

PAPER • OPEN ACCESS

Design exploration prior to blade multi-disciplinary optimisation

To cite this article: Fernando Echeverria *et al* 2018 *J. Phys.: Conf. Ser.* **1037** 042014

View the [article online](#) for updates and enhancements.

Related content

- [Progress in Multi-Disciplinary Data Life Cycle Management](#)
C Jung, M Gasthuber, A Giesler et al.
- [Crystal assisted experiments for multi-disciplinary physics with heavy ion beams at GANIL](#)
Denis Dauvergne
- [Cluster Optimisation using Cgroups at a Tier-2](#)
G Qin, G Roy, D Crooks et al.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Design exploration prior to blade multi-disciplinary optimisation

Fernando Echeverria^{1,2}, Fermin Mallor² and Unai San-Miguel¹

¹ACCIONA Windpower-NORDEX, Avenida Innovación 3, Pamplona, 31621, Spain

²UPNA Public University of Navarra, Pamplona, 31006, Spain

FEcheverria@nordex-online.com

Abstract. The approach of designing blades as a multi-disciplinary, holistic optimisation implies significant challenges owing to the high complexity of the involved factors such as aerodynamics, elasticity, controller and loads. Moreover, the large number of design variables complicates the intuitive analysis of the relationship between the design variables and responses. This paper presents the design variable exploration prior to blade optimisation, which reveals certain design variable combinations that lead to undesirable dynamic load amplification. Statistical tools, such as multiple logistic regression and fast and frugal decision trees, are applied to identify the conditions causing the phenomenon and predict the possible appearance under new design variable combinations.

1. Introduction

The significant increase in wind turbine sizes over the past several years makes efficient blade design fundamental. The traditional methodology of dividing the blade design into separate problems—airfoil geometry, spanwise distributions, structural layout design and load calculations—is evolving into a multi-disciplinary design optimisation that interconnects different disciplines [1][2][3][4]. The challenge of dimensionality [5] applied in optimisation problems refers to the complexity and high computational cost that arises when increasing the number of design variables.

Prior to optimisation, design exploration comprises several statistical techniques to investigate the design space, with the intent of obtaining knowledge to understand the relationships between design variables and outputs. Furthermore, this aids in detecting outliers and anomalies in the performance under determinate variable combinations [6].

Statistical methods have been used extensively for wind turbine load assessment in recent research, to overcome the complexity and high computational costs. For example, Monte Carlo techniques are frequently applied to deal with computationally expensive offshore fatigue assessment [7]. Long-term ultimate load assessment requires statistical support to estimate loads within a 50-year recurrence period [8][9][10]. In the research of Murcia et al. [11], polynomial surrogate models based on global sensitivity analysis were employed to characterise the energy production and fatigue loads as a function of inflow wind conditions.

In this study, statistical tools are applied to scan the design space, and identify and predict anomalous loading responses. Firstly, Latin hypercube sampling (LHS) [12] is applied to generate sets of design variables belonging to blade geometry, structural layout and wind turbine control efficiently. For each observation, the blade geometry and structure are constructed, and relevant blade properties



are calculated and introduced in an aeroelastic wind turbine model subjected to steady and dynamic load calculations following wind turbine design standards [13]. For these aero-elastic calculations, the ultimate and fatigue damage equivalent loads [14] are obtained from the main forces/moments in the wind turbine. Visual inspection with violin plots [15] reveals severe dynamic load amplification when complex design variable combinations take place. Statistical methods such as multiple logistic regression [16] and fast and frugal trees (FFTrees) [17] are applied to create a surrogate model and identify the design space that should be avoided to prevent undesirable responses in future blade optimisation problems.

2. Methods

2.1. Design variables

Blade geometry is conceived as a lofted B-spline [18] surface generated from aerofoil curves that are transformed according to longitudinal laws in the form of a B-spline: stacking, chord, twist, thickness, and pre-bending. The control points of the B-splines are set as design variables. Moreover, the thickness distribution of the main spar cap along the span is defined by three design variables. Finally, the torque-speed gain and gearbox ratio [19] are also established as design variables.

2.2. Latin hypercube sampling

LHS is a stratified sampling method that covers the design space efficiently with a reduced number of observations [12]. The statistical software R [20] is used for the LHS calculation, prior to which data scaling and centring are performed.

2.3. Calculation of loads

B-spline longitudinal laws are constructed for each observation provided by LHS, and the subsequent blade lofted surface is generated. In conjunction with the blade geometry, a new spar cap thickness distribution is introduced with a classic laminate theory method [21] in order to obtain the structural properties of the blade. The aeroelastic wind turbine model in Bladed [22] is updated according to the new blade geometry, structural properties, torque-speed gain and gearbox ratio. The model used in this research is a three-blade, horizontal-axis, upwind, variable speed and active pitch-regulated wind turbine.

Following the completion of the aeroelastic model, the Eigen-frequencies and modes are calculated and time domain simulations are run according to wind turbine design standards. The simulations are power production DLC1.2 load cases with a normal turbulent wind model [13]. Start-up, idling, normal shut-down and fault simulations are excluded from the present research. The simulations cover the wind turbine operation range and different wind inflow misalignments. The resultant loads in the main wind turbine locations, tower base, tower top, hub centre, blade root and several blade span stations are post-processed to obtain a set of ultimate loads. Moreover, load signals are processed in terms of constant amplitude fatigue cycles using the rainflow counting method. The damage equivalent load (DEL) is calculated on the basis of the Palmgren-Miner method, as the range of a sinusoidal load of constant frequency that would produce the same fatigue damage as the original signal [14]. Finally, steady calculations are completed for the wind speed range and loads of interest.

2.4. Violin plots

The ultimate loads and DELs are inspected by means of violin plots [15]. A violin plot is a combination of a box plot and a rotated kernel density plot. The bottom and top of the box represent the first and third quartiles, respectively, while the horizontal line inside the box represents the median. The ends of the whiskers represent the minimum and maximum of all data. The kernel density non-parametric function estimates the variable probability density function.

The violin plot of the DEL of the Mx moment (edgewise) in the blade root, presented in **Figure 1(a)**, reveals a high number of outliers within the observations. Furthermore, **Figure 1(b)** indicates

extreme Mx moments in the blade root, resulting from steady calculations, versus extreme values from the DLC1.2 calculations. Most of the observations exhibit an accurate linear correlation between the steady and dynamic conditions; however, a significant number of observations with high dynamic load amplification exist. Accordingly, these are classified as *normal* or *abnormal* when the linear correlation between the steady and dynamic extremes exists or not, respectively.

Figure 2(a) illustrates the Mx time history of a specific observation classified as *abnormal*. This specific load case exhibits an 8 m/s mean wind speed and 0° mean wind direction. **Figure 2(b)** presents the auto-spectral density of the same signal. The amplified response suggests the existence of a certain type of aeroelastic instability [23].

Following observation of the phenomenon, the assessment is focused on avoiding undesirable amplified responses by identifying the combination of design variables leading to its appearance. A detailed classification of the type of aeroelastic instability is beyond the scope of this paper and should be assessed separately.

Two statistical methods are employed to predict the existence of the abnormal situation: multiple logistic regression and FFTrees.

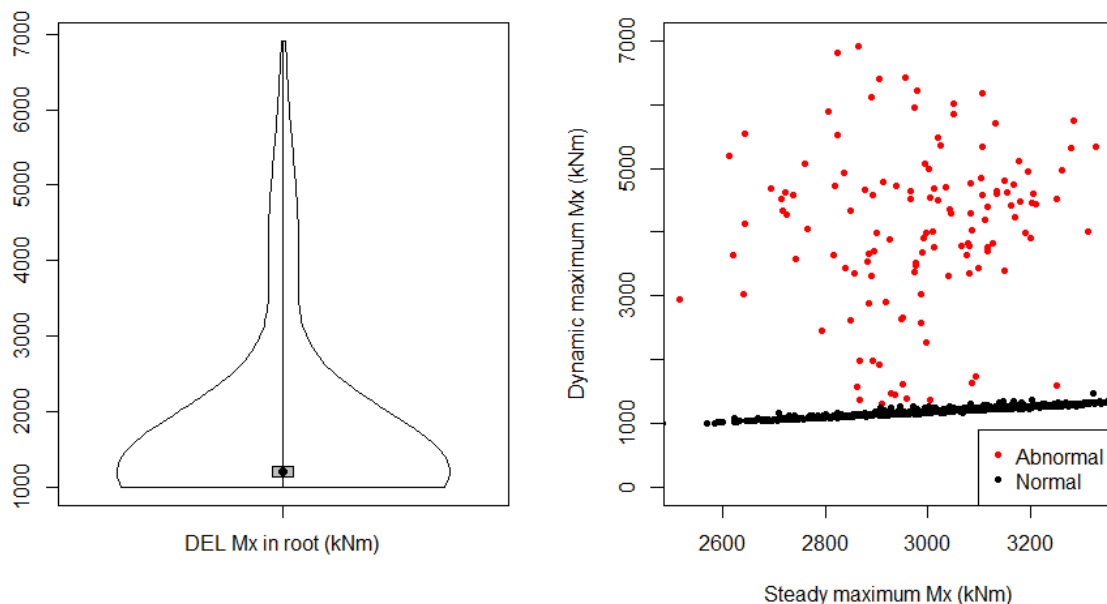


Figure 1 (a) Violin plot of DELs of Mx moment in blade root. (b) Maximum Mx moment in blade root from steady calculations versus maximum Mx moment in blade root from DLC1.2 calculations. Each point corresponds to one observation generated using LHS.

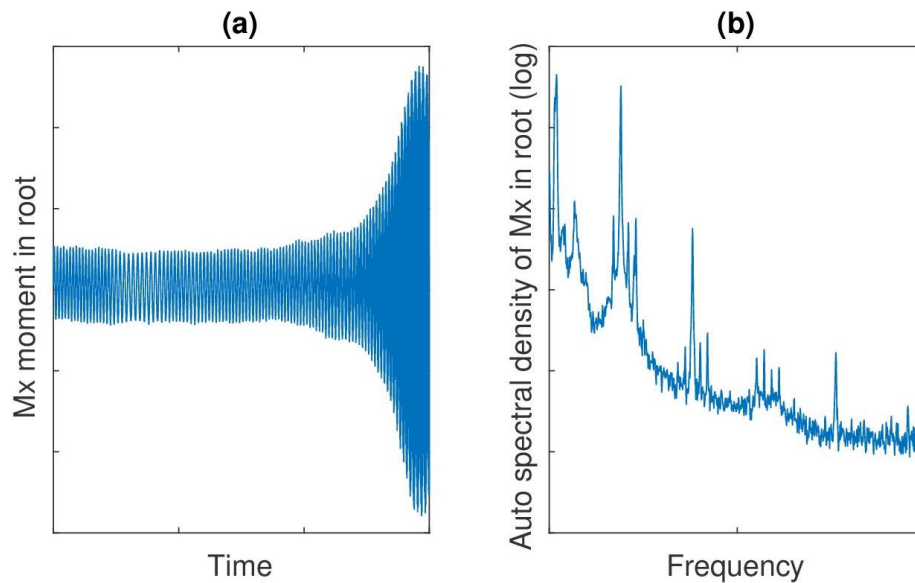


Figure 2 (a) Time history of Mx moment in root of specific observation classified as *abnormal* (8 m/s mean wind speed and 0° mean wind misalignment). (b) Auto-spectral density of Mx moment in root of specific observation classified as *abnormal* (8 m/s mean wind speed and 0° mean wind misalignment).

2.5. Overfitting in prediction models

Overfitting refers to the phenomenon in complex approximation models whereby the known data are fit with high accuracy, but new unknown observations are poorly predicted [16]. To avoid this phenomenon, the observations are randomly divided into two groups. The first, larger group serves to train the prediction model; the second group is used for accuracy validation and assessment under unknown data.

2.6. Multiple logistic regression

Prior to the statistical treatment, a binary response known as *abnormal* is artificially created and assigned to the calculated observations. The unity represents the existence of the dynamic amplification observed in subsection 2.4.

When assessing the relationship between the design variables and a binary response, linear regression is not an appropriate prediction model. In contrast, multiple logistic regression is convenient for this type of analysis [16][24].

The multiple logistic regression follows the equation:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p,$$

where X is the array of p design variables, β_i denotes the regression coefficients and $p(X)$ is the probability of obtaining an observation that produces a response classified as *abnormal*:

$$p(X) = \Pr(\text{abnormal} = 1|X).$$

The maximum likelihood method is used to estimate the regression coefficients. The significance of each design variable is determined by means of a contrast hypothesis test performed on the associated regression coefficients [16].

2.7. Fast and frugal trees

Decision trees are methods that stratify the design space into regions by following a tree scheme to predict a binary response. The cue-based questions in the tree are known as nodes, the answers to these questions are branches and the leaves are the decisions. The main advantage offered by these

methods is simplicity of interpretation. Nevertheless, the accuracy level is generally lower than that of more complex methods [16].

Among the different techniques related to decision trees, such as bagging, random forests and boosting, the novel FFTree method is selected [25]. For the fundamentals of this method, [25] or [26] can be consulted. The advantages of FFTrees with respect to other techniques are that they are fast, frugal, simple and transparent algorithms. Moreover, they are robust against overfitting [25].

2.8. Confusion table

The confusion table displayed in **figure 3** presents the results of a binary prediction model. The rows refer to decisions according to the model and the columns, to true values. Accordingly, the diagonal provides correct decisions and the off-diagonal provides erroneous values [25]. The results are classified according to the following nomenclature:

- Hits (hi) are response predictions as 1 that are correct.
- Correct rejections (cr) are response predictions as 0 that are correct.
- Misses (mi) are response predictions as 0 that are incorrect.
- False alarms (fa) are response predictions as 1 that are incorrect.

		Truth	
		1	0
Prediction	1	hi	fa
	0	mi	cr

Figure 3 General confusion tables of binary prediction models.

Based on the confusion table, the following metrics assess the predictor performance.

- Sensitivity (sens). Probability of correctly identifying a true positive case:

$$sens = \frac{hi}{hi + mi}$$

- Specificity (spec). Probability of correctly identifying a true negative case:

$$spec = \frac{cr}{cr + fa}$$

- Accuracy (acc). Probability of correctly identifying any case:

$$acc = \frac{hi+cr}{hi+cr+mi+fa}$$

At present, the threshold probability for assigning a prediction as unity is defined as 0.5. However, this probability can be increased to improve specificity, although this implies an associated sensitivity reduction.

3. Results

3.1. Multiple logistic regression

The model is fit using training data. A subsequent contrast hypothesis test performed reveals that variables involved in the pre-bending curve and blade thickness along the span present the highest impact in the binary response.

Regarding the prediction accuracy, **figure 4** displays the confusion table for a validation group of observations, while **table 1** presents the performance metrics.

		Truth	
		1	0
Prediction	1	18	6
	0	10	167

Figure 4 Confusion table of multiple logistic regression performed for validation observations.

3.2. *FFTree*

The *FFTree* obtained from the training data is presented in **figure 5**. Only three variables are relevant in the prediction model: one coefficient and knot of the pre-bending B-spline, and one coefficient of the B-spline thickness. **Figure 6** illustrates two pre-bending B-spline curves that are predicted as *abnormal* or *normal* according to the *FFTree*. Furthermore, the confusion table in **figure 7** displays the predictor performance for the validation observations.

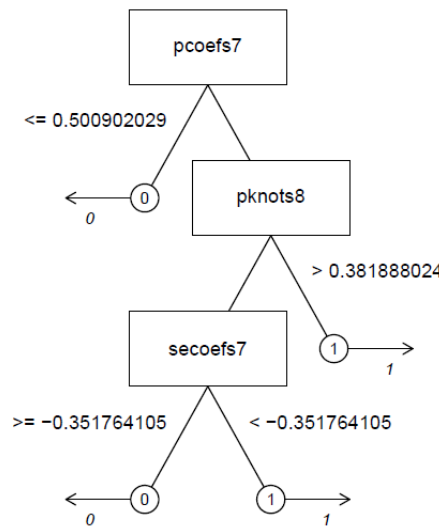


Figure 5 *FFTree* for predicting binary *abnormal* response (=1).

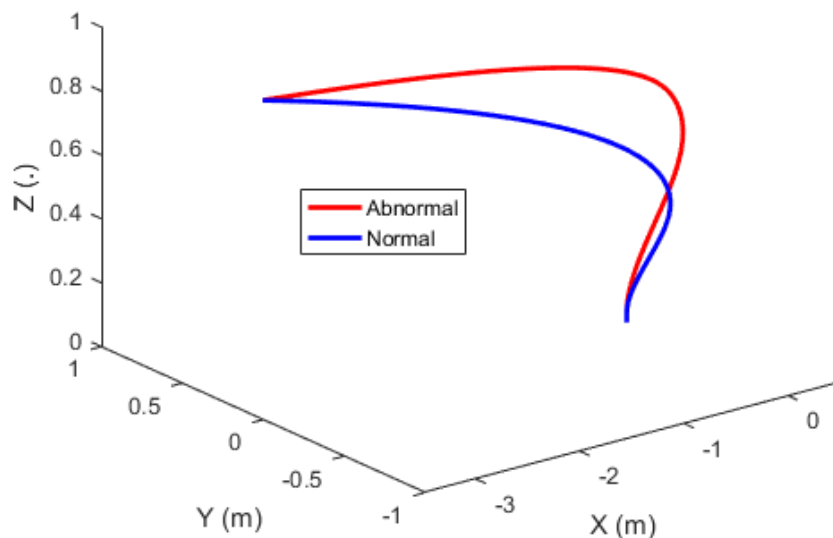


Figure 6 Example cases of pre-bending curves that represent associated abnormal or normal dynamic responses in the DLC1.2 load cases.

		Truth	
		1	0
Prediction	1	25	29
	0	6	140

Figure 7 Confusion table for *FFTree* analysis performed for validation observations.

The efficiency metrics for both prediction models are summarised in **table 1**.

Model	Sensitivity	Specificity	Accuracy
Logistic ($p=0.5$)	79%	94%	92%
Logistic ($p=0.85$)	64%	97%	93%
FFTree	81%	83%	83%

Table 1 Efficiency metrics of prediction models for validation observations (logistic regression with probability $p=0.5$, logistic regression with probability $p=0.85$ and FFTree).

4. Discussion

The appropriate prediction model sensitivity and specificity depend on the prediction purpose, and represent a solution of compromise [16]. If the model is used to specify a constraint in a subsequent optimisation problem, it is preferable to have a low number of false alarms (high specificity) and high number of misses (low sensitivity) than vice versa. One missed prediction provided incorrectly by the optimisation as a feasible solution can be identified afterwards by performing aeroelastic simulation to verify the prediction goodness. Nevertheless, a false alarm means a possible feasible solution that the optimisation has incorrectly rejected, without possible knowledge by the designer.

In terms of multiple logistic regression, the accuracy for unknown observations is high (92%), while the specificity (94%) is higher than the sensitivity (79%). When the threshold probability is changed to 0.85, the specificity changes from 94% to 97%; however, this implies a significant sensitivity loss, from 79% to 64%.

The FFTree provides a less accurate set of predictions (83%), with balanced sensitivity and specificity. However, the model simplicity offers an advantage for implementation in an optimisation problem constraint and for understanding the phenomenon.

5. Conclusions

Design exploration is performed as a prior stage to the multi-disciplinary optimisation of the blade, in order to obtain important information regarding the design space. Blade geometry, structural layout and control features are sampled using the LHS technique to cover the design space efficiently. As the result of aeroelastic calculations, a set of steady, ultimate and fatigue loads are obtained. A visual inspection with the aid of violin plots reveals high dynamic load amplification in certain complex design variable combinations. The consequent excessive loading should be avoided in the design.

Multiple logistic regression and the FFTree are fit and validated to predict the conditions causing the phenomenon. Although multiple logistic regression provides a more accurate prediction, the FFTree presents a visual tree scheme with easy implementation. In terms of the model sensitivity and specificity, further considerations should be made. For the model implementation in an optimisation problem, a high specificity is preferred as a constraint to avoid undesired load amplification, as one missed prediction can be identified afterwards.

The predicting models identify the pre-bending and blade thickness B-spline coefficients as those involved in the instability phenomenon.

Classification of the dynamic load amplification should be assessed in separate future research.

References

- [1] Bottasso CL, Campagnolo F, Croce A. Multi-disciplinary constrained optimization of wind turbines. *Multibody System Dynamics* [Internet]. Springer Netherlands; 2012;27(1):21–53.
- [2] Pavese C, Tibaldi C, Zahle F, Kim T. Aeroelastic multidisciplinary design optimization of a swept wind turbine blade. *Wind Energy* [Internet]. 2017;20(12):1941–1953. Available from: <http://dx.doi.org/10.1002/we.2131>

- [3] Echeverria F, Mallor F, San Miguel U. Global sensitivity analysis of the blade geometry variables on the wind turbine performance. *Wind Energy* 2017. DOI:10.1002/we.2111.
- [4] Echeverria F, Mallor FM, Sanz J. Fast Estimation of the Damage Equivalent Load in Blade Geometry Multidisciplinary Optimization. *Journal of Solar Energy Engineering* [Internet]. ASME; 2017 May;139(4):-10. Available from: <http://dx.doi.org/10.1115/1.4036636>
- [5] Bellman R, Corporation R. *Dynamic Programming* [Internet]. Princeton University Press; 1957. Available from: <http://books.google.it/books?id=wdtoPwAACAAJ>
- [6] Nist-Sematech. *Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/>, April 2012.
- [7] Graf PA, Stewart G, Lackner M, Dykes K, Veers P. High throughput computation and the applicability of Monte Carlo integration in fatigue load estimation of floating offshore wind turbines. *Wind Energy* [Internet]. 2016;19(5):861–872. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/we.1870>
- [8] Pan Q, Byon E. Adaptive Extreme Load Estimation in Wind Turbines. *AIAA SciTech Forum* [Internet]. American Institute of Aeronautics and Astronautics; 2017. Available from: <https://doi.org/10.2514/6.2017-0679>
- [9] Fogle J, Agarwal P, Manuel L. Towards an improved understanding of statistical extrapolation for wind turbine extreme loads. *Wind Energy* 2008;11(6):613–635. DOI:10.1002/we.303.
- [10] Graf P, Damiani R, Dykes K, Jonkman JM. Advances in the Assessment of Wind Turbine Operating Extreme Loads via More Efficient Calculation Approaches. *AIAA SciTech Forum* [Internet]. American Institute of Aeronautics and Astronautics; 2017. Available from: <https://doi.org/10.2514/6.2017-0680>
- [11] Murcia JP, Réthoré P-E, Dimitrov N, Natarajan A, Sorensen JD, Graf P, et al. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. *Renewable Energy* [Internet]. 2018;119:910–922. Available from: <http://www.sciencedirect.com/science/article/pii/S0960148117306985>
- [12] McKay MD. Latin Hypercube Sampling As a Tool in Uncertainty Analysis of Computer Models. *Proceedings of the 24th Conference on Winter Simulation* [Internet]. Arlington, Virginia, USA: ACM; 1992. p. 557–564. Available from: <http://doi.acm.org/10.1145/167293.167637>
- [13] IEC61400-1 Wind turbines part1 design requirements. Edition3. International Electrotechnical Commission; 2005
- [14] Passipoularidis V, Brøndsted P. *Fatigue Evaluation Algorithms: Review*. Danmarks Tekniske Universitet, Risø Nationallaboratoriet for Bæredygtig Energi; 2010.
- [15] Adler D. *vioplot: Violin plot* [Internet]. 2005. Available from: <http://wsopuppenkiste.wiso.uni-goettingen.de/dadler>
- [16] James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. G. Casella IO S. Fienberg, editor. Springer New York; 2013.
- [17] Gigerenzer G, Todd P. Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart*. Oxford University Press. 1999;3–34.
- [18] Piegel L, Tiller W. *The NURBS Book*. 2nd ed. Springer 1995, editor. Berlin, Heidelberg; 1997.
- [19] Munteanu I. Wind turbine control systems. Principles, modelling and gain scheduling design. *International Journal of Robust and Nonlinear Control* [Internet]. John Wiley & Sons, Ltd.; 2008;18(7):796–797. Available from: <http://dx.doi.org/10.1002/rnc.1263>
- [20] Team RC. *R: A Language and Environment for Statistical Computing* [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available from: <https://www.R-project.org/>
- [21] Bir G, Damiani R. *PreComp* [Internet]. 15013 Denver West Parkway Golden, CO 80401 303-275-3000: NREL National Renewable Energy Laboratory; Available from: <https://nwtc.nrel.gov/PreComp>
- [22] *BLADED* [Internet]. St Vincent’s Works, Silverthorne Lane, Bristol BS2 0QD England: Garrad

Hassan & Partners Ltd; Available from: <https://www.dnvgl.com/services/bladed-3775>

- [23] Holierhoek JG. Aeroelasticity of Large Wind Turbines. Technische Universiteit Delft; 2008.
- [24] Jr FEH. rms: Regression Modeling Strategies [Internet]. 2017. Available from: <https://CRAN.R-project.org/package=rms>
- [25] Phillips ND, Neth H, Woike JK, Gaissmaier W. FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. *Judgment and Decision Making*. 2017;12(4):344–368.
- [26] Gigerenzer G, Czerlinski L J.and Martignon. How good are fast and frugal heuristics? In *Decision science and technology*. Springer. 1999;81–103.