Soft Ordering 1-D CNN to Estimate the Capacity Factor of Windfarms for Identifying the Age-Related Performance Degradation

Manuel S Mathew[1], Surya Teja Kandukuri[2], Christian W Omlin[3]

[1,3]*University of Agder, Jon Lilletuns vei 9, 4879 Grimstad, Norway*
manuel.s.mathew@uia.no
christian.omlin@uia.no

[2]*Norwegian Research Centre, Energy & Technology Department, Tullins Gate 2, 0166 Oslo, Norway*
suka@norceresearch.no

ABSTRACT

Wind energy plays a vital role in meeting the sustainable development goals set forth by the United Nations. Performance of wind energy farms degrades gradually with aging. For deriving maximum benefits from these capital-intensive projects, these degradation patten should be analyzed and understood. Variations in the capacity factor over the years could be an indication of the age-related degradation of the wind farms. In this study, we propose a novel data-driven model to estimate the capacity factor of wind farms, which could then be used to estimate its age-related performance decline. For this, a 1-dimensional convolutional neural network (1-D CNN) is developed with a soft ordering mechanism under this study. The model was optimized using Huber loss to counteract the effects of outliers in data. The developed model could perform very well in capturing the underlying dynamics in the data as evidenced by a normalized root mean squared error (NRMSE) of 0.102 and a mean absolute error (MAE) of 0.035 on the test dataset.

1. INTRODUCTION

The United Nations and its member states have set forth the sustainable development goals (SDGs), in which SDG 7 outlines a commitment towards "*ensuring access to affordable, reliable, sustainable, and modern energy for all*" (Sachs, Kroll, Lafortune, Fuller, & Woelm, 2022). Five key targets have been identified towards attaining this goal. Targets 7.1 and 7.2 are of particular interest (*Goal 7: Affordable and clean energy*, 2024):

- Universal access to affordable and clean energy sources prioritizing the transition to renewable energy and energy-efficient technologies by 2030 (Target 7.1).
- Increasing the share of renewable energy in the global energy mix, encouraging the adoption of cleaner and greener alternatives to fossil fuels by 2030 (Target 7.2).

Wind energy, with its meteoric growth in recent years, will play a significant role in contributing towards these targets. For example, the share of wind energy in the global energy mix has increased from 342 TWh in 2010 to 2, 125 TWh in 2022 (International Energy Agency, 2023). With many large-scale wind projects in various stages of development, this trend is expected to continue in the coming years as well.

Wind turbines in a farm are often exposed to complex and harsh operational environments which adversely affects its health conditions and thereby its life expectancy. The average lifetime of wind turbines varies from 20 to 25 years, depending on the design features and operational environment (Adedipe, & Shafiee, 2021; Ziegler, Gonzalez, Rubert, Smolka, & Melero, 2018). During this period, wind turbines undergo gradual degradation in performance owing to the mechanical wear and tear over the years (Hamilton, Millstein, Bolinger, Wiser, & Jeong, 2020; Pan, Hong, Chen, Feng, & Wu, 2021), or the reduction in aerodynamic efficiency due to material erosion over the blade tips (Mathew, Kandukuri, & Omlin, 2022; Ravishankara, Ozdemir, & Weide; Sareen, Sapre, & Selig, 2014). It is estimated that, in Europe, nearly half of the wind turbines in operation will reach their end of designed life by 2030 (Windeurope Asbl/Vzw, 2024). Thus, estimation of the long-term performance of wind turbines in a farm is essential for identifying the possible system degradations over the years and thereby to plan the maintenance strategies and end-of-life decision support.

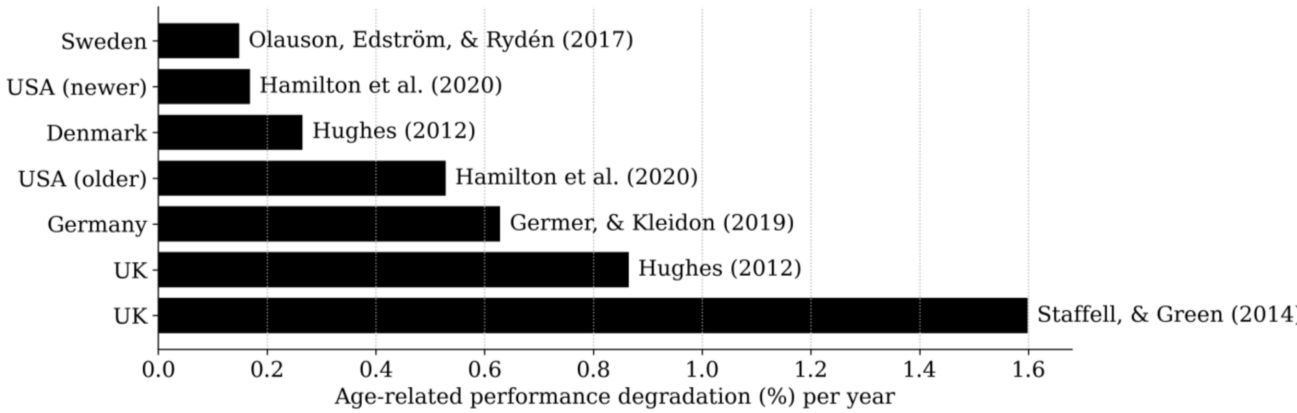

Figure 1. Age-related decline in capacity factor as reported by different studies.

Despite the significance of understanding the health status of wind turbines through its performance degradation during its lifetime, most of the earlier studies on condition monitoring focus solely on component level system reliability and availability (Staffell, & Green, 2014).

Wind turbines have several components integrated within the system and several of such turbines work together with mutual interactions in a wind farm. Hence, an analysis at turbine and farm level would help in giving a wholistic picture of the degradation issues. With the extensive deployment of supervisory control and data acquisition (SCADA) systems, time series performance of wind turbines and farms can be analyzed using data-driven models. Further, the degradation pattern in wind turbines over their life span is highly site-specific in nature (Mathew et al., 2022). Thus, data-driven models can help in estimating the performance degradation in wind turbines accurately and accounting for site-specific factors leading to their degradation. The authors have earlier developed a site-specific degradation estimation model for a wind turbine operating in Norway (Mathew et al., 2022). It was found that the reduction in performance of a wind turbine can be estimated using SCADA data and data-driven models. Further, it was estimated that on average, the performance of the wind turbine under study declined 0.64% every year of its operation. Similar studies have been carried out for turbine-level estimation of performance degradation at Irish, and Italian sites (Astolfi, Byrne, & Castellani, 2021; Byrne, Astolfi, Castellani, & Hewitt, 2020), showing degradation estimates of 8.8% and 1.5% over 12 years of operation. Such wide variation in the performance degradations of wind turbines further strengthens the argument for their site-specific analysis.

At a wind farm level, age-related decline in efficiency is quantified using the plant capacity factor (C_f), which is the ratio of the actual energy produced by the wind farm to the maximum possible energy it could have produced if it were operating at full capacity over the same period. In one of the earliest studies in estimating the wind farm level performance degradation, Hughes (2012) calculated the monthly capacity factor of wind farms operating in the UK and Denmark using 10 years of operational data, which was used to estimate the decline in performance of 13% in the UK and 4% in Denmark over the course of its operation, respectively. Similar results were reported by several studies (Germer, & Kleidon, 2019; Hamilton et al., 2020; Hughes, 2012; Olauson, Edström, & Rydén, 2017; Staffell, & Green, 2014) in the literature as illustrated in Figure 1. In the figure, the age-related decline in performance of wind farms estimated using capacity factor is normalized to per year values as reported in these studies.

These studies help in understanding the age-related performance decline in wind farm level and reiterate the regional and site-specific nature of the degradation phenomenon. However, most of these studies are based on cumulative data from different windfarms collected from public databases. Additionally, they depend on modelling the capacity factor based on meteorological reanalysis data and manufacturer's power curve (MPC) of the wind turbine. Hence, these studies are not based on the data measured from the specific wind farm site under study. The site-specific dynamics play a significant role in the age-related performance degradation of wind turbines, and the performance estimated using MPCs generally differ significantly from field performance of the turbines (Veena, Manuel, Mathew, & Petra, 2020). This could adversely affect the accuracy of these analyses. A more systematic and accurate analysis of the wind farm level performance degradation can be achieved through models based on the site-specific data, derived from the SCADA systems.

In this paper, we propose a deep neural network-based model to estimate the capacity factor of wind farms which can further be used for identifying the age-related performance degradation in wind farms. Apart from using the realistic data derived from SCADA for the site-specific analysis as discussed above, another novelty of the study is the use of convolutional neural network (CNN) model with the soft ordering mechanism. The remainder of the paper is organized

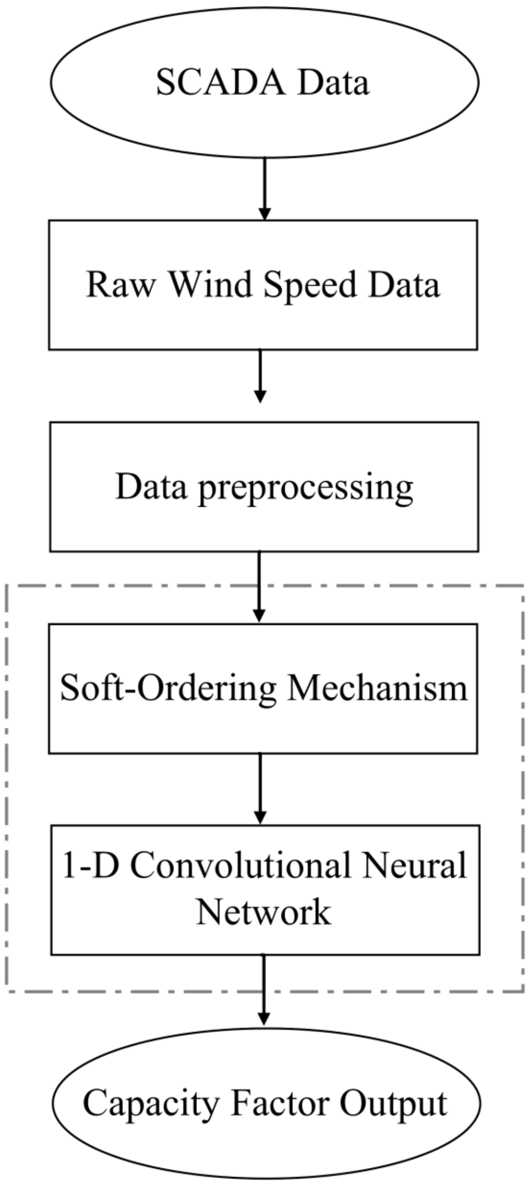

Figure 2. Schematic representation of modelling.

as follows: Section 2 starts by explaining the rationale behind using CNNs. In Section 2.1, the theoretical framework behind CNNs is briefly explained. The soft ordering mechanism employed in order to transform the input data into appropriate inputs to the CNNs is introduced in Section 2.2. Section 2.3 briefly discusses the 1-D CNN architecture and Section 2.4 describes the methodology followed in training, validating, and testing this model. The results from this study are detailed in Section 3 and finally Section 4 concludes this work and traces the next steps in this ongoing study.

2. METHODOLOGY

The performance of wind turbines in a wind farm is significantly influenced by the high spatial and local correlation of wind speed at each of the turbines through site-specific wake effects. But these correlations are further complicated due to the directional and stochastic nature of wind, making it harder for a straightforward analysis. Owing to their capability to extract salient feature representations from data with inherent spatial and local correlations, CNNs are a compelling approach to be explored. The overview of the methodology in estimating the capacity factor is shown in Figure 2.

2.1. Convolutional Neural Networks

The model for estimating the capacity factor in this study is based on CNN. CNNs are inspired by the natural vision in mammals and were popularized by Lecun et al. (1989) particularly for image recognition tasks. Even though the theoretical framework for CNNs predates this work, they used this architecture for automated extraction of features for vision related tasks.

Convolutional layers are the fundamental building blocks in CNN. They serve as the feature extractors exploiting local connectivity, and spatial locality (Kiranyaz et al., 2021; Rawat, & Wang, 2017). In convolutional layers, a learned kernel convolves with the input producing a feature map. The property of local connectivity arises from the fact that each element in the feature map is connected to a local subset of neurons in the previous layer or the input pixels. Spatial locality, on the other hand, is the result of the high correlation between the local subset of input to the convolutional layer. The feature map element at (i, j) in the kth feature map of the lth layer can be calculated as:

$$z_{i,j,k}^{l} = \boldsymbol{w}_{k}^{l} \boldsymbol{x}_{i,j}^{l} + b_{k}^{l} \tag{1}$$

where $\boldsymbol{w}_k^l$ and b_k^l are the weight vector and bias term of the k^{th} filter of the l^{th} layer, respectively. $\boldsymbol{x}_{i,j}^l$ is the local subset of input to the convolutional layer centered at (i, j). However, when used for tabular dataset, convolutional layers expect spatial and local correlation between the features. Non-linearity is generally introduced after convolution by using elementwise non-linear activation functions such as rectified linear unit (ReLU). ReLU outputs the input values as such if the input is positive and zero if the input value is negative.

$$a_{i,j,k}^{l} = max\left(0,\ z_{i,j,k}^{l}\right) \tag{2}$$

where $a_{i,j,k}^l$ is the activation at position (i, j, k) in layer l after applying ReLU function.

Pooling layers are an optional layer in CNNs which introduce shift-invariance to the feature maps produced by convolutional layers. Shift-invariance is achieved by reducing the resolution of the feature maps through average pooling or max-pooling depending on the task. The average pooling operation is given by:

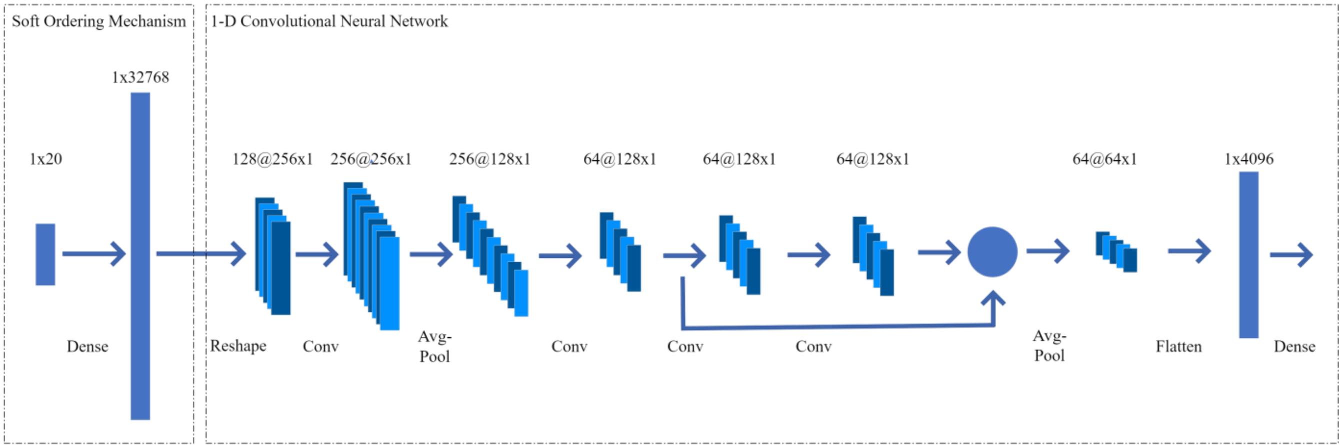

Figure 3. Architecture of the proposed 1-D CNN with soft ordering mechanism.

$$AvgPooling(\boldsymbol{A})_{i,j,k} = {}^{1}/_{N} \sum\nolimits_{m,n \in R_{i,j}} a^{l-1}_{m,n,k} \qquad (3)$$

where $\boldsymbol{A}$ is the activation map from layer (l-1), N is the number of elements in the pooling window $R_{i,j}$ of dimension (m, n), and $a^{l-1}_{m,n,k}$ is the activations from the pooling layer within the pooling window.

Several convolutional and pooling layers are stacked in a CNN to extract higher level feature representations. Further one or more fully connected (FC) layers are used to achieve higher level reasoning in CNNs (Simonyan, & Zisserman, 2014). The output layer is the final layer that uses task appropriate activation functions (e.g., sigmoid for classification or ReLU for regression).

However, a key challenge is that CNNs are designed to work with data presented in a uniform grid-like structure akin to images. The wind speed input from the wind farm cannot be fed directly to the CNNs assuming each point as a "pixel" in the pseudo-image, as generated from the wind farm layout. This is because the layout of wind turbines in a wind farm is non-uniform, often dictated by the availability of wind and other external factors such as terrain, land use regulations etc.

One solution for the irregular layout of wind farms is to pad the layout with zeros to make it a uniform grid like structure, which results in sparseness in the data. Sparseness in data may result in slowing down of training, reduction in model performance, and loss of spatial resolution.

To overcome these limitations, in this study, we propose a novel application of soft ordering mechanism for CNNs in estimation of the capacity factor. Under this method, the wind data, which is in tabular form, is reshaped into a multi-channel image format. The advantage of this method is that the spatial or sequential relationships of the data are preserved without the need for following a rigid order. This makes the proposed method unique and more suitable for modelling wind farms, which normally have nonrigid geometries. The proposed soft ordering 1-D CNN consists of two parts: a soft-ordering mechanism, and a 1-D CNN.

2.2. Soft ordering mechanism

Soft ordering is a technique to rearrange the data to introduce or preserve spatial or sequential relationships without following a rigid order. In this work, soft ordering is achieved by using an FC layer. The FC layer maps the input features into another higher dimensional feature space. This transformation helps in providing enough pseudo-pixels for the convolutional layers as well as to reorganize the features such that it mimics the spatial or sequential relationships in the data. The FC layer is followed by a non-linear activation function, ReLU in this work, for ensuring that the transformation can effectively learn a non-linear mapping.

Finally, the newly rearranged features are reshaped into multi-channel pseudo-images. Thus, the convolutional layer extracts the features from a rearranged non-linear transformation of the original data, and the model learns to effectively rearrange the features adaptively. Thus, the entire model can be trained in an end-to-end manner without significant preprocessing steps.

The soft ordering mechanism is shown in Figure 3, which takes in the input features and transforms them into non-linear higher-dimensional representations of size 32768. These representations are then reshaped into 128 channels with a signal size of 256 to be fed into the 1-D CNN.

2.3. 1-D CNN Architecture

As opposed to CNNs used for image tasks, where the convolution is applied to a 2-D tensor, a 1-D convolutional layer takes in a single dimensional signal and applies a convolutional kernel of similar dimensionality, typically smaller than the signal. This makes it suitable for applications like natural language processing, audio signal processing, and time series analysis.

In this work, the representations from the soft ordering mechanism are fed into the 1-D CNN. The 1-D CNN architecture is also shown in Figure 3. The first convolutional layer increases the number of feature channels to 256 while maintaining the size of each feature map at 256 by applying a convolution kernel of size 5. Subsequent adaptive average pooling layer reduces the feature map resolution to 128 x 1. The next three convolution layers apply a kernel of size 3 with a stride of length 1 and output 64 channels of feature maps of size 128. A skip connection is also added from the output of the second convolutional layer to the output of the fourth convolutional layer as shown in Figure 3 to solve the problem of vanishing gradients and hence network degradation (He, Zhang, Ren, & Sun, 2016). A second average pooling layer further reduces the size of the feature maps while ensuring enough receptive fields to facilitate learning. Finally, the output from the average pooling layer is flattened and fed into a fully connected layer which makes the capacity factor estimations. ReLU activation function is used throughout the network to introduce non-linearity except to the outputs of the FC layer in the soft ordering mechanism, where continuously differentiable exponential linear unit (CELU) activation (Barron, 2017) has been used. CELU ensures that non-linearity introduced is smooth and continuous for all values and helps in capturing the negative values effectively avoiding dying ReLU problem (Lu, Shin, Su, & Karniadakis, 2019).

Batch normalization has also been implemented to help the model learn faster and make training more stable by reducing internal covariate shift (Ioffe, & Szegedy, 2015). Further, weights normalization is also implemented to counteract vanishing or exploding gradients and improving generalization by preventing the weights from growing too large or too small (Salimans, & Kingma, 2016).

2.4. Network Training

The model was trained on a wind farm dataset operating at a Norwegian site, by collecting 13 years of operational data. Each of the twenty pitch-controlled wind turbines has a 2 MW rated capacity. The turbines have cut-in, rated, and cut-out velocities of 3 m/s, 18 m/s, and 25 m/s, respectively. The turbines had a rotor diameter of 82.4 m and were installed at a hub height of 70 m. The SCADA data from these turbines had a temporal resolution of 10 minutes (Under the non-disclosure agreement, the data cannot be shared with this paper). The wind speeds and power generated by these turbines were collected from the data and cleaned for missing data and outliers. The initial four years of data from 2007 to 2010 was used to train the model.

The capacity factor of the plant was calculated which served as the target variable and the individual wind speeds served as the features. The data was divided into training, validation, and testing sets in the ratio 3:1:1. Huber Loss was used to calculate the losses for back propagation. Huber Loss is given by:

$$l(y,x) = \begin{cases} \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} 0.5\,(\epsilon)^2, & if\ |\epsilon| < \delta \\ \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} \delta\,(|\epsilon| - 0.5\delta), & otherwise \end{cases} \tag{4}$$

where $\epsilon = y_n - x_n$, is the residual, δ is the threshold for switching between the δ-scaled L1 and L2 losses, and y_n is the model's estimation of x_n. The advantage of the Huber Loss is that it combines the benefits of both L1 loss (absolute error) and L2 loss (squared error) reducing the penalty for residuals less than the threshold and thereby making the model less sensitive to outliers than L2 loss. The Huber loss is sensitive to the threshold (δ) and was set as two times the standard deviation of the residuals from a basic regression model developed initially using inlier data. Additionally, L1 losses and L2 losses across the training epochs were monitored to ensure that the model's improvement on Huber loss is translated into real world improvement in the estimation of the model performance. Adam optimizer was used in this study for updating the parameters with $\beta_1 = 0.8$ and $\beta_2 = 0.999$. The learning rate (LR) for the optimizer was empirically set to 8×10^{-4}, with an exponential LR decay with $\gamma = 0.9$, meaning the LR would decay after each epoch gradually. This helps in having higher adjustments to the parameters in the beginning and relatively smaller ones towards the end of training. The model was trained over 200 epochs implementing an early stopping mechanism that monitors the validation losses with a patience of 25 to avoid overfitting. Further, L2 regularization was implemented to reduce the chances of overfitting. While dropout layers were investigated for better generalization, it was found that the performance of the model was worse, and convergence was very slow. In the next section, we discuss the results of this experiment in detail.

3. Results and Discussions

The various losses tracked during the training and validation phases are shown in Figure 4: (a) Huber loss, (b) L1 loss, and (c) L2 loss. As expected, the losses are high initially then quickly declining to a more gradual and stable loss condition.

The validation losses in all three of the metrics show high variability across the initial epochs as the model begins to learn from the training data quickly stabilizing showing improvements in generalizability of the model. The best performing model was detected at the 94th epoch with a training and validation loss (Huber loss) of 3.1×10^{-3} and 1.6×10^{-3}, respectively. The higher training loss observed

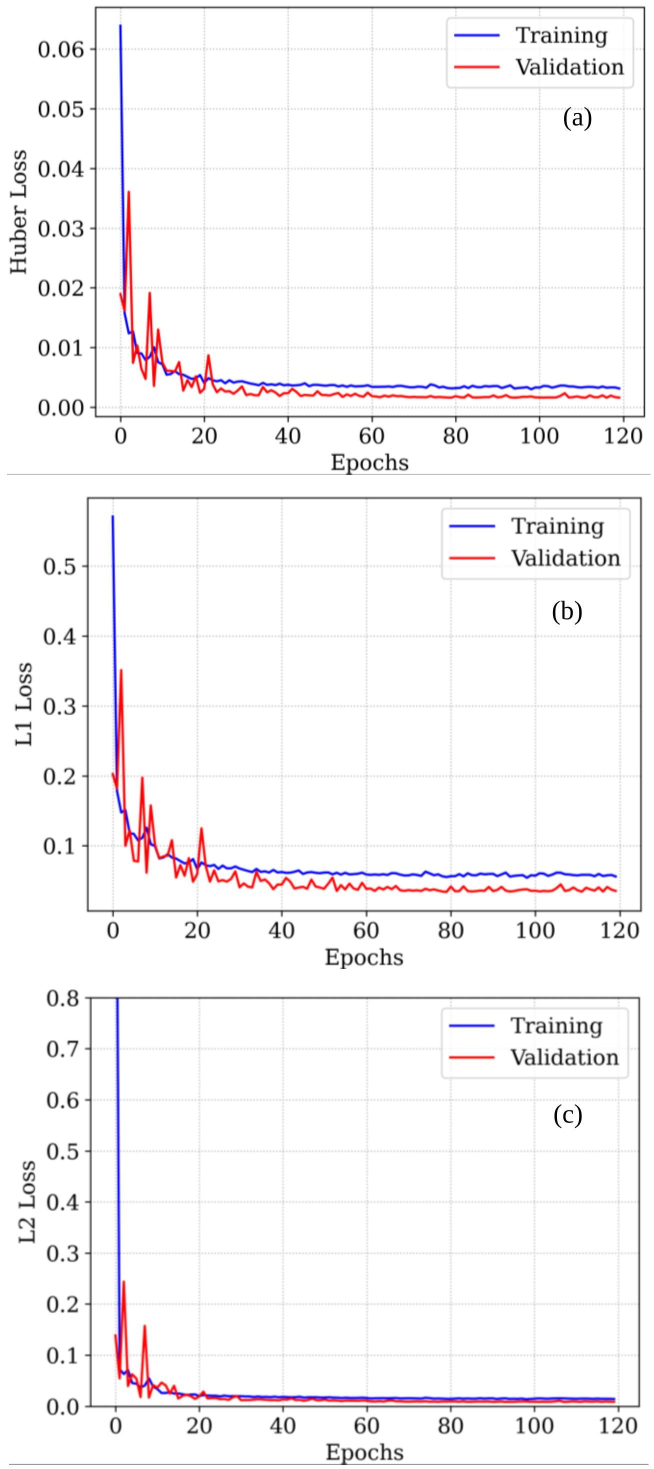

Figure 4. Various losses tracked during training: (a) Huber Loss, (b) L1 Loss, and (c) L2 Loss.

across the epochs are a result of the regularization methods applied during training. The corresponding MAE and mean squared error (MSE) for the training and validation phase can be seen in Table 1.

Table 1. Performance of the best model in training, validation, and test datasets.

Loss	Training	Validation	Test
Huber loss	3.1×10^{-2}	1.6×10^{-3}	1.7×10^{-3}
MAE	5.6×10^{-2}	3.5×10^{-2}	3.5×10^{-2}
MSE	1.4×10^{-2}	8.5×10^{-3}	1.0×10^{-2}

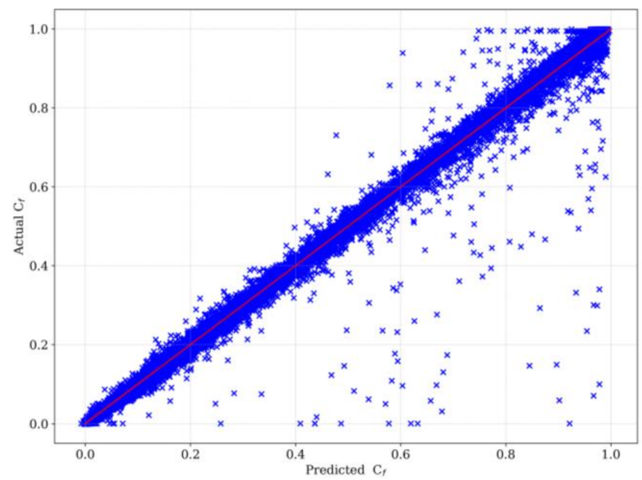

Figure 5. Comparison of the model predicted capacity factor to the measured capacity factor.

The model thus finalized, was tested with test data, which was not used in the training or validation phases to measure the generalizability of the model. Figure 5 shows the performance of the model on the test dataset. The blue scatter indicates the model prediction compared to the calculated values and the distance of these points from the red line indicates the residuals of the prediction model. The training curves (Figure 4) and the comparison in Figure 5, highlight the generalizability of the model to new data and performance of the model on new data, respectively.

The different error metrics in Table 1 quantifies this performance with a slightly higher Huber loss in predicting new datapoints. The normalized root mean squared error in predicting the capacity factor for the test dataset was 0.102. With only 0.363% of the test dataset having a residual value of more than 0.2, the model is found to be effective in capturing the plant capacity factor.

Figure 6 shows the actual and predicted power over different months in a year. It is evident that the predictions and the calculated values are in close agreement with each other. The yearly capacity factor of the farm was calculated as 0.298 against which the model prediction was 0.305. These results further support the argument that the model performs exceptionally well in predicting the wind farm capacity factor.

In previous studies on the age-related performance decline of windfarms, instead of the real data collected from the sites,

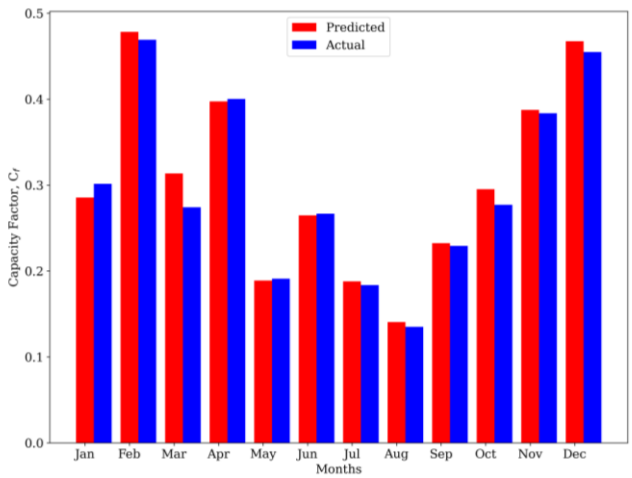

Figure 6. Comparison of actual and predicted power over different months in a year.

the wind estimates from the numerical weather prediction (NWP) models are used for estimating the capacity factors. Though the errors due this approximation are not specified in these studies, obvious differences between the NWP wind predictions and the real velocities available for the turbines could bias the model and thereby adversely affect the reliability of the results. In contrast, real wind measurements are used in the present work which resulted in accurate capacity factor predictions as evident from the low errors values. Similarly, while comparing with some CNN based studies for wind farm performance predictions (Chen et al., 2021; Kazmi, Gorgulu, Cevik, & Baydogan, 2023; Liu et al., 2021), the proposed soft ordering approach could improve the performance of the capacity factor estimations.

4. Conclusion

Wind turbines operating in a farm are exposed to complex operational conditions, causing degradation in their performance over the years of their operation. This age-related performance decline, if quantified at a wind-farm level, could contribute towards making efficient decisions at their end-of-life. As a first step towards this objective, we developed an intelligent algorithm for the estimation of wind farm capacity factor in this paper.

To predict the capacity factor of a wind farm, a 1-dimensional convolutional neural network has been trained exploiting the local connectivity inherent in wind farms. However, to sidestep the irregularity in wind farm layouts, while still using CNNs to model their performance, a soft ordering mechanism is used. The soft ordering mechanism in addition to the 1-D CNN, was able to effectively capture the inherent spatial dynamics in the wind farm as evidenced by the results discussed in the previous section. The model developed in this paper has a normalized root mean squared error of 0.102. This indicates that the errors in the model predictions are approximately 10.2 % of the range of the target values. This indicates that the proposed method could predict the capacity factor of the wind farm with high accuracy. Further, the performance of the model on previously unseen dataset (MAE: 0.035, MSE: 0.010), shows that the model can generalize well to newer data coming from the wind farm even though it was trained on data from earlier.

For developing the proposed model, high quality SCADA data are required, which may limit its applications in farms which do not have such systems in place. Nevertheless, most of the contemporary wind farms have implemented the SCADA systems and with the availability of the required data, the soft ordering 1-D CNN model developed under the study could further be used to estimate the age-related performance degradation in wind farms. This will be demonstrated by the authors through their ongoing research where logs on the turbine maintenance will also be considered.

Acknowledgements

This research work has been funded by Analytics for asset Integrity Management of Windfarms (AIMWind), under grant no. 312486, from Research Council of Norway (RCN).

AIMWind is collaborative research from University of Agder, Norwegian Research Center (NORCE), and TU Delft, with DNV and Origo Solutions as advisory partners.

References

Adedipe, T., & Shafiee, M. (2021). An economic assessment framework for decommissioning of offshore wind farms using a cost breakdown structure. *The international journal of life cycle assessment*, *26*(2), 344-370. https://doi.org/10.1007/s11367-020-01793-x

Astolfi, D., Byrne, R., & Castellani, F. (2021). Estimation of the performance aging of the vestas v52 wind turbine through comparative test case analysis. *Energies*, *14*(4), 915. https://www.mdpi.com/1996-1073/14/4/915

Barron, J. T. (2017). Continuously differentiable exponential linear units. *arXiv preprint arXiv:1704.07483*. https://doi.org/10.48550/arXiv.1704.07483

Byrne, R., Astolfi, D., Castellani, F., & Hewitt, N. J. (2020). A study of wind turbine performance decline with age through operation data analysis. *Energies*, *13*(8), 2086. https://www.mdpi.com/1996-1073/13/8/2086

Chen, X., Zhang, X., Dong, M., Huang, L., Guo, Y., & He, S. (2021). Deep learning-based prediction of wind power for multi-turbines in a wind farm. Frontiers in Energy Research, 9, 723775. https://doi.org/10.3389/fenrg.2021.723775

Germer, S., & Kleidon, A. (2019). Have wind turbines in Germany generated electricity as would be expected from the prevailing wind conditions in 2000-2014? *PLoS One*, *14*(2), e0211028. https://doi.org/10.1371/journal.pone.0211028

Goal 7: Affordable and clean energy. (2024). The Global Goals. Retrieved 13/03/2024 from https://www.globalgoals.org/goals/7-affordable-and-clean-energy/

Hamilton, S. D., Millstein, D., Bolinger, M., Wiser, R., & Jeong, S. (2020). How does wind project performance change with age in the united states? *Joule*, *4*(5), 1004-1020. https://doi.org/https://doi.org/10.1016/j.joule.2020.04.005

He, K., Zhang, X., Ren, S., & Sun, J. (2016, 27-30 June 2016). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, arXiv: 1512.03385. https://doi.org/10.48550/arXiv.1512.03385

Hughes, G. (2012). The performance of wind farms in the United Kingdom and Denmark. *Renewable Energy Foundation*, 48.

International Energy Agency. (2023). *World energy outlook 2023* (World Energy Outlook, Issue. https://www.iea.org/reports/world-energy-outlook-2023

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning, 448-456. https://doi.org/10.48550/arXiv.1502.03167

Kazmi, S., Gorgulu, B., Cevik, M., & Baydogan, M. G. (2023). A concurrent cnn-rnn approach for multi-step wind power forecasting. arXiv preprint arXiv:2301.00819. https://doi.org/10.48550/arXiv.2301.00819

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, *151*, 107398. https://doi.org/https://doi.org/10.1016/j.ymssp.2020.107398

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, *2*.

Liu, T., Huang, Z., Tian, L., Zhu, Y., Wang, H., & Feng, S. (2021). Enhancing wind turbine power forecast via convolutional neural network. *Electronics*, *10*(3), 261. https://doi.org/10.3390/electronics10030261

Lu, L., Shin, Y., Su, Y., & Karniadakis, G. E. (2019). Dying relu and initialization: Theory and numerical examples. *arXiv preprint arXiv:1903.06733*. https://doi.org/10.48550/arXiv.1903.06733

Mathew, M. S., Kandukuri, S. T., & Omlin, C. W. P. (2022). Estimation of wind turbine performance degradation with deep neural networks. 7th European Conference of the Prognostics and Health Management Society 2022, 7(1), 351-359. https://doi.org/10.36001/phme.2022.v7i1.3328

Olauson, J., Edström, P., & Rydén, J. (2017). Wind turbine performance decline in Sweden. *Wind Energy*, *20*(12), 2049-2053. https://doi.org/https://doi.org/10.1002/we.2132

Pan, Y., Hong, R., Chen, J., Feng, J., & Wu, W. (2021). Performance degradation assessment of wind turbine gearbox based on maximum mean discrepancy and multi-sensor transfer learning. *Structural Health Monitoring*, *20*(1), 118-138. https://doi.org/10.1177/1475921720919073

Ravishankara, A. K., Ozdemir, H., & Weide, E. v. d. Effect of leading edge erosion on wind turbine rotor aerodynamics. In *Aiaa scitech 2022 forum*. https://doi.org/10.2514/6.2022-0276

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, *29*(9), 2352-2449. https://doi.org/10.1162/neco_a_00990

Sachs, J. D., Kroll, C., Lafortune, G., Fuller, G., & Woelm, F. (2022). *Sustainable development report 2022*. Cambridge University Press. https://doi.org/10.1017/9781009210058

Salimans, T., & Kingma, D. P. (2016). Weight normalization: A simple reparameterization to accelerate training of deep neural networks. *Advances in neural information processing systems*, *29*. https://doi.org/10.48550/arXiv.1602.07868

Sareen, A., Sapre, C. A., & Selig, M. S. (2014). Effects of leading edge erosion on wind turbine blade performance. *Wind Energy*, *17*(10), 1531-1542. https://doi.org/https://doi.org/10.1002/we.1649

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556

Staffell, I., & Green, R. (2014). How does wind farm performance decline with age? *Renewable Energy*, *66*, 775-786. https://doi.org/https://doi.org/10.1016/j.renene.2013.10.041

Veena, R., Manuel, S. M., Mathew, S., & Petra, I. (2020). Parametric models for predicting the performance of wind turbines. *Materials Today: Proceedings*, *24*, 1795-1803. https://doi.org/https://doi.org/10.1016/j.matpr.2020.03.604

WindEurope asbl/vzw. (2024). *Wind energy today | windeurope*. Retrieved 13/03/2024 from https://windeurope.org/about-wind/wind-energy-today/

Ziegler, L., Gonzalez, E., Rubert, T., Smolka, U., & Melero, J. J. (2018). Lifetime extension of onshore wind turbines: A review covering germany, Spain, Denmark, and the UK. *Renewable and Sustainable Energy Reviews*, *82*, 1261-1271. https://doi.org/https://doi.org/10.1016/j.rser.2017.09.100

BIOGRAPHIES

Manuel S. Mathew is a PhD Research Fellow at the Information and Communication Technology department at the University of Agder, Norway. His interest is in the application of artificial intelligence in renewable energy systems particularly focusing on prognostics for wind farms. He completed his master's degree in Renewable Energy in 2021 from the University of Agder. In addition, he also holds a master's degree in systems engineering by research from the University of Brunei Darussalam. He did his bachelor's degree in electrical and electronics engineering from the Mahatma Gandhi University, India.

Surya Teja Kandukuri is a Senior Scientist at NORCE. He holds a part-time position as a Researcher at University of Agder, Grimstad, Norway. He obtained his PhD in condition monitoring from the University of Agder in 2018. He has over 12 years of experience in industrial research within aerospace, energy, marine and oil & gas sectors, developing condition monitoring solutions for high-value assets. He received his master's degree in systems and control engineering from TU Delft, The Netherlands, in 2006 and bachelor's in electrical engineering from Nagarjuna University in India in 2003.

Christian W. Omlin has been a professor of Artificial Intelligence at the University of Agder since 2018. He has previously taught at the University of South Africa, University of the Witwatersrand, Middle East Technical University, University of the South Pacific, University of the Western Cape, and Stellenbosch University. His expertise is in deep learning with a focus on applications ranging from safety to security, industrial monitoring, renewable energy, banking, sign language translation, healthcare, bio conservation, and astronomy. He is particularly interested in the balance between the desire for autonomy using AI technologies and the necessity for accountability through AI imperatives such as explainability, privacy, security, ethics, and artificial morality for society's ultimate trust in and acceptance of AI. He received his Ph.D. from Rensselaer Polytechnic Institute and his MEng from the Swiss Federal Institute of Technology, Zurich, in 1995 and 1987, respectively.