Point-to-point Responses to RC3

Title: Simulation performance of different planetary boundary layer schemes in WRFV4.3.1 on wind field over Sichuan Basin within “Gray zone” resolution
No.:egusphere-2024-1532

We are grateful for the thoughtful and constructive feedback from the referee. We have revised the text in the manuscript to answer the referee’s points and we believe this revision has improved the clarity and quality of our manuscript. This response provides a complete description of the changes that have been made in response to each comment. Referee comments are shown in black text, author responses are shown in blue text.
Once again, we sincerely appreciate the time and effort invested by the reviewers in evaluating our manuscript. We look forward to any additional feedback or suggestions.

The authors undertake a “gray zone” WRF simulation campaign in an understudied region of China (Sichuan Basin) using different PBL schemes compared to one airport meteorological measurement. Results using common statistical error metrics for wind speed and direction are shown, where results for the different PBL schemes show good agreement for wind direction but poor agreement for wind speed. A k-means clustering technique is leveraged to help group different error metrics together and gauge PBL performance.

Overall, while I appreciate the study the authors are trying to undertake, I feel the analysis is underwhelming in breadth and justifications for modeling choices made are weak. Only one observation site is chosen for comparison, and yet it is believed to be representative of the entire region. Additionally, I am still left questioning why such a high spatial resolution WRF simulation was conducted, especially when the comparison was only performed against one observation site. Discussion of relevant meteorological phenomena is vague. For example, stability is often mentioned and used to understand the results, but no mention of a stability metric is used or referenced.

Response:

Thank you for your detailed and constructive feedback. We really appreciate your acknowledgment of the study's intentions and understand the concerns you've raised regarding the scope of the analysis, the justification for modeling choices, and the depth of the discussion.

As you mentioned in the general comments that only one observation site is chosen for comparison, and yet it is believed to be representative of the entire region. We are sorry that the scope of our title is too large and the text does not explain clearly. We also realize that this will cause confusion to readers, so we will revise our title and further explain the purpose and area of this study in the introduction.

We appreciate your feedback and are committed to improving the manuscript accordingly.We believe our revisions will address the concerns raised and enhance the overall quality and rigor of our study. Below are our point-by-point replies, which we hope will address your comments satisfactorily.

2 Data and Methods – general comments/questions

What is the temporal output of the WRF data, and how often is the model updated?I don't think this is every mentioned.

Why use such a high-resolution inner domain? Is it to prove that such simulations are possible with a mesoscale model in this region? It is unclear why such a high resolution WRF simulation is performed, especially when only considering one measurement site.

Only one reference measurement is used, yet strong claims are made about PBL scheme performance for just one 10 m wind tower measurement.

Response:

Thank you for your insightful feedback.

Firstly, we are very sorry for our negligence of the temporal output of the WRF data in our manuscript. The temporal output of our WRF model is 10 minutes. We have added this information in the revised manuscript. However, in the original text, the information "the model is updated every 3 hours" was given at line 163.

Secondly, we appreciate your concern regarding the use of a high-resolution inner domain in our simulations.The decision to employ high-resolution simulations, specifically for the inner domain of the model, is motivated by the critical requirements of airport meteorological support services rather than solely demonstrating the feasibility of such simulations. The advancement of numerical models to a resolution of 1 km is a significant achievement, reflecting the current state-of-the-art in operational weather forecasting. However, as we look to the future, pushing towards even higher resolutions, becomes increasingly critical. Running models at these "gray zone" resolutions is essential for several reasons. For example, there is a need for more refined spatiotemporal resolution prediction of wind fields in many engineering applications.We will clarify in our manuscript.

Currently in the inner domain, it is difficult to collect stations with the wind data at a time resolution of 10 minutes and the access is open, so our research is focused on a single station, which limits our ability to generalize findings across different terrains and climates. However, when we use one single site, we also chose the approach that conduct multiple simulations covering different time periods and meteorological conditions to evaluate the performance of a PBL scheme. We believe that this approach not only strengthens the reliability of our findings but also demonstrates the feasibility and advantages of high-resolution mesoscale modeling in such complex and unique environments. Nonetheless, we acknowledge that further studies incorporating multiple sites would be beneficial to validate and expand upon these results.

What’s more, single site research is also seen in many other studies, such as Mantovani Júnior et al.(2023). In their research, the performance of eight PBL schemes is evaluated using detailed observations of the 2014 and 2015 dry season periods, specifically from 30 September to 2 October 2014, as well as from 14 to 16 October 2015.The observational data were collected at a research site named T3 (3.213°S, 60.598° W, 50 m) located nearby the confluence of the Negro and Solimoes rivers in Manacapuru City, Amazonas, Brazil, the result show that the local MYNN2.5 scheme showed the overall best performance for PBLH prediction, mainly at night.

In the research of Draxl et al.(2012), one coastal site over western Denmark was used to evaluate the wind speed and vertical wind shears simulated by different PBL schemes of WRF model. Dong et al.(2018) used the observation data from an one Arctic coastal meteorology station (named Tiksi Station) to evaluate high-resolution WRF simulations of strong surface wind for the Arctic region.

Moving forward, we plan to increase the number of observational sites to provide a more comprehensive analysis of the gray zone resolution issues across a broader region, including complex basins. This will enable us to refine our models further and address the resolution challenges specific to basin topographies.

References:

Mantovani Júnior, J.A.; Aravéquia, J.A.; Carneiro, R.G.; Fisch, G. Evaluation of PBL Parameterization Schemes in WRF Model Predictions during the Dry Season of the Central Amazon Basin. Atmosphere, 14, 850, https://doi.org/10.3390/atmos14050850,2023.

Draxl, C., Hahmann, A., Peña, A.,and Giebel, G.: Evaluating winds and vertical shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. Wind Energy. 17. 10.1002/we.1555, 2014.

Dong, H. T., Cao, S. Y., Takemi, T., and Ge, Y. Y.:WRF simulation of surface wind in high latitudes, Journal of Wind Engineering and Industrial Aerodynamics, 179,287-296,https://doi.org/10.1016/j.jweia.2018.06.009,2018.

2 Data and Methods – specific comments/questions

pg. 5, line 166: A spin-up period of 3-hours is short, especially with a domain with complex topography. What was the reason for such a short spin-up time? I’m concerned this could affect results for the case studies, at least in the first few hours after spin-up are thrown out.

Response:

Thank you for your valuable feedback regarding the spin-up period of 3 hours.

In the matter of fact, There is a lack of consensus and clear guidance on identifying the suitable length of spin-up time, the optimal spin-up time vary by event and situation. We acknowledge that the WRF spin-up time should not be too short as it is hard to develop the appropriate atmospheric circulations, but not the longer the better (Liu et al., 2023). The proper spin-up time depends on the time needed for initialization, which can be affected by the size of the domain and the local boundary perturbations (Warner et al., 1997; Kleczek et al., 2014), most studies have empirically chosen 6-12 hours as the spin-up time. However, their research area is large and the spatial resolutions is 1 km or lower, lacking the finer granularity of the present study. Empirically, the finer the grid size, the more time steps it would have in a given time window, hence faster spin-up. Given that our case studies primarily focus on short-term phenomena, and the model has a high spatial resolution of 0.3 km, and the analysis is based on 10-minute intervals, we choose a spin up time of 3hours.

To ensure that the short spin-up period did not adversely affect the outcomes of our study, we conducted sensitivity tests by extending the spin-up duration and comparing results (Figure 1 and Table 1). The differences in key output variables were minimal, which supports the adequacy of the 3-hour spin-up time for our specific application.

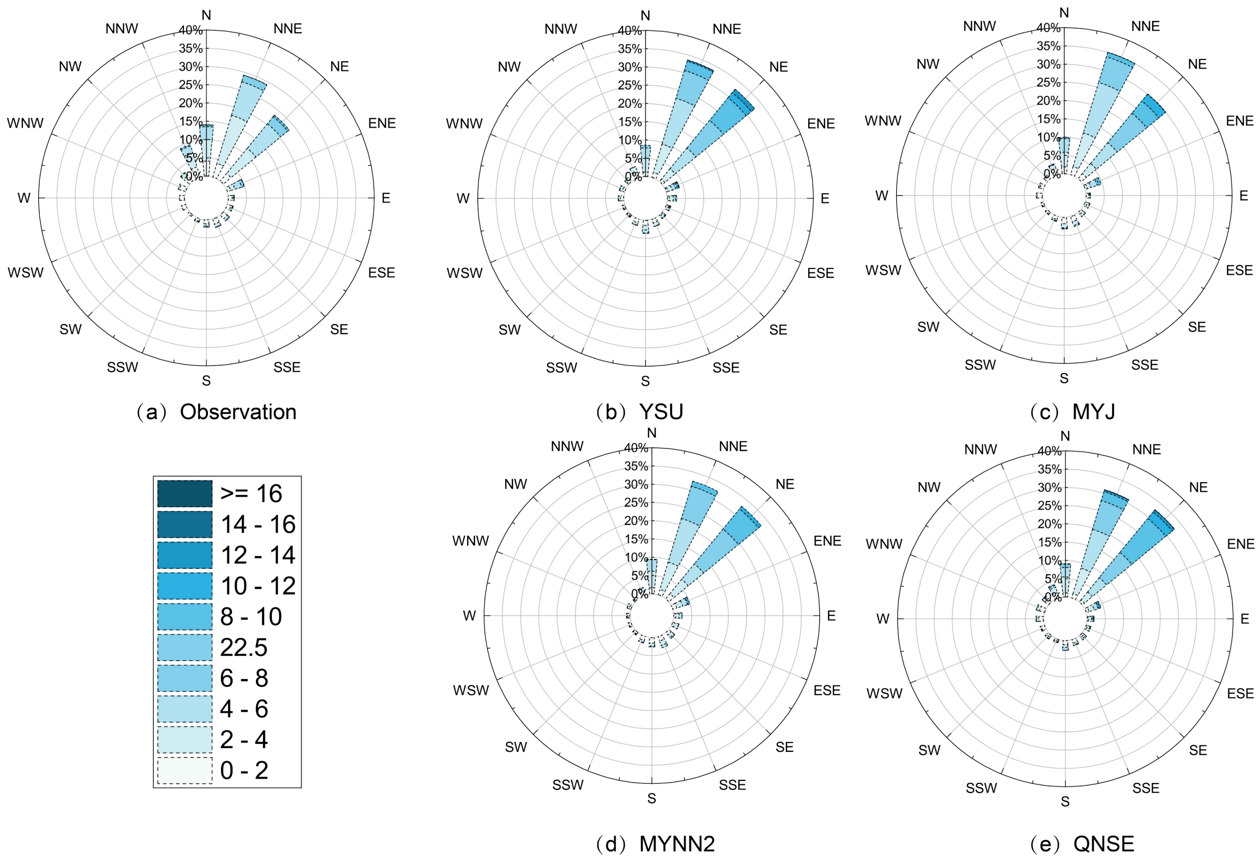

Figure 1. The wind rose for observation and simulated near-surface wind field corresponding to the four PBL schemes when considering spin up time of 6 hours, the circles represent the relative frequency (%), and the colors represent wind speed.

Table 1 Comparison of metrics for four PBL schemes between spin-up time of 3 hours and 6hours.

PBL Scheme	Metirc	Spin-up time	
		3h	6h
YSU	RMSE	2.62	2.84
	COR	0.60	0.58
	Bias	1.63	1.85
MYJ	RMSE	2.28	2.46
	COR	0.58	0.56
	Bias	1.18	1.36
MYNN2	RMSE	2.41	2.62
	COR	0.54	0.51
	Bias	1.34	1.51
QNSE	RMSE	2.75	3.01
	COR	0.61	0.59
	Bias	1.70	1.97

Reference:

Liu, Y., Zhuo, L., Han, D. W.: Developing spin-up time framework for WRF extreme precipitation simulations, Journal of Hydrology, 620, https:// doi.org/ 10.1016/ j.jhydrol. 2023.129443, 2023.

Warner, T. T., Peterson, R. A., and Treadon, R. E.: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction, B. Am. Meteorol. Soc., 78, 2599–2617, 1997.

Kleczek, M. A., Steeneveld, G. J., and Holtslag, A. A.: Evaluation of the weather research and forecasting mesoscale model for GABLS3: impact of boundary-layer schemes, boundary conditions and spin-up, Bound.-Lay. Meteorol., 152, 213–243, 2014.

3 Overview of historical cases and evaluation of simulations results – general comments/questions.

Throughout the results stability is mentioned many times by the authors, but it is never made clear how stability is defined in this study. If a discussion of model results compared to observations is going to take place, stability needs to be defined and/or referenced.

Response:

Thank you for highlighting the need for clarity regarding the concept of "stability" mentioned throughout the results section.

The stable mentioned in line 307-312, which means the stability of boundary layer. However, the stable mentioned in line 380-382, what we want to express through data analysis is that the error distribution obtained from the simulation of 28 individual cases by QNSE scheme has small changes, unlike the large changes in the error simulation of other schemes.We'll replace stable with a more accurate description in our manuscript.

3 Overview of historical cases and evaluation of simulation results – specific comments/questions

pg. 8, line 240: A more thorough description of the dominate atmospheric circulations for each event is needed. Where is the "cold air" coming from? Is it a frontal passage, low- level jet, local terrain flows, etc.? Just saying "cold air" is not informative.

Response:

We agree with the refree's suggestions and will incorporate the recommended changes into the manuscript.

Figure 2b: I appreciate and understand what the authors are trying to convey here, as trying to plot 28 different time-series in one plot is not easy. I would emphasize in the figure caption though that this is not a continuous time-series, as upon first glance, the figure can be misleading. Also, what is the significance of the 5 m/s dotted line?

Response:

Thank you for your insightful feedback regarding our figure. We will elaborate on Figure 2.
The shading in the figure was employed to highlight the time series of the 28 selected cases, which are discontinuous across days. To enhance clarity for the readers, we shaded every alternate case in pink. Regarding the dashed line, to facilitate understanding and to avoid confusion, we opted to include a contour line at 5 m/s. This choice was made because we selected individual cases where the 10-minute average wind speed was greater than or equal to 6 m/s. Since displaying all 28 cases with frequency of every 10 minutes in detail was unfeasible, we presented hourly data, which may inadvertently suggest that some individual cases did not reach the 6 m/s threshold. Therefore, we draw a dashed line of 5m/s instead of 6m/s in the picture to increase the readability of the picture .

Figure 3: Why is the color bar range for wind speed values different than those of Figure 2a? This makes it difficult to compare observations and model results. It would be more beneficial visually if the observational wind rose from Figure 2 is combined into one figure with the model results of Figure 3 to more easily compare.

Response:

We agree with the refree's suggestions and have incorporated the recommended changes into the manuscript.We believe these changes enhance the overall clarity and comparability of the data and hope they address your concerns.

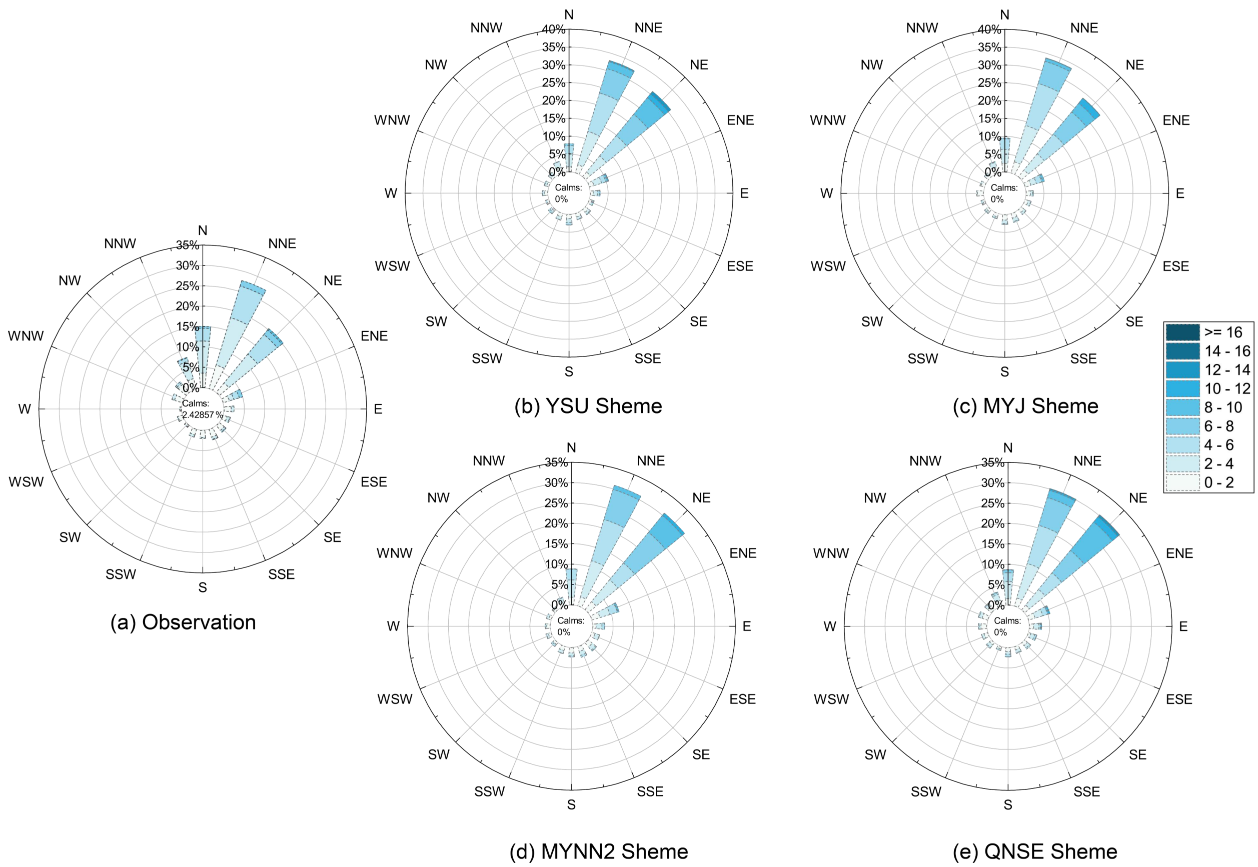

Figure 2. The wind rose for observation and simulated near-surface wind field corresponding to the four PBL schemes, the circles represent the relative frequency (%), and the colors represent wind speed.

pg. 11, line 285: Again, it'shard to compare the differences in wind speed with a different color bar range and not having the plots side-by-side. Additionally, what are these other studies showing similar results? Cite them at the very least, and perhaps include some number ranges for reference.

Response:

Thank you for your feedback regarding the comparison of wind speed differences and the references to other studies showing similar results. To address your concern, we have revised the figures to ensure a consistent color bar range across all wind speed plots. Furthermore, we have arranged the plots side-by-side within a single figure, facilitating a clearer and more direct comparison. We also have updated the manuscript to include citations to the relevant studies (e.g., Gómez et al.,2015; Denis et al.,2020;Yu et al., 2022).

We hope these revisions adequately address your concerns and improve the clarity and relevance of our work.

References:

Yu, E., Bai, R., Chen, X., and Shao, L.: Impact of physical parameterizations on wind simulation with WRF V3.9.1.1 under stable conditions at planetary boundary layer gray-zone resolution: a case study over the coastal regions of North China, Geosci. Model Dev., 15, 8111–8134, https://doi.org/10.5194/gmd-15-8111-2022, 2022.

Gómez-Navarro, J. J., Raible, C. C., and Dierer, S.: Sensitivity of the WRF model to PBL parametrisations and nesting techniques: evaluation of wind storms over complex terrain, Geosci. Model Dev., 8, 3349 – 3363, https://doi.org/10.5194/gmd-8-3349-2015, 2015.

Dzebre, D. E. and Muyiwa, S. A.:A preliminary sensitivity study of Planetary Boundary Layer parameterisation schemes in the weather research and forecasting model to surface winds in coastal Ghana.Renewable Energy,146,66-86,2020.

pg. 14, line 356: Quantitatively state what these deviations are instead of using qualitative language. This advice goes for the entire paper, where qualitative statements are often more common than quantitative.

Response：

Thank you for your valuable feedback regarding the use of qualitative language in our manuscript. We appreciate your suggestion to incorporate more quantitative statements throughout the paper. In response to your comment, we have thoroughly reviewed the manuscript and made the revisions.

Figure 7: There is a lot of information being shown here, which is tricky to do, but would this be better as a line plot where each line is a different PBL scheme, and the error bars are shading around those lines? That might be easier to read than ~100 bar charts.

Response:

Thank you for your suggestion to use a line plot with shading to represent the different PBL schemes. We understand that this approach could potentially make the comparison easier to read by reducing visual complexity. In fact, we tried implementing this idea (shown in Figure 2), but after careful consideration, we believe that our original method using box plots provides more detailed and comprehensive information.

In the original figure, the box plot is used to present the diurnal variation of wind speed errors corresponding to the four PBL schemes. This approach allows us to display not only the median error but also the spread of the data, including the interquartile range and outliers, which are crucial for understanding the variability and distribution of errors across different times of the day. The box plots effectively highlight the differences in performance among the PBL schemes, capturing both central tendencies and variations.

While the line plot with shading could simplify the visual presentation, it may also obscure some of the nuances in the data, particularly the distribution characteristics that are central to our analysis. Therefore, we believe that retaining the box plot format will provide a more informative and nuanced comparison of the PBL schemes, but we will add a detail description of this figure in the manuscript.

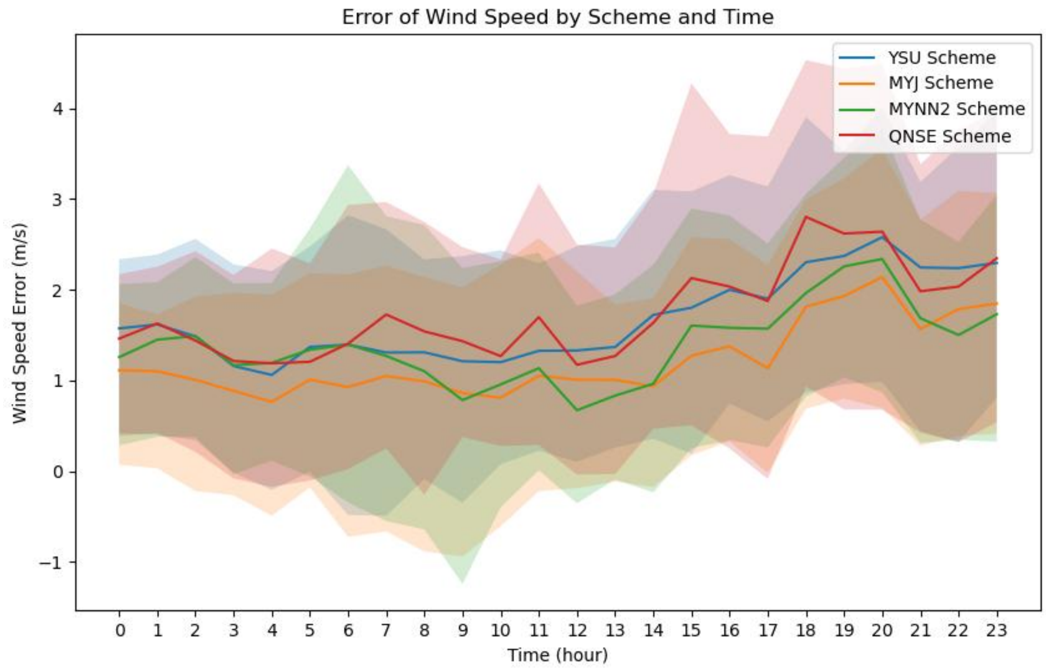

Figure 2. Diurnal variation of wind speed errors corresponding to four PBL schemes.

pg. 15, line 390: Perhaps the wrong word is being used here, but if the authors are going to make claims of significance, the authors should back up this statement with statistical significance tests. Otherwise, remove this statement and/or reword this sentence.

Response:

Thank you for pointing out the need for clarity regarding our claims of significance. we will remove such statements to avoid misinterpretation.

pg. 16, line 406: Unclassified results? What does this mean?

Response:

Thank you for pointing out the ambiguity in the term "unclassified results." What we meant by this was the simulation results obtained before applying the K-means clustering analysis. To clarify, the results from the simulations were first analyzed without any clustering (i.e., unclassified) and then compared to the results after applying the K-means clustering. In both cases, the QNSE and MYJ schemes were consistently found to be the most reliable for surface wind simulation in the Sichuan Basin.

Revised manuscript text as follows:

"This is consistent with the results obtained before applying K-means clustering, indicating that the QNSE and MYJ schemes are relatively stable and reliable choices for surface wind simulation in the Sichuan Basin with a model grid resolution of 0.3 km."

pg. 16, line 414: Are seasonal results not shown because there are no obvious seasonal differences?

Response:

Thank you for your comment. We have removed this statement in our manuscript.