# Wind turbine load estimation using machine learning and transfer learning

Xu, G.; Yu, W.; Kim, Taeseong

**DOI**
10.1088/1742-6596/2265/3/032108

**Publication date**
2022

**Document Version**
Final published version

**Published in**
Journal of Physics: Conference Series

# Wind turbine load estimation using machine learning and transfer learning

To cite this article: Guanqun Xu *et al* 2022 *J. Phys.: Conf. Ser.* **2265** 032108



# Wind turbine load estimation using machine learning and transfer learning

**Guanqun Xu[1,2], Wei Yu[2], Taeseong Kim[1]**

[1] Department of Wind Energy, Technical University of Denmark, Lyngby, Denmark

[2] Wind energy section, Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands

E-mail: `tkim@dtu.dk`

**Abstract.** Machine learning method has always been popular to solve wind turbine related problems at a data level. However, with the limitation of the availability of relevant data, transfer learning has gained increasing attention. In this study, traditional machine learning method of artificial neural networks (ANN), together with parameter-based transfer learning method has been used to estimate wind turbine load. First, ANN load model was built for DTU 10MW wind turbine as well as NREL 5MW wind turbine. Then, parameter-based transfer learning has been applied to the above-mentioned models to estimate load for a different turbine type or two mixed turbine types. Results indicate that ANN method provides good estimation on wind turbine fatigue load. For DTU 10MW ANN model, the trend of accuracy becomes steady as the number of input samples increases and 1500 samples is deemed as the optimal number of samples for training DTU 10MW. In addition, with transfer learning, it was succeeded in building NREL 5MW model with corresponding DTU 10MW pretrained model but failed in establishing mixed dataset model neither with DTU 10MW nor with NREL 5MW pretrained model.

## 1. Introduction

In wind turbine design process, the most important and most frequently appearing working condition should be simulated, which corresponds to a large amount of design load cases (DLCs) specified by IEC standard [1]. Normally, DLCs are simulated in aeroelastic code such as HAWC2 [2] with the establishment of physical wind turbine model. However, due to containing high level of design details [3], simulating integrated DLCs is often time-consuming.

Nowadays, with the rapid development of artificial intelligence (AI), an alternative way for wind turbine load simulation is from a numerical level by using AI method such as Artificial Neural Networks (ANN). ANN is a popular AI method that mimics neurons of human brains in terms of data processing and information transfer [4]. Furthermore, a trained AI model can be extended to predict loads for different field conditions, different types of turbines, etc. This can be realized through a transfer learning method. In transfer learning, a different but related problem can be tackled with an already solved problem [5]. Namely, through transfer learning, the question in the target domain can be addressed by training data in a related source domain [6].

The main aim of the study is to discover the possibility of simplifying wind turbine design procedure using transfer learning method. The general procedure of this study is that the trained ANN model is expanded via transfer learning to predict loads for different types of wind turbines under different wind conditions, with limited number of data. In this way, the data-level modelling approach can be validated and the accuracy of transfer learning method is checked.

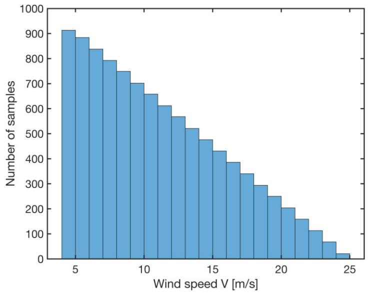
## 2. Methodology
In general, two types of load estimation model were established. One is the traditional model using ANN. Before building the model, several related parameters, e.g., learning rate, number of hidden layers, etc., need to be determined through hyperparameter optimization. Grid search approach was applied for the optimization where models with all possible combinations of hyperparameters are compared. k-fold cross validation method was used to evaluate model performance. In this study, both 5-fold and 10-fold cross validation were used depending on the input size. For 5-fold cross validation, each sub-net has five sub-dataset and four of them are for training and the rest one sub-dataset is for validation. 10-fold cross validation follows the same procedure except that it has ten sub-dataset.

Based on the ANN model, the other type of the load estimation model is built using transfer learning method. Through transfer learning, it is possible to train a model with good performance but with less time and data by “copying” the information from a pre-trained model. It is much faster and easier than training a new AI model from scratch. In [7], transfer learning approaches were divided into four categories: instance-transfer, feature-representation-transfer, parameter-transfer and relational-knowledge-transfer. In this study, “freeze weight” method was applied. It belongs to parameter-transfer and only the weights and bias of the output layer were trained in every epoch and other parameters were kept the same as the pretrained model.

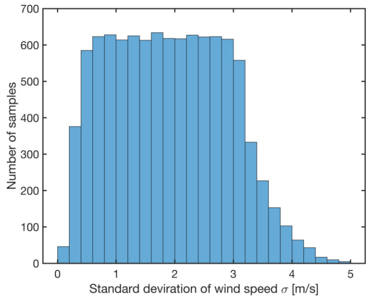
## 3. Data acquisition and model setup
### *3.1. Data acquisition*
In order to apply transfer learning, two types of wind turbine models were selected, one is DTU 10MW reference wind turbine [8] (called as DTU 10MW), the other one is NREL 5MW reference wind turbine [9] (called as NREL 5MW). The data for DTU 10MW was obtained from the database established in [10], both for input (wind speed, standard deviration of wind speed and wind shear) and output (fatigue loads (DEL) of blade root flapwise bending moment (BRFW) and tower-top fore-aft bending moment (TTFA)) of the model. Histograms of inputs and targets of all the 9980 samples are shown in Figure 1. NREL 5MW data were generated through HAWC2 simulations. DTU wind energy controller [11] was used in the simulation. It is noted that this controller is the same as the controller of DTU 10MW. In order to compare the performance of the trained AI models in different wind conditions, two wind classes (IA and IC) were considered. The 10-min short-term DEL was shown as histogram in Figure 2.

### *3.2. Model setup*
In all, three different transfer learning models were investigated. First, a classic transfer learning method was investigated where the pre-trained DTU 10MW model was trained with ANN, which is either one-output (BRFW) or twp-output (BRFW and TTFA). Then, limited number of NREL 5MW data was applied via transfer learning, extending the AI model to predict loads for NREL 5MW. This model is called as Case1. Second, in order to investigate the possibility that a single AI model can predict loads for both turbines simultaneously via transfer learning, two different AI models with mixed data of both DTU 10MW and NREL 5MW were studied. One is applying the mixed dataset to the pre-trained DTU 10MW model with full(9980) samples (called as Case2), the other one is applying the mixed dataset to the pre-trained NREL 5MW model with full(2577) samples (called as Case3). All the models are concluded in Table 1.

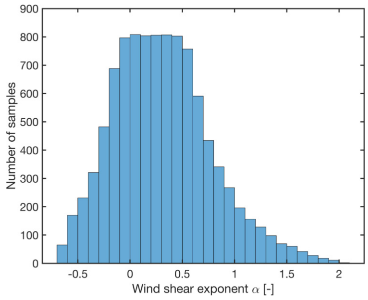
Journal of Physics: Conference Series **2265** (2022) 032108 doi:10.1088/1742-6596/2265/3/032108



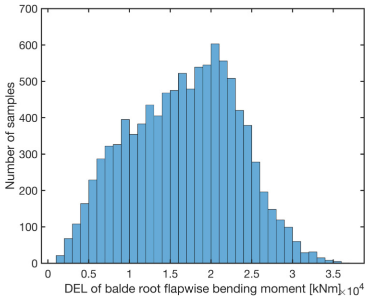

(a) Wind speed



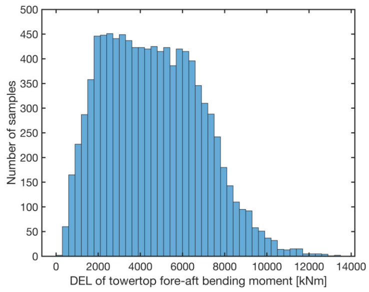

(b) Standard deviation




(c) Wind shear exponent

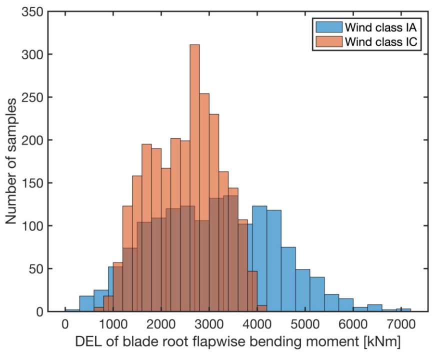



(d) DEL of BRFW

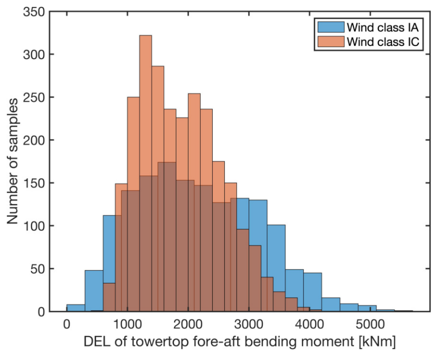



(e) DEL of TTFA

Figure 1: Input and target of DTU 10 MW load model




(a) DEL of BRFW




(b) DEL of TTFA

Figure 2: Target of NREL 5MW load model

Table 1: Model setup

| Case NO. | Pretrained model | | | | Transfer learning model | | |
|---|---|---|---|---|---|---|---|
| | *input type* | *input size* | *wind class* | *number of output(s)* | *input type* | *wind class* | *number of output(s)* |
| 1 | DTU 10MW | 500/1000/ 1500/9980 | IA | one/two | NREL 5MW | IA/IC | one/two |
| 2 | DTU 10MW | 9980 | IA | one | both turbines | both in IA or DTU 10MW in IA and NREL 5MW in IC | one |
| 3 | NREL 5MW | 2577 | IC | one | both turbines | both in IA or DTU 10MW in IA and NREL 5MW in IC | one |

## 4. Results of pretrained models

In the process of hyperparameter optimization, the evaluated hyperparameters were the number of hidden layer and the number of neurons per layer since these two parameters are influential in model accuracy. The number of hidden layer was varied from 1 to 5 and number of neurons

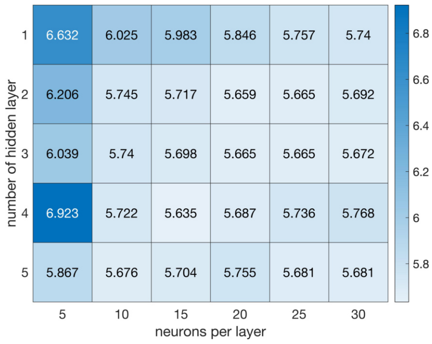



Figure 3: MAPE of NREL 5MW model

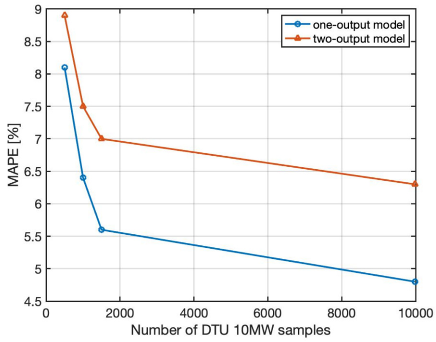



Figure 4: MAPE of DTU 10MW models

per layer was varied from 10 to 50 with an interval of 5. ReLU activation function was used during the training. 10% of the data used in each model were extracted for general testing. The rest were used in cross validation for training and validation. Mean absolute percentage error (MAPE) was chosen to represent the model accuracy.

NREL 5 MW (one-output) load model was built with 2577 (all) samples in wind class IC, shown in Figure 3. Each block corresponds to a selected combination of the number of neurons per layer and the number of hidden layer. And the number on each block is the corresponding MAPE in percentage. When the number of neurons per layer and the number of hidden layer increases, ANN model has more coefficients (bias and weights) to form the relation between input and output. This increases relation complexity and stability and results in the increase of accuracy (lower MAPE) in a global trend. The trend is not sharp since there are uncertainty in the data and the uncertainty does not correspond to any input. The lowest MAPE of one-output model for NREL 5MW is 5.6% where there are 4 hidden layers and 15 neurons per layer.

DTU 10MW models were established with 500, 1000, 1500 and 9980 samples. The optimal results of them are plotted in Figure 4. It is observed that MAPE decreases with increasing samples. This is reasonable because when there are more training data, the model is more systematic, thus higher accuracy. It is also shown that both one- and two-output models predict load(s) well where approximately 6% error rate is predicted.

## 5. Results of transfer learning models

### *5.1. Transfer learning Case1: trained with pure NREL 5MW dataset*

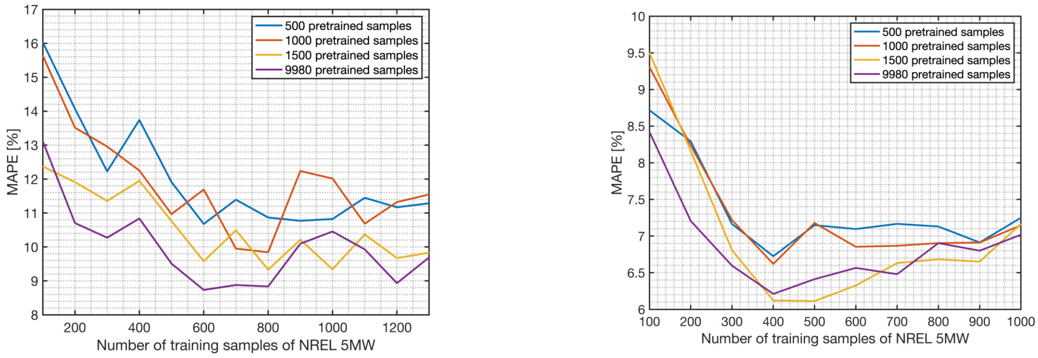
#### *5.1.1. One-output model*

MAPE of NREL 5MW one-output model is shown in Figure 5 with Figure 5a using data under the same wind class as the pretrained model and Figure 5b using the different. MAPE is plotted against the number of training samples of NREL 5 MW dataset. Different colors represent pretrained models with different numbers of samples.

Compared with the model using different wind classes, model using the same wind class in general has larger error. Pretrained model formalizes a relation between the input and the output of DTU 10MW. So when it is required to predict a different output distribution for NREL 5MW with exactly the same input, AI model can be confused.

Both figures in Figure 5 show that when the number of training samples increases, MAPE decreases in a global trend despite fluctuations. This is reasonable since when more NREL 5MW samples are involved, the model will be more systematic but the influence of data randomness cannot be eliminated. The fluctuation is more significant for same input cases, further verifying that "same input confusion" increases the difficulty of prediction.

Combining the result of models using data from same wind class and different wind classes, pretrained model with 500 samples has the worst performance while that with 9980 pretrained samples has the best performance. This matches the outcome that more samples leads to a more accurate prediction for pretrained model (Figure 4) and thus lower error for transfer learning model.

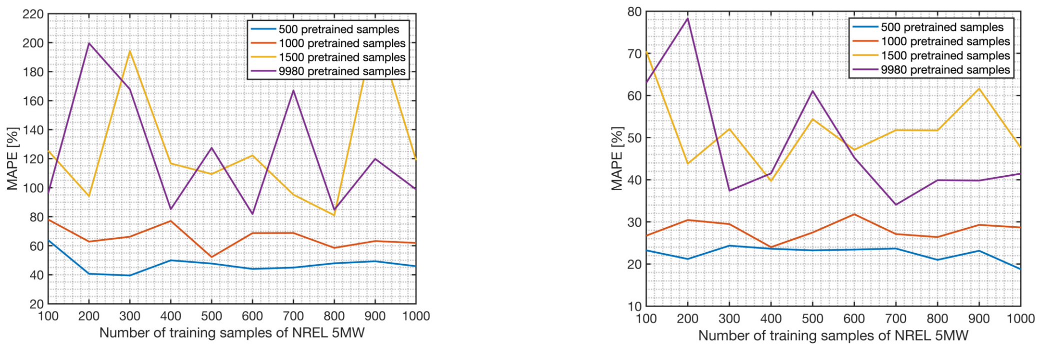



(a) same wind class (b) different wind class

Figure 5: MAPE of one-output 5 MW transfer learning model trained with pure NREL 5 MW dataset

*5.1.2. Two-output model*

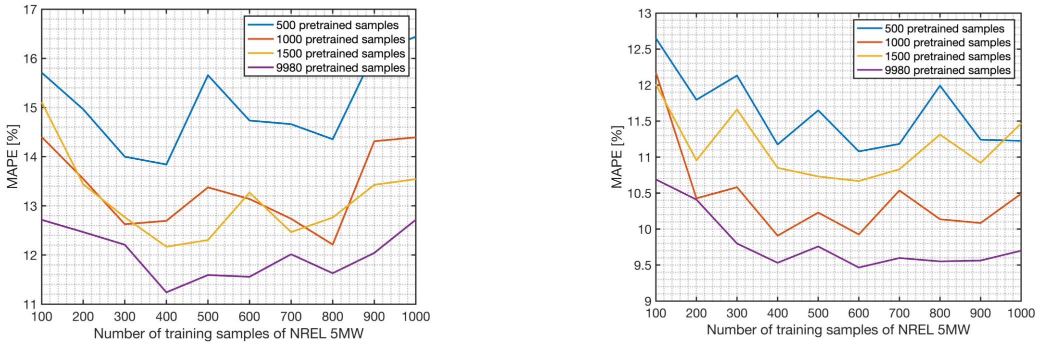
Two-output NREL 5MW model was built based on one-output and two-output DTU 10MW pretrained model. The result (MAPE) are shown in Figure 6 and 7 respectively. Overall, two-output NREL 5MW model based on one-output pretrained model (Figure 6) has unsatisfying performance. MAPE has large variations and they are even larger than 100% for some cases. A different result from NREL 5MW one-output model (Figure 5) is that when number of pretrained samples increases, MAPE increases in general. When there are more samples in the one-output pretrained model, the pretrained model has higher accuracy for one-output cases, but the weights and bias further deviate from two-output cases. Thus, the accuracy decreases for this two-output transfer learning model. It is also noted that MAPE are comparable when pretrained models are using 1500 and 9980 samples. A reason for this can be explained by Figure 4 that the corresponding one-output pretrained models have similar MAPE.




(a) same wind class (b) different wind class

Figure 6: MAPE of NREL 5MW two-output transfer learning model (one-output DTU 10MW pretrained model)

The second two-output NREL 5MW model was established with DTU 10MW two-output pretrained model and the result are shown in Figure 7. When training samples of NREL 5MW increases, the global trend of MAPE is decreasing regardless of fluctuations. In addition, compared with one-output NREL 5MW model (Figure 5), two-output model has larger MAPE, which is predictable since two output increases the complexity for prediction and is more likely to be biased. However, the error for two-output NREL 5MW model pretrained with two-output DTU 10 MW model is within an acceptable range. MAPE for all the cases are below 17% and are much better than that pretrained with one-output DTU 10MW model.




(a) same wind class (b) different wind class

Figure 7: MAPE of NREL 5MW two-output transfer learning model (two-output pretrained model)

### *5.2. Transfer learning Case2: trained with mixed dataset, using DTU 10MW pretrained model*

#### *5.2.1. Same wind class*

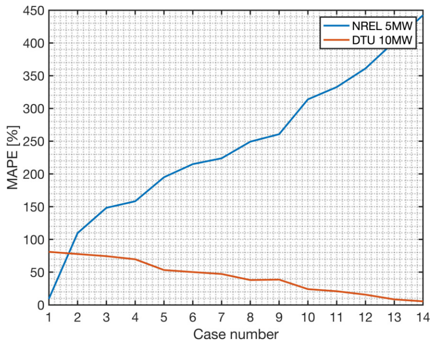
Model trained with mixed data with same input (both in class IA) was first established. Two mixing sample strategies were used: one is “fixed training samples” where the total training samples are fixed but the ratio of the two wind turbine samples is changing. The other is “increasing training samples” where the overall training samples are increasing with the two turbine samples having the same increment each time. The cases are shown in Table 2 and Table 3 and the results are shown in Figure 8.

Table 2: Fixed training sample for transfer learning (same wind class)

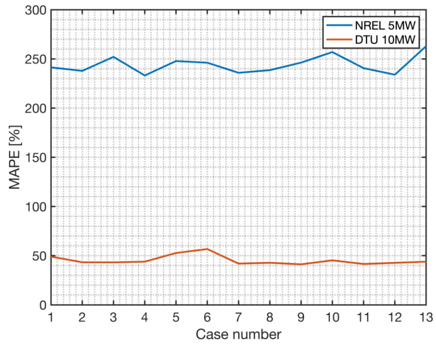
| Dataset number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of training samples of DTU 10MW | 0 | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 | 1100 | 1200 | 1300 |
| Number of training samples of NREL 5MW | 1300 | 1200 | 1100 | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 100 | 0 |

Table 3: Increasing training sample for transfer learning (same wind class)

| Dataset number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of training samples of DTU 10MW | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 | 1100 | 1200 | 1300 |
| Number of training samples of NREL 5MW | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 | 1100 | 1200 | 1300 |



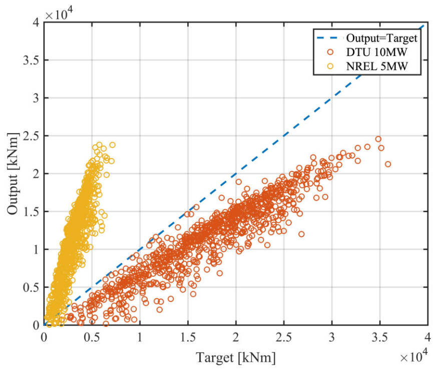

(a) fixed training sample



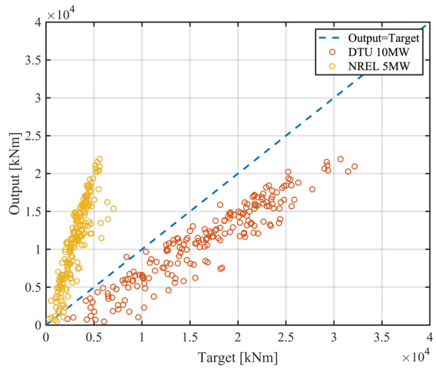

(b) increasing training sample

Figure 8: MAPE of mixed dataset transfer learning model (under same wind class)

As shown in Figure 8, mixing data from two turbine types does not lead to satisfying performance. In Figure 8a for fixed training sample cases, the prediction for NREL 5MW has acceptable MAPE (9.3%) only for dataset 1 (pure 5MW data). For other cases, MAPE is larger than 100%. Note that dataset 1 in Table 2 with 1300 NREL 5MW data is the same case as the corresponding one shown in Figure 5a, where the MAPE is 9.7%. In Figure 8b for increasing training sample cases, MAPE is relatively the same no matter how the size of training sample changes: approximately 40% for DTU 10MW and 240% for NREL 5MW. In addition, scatter plot of training and testing samples of dataset 13 in Table 3 are plotted in Figure 9. Most of the scatters are far away from the blue dashed line (zero MAPE), with NREL 5MW over predicted and DTU 10MW under predicted. Therefore, mixed turbine type model has bad performance when input of DTU 10MW and NREL 5MW are in the same class.




(a) training dataset




(b) test dataset

Figure 9: Scatter plot of prediction against database for mixed dataset case under same wind class (using dataset 13 in Table 3)

In order to further check this transfer learning method, a new ANN model trained with mixed dataset of 1300 DTU 10MW samples and 1300 NREL 5MW samples was tested with 130 DTU 10MW samples and 130 NREL 5MW samples. The result is shown in Table 6. Even when training a new ANN model with mixing dataset, the error is still large (130.6%), meaning that there is no suitable ANN model to predict accurate result for DTU 10MW and NREL 5MW simultaneously when input are the under the same wind class.

### 5.2.2. Different wind classes

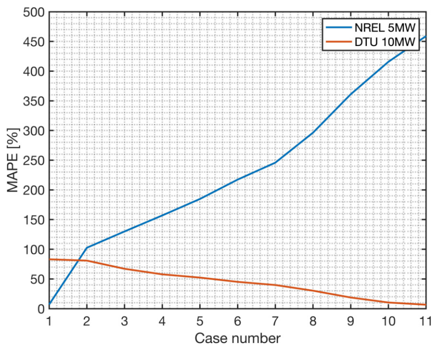
Then, the mixed turbine type model was trained with data from different wind classes: DTU 10MW data in class IA and NREL 5MW data in class IC. Similarly, as shown in Figure 10a, only dataset 1, where the model was trained with pure 5MW data, has relative good accuracy for 5MW. In Figure 11, scatter plot of dataset 10 in Table 5 also presents different prediction trends for the two wind turbine, which were under-predicted and over-predicted respectively. In all, this model failed to predict both turbine types.

Table 4: Fixed training sample for transfer learning (different wind class)

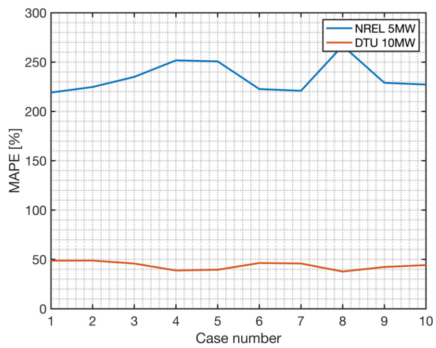
| Dataset number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of training samples of DTU 10MW | 0 | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
| Number of training samples of NREL 5MW | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 100 | 0 |

Table 5: Increasing training sample for transfer learning (different wind class)

| Dataset number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of training samples of DTU 10MW | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
| Number of training samples of NREL 5MW | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |




(a) fixed training sample



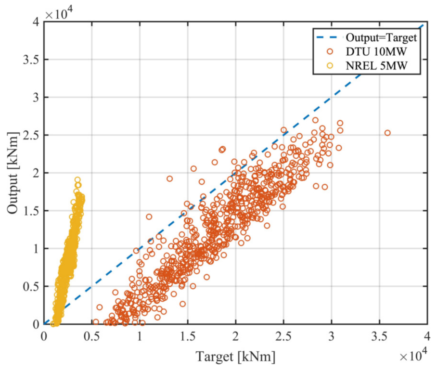

(b) increasing training sample

Figure 10: $MAPE$ of mixed dataset transfer learning model (under different classes)

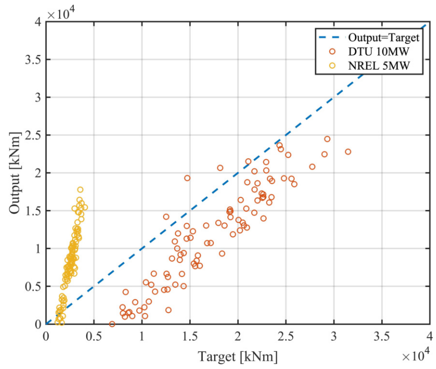
Moreover, using input from different wind classes, two new models were built with mixed dataset of 1000 samples for each turbine and 2500 samples for each turbine. The test result is shown in Table 6. MAPE of these two models are acceptable (18.7% and 15.8%), but when transfer learning was applied (Figure 10), error is quite large. Therefore, it is concluded that there exists an ANN model that can predict load for both DTU 10 MW and NREL 5 MW well. However, this "freeze weight" transfer learning method is not able to reach that goal.

### 5.3. Transfer learning Case3: trained with mixed dataset, using NREL 5MW pretrained model

In order to further check the performance of the "freeze weight" transfer learning method, source domain was switched to NREL 5MW. Namely, transfer learning method was applied on NREL




(a) training dataset




(b) test dataset

Figure 11: Scatter plot of prediction against database for mixed dataset case under different wind class (using dataset 10 in Table 5)

Table 6: MAPE of newly trained AI model with mixed dataset

| Input type for AI model | same | different | different |
|---|---|---|---|
| Number of training samples of NREL 5MW | 1300 | 1000 | 2500 |
| Number of training samples of DTU 10MW | 1300 | 1000 | 2500 |
| Total MAPE [%] | 130.6 | 18.7 | 15.8 |
| MAPE of NREL 5MW [%] | 221.3 | 26.0 | 20.5 |
| MAPE of DTU 10MW [%] | 40.0 | 11.4 | 11.1 |

5MW pretrained model to build mixed turbine type load model. Then, parametric study was carried out by mixing dataset of NREL 5MW and DTU 10MW data, with fixed training sample of 2200. All the parametric study cases are shown in Table 7 and the result for MAPE is shown in Figure 12. As shown in Figure 12, only when the model was trained with pure 5MW data (case 12), MAPE of NREL 5 MW is less than 100%, which further verifies that "freeze weight" method is not optimal when building a model with mixed dataset.

Table 7: Fixed training sample for transfer learning (different wind class, using NREL 5MW pretrained model)

| Dataset number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of training samples of DTU 5MW | 0 | 200 | 400 | 600 | 800 | 1000 | 1200 | 1400 | 1600 | 1800 | 2000 | 2200 |
| Number of training samples of NREL 10MW | 2200 | 2000 | 1800 | 1600 | 1400 | 1200 | 1000 | 800 | 600 | 400 | 200 | 0 |

### *5.4. Conclusion of transfer learning*

In conclusion, "freeze weight" transfer learning method provides desirable results partially. First, this transfer learning method has good performance when using corresponding DTU 10MW pretained model to build NREL 5MW model with samples in a different wind class. This is shown in Figure 5b and 7b for one-output and two-output model respectively. Second, this

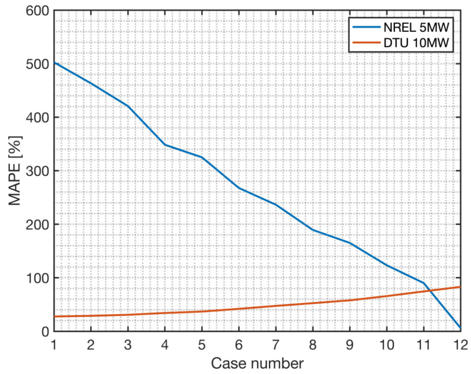



Figure 12: $MAPE$ of transfer learning model with mixed dataset of DTU 10MW and NREL 5MW under different class, using NREL 5MW pretrained model (different wind class)

transfer learning method also presents acceptable accuracy for NREL 5MW model with training samples in a same wind class of DTU 10MW (Figure 5a and 7a). Although maximum MAPE is approximately 16%, it is deemed as acceptable considering that it suffers the "same input confusion". In addition, this transfer learning method fails to predict two loads for NREL 5 MW if pretrained model can only predict one kind of load (Figure 6). Last but not least, it also fails to predict load for two turbine types simultaneously (Figure 8, 10 and 12).

## 6. Conclusion

In all, data-level methods of ANN and transfer learning have been established for wind turbine load estimation. The main conclusions of this study are:

- Traditional AI method of ANN can provide good estimation for wind turbine fatigue loads. 1500 samples is seen as the optimal number of samples for building DTU 10MW model.
- The aim of transfer learning, which is to tackle a different but related problem with solved problems, has been met by successfully building NREL 5MW model (with samples from a different wind class) with corresponding DTU 10MW pretrained model.
- Parameter-transfer based "freeze weight" method, failed to build mixed dataset model neither pretrained with DTU 10MW model nor NREL 5MW model.

Future work of this study could be focusing on other transfer learning methods such as instance-transfer. Also, apart from loads, other predictions can be made using transfer learning, e.g. power. In addition, another limitation of this research is that the model input is purely the characteristics of wind. Other important parameters, such as frequency of structural resonance, characteristic structure dimension and controller parameters can also be served as inputs. These turbine-specific parameters can provide physical insights from the data-level model.

## References

[1] International Electrotechnical Commission 2019 Wind Turbines Part 1: Design requirements, IEC 61400-1, 4th edition Standard IEC Geneva, CH URL `https://webstore.iec.ch`
[2] Larsen, Torben J and Hansen, Anders Melchior 2019 How 2 hawc2, the user's manual Tech. Rep. ISO/IEC TR 29110-1:2016 Department of Wind Energy, DTU Roskilde, Denmark URL `https://webstore.iec.ch`
[3] Vera-Tudela L and Kühn M 2017 *Renewable Energy* **107** 352–360 ISSN 18790682, 09601481
[4] Abiodun O I, Jantan A, Omolara A E, Dada K V, Mohamed N A and Arshad H 2018 *Heliyon* **4** e00938
[5] Wikipedia 2012 Transfer learning — Wikipedia, the free encyclopedia [Online; accessed 12-June-2021] URL `https://en.wikipedia.org/wiki/Transfer_learning#cite_note-1`
[6] Weiss K, Khoshgoftaar T M and Wang D 2016 *Journal of Big data* **3** 1–40
[7] Pan S J and Yang Q 2010 *IEEE Transactions on Knowledge and Data Engineering* **22** 1345–1359

[8] Bak C, Zahle F, Bitsche R, Kim T, Yde A, Henriksen L C, Natarajan A and Hansen M 2013 *DTU Wind Energy Report-I-0092* **5**

[9] Team D W E H D 2016 *NREL 5MW reference wind turbine HAWC2 model* Englewood, NJ, USA URL `https://www.hawc2.dk/Download/HAWC2-Model/NREL-5-MW-Reference-Wind-Turbine`

[10] Dimitrov N K, Kelly M C, Vignaroli A and Berg J 2018 *Wind Energy Science* **3** 767–790 ISSN 23667451, 23667443

[11] Hansen M H and Henriksen L C 2013 Basic DTU wind energy controller