General Comments

This article works to further the use and understanding of wind forecasting using machine learning algorithms. This work uses wind observations at lower levels to make timeseries predictions of winds higher above the surface and compares them to the established method, the Power Law. The authors chose locations with complex terrain to experiment with predictive methods to research where they have the greatest forecast skill. They also tested the viability of swapping models to different environments to look at their fit under different conditions.

The authors break down the model output as a function of the local environments at the different sites, taking the time to delve into the meteorological reasons why the outputs look the way that they do. Despite this, I have concerns about the application of the Power Law, with α being kept constant. As alpha is explained to be a proxy for atmospheric stability, the different terrain scenarios will greatly affect this (i.e. stability is rapidly affected by the onset and movement of the sea breeze and land breeze), especially because the replication of the diurnal cycle was important to the results. The period of record in the training datasets will likely not capture the environmental response to the changing of seasons. The article opens with a lot of connections to previous works but fails to discuss any connections between the research done here and those previous studies.

Most of the paper requires grammatical revisions. There are also portions of the manuscript that do not pertain to the subject matter at hand. This work tackles research that can benefit the scientific community and the furthering of wind forecasting and its application with machine learning methods and their portability, thus I would recommend significant revisions.

A. Dear referee,we acknowledge your comments. All the comments will surely enrich our work. We tried, on the best way, to attend all the contributions of both referees. Below, all the comments are answered. Concerning the grammar, if the draft is accepted, before publication, it will undergo a professional revision.

Specific Questions/Comments

Page 1, Lines 14-24: The first paragraph of the introduction lacks substance that is specific to the work being done here. It discusses the use of Machine learning and its applications, but the scope is too broad. This may be done more concisely by collapsing this paragraph into a sentence or two discussing the variety of machine learning applications with several examples, while transitioning quickly into your second paragraph, which is far more relevant.

A. Done

Page 4, Lines 88-94: Along with the governing equations of the LSTM cell, I suggest adding a diagram to help the reader. You can also connect to the different elements within the cell to strength your further explanation on lines 95-97.

A. Done

Page 5, Line 122: The explanation for Experiment 3 and 4 needs more clarity. After reading into the results, I understood that your goal was to apply each trained model to a

site that it was not originally trained for, however this was not clear when it was first proposed.

A. Rewritten

Page 5&6, Section 2.4: You lead off with the statement of "we chose the metrics that have typically been used in similar works". This statement is rather vague and could use some support. At the end of the paragraph, you say "Zhou et al. (2022) and Baquero et al. (2022) provide detailed explanations for those metrics". I suggest moving these references to the beginning of the paragraph and combining your statements… example: "We chose verification metrics like those used in Zhou et al. and Baquero et al., because…. For further information on these metrics, see Zhou and Baquero."

A. Rewritten

Page 6, Line 150: [This is a stylistic critique] You use the wording "a Metropolitan Region (São Paulo city)". Would it be better to say "the Metropolitan Region of São Paulo"? As the reader passes over parenthetical statements, they tend to cause a pause. This style of adding parenthetical statements for further description is used frequently in the paper. In my opinion, the number of parenthetical statements hurts the reading fluency and detracts from the reading experience. Please apply the stylistic switch where you deem necessary.

A. Rewritten

Page 6, Lines 157-166: The history of research on pollution in and around the Cubatão area is not greatly relevant to the background of this paper, nor its focus on wind forecasting. If the authors are trying to tie the importance of windspeed forecasting to the distribution of pollutants, this does not come across as a motivating factor clearly. If the authors did not intend for this to be a point of motivation for this research, I recommend the removal of this block of text.

A. Rewritten

Page 7, Lines 170-171: "Gasparoto et al. (2014) summarized some characteristics of the Ipanema National Forest." If there is further information that the reader should be aware of regarding the characteristics of the Ipanema National Forest that are relevant to the research being conducted at site 3, I recommend paraphrasing from the source provided or removing the sentence if the information is not overly relevant.

A. Done

Page 7, Line 179: "Both of them present a diurnal cycle" Your figures do not show a time series showing this result. Although I would caution adding more figures, you may put some thought into showing a time series here.

A. Time series were added as appendices

Page 8, Line 181: Could you elaborate on impact of the LLJ at the heights above ground you are using in this research?

A. rewritten

Page 8, Lines 184 and 185: [Klockow and Targa (1998) figure 2] The figure was *extremely* simplified, to the point that I was left wondering how it provided any insight into windspeed prediction or anything to do with the sea breeze. Could you please add further insight from the use of this referenced figure?

A. We agree that this a simplified conceptual model, but that was an important project in the sense of the first observational experiment including international cooperation. The intention of this reference is to illustrate the local circulation and its complexity due to the sea-land contrast and orography based on a scientific publication. This complexity may prejudice the model performance

Page 10, Lines 198-202: There are several references made in these lines, but they are not referencing the papers which foundationally developed the models. I would recommend adding a reference to the foundational works in addition to the references that guided the research done here thru the application and review of these models.

A. Done

Page 10, Line 206: "Thus, using the entire dataset for training and testing the model takes a while." Wording could be more professional here. As a technical comment, this is a good point to reference your computing resources. The use of Google Colab would obviously be different than if someone was training on new GPUs and a dedicated machine. The simplicity of this setup can be a benefit to further implementation on a wider scale by more users, as discussed on lines 294-297.

A. Rewritten

Page 11, Line 207: "At each step" At what steps? For each experiment?

A. Rewritten

Page 11, Line 211: "a short observational campaign can produce reliable results" Based on Table 1, your datasets only spanned certain times of the year sometimes missing seasons, and for site 3… entirely in the southern hemisphere winter. Is there any concern for the model's usability for prediction outside of these timeframes?

A. Yes, there is concern about the model's usability. Regarding this, we analyzed different seasons for the Site 3, which had the longest continuous campaign, covering a whole year, between 2017 and 2018. Those results weren't reported on the text, but we identified only small variations compared to the table 3. Despite we saw a reduction of the correlation, generally the LSTM still performed better than the PL for 2 different periods (December – January and March – April). For the June – July period, the methods converge (the R^2 tend to same value). The tests were done for 60, 80 and 100 m.

We believe that the success of these results stem from the LSTM abilities, as a RNN, to built a memory of time series.

Page 12, Lines 219-224: This paragraph covers a lot of different results. Due to the number of figures and tables, many of them are referenced in this paragraph in quick succession. This made things very difficult to keep track of. I suggest spacing out references and/or adding more elaboration on the results from each figure/table to increase the value of their inclusion.

A. Rewritten

Page 17, Line 261: "The forecast for the Site 1 highly improves when the 60 m wind speed is included on the input dataset for training the model at Site 3". Where is the point of origin of the 60 m wind dataset? This has been left a bit unclear. See the silly equation below that popped into my head as I read this sentence…

[Prediction for Site 3] = [Model trained at Site 1] + ([60m data from Site 3]? Or [60m data from Site 1]?)

A. The point of origin of the 60 m wind is the dataset for training:

[Prediction for Site 3] = [Model trained at Site 1 (40m wind speed + 40m wind dir + 40m TKE + 60m wind speed)

Additionally, have you considered adding 10m/2m data to any of these tests? These heights are the generally the most common in modelling and observed networks. When talking about applying a pretrained model from another region [maybe with similar terrain], it may be worthwhile to test this concept with data that would be more readily available at a wider range of locations.

A. We totally agree with the revisor concerning to modelling with 2m/10m data. Although these heights are the more commonly available, during our experiments, we haven´t simultaneous 2m/10m observations. As we believe it can improve the modelling, we suggest it as future work in the last paragraph of conclusions.

Page 27, Lines 276-277: "Ensemble and hybrid methods are strategies that also contribute to the model performances." Were these concepts tested in this study? If they were not, maybe this should be moved to an area discussing future work.

A. Yes, following Zhou et al. (2022), we tested the Complete Ensemble Empirical Mode Decomposition with Adaptative Noise (CEEMDAN) method. Zhou et al. (2022) applied the CEEMDAN to forecast the carbon price, which is also a random variable. According to Zhou et al. (2022), decomposition-integration methods increase the accuracy and save time when compared with other hybrid methods and the CEEMDAN presents many advantages over its precursors. Those reasons motivated us to choose CEEMDAN over other methods.

Despite the CEEMDAN features, we hadn't as good result as we got if the LSTM. Anyway, we agree with the revisor that it still is a point for investigation and that the hybrid and ensemble methods should be tested deeper for better results. Suggestion included as future work.

Page 27, Lines 279-281: “After testing some commonly used algorithms for the wind speed forecast (Random Forest Trees, Support Vector Regression and Multi-layer Perceptron), we found out the LSTM outperformed all of them. The LSTM outperformed even the decomposition methods.” These models were briefly discussed on page 9 and 10 starting at the end of line 196, but no results were ever shown to the reader, nor was it said that those models had poorer performance, only that “From this point, we refer to results with the LSTM RNN”. This needs to be revised in the results section before this statement is made in the conclusion.

A. Rewritten

Page 27, Line 283: “90%” How much of the remaining 10% of the data went towards validation and testing?

A. All the remaining 10% of the data were used for validation. Then, for testing the models we took more data. We included the explanation in the last paragraph of section 3.2 - Experiment 2.

Page 27, Line 288: “However, the improvement was better to the Sites 1 and 2 than for the Site 3” Is there any speculation or further explanation of why this was the outcome?

A. At this moment we don’t have further explanation for that behavior. Maybe more Planetary Boundary Layer analyses could help us understand those differences.

Page 27, Lines 292-293: “The Site 2 is strongly influenced by the sea and land breezes and the LSTM model captured the abrupt changes of the wind profile better than the Power Law” Please see my major comment in the ‘general comments' section about my concerns with keeping alpha constant when implementing the Power Law. This is one of the cases where there would be a rapid change in stability. You incorporated 60m data into the training of your models, so can you use the change in windspeed between 40-60m over a short timeseries to better estimate alpha ~windshear coefficient?

A. Suggestion accepted. The windshear coefficient was estimated and results are discussed in the last paragraph of Section 3.2 and Figure A5.

Page 27, Lines 294-297: Reword these statements for better clarity.

A. Done

Technical Corrections

Quick preface, if something is corrected here, it may have occurred more than once, however I will only be pointing out the first instance.

Page 1, Line 6: The use of the word “until” in this spot is grammatically incorrect. “Up to” may be a valid correction.

A. Done

Page 1, Line 8: “variables of 40 m” changed to “variables at 40 m”.

A. Done

Page 2, Lines 39-40: “Due the random feature of the wind speed” recommended rewording “Due to the random nature of windspeed when looking at data over short time scales”.

A. Done

Page 3, Lines 81-82: “state-of-art” changed to “state-of-the-art”.

A. Done

Page 4, Table 1: “31/Dec/2-16” changed to “31/Dec/2016”

A. Done

Page 4, Lines 106-107: “retrieve information each 10 minutes” changed to “retrieve information every 10 minutes”.

A. Done

Page 5, Lines 110-111: “the standard deviation of the horizontal (σu + σv) e vertical (σw) wind speed to forecast the wind speed at higher heights” Please revise, I am unsure what ‘e vertical’ is referring to.

A. Done

Page 6, Line 151: “the strong social difference” I am unsure what is meant here; maybe ‘strong social disparity’?

A. Done

Page 6, Line 157: “looks like a big wall.” Please revise with more scientific language or remove this statement.

A. Done

Page 8, Figure 2: There are numbers along the concentric circles of the wind rose that lack units. Based on other examples in literature, these are “%” percentages based on the normalized time period, i.e. relative frequency. See the pdf file for a screenshot that depicts the content of this comment.

A. The wind roses were corrected and now include the “%” signal.

Page 9, Figure 3: Values discussed in my comment above are overlapped by data, making the relative frequency values impossible to read. See pdf for the illustration of this comment.

A. The wind roses were corrected and now the values are visible

Page 9, Line 190: “along the day” change to “throughout the day”

A. Done

Page 9, Line 196: “until to reach the results that will be presented in this section” Please revise wording. Maybe? “the model configurations that had the best performance will be presented in this section”

A. Done

Page 10, Line 205: “50 thousand data” 50 thousand timesteps? Datapoints?

A. Done

Page 11, Lines 206-207: “Surprisingly, we found that the model only improves until a limited dataset size and was unnecessary to take the entire dataset.” Suggested edit for clarity “Surprisingly, we found that model improvement plateaued without using all of the datapoints of record.”

A. Done

Page 13, Line 238: The references used here should be prefaced as examples. ~ (e.g. Vassallo et al., 2020; Mohandes and Rehman, 2018)

A. Done

Page 16, Figure 6: The line representing the Power Law stops being visible around 170m along the x-axis. This would be more representative if we could see the behavior of all three models tested.

A. That happened because we cut the y-axis (R^2) at 60% and the Power Law underperforms dramatically above 170 m compared to the other two models. Keeping the same resolution for the three sites eases the reading..

Page 27, Lines 273-275: Grammar issues in first two sentences of the conclusion. Please revise for clarity.

A. Done