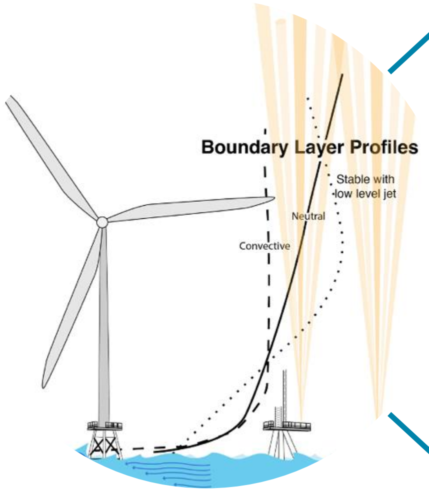
Comparison of wind speed profiles from a mesoscale model and a deep learning-based approach

Harish Baki
Atmospheric Sciences Research Center, University at Albany, NY, United States

Maximilian Pierzyna
Department of Geoscience and Remote Sensing
Delft University of Technology, The Netherlands

Sukanta Basu
Atmospheric Sciences Research Center,
Department of Environmental and Sustainable Engineering,
University at Albany

Objective



Tall towers, Lidars, Sodars

- High installation, procurement cost
- High O&M cost

Extrapolation of surface measurements using empirical equations

- Logarithmic law of wall
- Monin-Obukhov similarity theory
- Power law
- Limited validity and need additional variables

Numerical weather prediction models

- Mesoscale models
- Computationally demanding
- Dependence on resolution and model physics
- Sensitive to IC & BC

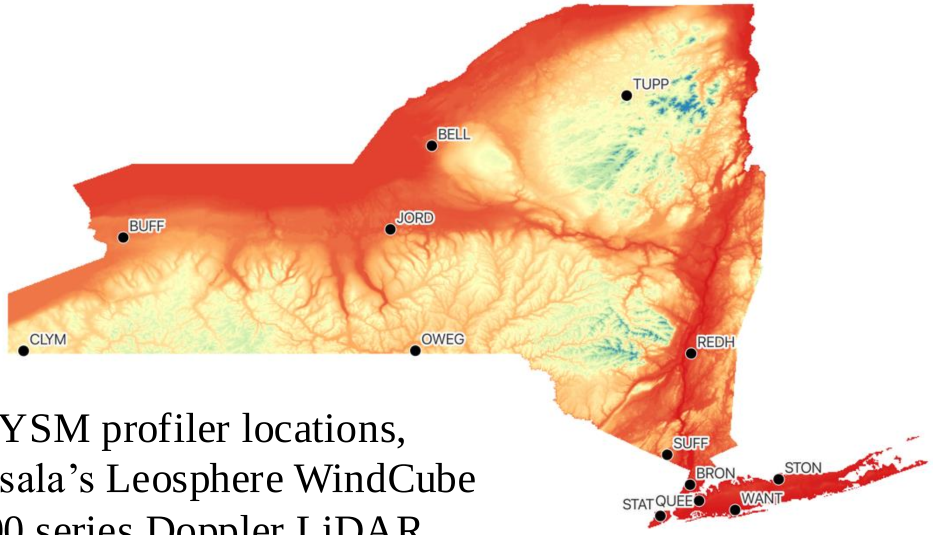


Fig 1. NYSM profiler locations, with Vaisala's Leosphere WindCube WLS-100 series Doppler LiDAR.

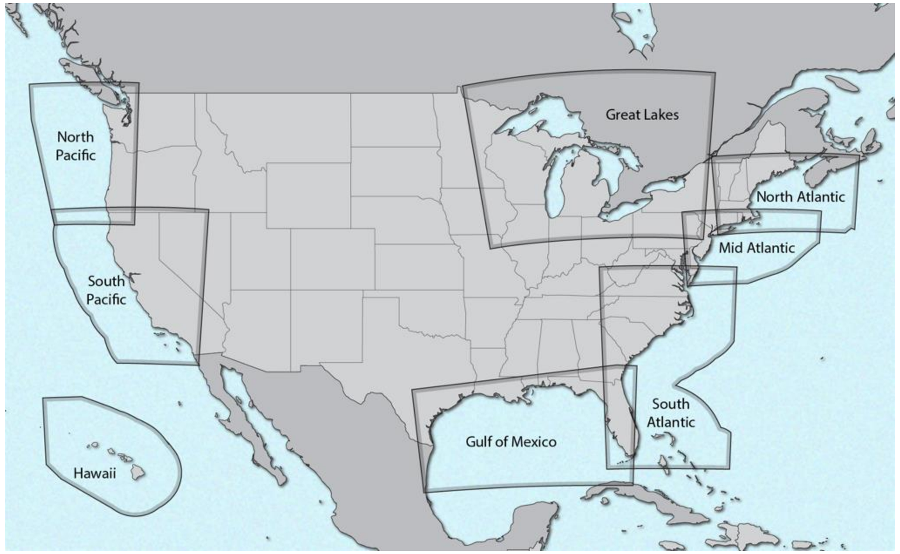


Fig 2. Domains of the 2023 National Offshore Wind dataset (NOW23). Source: National Renewable Energy Laboratory. (2020). 2023 National Offshore Wind data set (NOW-23) [data set]. Retrieved from https://dx.doi.org/10.25984/1821404.

University at Albany, State University of New York

NYS Mesonet

Deep learning-based alternative approach

Environmental Data Science (0000), xx: 1–10
doi:10.1017/xxxx

CAMBRIDGE UNIVERSITY PRESS

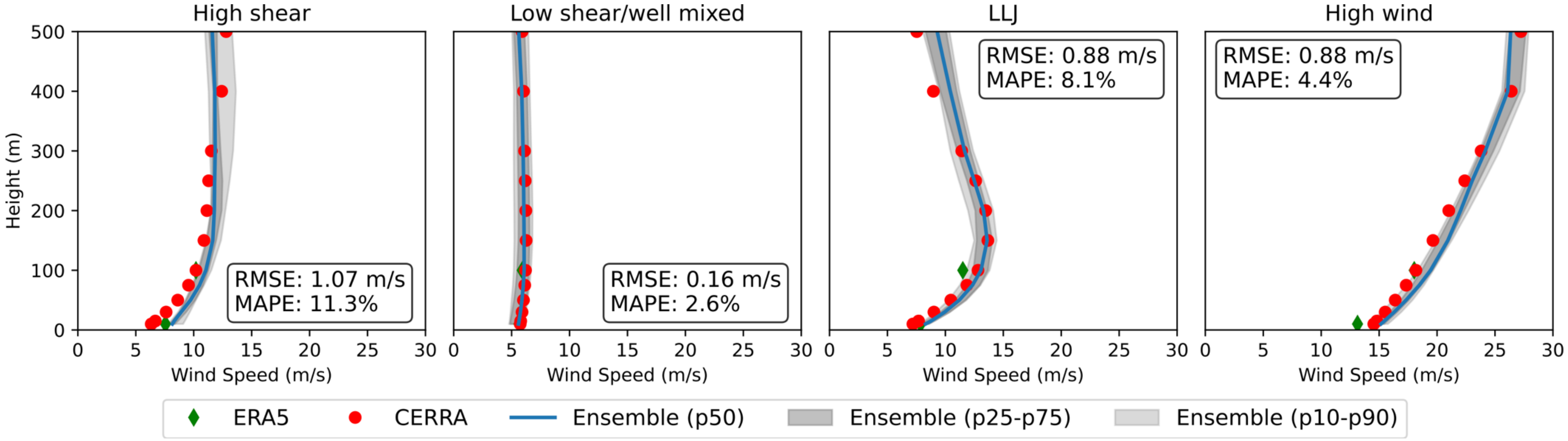
APPLICATION PAPER

Estimating high-resolution profiles of wind speeds from a global reanalysis dataset using TabNet

Harish Baki[1]* and Sukanta Basu[2,3]

[1]Faculty of Civil Engineering and Geosciences, TU Delft, Delft, The Netherlands
[2]Atmospheric Sciences Research Center, University at Albany, Albany, New York, United States
[3]Department of Environmental and Sustainable Engineering, University at Albany, Albany, New York, United States
*Corresponding author. Email: h.baki@tudelft.nl

- 34 variables from **ERA5** as **inputs** and wind speed at 12 vertical levels from **CERRA** reanalysis as **targets,** at FINO 1.
- **Measure-correlate-predict (MCP)** approach: trained for one year (2000) and predict for another (2001)
- An **attention-based** sequential deep learning model – **TabNet**
- How do we make the methodology generic across different wind profile datasets?



UNIVERSITY AT ALBANY STATE UNIVERSITY OF NEW YORK

NYS Mesonet

Chebyshev polynomial approximation

- Wind speed (u) as a function of height (z)

$$u(z) = \sum_{n=0}^{\infty} C_n T_n(z)$$

- Chebyshev polynomials

$$T_0(z) = 1$$
$$T_1(z) = z$$
$$T_{n+1}(z) = 2zT_n(z) - T_{n-1}(z)$$

- 4th order Chebyshev polynomials

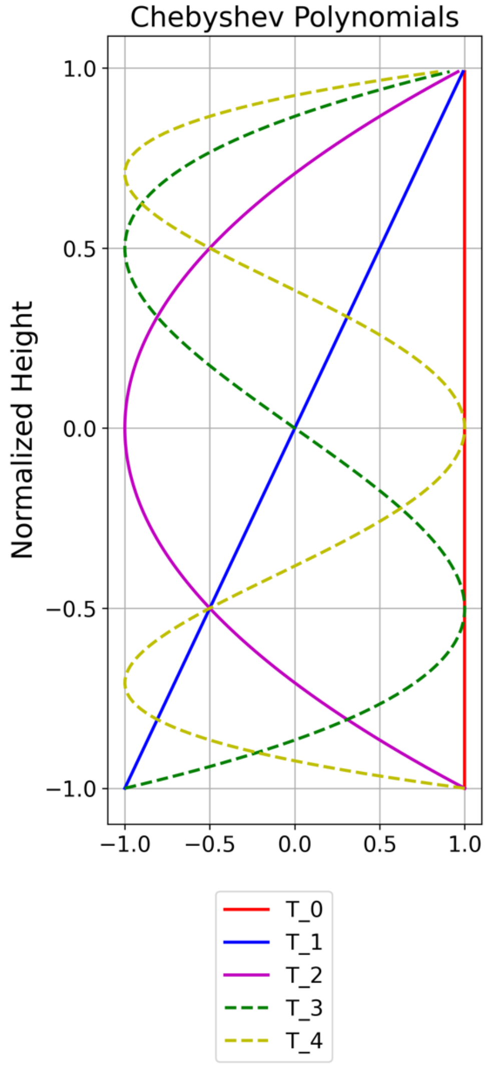


Fig 6. Chebyshev polynomials.

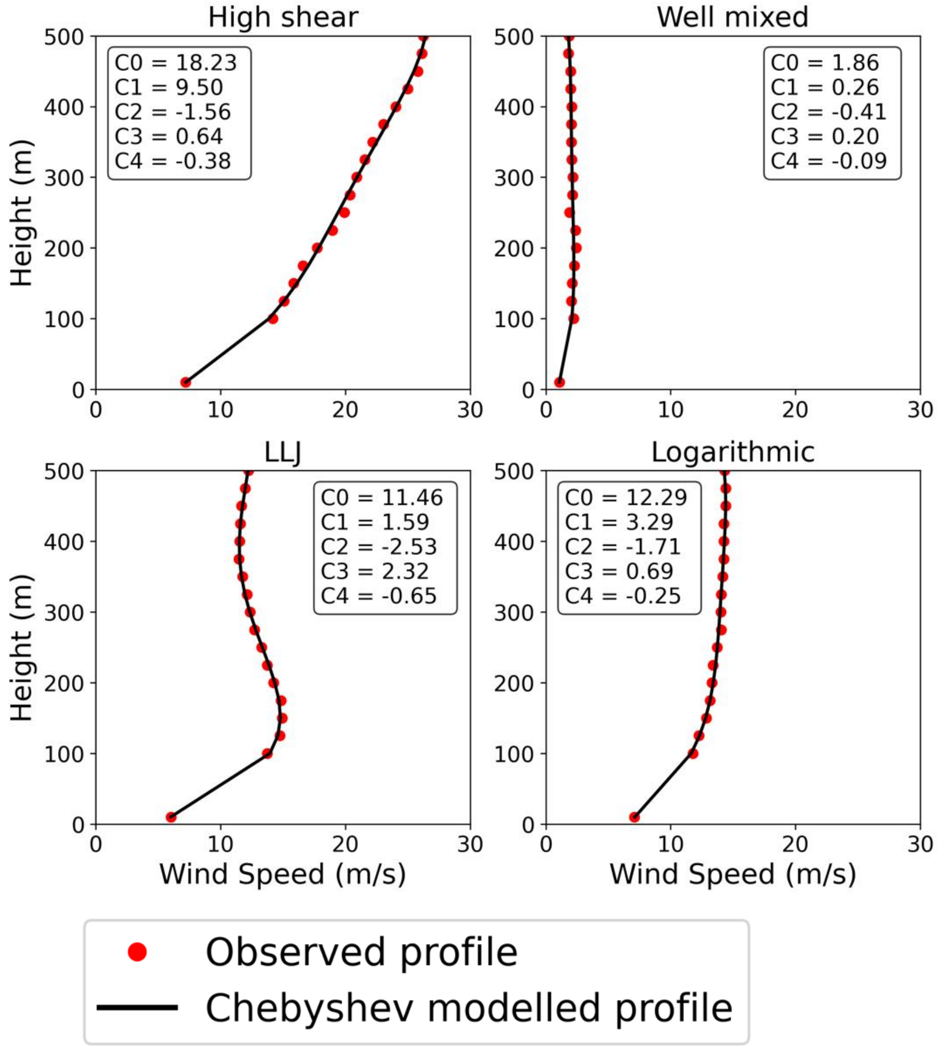
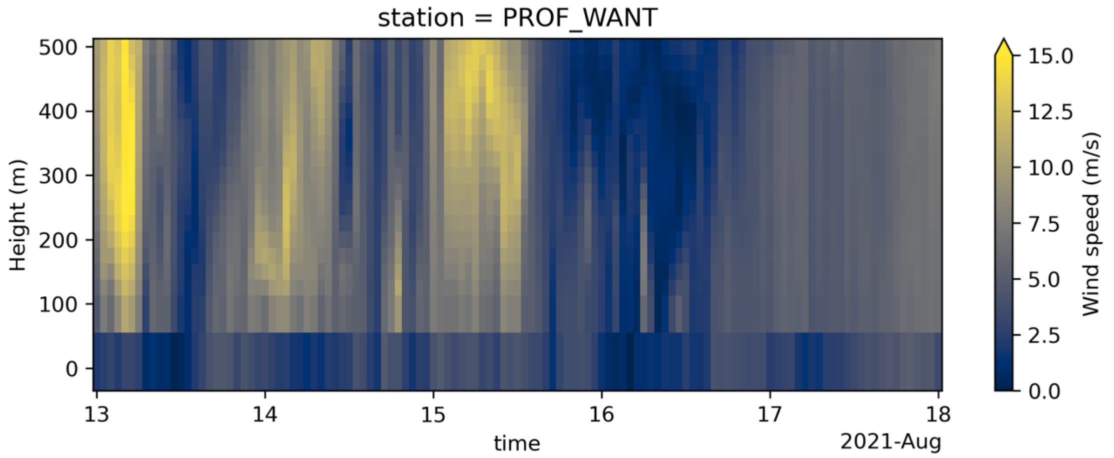


Fig 7. Chebyshev approximation of NYSM profiler wind profiles for four wind regimes.

Datasets

Table 1. Description of the ERA5 meteorological variables as inputs.

Type	Variable	Equation	Description	Units
Derived	$\mathbf{W}_{10}$	$\sqrt{U_{10}^2 + V_{10}^2}$	Wind speed at 10 m a.g.l. computed from zonal and meridional components	m s-1
Derived	$\mathbf{W}_{100}$	$\sqrt{U_{100}^2 + V_{100}^2}$	Wind speed at 100 m a.g.l. computed from zonal and meridional components	m s-1
Derived	α	$\frac{\log \mathbf{W}_{100}/\mathbf{W}_{10}}{\log(100/10)}$	Power-law exponent of wind profile within 10–100 m a.g.l.	–
Derived	$\mathbf{W}_{975}$	$\sqrt{U_{975}^2 + V_{975}^2}$	Wind speed at 975hPa computed from zonal and meridional components	m s-1
Derived	$\mathbf{W}_{950}$	$\sqrt{U_{950}^2 + V_{950}^2}$	Wind speed at 950hPa computed from zonal and meridional components	m s-1
Raw	$\mathbf{u}_*$		Friction velocity	m s-1
Raw	$\mathbf{W}^i_{p10}$		Instantaneous wind gust at 10 m a.g.l.	m s-1
Raw	T_2		Air temperature at 2 m a.g.l.	K
Raw	T_0		Skin temperature	K
Raw	T_s		Upper-level soil temperature	K
Raw	T_{d2}		Dew point temperature at 2 m a.g.l.	K
Raw	P_0		Mean sea level pressure	Pa
Raw	$\mathbf{H}$		Boundary layer height	m
Raw	h_{cb}		Cloud base height	m
Raw	H_S		Instantaneous surface sensible heat flux	W m-2
Raw	H_L		Instantaneous moisture flux	Kg m-2 s-1
Raw	TCC		Total cloud cover	–
Raw	LCC		Low-level cloud cover	–
Raw	CAPE		Convective available potential energy	J kg-1
Raw	CIN		Convective inhibition	J kg-1
Raw	$\bar{\epsilon}$		Energy dissipation rate in boundary layer	J m-2
Raw	T_{975}		Air temperature at 975hPa	K
Raw	T_{950}		Air temperature at 950hPa	K



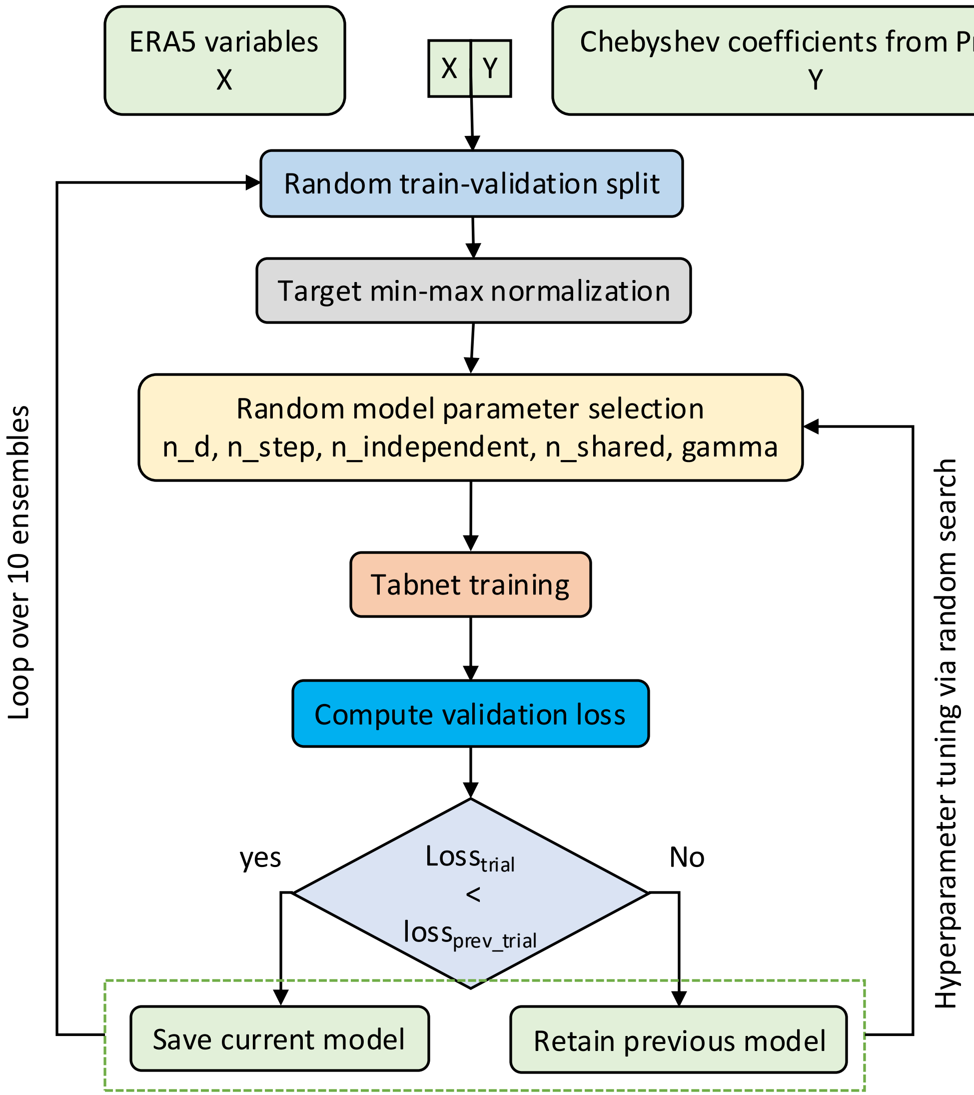
Vertical profiles of wind speed from NYSM profilers are used for ML training.

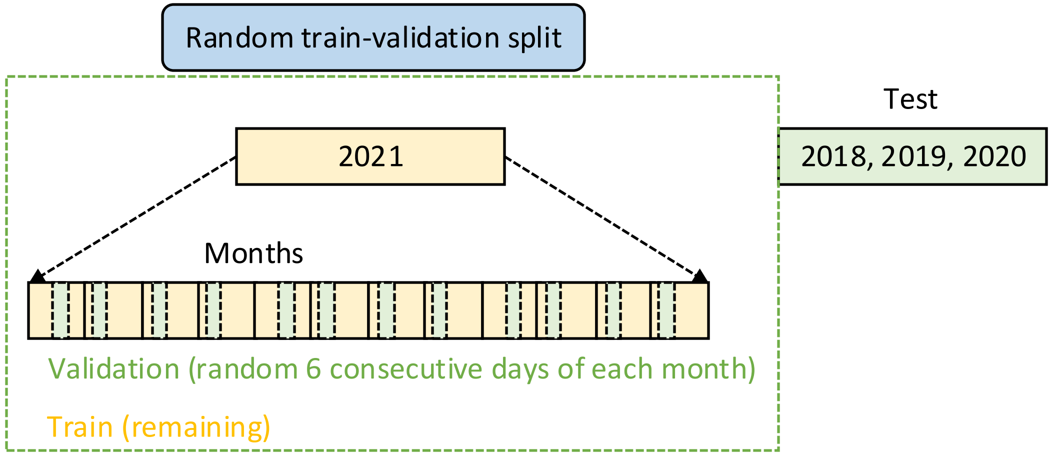
An intersect of 13 profiler locations between NYSM and NOW23 datasets are identified. The NOW23 wind profiles are used for comparison.

UNIVERSITY AT ALBANY
STATE UNIVERSITY OF NEW YORK

NYS Mesonet

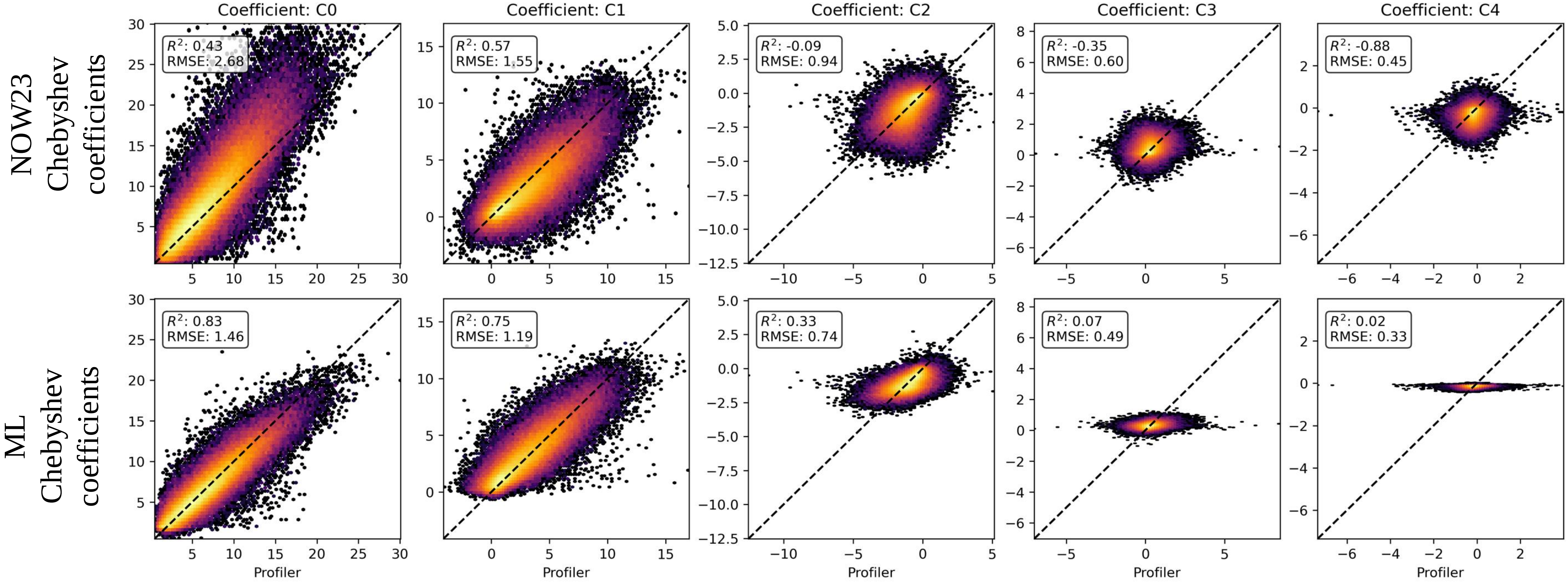
DL training procedure





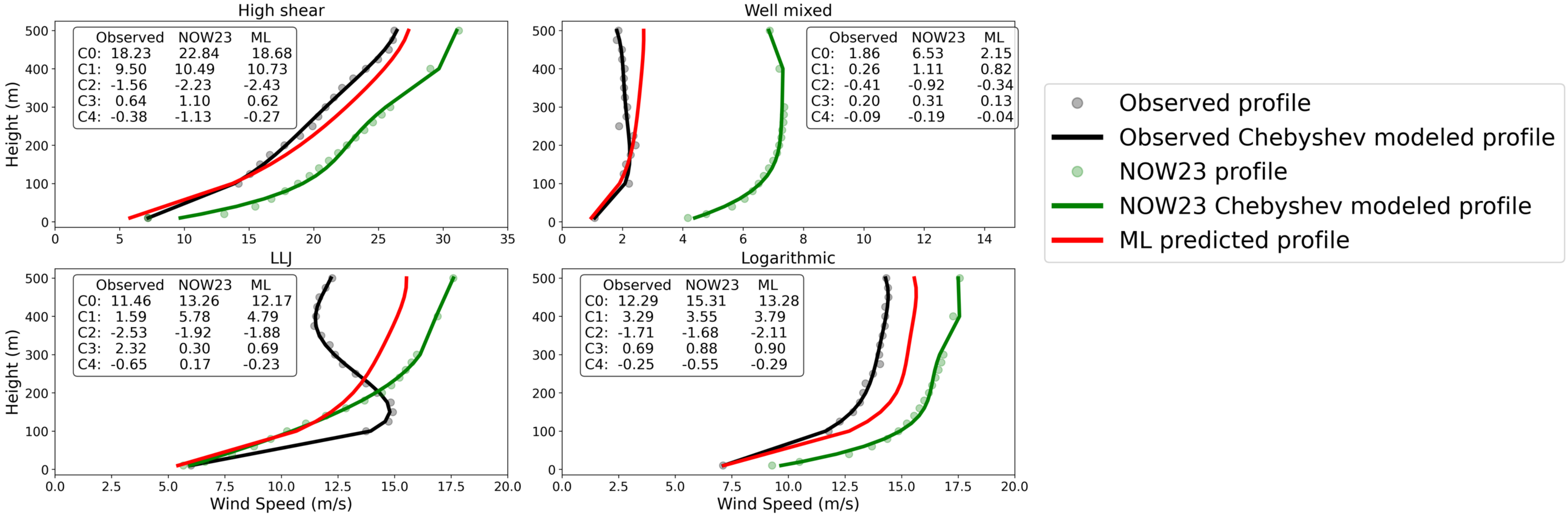
Among 13 profilers, 9 are used in training/validation and 4 are used in testing.

Comparison of Chebyshev coefficients



Bivariate histograms of Chebyshev coefficients between profiler, NOW23, and ML predictions, from four test stations and three test years.

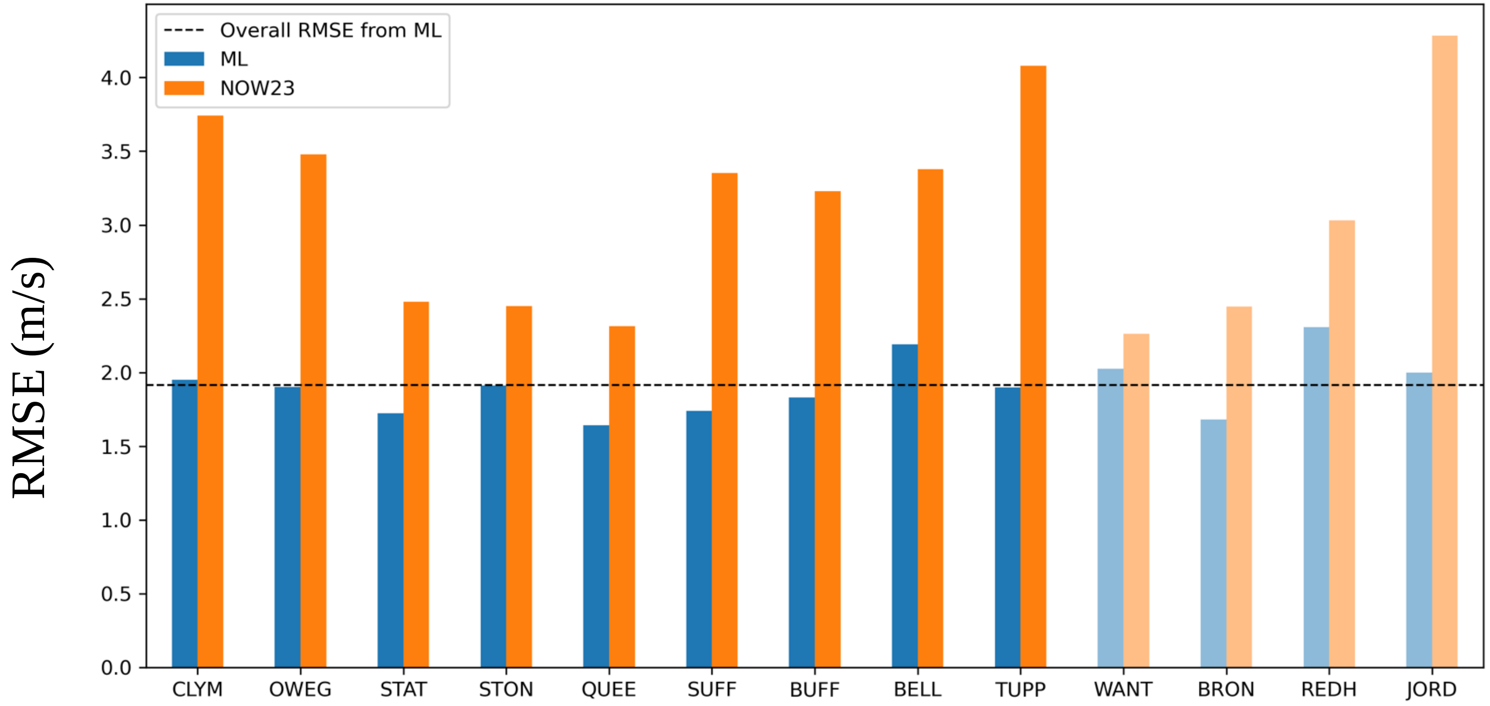
Comparison of wind speed profiles



Comparison of wind speed profiles from profiler (actual and Chebyshev modeled), NOW23 (actual and Chebyshev modeled), and ML predictions (Chebyshev modeled), across four wind regimes.

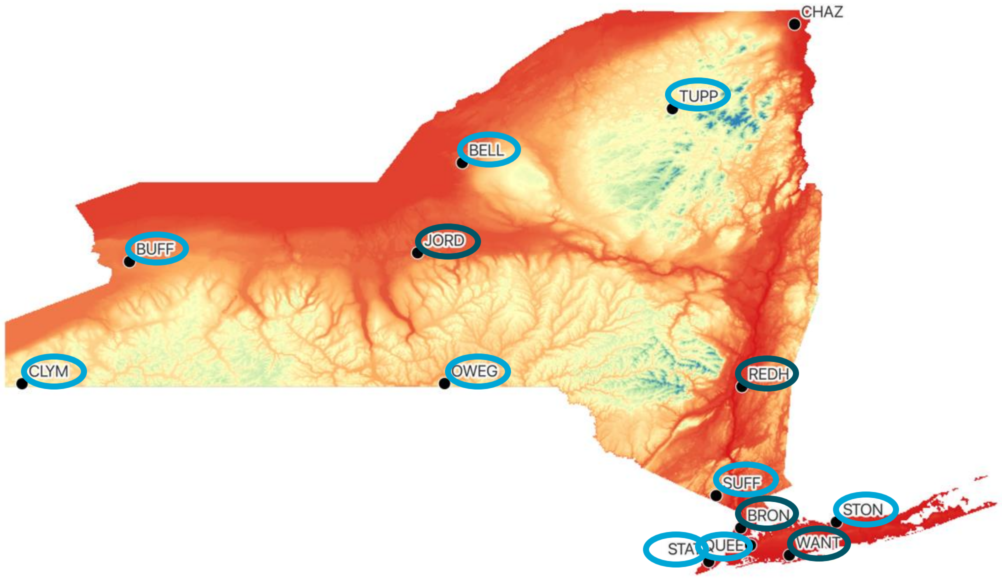
University at Albany, State University of New York — NYS Mesonet

Generalization



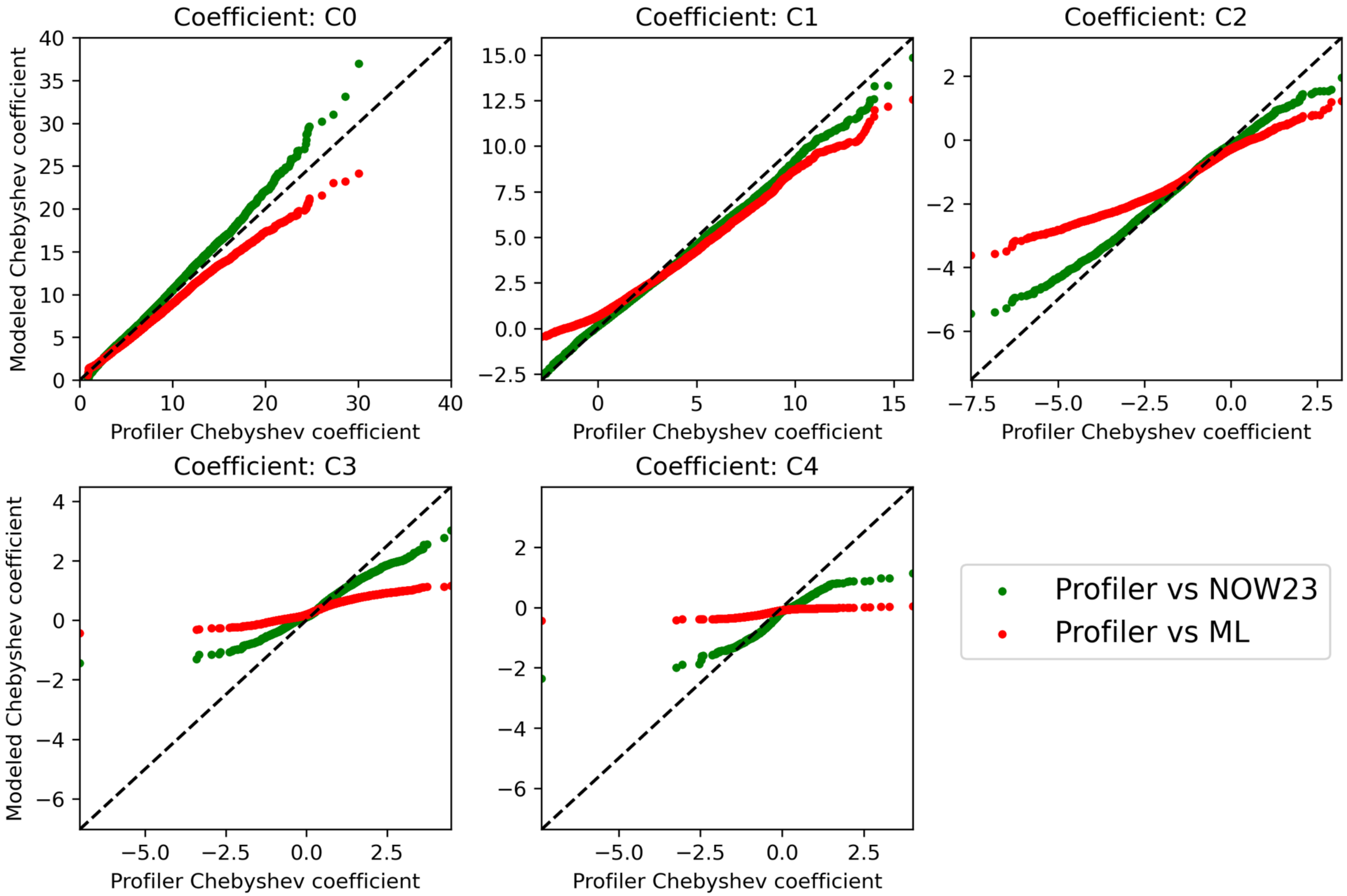
Temporal generalization

Spatio-temporal generalization



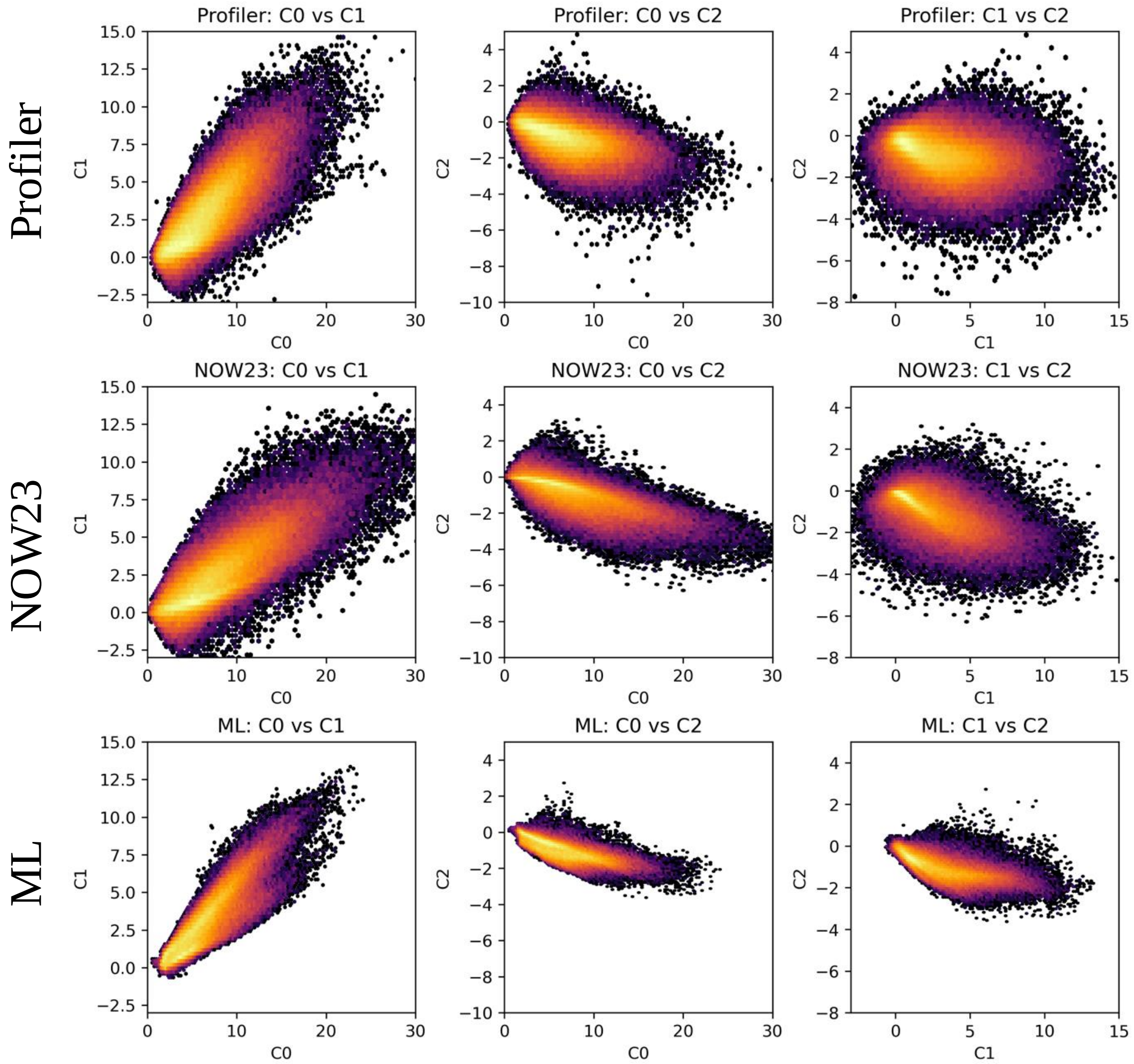
Comparison of wind profile RMSE scores from NOW23 and ML predictions, with respect to the NYSM profilers, across the stations.

Quantile-quantile comparison of Chebyshev coefficients



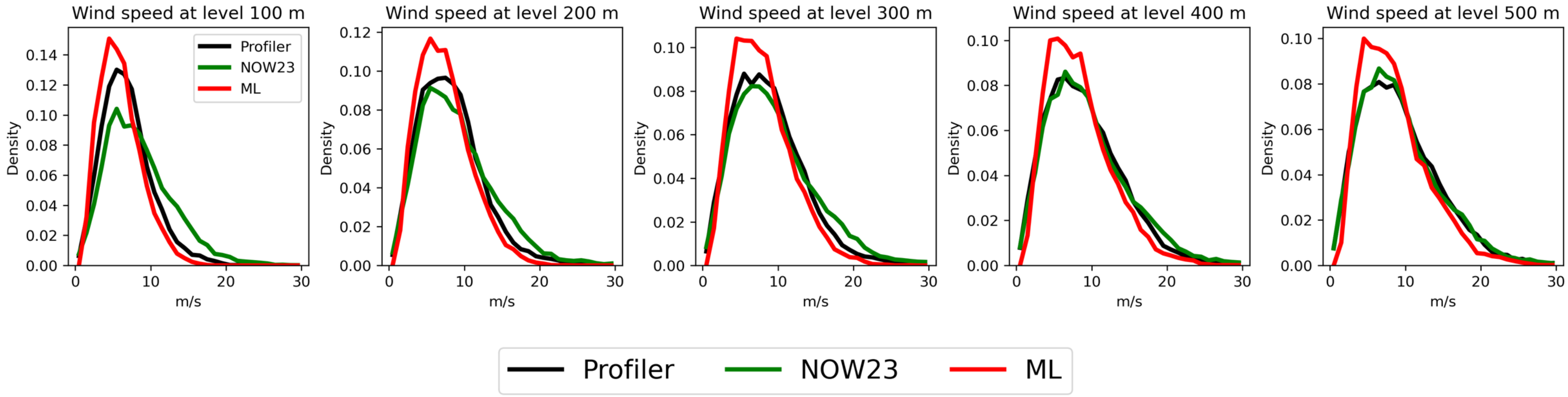
Quantile-quantile comparison of Chebyshev coefficients between profiler, NOW23, and ML predictions, at station PROF_WANT, for three test years.

Intercomparison of Chebyshev coefficients

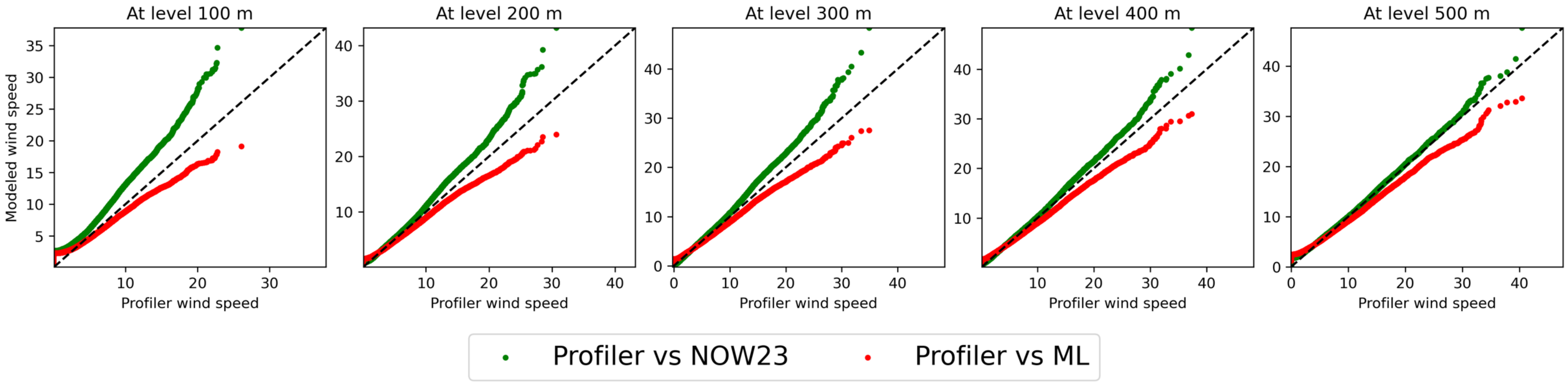


Bivariate histograms of Chebyshev coefficients within profiler, NOW23, and ML predictions, from 13 stations and three test years.

Comparison of wind speed distribution and quantiles



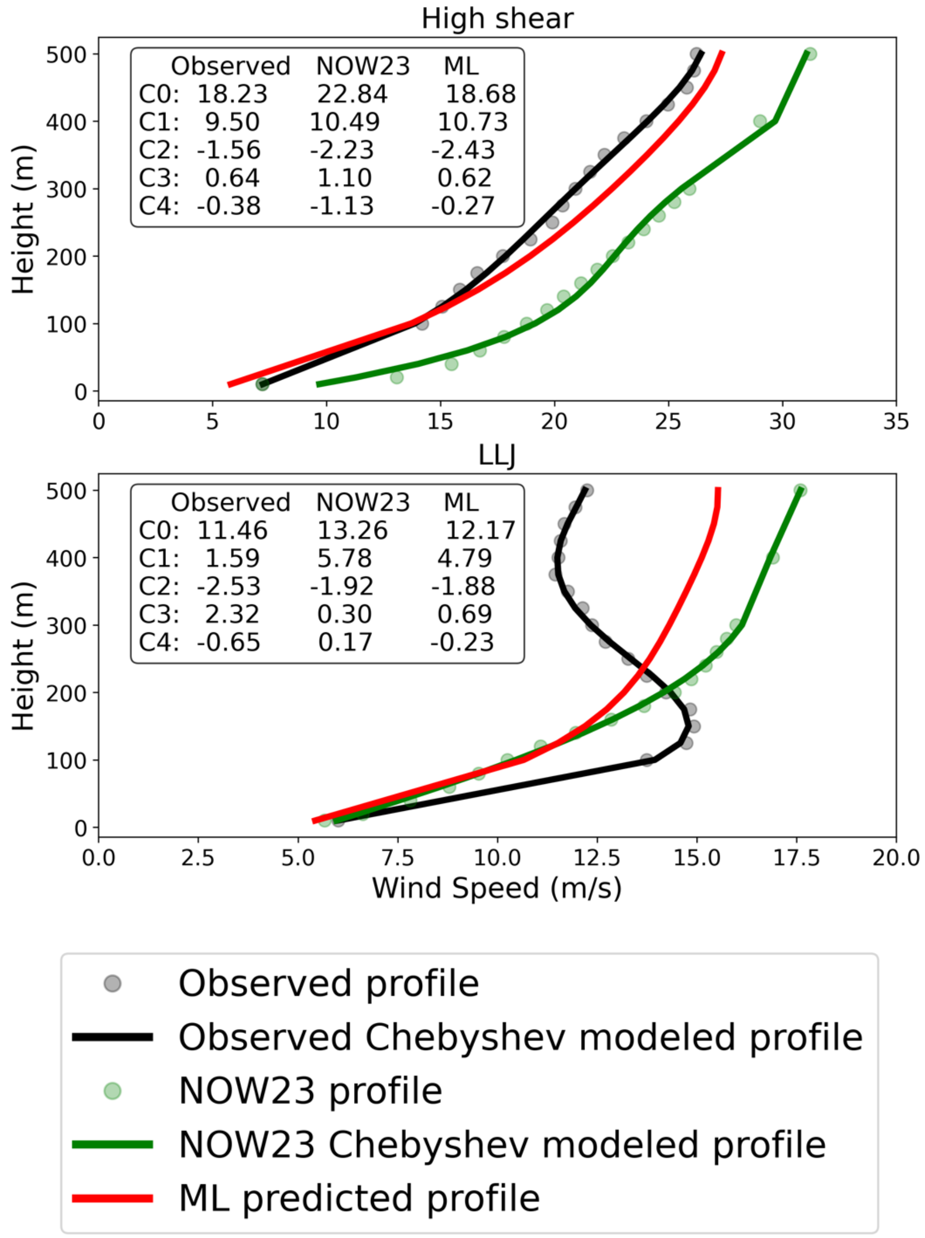
Distribution of wind speed at different levels, from profiler, NOW23, and ML predictions, at station PROF_WANT, for three test years.



Quantile-Quantile comparison of wind speed at different levels between profiler, NOW23, and ML predictions, at station PROF_WANT, for three test years.

Conclusions

- Robust validation approach based on Chebyshev polynomial
- Concurrent predictions are better in ML than physics-based approach, since WRF introduces temporal shift
- WRF performs better than ML, preserving diverse characters of wind profiles – highly important for wind resource assessment



Future work

- ML does not include any temporal dynamics, namely inertial oscillation
- ML lacks spatial features (land-sea temperature gradient)

University at Albany, State University of New York

NYS Mesonet

Thank you
for your attention!

UNIVERSITY AT ALBANY
STATE UNIVERSITY OF NEW YORK

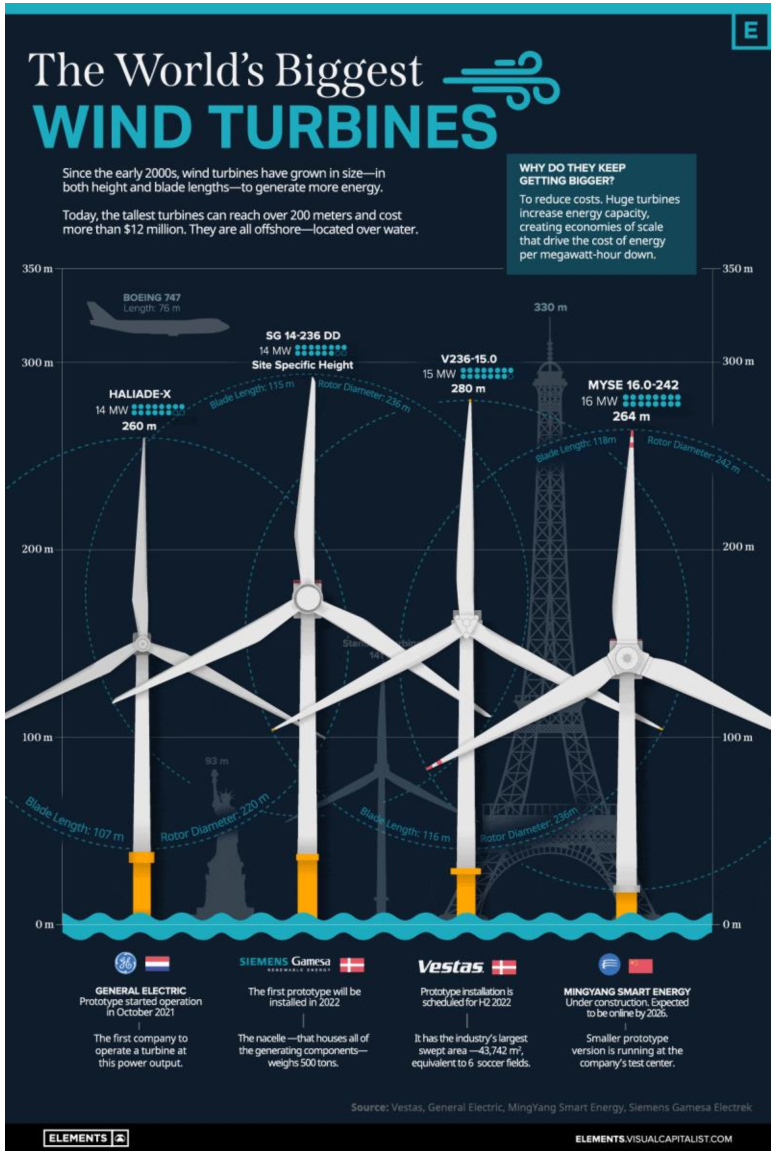
Estimating high-resolution profiles of wind speeds from a global reanalysis dataset using TabNet

Harish Baki
Atmospheric Sciences Research Center, University at Albany, NY, United States

Maximilian Pierzyna
Department of Geoscience and Remote Sensing
Delft University of Technology, The Netherlands

Sukanta Basu
Atmospheric Sciences Research Center,
Department of Environmental and Sustainable Engineering,
University at Albany

Background



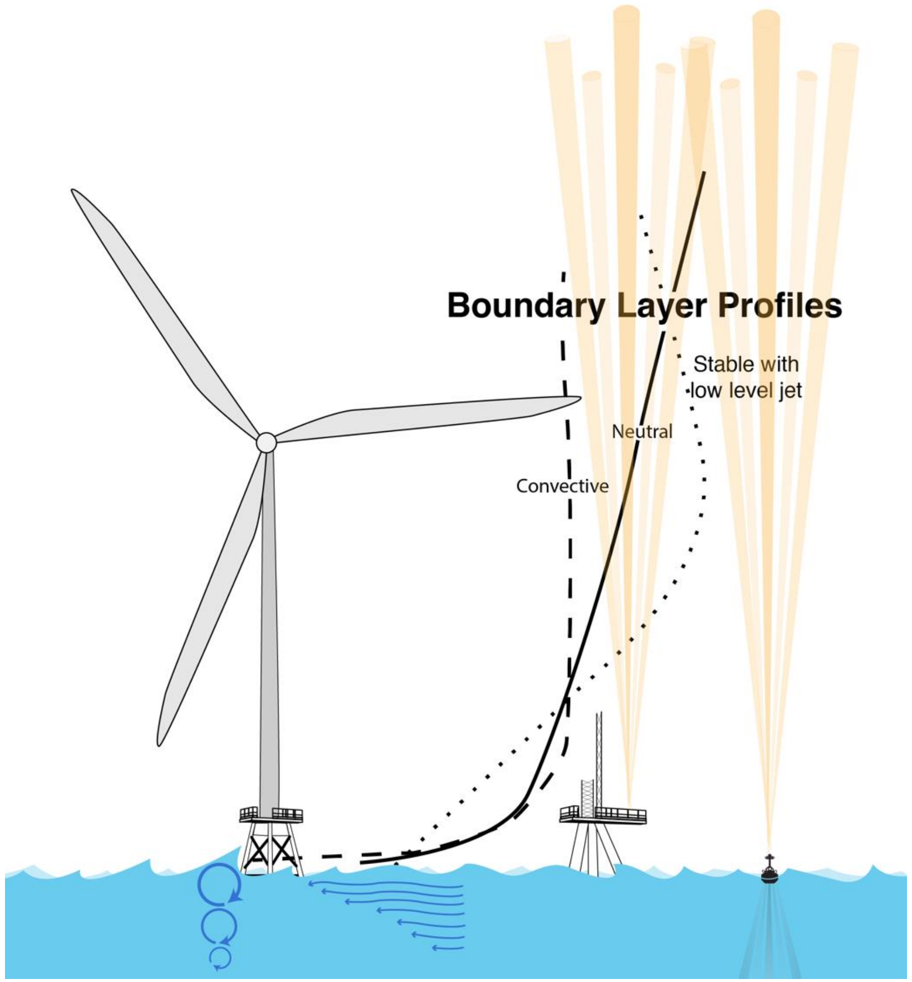


Fig 1. Different types of wind profiles in the atmospheric boundary layer.

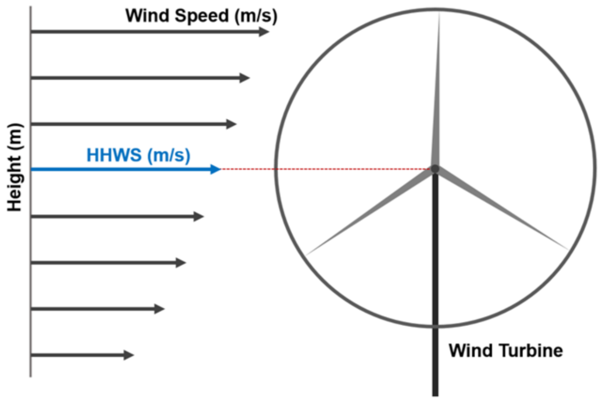


Fig 2. Turbine power as a function of hub height wind speed.

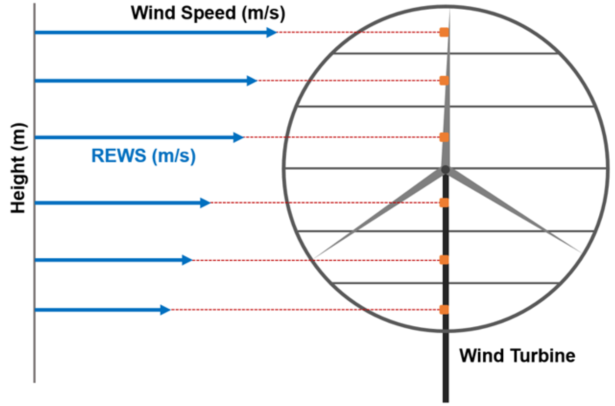
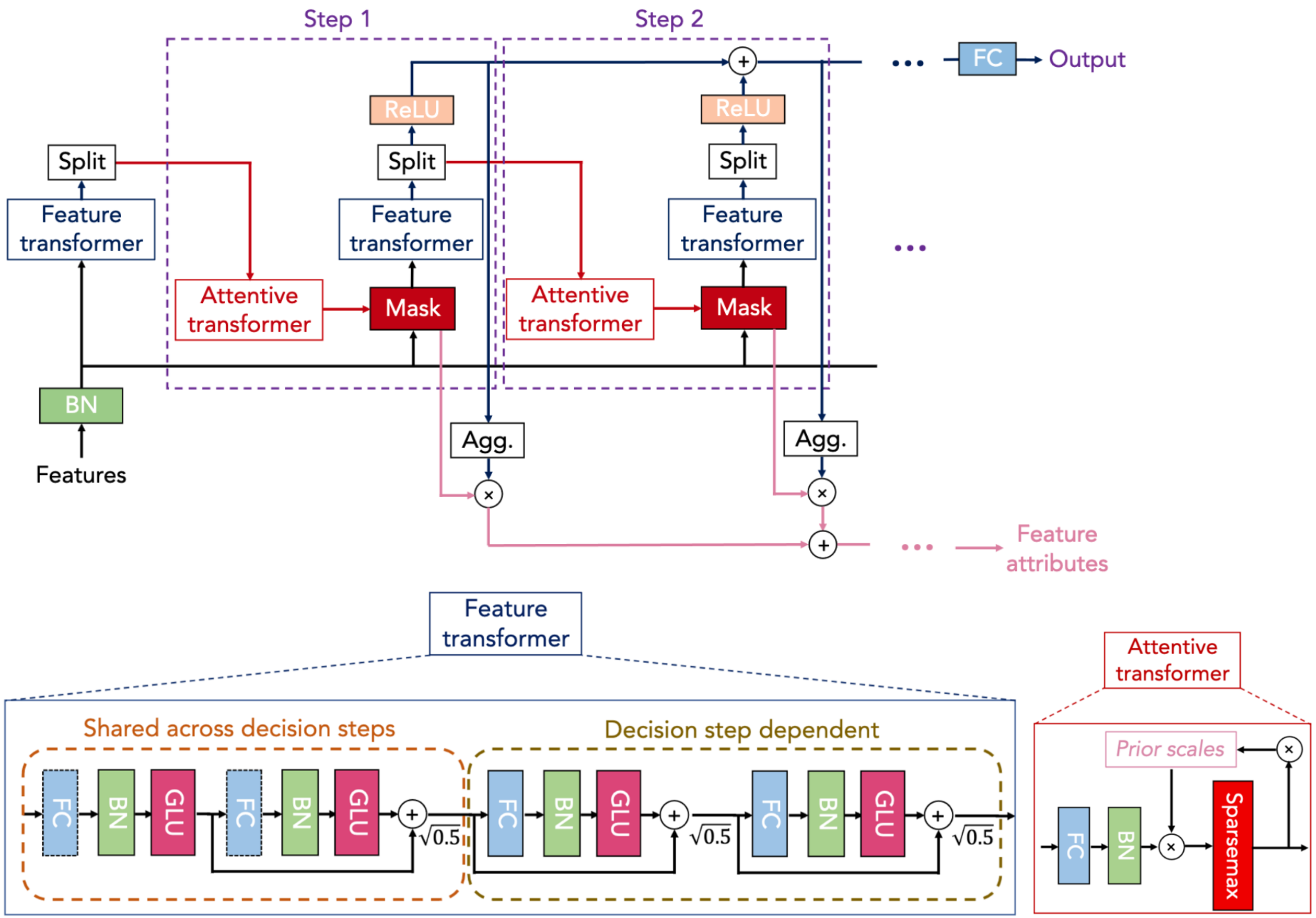


Fig 3. Turbine power as a function of rotor equivalent wind speed.

TabNet: Attentive Interpretable Tabular Learning



Source: Sercan O. Arık and Tomas Pfister (2020). TabNet: Attentive Interpretable Tabular Learning.

Model parameters:

- N_{steps}: decision steps, [3,10]
- n_d: width of decision prediction layer, [8,64]
- n_a: width of attention embedding layer, [8,64]
- n_{shared}: number of shared GLU blocks, [1,5]
- $n_{independent}$: independent GLU blocks, [1,5]
- γ: feature reusage coefficient, [1.0, 2.0]

Feature interpretability

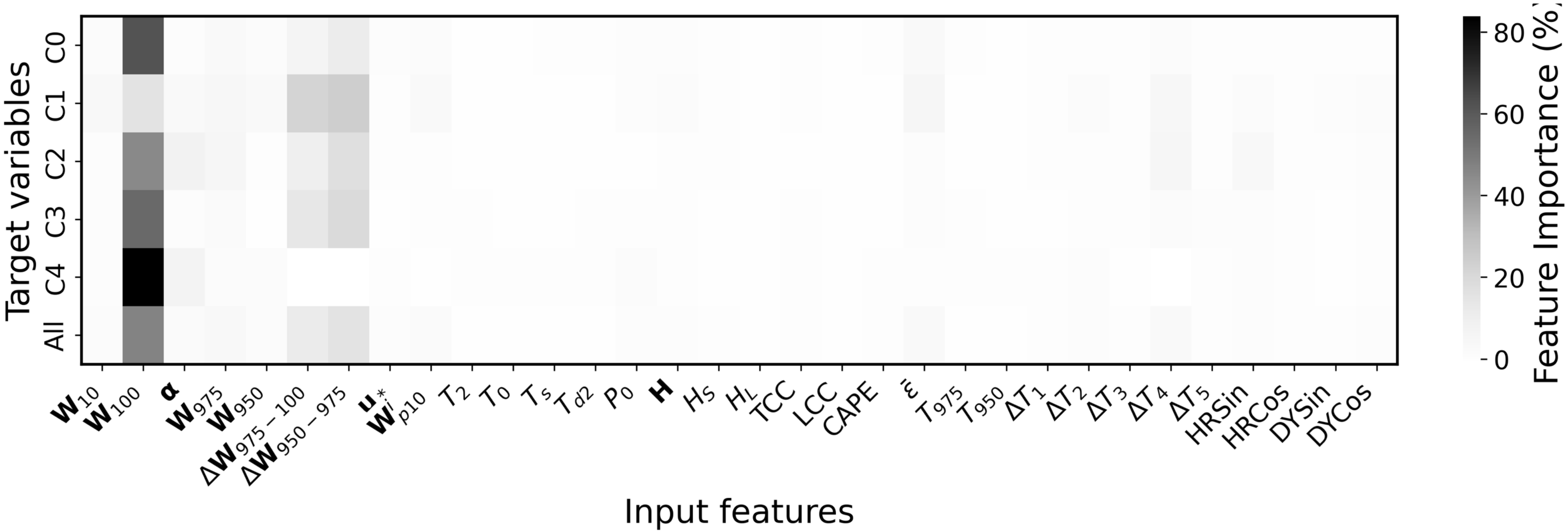


Fig 13. Combined feature importance of input variables on the target coefficients.

Sensitivity to the input data

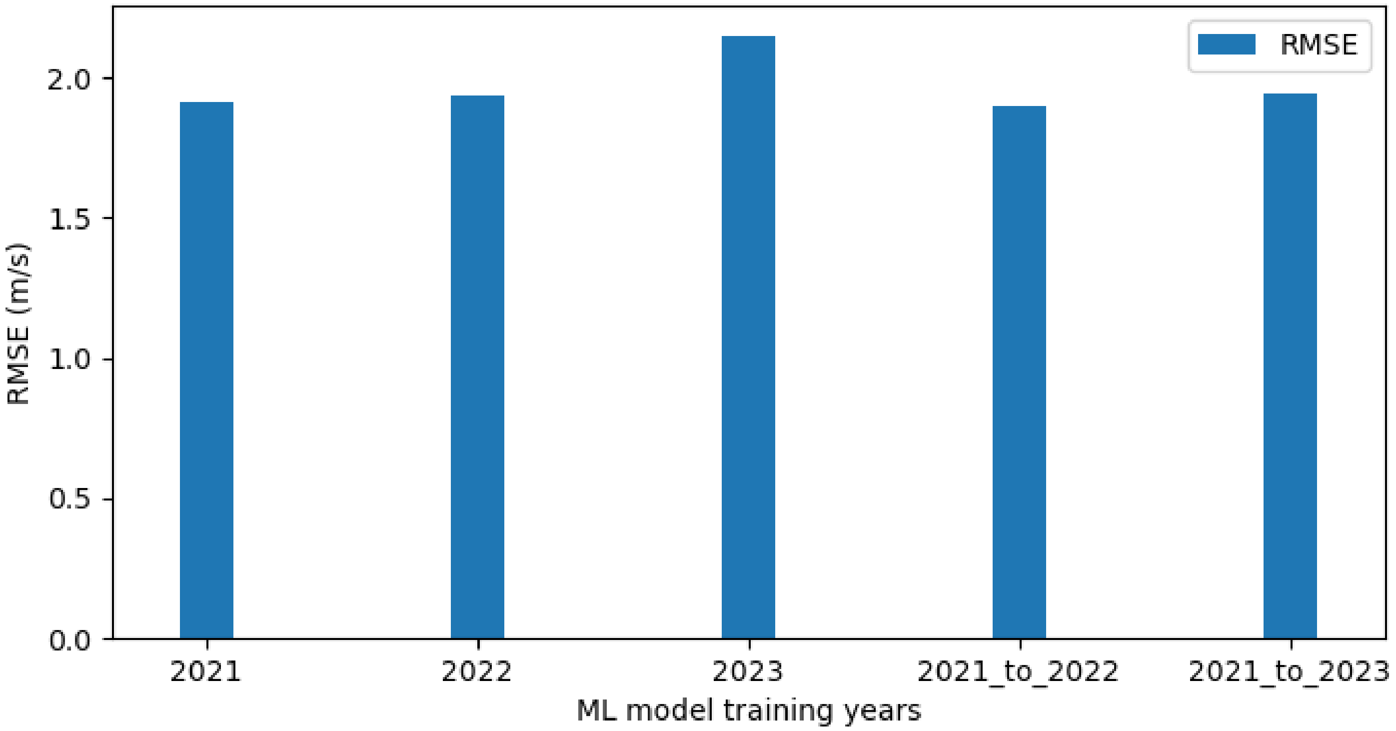


Fig 12. Wind profile RMSE scores from ML predictions, with respect to the NYSM profilers, across different training sizes.