This is a relevant study for the Wind Energy community, presenting a machine learning approach to estimate wind profiles over complex terrain, using valuable input data from field campaigns in the global South (where observations are typically more sparse).

My comments and suggestions are intended to improve the clarity of the paper.

A. Dear referee, we acknowledge all your valuable comments. We are sure that it´s a worthy contribution to our work. We tried, on the best way to attend all the contributions of both referees. Following, all the comments are answered.

Minor Comments - language and typos:

Lines 6-7: Replace 'until' with 'up to'.

A. Done

Line 8: Replace ' when the input dataset included only variables of 40 m height' with 'for input variables up to 40 m high only'.

A. Done

Line 10: Replace 'others' with 'other'.

A. Done

Line 14: Replace 'is' with 'are'; and 'adopted as a power tool' with 'adopted as power tools'.

A. Done

Lines 38-40: These two sentence could be improved. Here's a suggestion: 'Wind forecasts underpin wind power prediction, which is essential to support wind energy production in the short-term. Although winds have been traditionally forecasted with Numerical Weather Prediction models, the use of Machine Learning has become more widespread, not only to correct the biases derived from the highly variable nature of the winds, but also as stand-alone prediction models. Wang et al. (2021) ...'.

A. Done

Line 43: Replace 'some works' with ' a few studies'.

A. Done

Lines 44-45': Replace sentence as: 'They proposed the use of Long-Short Term Memory (LSTM) to improve wind speed forecasting for power prediction.'

A. Done

Lines 48-49: Replace 'and still recommend that the adaptiveness of the hybrid models needs to be further researched' with 'whilst recommending further investigation on the capabilities of hybrid model approaches'.

A. Done

Lines 53-54: Delete all 'the'.

A. Done

Line 62: Replace 'others' with 'other'.

A. Done

Line 64: Replace 'Following the same tendency' with 'Similarly, '.

A. Done

Line 66: Replace as 'and observations at four different heights as input.

A. Done

Lines 67-69: Replace sentence as: 'Bodini and Optis (2020a) and Bodini and Optis (2020b) found that random forests outperform standard wind extrapolation approaches, using a "round-robin" validation method. They highlighted the benefits of including observational data capturing the diurnal variability of the atmospheric boundary layer, namely the Obukhov length (...) at 4 m high.'.

A. Done

Lines 72- 73: Replace as: ', advising to carefully select the input data and emphasizing the importance of normalization'.

A. Done

Line 74: Replace 'over almost plain terrains' with 'over low complex terrain'.

A. Done

Line 75: Replace 'to' with 'in'; 'the most of the studies' with 'most studies'.

A. Done

Line 76: Replace 'for plain terrains' with 'low complexity orographies'; 'that analyzed different surfaces' with 'who analyzed different types of terrain complexity'.

A. Done

Line 79: Replace as: Recurrent Neural Networks (RNNs) are a type of artificial neural networks where the output of one time step is used as an input ...'.

A. Done

Line 109,120: Where you say 'levels' do you mean 'heights'?

A. Corrected

Line 111: Replace 'e' with 'and'.

A. Done

Lines 128-129: Replace start of the sentence: 'The authors state that:'.

A. Done

Lines 146: Replace 'occurred along three years' with 'took place over a three-year period'.

A. Done

Lines 146-147: Replace sentences as: 'All three observational sites are within x km from the coast, and clearly marked on the map.' - it'd be useful to know of far exactly they are from the coast, as proximity to the sea has a great influence over sea breezes and consequently the wind profile'.

A. Done

Line 148-149: Replace 'among each other' with 'between sites'; 'the surfaces' with 'types of terrain'; 'starting by the altitude ...' with 'namely the height and surface roughness'. - My understanding is that here you mean height (above ground) and not altitude (above mean sea level), although altitude is also a relevant factor to consider.

A. Done

Line 150-152: Replace 'by a chaotic constructive pattern that mixes high buildings common to a big city and simple residences' with 'by a densely mixed urban matrix'. (As is Sao Paulo, with a mixture of densely packed high-rise buildings and lower residential dwellings). I'd also remove the last sentence.

A. Done

Lines 153-154: Delete the first 'The'; '... which is characterized by a large number of industries from the largest harbor complex in Latin America.'.

A. Done

Line 155: Replace 'mountain chain ...' with 'mountain range, where Serra do Mar rises sharply, up to more than 700 m high, across 5 km wide.'.

A. Done

Line 167: Delete the first 'The'.; replace as ', as shown in Fig. 1.'

A. Done

Line 169: Replace as ' is the Aracoiaba Hill to the southeast, rising up more than 300 m high up to 900 m altitude.'

A. Done

Lines 195-222: Replace simply as: '... testing different machine learning models with multiple configurations, namely: Random Forests (...). Here we only present results for the best performing model LSTM RNN ...'.

A. Done. More references were included following the suggestions of the other Revisor.

Lines 204-205: Replace as: 'Data from Site 3 was first used to train the model; starting with wind speeds at 40 m to predict speeds at higher heights'.

A. Done

Line 209-210: Replace as: 'As the time series is comprised of 10-min temporal averages, that corresponds to roughly two months worth of observational data.'.

A. Done

Line 212: Delete 'The' before tables.

A. Done

Line 215: Replace 'As we see, for the three sites' with 'For all three sites'.

A. Done

Lines 294-297: Replace as: 'The results found from this observational campaign, albeit short, show the benefits of Doppler Lidar in improving model results to estimate winds at height. This is particularly relevant to help support the Energy transition and Net Zero targets. Despite the costs associated with Doppler Lidars, the authors would encourage further strategic collaborations to drive observational data improvements leading to advances in model prediction.'

A. Done

Major comments - literature review:

Lines 15-37: Although the literature review on Machine Learning in the environmental sciences is interesting, I am struggling to see how this is relevant to the main focus of

the paper (apart from the Cubatao context given in page 6). I'd recommend replacing these paragraphs with a more significant state-of-the-art summary, for example about Machine Learning applied to wind profile prediction, wind resource assessment, and wind power forecasting. Here are a few suggestions:

A. As suggested by the other referee, we reduced the literature review focusing mainly on wind forecasts.We acknowledge all those valuable recommendations, as we could cite them through the text.

Lines 75-76: Regarding the statement made here about vertical wind extrapolation studies, I'd refer the authors to these works looking at a range of terrain complexities, as well as Machine Learning methods: https://doi.org/10.1002/we.2013; https://doi.org/10.1016/j.egyai.2022.100209. I'd suggest a read through these papers and references to complement the Introduction.

A. Done

Lines 156-168: This is really interesting for socio-economic context and from a geographic point of view, but again I am struggling to see how it is relevant to the overarching theme of the paper. I think the authors could simply refer the reader to the Vieira and Gramani (2015) paper for further reading. Instead, I'd suggest to add a few notes on the climatic characteristics for that Sao Paulo region (where the 3 sites are located), as well as orographic complexity indices at each site (e.g. roughness length or silhouette area per unit area, if available). How are seasons typically classified for those areas and terrains - this would help complement the interpretation stated in page 9.

A. We removed part of the text, as suggested.

Major comments - graphics:

Lines 88-97: The mathematical explanation of the LSTM could benefit from a diagram of the network, to help readers understand visually the in/out/update mechanisms.

A. Done

Figures 2,3,4: Would it be possible to round the edges, so that they match the circle lines? Also, it'd be useful to see seasonal results for wind roses per month or season - these could be added to the appendix.

A. Those figures were redone, however, it was not possible to round the edges using the windrose library from Python. The seasonal wind roses were also included for the Site 3, because that Site has a continuous year of observational data.

Figures 5,6,7: Please add errors bars. Could you also add subplots for RMSE, MAE and F-score?

A. RMSE and MAE subplots were included with R2. As the error bars overlapped, we presented them on the appendix, as a subplot for each curve.

Regarding to F_score, this metric is mostly used for classification algorithms, while R2 is mostly used for regression.

Major comments - requiring additional input:

Lines 99-100: Can you please expand how the interpolation was performed and how the data was normalized? For example, was it a simple linear interpolation, or normalized by the maximum wind speed value? This is important to understand how the input data was processed (e.g. were diurnal cycle or spatial corrections accounted for) and how that could be affecting the representativeness of the outputs. It would also be useful to know the percentage of missing data that was processed in this way.

A. For the normalization we used the StandardScaler function from the Scikit Learn Library, as referenced (Pedregosa et al., 2011). This function standardizes the sample by removing the mean and scaling to unit variance. For a observed data (x), its standard value is:

z = (x - u)/s, where u is the mean and s is the standard deviation

The StandardScaler function was also applied by Schwegmann et al. 2023 (https://doi.org/10.1016/j.egyai.2022.100209)

The interpolation was done with the interpolate function from Pandas using the linear method.

These information were included in Section 2.1.

Table A3 exhibits the data availability.

Section 2.4 and Table 4: Can I suggest adding results for Root Mean Square Error (RMSE), so that error values are in the same scale as MAE and MAPE? Also, F-score is widely used in ML model evaluation and comparison - please consider computing it.

A. RMSE included.

Lines 206-208; 219-220; 225-226: Regarding the statement "Thus, using the entire dataset for training and testing the model takes a while. (...) was unnecessary to take the entire dataset" - how long is 'a while', and did you consider compute optimization options such as parallelisation? This could help understand why a shorter input dataset was better - was it just due to compute constraints, or was it the quality of the data? I'd second these questions for Sites 1 and 2.

A. To answer this question, we set the run environment to TPU (https://cloud.google.com/tpu/docs/intro-to-tpu?hl=pt-br#cloud_tpus_are_not_suited_to_the_following_workloads) and ran the Site 3 dataset, changing the data points. Each time we added 1000 data points, the execution time increased by 20%. We didn´t appliy compute optimization, but now we consider it´s as option for future works.

Line 209: Can you please clarify about the train-test-validation split ratios for each site experiment (i.e. was it just train-test or did you reserve some for validation)? Also could you highlight in the manuscript whether the LSTM model was trained and optimised on a site-basis, rather than for all 3 sites simultaneously?

A. As indicated in Table A1, 90% was used for training and 10% for validation. For the tests we took three more samples for each site with 2000, 4000 and 7000 data points. Results are shown in Fig. A8.

The model was trained and optimized on a site-basis and the hyperparameters for each one are indicated at Table A1.

Sections 3.1-3.4: A list (e.g. adding to Table A1) of all input variables for each experiment, as well as any relevant LSTM hyperparameters and network functions, would be useful to share.

A. We added Tables A2 and A4.

Lines 210-211: Regarding the statement: "This means that a short observational campaign can produce reliable results", I think it's a bit premature to reach that conclusion, without seeing results for at least a whole annual cycle. My recommendation is to delete this.

A. Removed

Lines 298-300: For your final remarks, I'd encourage emphasizing the outcomes in two parts: obs campaings and ML model improvements. Both of which are valuable and complementary, although your ML modelling could be further improved if considering other sources of input data (e.g. reanalysis, NWP, etc.) and additional parameters (e.g. humidity, air density, pressure, temperature, orography and land-use).

A. Done