

Enhancing Wind Forecasting Accuracy Using a Decomposition-Based Hierarchical LSTM Model

Ramon Gomes da Silva, Victor H. A. Ribeiro, Doglas Q. dos Anjos, Matheus H. D. M. Ribeiro and Gilberto Reynoso-Meza

Industrial and Systems Eng. Graduate Program (PPGEPS) · PUCPR · UTFPR

Why decompose wind speed before forecasting

Renewable planning

Robust wind speed forecasts support safer and more stable power system operation.

A hard signal

Wind speed is non-stationary and nonlinear, with patterns from seconds to days.

Decomposition helps

Splitting the signal into components lets each feature be modeled separately.

Recent related work

Moreno et al. (2024): multi-stage decomposition mitigates forecast errors.

Jia et al. (2025): SSA + CEEMDAN dual decomposition with an optimized Informer.

Fu et al. (2025): sequential VMD and wavelet decomposition with a bi-GRU; R^2 above 0.99 reported.

PUCPR
GRUPO MARISTA

THE QUESTION

Can reconciliation improve multi-stage decomposition forecasts?

Objective: treat the SSA and VMD components as levels of a hierarchical time series (HTS), forecast each with an LSTM and compare reconciliation methods.

Data: wind speed at a Brazilian wind farm

13,248

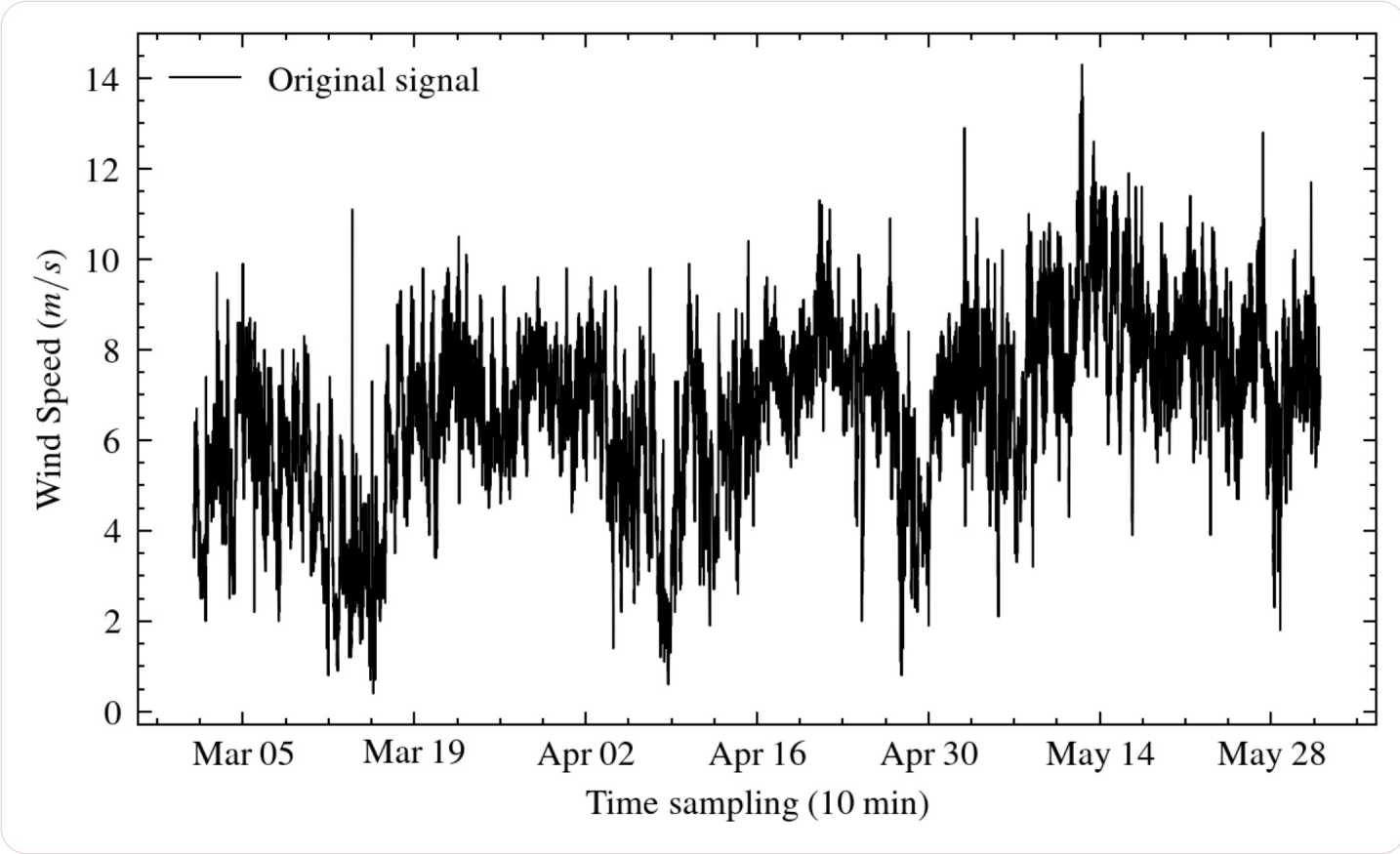
samples, 10-minute sampling, March to May 2020

95 m

turbine height, wind farm in Parazinho, Northeastern Brazil

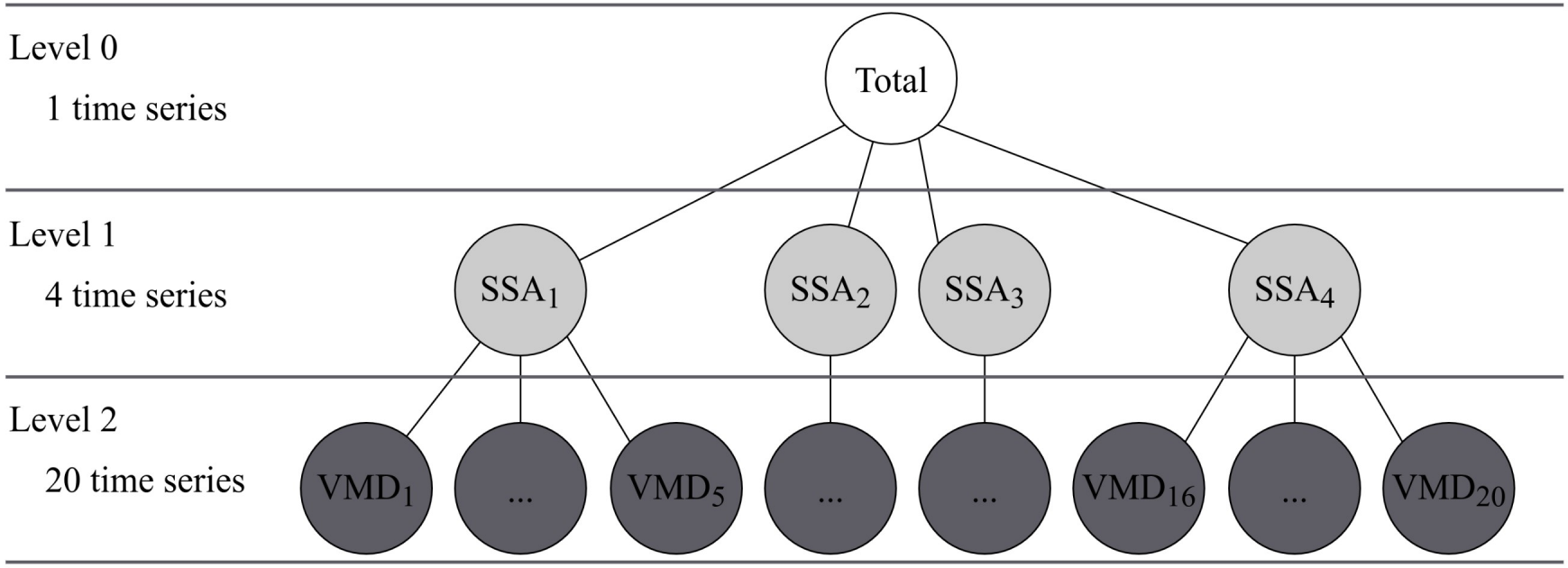
6.68 m/s

mean wind speed (min 0.4, max 14.3, SD 1.95)



Wind speed (m/s), 10-minute series

Hierarchical structure: 25 series in 3 levels



Two-stage decomposition

SSA splits the signal into 4 components; VMD splits each into 5 modes (20 in total)

Reconciliation

Adjusts the base forecasts so that the levels are coherent

Experimental workflow

1	**Data**	13,248 wind speed samples, 10-minute sampling, 2020
2	**Decomposition**	SSA into 4 components, then VMD into 5 modes each
3	**Hierarchy**	Signal → 4 SSA → 20 VMD: 25 series in 3 levels
4	**Base models**	One LSTM per series; HUBER, RIDGE and SVR forecast the signal; Optuna, 50 trials each
5	**Reconciliation**	BU, MINTS, OLS, TDGSA and TDGSF over the LSTM forecasts
6	**Evaluation**	Last 288 observations, horizons of 10, 30 and 60 min; MAPE, RMSE, IP, DM test

PUCPR
GRUPO MARISTA

Five reconciliation approaches

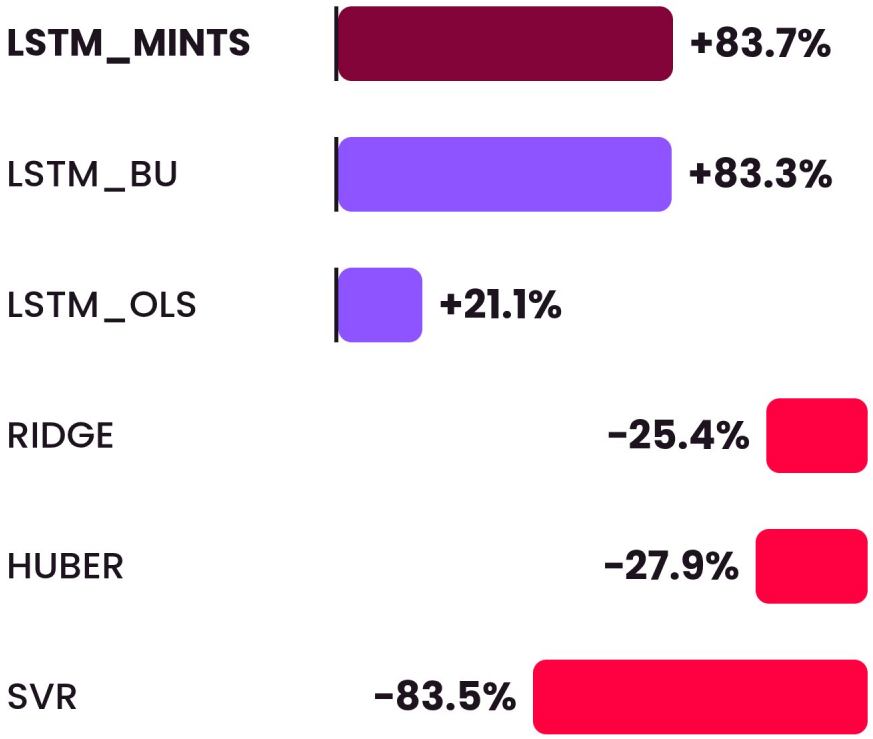
Method	How it reconciles
Bottom-Up (BU)	Sums the forecasts of the 20 VMD modes upward through the hierarchy.
Top-Down (TDGSA)	Splits the top-level forecast by the average of historical proportions (Gross–Sohl A).
Top-Down (TDGSF)	Splits the top-level forecast by proportions of historical averages (Gross–Sohl F).
MinT-OLS (OLS)	Least squares: treats the errors of all series as equal and independent.
MinT-Shrink (MINTS)	Uses the estimated error covariance, with shrinkage, to minimize the variance of the reconciled errors.

All start from the same LSTM forecasts. The reconciled models are named LSTM_BU, LSTM_TDGSA, LSTM_TDGSF, LSTM_OLS and LSTM_MINTS. Top-down methods keep the top-level forecast, so at that level they match the LSTM.

PUCPR
GRUPO MARISTA

MINTS and BU lead the average gain over the LSTM

Average improvement index (IP, %) over the LSTM baseline



BEST METHOD

+83.7%

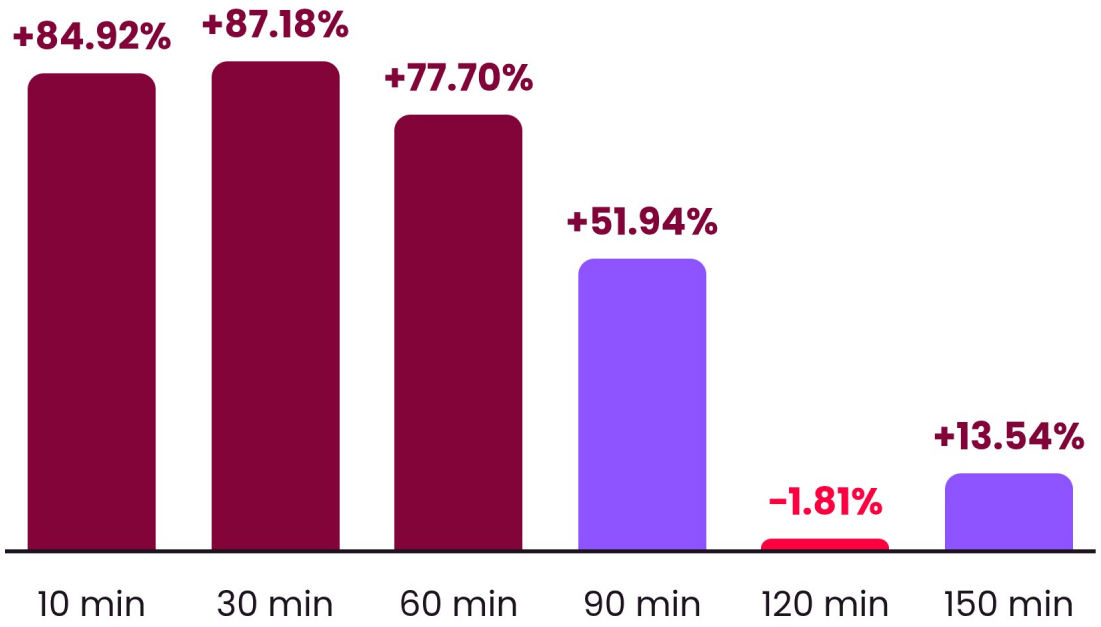
average gain of LSTM_MINTS over the LSTM, across MAPE, RMSE and the three horizons

IP = (M_baseline - M_model) / M_baseline × 100

PUCPR
GRUPO MARISTA

Large gains, but not uniform across horizons

LSTM_MINTS: reduction in MAPE relative to the LSTM (IP)



Burgundy: main horizons. Violet: additional horizons (robustness).
Red: below the baseline.

Best at 30 min

MINTS: 87.18% (MAPE) and 88.51% (RMSE).

BU is equivalent

DM p = 0.097 at 30 min and 0.281 at 60 min.

At 120 min

BU gains 0.99% in MAPE; MINTS falls to -1.81%.

PUCPR
GRUPO MARISTA

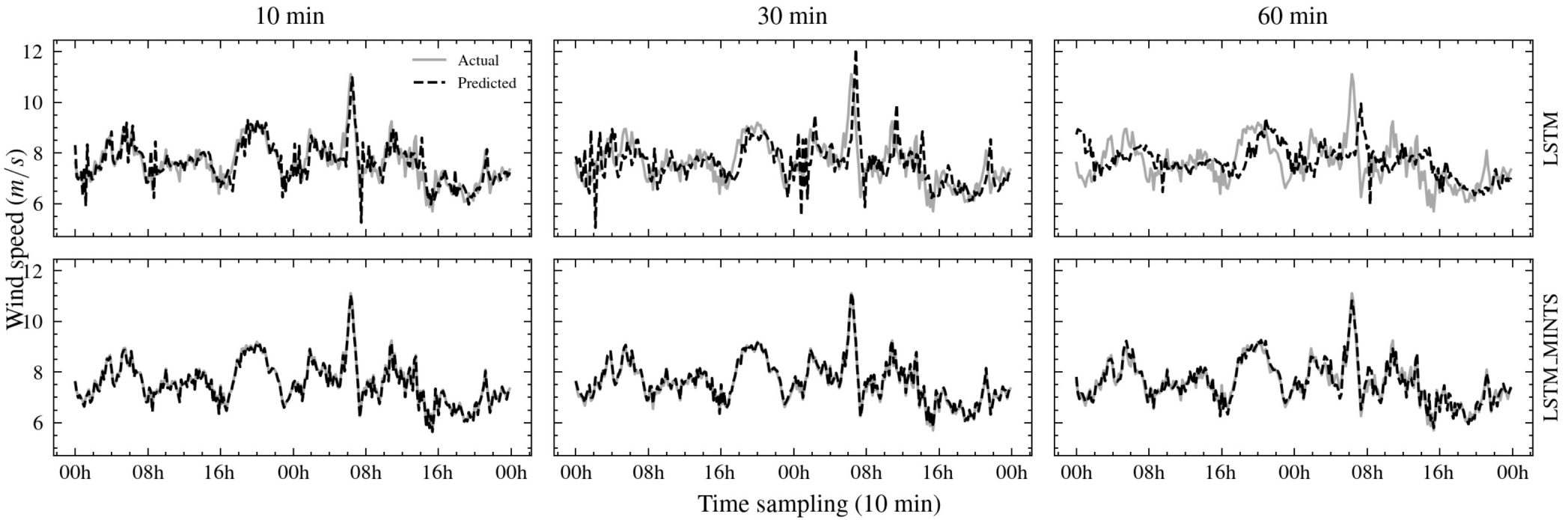
Top-level accuracy of all models

Model	10 min MAPE	10 min RMSE	30 min MAPE	30 min RMSE	60 min MAPE	60 min RMSE
LSTM	4.987%	0.4850	7.424%	0.7463	8.592%	0.8631
LSTM_BU	**0.775%**	**0.0716**	**0.973%**	**0.0872**	**1.954%**	**0.1908**
LSTM_MINTS	**0.752%**	**0.0703**	**0.952%**	**0.0857**	**1.916%**	**0.1872**
LSTM_OLS	3.811%	0.3727	5.779%	0.5809	7.083%	0.7099
LSTM_TDGSA	4.987%	0.4850	7.424%	0.7463	8.592%	0.8631
LSTM_TDGSF	4.987%	0.4850	7.424%	0.7463	8.592%	0.8631
HUBER	8.655%	0.8257	8.801%	0.8328	8.079%	0.8604
RIDGE	9.196%	0.8539	7.660%	0.7935	7.701%	0.8017
SVR	10.118%	1.0790	13.572%	1.3412	13.572%	1.3412

Top-level accuracy on the last 288 observations. Shaded: best model; bold: best and second best. RMSE in m/s.

PUCPR
GRUPO MARISTA

LSTM_MINTS tracks the wind speed at all three horizons



Actual (grey) vs predicted (dashed) wind speed, last 288 observations. Top: LSTM. Bottom: LSTM_MINTS.

Significance: MINTS against every other model

LSTM_MINTS (best model) against all other models; Diebold-Mariano test, squared-error loss, Harvey correction, significance at 5%

Horizon	Significantly beaten ($p < 0.05$)	Not distinguishable
10 min	All 8 models ($p \leq 0.015$)	None
30 min	7 of 8 ($p < 0.001$)	LSTM_BU ($p = 0.097$)
60 min	7 of 8 ($p \leq 0.004$)	LSTM_BU ($p = 0.281$)

At 30 and 60 min, BU and MINTS are statistically equivalent: the evidence does not support MINTS as uniformly superior.

PUCPR
GRUPO MARISTA

Conclusions and next steps

Reconciliation pays off

MINTS and BU cut the LSTM error by up to 88%; MAPE at 10 min falls from 4.99% to 0.75%.

The hierarchy adds value

Organizing SSA and VMD components as an HTS enables reconciliation beyond simple summation.

Read with care

BU matches MINTS at 30 and 60 min; MINTS is below the baseline at 120 min; test window of 288 observations.

Next steps

- Sensitivity of the SSA window and the VMD mode count; decomposition reconstruction error
- Middle-out and probabilistic reconciliation; other neural base models
- Classical benchmarks: naive, ETS and ARIMA/SARIMA
- Multiple, longer test windows

PUCPR
GRUPO MARISTA

Thank you!

Questions and comments

ramon.gsilva@pucpr.br
ramongss.com

