Transfer Learning of Model-Based Neural Network for Transfer Function Inversion and Load Monitoring of Wind Turbines

Delbos F, Caleyron F., Leroy J-M.

IFP Energies nouvelles, 1-4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France

frederic.delbos@ifpen.fr

Abstract. One main concern when using a generic model of a wind turbine is how can we apply it on real measurements for different turbines of the same type. This paper gives a first answer by proposing an inversion approach of a transfer function. Then that approach is tested on a real test case. Two transfer functions from two different turbines of the same farm are inverted on SCADA measurements with transfer learning of a model-base neural network. The inversion results show better predictions than ones with a transfer function classically trained with LIDAR data. An application to the prediction of damage equivalent load with a generic model is showed: the results of applying the inverted transfer function is compared to real DEL measurements and to the transfer function trained from LIDAR data.

1. Introduction

There are three main approaches for loads monitoring of wind turbines. The first approach consists in recording real load measurements of the structural components of interest. The results are reliable but at a prohibitive cost per wind turbine and for a limited time of monitoring. Hence load measurements can only be used to validate results of other approaches or to build a data-based surrogate model. The latter, which is the second approach, enables to build a proxy model from available SCADA (Supervisory Control and Data Acquisition) turbine data to loads [1]. A first difficulty is to ensure that the training data (limited measurements in time) contains a representative amount of the diversity of real data. For instance, the proxy model may behave poorly for environmental conditions which are not met in the training data. Another pitfall of this approach is the difficulty to generalize proxies from one turbine to other turbines of the same type in the same wind farm. The last approach is the model-based approach [2] where an aero-servo-elastic dynamic model of the turbine is designed to predict loads for different environmental conditions. One difficulty relies on the modelling error. A careful choice of model assumptions is required to produce reliable results. But once this step is validated, there is still a gap between synthetic and real data. For example, some of the model input parameters are the free wind mean and standard deviation values in front of the turbine while wind data is generally recorded in the lee side of the turbine. Therefore, mean and standard deviation of the wind speed obtained from SCADA are tainted with uncertainties and should be corrected with a wind transfer function.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

2. Objectives

This paper focus on the model-based approach and the computation of the wind transfer function (TF). That function is traditionally computed using additional data. For instance, one can use LIDAR data that record the free wind statistics in front of wind turbines. But LIDAR data are still expensive to use and are not widely available in wind farms. This is the reason why the main first objective of this paper is to address the wind TF inversion without the need of additional data. And that is the purpose of the transfer learning method described in the next section. Transfer learning see [5] for a recent review, is not a new method and has been investigated for several years for improving surrogate model outputs of wind turbine. For instance, in [6] transfer learning is used for an application in wind turbine fault detection. But its use for inversion as described in this paper is a novel concept not found in the literature. That concept is also a first step for the use of one generic model that enable to predict reliable outputs for many turbines of the same type. This is showed in the fourth section by inverting two TF associated with two turbines of the same farm (using the same generic model). Then, a second objective is to apply that new TF function and show its effects on loads prediction. In the fifth section, measurements results from [6] are compared to DEL predictions using a standard TF (from LIDAR data) or an inverted one (from transfer learning).

3. Methodology

The first step of the workflow consists in generating a model database. An aero-servo-elastic dynamic model representative of 2 MW MM82 Senvion turbines is used whose control, aerodynamics and structural parameters are calibrated on available data. Then, bound constraints and a non-linear constraint are applied on the input parameters to cover all the situations the asset is expected to experience. A maximin LHS (Latin Hypercube Sample) is performed to get an optimal design of experiments (DOE). That DOE is in dimension 3 with the following parameters: the free mean wind speed U (ranging from 4m/s to 22m/s), the free wind standard deviation σ_U (from 0m/s to 3.5m/s) and the mean nacelle yaw angle γ (from -20° to 20°). The DOE is composed of 600 points: 500 for training models and 100 for testing (designed with an augment LHS). For each point, 6 wind realizations are randomly chosen each respecting mean and standard deviation parameters. Based on this design, 10 min dynamic simulations in various operational and environmental conditions are performed. That represents a total of 3600 simulations. The results of all these simulations are recorded in the model database.

In the second step, we start by dividing the SCADA scalar data into 2 classes. The first class is composed of uncertain real data or real data which does not exactly conform with the input variables used for the physical model. For our case, we grouped in this class U, the mean wind speed, σ_U the wind standard deviation, and γ the mean nacelle yaw angle. The second class contains the last 3 SCADA scalars with smaller uncertainties: Pw, the generated power, ω, the rotational speed and θ, the pitch angle of the blades. Then, as depicted in figure 1, 3 independent MLP NN (Multi-Layer Perceptron Neural Network) denominated NN1 are trained to link data of the first class to each data of the second class based on the model database. By convention, for the rest of this paper tilde variables refer to predictions. Two of the NN hyper-parameters describe the NN design: the number of hidden layers can range from 1 to 3 and the number of neurons per layer from 10 to 20. The best model for each output is selected using a 5-fold cross-validation method.

The third step consists in applying a transfer learning method that enables to fit a wind transfer function (called NN2 in figure 1) that will be used to correct uncertain SCADA data, namely U, σ_U and γ. That function applies a correction to the real SCADA data of the first class so that they can be passed as input of the 3 MLP NN (NN1) already trained in the third step. Figure 1 describes the overall NN architecture. The grey box on the right part of the figure corresponds to the already trained NN models NN1. Its weights are kept constant in the learning process. The residual network NN2 presented on the left part of the figure represents the wind transfer function. As it should only learn a correction of the

input data, a switch connection is added between the input layer and its output layer. The hyper-parameters describing its design are the same than the ones used in the second step. A regularization connection is also added: the outputs of the residual NN are concatenated with the general output layer. The loss function to be minimized is defined as

$$l(w) = (1-\lambda)f(w) + \lambda r(w),$$
$$\text{with } f(w) = \|Pw - \widetilde{Pw}\|^2 + \|\omega - \tilde{\omega}\|^2 + \|\theta - \tilde{\theta}\|^2,$$
$$\text{and } r(w) = \|U - \widetilde{U}\|^2 + \|\sigma_U - \widetilde{\sigma_U}\|^2 + \|\gamma - \tilde{\gamma}\|^2,$$

where w are the unknown weights of the residual NN, $\lambda \in [0,1]$ is the regularization parameter, f is the cost function (or the calibration function) and r is the regularization function. The value of the regularization parameter λ is not obvious to fix. One well known method to overcome this problem is to use a continuation method [4]. Starting from a large regularization weight, the regularization parameter is regularly decreased to retrieve more and more corrections until the data are fitted with the sufficient accuracy. The solution is thus an acceptable level of corrections that fits the data up to a certain accuracy. Applying the continuation method in the NN training process, consists then in starting with a value of 1 for the regularization parameter and dividing it by 2 for each passed 10 epochs of training.

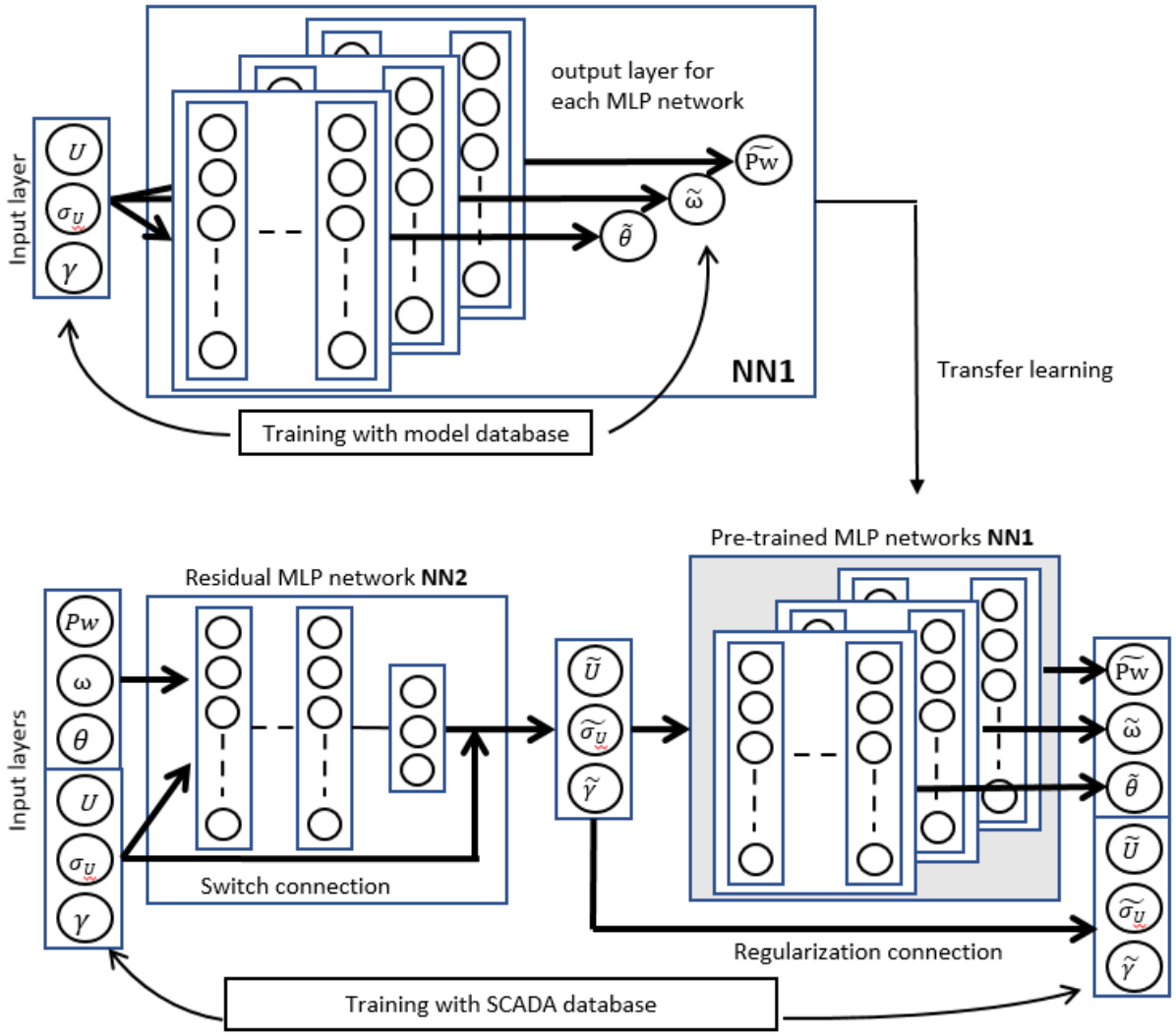


Figure 1. Diagram of the TF-TL method

The last step consists in building surrogate load models from the model database. Surrogate models that link available real data (typically 10 min SCADA statistics, sensors time series) to required outputs (typically Damage Equivalent Loads, DELs) are created. These surrogate models can be of any type because they will only be used for prediction. Finally, the inverted transfer function is combined with that surrogate model to predict load or fatigue of the wind turbine (not represented on Figure 1).

4. Transfer function inversion

Before carrying the inversion process, we first need to build models NN1 for the target parameters (second step of the previous section). In order to do that, we use the model database to train independent MLP NN regressions of the generated power, the rotational speed and the pitch angle. The model database is splitted in 500 training samples and 100 test samples. Table 1 synthetize results obtained for linear regression and the best MLP NN. The later gives very good results in train or test phase with small RMSE and R2 closed to 1. It outperforms the linear regression with a factor of 10. The best MLP model obtained for the generated power (resp. the rotational speed and the pitch angle) is composed of 2 hidden layers with 20 neurons (resp 14 and 19).

Table 1. Training and test results for predictions of generated power, rotational speed and pitch of a generic 2MW MM82 Senvion turbine

Regression method	Generated power (Pw)				Rotational speed (ω)				Pitch (θ)			
	Train		Test		Train		Test		Train		Test	
	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2
Linear Regression	255.01	0.857	245.43	0.865	1.279	0.517	1.240	0.514	1.600	0.948	1.546	0.950
MLP	24.72	0.999	23.53	0.999	0.065	0.999	0.098	0.997	0.238	0.999	0.263	0.999

The inversion workflow has been applied on an onshore wind farm owned and operated by ENGIE Green. The farm is composed of 7 x 2 MW MM82 Senvion turbines for which 10 years of 10 minutes SCADA statistics are available. Two turbines have been heavily instrumented for research purpose which allows to compare measured and calculated fatigue loads at the blades root, see [3]. In table 1, one can see that the total number of available SCADA data is large for both turbines (greater than 2E+5 raw data). LIDAR data mounted on the SMV6 nacelle are also available. That allows to build a wind TF-LID transfer function to recover the free wind 10 minutes statistics. And then enable to do comparisons between results with initial predictions (predictions without wind TF), wind TF-TL predictions (predictions with the wind TF from the transfer learning method) and wind TF-LID predictions (predictions with the wind TF from LIDAR data).

Table 2. Available data for SMV5 and SMV6 turbines

Turbine	10 minutes SCADA data in "normal operation"	10 minutes DEL for turbine "normal operation" and outside wake sectors	LIDAR
SMV5	340277	-	NO
SMV6	244427	3772	YES

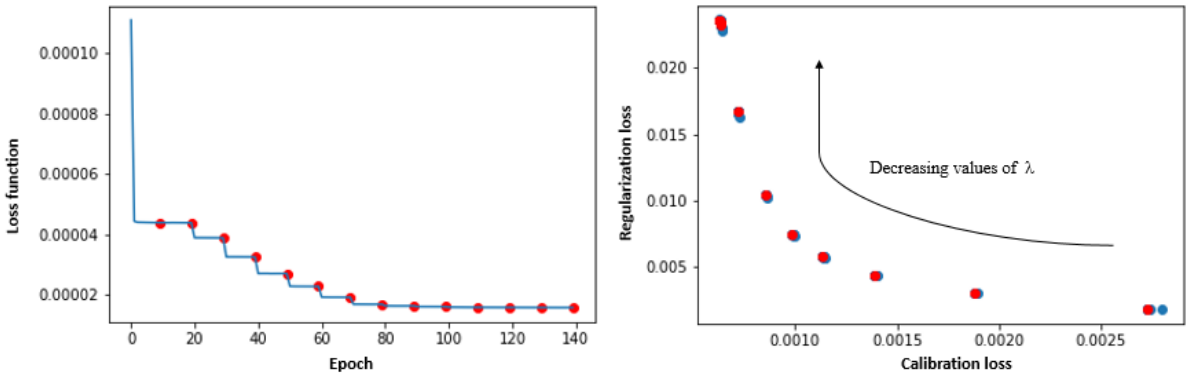


Figure 2. SMV6 turbine - left: loss function versus epoch during training of the TF-TL method for the best design configuration; right: Regularization versus calibration loss during training of the TF-TL method for the best design configuration

In this study, the results are computed for the SMV6 and SMV5 turbines. For these turbines, there is one large mode in the wind distribution between 5 and 10m/s which means that the repartition of data with respect to U is highly imbalance (compared to the model database). Hence, to have a more representative inversed TF for all values of U we have used weight sampling during training. The allocated weight for each sample was fixed as a proportional value to the inverse to its bin frequency (up to a minimum frequency). As described in previous section we then apply a continuous approach to find the best regularization parameter. Figure 2, left, shows the decreasing of the loss function with respect to epoch (each red dot corresponds to a decrease by 2 of the regularization parameter) for the SMV6 turbine. And figure 2 right, shows the "L curve" plot which is difficult to interpret.

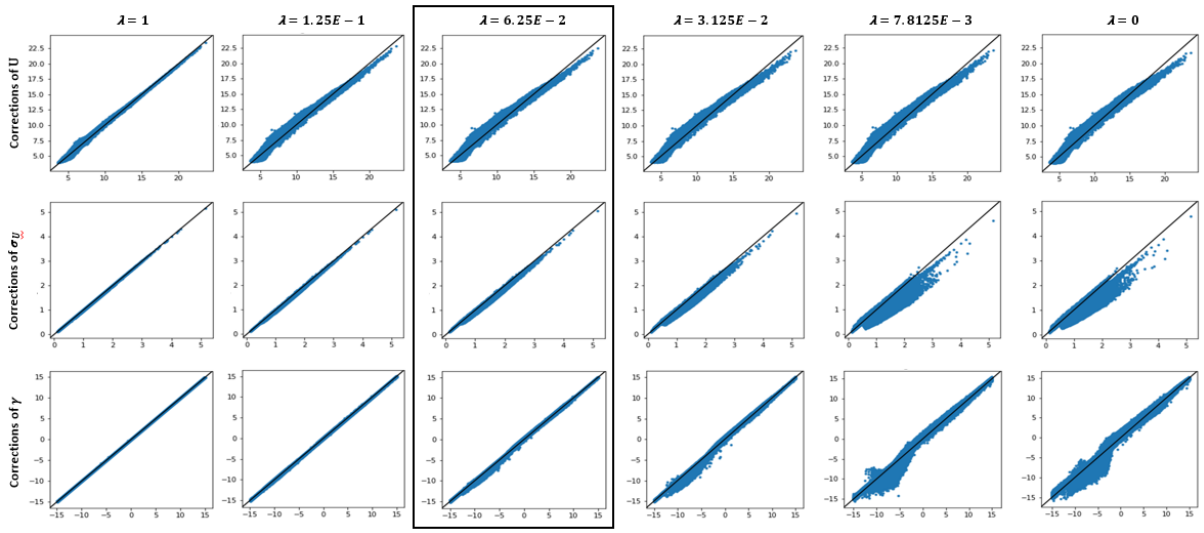


Figure 3. SMV6 turbine – corrections result for the TF-TL method with decreasing values of λ

In figure 3, the parity plots of the corrected parameters have been represented for decreasing values of λin order to give more significant insight of the inversion process. One can see that the mean wind speed U (plots of the first line) is well recovered. A good solution appears for a large λ (1.25E-1) and then the corrections for U change slowly with decreasing values of λ. At the opposite the wind standard deviation σ_U (plots of the second line), and the mean nacelle yaw angle γ (plots of the third line) are not well recovered: their corrections continue to grow while λ is decreasing. There is mainly two explanation: first the information from the fitted data may not sufficient to invert these two

parameters and second there may be also an indetermination between the values of σ_U and γ. That is the reason why we have chosen to keep an intermediate value for λ of 6.25E-2 (plots in the frame column of figure 3): that enables to have a good correction for U and one not too large for the other two parameters. Note that the inversion of the SMV5 turbine data shows the same behavior with λ. For the following of this paper the results of the TF-TL method refer to the inversion results for λ=6.25E-2.

We can now compare the wind TF-TL inversion results with the initial predictions and with the wind TF-LID results. Table 3 and figure 3 (resp. table 4 and figure 4) show these comparisons for the SMV5 turbine (resp. for the SMV6 turbine). The left-hand side of figure 4 (resp. figure 5), shows the corrections applied by the two TF method for the uncertain parameters. The TF-TL methods seems to well capture corrections for the wind speed U: the shape of the correction looks like the ones of the TF-LID method. Concerning σ_U both methods give in general lower values than measurements. That is what is expected because the measurements recorded in the lee side of the turbine are submitted to additional blade turbulences. But TF-LID corrections are larger: TF-TL method is not able to completely capture these corrections. One main reason is that the fitted SCADA data may not be very sensitive to variations of σ_U. One advantage of the TF-TL method is that other type of data can be corrected like here the nacelle yaw angle γ. On the right-hand side of figure 4 (resp. figure 5) the different predictions for the other data are plotted. Without great surprises the initial predictions show values that are not well fitted, and the TF methods shows better match of Pw, ω, and θ. But the TF-TL prediction plots show a narrower spread of values compared to the ones of the TF-LID which means in general more accurate predictions. We can also observe it on table 3 (resp. table 4) where we note a factor 2 of improvements for Pw and a factor 3.5 for ω. The predictions of θ are of the same level of accuracy for both methods.

Table 3. SMV5 turbine - RMSE values for corrections and predictions results for TF-LID and TF-TL methods

	U	σ_U	γ	Pw	ω	θ
Initial	-	-	-	157.7	0.724	1.247
TF-LID	0.379	0.190	-	98.6	0.288	1.141
TF-TL	0.557	0.067	0.077	56.3	0.079	1.156

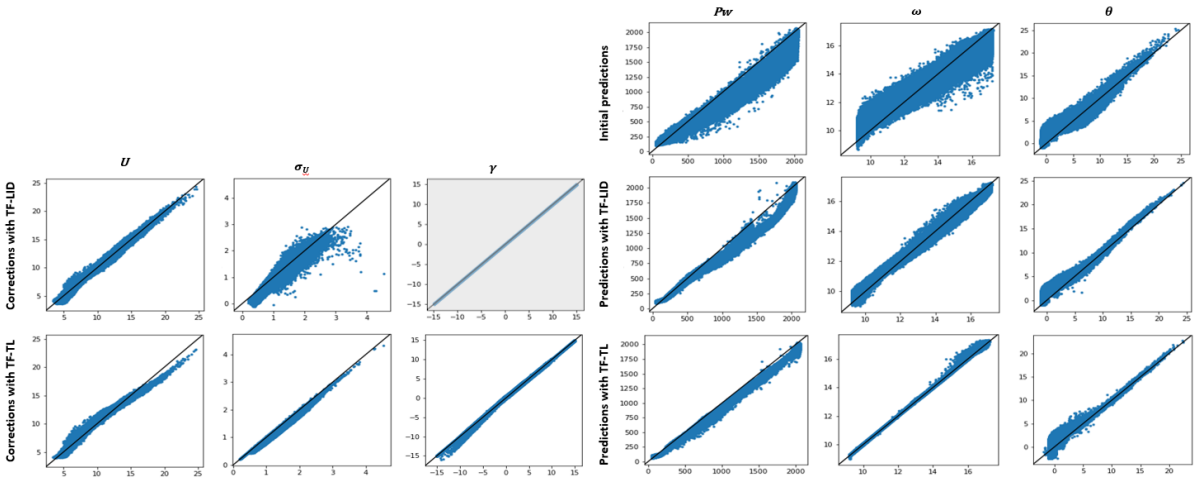


Figure 4. SMV5 turbine - corrections and predictions results for TF-LID and TF-TL methods

Table 4. SMV6 turbine - RMSE values for corrections and predictions results for TF-LID and TF-TL methods

	U	σ_U	γ	Pw	ω	θ
Initial	-	-	-	178.8	0.927	1.299
TF-LID	0.497	0.194	-	102.2	0.284	1.168
TF-TL	0.700	0.056	0.127	54.1	0.083	1.122

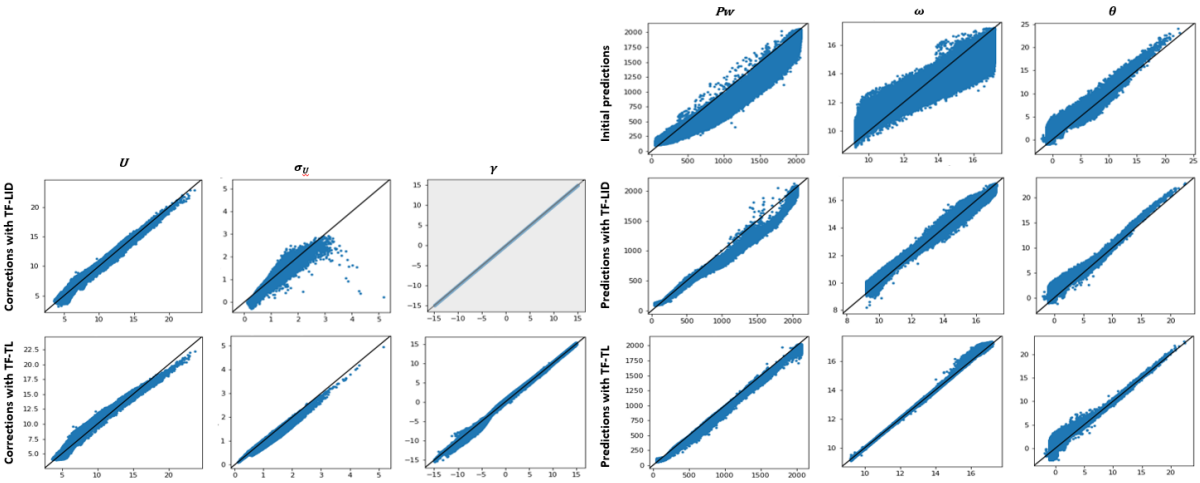


Figure 5. SMV6 turbine - corrections and predictions results for TF-LID and TF-TL methods

5. Application to blade root flap damage equivalent load predictions

The available wind data (mean and standard deviation) allow to estimate damage which are statistically satisfactory for a long-term estimation. To have more robust results for short term efforts (over a given 10 min period), it is preferable to have high frequency data. In this application we only use 10 min SCADA data.

To predict DEL we first need to build an accurate synthetic model that takes SCADA values as input and gives DEL as output. Using the model database (composed of 500 training samples and 100 test samples), we have trained 6 different types of regression models to predict normalized blades root flap DELs. Table 5 synthetize the results of the best results (among many varying hyperparameters) obtained by each regression method. The best method when looking for small RMSE in train and test results is the Support Vector Machine (SVM).

Table 5. Training and test results for prediction of normalized blades root flap damage equivalent loads

Regression method	Train		Test	
	RMSE	R2	RMSE	R2
Linear Regression	48.43	0.931	49.81	0.914
SVM	**16.51**	**0.992**	**20.56**	**0.985**
Random Forest	19.08	0.989	38.39	0.949
Gradient Boosting	4.65	0.999	26.72	0.975
MLP	17.17	0.991	22.27	0.983
Gaussian Process	29.87	0.974	32.24	0.964

From table 1, one can observe that load measurements, and their associated 10 minutes statistics DELs, are available for the SMV6 turbine and within a 3-month period from October 2017 to December 2017. That leads to 3772 available measured DEL values for records in “normal operation” and outside wake

sectors. In figure 6 the comparison between measured and predicted normalized blades root flap damage equivalent loads is showed for 3 months of operations as a function of the mean wind speed. The plot allows to compare the statistical distribution of each variable as it represents the median and the quantiles 1%. 25%. 75% and 99%. It is shown that TF-TL predictions give better estimates of the fatigue loads than the initial predictions with only 10 minutes SCADA inputs. The TF-LID predictions are even better because they use the additional LIDAR data to build the TF. Load predictions for wind speed greater than 12m/s are globally underestimated due to bad prediction of the underlying physical model.

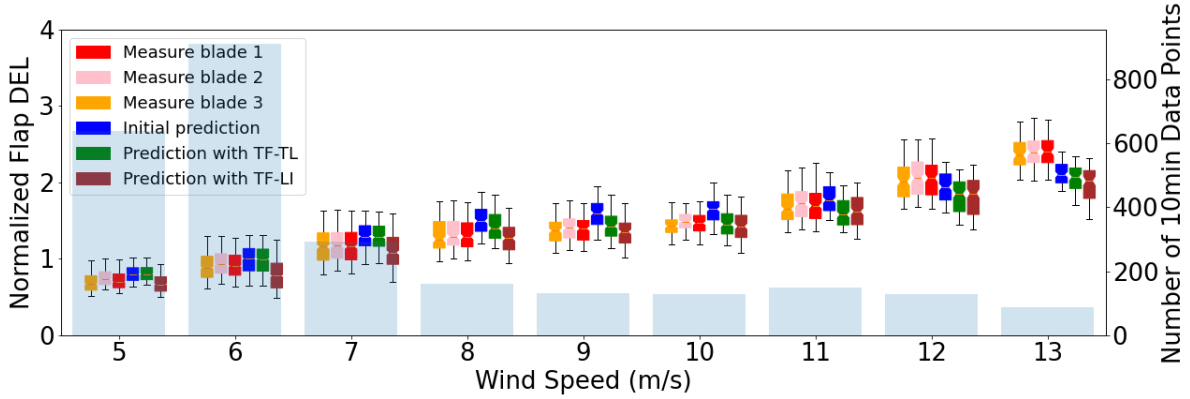


Figure 6. Comparison between measured and predicted normalized blades root flap damage equivalent loads

6. Conclusions

In this article we present the TF-TL method. a novel method that enables to reconciliate model-based results with real data using only SCADA data. Inversion results show better predictions of the TF-TL method than the TF-LID method. As shown in the results for two different turbines of the same farm, the TF-TL method can easily be applied for all turbines with the same underlying physical model. It is a first step on how to use of generic synthetic model on real measurements data. Results on load predictions show small improvements compared to initial predictions. That work open many perspectives. First. assuming LIDAR data are available. more accurate results could be obtained by applying the TF-TL method to data firstly corrected by the TF-LID method. Second. assuming load measurements data are available a RNN can be added at the end of the network to decrease differences between current predictions and real measurements. That may lead to better integration of model and data knowledge. And finally. other type of SCADA data can be also incorporated to enhance the TF and more particularly to recover a better correction for the wind standard deviation.

References

[1] Venu A et al 2020. Prediction and Validation of Fatigue loads using Artificial Intelligence on Real World Measurement Data. *J. Phys.: Conf. Ser.* 1618 022006.
[2] Dimitrov N K and Natarajan A 2019 From SCADA to lifetime assessment and performance optimization: how to use models and machine learning to extract useful insights from limited data. *J. Phys.: Conf. Ser.* (Online). 1222.
[3] Tagliatti F 2019 Investigations of Wind Turbine Fatigue Loads under Wind Farm Control: Analysis of Field Measurements. *Master of Science. Delft University of Technology*.
[4] Hansen P C 1992 Analysis of discrete ill-posed problems by mean of the Lcurve. *SIAM Review, 34, 561–580.*
[5] Zhuang, F. et al. A comprehensive survey on transfer learning. *Proc. IEEE 2020*, 109, 43–76.
[6] Schröder, L. et al 2022. Using Transfer Learning to Build Physics-Informed Machine Learning Models for Improved Wind Farm Monitoring. *Energies 2022*, 15, 558.

[7] project SMARTEOLE (ANR-14-CE05-0034).

The present work has used experimental data from the French national project SMARTEOLE (ANR-14-CE05-0034) measurement campaigns.