How generalizable is a machine-learning approach for modeling hub-height turbulence intensity?

Nicola Bodini[1], Julie K. Lundquist[1,2,3], Hannah Livingston[2] and Patrick Moriarty[1]

[1] National Renewable Energy Laboratory, Golden, Colorado 80401, USA
[2] University of Colorado Boulder, Boulder, Colorado 80302, USA
[3] Renewable and Sustainable Energy Institute, Boulder, Colorado 80303, USA

E-mail: `nicola.bodini@nrel.gov`

Abstract. Hub-height turbulence intensity is essential for a variety of wind energy applications. However, simulating it is a challenging task. Simple analytical models have been proposed in the literature, but they all come with significant limitations. Even state-of-the-art numerical weather prediction models, such as the Weather Research and Forecasting model, currently struggle to predict hub-height turbulence intensity. Here, we propose a machine-learning-based approach to predict hub-height turbulence intensity from other hub-height and ground-level atmospheric measurements, using observations from the Perdigão field campaign and the Southern Great Plains atmospheric observatory. We consider a random forest regression model, which we validate first at the site used for training and then under a more robust round-robin approach, and compare its performance to a multivariate linear regression. The random forest successfully outperforms the linear regression in modeling hub-height turbulence intensity, with a normalized root-mean-square error as low as 0.014 when using 30-minute average data. In order to achieve such low root-mean-square error values, the knowledge of hub-height turbulence kinetic energy (which can instead be modeled in the Weather Research and Forecasting model) is needed. Interestingly, we find that the performance of the random forest generalizes well when considering a round-robin validation (i.e., when the algorithm is trained at one site such as Perdigão or Southern Great Plains) and then applied to model hub-height turbulence intensity at the other location.

1. Introduction

As wind energy expands and becomes an increasingly large portion of energy portfolios, collecting data on atmospheric turbulence at turbine hub height is essential for optimizing the performance of wind power plants and their successful integration into the electric system. Hub-height turbulence intensity (TI) impacts turbine power generation [1, 2, 3, 4, 5, 6], and is instrumental in the control strategy for wake modeling and wake-steering strategies [7, 8]. Turbine siting and power forecasting also rely on hub-height turbulence and are determining factors of whether a wind power plant will be financially advantageous [9, 10].

Various formulations have been proposed for estimating hub-height TI, or TI profiles, as a function of surface quantities. The Mann model [11, 12] provided an analytic characterization of turbulence profiles in the presence of uniform linear shear. Cheung et al. [13] proposed and tested a formulation based on assumptions for turbulent boundary layers in a range of stability conditions. Building from the power-law approach for extrapolating wind speeds from surface-level measurements to hub height, Gualtieri [14] proposed and tested an empirically based approach for extrapolating 30-m measurements to 100-m altitudes and found that it is critical to consider atmospheric stability. Türk and Emeis [15] showed how offshore turbulence intensity has a strong dependence on wind speed and wave height, and proved that specifications given by the International Electrotechnical Commission (IEC) wind turbine standards 61400-1 [16] and 61400-3 [17] are inaccurate for their test location. All these efforts come with specific limitations in their range of applicability, so a more universal model for hub-height TI is currently not available.

Further, numerical weather prediction models do not routinely predict TI. For example, the state-of-the-art numerical weather prediction model, the Weather Research and Forecasting (WRF) model, [18, 19]), does not provide TI as output. Only some choices of planetary boundary layer schemes (i.e., the MYNN scheme [20, 21]) even provide a prognostic calculation of turbulence kinetic energy (TKE), while others widely used for wind energy applications (i.e., the YSU scheme [22]) do not even estimate TKE. And in any case, the spatial scales at which numerical weather prediction models are run are too coarse for some applications, such as turbine control, where the knowledge of hub-height TI is essential.

Given the complexity and inherent nonlinearity of atmospheric turbulence, machine-learning techniques could be leveraged for the derivation of a model for hub-height TI. Machine-learning approaches have been successfully applied to several wind energy applications, ranging from power forecasting [23, 9] to issues at the wind turbine scale, and for turbine power curve modeling [24], turbine faults and controls [25], and turbine blade management [26]. Recently, machine learning has also been successfully applied to vertically extrapolate wind speed [27, 28, 29], both on land and offshore, to better represent the turbulence dissipation rate [30] and to better parameterize surface fluxes [31, 32].

In this analysis, we develop and perform a thorough validation of a machine-learning approach to model hub-height TI, calculated as a 30-minute average as:

$$TI = \frac{\sigma_U}{\overline{U}}, \tag{1}$$

where σ_U is the standard deviation of the horizontal wind speed, and $\overline{U}$ is the 30-minute average horizontal wind speed. We consider observations from the Perdigão field campaign [33] and the U.S. Department of Energy Atmospheric Radiation Measurement Southern Great Plains (SGP) observatory [34], which are both presented in Section 2, together with a description of the proposed machine-learning method. In Section 3, we first derive physical insights on the variability of hub-height TI from the application of the proposed approach at a single site. Next, we test the generalization accuracy of the model when adopting a more practical validation setup, which resembles the use of the WRF model, where the algorithm is trained at one location and then applied to model hub-height TI at a different site. Finally, we summarize our results and suggest future work in Section 4.

2. Data and methods

2.1. Observational data sets

The first set of observations we consider is taken from the Perdigão field campaign [33], an international effort that brought scientists from multiple institutions to a valley in central Portugal (see map in figure 1a), where an intensive observational period was completed from 1 May 2017 to 15 June 2017. The site is characterized by a complex topography, with a difference in altitude of about 200 m from the bottom of the valley to the top of the two almost-parallel ridges, which are separated by 1.5 km. The site is largely covered in trees, which increase the surface roughness. The areas outside the ridges and valley are largely farmland and eucalyptus groves. For this analysis, we use data from three 100-m meteorological towers (tse04, tse09,

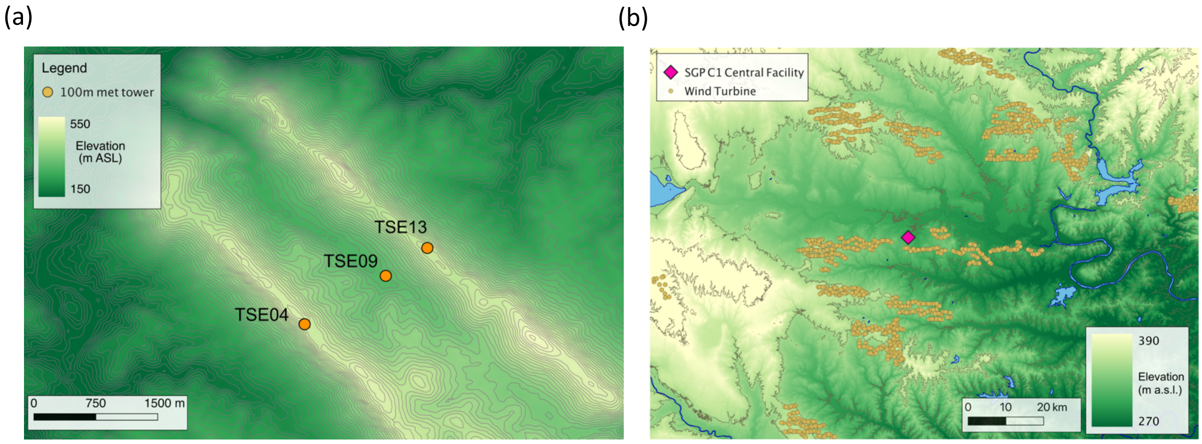

Figure 1. Map of (a) the Perdigão field campaign and (b) SGP locations, with the locations of the meteorological towers considered in this study.

and tse13—see map in figure 1a). At each tower, we consider observations from the 20-Hz sonic anemometers at 10 m and 60 m above ground level (AGL). To exclude tower wake effects, we discard from analysis those time stamps when the wind direction was ±30° from the direction of the tower boom. We also discard periods of precipitation to avoid data contamination [35], and periods when the 60-m wind speed was lower than 3 m s^{-1}, which are not relevant for wind energy purposes.

Next, we use observations from the main meteorological tower at the SGP observatory [34] in northern Oklahoma in the United States. The rural region is characterized by relatively simple topography, and its land use is primarily cattle pasture and wheat fields (figure 1b). Observations are from sonic anemometers at 10 m and 60 m AGL, publicly available as 30-minute averages, and span from 22 July 2015 to 15 March 2022.

2.2. Regression algorithms

In our analysis, we consider two regression algorithms, and we will compare their performance in predicting 30-minute average TI at 60 m AGL. For both regression models, we use as inputs the following variables: wind speed at 10 m AGL, TKE at 10 m AGL, TKE normalized by wind speed (TKE/WS) at 10 m AGL, Obukhov length (L) at 10 m AGL, proxy for atmospheric stability, wind speed at 60 m AGL, TKE at 60 m AGL, TKE normalized by wind speed (TKE/WS) at 60 m AGL. Although obviously correlated with TKE and wind speed, we decide to include a normalized version of TKE as an additional input because it mimics the functional structure of TI, where horizontal wind speed is at the denominator. As already mentioned, we discard from the analysis all time stamps where 60-m wind speed is lower than a typical cut-in wind speed for modern wind turbines, here set to 3 m s^{-1}, a wind regime not relevant for wind energy purposes.

As a baseline, we consider a multivariate linear regression. To avoid training a model that overfits the data, regularization techniques need to be implemented so that the learning model is constrained: the fewer degrees of freedom the model has, the harder it will be for it to overfit the data. We use Ridge regression [36] (`Ridge` in Python's library scikit-learn) to constrain the multivariate regression. The Ridge regression is achieved by adding a regularization term to the cost function (mean square error), with a hyperparameter α that controls how much the model will be regularized. The optimal value of α is determined by cross validation, with values sampled in the range from 0.1 to 10.

Next, as a proof of concept of the skills of more sophisticated machine-learning algorithms, we use a random forest regression model (`RandomForestRegressor` module in Python's scikit-learn [37]). We note here that ensemble models such as the random forest have shown strong predictive power in previous work [9, 27], and we defer a more exhaustive analysis of the predictive skills of other learning algorithms to a later study. The set of algorithm hyperparameters considered in the training and validation process and their sampled ranges used in the cross validation are listed in table 1.

Table 1. Algorithm hyperparameters considered for the random forest and their considered values in the cross validation.

Hyperparameter	Sampled Values
Number of estimators	10–800
Maximum depth	4–40
Maximum number of features	1–7
Minimum number of samples to split	2–11
Minimum number of samples for a leaf	1–15

For each regression algorithm, we adopt a nested cross-validation approach. We divide the data set into three subsets: the training, validation, and testing sets [38]. First, the learning algorithm is trained multiple times with different hyperparameters on the training set. Second, the best set of hyperparameters is selected as the one that gives the best results when applying the model on the validation set. Finally, the generalization skills of the learning model are evaluated on the testing set, which was not seen by the model during the optimization of the regression algorithm. Given the time-series nature of our problem, the construction of the training, validation, and testing sets needs to be carefully considered. If the data in each of the three sets were picked randomly, the performance of the machine-learning model would likely be artificially improved. In fact, if created randomly, we could end up in a situation where the observations at a given time stamp would be in the training set, and the observations at the subsequent time stamp, which are likely highly correlated with those from the previous one, could instead be part of the validation or testing set. Therefore, we use contiguous data to build the three sets: we keep a contiguous 20% of the data for testing, whereas the remaining 80% are used in a five-fold cross validation to optimize the model hyperparameters. To obtain the best possible estimate of the generalization skill of the model, in this nested approach we repeat the training process by shifting the testing set five times, so that it can cover the full period of record. We therefore evaluate the overall performance based on the root-mean-square error (RMSE) between the actual and predicted hub-height TI, averaged over the five different testing periods.

3. Results

We first consider the predictive skills of the multivariate linear regression approach when it is trained and tested at the same site. Figure 2 shows scatterplots of observed and linear-regression-predicted (over the testing set) TI at 60 m AGL at the three Perdigão towers and the SGP main tower. We quantify the skills of the considered regression algorithm in terms of the RMSE normalized by the range of observed TI values, to be able to make a fair comparison among different sites. The results show NRMSE between 0.067 and 0.086 at the Perdigão towers and 0.020 at SGP.

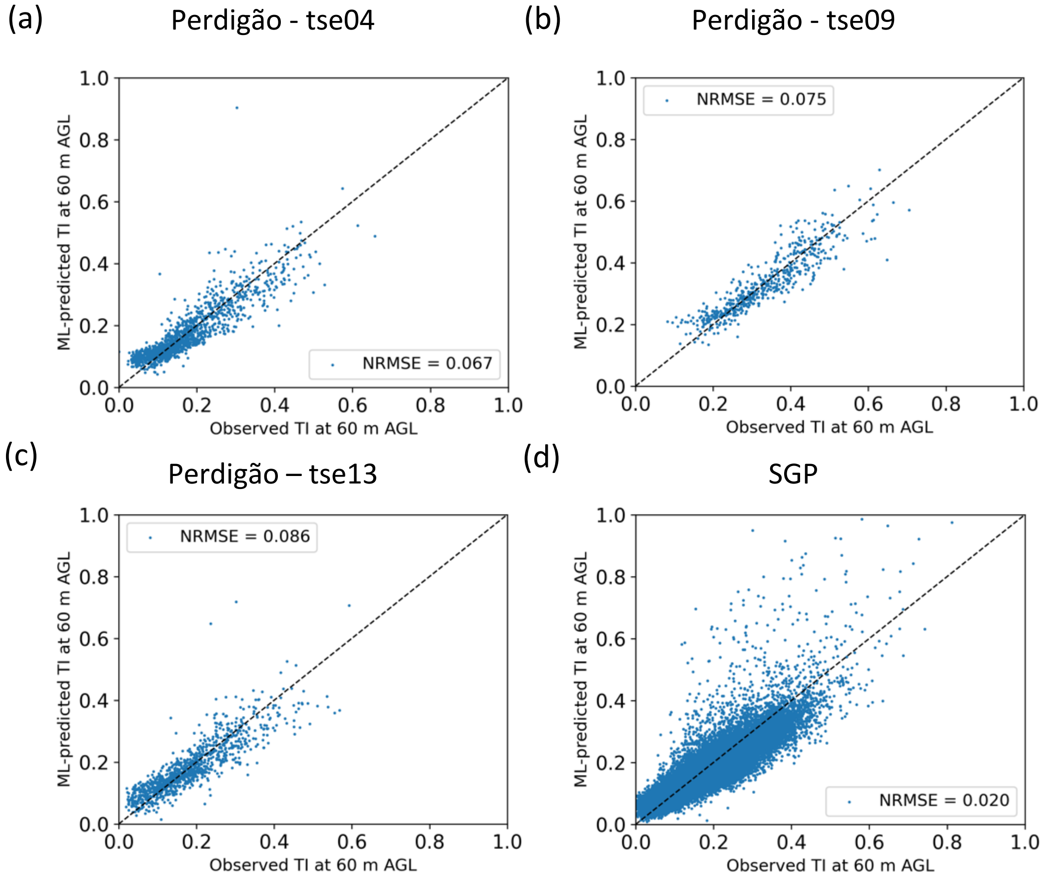

Figure 2. Scatterplots of observed and linear-regression-predicted 30-minute average TI at 60 m AGL at (a–c) the three Perdigão towers and (d) the SGP tower, when the regression model is trained and tested at the same location.

We then compare these baseline results to what can be achieved when adopting the more sophisticated random forest approach. We find good performance at all sites, with NRMSE ranging from 0.012 (in simple terrain and with a much bigger training set) to 0.071 (in the very complex terrain of the bottom of the Perdigão valley, Tower tse09). At Perdigão, these results represent a reduction of 18-25% compared to the multivariate linear regression results for the two towers on the ridges, and a reduction of just 6% for tse09, where the reduced number of available data and the impact of the ultra-complex terrain and canopy make it a challenging site to model, even for a more sophisticated random forest algorithm. At SGP, where the terrain is simpler and the training period includes multiple years of observations, we observe a 30% reduction in NRMSE when moving from a linear regression to the random forest approach.

Next, we assess the importance of the various input variables used to feed the random forest. First, we determine how the testing NRMSE varies as the random forest is fed with more and more input variables (figure 4). We start by training a random forest with 10-m wind speed as the only input, and then we progressively add more near-surface input variables before also adding hub-height quantities. In general, using near-surface wind speed as the only input does not allow for an accurate modeling of hub-height TI. When including near-surface turbulence and stability metrics as inputs, the NRMSE drops, but it still remains significantly larger than the values seen when including variables observed at hub-height as additional inputs. Finally, if

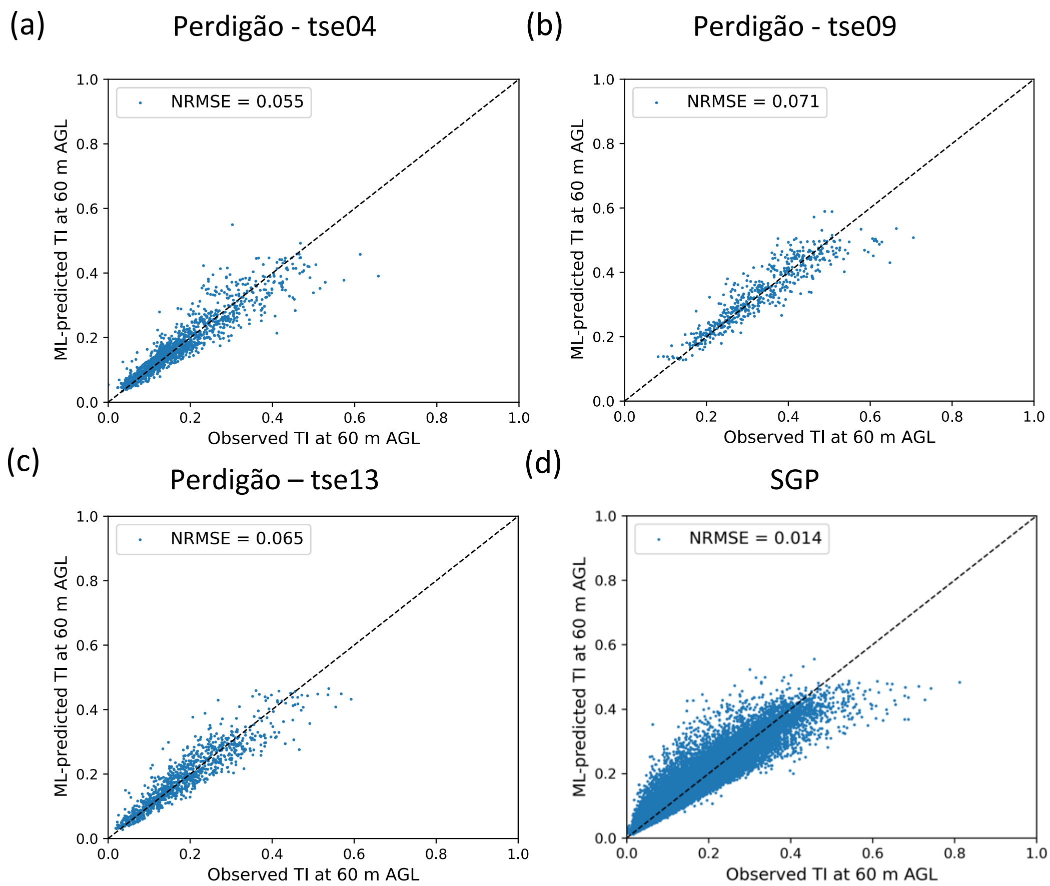

Figure 3. Scatterplots of observed and random-forest-predicted 30-minute average TI at 60 m AGL at the three Perdigão towers and the SGP one, when the regression model is trained and tested at the same location, for the testing set.

no data on hub-height turbulence are given to the model, and wind speed is the only hub-height input feature, we still find good performance, with testing NRMSE less than 0.1, but still higher than the lowest values found with the full set of inputs, especially in simple terrain at SGP.

As a way to summarize this analysis, we report in figure 5 the relative feature importance of each input variable at the Perdigão tse09 and SGP towers (results from the two other towers at Perdigão were similar). This metric is calculated by assessing to what degree the random forest tree nodes that use each feature reduce the mean square error on average (across all trees in the forest), weighted by the number of times each feature is selected. At both locations, we see how hub-height variables are the most influential in predicting hub-height TI, which is consistent with the results of the analysis in figure 4. Still, near-surface variables help further improve the accuracy of the random forest, especially at the Perdigão site. Although useful, we note how this metric is affected by the existing correlation between different inputs: for example, if the 60-m normalized TKE was not included as input, we would likely see a significantly higher importance for the other two hub-height variables (wind speed and TKE).

Finally, we leverage the data sets from both locations to understand the generalization skills of the proposed random forest approach. Specifically, we apply a round-robin validation, where we train the learning model at one site (either SGP or Perdigão), and then apply it to model 60-m TI at the other location, and vice versa. Figure 6 shows the results of this validation exercise in terms of scatterplots of observed and machine-learning-predicted 60-m TI. While the

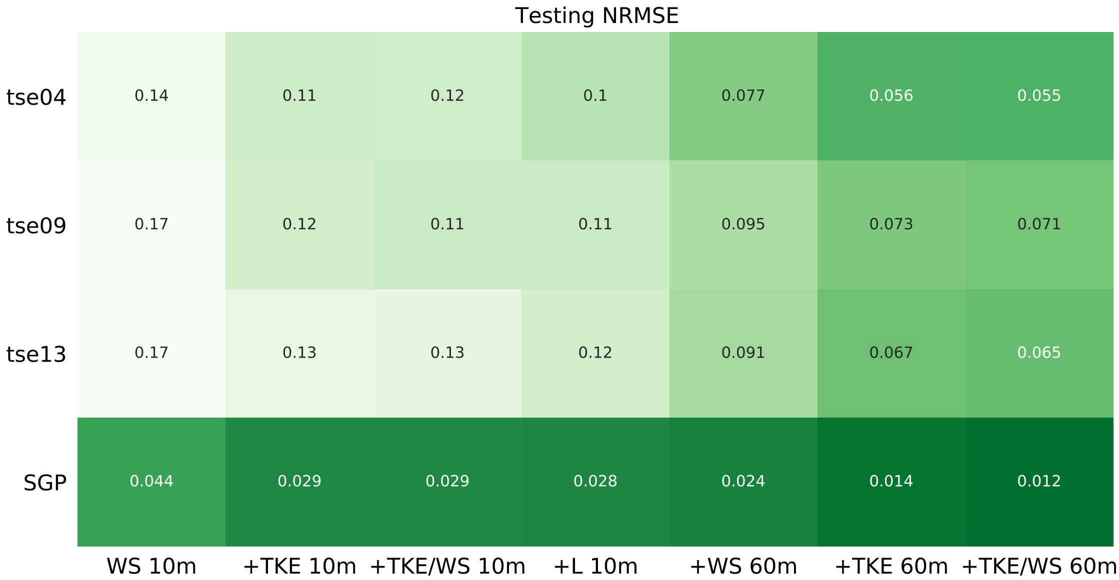

Figure 4. Testing NRMSE when a random forest is used to model 30-minute average 60-m TI at the Perdigão and SGP towers, when multiple variables are progressively added as inputs to the learning algorithm.

NRMSE values are larger than the optimal ones found for a same-site exercise (see figure 3), we still see a remarkable accuracy for both cases, with NRMSE values of 0.03 and 0.06 when the model is applied at SGP and Perdigão, respectively. We find a larger increase, compared to the same-site scenario, when training the model at Perdigão and applying it at SGP, possibly because of the very limited length of the training data at Perdigão. On the other hand, when training the model at SGP and applying it at Perdigão, the NRMSE is not significantly different than what observed in figure 3, under the same-site scenario.

4. Conclusions

The ability to model hub-height turbulence intensity would be extremely useful in all applications where hub-height turbulence is crucial, ranging from wind turbine control to power forecasting and—in a broader perspective—pollutant dispersion and drone flight forecasting. Such a model, if incorporated in the state-of-the-art numerical weather prediction model at the mesoscale (WRF), would represent a huge advancement for the wind energy modeling community. Despite this intense motivation to find accurate approaches for estimating turbulence conditions at hub height, few models have been proposed to date for such application, each with their inherent uncertainty and limitations. The challenge is even larger for sites in complex terrain.

We analyzed atmospheric data at 10 m and 60 m above ground level from four meteorological towers at two sites: three towers from the Perdigão field campaign, and a fourth tower at the SGP site. We considered how a machine-learning-based algorithm could be derived and applied to successfully model hub-height TI. We focused our validation efforts first using a same-site approach, where the random forest is trained and then tested at the same site, which is useful to derive several physical insights into the proposed model. We find that the proposed random forest is successful in modeling hub-height TI, with testing NRMSE as low as 0.014 when applied on 30-minute average data. In order to obtain such low values of testing NRMSE, the knowledge of hub-height TKE (which can instead be modeled in WRF) is needed. Next, we considered a more practically useful (and scientifically fair) round-robin validation, where

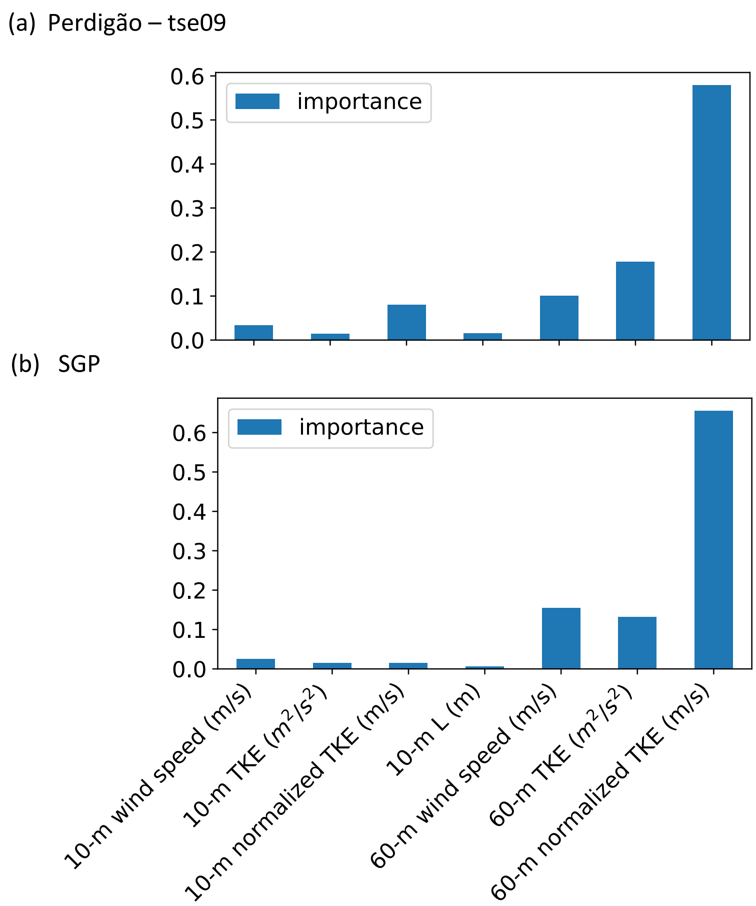

Figure 5. Relative importance of the various input variables used to feed the random forest in predicting 10-minute average 60-m TI at (a) the Perdigão tse09 tower and (b) the SGP tower.

the random forest is trained at one site and then applied to predict hub-height TI at a different site. Notably, we find a good generalization performance of our proposed model even under this scenario, especially when the training set includes many years of data (in our case, at SGP). Our validation results confirm the good generalization skills of the proposed algorithm, despite the significant differences (in terrain, canopy, and duration of the period of record) between the two considered field campaigns.

While promising, several opportunities exist to further expand our work in the future. A more complete survey of additional learning algorithms and/or input variables could be completed to optimize the performance of the machine-learning approach. Also, one could expand the training and testing validation by considering different sites, such as those offshore. Finally, this more complete and broadly validated model could be incorporated in WRF to add to it a capability to output TI.

Acknowledgments

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S.

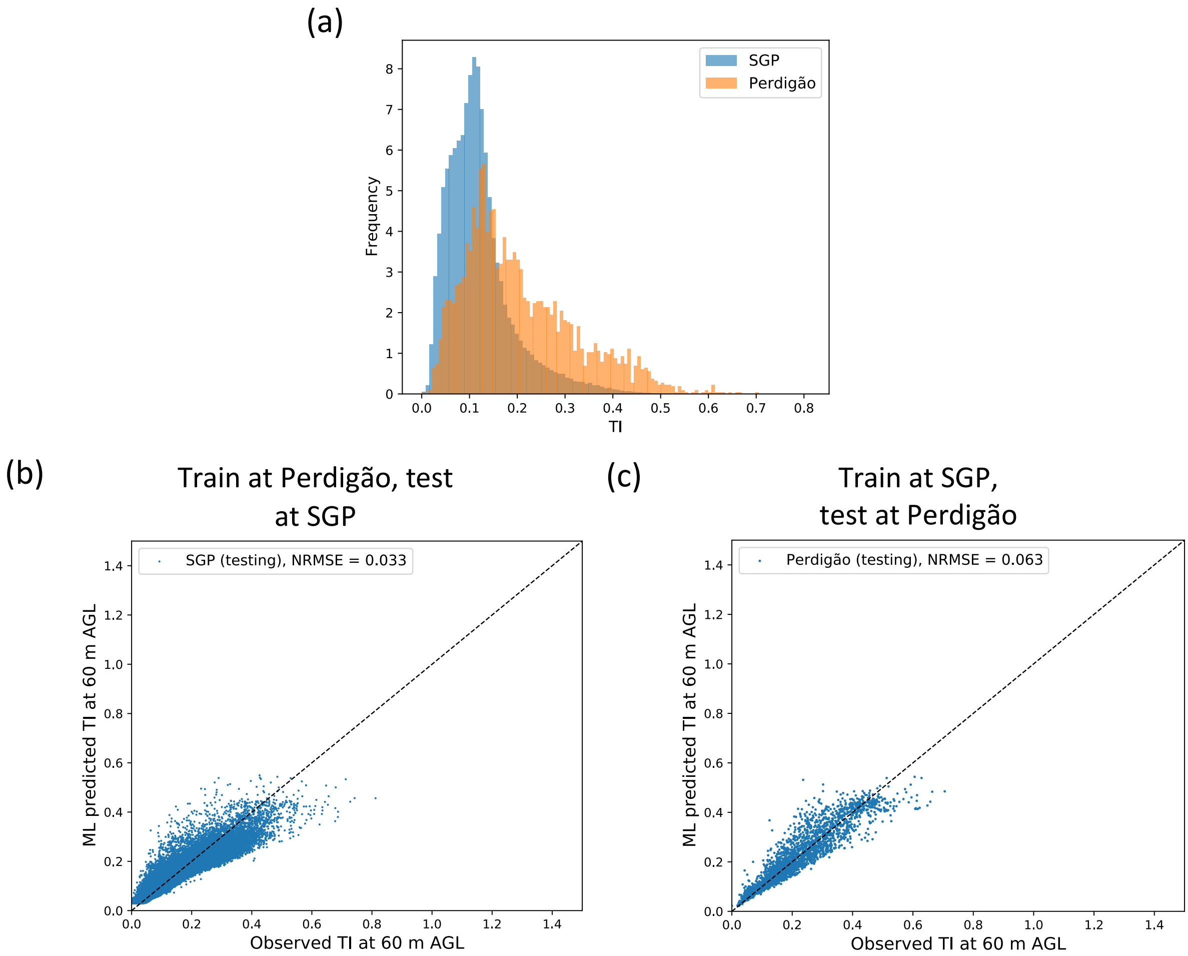

Figure 6. (a) Histograms of 30-minute average TI at 60 m AGL for the three Perdigão towers combined and the SGP one. (b) Scatterplot of observed and random-forest-predicted 30-minute average 60-m TI when the random forest is trained at the three Perdigão towers combined and tested at SGP. (c) Same as (b) but with training and tested sets switched.

Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.
We thank the residents of Alvaiade and Vale do Cobrão for their essential hospitality and support throughout the Perdigão field campaign. In particular, we are grateful for the human and logistic support Felicity Townsend provided to our research group in the field. We express appreciation to the Earth Observing Laboratory at the National Center for Atmospheric Research, the Danish Technical University, and INEGI for their herculean efforts to construct and maintain the data streams from the networks of towers at Perdigão.

References

[1] Elliott D L and Cadogan J B 1990 Effects of wind shear and turbulence on wind turbine power curves Tech. rep. Pacific Northwest Lab., Richland, WA (USA)

[2] Wharton S and Lundquist J K 2012 *Wind Energy* **15** 525–546 ISSN 1095-4244 URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/we.483`
[3] Wharton S and Lundquist J K 2012 *Environmental Research Letters* **7** 014005 ISSN 1748-9326
[4] Vanderwende B J and Lundquist J K 2012 *Environmental Research Letters* **7** 034035 ISSN 1748-9326
[5] Murphy P, Lundquist J K and Fleming P 2020 *Wind Energy Science* **5** 1169–1190 ISSN 2366-7443 publisher: Copernicus GmbH URL `https://wes.copernicus.org/articles/5/1169/2020/`
[6] Vahidzadeh M and Markfort C D 2019 *Energies* **12** ISSN 1996-1073 URL `https://www.mdpi.com/1996-1073/12/9/1805`
[7] Fleming P, King J, Dykes K, Simley E, Roadman J, Scholbrock A, Murphy P, Lundquist J K, Moriarty P, Fleming K, van Dam J, Bay C, Mudafort R, Lopez H, Skopek J, Scott M, Ryan B, Guernsey C and Brake D 2019 *Wind Energy Science* **4** 273–285 ISSN 2366-7443 publisher: Copernicus GmbH URL `https://wes.copernicus.org/articles/4/273/2019/`
[8] Fleming P, King J, Simley E, Roadman J, Scholbrock A, Murphy P, Lundquist J K, Moriarty P, Fleming K, van Dam J, Bay C, Mudafort R, Jager D, Skopek J, Scott M, Ryan B, Guernsey C and Brake D 2020 *Wind Energy Science* **5** 945–958 ISSN 2366-7443 publisher: Copernicus GmbH URL `https://wes.copernicus.org/articles/5/945/2020/`
[9] Optis M and Perr-Sauer J 2019 *Renewable and Sustainable Energy Reviews* **112** 27–41
[10] Bodini N, Optis M, Perr-Sauer J, Simley E and Fields M J 2021 *Wind Energy*
[11] Mann J 1994 *Journal of Fluid Mechanics* **273** 141–168 ISSN 1469-7645, 0022-1120 publisher: Cambridge University Press
[12] Mann J 1998 *Probabilistic Engineering Mechanics* **13** 269–282 ISSN 0266-8920 URL `https://www.sciencedirect.com/science/article/pii/S0266892097000362`
[13] Cheung L C, Premasuthan S, Davoust S and von Terzi D 2016 *Journal of Physics: Conference Series* **753** 032008 ISSN 1742-6588, 1742-6596 URL `https://iopscience.iop.org/article/10.1088/1742-6596/753/3/032008`
[14] Gualtieri G 2017 *Renewable Energy* **111** 175–186 ISSN 0960-1481 URL `https://www.sciencedirect.com/science/article/pii/S0960148117302872`
[15] Türk M and Emeis S 2010 *Journal of Wind Engineering and Industrial Aerodynamics* **98** 466–471
[16] 61400-1 I 2019 Wind turbines—part 3: design requirements
[17] 61400-3 I 2019 Wind turbines—part 1: design requirements for offshore wind turbines
[18] Powers J G, Klemp J B, Skamarock W C, Davis C A, Dudhia J, Gill D O, Coen J L, Gochis D J, Ahmadov R, Peckham S E *et al.* 2017 *Bulletin of the American Meteorological Society* **98** 1717–1737
[19] Skamarock W C, Klemp J B, Dudhia J, Gill D O, Liu Z, Berner J, Wang W, Powers J G, Duda M G, Barker D M *et al.* 2019 *National Center for Atmospheric Research: Boulder, CO, USA* 145
[20] Nakanishi M and Niino H 2009 *Journal of the Meteorological Society of Japan. Ser. II* **87** 895–912
[21] Olson J B, Kenyon J S, Angevine W, Brown J M, Pagowski M, Sušelj K *et al.* 2019
[22] Hu X M, Klein P M and Xue M 2013 *Journal of Geophysical Research: Atmospheres* **118** 10–490
[23] Giebel G, Brownsword R, Kariniotakis G, Denhard M and Draxl C 2011 *ANEMOS. plus*
[24] Clifton A, Kilcher L, Lundquist J and Fleming P 2013 *Environmental research letters* **8** 024009
[25] Leahy K, Hu R L, Konstantakopoulos I C, Spanos C J and Agogino A M 2016 *2016 IEEE International Conference on Prognostics and Health Management (ICPHM)* (IEEE) pp 1–8
[26] Arcos Jiménez A, Gómez Muñoz C and García Márquez F 2018 *Energies* **11** 13
[27] Bodini N and Optis M 2020 *Wind Energy Science* **5** 489–501
[28] Bodini N and Optis M 2020 *Journal of Physics: Conference Series* vol 1618 (IOP Publishing) p 062037
[29] Optis M, Bodini N, Debnath M and Doubrawa P 2021 *Wind Energy Science Discussions* 1–26
[30] Bodini N, Lundquist J K and Optis M 2020 *Geoscientific Model Development* **13** 4271–4285
[31] Leufen L H and Schädler G 2019 *Geoscientific model development* **12** 2033–2047
[32] McCandless T, Gagne D, Kosovic B, Haupt S E, Yang B, Becker C and Schreck J in review *Boundary Layer Meteorology*
[33] Fernando H J S, Mann J, Palma J M L M, Lundquist J K, Barthelmie R J, Belo-Pereira M, Brown W O J, Chow F K, Gerz T, Hocut C M, Klein P M, Leo L S, Matos J C, Oncley S P, Pryor S C, Bariteau L, Bell T M, Bodini N, Carney M B, Courtney M S, Creegan E D, Dimitrova R, Gomes S, Hagen M, Hyde J O, Kigle S, Krishnamurthy R, Lopes J C, Mazzaro L, Neher J M T, Menke R, Murphy P, Oswald L, Otarola-Bustos S, Pattantyus A K, Rodrigues C V, Schady A, Sirin N, Spuler S, Svensson E, Tomaszewski J, Turner D D, Veen L v, Vasiljević N, Vassallo D, Voss S, Wildmann N and Wang Y 2019 *Bulletin of the American Meteorological Society* **100** 799–819 ISSN 0003-0007, 1520-0477 publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society URL `https://journals.ametsoc.org/view/journals/bams/100/5/bams-d-17-0227.1.xml`
[34] Sisterson D, Peppler R, Cress T, Lamb P and Turner D 2016 *Meteorological Monographs* **57** 6–1

[35] Zhang R, Huang J, Wang X, Zhang J A and Huang F 2016 *Journal of Ocean University of China* **15** 389–398
[36] Hoerl A E and Kennard R W 1970 *Technometrics* **12** 55–67
[37] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V *et al.* 2011
[38] Friedman J, Hastie T and Tibshirani R 2001 *The elements of statistical learning* vol 1 (Springer series in statistics New York)