRT algorithm is selected to convert the scenarios of wind speed and direction (generated by the VARMA model) into wind power at each iteration of the Monte-Carlo simulations.

5.3. Impact of aerodynamic losses on adequacy indices

Sequential Monte-Carlo simulations were run using the methodology described in Section 4. In order to quantify the impacts of disregarding aerodynamic losses, the offshore wind generation is computed using on one hand the GBRT model, and on the other hand the simple aggregated power curve. The results are shown in Table 2. It can be observed that the reliability indices change significantly: the relative difference reaches approximately 12%, depending on the method used to model the offshore generation. In particular, ignoring the aerodynamic losses caused by wake effects leads to an overestimation of annual offshore energy production of 7.4%. This is a bit lower than what is usually found in the current literature, where average power losses due to wind turbine wakes are of the order of 10 to 20% of total power output in large offshore wind farms. The lower contributions of aerodynamic effects in our case can be explained by the fact that Alpha Ventus is a very small offshore farm wherein wind turbines are quite spaced from one another.

Table 2
Reliability indices, with the ML and the traditional power curve approach.

	LOLE	LOEE	Annual energy
With ML proxies	0.273 h/year	2.162 MWh/year	235.2 GWh
With power curves	0.243 h/year	1.92 MWh/year	254.2 GWh
Relative difference	12.17%	12.49%	7.4%

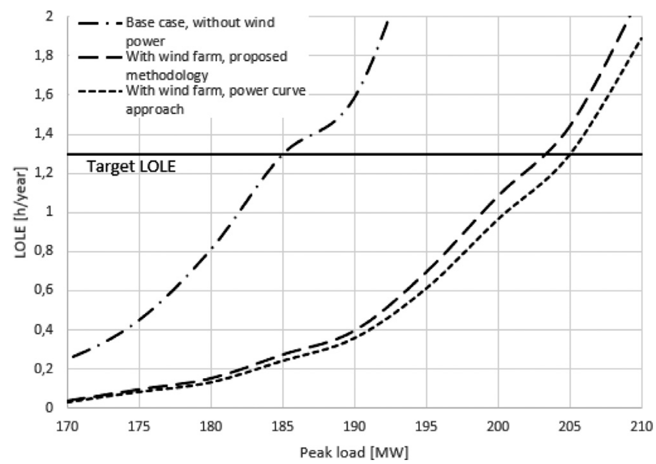

Fig. 5. Evolution of the LOLE when the load varies.

5.4. Sensitivity analysis on the capacity credit of offshore wind

The contribution of wind power to the system adequacy, that is the capacity credit of wind, is estimated by determining the capacity of conventional plants displaced by wind power, whilst maintaining the same degree of system security (i.e., the reliability indices should remain unchanged). In order to so, the reliability indices are first computed for the base RBTS, without any offshore generation. The annual peak load is then varied and the reliability indices are plotted against it. Finally, the wind farm based on Alpha Ventus is added to the system, and the process is repeated.

The effective load carrying capability (ELCC) of the wind farm is the amount of extra load that can be served without degrading a chosen reliability index. The latter is usually the LOLE of the system before the addition of the wind farm. Fig. 5 shows the evolution of the LOLE when the peak load is increased and decreased from the RBTS, with and without offshore generation. Again, the LOLE is computed using both our methodology and the power curve approach. The ELCC corresponds to the peak load difference ΔL between the curves at the target reliability level (computed when the peak load is that of the base system, i.e., 185 MW for the RBTS).

It can be seen that the LOLE reaches the target LOLE when 18 MW of load is added with wake effects taken into account, and 20 MW otherwise. The ELCC of a wind farm is usually expressed as the percentage of the extra load over the added generator's installed capacity (60 MW for Alpha Ventus):

$$ELCC[\%] = \frac{\Delta L}{Capacity} * 100 \tag{11}$$

Therefore, the effective load carrying capability (ELCC) of the offshore wind generation is 30% for our methodology and 33.33% with a simplified method. This means that ignoring the aerodynamic losses in offshore wind farms also leads to an overestimation of the capacity credit of offshore generation, which misleads adequacy results.

The difference between the estimation of the ELCC with the proposed approach and the simplified one could be seen as rather limited in this particular case. Indeed, losses due to wake effects are lower in a small wind farm such as Alpha Ventus, where only a few wind turbines are waked by the upfront ones. This is not the case in larger wind farm, where numerous wind turbines are impacted by others and subjected to deep-array effects. Therefore, this difference in ELCC would be further increased if larger offshore wind farms were considered.

6. Conclusions

In this paper, a new methodology is presented to model the offshore wind generation in a fast and accurate way, in the context of adequacy studies. In particular, a machine learning model is trained to account for aerodynamic phenomena arising from wake effects in the power output of wind farms. This model is integrated in sequential Monte-Carlo simulations. The proposed methodology is applicable to any wind farm, even if it does not exist in real life. Indeed, the input data required to apply the proposed methodology are the layout of wind farm, wind turbine characteristics (hub height, radius, nominal power) as well as correlated hourly wind speeds and wind directions, at the location of the wind farm. These data can be easily gathered, as characteristics of existing wind farms are easily found online and datasets of wind data are widely available as well, as opposed to wind farm output power data. It should be noted that the methodology could be applied to other situations: e.g., it could be used to assess several designs of wind farm in an offshore concession, by computing the expected annual energy. Therefore, it could also bring some insight on the levelized cost of energy by providing an estimation of the expected electrical energy produced over lifetime.

The results of the case study show that improving the offshore wind energy modelling has a large impact on the reliability indices. Indeed, when compared to our method, the traditional power curve approach for wind farms, where aerodynamic losses are ignored, leads to an underestimation of LOLE and LOEE values. The difference in the reliability indices indicate that the adequacy of the considered power system is overestimated. This is problematic because it means that future power systems may not receive the right investments to keep the targeted level of adequacy. This in turn could lead to difficulties to supply electricity in the future. This proves that in the current energy transition context, where offshore wind generation becomes increasingly represented, it is important to improve the way it is considered when assessing the adequacy of power systems.

A possible perspective to this work would be to use larger wind farms with different characteristics and for which no measurements are available (but this would require to run even more RANS simulations). Moreover, several wind farms could be added to the reliability test system, to study inter-park effects. Indeed, in case of neighbouring wind farms, the wind turbines of one park could affect the closest wind turbines of a neighbouring wind farm.

CRediT authorship contribution statement

Thuy-hai Nguyen: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Jean-François Toubeau:** Conceptualization, Writing – review & editing. **Emmanuel De Jaeger:** Supervision, Writing – review & editing. **François Vallée:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS), Belgium under Grant No. 2.5020.11 and by the Walloon Region.

Data was made available by the RAVE (Research at Alpha Ventus) initiative, which was funded by the German Federal Ministry of Economic Affairs and Energy on the basis of a decision by the German Bundestag and coordinated by Fraunhofer IWES (see: www.rave-offshore.de)

This research is supported by the energy transition funds project "PhairywinD, Belgium" organized by the Belgian FPS economy.

References

[1] M.P. van der Laan, N.N. Sørensen, P.-E. Réthoré, J. Mann, M.C. Kelly, N. Troldborg, The k-ϵ-fp model applied to double wind turbine wakes using different actuator disk force methods, Wind Energy 18 (12) (2015) 2223–2240.
[2] N.B. Negra, O. Holmstrm, B. Bak-Jensen, P. Sorensen, Aspects of relevance in offshore wind farm reliability assessment, IEEE Trans. Energy Convers. 22 (2007).
[3] H. Kim, C. Singh, A. Sprintson, Simulation and estimation of reliability in a wind farm considering the wake effect, IEEE Trans. Sustain. Energy 3 (2) (2012) 274–282.
[4] T.-h. Nguyen, J.-F. Toubeau, E. De Jaeger, F. Vallée, Machine learning proxies integrating wake effects in offshore wind generation for adequacy studies, in: International Conference on Environmental and Electrical Engineering, 2021.
[5] Y.-T. Wu, T.-L. Liao, C.-K. Chen, C.-Y. Lin, P.-W. Chen, Power output efficiency in large wind farms with different hub heights and configurations, Renew. Energy 132 (2019) 941–949.
[6] S.-P. Breton, J. Sumner, J.N. Sørensen, K.S. Hansen, S. Sarmast, S. Ivanell, A survey of modelling methods for high-fidelity wind farm simulations using large eddy simulation, Phil. Trans. R. Soc. A 375 (2091) (2017) 20160097.
[7] A.M.I. Nodeland, Wake Modelling Using an Actuator Disk Model in OpenFOAM (Master's thesis), Institutt for energi-og prosessteknikk, 2013.
[8] N. Troldborg, F. Zahle, N.N. Sørensen, P.-E. Réthoré, Comparison of wind turbine wake properties in non-uniform inflow predicted by different rotor models, 555, (1) 2014, 012100, IOP Publishing,
[9] M.P. van der Laan, N.N. Sørensen, P.-E. Réthoré, J. Mann, M.C. Kelly, N. Troldborg, K.S. Hansen, J.P. Murcia, The k-ϵ-fp model applied to wind farms, Wind Energy 18 (12) (2015) 2065–2084.
[10] RAVE: REsearch at ahlpa ventus, 2022, [Online]. Available: https://rave-offshore.de/en/data.html last access: April 2022.
[11] ESI opencfd release openfoam v2106, 2021, [Online]. Available: https://www.openfoam.com/news/main-news/openfoam-v2106 last access: October 2021.
[12] N. Sedaghatizadeh, M. Arjomandi, R. Kelso, B. Cazzolato, M.H. Ghayesh, Modelling of wind turbine wake using large eddy simulation, Renew. Energy 115 (2018) 1166–1176.
[13] A. Niayifar, F. Porté-Agel, Analytical modeling of wind farms: A new approach for power prediction, Energies 9 (9) (2016) 741.
[14] J.-F. Toubeau, J. Bottieau, F. Vallée, Z. De Grève, Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets, IEEE Trans. Power Syst. 34 (2) (2018) 1203–1215.
[15] F. Vallee, J. Lobry, O. Deblecker, System reliability assessment method for wind power integration, IEEE Trans. Power Syst. 23 (3) (2008) 1288–1297.
[16] R. Billinton, H. Chen, R. Ghajar, Time-series models for reliability evaluation of power systems including wind energy, Microelectron. Reliab. 36 (9) (1996) 1253–1261.
[17] R. Billinton, R.N. Allen, Reliability Evaluation of Power Systems, second ed., Plenum Press, Reading, MA, 1996.
[18] KNW Data, 2022, [Online]. Available: https://www.knmiprojects.nl/projects/knw-atlas/knw-data last access: April 2022.