

Article

Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast

Lorenzo Donadio , Jiannong Fang *  and Fernando Porté-Agel * 

Wind Engineering and Renewable Energy Laboratory (WiRE), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland; lorenzo.donadio@epfl.ch

* Correspondence: jiannong.fang@epfl.ch (J.F.); fernando.porte-agel@epfl.ch (F.P.-A.)

Abstract: In the past two decades, wind energy has been under fast development worldwide. The dramatic increase of wind power penetration in electricity production has posed a big challenge to grid integration due to the high uncertainty of wind power. Accurate real-time forecasts of wind farm power outputs can help to mitigate the problem. Among the various techniques developed for wind power forecasting, the hybridization of numerical weather prediction (NWP) and machine learning (ML) techniques such as artificial neural networks (ANNs) are attracting many researchers world-wide nowadays, because it has the potential to yield more accurate forecasts. In this paper, two hybrid NWP and ANN models for wind power forecasting over a highly complex terrain are proposed. The developed models have a fine temporal resolution and a sufficiently large prediction horizon (>6 h ahead). Model 1 directly forecasts the energy production of each wind turbine. Model 2 forecasts first the wind speed, then converts it to the power using a fitted power curve. Effects of various modeling options (selection of inputs, network structures, etc.) on the model performance are investigated. Performances of different models are evaluated based on four normalized error measures. Statistical results of model predictions are presented with discussions. Python was utilized for task automation and machine learning. The end result is a fully working library for wind power predictions and a set of tools for running the models in forecast mode. It is shown that the proposed models are able to yield accurate wind farm power forecasts at a site with high terrain and flow complexities. Especially, for Model 2, the normalized Mean Absolute Error and Root Mean Squared Error are obtained as 8.76% and 13.03%, respectively, lower than the errors reported by other models in the same category.



Citation: Donadio, L.; Fang, J.; Porté-Agel, F. Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast. *Energies* **2021**, *14*, 338. <https://doi.org/10.3390/en14020338>

Received: 17 December 2020

Accepted: 6 January 2021

Published: 9 January 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 is promoted worldwide as the core technology for the mitigation of climate change, which is one of the global challenges faced by humanity. That is because renewable energy sources such as wind and solar do not emit greenhouse gases that contribute to global warming. In the past two decades, wind energy has been under fast development in many countries. The cumulative wind power installed is expected to reach 908 GW by the end of 2023 according to the 2018 Annual Report of Global Wind Energy Council (GWEC). However, the dramatic increase of wind power penetration in electricity production has posed some new challenges to grid integration due to the intermittent and non-schedulable nature of wind power: it varies with weather conditions, and is not always available when electricity is needed. Accurate real-time wind power forecasting can help to mitigate the problem, especially to reduce the integration cost [1]. The applications of forecasts targeted at various time scales include wind turbine regulation and control (0–30 min), load dispatch planning (30 min–6 h), reducing the imbalance penalties for wind

farm operators and determining the reserve capacity to ensure grid stability (6 h–1 day ahead), and maintenance scheduling and wind farm designing (>1 day ahead) [2]. Overall, forecasting plays an important role in the integration of renewables into a sustainable and resilient energy system.

Wind power forecasting models can be clustered into three main groups: physics-driven, data-driven, and combined. Detailed classification and review of various methods can be found in references [2–4]. In short, the first group is based upon forecasted values from Numerical Weather Prediction (NWP) models and utilizes physical models/considerations such as local terrain, wind farm layout and power curve to reach power estimates at specific sites. The second group applies statistical methods to find the relationships between a wealth of explanatory variables (including time series data from meteorological stations, online measured data from wind farms, and NWP results) and the power to be predicted. The third group tries to combine different strategies or models with the idea of improving model performance by retaining advantages of each approach. The models in the first group usually need considerable computation times, hence, only suitable for applications of time horizons longer than 6 h. The models in the second group are applicable for all time horizons. Nevertheless, they are subject to the availability of long enough historical data.

Among the various techniques adopted in data-driven wind power forecasting models, machine learning (ML) techniques are attracting many researchers world-wide nowadays. This is because they are capable of finding complex nonlinear relationships between inputs and outputs. Some studies have shown that ML-based models outperform conventional statistical methods (i.e., time series approaches) in which the inputs are based on time series data from measurements [5–8]. It is known that, with the increase of time horizon, the inclusion of NWP inputs becomes more and more important, especially for time horizons longer than 6 h. In recent years, research works on developing NWP+ML models for wind power forecasting have emerged [9–16]. Mostly, the role of ML can be regarded as statistical downscaling, which downscale NWP results into power through a transfer function trained with historical data. Compared to physical downscaling which performs higher resolution simulations by adding microscale physics to capture more local conditions, ML-based downscaling is computationally more efficient and has the potential to yield more accurate wind power forecasts because, unlike physical downscaling, it is not subject to various systematic errors due to physical assumptions and simplifications. Ramirez-Rosado et al. (2009) [9] compared two forecasting systems in which NWP results were enhanced by various artificial neural networks (ANNs) to provide forecasted values for horizons beyond 12 h, and demonstrated significant improvements over the persistence model in a case study of a wind farm situated in a mountainous region in the north of Portugal (e.g., for the wind farm power-generation forecasted with horizons between 12 and 24 h, the normalized Root Mean Square Error (RMSE) is 14% for the best performing NWP+ML model and 31.2% for the persistence model). Vaccaro et al. (2012) [10] proposed a framework for one-day-ahead wind power forecasting, which combines multiple physical forecasting models and an adaptive machine learning technique called the local Lazy Learning. Experimental results for a wind farm showed that the normalized Mean Absolute Error (MAE) of the power forecasts obtained by the framework is about 13%. The paper of Zhao et al. (2012) [11] reported the development and application of a day-ahead wind power forecasting system in China, which mainly consists of a NWP model and an ANN model. Results from a real case validated the effectiveness of the forecasting system with a normalized RMSE of 16.47%. Xu et al. (2015) [12] introduced a NWP data adjustment approach which first identifies and clusters the errors of NWP using data-mining techniques, and then adjusts the abnormal raw NWP data before it is sent to NWP+ML models. Tests using the data of wind farms from Global Energy Forecasting Competition 2012 [17] demonstrated that the proposed approach can effectively reduce the normalized RMSE by 2–3%. Men et al. (2016) [13] proposed to use an ensemble of mixture density neural networks to quantify the uncertainties arising from both struc-

ture and output of a NWP+ML model. An application of the proposed methodology for a selected wind turbine in a wind farm in Taiwan demonstrated that it works for the purpose of uncertainty quantification, but the errors of deterministic forecasts are high (e.g., the normalized RMSE is 26.8%). Mana et al. (2017) [14] evaluated two NWP+ML approaches: a pure ANN method and a hybrid method based on the combination of ANN and microscale computational fluid dynamics. Power forecast results for a wind farm sited at an Italian mountainous site with complex terrain conditions showed that pure ANN performs better than the hybrid method as far as the RMSE is concerned and the errors depend on the layouts (e.g., the normalized RMSE of pure ANN is about 25% for layout 1 and 21% for layout 2). In a recent paper of Mana et al. [15], the authors discussed two different techniques of downscaling the NWP data for wind farm power forecast: a purely deterministic method based on microscale computational fluid dynamics and a purely statistical approach based on ANN. It was shown that, for a test case of moderate terrain complexity, the deterministic forecast, appropriately tuned with the NWP nodes selection, performs as good as the ANN-based forecast with the normalized MAE around 10% and RMSE around 17%. Ref. [16] describes an integrated wind and solar energy forecasting system which is developed for the Shagaya Renewable Energy Park in Kuwait by blending physical models with artificial intelligence. Testing results showed that the normalized RMSE for wind power forecasts at horizons between 12 and 24 h is about 20%.

It is known that, among the various NWP+ML models developed, it is difficult to draw conclusions as to which model is the best because each model is limited to a specific site. The site features such as topography and wind farm layout are different and the datasets are non-identical. The choice of the NWP model and the prediction length could also be different. Thus, a NWP+ML model developed at its site, in general, cannot have the optimal performance when applied to the other regions.

This paper presents two hybrid NWP and ANN models for wind power forecasting over a highly complex terrain. The developed models have a fine temporal resolution and a sufficiently large prediction horizon (>6 h ahead) to be used in the aforementioned relevant applications. Model 1 uses the mesoscale Weather Research and Forecasting Model (WRF) predictions coupled with ANNs to directly forecast the energy production of each wind turbine. An intermediate calculation of the wind speed is not necessary, in this way the use of power curves is evaded. Model 2 forecasts first the wind speed, then converts it to the power using a power curve derived from on-site measurements. The case study is the Juvent wind farm in the Jura mountains in Switzerland, composed of sixteen turbines as shown in Figure 1. The site features the combined presence of three complexities: topography, heterogeneous vegetation varying from grassy to forested, and interactions between wind turbine wakes. Hence, physics-based wind power forecasting at such a site is very challenging [18].

In this research work, the performance of each model is optimized by systematically varying the combination of modeling options (selection of inputs, network structures, and parameters) and choosing the best performing one. Performances of different models are evaluated based on four normalized error measures with statistical results. Important factors influencing model performance in the context of complex terrain are discussed. Python is utilized for task automation and machine learning. The end result is a fully working library for wind power predictions and a set of tools for running WRF in forecast mode and requesting data from the Juvent wind farm in real time for online performance tuning and retraining if necessary.



Figure 1. Location of the case study site in the Swiss Jura mountains and numbering of the turbines in the Juvent wind farm (source: www.juvent.ch).

2. Materials and Methods

2.1. Prediction Pipeline

The goal of this research work is to predict wind power production from WRF results using machine learning. As is best practice with machine learning projects, one must establish a reliable pipeline so the prediction can be run systematically, and thus provide a real-time forecast.

The prediction pipeline (Figure 2) begins with the Global Forecast System (GFS), a global numerical weather prediction model, which is run by the United States' National Weather Service. The mathematical model is run four times a day, at 00, 06, 12 and 18 h UTC time, and the output data is available online. The relevant GFS data for our project consists of *grib2* files that are used by the WPS (WRF pre-processing) process, to build the boundary and initial conditions for the WRF forecast. The WRF model setup uses a triple nesting technique and provides a forecast with a horizontal spatial resolution of 3.3×3.3 km at 10 min intervals inside the third domain. Simulations are run in parallel at the cluster. To reduce the size of the WRF data to be transferred to the local computer, the post-processing step extracts a set of relevant variables, namely temperature, pressure, wind direction and speed from the raw WRF output. These variables are located in a grid of points around the wind farm, as illustrated in Figure 6. Such variables are extracted by interpolation at four different heights: 50, 95, 150 and 200 m above the ground surface.

After the processes that take place in the remote cluster are finished, the output file from the post-processing is copied to the local computer for the prediction step of the pipeline. The data are prepared and then used as an input for the pre-trained Artificial Neural Networks (ANNs). The final output of the pipeline is the power prediction generated for the next 12 h. The ANNs were trained using archived Supervisory Control And Data Acquisition (SCADA) data from the Juvent wind farm, as well as WRF data. The SCADA data contains measured meteorological variables such as temperature, wind speed and direction as well as the target variable: generated power. Only the simulated meteorological variables from WRF are used as predictors, as detailed in Section 2.5.

2.2. WRF

The Weather Research and Forecasting model (WRF) is one of the most popular numerical weather prediction systems. The WRF model is adopted here as the meso-scale model, using a triple nesting technique, with one-way interactions, as shown in Figure 3. For this particular project WRF is used to the fullest of its capabilities. As a first research step, one year of WRF simulations were run to obtain data that were overlapped with the archived data from the wind farm in order to train the neural network. Simulations were run for the period from June 2019 to June 2020. Following this step, WRF was run in forecasting mode, every six hours, by means of a cron scheduler, and to use its output to generate a power forecast using ML. As Figure 2 shows, all the processes involving WRF: downloading the GFS data, pre and post process and the actual WRF execution take place in the remote cluster, but the commands required to initiate such processes are completely automated and are run from a Python terminal in the local computer.

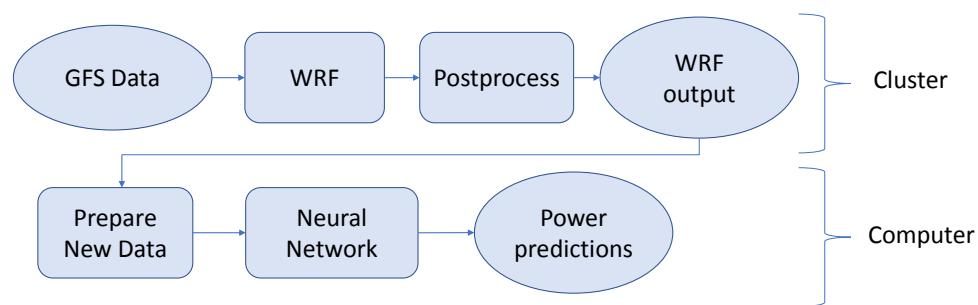


Figure 2. Prediction pipeline.

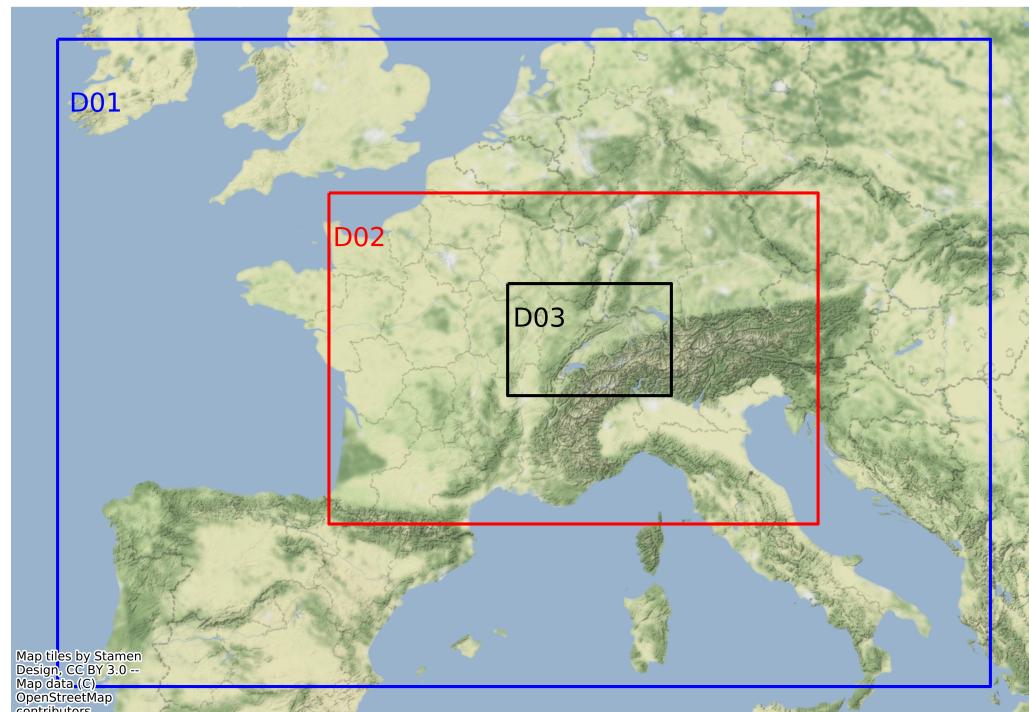


Figure 3. An illustration of the three nested domains (D01,D02,D03) used by WRF to run the meso-scale meteorological model. The Juvent wind farm is located at the center of the third domain on the crest of the Jura mountains in Switzerland.

2.3. Baseline Model

For the evaluation of machine learning models, a baseline model based on WRF predicted wind speeds and a power curve is used. The power curve is fitted to all the wind and power data of each wind turbine, then using the WRF predicted wind speed interpolated at the turbine location the power output of each turbine is predicted. This method needs no training at all and it performs decently as it uses the fitted power curve without outliers.

2.4. Artificial Neural Networks

WRF and ANNs are coupled to directly predict wind power. A feed-forward neural network was developed to take as inputs the temporally lagged meteorological variables extracted from the WRF output, as shown in Figure 4. The idea of using machine learning is to allow the ANN to learn relationships in the data that might be hidden, this will help to tackle power prediction over complex terrain. An ANN is trained for each individual wind turbine with the attempt to infer details from the meso-scale predictions and correct eventual systematic errors from the WRF model. In this study, the focus is on the direct prediction of the generated power and wind speed, but the ANN can be easily trained to predict any other variable present in the Juvent archived data, such as temperature and pressure.

In this research, two models were developed. The first model predicts directly the generated power, while the second model predicts first the wind speed and then uses the same methodology as the baseline model (i.e., use the fitted power curve to finally predict the generated power). So, the final output of both models is power at 10 min interval for the next 12 h.

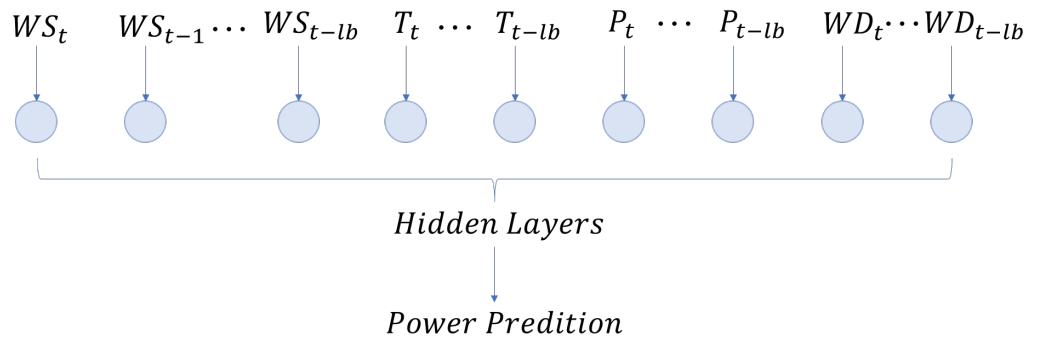


Figure 4. Scheme of the input layer of the ANN. WS stands for wind speed, T for temperature, P for pressure and WD for wind direction. The subscript t is the current time, and each value is lagged until reaching the time corresponding to the subscript $t - lb$; where lb stands for the look-back parameter.

As a general best practice rule when training an ANN, the data should be scaled. The main reasons are that the activation functions of the ANNs do not handle raw data very well, and scaling data leads to faster learning and convergence. The developed model includes the possibility to scale data in three different ways: Robust scaling, MinMax scaling and Standard scaling (for more details, see the Scikit-Learn preprocessing library [19]). Only one type of scaling was used in this study: Standard scaling, the impact of using a different one could be studied. The developed model also provides the ability to automatically re-scale the predictions to return the predicted values in the original units regardless of the scaling choice.

2.5. Data and Training

Data from June 2019 to June 2020 were used to train the different models. The data acquisition from the wind-farm and the WRF simulations are automated processes, the data collection is still ongoing to enlarge the data set for possible future research. WRF version 4.01 was used for the simulations. For each WRF simulation the output file was post-processed

and copied from the remote cluster to the local computer. At the end close to 800 post-processed text files were available for training of the ANN, those files were then merged with the Juvent data (for each wind turbine) by date. Figure 5 presents the normalized power data from all the wind turbines and the simulated wind speed and direction for one sample day to illustrate the data set for the sake of completeness. The final data set consists of sixteen different DataFrames, one for each wind turbine, with columns corresponding to four meteorological variables from WRF simulations and all the observed data from the Juvent wind farm and each row corresponding to a unique time stamp. The WRF wind field in Figure 6 shows the location of the data points. The selected meteorological variables (wind speed and direction, temperature and pressure) are extracted from WRF simulations at each location marked by a black arrow. That information is then interpolated at every wind turbine's location.

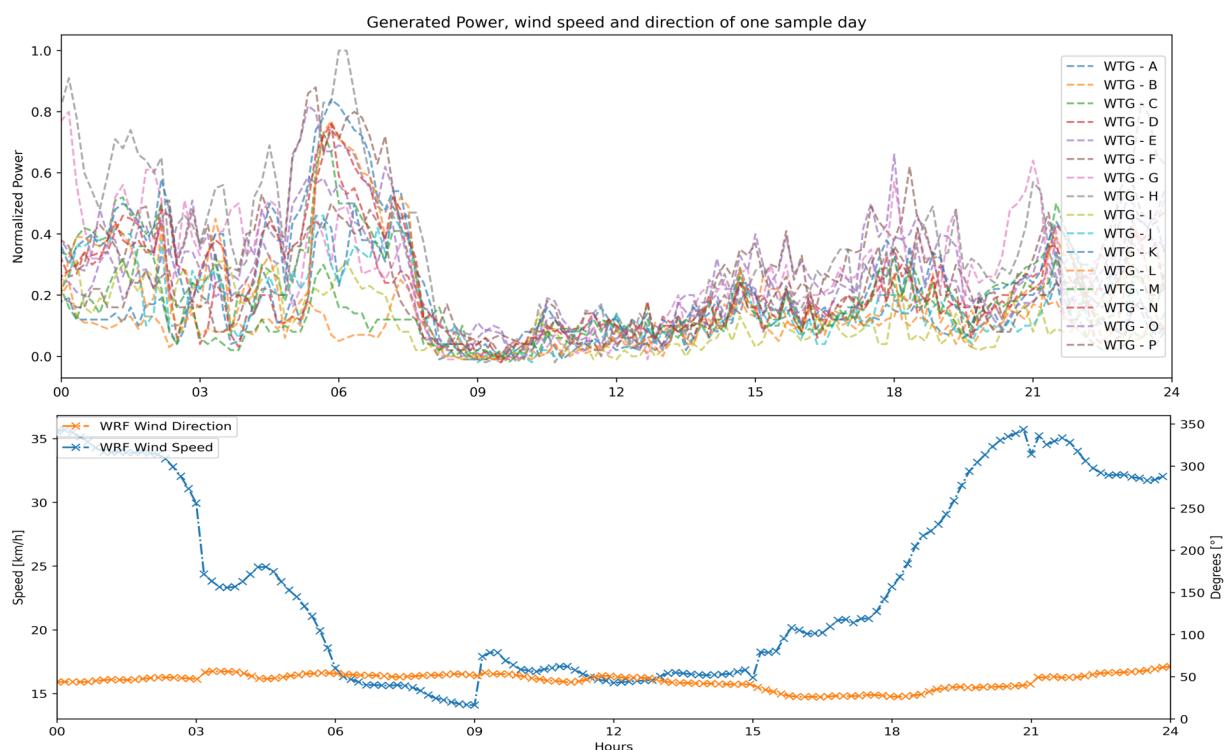


Figure 5. The generated power of each wind turbine normalized by the maximum of measured values is presented in the upper panel for a sample day. Shown in the lower panel are the WRF simulated wind speed and direction for the same day at the wind farm location.

Before training the models, the power curve for each turbine was inspected. The outliers were removed by applying a recursive fitting of a theoretical power curve and eliminating points far away from the fitted curve from the data set. The process first fits the five parameter logistics function [20] to observed data via a least-squares optimization, then the first set of points is removed. Then a generalized additive model is used to fit the remaining points and the elimination process is repeated for points that are far from the fitted curve. This simple approach does not aim to perfectly remove all the disturbances in the data but to remove the big bulk of the outliers, mainly caused by operational stopping of wind turbines. After the outlier removal and as a final pre-processing step the days that were left with very less than 50 measurements were entirely removed from the data set.

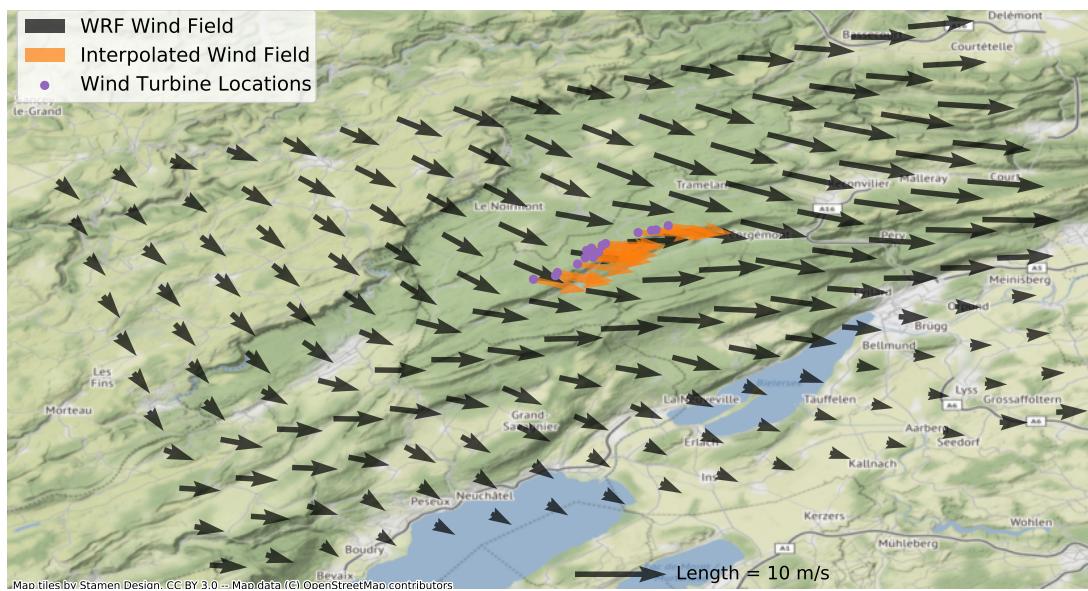


Figure 6. Wind field map. Four meteorological variables (temperature, pressure, wind speed and direction) are retrieved from WRF at each point designated by a black arrow with the scale indicated at the bottom of the figure. Those variables are then interpolated at the wind turbine locations, at four given heights: 50, 95, 150 and 200 m above ground level.

The data was then split into three sets for testing, training and validation. From the entire data set four random days per month were chosen to be part of the validation set. This ensures a validation that is close to the real use of the models: predicting wind power for the next dozen of hours. The remaining data is automatically and randomly split by the ANN training process with 15% of the remaining data assigned to the test set, and the other 85% to the train set.

2.6. Metrics of Model Performance

In order to evaluate the performance of the ANN models for wind power forecasting, four different metrics were used. They are the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), the Median Absolute Error (MedAE) and the Maximum Error (ME). Each of this broadly used metrics provides similar yet unique insight on the performance of the model. The formulas are given in Equations (1)–(4), where y corresponds to the real data and \hat{y} to the predicted data.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (2)$$

$$\text{MedAE} = \text{Median}(|y_1 - \hat{y}_1|, \dots, |y_n - \hat{y}_n|), \quad (3)$$

$$\text{ME} = \text{Max}(|y_1 - \hat{y}_1|, \dots, |y_n - \hat{y}_n|). \quad (4)$$

In this paper, the resulting metric is normalized by the nominal power of the wind turbine in question so that the model performances can be compared fairly. To calculate the metrics, the Scikit-Learn metrics package [19] is adopted.

2.7. Optimization

The ANN has several parameters that need to be adjusted/optimized. They have influences on the way the data is prepared and how the ANN is trained. Those parameters are the temporal lag (*look_back* and *time_jump*), the sample weight (*quantile* and *q_weight*),

and the architecture itself of the ANN. A more detailed explanation of these parameters is provided in the developed *WNN.py* library as well as in the tutorials on how to use this library.

The generated power distribution is skewed and since it is important to predict well the peaks in production, a weighting scheme was implemented during training. Higher weights will be assigned to the power values that are above certain *quantile* of the power distribution. The aforementioned process is called sample weighting (for more details see the *sample_weight* method in the Keras library [21]). It is necessary to adjust the precise value of *quantile* and how important the weight (*q_weight*) will be after that value. In other words the samples are weighted with a step function whose value goes from one to *q_weight* for samples greater than a certain *quantile*. This weighting scheme influences the loss function during training of the ANN, and therefore the back-propagation.

The choice of using a different time step between WRF results is available but for the rest of the paper the time step is 10 min, as is standard in meteorology. The time interval from which the model incorporates past WRF results is determined by the *look_back* parameter. Thus the *look_back* also influences the size of the input layer and therefore the whole architecture (i.e., the number of neurons per layer). For example, assume that *look_back* = 9, the model will look 90 min into the past WRF results.

The ANN prototype has seven layers in total (not including the Dropout): the aforementioned input layer, five hidden layers and the output layer that consists of only one neuron. The architecture in Table 1 details the chosen ANN for the optimization of hyper-parameters. This architecture corresponds to *look_back* = 12 and *time_jump* = 1. It utilizes the wind speed and direction, temperature and pressure from WRF at two different heights (95 and 150 m above ground level) as predictors. Hence the number of neurons in the first layer is equal to the number of features multiplied by the value of *look_back*, namely, four features at two different heights multiplied by twelve.

Table 1. This ANN is used to predict directly the generated power. The chosen architecture with *look_back* = 12, *q_weight* = 1.55 and *quantile* = 0.55. The ANN has a total of 31,336 parameters. It takes into account the wind speed, wind direction, temperature and pressure. It uses a Tanh activation in one layer as an attempt to mimic the nonlinear behavior of the power curve.

Layer Type	Neurons	Parameters
ReLU	104	10,920
ReLU	69	7245
ReLU	69	4830
Dropout	69	0
LeakyReLU	52	3640
Tanh	52	2756
ReLU	36	1908
ReLU	1	37

Since the size of the input layer changes when the *look_back* parameter is modified, the hidden layers should change size as well. With respect to the number of neurons in the input layer, the first layer has two thirds, the second hidden layer has half and the last two hidden layer have one third. Using this varying architecture, one could asses different models with different *look_back* values.

The ANN illustrated in Table 2 does not include the Tanh layer since there is no need to mimic the nonlinear behavior of the power curve because the ANN predicts the wind speed instead of the power. This ANN is the first step in the hybrid predictive model. The wind speed is first predicted using the ANN and then with the fitted power curve, the generated power is predicted. The influence of using a single ANN to directly predict wind power and this hybrid method on the model performance will be discussed in the results section.

Table 2. This ANN is used to predict the wind speed and then a fitted power curve is used to predict power. The chosen architecture corresponds to *look_back* = 9 and no weighting scheme. The ANN has a total of 16,972 parameters. It takes the wind speed, wind direction, temperature and pressure as input variables.

Layer Type	Neurons	Parameters
ReLU	80	6480
ReLU	53	4293
ReLU	53	2862
Dropout	53	0
LeakyReLu	40	2160
ReLU	28	1148
ReLU	1	29

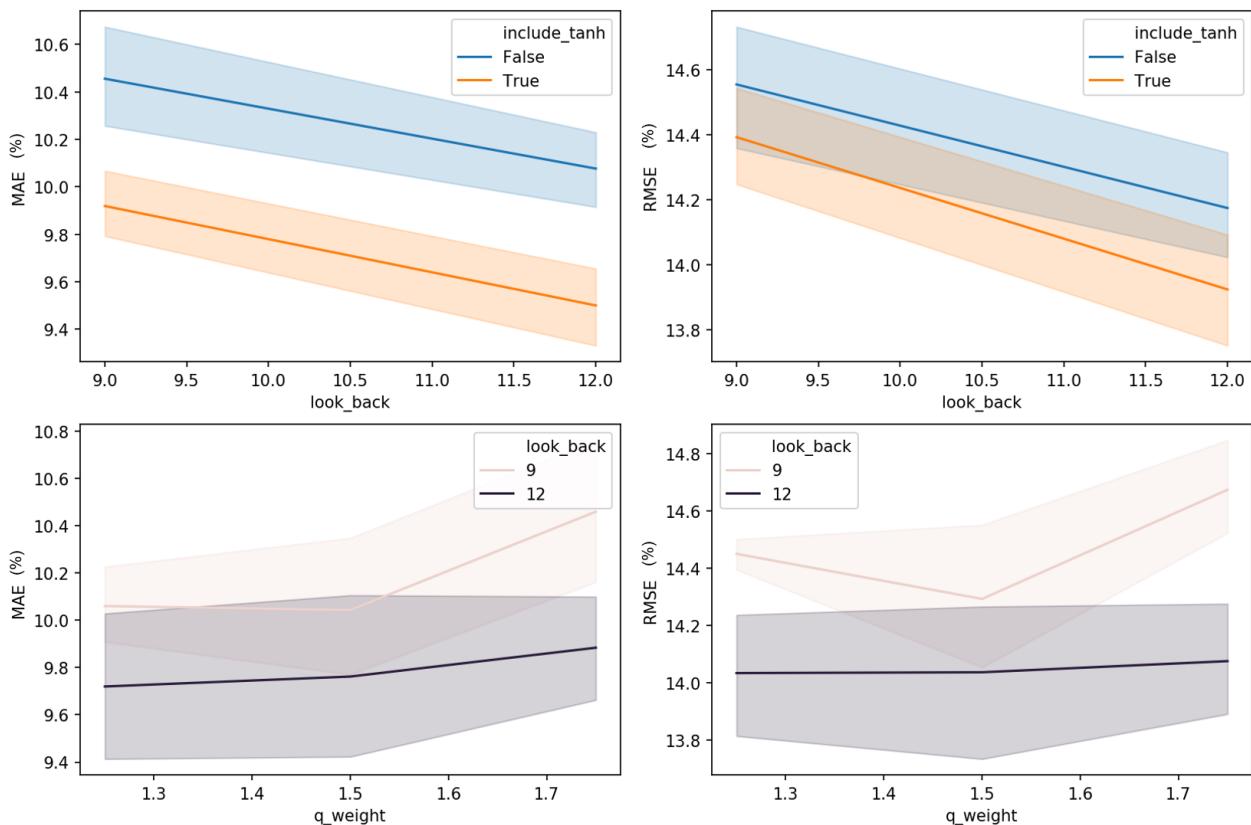


Figure 7. Sensitivity analysis results of the ANN presented in Table 1. This figure shows the optimization process for different parameters such as *look_back*, which means how many past WRF results are used as inputs to the ANN model; a *look_back* of twelve would mean that the WRF data from the past two hours are used to predict the electricity production because the WRF output time step is 10 min. The logical variable *include_tanh* indicates whether or not to include a layer in the ANN with the hyperbolic tangent as activation function.

Both architectures are the results of a sensitivity analysis and multiple tests in order to find the optimal parameters and structure. Some of the performed experiments to optimize the parameters in the direct model are shown in Figure 7. The two subplots at the top clearly show that including a tanh layer to mimic the nonlinear behavior of the power curve yields an improvement, since it reduces both MAE and RMSE consistently. For the two parameters *q_weight* and *quantile* controlling the sample weighting, the optimal values are chosen to be 1.55 and 0.55. This is based on the consideration that the associated error metrics are at the local minimums in most parametric studies (but not always, as shown in the two

subplots at the bottom of Figure 7). The optimal value of *look_back* is found to be around 12, corresponding to a correlation time of 2 h. It is worth mentioning that such a time scale is closely related to the existence of large-scale motions in the atmospheric boundary layer from which turbines extract wind powers. This observation implies that although ANNs are often criticised as black-box models lacking physical grounds, certain optimizations could bring back some physical considerations.

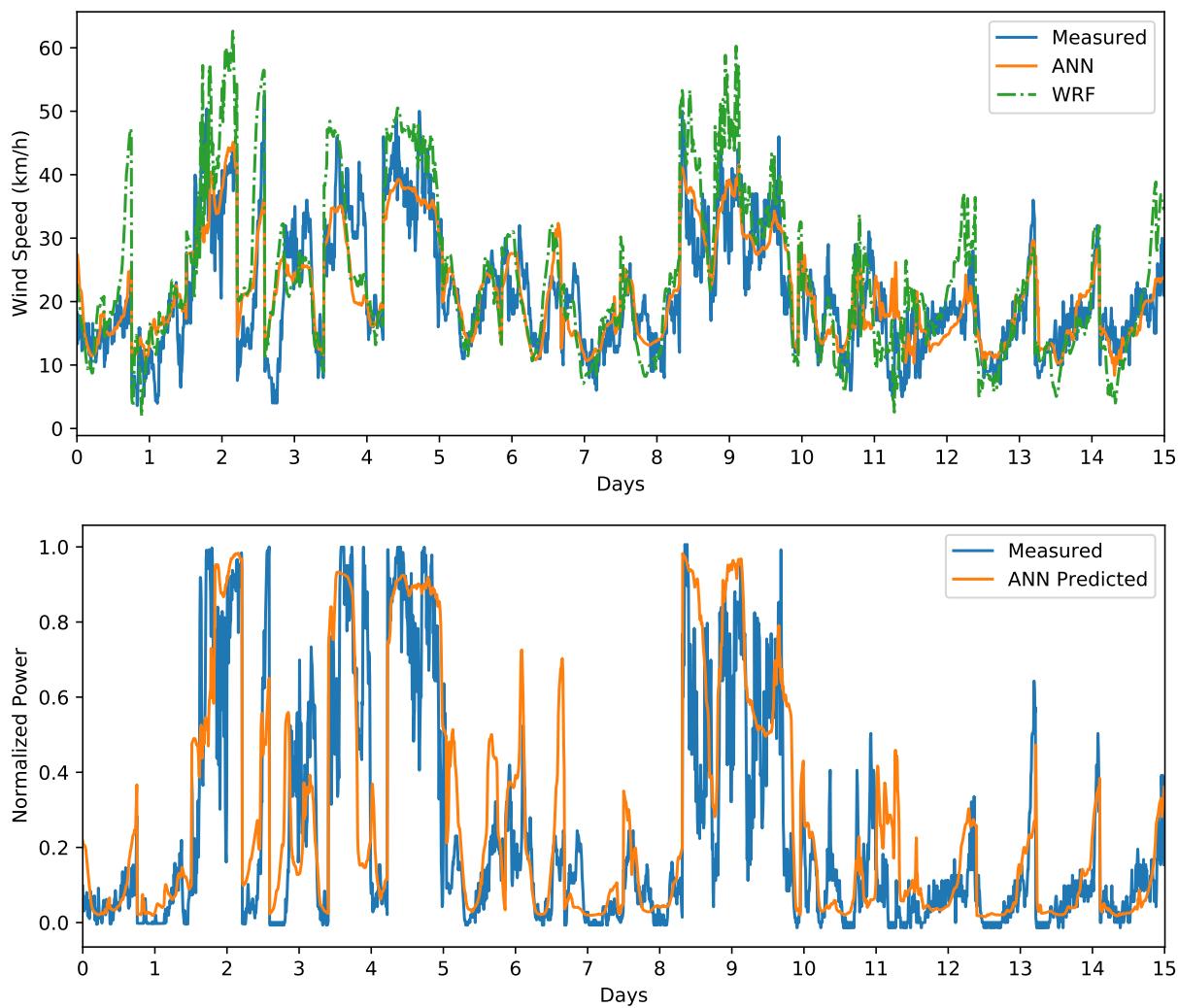


Figure 8. Time series results of measurement and two different ANNs. The upper panel shows the wind speed measured and predicted by WRF and the ANN of the hybrid model, whereas the lower panel shows the power measured and predicted by the direct ANN model. The data shown here corresponds to the validation set of one turbine and the power is normalized by the maximum of measured values. The days illustrated in the figures are not continuous since the validation set is composed of randomly selected days from the original data set.

3. Results and Discussion

This section presents the performance comparison of the baseline, the direct power prediction with an ANN and the hybrid model (ANN plus fitted power curve). The upper panel of Figure 8 shows the wind speed predicted by the ANN in the hybrid model, compared with the WRF results and the measurements. One can clearly see that the ANN improves the WRF prediction on wind speed significantly. The lower panel of Figure 8 shows the direct prediction of generated power by the ANN and the measured power, which allows a qualitative evaluation of the forecasting ability of the model. It can be

seen that the ANN model is able to capture accurately the general behaviour of the power production but quick oscillations in generated power remain hard to predict, for example the patterns in days seven and eleven. It is visible in both panels that the measured data, whether it is wind power or wind speed has high frequency and low amplitude fluctuations that are hardly followed by the predictions. The general trend and behaviour of the variables of interest are nevertheless well captured by the ANN models. The oscillations in the time scale of tenths of minutes can be better predicted using other techniques such as ML models based on time-series data. Since the proposed models aim for forecast horizons beyond 6 h, the behavior of smoothing out such oscillations does not pose a problem in their practical applications.

For a more quantitative analysis, Figure 9 shows the normalized MAE and RMSE of the three models for each wind turbine. For the other two metrics (MedAE and ME), the results are presented in Figure 10. For the baseline model, its MAE ranges from 11.7% to 18.9% and its RMSE ranges from 18.3% to 28.3%. For both metrics, the maximum occurs at turbine 2 and the minimum occurs at turbine 6. So, the baseline model for turbine 2 performs the worst. The direct ANN model improves the baseline model almost at every wind turbine, and the hybrid model clearly outperforms the other two models (Figures 9 and 10). Compared to the baseline model, the hybrid model improves consistently the predictions for all the wind turbines. Especially, for turbine 2, the RMSE is reduced by 13.3% and the MAE is reduced by 9.2%, hence the relative improvement of forecast accuracy is nearly 50% (in another word, the accuracy is doubled in the hybrid model).

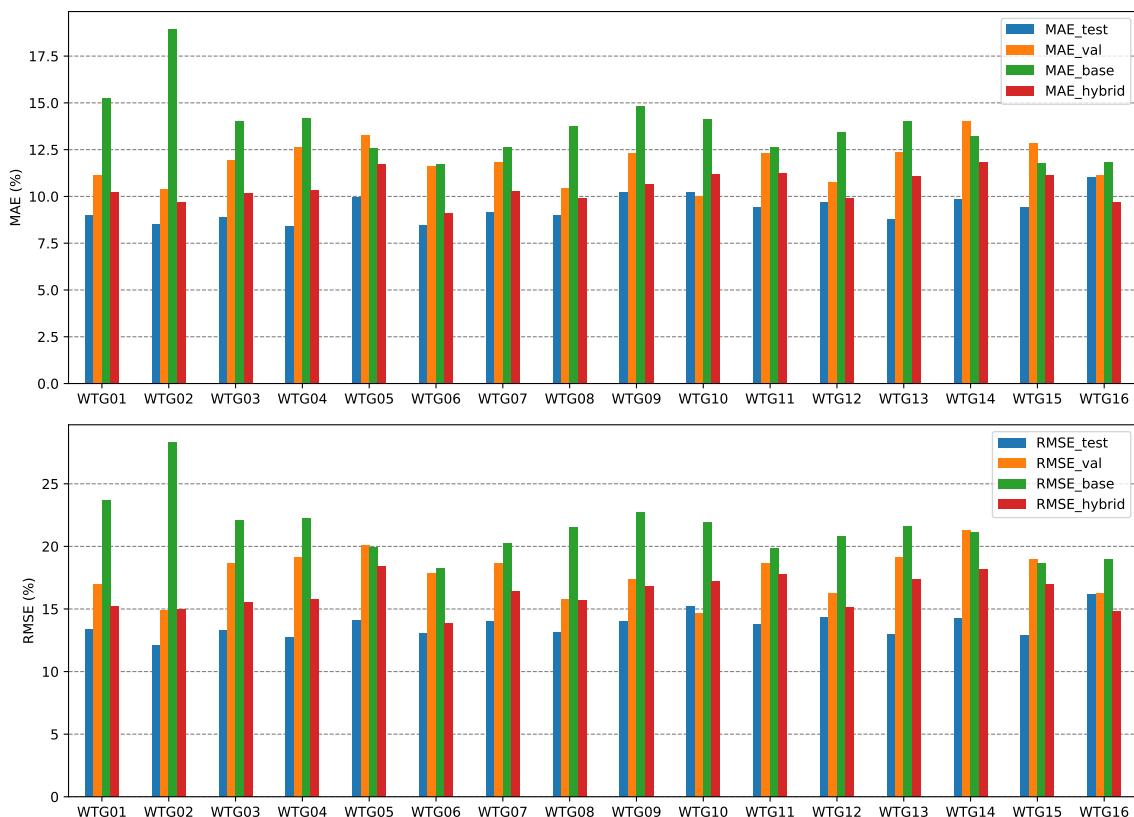


Figure 9. Normalized Mean Average Error and Root Mean Squared Error of the generated power predictions for each wind turbine. The predictions are obtained by the three models considered. The first is the direct power prediction using an ANN (including both test and validation results, in blue and orange respectively). The second is the validation set of the baseline model, in green. The third is the validation set of the hybrid model, depicted in red. It is worth mentioning that for the two last models it is not relevant to show the test set, because the baseline has no training and the test of the hybrid model has different units, since the ANN is trained to predict wind speed.

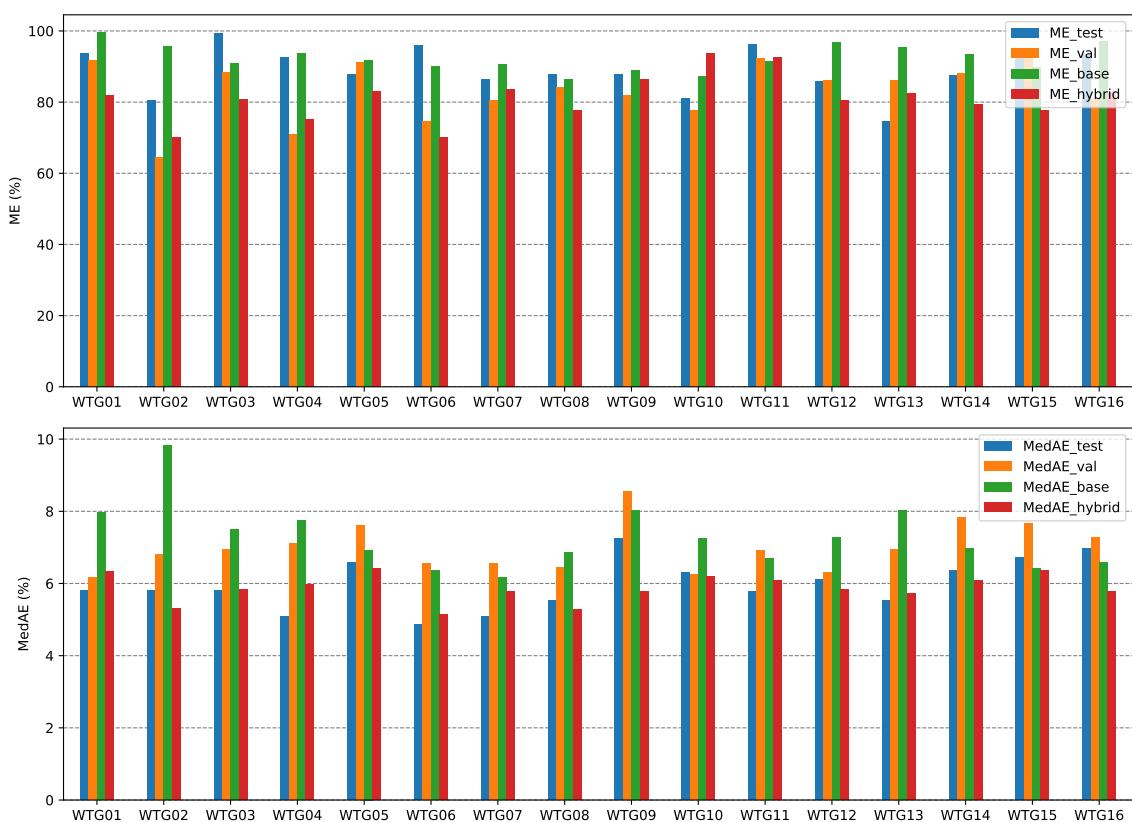


Figure 10. Normalized Median Absolute Error and Maximum Error for model predictions at all wind turbines. It is again clear that the hybrid model outperforms the rest.

The effect of complex terrain on the WRF simulations can be observed in the high standard deviation of the baseline predictions. For some of the turbines such as 5, 6, 7, 11, 15, and 16, power outputs are predicted with acceptable accuracy (namely, MAE lower than 12%, comparable to the ANN-based models), whereas for the others, they are not. This can be explained by the fact that WRF is a mesoscale model in which local terrain features such as topography and vegetation and micro-scale flow complexities such as wind turbine wakes are not or poorly represented. For example, turbine 2 is located at the edge of the forest and on the leeward side of the hill Mont-Soleil when wind blows from the predominant sector (south-west). It was shown in [18] that high-fidelity modeling of the recirculating zone and the forest is necessary for accurately predicting the wind field around turbine 2. That is why the baseline model performs the worst at this location. Indeed, this is one aspect that the use of machine learning aims to improve, by learning the systematic errors or biases in the physics-based predictions regarding each individual wind turbine and correcting them in the prediction phase. Significant improvements brought by the proposed hybrid model confirm the expectation. Nevertheless, the performance of the hybrid model still has considerable variations at different locations. A recent study [15] showed that an optimal selection of the NWP nodes (based on the wind farm layout, the wind rose and the terrain) improves the forecast performance of a physical downscaling method. Similarly, the ANN models developed here can be further improved by adding the selection of the NWP nodes into the optimization process.

Table 3 summarizes the turbine-averaged metrics for the evaluation of the overall performance of the three models considered in this research. The hybrid model stands out with consistent improvements over the other two models. For example, taking the baseline model as a reference, it reduces almost half the standard deviation for both RMSE and MAE. As revealed by the short review of relevant references in the introduction, the reported values for the MAE range from 10% to 13% and those for the RMSE range from 14% to 26%. Most of them are for the total wind farm production forecast (i.e., the sum of the production

forecast at each turbine). To have a fair comparison of our work with the others, the total production forecast is obtained by summing the forecast given by the hybrid model for each turbine. After doing so, the normalized MAE and RMSE are found to be 8.76% and 13.03%, respectively. It is important to note that a lower value of the normalized MAE for the total production forecast is expected because instantaneous errors of individual forecasts at different turbines may cancel out from time to time. Overall, the high accuracy attained by the hybrid model makes it amongst the state-of-the-art NWP+ML models developed so far. In the context of complex terrain, the hybrid model considerably outperforms the other ANN models developed for sites with similar level of complexity. For examples, the NWP+ML model developed for the wind power prediction of a wind farm located at the eastern coastal area of Nantong city in Jiangsu province, China [11] reported a value of 16.47% for the normalized RMSE; the ANN model developed for a wind farm at an Italian mountainous site [14] was shown to give the normalized RMSE around 25% for one cluster of turbines and 21% for another cluster of turbines; the ANN-based power forecast performed at a site on the west coast of Norway with moderate terrain complexity [15] yields the normalized MAE around 10% and the normalized RMSE around 16%. Considering the fact that the site in this study possesses an additional complexity caused by the heterogeneous land cover (varying from grassy to forested), the effectiveness of performed optimizations and the superior forecasting capability of the hybrid model are validated.

Table 3. Summary of the metrics averaged over the 16 turbines for the three considered models. For the direct power prediction with an ANN, results from both test set and validation set are included, whereas for the other two only the validation set is relevant.

	RMSE	MAE	MedAE	ME
Averaged Error (% kW/kW nominal)				
Test	13.7	9.4	6.0	89.1
Validation	17.8	11.8	7.0	83.3
Baseline	21.4	13.7	7.3	92.4
Hybrid	16.3	10.5	5.9	81.2
Standard Deviation (% kW/kW nominal)				
Test	1.0	0.7	0.7	6.7
Validation	1.9	1.1	0.7	8.2
Baseline	2.4	1.8	0.9	3.7
Hybrid	1.3	0.8	0.4	6.5

4. Conclusions

This research work has successfully performed a proof of concept on the WRF and ANN coupling to directly predict wind power production over a highly complex terrain. Thanks to intensive parameter/architecture optimizations, the resulting prediction pipeline is able to produce real time forecasts with a higher accuracy (MAE being 8.76% and RMSE being 13.03%) compared to the existing forecast models in the same category. The developed library provides the user with flexibility in the choice of parameters, features, target variables and machine learning architectures. This data-driven approach into wind power prediction proves useful since it can be easily and constantly retrained and upgraded. The use of ANN was proven a valuable addition to the existing forecasting systems, hopefully providing unique insights on another application of machine learning and solving some of the challenges that complex terrain poses when predicting wind fields and power outputs, without relying on computationally heavy physics-driven approaches.

The main findings of this research work can be summarized as follows:

1. The hybrid model combining an ANN for wind speed forecast with an on-site calibrated power curve performs better than the direct model using an ANN with power as its output;
2. For the direct model, including a tanh layer to mimic the nonlinear behavior of the power curve improves the forecast accuracy;
3. In the ANN inputs, it is preferable to include the NWP results in the past (up to 2 h for a good balance between performance gain and computation cost);
4. A proper sample weighting scheme taking into account the skewed power distribution can improve the model performance;
5. For a site with complex terrain conditions, the model performance is also heterogeneous. The effectiveness of ANN models in correcting systematic errors of physics-based models is proven.

Despite the model optimization and validation studies are performed at a specific site in the Swiss Jura mountains, some of the findings above may be generalized at least qualitatively.

Author Contributions: Conceptualization, L.D., J.F. and F.P.-A.; Data curation, L.D. and J.F.; Formal analysis, L.D.; Funding acquisition, F.P.-A.; Investigation, L.D.; Methodology, L.D. and J.F.; Project administration, J.F. and F.P.-A.; Resources, F.P.-A.; Software, L.D.; Supervision, J.F. and F.P.-A.; Validation, L.D.; Visualization, L.D.; Writing—original draft, L.D. and J.F.; Writing—review and editing, J.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swiss Federal Office of Energy (Grant SI/502135-01) and the Swiss Innovation Agency (grant number 1155002544) within the frame of the Swiss Centre for Competence in Energy Research on the Future Swiss Electrical Infrastructure (SCCER-FURIES).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is not made available publicly.

Acknowledgments: The authors would like to thank Sophie Bosse and BKW Energie SA for providing the Juvent wind farm data.

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

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
D01	Domain 1 of Weather Model
D02	Domain 2 of Weather Model
D03	Domain 3 of Weather Model
GFS	Global Forecast System
MAE	Mean Absolute Error
ME	Maximum Error
MedAE	Absolute Error
ML	Machine Learning
NWP	Numerical Weather Prediction
P	pressure
RMSE	Root Mean Squared Error
SCADA	Supervisory Control And Data Acquisition
T	temperature
WD	wind direction
WRF	Weather Research and Forecasting Model
WS	wind speed
WTG	Wind Turbine Generator

References

1. Notton, G.; Nivet, M.L.; Voyant, C.; Paoli, C.; Darras, C.; Motte, F.; Fouilloy, A. Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting. *Renew. Sustain. Energy Rev.* **2018**, *87*, 96–105. [[CrossRef](#)]
2. Yousuf, M.; Al-Bahadly, I.; Avci, E. Current Perspective on the Accuracy of Deterministic Wind Speed and Power Forecasting. *IEEE Access* **2019**, *7*, 159547–159564. [[CrossRef](#)]
3. Jung, J.; Broadwater, R. Current status and future advances for wind speed and power forecasting. *Renew. Sustain. Energy Rev.* **2014**, *31*, 762–777. [[CrossRef](#)]
4. Foley, A.; Leahy, P.; Marvuglia, A.; McKeogh, E. Current methods and advances in forecasting of wind power generation. *Renew. Energy* **2012**, *37*, 1–8. [[CrossRef](#)]
5. Colak, I.; Sagiroglu, S.; Yesilbudak, M. Data mining and wind power prediction: A literature review. *Renew. Energy* **2012**, *46*, 241–247. [[CrossRef](#)]
6. Manero, J.; Béjar, J.; Cortés, U. Wind energy forecasting with neural networks: A literature review. *Comput. Sist.* **2018**, *22*, 1085–1098. [[CrossRef](#)]
7. Marugán, A.; Márquez, F.; Pérez, J.; Ruiz-Hernández, D. A survey of artificial neural network in wind energy systems. *Appl. Energy* **2018**, *228*, 1822–1836. [[CrossRef](#)]
8. Liu, H.; Chen, C.; Lv, X.; Wu, X.; Liu, M. Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods. *Energy Convers. Manag.* **2019**, *195*, 328–345. [[CrossRef](#)]
9. Ramirez-Rosado, I.; Fernandez-Jimenez, L.; Monteiro, C.; Sousa, J.; Bessa, R. Comparison of two new short-term wind-power forecasting systems. *Renew. Energy* **2009**, *34*, 1848–1854. [[CrossRef](#)]
10. Vaccaro, A.; Mercogliano, P.; Schiano, P.; Villacci, D. An adaptive framework based on multi-model data fusion for one-day-ahead wind power forecasting. *Electr. Power Syst. Res.* **2011**, *81*, 775–782. [[CrossRef](#)]
11. Zhao, P.; Wang, J.; Xia, J.; Dai, Y.; Sheng, Y.; Yue, J. Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China. *Renew. Energy* **2012**, *43*, 234–241. [[CrossRef](#)]
12. Xu, Q.; He, D.; Zhang, N.; Kang, C.; Xia, Q.; Bai, J.; Huang, J. A short-term wind power forecasting approach with adjustment of numerical weather prediction input by data mining. *IEEE Trans. Sustain. Energy* **2015**, *6*, 1283–1291. [[CrossRef](#)]
13. Men, Z.; Yee, E.; Lien, F.S.; Wen, D.; Chen, Y. Short-term wind speed and power forecasting using an ensemble of mixture density neural networks. *Renew. Energy* **2016**, *87*, 203–211. [[CrossRef](#)]
14. Mana, M.; Burlando, M.; Meißner, C. Evaluation of two ANN approaches for the wind power forecast in a mountainous site. *Int. J. Renew. Energy Res.* **2017**, *7*, 1629–1638.
15. Mana, M.; Astolfi, D.; Castellani, F.; Meißner, C. Day-ahead wind power forecast through high-resolution mesoscale model: Local computational fluid dynamics versus artificial neural network downscaling. *J. Sol. Energy Eng.* **2020**, *142*, 034502. [[CrossRef](#)]
16. Haupt, S.; McCandless, T.; Dettling, S.; Alessandrini, S.; Lee, J.; Linden, S.; Petzke, W.; Brummet, T.; Nguyen, N.; Kosović, B.; et al. Combining artificial intelligence with physics-based methods for probabilistic renewable energy forecasting. *Energies* **2020**, *13*, 1979. [[CrossRef](#)]
17. Hong, T.; Pinson, P.; Fan, S. Global energy forecasting competition 2012. *Int. J. Forecast.* **2014**, *30*, 357–363. [[CrossRef](#)]
18. Tabas, D.; Fang, J.; Porté-Agel, F. Wind energy prediction in highly complex terrain by computational fluid dynamics. *Energies* **2019**, *12*, 1311. [[CrossRef](#)]
19. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
20. Yesilbudak, M. Partitional clustering-based outlier detection for power curve optimization of wind turbines. In Proceedings of the 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, UK, 20–23 November 2016; pp. 1080–1084. [[CrossRef](#)]
21. Chollet, F. Keras. 2015. Available online: <https://keras.io> (accessed on 8 January 2021).