Deep Learning Enabled Data-Driven Detection of Wind Turbine Operating Conditions Using Transformer Networks

Fraunhofer IWES

Rubbiya Sultan[1,2], Nourhan Ahmed[2], Luca Faller[1], Timo Lichtenstein[1]

[1] Fraunhofer Institute for Wind Energy Systems IWES, Postkamp 12, 30159 Hannover, Germany
[2] University of Hildesheim, Universitätsplatz 1, 31141 Hildesheim, Germany

Contact:
timo.lichtenstein@iwes.fraunhofer.de

Motivation

Development of Early Fault Detection Methods
Wind turbine operating data is a promising prerequisite for early fault detection methods. Such data is available from the Supervisory Control and Data Acquisition (SCADA) system. However, method development requires labeled data containing component failures and indicating periods of normal operation, commonly derived through laborious manual pre-processing of alarm logs and service reports.

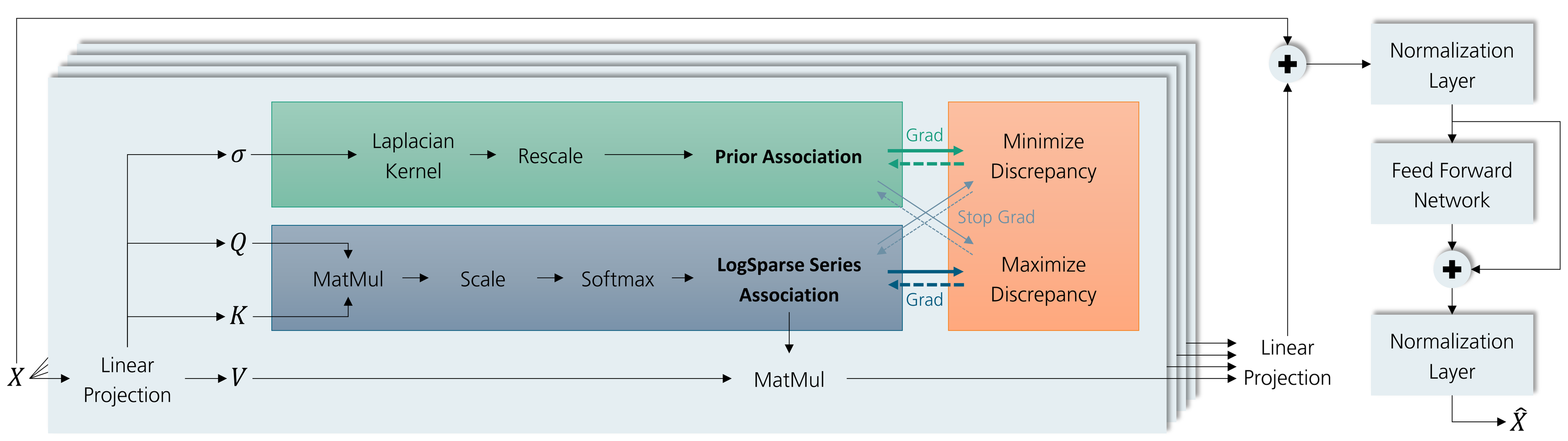
Automatic Deduction of Deviating Operating Condition Periods
Regular operational states dominate industrial operating data. Thus, novelties are rare and hard for machine learning models to learn, especially in the highly dynamic system of a wind turbine. Our solution is an unsupervised framework to classify a time series dataset of SCADA operating data from turbines in an offshore wind farm into normal and deviating operating conditions.

Transformers

Transformers have advanced significantly since they initially emerged for neural machine translation, a challenging task in natural language processing (NLP)[1]. They are also effectively employed in a variety of time series tasks, demonstrating their outstanding modeling capabilities for long-range dependencies in sequential data[2]. The primary operation of any transformer architecture, the so-called self-attention, is the scaled dot product to identify associations among distinct input segments.

Anomaly Transformer
Industrial data is characterized by the dominance of regular operating time points. To learn informative representations from complex dynamics of such data, the approach of the Anomaly Transformer[3] uses an association discrepancy to distinguish between time points of regular and anomalous or deviating behavior.

Topology of the LogSparse Novelty Transformer

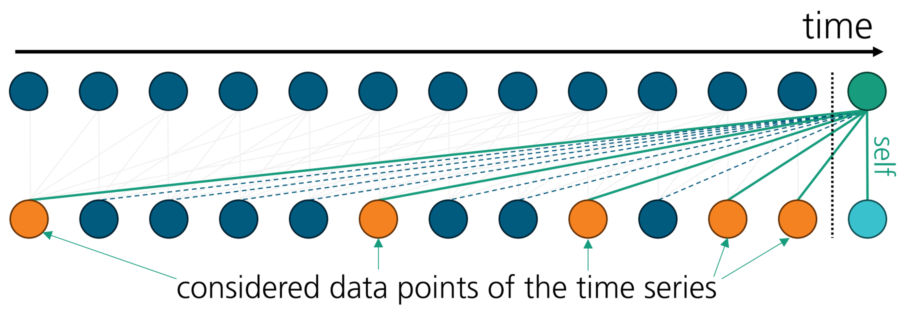


LogSparse Novelty Transformer
New architecture for the identification of time points that deviate from the regular operational pattern of a wind turbine, which may be referred to as novelties or deviations

- Added dilations to the convolution layers within the encoder of the network enhancing the receptive field of the model
- Prior association: Laplacian kernel with learnable scale parameter α instead of a Gaussian kernel, making the model more resilient to outliers
- Series association: LogSparse attention ignoring an exponentially growing number of data points with increased temporal distance for reduced memory load

Incorporating the aforementioned modifications outperforms the original Anomaly Transformer on both datasets, yielding better performance than the conventional approach.

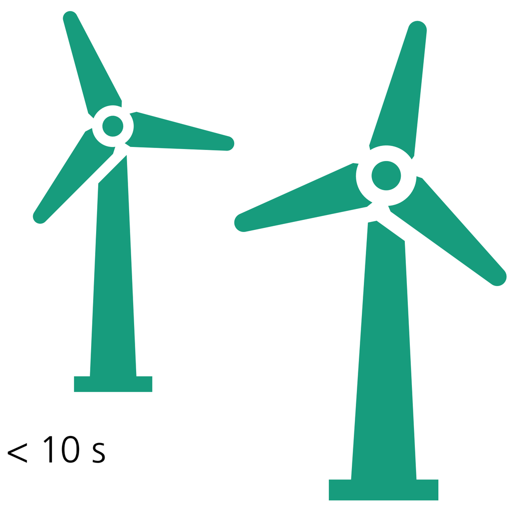
Skematic of the LogSparse Attention



Ablation Study of Modifications

Architecture	F1-Score SCADA 30 min
Anomaly Transformer (AT)	0.9652
AT + dilation d=2	0.9657
AT + dilation d=3	0.9737
AT + dilation d=4	0.9667
AT + Laplacian prior association	0.9683
AT + LogSparse series association	0.9737
LogSparse Novelty Transformer	**0.9746**

Data Basis



Wind Farm SCADA Operating Data
- 3 years of operating data of an offshore wind farm with 100 turbines of 3 MW nominal power in the North Sea
  - Two-stage planetary gearbox with high-speed shaft
  - Doubly fed induction generator (DFIG)
- Roughly 100 signals of which 19 are available with resolutions < 10 s

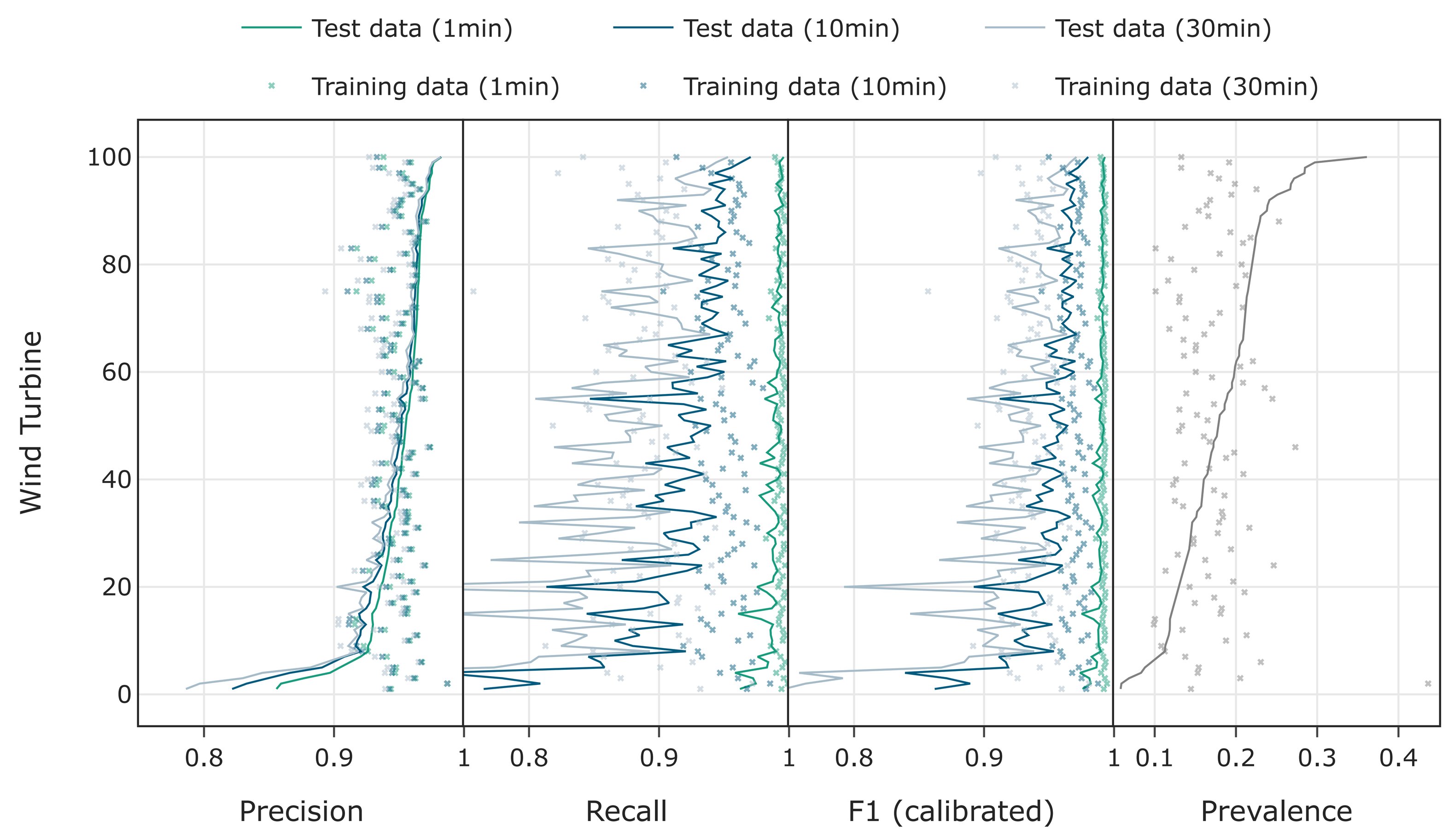
Wind Farm Event Data As Ground Truth for Deviations
- SCADA status logs that detail manual and automated actions, such as starts, stops, and pauses
- Automated alarm logs highlighting deviations in component conditions
- Maintenance dataset including task specifics, affected components, turbines, and timestamps
- The event data is not used during training, but only when assessing the performance

Data Selection
- Omission of all setpoint and control values as they do not reflect the actual turbine behavior except for the wind turbines active power output capability, resulting in 24 measured signals
- First two years for training, 80% as training, 20% as cross-validation data
- Third year of data for prediction on unseen data
- Data with the highest possible resolution was used and resampled to 1 min, 10 min, and 30 min

Evaluation Metrics

Precision	Recall	Calibrated Precision	F1-Score
Fraction of correctly detected deviating operating conditions among all predictions	Fraction of correctly detected deviating operating conditions among all real deviations	Precision normalized to the prevalence or class ratio π to allow for comparison on varying datasets[4]	The harmonic mean of precision and recall as a measure of prediction performance
$P = \frac{TP}{TP + FP}$	$R = \frac{TP}{TP + FN}$	$P' = \frac{TP}{TP + \frac{\pi}{1-\pi} \cdot FP}$	$F = \frac{2}{P^{-1} + R^{-1}}$

Prediction Performance on Data of the Whole Wind Farm



Results and Conclusion

- The LogSparse Novelty Transformer has achieved noteworthy F1-scores across all wind turbines both on the training dataset as well as on unseen data in the test set, exhibiting its ability to distinguish between normal turbine operation and deviation from expected behavior.
- Training times per turbine vary strongly with resolution from ≈ 120 s for data in 30 min resolution to ≈ 3400 s for data in 1 min resolution
- Precision is highly dependent on the prevalence, i.e., the class ratio of the data
- Recall is improved when increasing the resolution of the underlying data
  - Recall is around 0.99 for all turbines when using data with 1 min resolution
- The automated labeling derived from the proposed method may be used to directly conduct analyses of turbine data, eliminating the need for laborious annotations based on service reports

1 Vaswani, A. *et al.* (2017) 'Attention Is All You Need', Advances in Neural Information Processing Systems 30 (NIPS 2017). Available at arXiv: 1706.03762

2 Wen, Q. *et al.* (2023) 'Transformers in Time Series: A Survey', in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)*, Macau SAR China. DOI: 10.24963/ijcai.2023/759

3 Xu, J. *et al.* (2022) 'Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy', *International Conference on Learning Representations (ICLR)*. Available at arXiv: 2110.02642

4 Siblini, W. *et al.* (2020) 'Master Your Metrics with Calibration', *International Symposium on Intelligent Data Analysis, Advances in Intelligent Data Analysis XVIII. IDA 2020. Lecture Notes in Computer Science, vol 12080. Springer, Cham.* DOI: 10.1007/978-3-030-44584-3_36