Severe Wind Event Prediction with Multivariate Physics-Informed Deep Learning

Willa Potosnak, Cristian Challu, Kin G. Olivares, James K. Miller, Artur Dubrawski
Auton Lab, School of Computer Science, Carnegie Mellon University
{wpotosna, cchallu, kdgutier, mille856, awd}@cs.cmu.edu

Abstract

Wind turbines play a crucial role in combating climate change by harnessing the force of the wind to generate clean and renewable energy. One key factor in ensuring the long-term effectiveness of wind turbines is the reduction of operating costs due to maintenance. Severe weather events, such as extreme changes in wind, can damage turbines, resulting in costly maintenance and economic losses in power production. We propose a preliminary physics-informed deep learning model to improve predictions of severe wind events and a multivariate time series extension for this work.

1 Introduction

Wind turbines play a crucial role in combating climate change by harnessing the force of the wind to generate clean and renewable energy. The energy produced by turbines is not only sustainable but also does not result in greenhouse gas emissions, which are major contributors to global warming. As such, wind turbines can help reduce reliance on other energy sources that produce greenhouse gas emissions and contribute to climate change. This clean energy impact can help mitigate the adverse effects of climate change, such as extreme weather events, rising sea levels, and disruptions to ecosystems.

One key factor in ensuring the long-term effectiveness of wind turbines is the reduction of operating costs due to maintenance. The importance of minimizing turbine need for maintenance is twofold. Firstly, it directly contributes to making wind energy more economically viable. Turbine repairs, replacements, and lost power production time due to wind turbine shutdowns (WTS) required for maintenance can offset the economic benefits of wind energy. Secondly, by keeping maintenance costs down, wind energy becomes a financially attractive option compared to energy sources that are more harmful to the environment. This competitiveness is vital for the widespread adoption of renewable energy, as it encourages businesses and governments to invest in sustainable practices.

Severe weather events, particularly gusts of wind, pose a significant threat to the structural integrity of wind turbines. The force exerted by sudden strong winds can lead to excessive strain on the turbine components, potentially causing damage and compromising their operational efficiency. In response to such adverse weather conditions, turbine manufacturers typically shut down turbines temporarily when the wind speeds exceed predetermined safety thresholds. By temporarily shutting down wind turbines during severe wind events, manufacturers aim to safeguard the turbines against damage, ensuring the longevity of the equipment and minimizing the need for costly repairs.

Using predetermined safety thresholds for WTS is a reactive approach to prevent turbines from operating under conditions that could lead to structural stress or mechanical failure. However, a reactive approach is not optimal as waiting until a severe wind gust is imminent before taking action can still result in physical turbine damage. To prevent turbine damage due to severe wind events, a proactive WTS is needed to shutdown the turbine prior to the event. However, accurate event time estimates are crucial to mitigate lost power production due to premature WTS.

Machine learning (ML) can enable proactive WTS by providing accurate predictions of severe wind events. By leveraging historical weather data, real-time atmospheric conditions, and advanced algorithms, deep learning models can be used to predict the time until a severe wind gust reaches a turbine, or time-to-event, allowing for proactive and strategic WTS.

2 Related Work

Prior work to predict severe wind events has predominantly focused on leveraging forecasting models based on historical wind speed data [7]. Training ML models with high-frequency data becomes particularly crucial in domains like wind energy, where the need for fine-resolution predictions is imperative for optimizing turbine power output. Forecasting high-frequency data, such as on the scale of seconds, presents challenges in time-series analysis due to inherent noise associated with data at such fine temporal resolutions. Rapid fluctuations and short-lived patterns characterize second-level data, making it vulnerable to unpredictable external factors. Traditional forecasting models may struggle to accurately capture these dynamics, resulting in increased forecast errors.

Labeled data provides a means to train ML models on specific events of interest, which can mitigate model overfitting to noise and artifacts. Temporal classification is one approach to predict weather events using labeled time series data [1]. However, its application in practice may be hindered by several factors, including subjective determination of a model score threshold to define event classes. An incorrectly low threshold can result in false positive predictions and premature WTS while an incorrectly high threshold can result in false negative predictions and no WTS. Additionally, the scarcity of severe weather events in data can result in label imbalances within the dataset, leading to suboptimal performance in accurately predicting rare but critical events.

Time-to-event, or survival, analysis is an alternative ML approach that leverages labeled event data to predict time until an event. Generally, time-to-event models, such as Random Survival Forests (RSF) [2] and Deep Survival Machines (DSM) [3], require the provision of specific times for the model to evaluate. This characteristic renders these model less appropriate for applications that require high-resolution timescales. Similar to temporal classification, time-to-event models can also suffer errors based on the subjective determination of the model score threshold. Furthermore, assumptions, such as the independence of survival times and censoring mechanisms, might be violated by the complex interactions and dependencies inherent in weather patterns. This highlights the need for tailored approaches for time-to-event prediction concerning weather events.

3 Proposed Methods

Preliminary Model We propose a preliminary physics-informed deep learning model that can improve time-to-event predictions of severe wind events affecting wind turbines as shown in **Fig.** 1. To the best of our knowledge, no prior work has developed a physics-based deep learning time-to-event prediction model for severe weather events, such as wind.

Incorporating fundamental principles from physics in deep learning models enables the model to develop a systematic comprehension of weather dynamics. Moreover, this incorporation can function as a regularization technique, guiding the model to produce more precise predictions that adhere to physical constraints. The behavior of wind can be elucidated through fundamental principles of physics in the form of equation $d = vt$ that relates distance d, velocity, v, and time, t. By rearranging this equation to $t = \frac{d}{v}$, we gain a valuable tool for estimating the time it takes for a weather event to reach subsequent turbines. For example, when a weather event is initially monitored at a specific turbine, the distance, d, between this monitoring point and subsequent turbines, as well as the velocity, v, of the wind recorded at the event, can be used to calculate the time it will take for the weather event to impact subsequent turbines. Furthermore, the wind velocity recorded at the initial time of a wind event may vary for subsequent turbines due wind direction and terrain, among other factors. To account for discrepancies in velocity, we extend the equation $t = \frac{d}{v}$ to incorporate a velocity scaling factor, s, that is learned using a deep learning model, $f_\theta(x)$:

$$\hat{t} = \frac{d}{v \cdot s}. \tag{1}$$

To learn s, we employ a fully-connected Multilayer Perceptron (MLP) [4; 5] as $f_\theta(x)$ with sigmoid activations. Here, $x \in R^{N \times C}$ is a feature vector, where N is the number of samples, or turbines, and C is the number of features.

Preliminary results demonstrate that a physics-based time-to-event model outperforms both purely physics and ML baselines in terms of Mean Absolute Error (MAE) (standard deviation) as shown in **Table** 1. More information regarding data, baselines, model hyperparameters, and additional results can be found in **Appendix** A.

Baselines					**Proposed**
$\hat{t} = \text{Avg. } t$	$\hat{t} = \frac{d}{v}$	$\hat{t} = \frac{d_t}{v_t}$	$\hat{t} = f_\theta(x)$	$\hat{t} = \text{DSM}(x)$	$\hat{t} = \frac{d}{v \cdot s}$
45.57 (7.03)	6.79 (6.19)	5.03 (4.45)	12.72 (6.94)	6.62 (5.43)	**2.61 (4.02)**

Table 1: Preliminary Results

Multivariate Time Series Extension The preliminary model has demonstrated improved performance over baselines while only leveraging aggregate information 1-minute prior to the first turbine event. However, wind speed exhibits time-varying characteristics. As such, we propose to extend our preliminary physics-informed time-to-event model to leverage time series data across multiple turbines to infer a time-dependent velocity representation, or scaling factor, s. By leveraging time series data across turbines, the model may be able to capture the complex dynamics of the storm system as it moves through the park and generate more accurate velocity representations.

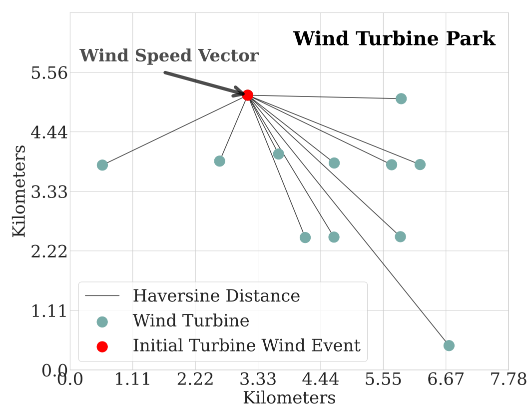

Figure 1: Principles of physics in the form of equations governing motion explained by distance, velocity, and time can inform ML approaches for time-to-event prediction.

Graph Attention Networks (GATs) [6] offer a promising deep learning approach to harness spatial-temporal information recorded across wind turbines. Given a graph, $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, individual turbines would be represented as nodes, $\mathcal{N}$, connected by edges, $\mathcal{E}$. Despite the potential of graphical models like GATs, they may necessitate substantial data for effective training where the availability of severe weather event data could be limited. Leveraging fundamental principles from physics may reduce data requirements and improve model convergence while maintaining predictive accuracy. The utility of GATs for this challenge lies in its ability to assign different weights to different neighboring nodes. As such, a physics-based graphical model can be employed to learn relevant neighboring turbines for which to inform time-dependent velocity representations used for time-to-event prediction of wind events.

4 Conclusions

We proposed a physics-informed multivariate deep learning model for predicting severe wind events. Our preliminary results demonstrate the potential effectiveness of our proposed approach for implementing proactive WTS. Implementing such a model in practice could reduce turbine operating costs due to power production losses, maintenance, and premature WTS, which can help support wind turbines as a financially attractive clean and renewable energy source.

References

[1] Shubhi Harbola and Volker Coors. Deep learning model for wind forecasting: Classification analyses for temporal meteorological data. pages 211–224, 2022.

[2] Hemant Ishwaran, Udaya B. Kogalur, Eugene H. Blackstone, and Michael S. Lauer. Random survival forests. *The Annals of Applied Statistics*, 2(3), September 2008.

[3] Chirag Nagpal, Xinyu Li, and Artur Dubrawski. Deep survival machines: Fully parametric survival regression and representation learning for censored data with competing risks. *IEEE Journal of Biomedical and Health Informatics*, 25(8):3163–3175, 2021.

[4] Vinod Nair and Jeoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *ICML-23*, 2010.

[5] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386—-408, 1958.

[6] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018.

[7] Bo Yang, Linen Zhong, Jingbo Wang, Hongchun Shu, Xiaoshun Zhang, Tao Yu, and Sun Liming. State-of-the-art one-stop handbook on wind forecasting technologies: An overview of classifications, methodologies, and analysis. *Journal of Cleaner Production*, 283, 2021.

A Appendix

A.1 Data

Preliminary work was performed on a wind turbine dataset consisting of 50 turbines with 1 month of data sampled at 5 Hz. The data was downsampled to 0.1 Hz (10 second timestamps) to facilitate processing data with ML models.

A severe wind event was defined as the first timestamp with windspeed greater than 20 meters per second. Following events in the same 24-hour period were considered part of the same wind event. Several wind events were observed for each turbine within a 1-month period as shown in **Figure** 2. Data from the first 3 wind events across turbines prior to day 22 were used to train the model and the wind event on day 22 was used for the final evaluation to maintain temporal ordering and prevent data leakage.

Features used to train the ML models consist of meteorological data including, ambient temperature, wind speed and wind direction, recorded at the initial event time as well as distance and angle between turbines.

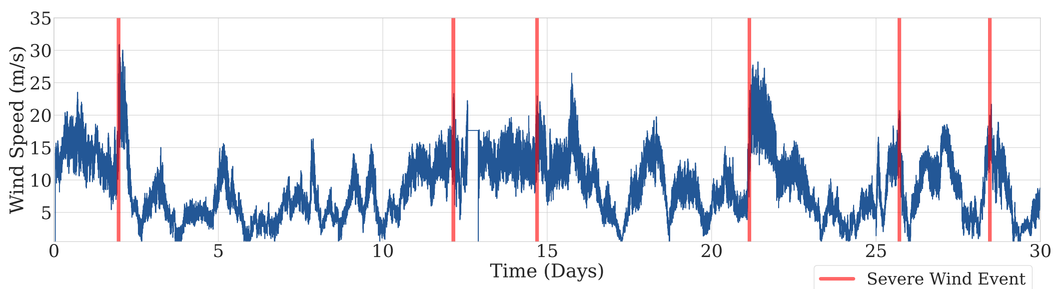

Figure 2: Several wind events were observed for each turbine within a 1-month period.

A.2 Baselines

We evaluate 5 baseline models including 2 physics-based models and 2 ML models:

Average Heuristic

- Avg. t: The average time-to-event for events 1-3 was used to estimate $\hat{t}$.

Physics-based Models

- $\hat{t} = \frac{d}{v}$: Time-to-event, $\hat{t}$ was predicted based on the velocity, v, recorded at the turbine that experienced an event prior to other turbines and the Haversine distance, d between the event turbine and subsequent turbines in the park.
- $\hat{t} = \frac{d_t}{v_t}$: This physics-based model is designed to emulate a real-time monitoring system that continually updates predictions as soon as a wind event impacts a new turbine. Time-to-event, $\hat{t}$, was predicted based on the velocity, v, recorded at the turbine that experienced an event prior to other turbines and the Haversine distance, d between the event turbine and subsequent turbines in the park. For each subsequent turbine to experience the event, the velocity and distance from the last turbine was updated and used to generate new predictions for turbines yet to experience the event or be shutdown by prior time-to-event predictions.

ML Models

- $\hat{t} = f_\theta(x)$: A fully-connected Multilayer Perceptron (MLP) [4; 5] with ReLU activations was used to predict Time-to-event, $\hat{t}$, given covariates, x.

- Deep Survival Machines (DSM) [3]: DSM is a survival analysis model that estimates the conditional distribution $P(T > t|X = x)$ as a mixture over k parametric distributions using representation of the individual covariates, x, learned with an MLP. The model outputs the likelihood that an event has not yet occurred at a specified time, t, given covariates, x. We take time where $P(T > t|X = x) = 0.5$ as the estimated time-to-event based on a decision threshold of 0.5.

A.3 Hyperparameters

Hyperparameter selection was performed using using 3-fold cross validation blocked by grouped wind events across turbines. Final hyperparameters were selected that minimized mean squared error across held-out folds. A table outlining hyperparameter options for the $\hat{t} = f_\theta(x)$ baseline and proposed preliminary model are listed below:

Hyperparameter	**Considered Values**
Learning rate	{1e-3, 1e-2}
Number of layers	{2, 3}
Hidden size	{64, 128}
Batch size	32
Training steps	5000

Table 2: Common hyperparameter search space

A.4 Additional Results

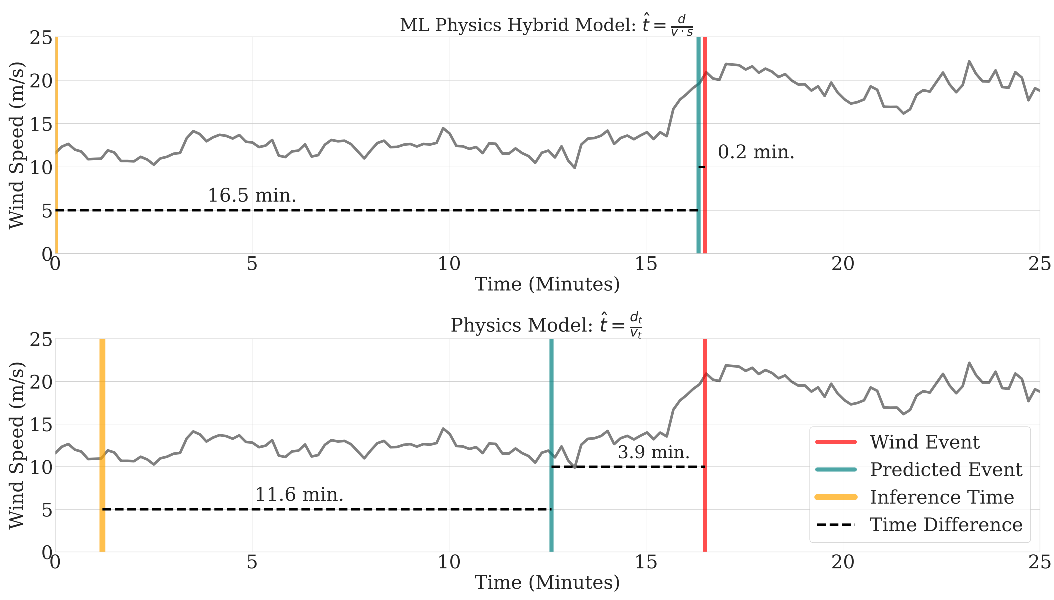

Figure 3: Comparing both physics and ML physics hybrid models for turbine example #1, the ML Physics Hybrid model (top), $\hat{t} = \frac{d}{v \cdot s}$, provides a more accurate event prediction and a longer lead time from model inference to predicted event.

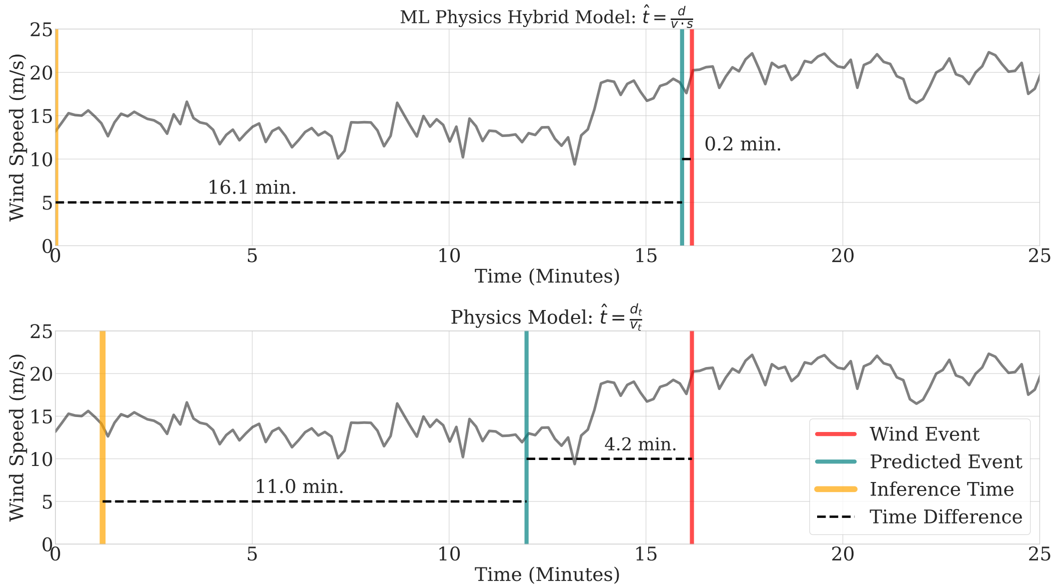

Figure 4: Comparing both physics and ML physics hybrid models for turbine example #2, the ML Physics Hybrid model (top), $\hat{t} = \frac{d}{v \cdot s}$, provides a more accurate event prediction and a longer lead time from model inference to predicted event.

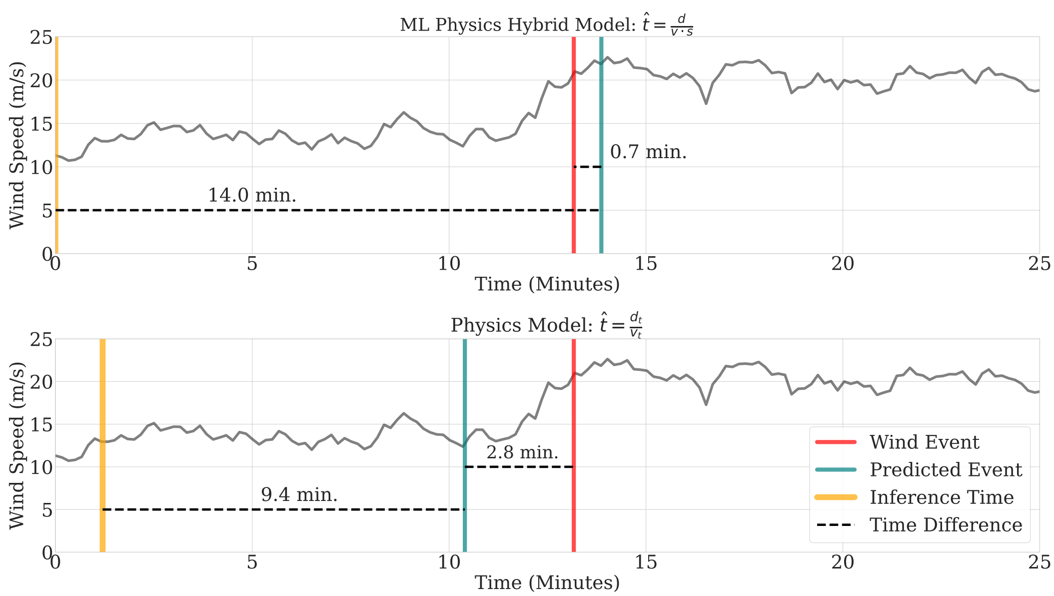

Figure 5: Comparing both physics and ML physics hybrid models for turbine example #3, the ML Physics Hybrid model (top), $\hat{t} = \frac{d}{v \cdot s}$, provides a more accurate event prediction and a longer lead time from model inference to predicted event.