IN-VENTO 2026
International Conference of the Italian Association for Wind Engineering
Bari, September 13th – 16th, 2026

Wind profiles for real-time digital twins

Jingyu Wei[a,b,d], Yasutaka Narazaki[b,c], Giuseppe Quaranta[d], Qingshan Yang[e], Christos T. Georgakis[f], Cristoforo Demartino[g]

[a]*College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, 310058 Zhejiang, PR China*
[b]*Zhejiang University–University of Illinois Urbana-Champaign Institute, Zhejiang University, Haining 314400 Zhejiang, PR China*
[c]*Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*
[d]*Department of Structural Engineering and Geotechnics, Sapienza University of Rome, Rome, Italy*
[e]*School of Civil Engineering, Chongqing University, Chongqing 400044, China*
[f]*Department of Engineering, Aarhus University, Aarhus, Denmark*
[g]*Department of Architecture, Roma Tre University, Rome 00153, Italy*

SUMMARY:
Vertical wind profiles are important for characterizing height-dependent wind conditions in wind-energy systems, wind-sensitive structures, and real-time digital-twin applications. This study develops a data-driven approach for estimating the current wind-profile shape and forecasting its short-term evolution. Each measured profile is normalized by the wind speed at 10 m and described by three coefficients of a cubic polynomial, reducing the full-profile prediction task to a low-dimensional regression problem. XGBoost is used for nowcasting the current profile from recent ground-level meteorological observations, while an LSTM network forecasts the following six time steps, corresponding to one hour, using meteorological variables and previously observed profile coefficients. The approach is evaluated using 114,274 complete 10-minute wind profiles measured at seven heights over 843 days at the Cabauw experimental site. The cubic polynomial offers a suitable balance between fitting accuracy and model complexity. Among the tested configurations, the best results are obtained with a one-day input history for nowcasting and a half-day history for forecasting. Both models show higher accuracy when the wind speed at 10 m exceeds 3.5 m/s. The predicted profiles provide compact, time-varying wind information for wind-engineering and digital-twin applications.

Keywords: wind profiles, nowcasting, forecasting, machine learning, wind lidar, digital twins

1. INTRODUCTION

Height-resolved wind information is required to quantify wind loads on tall buildings, bridges, towers, and wind turbines, and to provide time-varying environmental input for operational digital twins. Continuous wind-profile measurements from lidar systems or instrumented towers are not always available because of instrument limitations, data gaps, and operating costs. This study represents nondimensional wind profiles using cubic-polynomial coefficients, applies XGBoost to estimate the current profile from ground-level meteorological observations, and uses an LSTM network to forecast short-term profile changes from recent meteorological variables and historical profile information, as summarized in Fig. 1 (Wei et al., 2025).

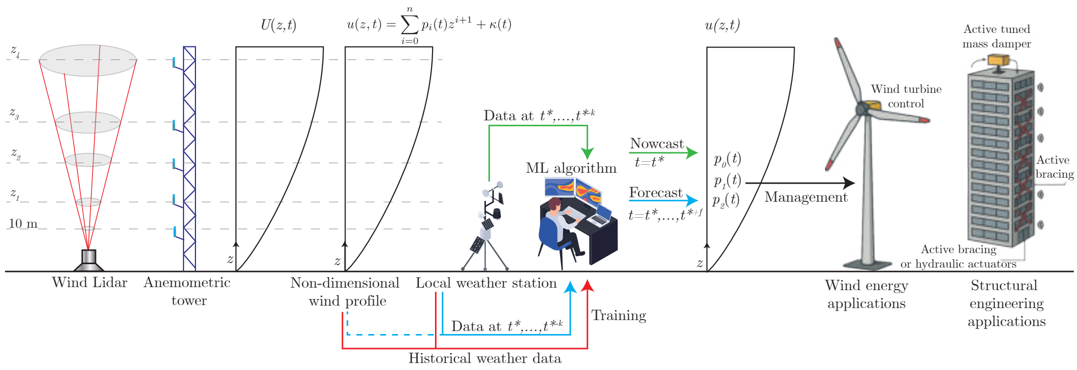


Figure 1. Proposed framework for wind-profile nowcasting and forecasting using machine-learning techniques.

2. METHODOLOGY

The measured mean wind speed $U(z,t)$ is first converted into a non-dimensional wind profile by normalizing it with the wind speed at the reference height $z_{\text{ref}} = 10$ m:

$$u(z,t) = \frac{U(z,t)}{U(z_{\text{ref}},t)}. \tag{1}$$

The resulting non-dimensional profile is represented using the following polynomial form:

$$u(z,t) = \sum_{i=0}^{n} p_i(t) z^{i+1} + \kappa(t), \qquad \kappa(t) = 1 - \sum_{i=0}^{n} p_i(t) z_{\text{ref}}^{i+1}, \tag{2}$$

where n controls the polynomial order, $p_i(t)$ are the time-dependent profile coefficients, and $\kappa(t)$ is determined by the constraint $u(z_{\text{ref}},t) = 1$. The coefficients are identified for each measured profile through least-squares fitting. Polynomial representations from linear to fifth degree were examined. The model with $n = 2$, corresponding to a cubic polynomial with three coefficients p_0, p_1, and p_2, was selected because it substantially reduced the fitting error relative to the lower-order models, whereas the quartic and quintic models provided only limited additional improvement. The prediction of a complete vertical wind profile is therefore reduced to the prediction of three time-dependent coefficients.

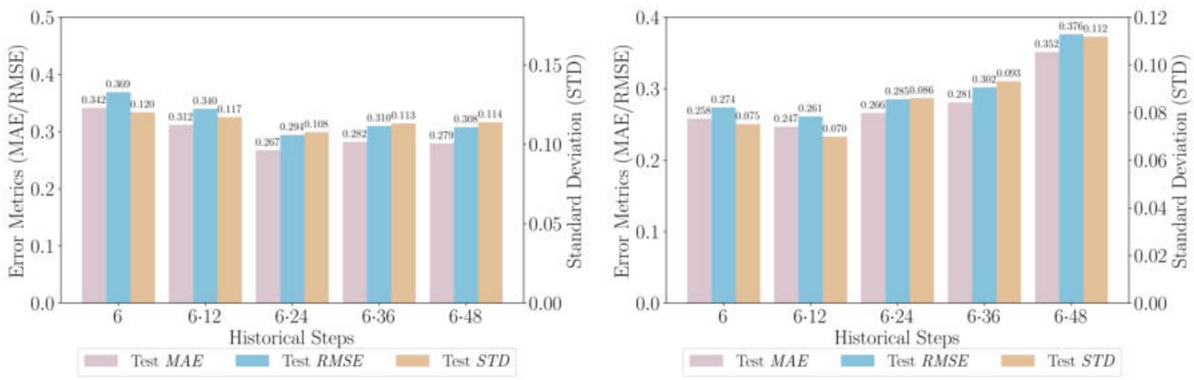
Two prediction tasks are considered. For nowcasting, XGBoost estimates the current profile coefficients from recent ground-level meteorological observations, including air temperature, atmospheric pressure, 10 m wind speed and direction, month, hour, and the associated north–south and east–west wind components. Historical wind-profile measurements are used to define the training targets, but they are not needed as inputs when the trained model is used operationally. XGBoost is selected for its efficient and robust regression performance with structured meteorological data (Chen and Guestrin, 2016). For forecasting, an LSTM network uses recent meteorological variables and previously observed profile coefficients to predict wind profiles over six time steps, corresponding to a one-hour horizon. The network captures changes in wind-profile shape over time and relates them to recent atmospheric conditions (Hochreiter and Schmidhuber, 1997).

Different input-history lengths are examined for both prediction tasks. For each candidate history window, Bayesian optimization with five-fold cross-validation is used to tune the model hyperparameters on the training set. The resulting models are evaluated on the test set using the mean absolute error (MAE), root mean square error (RMSE), and standard deviation (STD) of the prediction errors.

3. APPLICATION AND RESULTS

The proposed framework is evaluated using measurements from the Cabauw Experimental Site for Atmospheric Research in the Netherlands (Knoop and Jong, 2020). Wind profiles were measured with a ZephIR 300M wind lidar from February 2018 to June 2020. The dataset contains 10-minute mean wind speeds at seven heights: 10, 19, 38, 79, 139, 199, and 251 m. Profiles with a missing measurement at any height were excluded, leaving 114,274 complete profiles, or approximately 94% of the original records.

The effect of input-history length is shown in Fig. 2. For nowcasting, the lowest test-set errors were obtained when the training samples were shuffled and the meteorological input history covered $6 \cdot 24$ time steps, or 24 h. This configuration achieved an MAE of 0.267 and an RMSE of 0.294. The result suggests that XGBoost primarily learns the relationship between meteorological conditions and the current profile shape, rather than depending on the strict sequence of adjacent samples. For forecasting, the temporal order of the data was retained. The lowest errors were obtained with an input history of $6 \cdot 12$ time steps, or 12 h, giving an MAE of 0.247 and an RMSE of 0.261. Extending the input window did not improve the six-step forecasts, which indicates that recent meteorological conditions and profile coefficients contain the information most relevant to the one-hour prediction horizon.

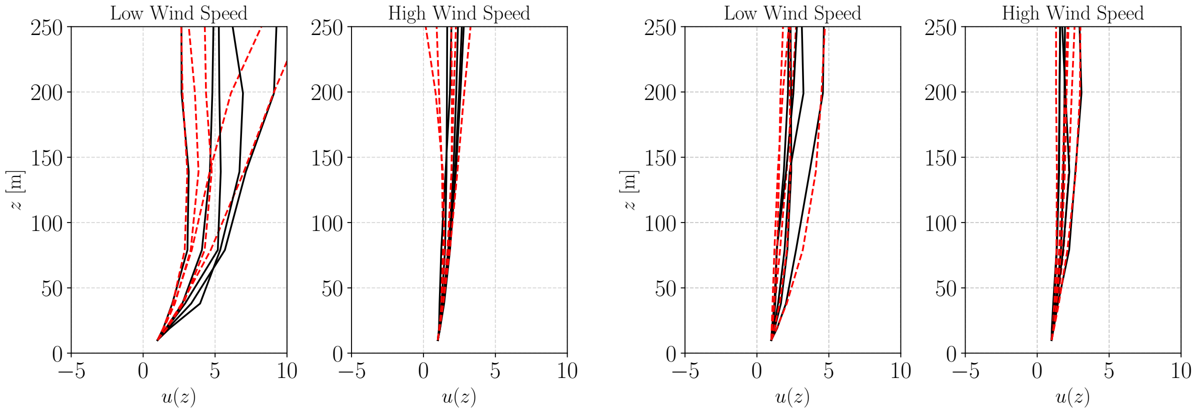


(a) Nowcasting (b) Forecasting

Figure 2. Effect of the input-history length on test-set performance for (a) XGBoost-based nowcasting and (b) LSTM-based forecasting. Each time step corresponds to 10 min. The lowest errors are obtained using a 24-h history for nowcasting and a 12-h history for forecasting.

To assess prediction performance under different inflow conditions, the test profiles were grouped into eight 45-degree wind-direction sectors and separated into low- and high-wind-speed cases using a threshold of $U_{10} = 3.5$ m/s. Representative results from this sector-wise analysis are presented in Fig. 3. The nowcasting results show how the current nondimensional profile is reconstructed from ground-level meteorological observations. The forecasting results show how future profile shapes are predicted from recent meteorological variables and previously observed polynomial coefficients.

The predicted profiles agree well with the measurements for both nowcasting and forecasting, showing that the polynomial representation captures the main vertical features of the wind field. The agreement is especially good when the wind speed at 10 m exceeds 3.5 m/s. At these higher wind speeds, the profiles are generally more coherent, the lidar data are of better quality, and the fitted polynomial coefficients are more stable, leading to more accurate profile predictions. Overall, the method provides reliable wind-profile estimates using routinely available



(a) Nowcasting (b) Forecasting

Figure 3. Representative nowcasting and forecasting results under low- and high-wind conditions. Black solid lines denote measured non-dimensional wind profiles, while red dashed lines denote predicted profiles.

meteorological observations and recent profile data.

4. CONCLUSIONS

This study presents a compact framework for wind-profile nowcasting and short-term forecasting using lidar and meteorological data. Nondimensional wind profiles are described by three cubic-polynomial coefficients. XGBoost is used for nowcasting, while an LSTM network predicts six future time steps. Of the configurations examined, a 24-h input history produced the best nowcasting results, and a 12-h history produced the best forecasting results. Both models had lower prediction errors under higher-wind conditions. The predicted profiles provide time-varying wind information for wind-energy, structural-assessment, and digital-twin applications.

ACKNOWLEDGEMENTS

This work was supported by the Italian Ministry of University and Research (MUR) under the Fondo Italiano per la Scienza 2024–2025, Bando FIS 3, project “FAIR – Forecasting and Adaptive Infrastructure Resilience”, grant no. FIS-2024-02996, Consolidator Grant, ERC sector PE8 – Products and Processes Engineering.

REFERENCES

Chen, T. and C. Guestrin (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

Hochreiter, S. and J. Schmidhuber (1997). Long short-term memory. Neural Computation 9, 1735–1780.

Knoop, S. and A. de Jong (2020). *CESAR wind lidar data at Cabauw*. Cabauw Experimental Site for Atmospheric Research dataset.

Wei, J., Y. Narazaki, G. Quaranta, Q. Yang, C. T. Georgakis, and C. Demartino (2025). Wind profile nowcasting and forecasting using machine learning. Journal of Wind Engineering and Industrial Aerodynamics 266, 106162.