Exploring the Interdependence of Vertical Extrapolation Uncertainties in Repowering Wind Farms

Paul Mazoyer, Thomas Duc, Georges Kariniotakis, Andreas Bechmann

Exploring the Interdependence of Vertical Extrapolation Uncertainties in Repowering Wind Farms

P Mazoyer[1,2], T Duc[2], A Bechmann[3] and G Kariniotakis[1]

[1] Mines Paris, PSL University, Centre for Processes, Renewable Energy and Energy Systems (PERSEE), 1 rue Claude Daunesse, Sophia Antipolis 06904, France
[2] ENGIE Green, 6 Rue Alexander Fleming, Lyon 69007, France
[3] DTU Wind and Energy Systems, Frederiksborgvej 399, Roskilde 4000, Denmark

E-mail: paul.mazoyer@minesparis.psl.eu

Abstract. Assessing a wind farm's annual energy production (AEP) involves modelling the wind resource and the wind-to-power conversion at the site. The greenfield pre-construction phase generally comprises the installation of wind measurement devices. For repowering projects, the wind data from the pre-construction phase of the existing farm can be used as wind input to assess the energy yield of the repowered wind farm. Indeed, one study demonstrates that when the modelling error correlations are known, the AEP prediction uncertainty of the repowered farm can be reduced by combining the energy production records of the existing farm with the AEP assessment for both farms. Previous studies have successfully identified the correlation structure for certain errors, especially for horizontal flow modelling, but not for vertical flow modelling. However, vertical extrapolation is essential, as the wind measurement heights are generally lower than the hub height on the repowered farm. This paper bridges this research gap and demonstrates that the correlation structure of errors in vertical profile modelling is Gaussian, with parameters dependent on shear values and heights. The distribution is validated against site data from simple to moderately complex sites in France.

1. Introduction

Assessing a wind farm's annual energy production (AEP) involves modelling the wind resource and the wind-to-power conversion at the site. This integrated model chain is commonly called a physics-driven model (PDM). A greenfield pre-construction phase generally comprises the installation of wind measurement devices. For repowering projects, the wind data from the pre-construction phase of the existing farm can be used as wind input to assess the energy yield of the repowered wind farm. Indeed, uncertainty can be reduced by combining the PDM of the existing and repowered farms with the same wind input and records of the existing farm's energy production [1]. However, this method requires capturing the correlation between the various PDM modelling errors when applied to two farm configurations (the existing and the repowered). Models exist for the horizontal flow modelling error [2] and the wind measurement error [3], but in general, assessing error correlations remains a challenge [4,5]. In particular, the vertical extrapolation modelling error correlation has not been investigated, which poses a significant challenge when applying the repowering method [1] because 1) repowered wind turbines are generally taller than existing turbines and 2) pre-construction wind measurement heights are generally lower than hub heights as allowed by guidelines [6]. Given these considerations, this paper introduces a model explicitly

tailored to the vertical extrapolation modelling error. It focuses on the power law model (PL) [7] that performs well in simple terrain [8, 9, 10]. Still, experimental [11, 12, 13] and theoretical studies [14, 15] show that PL error is non-negligible.

Besides the vertical error correlation coefficient, this work also assesses the error correlation structure, i.e. the univariate and joint distribution of the PL error. There is no consensus on the distributions that best describe the PDM chain's various (univariate) modelling errors. Standards and a few references [2, 16, 17, 18] assume they are null-mean Gaussian distributed, while other references [19, 20, 21] consider alternative distributions. In this paper, we demonstrate that the distribution that best represents the PL error is a Gaussian distribution. We extend the reasoning to the joint distribution that we also demonstrate to be Gaussian. We provide a model to calculate the joint distribution's mean and covariance matrix parameters.

Section 2 describes the dataset and the method used to observe PL errors. Section 3 describes the identification results for the distribution of the PL error using a pure regression model for the mean and a physics-based model for the standard deviation from [16]. Section 4 identifies the joint distribution using the Copula framework and the correlation coefficient that parametrises this distribution. Finally, section 5 concludes this work.

2. Experimental evaluation framework for the datasets

The experimental dataset consists of 12 Lidar wind measurement campaigns at different locations in France, covering from 4 to 12 months . The sites range from simple to moderately complex, possibly including patches of forests and small buildings close to the Lidar device. Data were quality-checked using expert-based visual analysis and quality checks recommended by the device manufacturer. The filtered datasets contain wind speeds as low as 0 m/s.

2.1. Power law method application to Lidar data

PL models the wind speed at height *h* using as input wind speed measurement:

$$U(h,t) = \mathrm{U}(\mathrm{h}_2,t) * \left(\frac{h}{h_2}\right)^{\alpha(t)} \tag{1}$$

where h is the target height, h_2 is one measurement height, t is the time of measurement, and α is the shear coefficient calculated as:

$$\alpha(t) = \ln\left(\frac{\mathrm{U}(h_2,t)}{U(h_1,t)}\right) * \left(\ln\left(\frac{h_2}{h_1}\right)\right)^{-1} \tag{2}$$

where h_1 is a measurement height different than h_2. Finally, the wind speed $U(h)$ and the shear coefficient α are defined as the mean over the measurement period:

$$U(h) = N_t^{-1} \sum_{t=1}^{N_t} U(h,t) \tag{3}$$

$$\alpha = N_t^{-1} \sum_{t=1}^{N_t} \alpha(t) \tag{4}$$

where N_t is the number of measurement instances. The PL error ΔU is defined as the difference between $U(h)$ and $U_{real}(h)$, the modelled and true mean wind speed at height h:

$$\Delta U(h) = U(h) - U_{real}(h) \tag{5}$$

Using the mean error instead of the error for each modelled wind speed at time *t* greatly simplifies the calculation while still providing most of the required information to model the whole wind speed time-series uncertainty. A possible addition to this analysis is modelling the error on the scale factor of the Weibull wind speed distribution, but this has not been considered here.

Each Lidar dataset is decomposed in test sets containing only three heights $\{h_1, h_2, h\}$ where $h_1 < h_2 < h$. The measurements at height $h_1$and h_2 are used to calibrate α and model $U(h)$ while the

measurements at height h are used as the reference wind speed to compute the modelling error. For example, if a Lidar measurement campaign is configured with heights of 60, 80, 100 and 120 meters, there are four possible sets:

$$\{h_1, h_2, h\} \in \{\{60,80,100\},\{60,80,120\},\{60,100,120\},\{80,100,120\}\}.$$

Given the various heights involved in the analysis, we define a normalised height called the inter-range ratio (IR), defined as the height ratio between the target height h and the measurement height h_2:

$$IR = \frac{h}{h_2} \tag{6}$$

2.2. Datasets description

For the 12 campaigns, there are 843 sets of heights. The minimum measurement height h_1 is 50 meters, the maximum target height h is 160 meters, and the IR ranges from 1.10 to 1.77. The mean shear coefficient varies from 0.15 to 0.38 (Figure 1); however, 50% of the values range between 0.22 and 0.28.

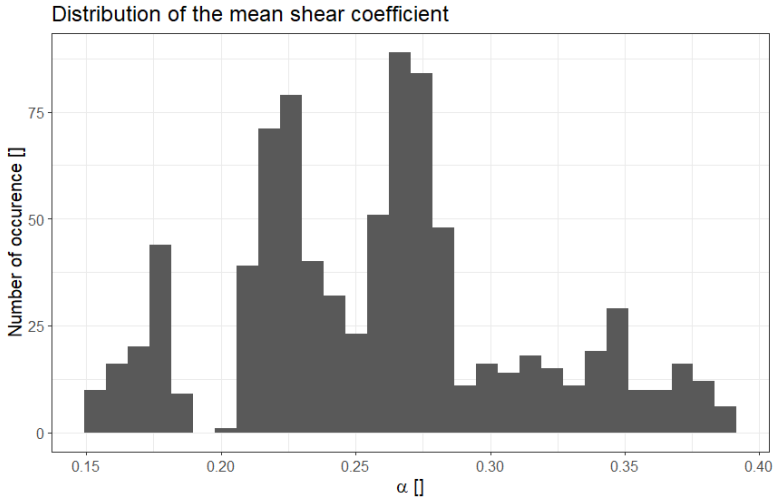

Figure 1. Histogram of the mean shear values.

2.3. Probabilistic approach

The approach to identifying the PL error distribution consists of fitting probability density function (PDF) models to the observed errors using a maximum likelihood estimation. With the PL error distribution, the joint error can be identified. The joint error is described by the bivariate cumulative distribution function (CDF) H. H outputs the probability that the modelling error at height h_e is lower than the value e_e while the modelling error at height h_r is lower than the value e_r:

$$H(e_e, e_r) = Prob(\Delta U(h_e) < \mathrm{e_e}, \Delta U(h_r) < e_r) \tag{7}$$

The Copulas framework [22] offers a convenient way to identify the joint CDF. It decomposes the CDF into a function C, that captures the correlation and the marginal distributions F:

$$H(e_e, e_r) = C(F_{\Delta U(h_e)}(e_e), F_{\Delta U(h_r)}(e_r)) \tag{8}$$

Where $F_{\Delta U(h_e)}$ and $F_{\Delta U(h_r)}$ are the marginal CDFs of $\Delta U(h_e)$ and $\Delta U(h_r)$ respectively. The following section describes the identification of F, while section 4 describes the identification of C.

3. Identification of the distribution of the PL error

3.1. Asymmetry and bias of the observed PL model error

Figure 2 shows the observed error histogram. Errors range from -0.93% (underestimation) to +3.73% (overestimation). The mean value is 0.45%, the standard deviation is 0.74%, and the

skewness is 1.66.

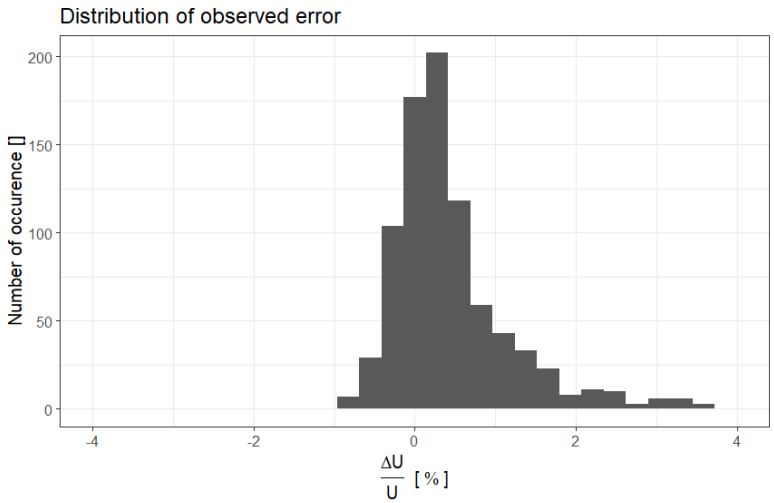

Figure 2. Histogram of the observed PL model errors.

Given the asymmetry and bias of the observed errors, the PL error is unlikely to follow a null-mean Gaussian distribution. In fact, micro-meteorological models suggest that the PL error depends on the shear values and the inter-range ratio [14], as observed when plotting the error as a function of the shear and IR values. Figure 3 shows the scatter plot of error for various shear values. For shear values above 0.2, the mean error is often above 0, while for values below 0.2, the mean error is negative. Figure 4 shows the scatter plot of error for various IR values. The mean error increases linearly from 0% to 1.25%, and the spread increases from almost 0% to 3%. There is a linear sensitivity to IR and a non-linear sensitivity to the shear values. Such systematic influences must be removed to obtain a residual error that behaves aleatory so that the central limit theorem ensures a Gaussian residual distribution.

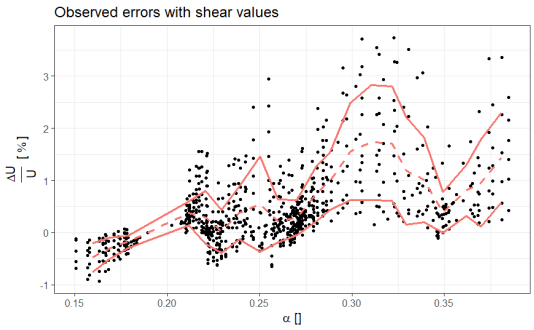

Figure 3. Scatter plot of error vs shear.

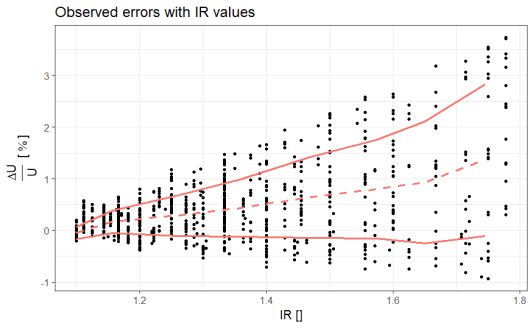

Figure 4. Scatter plot of error vs IR.

3.2. Regression model for the systematic error

The regression model f captures the expected values of PL error while leaving the residual ξ independent and Gaussian distributed:

$$\frac{\Delta U}{U} = \mathrm{f}(\mathrm{IR}, \alpha) + \xi \tag{9}$$

Among the models tested, the bilinear model fits best the PL error with a coefficient of determination of 60%:

$$f(IR, \alpha) = 0.06 - 0.05 * IR - 0.30 * \alpha + 0.27 * IR * \alpha \quad (10)$$

Figure 5 shows the distribution of the residuals ξ: the distribution is centred, the standard deviation is reduced from 0.74% to 0.47%, and the skewness (0.08) is significantly reduced compared to unadjusted errors. The distribution is close to a null-mean Gaussian distribution, as the dashed line depicts. It can reasonably be modelled by its standard deviation. Further hypothesis testing is performed in Section 3.4.

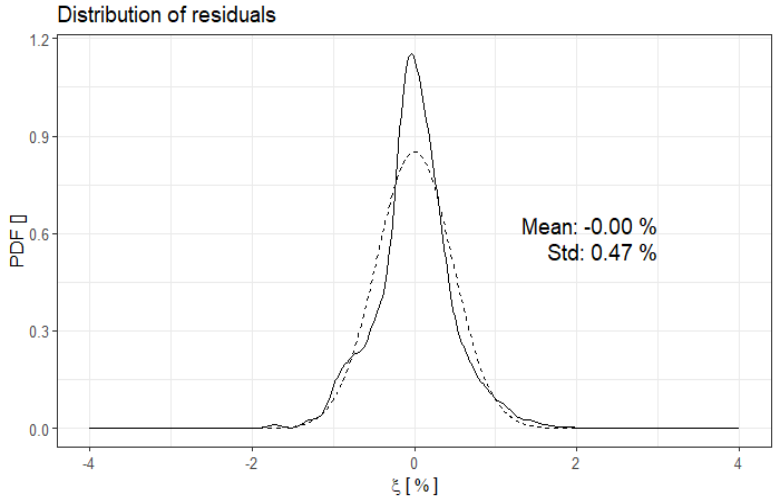

Figure 5. The experimental probability density function of the adjusted error (full line) and the best Gaussian fit (dashed line).

Figure 6 shows the scatter plot of the residuals for various shear values. The mean curve (dashed line) is flat compared to Figure 3; however, there is still a large variability. The 1-sigma envelope (full lines) shows a slight non-linear dependence of the residual standard deviation with the shear values. However, it is difficult to identify a trend visually. Figure 7 shows the scatter plot of the residuals for various IR values. The mean curve is flat; however, the 1-sigma envelope increases linearly with IR values. The residual standard deviation depends on the shear and the IR values thus, we also look for a standard deviation model.

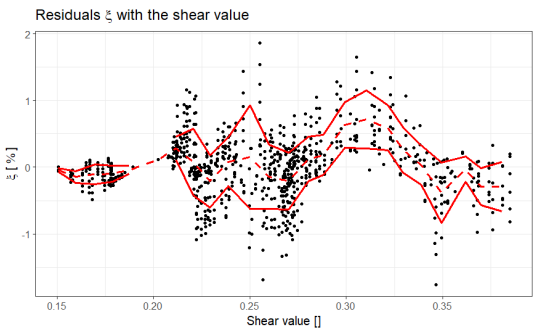

Figure 6. Scatter plot of the residuals per shear values.

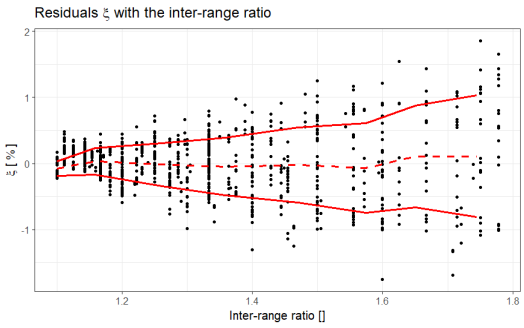

Figure 7. Scatter plot of the residuals per IR values.

3.3. Physics-driven model for the residual standard deviation

We leverage the physics-driven uncertainty model introduced in [16] to model the residual standard deviation, allowing for better generalisation of the findings compared to pure statistical regression models. The first part of this model focuses on the propagation of measurement uncertainty when calibrating the shear value and extrapolating wind speeds. The second part, stands for the uncertainty representativeness and models the error arising by assuming that the shear value is the same between the measurement heights and the hub height. The model equation is:

$$\sigma^2(\xi|\alpha, h_1, h_2, h) = \left(\frac{unc_{meas}}{U}\right)^2 \left(1 + 2 * (1 - \rho_{meas}(h_1, h_2)) \ln^2\left(\frac{h}{h_{obs}}\right) \ln^{-2}\left(\frac{h_2}{h_1}\right)\right) + \alpha^2 \sigma_{rep}^2 \ln^2\left(\frac{h}{h_{obs}}\right) \quad (11)$$

where unc_{meas} is the measurement uncertainty in metres per second, $\rho_{meas}(h_1, h_2)$ is the measurement error correlation for two different measurement heights h_1 and h_2, h_{obs} is defined as $\sqrt{h_1 h_2}$, and σ_{rep}^2 is the model for the representativeness uncertainty:

$$\sigma_{rep} = \ln^{-1}\left(\frac{\sqrt{h \cdot h_{obs}}}{z_{o,eff}}\right) c_r \, \alpha \ln\left(\frac{h}{h_{obs}}\right) + B_{os}|\alpha|^{-1}\left(1 + \tanh\left(-\frac{\alpha}{\alpha_{ref}}\right)\right) \quad (12)$$

c_r and B_{os} are heuristic parameters calibrated on data, $z_{o,eff}$ is the effective roughness length of the site, and α_{ref} is the reference shear value.

We maximise the likelihood of the residuals by tuning c_r and B_{os} and the measurement error correlation ρ_{meas} to calibrate the model. The measurement uncertainty $unc_{meas} * U^{-1}$ is set to 2% according to [3], the effective roughness length $z_{o,eff}$ to 0.8 meters, corresponding to a simple site, and the reference shear α_{ref} to 0.2 as suggested in [16]. Table 1 indicates the optimal parameters.

Table 1. Optimal parameters for the uncertainty model

Parameter	Optimal value
B_{os}	10^{-5}
c_r	1.85
ρ_{meas}	0.992

The optimal offshore parameter B_{os} is significantly lower than 0.05, the value identified in [16]. However, the reference value $B_{os} = 0.05$ found results in a very low likelihood, likely because our dataset only contains onshore data. The complex terrain parameter c_r and the correlation of the measurement error is almost 1, consistent with the findings of [16]. With such calibrated values, the uncertainty model becomes:

$$\sigma^2(\xi|\alpha, \mathrm{h}_1, h_2, h) = \left(\frac{\mathrm{unc}_{\mathrm{meas}}}{U}\right)^2 + \alpha^4 1.85 \ln^4\left(\frac{h}{\sqrt{h_1 h_2}}\right) \ln\left(\frac{\sqrt{h\sqrt{h_1 h_2}}}{0.8}\right)^{-2} \quad (13)$$

3.4. Normality assessment

Figure 8 shows that the PDF of the transformed errors (full line) and the standard Gaussian PDF (dash line) are close. The transformed error consists of the error adjusted from the systematic error and divided by the residuals standard deviation model. It is expected that the transformed error histogram resembles a Gaussian PDF if the PL error is Gaussian distributed. The transformed error distribution is slightly larger, indicating that the residual standard deviation model underestimates the errors. The normality test (Shapiro test) applied to the transformed errors gives a 10% p-value, indicating that the transformed data is likely Gaussian distributed and the PL error is likely Gaussian distributed.

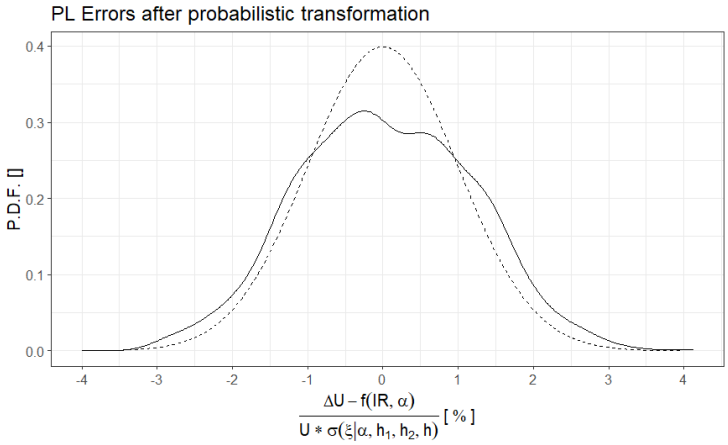

Figure 8. Experimental PDF of the transformed PL error (full line) and standard Gaussian PDF (dashed line)

Finally, this demonstrates that a reasonable model for the PL error CDF is:

$$F_{\Delta U(h)}(e) = \mathcal{N}(e|\,\mu = \mathrm{f(IR,\alpha)}, \sigma = \sigma(\xi|\alpha, \mathrm{h}_1, h_2, h)) \tag{15}$$

Where $\mathcal{N}$ refers to the Gaussian CDF.

3.5. Discussion

The calibration of the physics-based model involves arbitrary selection of values for certain parameters. Especially, the wind speed measurement uncertainty of Lidar is still an area of discussion and a value of 1% for $\mathrm{unc}_{\mathrm{meas}} * U^{-1}$ would be also plausible. The calibrated results are only slightly dependent on the latter and therefore the authors are confident on the parametrised values. Besides, the value of the optimal correlation coefficient is unexpected as it suggests that measurement errors are fully correlated across each height. The main drivers of the wind speed measurement errors are environmental parameters, such as turbulence intensity and shear value [22,23]. However, the authors also expect aleatoric errors arising from the measurement principles of Lidar. These results should be further analysed, comparing Lidar to Lidar at different heights.

The distribution parameters are strongly dependent on the shear coefficient. This implies that the PL model is prone to fail in large shear context such as very stable atmosphere while it is more consistent in small shear context such as neutral or unstable conditions. The attempt of the author to further explain the behaviour of the systematic error with the shear values remains unsuccessful. The mechanism that links stability to PL model error have theoretically been explained [14] however when assuming that shear increases from unstable, to neutral and finally stable conditions, we draw opposite conclusions to [14]. This is most likely explained by the inability to solely describe stability from the shear values, especially when using the mean shear value.

This distribution is fitted on Lidar data and similar analysis should be conducted with mast data to validate the model described here. Indeed, mast wind measurement error slightly differs to Lidar however the optimum parameters should be similar.

4. Identification of the correlation function

In this section, we identify the Copula model that best fits our dataset. Using the identified Copula and the marginal distribution identified in section 3.4, we can deduct a model for the joint PDF of PL errors.

4.1. Method and dataset for Copula identification

To identify the Copula that links the PL error at two different heights, we first pair all possible combinations of observed error for each Lidar campaigns. Considering the four-heights example of section 2.1, the pairs consist of the sets {60,80,100} and {60,80,120} then the sets {60,80,100}

and {60,100,120} and so on until all possible combinations. For such example, there are 6 possibles pairs. For the 12 campaigns, there are 1240 possibles pairs. We filter out pairs for which the difference between target height is less than 10 meters as the error correlation is likely fully correlated: {60,80,100} and {60,80,110} would be filtered out. Eventually, for the 12 Lidar campaigns, there are 835 pairs of PL errors remaining. Copulas applies to the rank of the values rather than the values themselves (see equation 8): we apply the CDF identified in section **Error! Reference source not found.** to transform values into their rank. Finally, we test several function C to find the best matching Copula. The Akaike Information Criterion (AIC) is a quality indicator that optimises both the likelihood and the complexity of the Copula model [25]. We select the model that achieves the lowest AIC.

4.2. Dataset description

Figure 9 shows the scatter plot of the pairs of errors. There is a clear correlation between the errors. Figure 10 shows the scatter plot of the transformed errors. It exhibits a clear linear trend and symmetric distribution with small tails. Only a few points are slightly outside the linear trend.

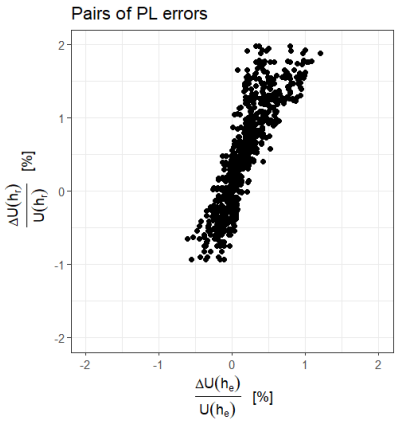

Figure 9. Scatter plot of PL error pairs.

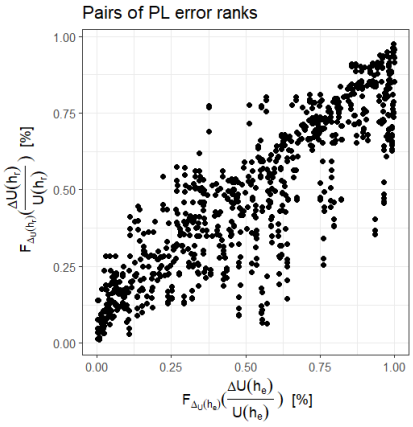

Figure 10. Scatter plot of pairs of PL errors ranks.

4.3. Copula fitting and joint CDF identification

The Gaussian Copula best fits the data of Figure 10 with an AIC of -800. The t-Student follows closely with an AIC of -799. The best AIC for the asymmetric Copulas (Joe, Gumbel, Clayton) is -633, and thus has a lower match to data than the Gaussian result. Given that the marginals are Gaussian (see section 3.4) and that the best matching Copula is Gaussian, the Sklar's theorem [26] indicates that the joint CDF is Gaussian:

$$H(e_e, e_r) = \mathcal{N}(\mathrm{e_e}, \mathrm{e_r} \mid \mu, \Sigma) \tag{16}$$

The mean values are obtained with f, the diagonal element of the covariance matrix with $\sigma(\xi)$ yet the off-diagonal element (covariance) are unknown. It is likely that the correlation decreases with height configuration: the further the extrapolation heights h_e and h_r, the lower the correlation. To reflect this, we evaluate the correlation coefficient sorting the dataset per ratio of extrapolation heights (EH):

$$EH = \frac{h_r}{h_e} \tag{17}$$

Table 2 shows the correlation coefficients derived for three intervals of EH. As intuitively expected, it decreases 95% to 80% with increasing EH.

Table 2. Pearson correlation coefficient of PL model errors.

Ratio of extrapolation height ()	Correlation coefficient (%)	Number of points ()

[1,1.2]	97%	360
[1.2,1.4]	90%	369
[1.4,1.6]	82%	106

4.4. Conclusion and discussion

The joint Gaussian CDF fits the observed PL errors. The mean and the covariance matrix parameters are dependent on the site characteristics, the mean wind profile and the heights of measurement and extrapolation. The joint CDF permits to draw sampling of the PL error when running Monte Carlo simulation of PDM. However, it turns also useful when running uncertainty propagation of PDM using Taylor's series expansion, as performed in [2]. This model only applies for onshore simple to moderately complex sites.

5. Conclusion

Characterising and mitigating uncertainties in wind farm AEP predictions represents a pivotal challenge for both industrial stakeholders and researchers. The financial viability of wind farms is intricately tied to the accuracy of prediction uncertainties. Consequently, the endeavours of ongoing research are bifurcated, with one faction concentrating on the advancement of sophisticated models for climate prediction, wind farm flows, and wind-to-power conversion. Simultaneously, another segment of the research community dedicates its efforts to comprehensively capturing the uncertainties inherent in existing models. This dual focus not only aims to enhance the prediction of uncertainties but also seeks to leverage available data to effectively diminish uncertainties, particularly in scenarios such as wind farm repowering, as elucidated in [1].

Within the context of this study, we have demonstrated that the uncertainty associated with vertical extrapolation follows a bivariate Gaussian distribution, particularly applicable to onshore sites of varying complexity. A model encapsulating all pertinent parameters of this distribution has been derived and is readily deployable. Our assessment of the model's quality, supported by statistical quality indicators, affirms its reliability. This uncertainty distribution proves particularly valuable in the context of Monte Carlo simulations for wind farm AEP prediction, especially when applying an energy yield assessment method that leverages historical data for repowering farms [1].

It is imperative to note that our results predominantly focus on the uncertainty of mean wind speed, with limited exploration into other moments of wind speed distribution. The scope of our uncertainty analysis pertains specifically to the scale parameter of a Weibull distribution of wind, omitting considerations of the shape parameter inherent in the typical Weibull distribution of wind speed. Notwithstanding, the shape parameter exerts a substantial influence on AEP values. Consequently, a prospective study could extend our analysis to incorporate this parameter, thereby enhancing the applicability of Monte Carlo simulations for wind farm AEP predictions.

Acknowledgments

The authors would like to thank Mark Kelly from DTU for his help in analysing the results and in understanding the surface layer wind profile theory.

References

[1] Mazoyer P, Duc T, Bechmann A and Kariniotakis G 2024 CONWEY – A new method for assessing the energy yield of repowered wind farms with quantified reduced uncertainty *preprint*

[2] Clerc A, Anderson M, Stuart P and Habenicht G 2012 A systematic method for quantifying wind flow modeling uncertainty in wind resource assessment *J. Wind Eng. Ind. Aerodyn.* **111** 85–94

[3] International Standard IEC61400-12-1 : Wind Turbines-Part 12-1: Power performance measurements of electricity producing wind turbines. IEC.com [Internet]. 2012 Jan 1; Available from: https://webstore.iec.ch/publication/68499

[4] Bodini N and Optis M 2020 Operational-based annual energy production uncertainty: are its components actually uncorrelated? *Wind Energy Sci.* **5** 1435–48

[5] Lee J C Y and Fields M J 2021 An overview of wind-energy-production prediction bias, losses, and uncertainties *Wind Energy Sci.* **6** 311–65

[6] Measnet: evaluation of site-specific wind conditions, version 3. Measnet.com [Internet]. 2022 Sept. Available from : https://www.measnet.com/

[7] Hellmann G Über die Bewegung der Luft in den untersten Schichten der Atmosphäre: Kgl. Akademie der Wissenschaften [G.] Reimer 1914

[8] Gualtieri G 2019 A comprehensive review on wind resource extrapolation models applied in wind energy *Renewable Sustainable Energy Rev.* **102** 215–33

[9] Lackner MA, Rogers A, Manwell JF and McGowan JG. 2010 A new method for improved hub height mean wind speed estimates using short-term hub height data *Renewable Energy* **35** 2340–7

[10] Basse A, Pauscher L and Callies D 2020 Improving vertical wind speed extrapolation using Short-Term LIDAR measurements *Remote Sensing* **12** 1091

[11] Irwin JS 1979 A theoretical variation of the wind profile power-law exponent as a function of surface roughness and stability. *Atmos. Environ.* **13** 191–4.

[12] Touma J S 1977 Dependence of the Wind Profile Power Law on stability for various locations *Journal of the Air Pollution Control Association* **27** 863–6

[13] Gualtieri G 2016 Atmospheric stability varying wind shear coefficients to improve wind resource extrapolation: A temporal analysis. *Renewable Energy* **87** 376–90

[14] Emeis S 2018 *Wind Energy Meteorology* Springer (Berlin, Heidelberg)

[15] Sedefian L 1980 On the Vertical Extrapolation of Mean Wind Power Density *J. Appl. Meteorol.* **9** 488–93

[16] Kelly M, Kersting G, Mazoyer P, Yang C, Fillols FH, Clark S, et al. 2019 Uncertainty in vertical extrapolation of measured wind speed via shear. *DTU Wind Energy E* E-**0195**

[17] Villena-Ruiz R, Ramírez F J, Honrubia-Escribano A and Gómez-Lázaro E 2018 A techno-economic analysis of a real wind farm repowering experience: The Malpica case. *Energy Convers. Manage.* **172** 182-99

[18] Liu R, Peng L, Huang G, Zhou X, Yang Q and Cai J. 2023 A Monte Carlo simulation method for probabilistic evaluation of annual energy production of wind farm considering wind flow model and wake effect *Energy Convers. Manage.* **292** 117355

[19] Bleeg J, Purcell M, Ruisi R and Traiger E 2018 Wind farm blockage and the consequences of neglecting its impact on energy production. *Energies* **11** 1609

[20] Afanasyeva S, Saari J, Kalkofen M, Partanen J and Pyrhönen O 2016 Technical, economic and uncertainty modeling of a wind farm project. *Energy Convers. Manage.* **107** 22–33.

[21] Dong M, Li Y, Song D, Yang J, Su M, Deng X, et al. 2021 Uncertainty and global sensitivity analysis of levelized cost of energy in wind power generation. *Energy Convers. Manage.* **229** 113781

[22] Nelsen R B 1999 *An Introduction to Copulas* Springer (Berlin, Heidelberg)

[23] Clive P. 2008 Compensation of vector and volume averaging bias in lidar wind speed measurements *J. Phys. Conf. Ser.: Earth and Environmental Sci.* **1** 012036.

[24] Rosenbusch P, Mazoyer P, Pontreau L, Allain P and Cariou JP 2021 Wind speed reconstruction from mono-static wind lidar eliminating the effect of turbulence. *J. Renewable Sustainable Energy* **13** 063301.

[25] Akaike H 1974 A new look at the statistical model identification. *Trans. Autom. Control* **19** 716–23.

[26] Härdle W K and Simar L 1988 *Applied Multivariate Statistical Analysis*. Springer (Berlin, Heidelberg)