Evaluation of Wind Turbine Power Outputs with and without Uncertainties in Input Wind Speed and Wind Direction Data

M. Zou*, D. Fang*, S. Z. Djokic*, V. Di Giorgio**, R. Langella** and A. Testa**
*School of Engineering, The University of Edinburgh, (*UK*)
**University of Campania "Luigi Vanvitelli", Aversa (CE), (*Italy*)

***Abstract*—This paper analyses importance of including wind direction (WD) as an additional explanatory variable to the wind speed (WS) for evaluating uncertainty in wind turbine (WT) power output (P_{out}). Using available measurements of an actual WT, the paper compares a "two-dimensional" (2D) P_{out}-WS model with a "three-dimensional" (3D) P_{out}-WS-WD model for two general cases: a) for the specific input WS and WD values (i.e. WS and WD without uncertainties), and b) for the forecasted input WS and WD values (i.e. WS and WD with uncertainties). In paper, 2D and 3D Gaussian mixture Copula model and vine Copula framework are combined with 2D and 3D Markov chain models, which are used to forecast input WS and WD data with uncertainties. The obtained results show that inclusion of WD will provide noticeable improvement for models with no uncertainties in input WS and WD data, while in the case of forecasted WS and WD data with uncertainties, WS is a much stronger contributor to the total WT P_{out} uncertainty than WD.**

***Index Terms*—Copula model, Markov chain model, uncertainty, wind direction, wind speed, wind turbine.**

I. Introduction

Over the recent years, there was a significant increase in numbers and sizes of wind-based electricity generation systems (WEGS), which are expected to grow even further in the future. Wind energy is a "clean" and renewable energy resource, which however exhibits strong and stochastic spatio-temporal variations, therefore resulting in uncertain and difficult to predict outputs of WEGS. The assessment and quantification of uncertainties in power outputs of WEGS's are important not only for wind farm (WF) owners, but also for system operators, due to their impact on the variations of network power flows and related uncertainties in network operating conditions [1-3].

This paper extends the initial work by the authors in [4], in which the conversion of input wind energy into the output electrical power by a wind turbine (WT) is approached as an inherently stochastic process, whose statistical properties can be described and modelled using a suitable set of probabilistic input and output variables. In [4], the authors investigated importance of including wind direction (WD) as an additional input explanatory variable to the wind speed (WS) for assessing uncertainty in the WT power output (P_{out}). This is done by comparing the results of a "two-dimensional" (2D) P_{out}-WS model with a "three-dimensional" (3D) P_{out}-WS-WD model for the *specific* (e.g. measured, or known) input WS and WD values, demonstrating that the presented 3D model allows for a more confident evaluation of WT's P_{out} uncertainty.

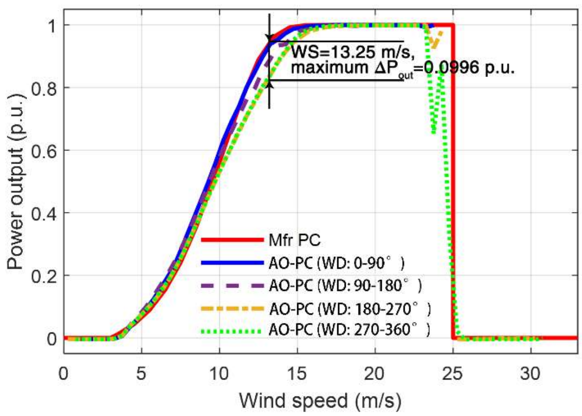
The analysis in this follow-up paper provides a more comprehensive evaluation of uncertainties in WT power outputs, as it takes into account uncertainties in the input wind energy, i.e. uncertainties in the input WS and WD values. For that purpose, 2D and 3D Copula based models from [4] are combined with 2D and 3D Markov chain (MC) models [5] , which are in this paper developed using two years of available WS and WD measurements and then used to forecast the WS and WD values, together with their joint and marginal probabilities, in the third year. For a given averaging window (e.g. 10 minutes), the uncertainty in WT's power output is defined in this paper as the 5^{th}-95^{th} percentile range of variations of P_{out} values for two general cases: a) for specific or known input WS and WD values (i.e. WS and WD are without uncertainties, as in [4]), or b) for estimated or forecasted input WS and WD values (i.e. WS and WD are both with associated uncertainties). This definition allows for a systematic analysis and direct comparison of different causes of uncertainties in WT power output in Case a) and Case b), as well as for a consistent evaluation of uncertainties in both WEGS forecasting and WEGS hindcasting studies, which basically differ only in the way how the input WS and WD data are obtained and used in the WT model.

II. WT Power Curve Models, Importance of Including WD and Empirical P_{OUT} Uncertainty

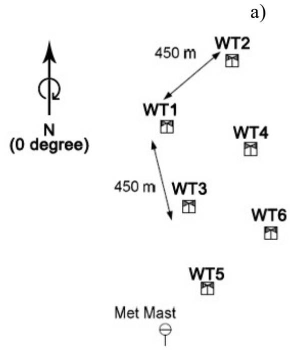
The most common approach in assessing outputs of WEGS's is to correlate WT power output with only one parameter: input WS. A widely-used examples are power curves provided by WT manufacturers, which are obtained in air-tunnels and therefore cannot correctly represent many site-specific and application-specific factors, such as the wake or terrain effects, or WT dynamics due to WS and/or WD variations [6]-[7]. The impact of WD is often neglected in both deterministic and probabilistic studies, [3], [8-11], assuming that, for the WS between the cut-in and cut-out speed, yawing control will orient WT so that the WD ("attacking angle") is perpendicular to the rotor blades plane, [12], [13]. However, yawing mechanisms are relatively slow (e.g. 0.5 °/s in [11]), so rapid WD changes typically result in a reduced P_{out} of a WT. Furthermore, the conventional WT models do not provide information on the impact of WD and although [14]-[15] demonstrated that WT model can be improved by considering both WS and WD, this work mostly concentrated on assessing impact of stochastic WS and WD variations, not on correlating P_{out} uncertainties for specific WS and WD data.

A. An Example of WD Impact

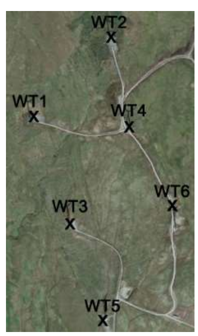
Fig. 1a compares manufacturer power curve (Mfr-PC) model and average operational power curve (AO-PC) models for one WT (WT3) in a WF shown in Figs 1b and 1c. The AO-PC plots are obtained using synchronous WS, WD and P_{out} measurements for winter seasons over the period of three years. As the WF layout and AO-PC results in Fig. 1a indicate, WT3 has the highest average P_{out} for WDs between 0-90 ^{o}, corresponding to the minimum turbulence and wake effect conditions. It is also clear that Mfr-PC does not correctly represent the actual operating performance of WT3: the maximum difference from Mfr PC in the WS region between cut-in and rated P_{out} speed is around 10% for WS=13.25 m/s. (Note: WD is measured as a clockwise relative angle between the zero degree North direction, as indicated in Fig. 1b.)



a)



b)



c)

Fig. 1. An illustration of the impact of WD on WT power outputs: a) comparison of Mfr-PC and AO-PC models for the four main WD quadrants: b) and c) WF layout, with positions of WTs within the WF.

B. Manufacturer and Operational WT Power Curves

A typical Mfr-PC WT model specifies the power output generated by the WT for the corresponding input WS using a deterministic segmented function:

$$P = \begin{cases} 0 & v < v_i, v \geq v_o \\ f(v) & v_i \leq v < v_r \\ P_{\text{rated}} & v_r \leq v < v_o \end{cases} \tag{1}$$

where: v is the wind speed; $P=P_{out}$ is the corresponding (synchronous) WT power output; P_{rated} is the rated WT power output; v_i, v_r, and v_o are cut-in, rated and cut-out WS values, respectively. The nonlinear function *f(v)*, which represents the WS-P_{out} relationship, is in the most of Mfr PC models limited to a region between v_i and v_r speeds.

However, in actual operating conditions, a range of P_{out} values is measured for a range of specific input WS values, as indicated in Fig. 2, showing actually measured 10-min P_{out} values, the Mfr PC, the corresponding 2D AO-PC and 5th-95th percentile range of variations around it.

According to Fig. 2 and contrary to (1), P_{out}-WS relationship can be divided into five indicated WS regions. No uncertainty model is required in Regions I, III and V. In Regions II and IV, for the same input WS P_{out} can vary significantly around the mean value and its uncertainty is indicated with a 5th-95th percentile range around the mean. Although two uncertainty models are required in Regions II and IV, less than 0.5 % of total measurements is found in Region IV, so only uncertainties in P_{out} values in Region II are included in the analysis presented in this paper (in Regions I, III V P_{out} is assumed to follow Mfr-PC).

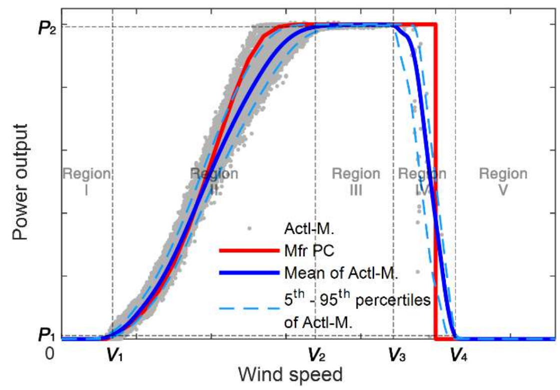


Fig. 2. Comparison of Mfr PC and actually measured values (Actl-M), with corresponding AO-PC and 5th-95th percentile range of variations (uncertainty in P_{out}) within the five indicated WS regions.

C. Empirical (Actual) WT P_{out} Uncertainty

For the actual synchronous measurements of WS, WD and P_{out}, "empirical WT power output uncertainty" is defined as the 5th and 95th percentile range of variations around the measured mean P_{out} values for the specific WS value (e.g. for WS in bins of 0.5 m/s) and for the specific WD value (e.g. for WD in bins of 30 ^{o}), Fig. 3. If there are not enough P_{out} data for particular WS and WD bins, the existing bins are extended over the adjacent bins, so statistically significant number of data is obtained. This "empirical" WT P_{out} uncertainty is used as a reference for evaluating "modelled" WT P_{out} uncertainty for both specific and forecasted WS and WD values (Section V). Fig. 3 illustrates empirical P_{out} uncertainty and possible significant errors if Mfr-PC model is used.

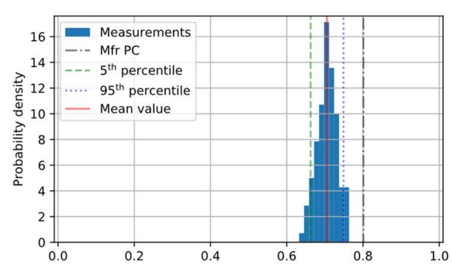


Fig. 3. An example of empirical/actual WT P_{out} uncertainty

III. COPULA-BASED MODEL

The models presented in this section are based on Copula theory [16] and conditional 2D P_{out}-WS and 3D P_{out}-WS-WD probability models. In Copula models with deterministic (i.e. known or measured) input values, the actual measured WS and WD data (without uncertainties) are used as inputs to determine mean P_{out} values and corresponding 5^{th}-95^{th} percentile range (i.e. uncertainty in P_{out}). In Copula models with probabilistic input values, the forecasted WS and WD values with their uncertainty ranges obtained by the MC model (see next section) are used to calculate conditional probability density of P_{out}, its mean value and 5^{th}-95^{th} percentile uncertainty ranges.

Widely used Copula families in 2-D modelling are elliptical Copulas (e.g., Gaussian Copula and *t*-Copula) and Archimedean Copulas (e.g., Clayton Copula, Frank Copula and Gumbel Copula), which, however, cannot represent multimodal correlation well. Ref. [14] introduces a Gaussian mixture Copula model (GMCM), which provides improved performance and flexibility, compared to a more widely-used, but simpler, elliptical Copulas and Archimedean Copulas, especially for describing highly nonlinear dependencies [15]. Ref. [16] proposes a vine Copula model to model high-dimensional dependency by using several Copulas in pair correlations to analyse the whole correlation structure. Applications of Copula are attracting increasing attention in the areas of power system analysis and renewable energy research, such as load profile clustering [17], forecasting of WEGS's [18-20] and related aspects of uncertainty and correlation [2], [21-23].

According to [24], the joint distribution of multiple variables can be expressed by marginal distributions and a Copula function that describes the dependency structure among the variables. For known marginal distributions F_1, $F_2,\ldots, F_d$, their joint distribution $F_{1,2,\ldots,d}$ can be written as $F_{1,2,\ldots,d} = C(F_1, F_2,\ldots, F_d)$, where $C(\cdot)$ is the Copula distribution function. If $F_1, F_2,\ldots, F_d$ are continuous, then $C(\cdot)$ is unique and for a d-dimensional variable structure:

$$C\left(F_1,F_2,\ldots,F_d\right)=F_{1,2,\ldots,d}\left(x_1,x_2,\ldots,x_d\right) \tag{2}$$

$$c(F_1,F_2,\ldots,F_d)=\frac{\partial^d C(F_1,F_2,\ldots,F_d)}{\partial F_1\partial F_2\ldots\partial F_d}= \\ =\frac{\partial^d F_{1,2,\ldots,d}(x_1,x_2,\ldots,x_d)}{\partial x_1\partial x_2\ldots\partial x_d}\times\prod_{i=1}^{d}\frac{\partial x_i}{\partial F_i}=\frac{f_{1,2,\ldots,d}(x_1,x_2,\ldots,x_d)}{\prod_{i=1}^{d}f_i(x_i)} \tag{3}$$

where: $C(\cdot)$ and $c(\cdot)$ are Copula distribution and density functions; $F_i(\cdot)$ and $f_i(\cdot)$ are marginal distribution and density functions of x_i and $F_i(\cdot)\sim U(0,1)$; $F_{1,2,\ldots,d}(\cdot)$ and $f_{1,2,\ldots,d}(\cdot)$ are joint distribution and density functions.

Specifically, GMCM can be written as:

$$\begin{cases} C_{\text{GMCM}}\left(F_1,F_2,\ldots,F_d\right)=\Psi\left(\Psi_1^{-1}\left(F_1\right),\Psi_2^{-1}\left(F_2\right),\ldots,\Psi_d^{-1}\left(F_d\right)\right) \\ c_{\text{GMCM}}\left(F_1,F_2,\ldots,F_d\right)=\dfrac{\psi\left(\Psi_1^{-1}\left(F_1\right),\Psi_2^{-1}\left(F_2\right),\ldots,\Psi_d^{-1}\left(F_d\right)\right)}{\prod_{j=1}^{d}\psi_j\left(\Psi_j^{-1}\left(F_j\right)\right)} \end{cases} \tag{4}$$

where $\psi(\cdot)$ and $\Psi(\cdot)$ are the probability density function and probability distribution function of Gaussian mixture model, respectively, and $\psi_j(\cdot)$ and $\Psi_j^{-1}(\cdot)$ are the corresponding marginal density function and inverse distribution function at j-th dimension.

A. 2-D P_{out}-WS Model: GMCM Framework

Let x_1 represents P_{out} and x_2 represents WS. As P_{out} is a truncated variable, with minimum equal to 0 and maximum equal to rated power output, it is transformed into unconstrained linear space as:

$$x_{1;\text{new}}=\log\left(\frac{x_{1;\text{old}}-0}{\text{rated power output}-x_{1;\text{old}}}\right) \tag{5}$$

WS is also a truncated variable (it cannot be less than 0) and similar transformation is applied: to build a P_{out}-WS model with GMCM, first the GMMs are used as marginal distribution functions, $F_1(\cdot)$ and $F_2(\cdot)$, to transfer x_1 and x_2 into cumulative distribution function domain, i.e., U (0,1). The joint probability density of x_1 and x_2 is then:

$$f_{12}\left(x_1,x_2\right)=c_{\text{GMCM}}\left(F_1,F_2\right)f_1\left(x_1\right)f_2\left(x_2\right) \tag{6}$$

For deterministic WS input, the conditional probability density of P_{out} is obtained as:

$$f_{1|2}\left(x_1\mid x_2\right)=\frac{f_{12}\left(x_1,x_2\right)}{f_2\left(x_2\right)}=c_{\text{GMCM}}\left(F_1,F_2\right)f_1\left(x_1\right) \tag{7}$$

If the given WS is not deterministic, but features uncertainties, e.g. in a range from $x_{2;\text{lower}} \le x_2 < x_{2;\text{upper}}$, then the theoretical conditional probability density of P_{out} can be derived as in (8).

This paper, however, proposes another numerical method to calculate $f_{1|2}(x_1|x_{2;\text{lower}} \le x_2 < x_{2;\text{upper}})$, as in the processed data $x_{2;\text{upper}}$ and $x_{2;\text{lower}}$ may be lower and outside the boundary of Region II, i.e. it may belong to adjacent deterministic regions (without uncertainty), which are not included in the GMCM (see Fig. 2). The proposed method is "grid search method" which is based on a mixture distribution combining both deterministic and probabilistic model, which can be formulated as in (9).

$$\begin{aligned} f_{1|2}\left(x_1\mid x_{2;\text{lower}}\le x_2<x_{2;\text{upper}}\right)&=\frac{f_{12}\left(x_1,x_{2;\text{lower}}\le x_2<x_{2;\text{upper}}\right)}{f_2\left(x_{2;\text{lower}}\le x_2<x_{2;\text{upper}}\right)}=\frac{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_{12}\left(x_1,x\right)dx}{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_2\left(x\right)dx} \\ &=\frac{1}{F_{2;\text{upper}}-F_{2;\text{lower}}}\frac{F_{12}\left(x_1,x_{2;\text{lower}}\le x_2<x_{2;\text{upper}}\right)}{\partial x_1}=\frac{1}{F_{2;\text{upper}}-F_{2;\text{lower}}}\left(\frac{\partial C_{\text{GMCM}}\left(F_1,F_{2;\text{upper}}\right)}{\partial x_1}-\frac{\partial C_{\text{GMCM}}\left(F_1,F_{2;\text{lower}}\right)}{\partial x_1}\right) \end{aligned} \tag{8}$$

$$f_{1|2}\left(x_1\mid x_{2;\text{lower}}\le x_2<x_{2;\text{upper}}\right)=\frac{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_{1,2}\left(x_1,x\right)dx}{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_2\left(x\right)dx}=\frac{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_2\left(x\right)f_{1|2}\left(x_1\mid x\right)dx}{\int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}}f_2\left(x\right)dx}=\lim_{n\to\infty}\frac{\sum_{i=1}^{n}f_2\left(x_i\right)f_{1|2}\left(x_1\mid x_i\right)}{\sum_{i=1}^{n}f_2\left(x_i\right)} \tag{9}$$

where $x_{2;\text{lower}} \leq x_i < x_{2;\text{upper}}$, so a suitable "data grid" is formed. Eqn (9) is mathematically equivalent to a mixture of distributions, where the weights are $f_2(x_i)$ and the normalized factor is $\sum_{i=1}^{n} f_2(x_i)$. In actual calculation, since n cannot be very large due to the heavy computational burden, there may be some biases for small n, even at a risk of overestimating actual P_{out} uncertainty.

B. 3-D P_{out}-WS-WD model: Vine-GMCM framework

In addition to x_1 (P_{out}) and x_2 (WS), let x_3 represent WD, which is a circular variable and usually cannot be handled as linear variables [25]. Therefore, x_3 is firstly transformed to a linear domain using (10), with output bounded between -√2 and √2, also requiring to perform truncated-to-unconstrained transformation.

$$x_{3;\text{new}} = \cos\left(\frac{2\pi \times x_{3;\text{old}}}{360}\right) + \sin\left(\frac{2\pi \times x_{3;\text{old}}}{360}\right) \qquad (10)$$

To build a correlational P_{out}-WS-WD model, as in the P_{out}-WS model, the GMMs are used as marginal distribution functions, $F_1(\cdot)$, $F_2(\cdot)$ and $F_3(\cdot)$, to transfer x_1, x_2 and x_3 into U (0,1) domain. The joint probability density of x_1, x_2 and x_3, $f_{123}(x_1, x_2, x_3)$ can be derived using chain rules: $f_{3|12}(x_3|x_1, x_2)\times f_{2|1}(x_2|x_1)\times f_1(x_1)$, in which $f_{3|12}(x_3|x_1, x_2)$ is equal to (11), with $f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)$. Therefore, $f_{123}(x_1, x_2, x_3)$ can be expressed as in (12). The key in (12) is to estimate conditional Copula density, i.e., $c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1))$, which is modelled through its arguments, i.e., $F_{2|1}(x_2|x_1)$ and $F_{3|1}(x_3|x_1)$:

$$\begin{cases} F_{2|1}(x_2 \mid x_1) = \dfrac{\partial C_{12}\left(F_1(x_1), F_2(x_2)\right)}{\partial F_1(x_1)} \\[2ex] F_{3|1}(x_3 \mid x_1) = \dfrac{\partial C_{13}\left(F_1(x_1), F_3(x_3)\right)}{\partial F_1(x_1)} \end{cases} \qquad (13)$$

Clearly, the full correlation structure is made from the three Copula pairs: C_{12}, C_{13}, and $C_{23|1}$, which are all fitted using GMCM. To estimate the conditional probability density of power output for given WS and WD values, standard conditional probability formulation and further Copula derivation can be applied, with the final full expression for $f_{1|23}(x_1|x_2, x_3)$ given in (14).

$$f_{3|12}(x_3 \mid x_1, x_2) = c_{23|1}\left(F_{2|1}(x_2 \mid x_1), F_{3|1}(x_3 \mid x_1)\right) f_{3|1}(x_3 \mid x_1) = c_{23|1}\left(F_{2|1}(x_2 \mid x_1), F_{3|1}(x_3 \mid x_1)\right) c_{13}\left(F_1(x_1), F_3(x_3)\right) f_3(x_3) \qquad (11)$$

$$f_{123}(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3) c_{12}\left(F_1(x_1), F_2(x_2)\right) c_{13}\left(F_1(x_1), F_3(x_3)\right) c_{23|1}\left(F_{2|1}(x_2 \mid x_1), F_{3|1}(x_3 \mid x_1)\right) \qquad (12)$$

$$\begin{aligned} f_{1|23}(x_1 \mid x_2, x_3) &= \frac{f_{123}(x_1, x_2, x_3)}{f_{23}(x_2, x_3)} = \frac{f_1(x_1) c_{12}\left(F_1(x_1), F_2(x_2)\right) c_{13}\left(F_1(x_1), F_3(x_3)\right) c_{23|1}\left(F_{2|1}(x_2 \mid x_1), F_{3|1}(x_3 \mid x_1)\right)}{c_{23}\left(F_2(x_2), F_3(x_3)\right)} \\ &= \frac{f_1(x_1) c_{12}\left(F_1(x_1), F_2(x_2)\right) c_{13}\left(F_1(x_1), F_3(x_3)\right)}{c_{23}\left(F_2(x_2), F_3(x_3)\right)} c_{23|1}\left(\frac{\partial C_{12}\left(F_1(x_1), F_2(x_2)\right)}{\partial F_1(x_1)}, \frac{\partial C_{13}\left(F_1(x_1), F_3(x_3)\right)}{\partial F_1(x_1)}\right) \end{aligned} \qquad (14)$$

The correlation between WS and WD is directly modelled with the additional Copula C_{23}, which is also fitted with a GMCM. If the inputs x_2 and x_3 of the model have been specified by deterministic series of values, the calculation of (14) is straightforward. If the input WS and WD are both uncertain and in forms of ranges, then similar grid search approach as in (9) could be applied:

$$f_{1|2,3}\left(x_1 \mid \begin{matrix} x_{2;\text{lower}} \leq x_2 < x_{2;\text{upper}}, \\ x_{3;\text{lower}} \leq x_3 < x_{3;\text{upper}} \end{matrix}\right) = \frac{\int_{x_{3;\text{lower}}}^{x_{3;\text{upper}}} \int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}} f_{23}(x, y) f_{1|2,3}(x_1 \mid x, y)\, dx dy}{\int_{x_{3;\text{lower}}}^{x_{3;\text{upper}}} \int_{x_{2;\text{lower}}}^{x_{2;\text{upper}}} f_{23}(x, y)\, dx dy} = \lim_{\substack{n_1 \to \infty, \\ n_2 \to \infty}} \frac{\sum_{j=1}^{n_2} \sum_{i=1}^{n_1} f_{23}(x_i, y_j) f_{1|2,3}(x_1 \mid x_i, y_j)}{\sum_{j=1}^{n_2} \sum_{i=1}^{n_1} f_{23}(x_i, y_j)} \qquad (15)$$

where $x_{2;\text{lower}} \leq x_i < x_{2;\text{upper}}$, $x_{3;\text{lower}} \leq y_i < x_{3;\text{upper}}$, and the space determined by $x_{2;\text{lower}}$ and $x_{2;\text{upper}}$, and $x_{3;\text{lower}}$ and $x_{3;\text{upper}}$ is divided into a limited number of grids. It is also not required for the input WS and WD to be already correlated with each other, because $f_{23}(x_i, y_j)$, which may be calculated by Copula as $c_{23}(F_2(x_2), F_3(x_3))f_2(x_2)f_3(x_3)$, will model the correlation between the WS and WD by assigning different weights to different (WS, WD) pairs in the "mixing" process. Another reason why analytical approach as in (8) is not applied is that the Copula distribution function for Vine Copula is unknown and it is assumed to be intractable, therefore requiring numerical integration; also, the grid search method may suffer bias issues, if values of n_1 and/or n_2 are not sufficiently large.

The only requirement during the training process is to fit marginal GMM distributions and the Copula GMCM models, [14], to evaluate (7)-(9) and (14)-(15). To have a valid distribution on the full range of x_1 (P_{out} range), array values between $x_{1;\text{lower}}$ and $x_{1;\text{upper}}$ are created, $\boldsymbol{x_1}$=[$x_{1;\text{lower}}$, $x_{1;\text{lower}}+step$, $x_{1;\text{lower}}+2step$, …, $x_{1;\text{upper}}$] and conditional distributions, (7)-(9) and (14)-(15), generalised as $f(x_1|inputs)$, are calculated for this $\boldsymbol{x_1}$ array. The results are PDF arrays for $x_{2;\text{lower}} \leq x_i < x_{2;\text{upper}}$, where renormalisation is done to ensure the integration of PDF is equal to 1. Finally, trapezoidal numerical integration is used to calculate the corresponding CDF $F(x_1|inputs)$ from the above PDF and statistical inference in terms of mean value and 5^{th}-95^{th} percentiles for $x_1|inputs$ is obtained directly from obtained PDF $f(x_1|inputs)$ and CDF $F(x_1|inputs)$.

IV. Markov Chain (MC) Model Forecasted Input WS and WD Data with Uncertainties

The MC based models presented in this section as input data at a given instant t_h use the actual measured WS(t_h) and WD(t_h) to forecast the corresponding WS and WD values at the instant t_h+mΔt and then to determine the mean P_{out} value, denoted as E[$P_{\text{out}}(t_h + \text{m}\Delta\text{t})$]. The models also calculate related uncertainty as the 5^{th}-95^{th} percentile range around mean P_{out} for given measured WS(t_h) and WD(t_h).

A. Discretization Process

Table I lists MC model discretisation classes: WS has been discretized in $N_{WS} = 27$ classes, where each class is represented by its central value, ws_i. In order to adapt WS classes to the power curve behaviour, the first and the last classes are chosen so the corresponding P_{out} is null. On the other hand, WD has been discretized in $N_{WD} = 12$ classes, each of 30 ° length and centred around the cardinal points.

TABLE I
DISCRETIZATION CLASSES FOR WD (COLUMNS 1-3) AND WS AND POUT (COLUMNS 4-7)

Class	WD [°]	CCP[a]	Class	WS [m/s]	Class	WS [m/s]
1	[345;15]	N	**1**	[0; 3]	**14**	[9; 9.5]
2	[15;45]	N-NE	**2**	[3; 3.5]	**15**	[9.5; 10]
3	[45;75]	NE-E	**3**	[3.5; 4]	**16**	[10; 10.5]
4	[75;105]	E	**4**	[4; 4.5]	**17**	[10.5; 11]
5	[105;135]	SE-S	**5**	[4.5; 5]	**18**	[11; 11.5]
6	[135;165]	S-SE	**6**	[5; 5.5]	**19**	[11.5; 12]
7	[165;195]	S	**7**	[5.5; 6]	**20**	[12; 12.5]
8	[195;225]	S-SW	**8**	[6; 6.5]	**21**	[12.5; 13]
9	[225;255]	SW-W	**9**	[6.5; 7]	**22**	[13; 13.5]
10	[255;285]	W	**10**	[7; 7.5]	**23**	[13.5; 14]
11	[285; 315]	NW-W	**11**	[7.5; 8]	**24**	[14; 14.5]
12	[315;345]	N-NW	**12**	[8; 8.5]	**25**	[14.5; 15]
			13	[8.5; 9]	**26**	[15; 25]
					27	[25; +∞]

[a] CCP denotes "Central Cardinal Point"

B. Forecasted WS Data

For WS forecasting, the first order MC (FOMC) model from [5] is used, where the state variable is the average WS over time interval $[t_{h-1\Delta t}, t_h]$ and $t_h = h\Delta t$. The FOMC model is based on the construction of a one-step transition probabilities matrix, $\bar{\bar{P}}_{WS}$, whose generic element p_{ij} represents the transition probability between the class *i* and class *j* in one Δt. The dimension of this matrix is $N_{WS}xN_{WS}$, equal to 27x27 in the considered case.

The method starts from the observed state probability vector at time t_h, whose elements are all zero, except the one corresponding to the measured value, e.g. $WS(t_h)$:

$$\pi(t_h) = \left[\pi_1(t_h), \dots, \pi_{N_{WS}}(t_h)\right]. \quad (16)$$

Afterwards, the method forecasts the state probability vector at *m* steps ahead (one step-ahead for *m*=1) as:

$$\hat{\pi}(t_{h+m\Delta t}) = \pi(t_h)[\bar{\bar{P}}_{WS}]^m. \quad (17)$$

The conditional point predictor of WS is obtained by:

$$WS(t_{h+m\Delta t})|WS(t_h) = \sum_{i=1}^{N_{WS}} ws_i\, \hat{\pi}_i(t_{h+m\Delta t}). \quad (18)$$

Eqn (18) represents the probability mass function (PMF) of the forecasted WS and can therefore be used to obtain the cumulative mass function (CMF) and the required 5th and 95th percentiles uncertainty range.

C. Forecasted WS-WD Data

Introducing a new bivariate random variable, $\overline{W} = [WS, WD]$, it is possible to obtain a one-dimensional form of random variable W, characterised by a number of classes $N_{WS}N_{WD}$, each representative of a given pair of WS and WD classes [26]. In this way, a Markov transition probability matrix can be constructed, taking as state variable W. The result is a matrix $\bar{\bar{P}}_{W}$ of dimension $(N_{WS} \cdot N_{WD})x(N_{WS} \cdot N_{WD})$, i.e. 324x324 in the considered case, whose generic element p_{ij} represents the transition probability between class *i* and class *j* of W. The introduction of a one-dimensional random variable W allows to use (16)-(18) to obtain input data for the 2D and 3D Copula models with WS and WD uncertainties.

V. CASE STUDY AND OBTAINED RESULTS

A. Measurement Datasets Available for the Analysis

Available measurements are 3-year simultaneously recorded (at WT nacelle) 10-minute average values of WS, WD and P_{out} for six individual WTs (a 3 MW doubly-fed induction generator from [13]) in an on-shore WF in Scotland, UK, Fig. 1 . The first two years of measurements are used to build the models and then MC model is used to forecast WS and WD values, together with their joint and marginal probabilities, in the third year. The data from the third year are also used to test/validate presented models.

The available measurements are post-processed in order to remove outliers and to replace missing data. The outliers in the recorded data (e.g. due to monitoring system errors) are removed (i.e. "filtered") using methodology in [27], while missing WT measurements are replaced by available WT data with the same data/time stamps. For every year, data are separated into four sets, corresponding to four seasons: "spring" (Mar-May), "summer" (Jun-Aug), "autumn" (Sep-Nov), and "winter" (Dec-Feb).

B. Metrics for Comparison of Uncertainty Ranges

The accuracy of uncertainty ranges estimated by different models is quantified in two ways: one focuses on the errors in capturing two uncertainty interval values, while another attaches more importance to the length of the uncertainty range. The following quantities are calculated by the different models and for each season.

- $\hat{P}_{out,mean}$, $\hat{P}_{out,5\%}$ and $\hat{P}_{out,95\%}$ which are the mean value predicted by the model, the 5th and 95th percentile values, respectively. These values can be assessed using goodness-of-fit indices, such as mean absolute error (MAE) and root-mean-square error (RMSE) [28]. The corresponding errors are denoted as: δu_{mean}, $\delta u_{5\%}$, and $\delta u_{95\%}$, respectively.
- In addition, the arithmetic mean value of $\delta u_{5\%}$, and $\delta u_{95\%}$, associated with the interval values estimation are calculated as the errors of estimated uncertainty interval values, using the following expression:

$$\delta u = (\delta u_{5\%} + \delta u_{95\%})/2 \quad (19)$$

If the empirical uncertainty range (see Section II) length is R_e=($P_{out,95\%}$ - $P_{out,5\%}$), and the modelled uncertainty range length is R_M=($\hat{P}_{out,95\%}$ - $\hat{P}_{out,5\%}$), the normalized uncertainty range estimation error is defined as:

$$\Delta u = (R_M/R_e) - 1 \quad (20)$$

Finally, the absolute and percentage errors in the total, overestimated and underestimated WT energy productions, denoted as E_T, E_O and E_U respectively, are also compared with the actual WT production, assuming that the mean P_{out} values are produced over each averaging window of 10 minutes.

Defining ΔP as difference between the expected and actually measured power output values at time stamp t_h:

$$\Delta P = \hat{P}_{out}(t_h) - P_{out}(t_h) \tag{21}$$

the estimation of the total energy production (as algebraic sum and sum of absolute values) is calculated for $\Delta t = 10\ min$ by trapezoidal integration formula:

$$E_T = \sum_h \frac{\Delta P(t_h)+\Delta P(t_{h+1})}{2}\Delta t, \tag{22}$$

$$E_{T(Abs.)} = \sum_h \frac{|\Delta P(t_h)|+|\Delta P(t_{h+1})|}{2} \tag{23}$$

If $A(t_h)$ is an infinitesimal area calculated between the two values, it is possible to distinguish total overestimated and underestimated productions defining two different sets of time stamps: T_o where $\hat{P}_{out}$ is greater than measured P_{out}, and T_u in opposite case.

$$E_o = \sum_{h=1}^{T_o} A(t_h)\ ;\text{and } E_u = \sum_{h=1}^{T_u} A(t_h) \tag{24}$$

C. *Results of Copula Based Models without Uncertainties in WS and WD*

Comparison of a one-day time series for a day in winter season of 2D P_{out}-WS GMCM and 3D P_{out}-WS-WD Vine-GMCM Copula model results for mean P_{out} values and corresponding uncertainty ranges when measured WS and WD values are used as model inputs (without WS and WD uncertainties) is shown in Fig. 4, together with the results obtained by using manufacturer power curve (Mfr-PC).

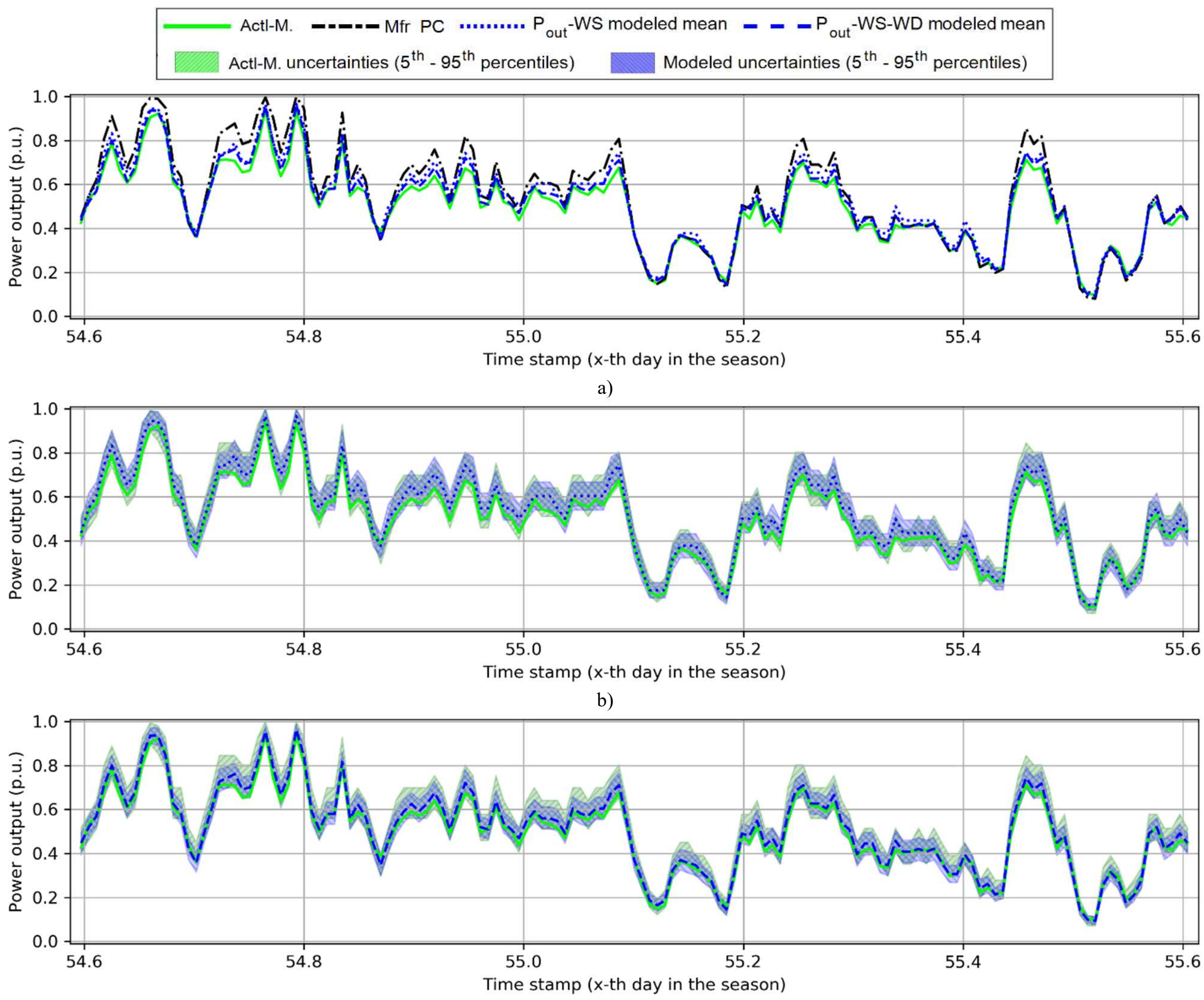


Fig. 4. Comparison of results for one whole day in winter for input WS and WD without uncertainties: a) mean values; b) 2D Copula P_{out}-WS GMCM model with P_{out} uncertainties; and c) 3D Copula P_{out}-WS-WD with P_{out} uncertainties

TABLE II

COMPARISON OF P_{OUT}-WS MODEL, P_{OUT}-WS-WD MODEL, MANUFACTURER POWER CURVE AND ACTUAL MEASUREMENTS FOR THE SAME DAY IN FIG. 4

Model	MAE (p.u.)	RMSE (p.u.)	δu (MAE)	δu (RMSE)	Δu	E_O (MWh)	E_O (%)	E_U (MWh)	E_U (%)	E_T (MWh)	E_T (%)	$E_{T(Abs.)}$ (MWh)	$E_{T(Abs.)}$ (%)
Mfr-PC	0.057	0.072	-	-	-	3.901	12.123	0.199	5.068	3.703	10.258	4.100	11.358
P_{out}-WS	0.023	0.027	0.018	0.021	-0.122	1.591	4.907	0.067	1.834	1.524	4.221	1.658	4.594
P_{out}-WS-WD	0.019	0.023	0.023	0.027	-0.221	1.257	4.120	0.126	2.256	1.130	3.132	1.383	3.831

TABLE III

COMPARISON OF P_{OUT}-WS MODEL, P_{OUT}-WS-WD MODEL, MANUFACTURER POWER CURVE AND ACTUAL MEASUREMENTS FOR WHOLE SEASONS

Model	MAE (p.u.)	RMSE (p.u.)	δu (MAE)	δu (RMSE)	Δu	E_O (MWh)	E_O (%)	E_U (MWh)	E_U (%)	E_T (MWh)	E_T (%)	$E_{T(Abs.)}$ (MWh)	$E_{T(Abs.)}$ (%)
Spring													
Mfr-PC	0.026	0.041	-	-	-	73.326	6.033	31.505	5.682	41.821	2.363	104.832	5.923
P_{out}-WS	0.019	0.028	0.014	0.019	-0.205	55.775	5.178	21.100	3.046	34.675	1.959	76.875	4.344
P_{out}-WS-WD	0.020	0.030	0.016	0.023	-0.222	58.167	5.605	22.096	3.018	36.071	2.038	80.263	4.535
Summer													
Mfr-PC	0.030	0.049	-	-	-	151.525	10.541	37.676	7.682	113.849	5.905	189.201	9.813
P_{out}-WS	0.017	0.025	0.013	0.018	-0.311	90.338	6.111	18.964	4.217	71.374	3.702	109.302	5.669
P_{out}-WS-WD	0.016	0.024	0.014	0.020	-0.328	83.514	5.758	18.298	3.831	65.216	3.383	101.812	5.281
Autumn													
Mfr-PC	0.037	0.059	-	-	-	195.288	8.476	36.986	7.800	158.302	5.698	232.274	8.360
P_{out}-WS	0.018	0.031	0.015	0.021	-0.352	84.975	4.405	28.466	3.352	56.509	2.034	113.441	4.083
P_{out}-WS-WD	0.016	0.029	0.019	0.026	-0.421	75.371	4.073	26.129	2.816	49.242	1.772	101.500	3.653
Winter													
Mfr-PC	0.029	0.044	-	-	-	92.428	6.594	66.920	8.513	25.508	1.166	159.348	7.283
P_{out}-WS	0.020	0.032	0.013	0.019	-0.235	32.164	3.801	76.583	5.708	-44.419	-2.030	108.748	4.971
P_{out}-WS-WD	0.018	0.029	0.014	0.020	-0.275	31.989	3.787	68.939	5.133	-36.950	-1.689	100.927	4.613
Whole Year													
Mfr-PC	0.031	0.050	-	-	-	512.567	8.061	173.088	7.508	339.479	3.918	685.655	7.914
P_{out}-WS	0.018	0.029	0.014	0.019	-0.284	263.252	4.938	145.113	4.354	118.139	1.364	408.365	4.713
P_{out}-WS-WD	0.017	0.028	0.016	0.022	-0.322	249.041	4.805	135.462	3.892	113.579	1.311	384.502	4.438

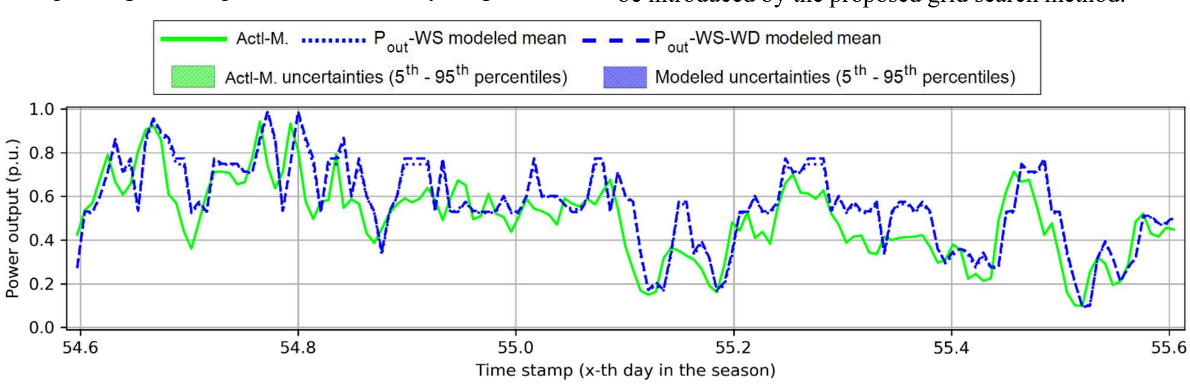
Table II quantifies the benefits of the P_{out}-WS-WD model over P_{out}-WS model for the same day in Fig. 4, while Table III compares performance of both models in separate seasons and for the whole year.

The results of 2D and 3D models are numerically close to each other, but 3D P_{out}-WS-WD model is more accurate for mean values and gives smaller uncertainty ranges, compared to 2D P_{out}-WS model, for all seasons (except for spring) and also for the whole year-length datasets. This confirms that introducing WD as an additional explanatory variable can increase the confidence in the evaluation of uncertainties in the WT's P_{out} values. For the spring season results, although P_{out}-WS-WD model still provides smaller uncertainties, its mean outputs are worse than P_{out}-WS model. One possible reason is that the WT3 yawing mechanism was faulty or not operating in the same way as during the spring seasons in the first two years; however, the WF log of events and maintenance record was not available to confirm that assumption.

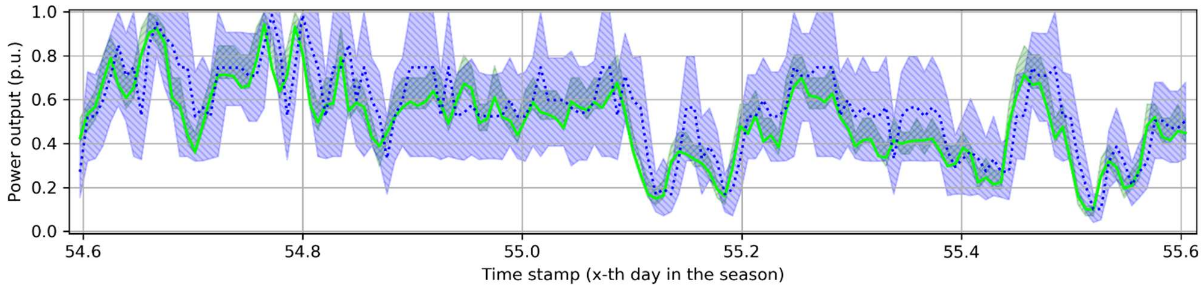
D. Results of Copula Based Models with Uncertainties in WS and WD

Next set of Copula based model results compares estimated time series for mean P_{out} values and corresponding 5th-95th percentile uncertainty ranges when input WS and WD values are provided with uncertainties (obtained from MC model). Fig. 5 and Table IV give results for the same winter day in Fig. 4 (there is no results for Mfr-PC, as it is a deterministic model and cannot handle uncertain inputs). The results confirm that the proposed grid search method allows the Copula models with uncertain WS and WD inputs to estimate related uncertainties in P_{out} values. The lag in the P_{out} series, for both mean values and uncertainty ranges, is as expected and due to the inherent feature of the forecasting processes.

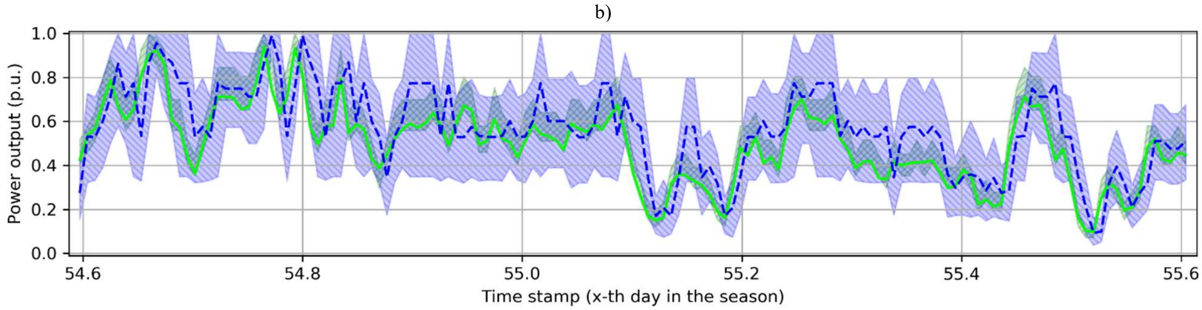
In this case, P_{out} uncertainty is evaluated considering both uncertainties in input wind energy resource (i.e. WS and WD) and uncertainties of the conversion of wind energy into electricity by WT. Consequently, 5th-95th percentile range is much wider than for the results in Fig. 4. However, the 3D P_{out}-WS-WD Copula model still provides smaller uncertainty estimation ranges, partly due to the fact that the uncertainty range is an inherent feature of Copula model. The worse estimation about mean values may be because the accuracy of mean values is more determined by the accuracy of input data: 2D P_{out}-WS model considers only WS, and therefore only errors/bias in WS data, while 3D P_{out}-WS-WD model includes errors/bias in both WS and WD data. Also, some bias may be introduced by the proposed grid search method.



a)



b)



c)

Fig. 5. Comparison of results for one whole day in winter (same as in Fig. 4) for input WS and WD with uncertainties: a) mean values; b) 2D Copula P_{out}-WS GMCM model with P_{out} uncertainties; and c) 3D Copula P_{out}-WS-WD with P_{out} uncertainties

TABLE IV
COMPARISON OF P_{OUT}-WS MODEL AND P_{OUT}-WS-WD MODEL FOR THE SAME DAY IN FIG. 5

Model	**MAE (p.u.)**	**RMSE (p.u.)**	δu **(MAE)**	δu **(RMSE)**	Δu	E_O **(MWh)**	E_O **(%)**	E_U **(MWh)**	E_U **(%)**	E_T **(MWh)**	E_T **(%)**	$E_{T(Abs.)}$ **(MWh)**	$E_{T(Abs.)}$ **(%)**
P_{out}**-WS**	0.108	0.133	0.161	0.195	2.512	6.048	24.471	1.812	15.917	4.236	11.735	7.859	21.774
P_{out}**-WS-WD**	0.114	0.141	0.160	0.194	2.494	6.606	26.930	1.691	14.615	4.915	13.617	8.296	22.984

VI. CONCLUSIONS

Due to strong and stochastic spatio-temporal variations of wind energy resource, the outputs of wind-based electricity generation systems (WEGS) are difficult to predict and are typically evaluated with associated uncertainty levels. As the WEGS contribution the total system generation capacity increases, the assessment and quantification of these uncertainties is becoming increasingly important due to their greater impact on the variations of network power flows and related uncertainties in network operating conditions.

This paper extends the initial work by the authors in [4], which investigated importance of including wind direction (WD) as an additional input explanatory variable to the wind speed (WS) for assessing uncertainty in the wind turbine (WT) power output (P_{out}). The work in [4] is extended in this follow-up paper, by providing a more comprehensive evaluation of uncertainties in WT's P_{out} by comparing a "two-dimensional" (2D) P_{out}-WS model with a "three-dimensional" (3D) P_{out}-WS-WD model for two general cases: a) for the specific or known input WS and WD values (i.e. WS and WD are without uncertainties, as in [4]), and b) for the forecasted or estimated input WS and WD values (i.e. WS and WD are both with associated uncertainties). For that purpose, 2D and 3D Copula based models from [4] are combined with 2D and 3D Markov chain (MC) models from [5], which are in this paper used to forecast the mean WS and WD values, together with their 5^{th}-95^{th} percentile ranges of variations, representing uncertainties in input WS and WD data.

The results for an actual WT demonstrate that although input WS has a stronger impact on the WT P_{out} values, the inclusion of WD in the analysis might provide noticeable improvements in models with no uncertainties in input WS and WD data, e.g., in models used for hindcasting studies.

The results for input WS and WD with uncertainties confirm that WS is a much stronger contributor to the total P_{out} uncertainty than WD, which for the considered WT provided at best only a marginal improvement. The main reason are wide ranges of variations (i.e. large spreads of uncertainties) of WD values associated with all considered WS uncertainty ranges. This part of the presented analysis and developed models allows to obtain useful information for uncertainty importance analysis, as required in e.g. wind energy forecasting studies.

The results for two considered cases (with and without uncertainties in input WS and WD data) are also useful for aleatory uncertainty evaluations, i.e. for quantifying the impact of variations of input parameters of the considered model, while the results of 2D and 3D models can be considered in terms of epistemic uncertainty evaluations, i.e. increased "knowledge" on wind energy resource through the inclusion of WD as an additional explanatory variable. In this context, separation of obtained results into seasons allowed to implicitly include temperature in the analysis, as temperature changes in different seasons will impact variations in air density, which will ultimately influence changes in WT's P_{out} for the same input WS and WD values in different seasons. Further analysis of these aspects is subject of the future work by the authors.

REFERENCES

[1] Y. Dvorkin, M. Lubin, S. Backhaus, and M. Chertkov, "Uncertainty Sets for Wind Power Generation," *IEEE Transactions on Power Systems,* vol. 31, no. 4, pp. 3326-3327, 2016.

[2] G. Papaefthymiou and D. Kurowicka, "Using Copulas for Modeling Stochastic Dependence in Power System Uncertainty Analysis," *IEEE Transactions on Power Systems,* vol. 24, no. 1, pp. 40-49, 2009.

[3] D. Fang, M. Zou, and S. Djokic, "Probabilistic OPF Incorporating Uncertainties in Wind Power Outputs and Line Thermal Ratings," in *International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, June, 2018.

[4] S. Z. Djokic, M. Zou, D. Fang, V. D. Giorgio, R. Langella, and A. Testa, "Importance of Correlating Wind Speeds and Wind Directions for Evaluating Uncertainties in Wind Turbine Power Outputs," in *International Conference on Clean Energy and Power (ICCEP)*, July, 2019.

[5] A. Carpinone, M. Giorgio, R. Langella, and A. Testa, "Markov chain modeling for very-short-term wind power forecasting," *Electric Power Systems Research,* vol. 122, pp. 152-158, May 2015.

[6] *Wind turbines –Part 12-1: Power performance measurements of electricity producing wind turbines, International Electrotechnical Commission, IEC 61400-12-1*, 2005.

[7] H. Kim, C. Singh, and A. Sprintson, "Simulation and Estimation of Reliability in a Wind Farm Considering the Wake Effect," *IEEE Transactions on Sustainable Energy,* vol. 3, no. 2, pp. 274-282, 2012.

[8] M. Lydia, A. I. Selvakumar, S. S. Kumar, and G. E. P. Kumar, "Advanced Algorithms for Wind Turbine Power Curve Modeling," *IEEE Transactions on Sustainable Energy,* vol. 4, no. 3, pp. 827-835, 2013.

[9] V. Thapar, G. Agnihotri, and V. K. Sethi, "Critical analysis of methods for mathematical modelling of wind turbines," *Renewable Energy,* vol. 36, no. 11, pp. 3166-3177, 2011.

[10] J. Tongdan and T. Zhigang, "Uncertainty analysis for wind energy production with dynamic power curves," in *2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems*, 2010, pp. 745-750.

[11] D. Villanueva and A. Feijóo, "Normal-Based Model for True Power Curves of Wind Turbines," *IEEE Transactions on Sustainable Energy,* vol. 7, no. 3, pp. 1005-1011, 2016.

[12] K. H. Hohenemser and A. H. P. Swift, "Wind Turbine Speed Control by Automatic Yawing," *Journal of Energy,* vol. 7, no. 3, pp. 237-242, 1983/05/01 1983.

[13] Vestas, "General Specification V90 - 3.0 MW," Vestas Wind Systems A/S, Report 950010.R12013.

[14] A. Tewari, M. J. Giering, and A. Raghunathan, "Parametric Characterization of Multimodal Distributions with Non-gaussian Modes," in *2011 IEEE 11th International Conference on Data Mining Workshops*, 2011, pp. 286-292.

[15] Y. Wang, D. G. Infield, B. Stephen, and S. J. Galloway, "Copula based model for wind turbine power curve outlier rejection," *Wind Energy,* Article vol. 17, no. 11, pp. 1677-1688, 2014.

[16] H. Joe, *Multivariate Models and Multivariate Dependence Concepts*. Taylor & Francis, 1997.

[17] M. Sun, I. Konstantelos, and G. Strbac, "C-Vine Copula Mixture Model for Clustering of Residential Electrical Load Pattern Data," *IEEE Transactions on Power Systems,* vol. 32, no. 3, pp. 2382-2393, 2017.

[18] Z. Wang, W. Wang, C. Liu, Z. Wang, and Y. Hou, "Probabilistic Forecast for Multiple Wind Farms Based on Regular Vine Copulas," *IEEE Transactions on Power Systems,* vol. 33, no. 1, pp. 578-589, 2018.

[19] M. Cui, V. Krishnan, B. Hodge, and J. Zhang, "A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps," *IEEE Transactions on Smart Grid,* pp. 1-1, 2018.

[20] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renewable Energy,* vol. 40, no. 1, pp. 29-39, 2012/04/01/ 2012.

[21] Q. Lu, W. Hu, Y. Min, F. Yuan, and Z. Gao, "Wind power uncertainty modeling considering spatial dependence based on Pair-copula theory," in *2014 IEEE PES General Meeting | Conference & Exposition*, 2014, pp. 1-5.

[22] Y. Wang, N. Zhang, C. Kang, M. Miao, R. Shi, and Q. Xia, "An Efficient Approach to Power System Uncertainty Analysis With High-Dimensional Dependencies," *IEEE Transactions on Power Systems,* vol. 33, no. 3, pp. 2984-2994, 2018.

[23] Z. Q. Xie, T. Y. Ji, M. S. Li, and Q. H. Wu, "Quasi-Monte Carlo Based Probabilistic Optimal Power Flow Considering the Correlation of Wind Speeds Using Copula Function," *IEEE Transactions on Power Systems,* vol. 33, no. 2, pp. 2239-2247, 2018.

[24] A. Sklar, *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959, pp. 229-231.

[25] N. I. Fisher, *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press, 1993.

[26] V. D. Giorgio, S. Z. Djokic, R. Langella, A. Testa, and M. Zou, "Two Dimensional Very Short-term Probabilistic Wind Power Forecasting based on Markov Chain Models," in *International Conference Probabilistic Methods Applied to Power Systems (to be submitted)*, August, 2020.

[27] M. Zou, D. Fang, S. Djokic, and S. Hawkins, "Assessment of wind energy resources and identification of outliers in on-shore and off-shore wind farm measurements," in *3rd International Conference on Offshore Renewable Energy (CORE)*, 2018.

[28] J. W. Cort and M. Kenji, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research,* vol. 30, no. 1, pp. 79-82, 2005.