# Point-to-point Responses to RC2

**Review of "Simulation performance of different planetary boundary layer schemes in WRF V4.3.1 on wind field over Sichuan Basin within "Gray zone" resolution" by Want et. al.**

This study performed sensitivity experiments using four PBL schemes over the complex terrain in Sichuan Basin at the "Gray zone" resolution. The results show that while wind direction can be well reproduced and is not very sensitive to the PBL schemes, wind speed shows more sensitivity. The QNSE scheme had the best performance in reproducing the temporal variation, whereas the MYJ scheme had the smallest model bias. Using K-means classification, the authors concluded that the performance of the schemes is influenced by circulations. Predicting near-surface winds has practical importance and remains an ongoing challenge, especially over complex terrains. The choice of PBL has a significant impact on model performance. Therefore, this study is significant in this regard. However, the present form of analysis can be improved. I would overall recommend a major revision before it can be considered for possible publication.

Response:

Thank you for your thoughtful review and constructive comments on our manuscript. We have carefully considered your comments and suggestions for improvement. we appreciate your recognition of the significance of our study and your recommendation for a major revision. We are committed to making the necessary improvements and will ensure that the revised manuscript addresses all your comments comprehensively. Thank you once again for your valuable feedback, which will undoubtedly strengthen our work. Below are our point-by-point replies.

**Major comments**

1. Since the authors emphasize this is a case study, one would expect case-by-case analysis. However, most analyses focus on bulk statistics or aggregate the data in some ways. The cases were selected solely based on wind speed exceeding 6 m/s. Is there any reason why this threshold is used? The length of each case should also be clarified.

Response:

Thank you for your insightful comment.

We acknowledge the importance of conducting a case-by-case analysis, particularly in the context of a case study, as it allows for a deeper understanding. We will add analysis of different individual weather conditions in our manuscript.

However, it is true that the performance of a PBL scheme in simulating near-surface winds based on a single case study has inherent limitations.The success of a single case study might depend on specific initial conditions, boundary conditions, and external forcing factors. These conditions can affect the performance of the PBL scheme. So, we think it is necessary to conduct multiple simulations covering different time periods and meteorological conditions to evaluate the performance of a PBL scheme at a single site.

For the present study, we primarily focus on bulk statistics and aggregate data to provide an initial broad assessment across multiple cases. This approach offers a comprehensive overview of the model's performance under varying conditions, which we believe is essential for establishing a robust baseline.We appreciate your suggestion, and it aligns well with the next phase of our research, where we will focus on in-depth case-by-case analyses to further refine our understanding of the model's capabilities and limitations over western Sichuan Basin.

The threshold of 6 m/s for wind speed was selected based on its relevance to operational forecasting needs, especially in regions where wind speed significantly influences aviation safety and efficiency. This threshold aligns with established criteria in similar studies and ensures that the cases considered are both meteorologically significant and operationally relevant.

Regarding the length of each case, given that there are no instances of average wind speeds exceeding 6 m/s lasting longer than 24 hours in the western Sichuan Basin, we have therefore chosen a 24-hour length for each case.

2. The distribution probability analysis is a good way to evaluate the bulk features. How are the two parameters used in the Weibull distribution function connected to the distribution properties? Is the 10-min or event average used in the Weibull analysis? Please clarify.

Response:

Thank you for your insightful comments. We will incorporate these clarifications and ensure that the revised version addresses your concerns. Thank you for your valuable feedback.

The two parameters of the Weibull distribution, typically referred to as the shape parameter ($k$) and the scale parameter ($\lambda$), play crucial roles in defining the distribution's characteristics. The shape parameter ($k$) indicates the distribution's variability and can provide insights into the nature of the wind speed distribution (e.g., whether it is more uniform or skewed). A value of $k < 1$ suggests that the distribution has a decreasing hazard function, while $k > 1$ indicates an increasing hazard function. The scale parameter ($\lambda$) provides a measure of the distribution's scale, representing the characteristic wind speed. We will expand the discussion in the revised manuscript to clarify how these parameters relate to the bulk features being analyzed.

In our analysis, the 10-min winds were used in the Weibull analysis. This choice was made to ensure that the results accurately reflect the temporal variability of the wind speeds and provide a robust representation of the data. We will clarify this point in the revised manuscript to ensure that readers understand the basis for our analysis.

3. The performance of PBL schemes can be influenced by many factors such as model assumptions, weather conditions, and local stability. Events with similar statistical errors do not directly reflect that they resulted from similar driving factors. Instead of classifying the events based on their statistical errors, I would suggest the opposite approach – classify the weather conditions and link the model errors to them.

Response:

Thank you for your valuable comments.

We appreciate your suggestion and recognize its potential relevance, we will consider incorporating a discussion of how future research could integrate both approaches—classifying events by weather conditions and analyzing statistical errors—to provide a more comprehensive understanding of PBL scheme performance over Sichuan Basin.

While we recognize that the performance of PBL schemes is indeed influenced by various factors, our primary objective in this study is to evaluate the performance of model outputs specifically in relation to statistical errors (RMSE and COR). Classifying events based on their statistical errors can help us distinguish between different cases that may have similar RMSE but different trends COR, revealing patterns in the performance of the model. For instance, a high RMSE with a low COR might indicate systematic errors rather than random fluctuations, suggesting specific adjustments may be needed in model parameters.This will provide a basis for subsequent mechanism analysis of these 28 individual simulations based on weather conditions and stability.

Therefore, we believe that our approach is suitable for the objectives of this study and provides valuable insights into the sensitivity of PBL schemes to statistical errors.

Thank you once again for your thoughtful suggestions, which will help enhance the clarity of our manuscript, and we are open to any further suggestions you may have.

**Specific comments**

Line 110-113: Please elaborate on why it is important to run the model at the “gray zone” resolutions?

Response:

Thank you for your valuable comments and suggestions.

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.

Line 45: change to “winds”.

Response:

Suggestion taken, thank you for your comment.

Line 69: Please add WRF version.

Response:

Suggestion taken, thank you for your comment.

Line 83 and other places: Add a space between the number and units.

Response:

Thank you for your comments and suggestions. We acknowledge this typo, and we appreciate your input.

However, based on the recommendations of another reviewer, we have incorporated a more detailed mechanistic analysis. As a result, we have revised the original statement, which you will see in the updated manuscript.

Line 105-107: Please add reference to this statement.

Response:

Suggestion taken, thank you for your comment.

Line 126-127: Why the case study is novel?

Response:

Thank you for your comments.

In the original manuscript, Line 126-127: “In this study, the experimental approach is different from what has been used in other studies, where one case or long continuous time is simulated.”

So far, there is a substantial amount of research employing the WRF model to simulate wind fields, predominantly concentrating on one case study or use continuous long-run, large-scale regional assessments, or evaluations at single station. However, studies that assess WRF simulations through multiple case studies (Gómez et al., 2015) remain relatively scarce in the literature.

As we stated in the above response, to evaluate the performance of a PBL scheme over a large area, it is usually necessary to conduct multiple simulations covering different time periods and meteorological conditions. This approach provides a broader data sample, ensuring that the evaluation results have statistical significance, rather than being based on the outcomes of a single case.

Line 133: Please replace “*” with “×”.

Response:

Suggestion taken, thank you for your comment.

Line 167: Change to “model configuration” .

Response:

Thanks for pointing out this typo, we have made correction in the manuscript.

Table 1: What surface scheme was used?

Response:

Thank you for pointing this out. We will update Table 1 to include the surface layer scheme and the surface physic scheme used in our simulations to provide greater clarity and completeness of the methodology.

Table 1 PBL schemes and surface layer scheme

| PBL scheme | Surface layer scheme |
|---|---|
| YSU | Revised MM5 Monin-Obukhov scheme |
| MYJ | MYJ |
| MYNN2 | MYNN |
| QNSE | QNSE |

Line 189: Why is 6 m/s selected as a threshold to select the cases? How long does a case last a day?

Response:

Thanks for your comment regarding the selection of the 6 m/s threshold. The choice of this particular threshold is rooted in its significance to aviation operations, which is the primary focus of this study. Specifically, wind speeds around 6 m/s are critical for aircraft during takeoff and landing, where maintaining control and ensuring safety are of utmost importance.

By selecting 6 m/s as the threshold, we aimed to ensure that the cases analyzed in this study are directly relevant to these crucial phases of flight. This choice enables the study to provide findings that are not only scientifically rigorous but also practically applicable to the field of aviation weather, thereby enhancing the relevance of the results to real-world aviation scenarios.

We hope this explanation clarifies the rationale behind our decision, and we are open to any further suggestions you may have.

Line 208: Suggest using Bias which is more commonly used.

Response:

Suggestion taken, thank you for your comment.

Figure 2: What do the shading mean and dashed line mean?I assuming the dashed line is the threshold, which is 6 m/s in the text, but 5 m/s is showing in the figure. Please clarify. Again, from this figure, many of the cases were associated with diurnally varying winds while some cases were not. It would be interesting to see what synoptic scale/local conditions drive those wind patterns, and evaluate the PBL schemes' performance associated with those conditions.

Response:

Thank you for your insightful comments regarding Figure 2, and we'll add some clarification here.

We appreciate your observations and would like to clarify a few points.

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 .

It is important to note that in the Chengdu Plain areas of the western Sichuan Basin, the average wind speed and climate conditions are typically below 3 m/s. Nonetheless, it is possible for the 10-minute average wind to exceed 6 m/s induced by synoptic regimes.

We also appreciate your suggestion to investigate the synoptic scale and local conditions that drive the observed diurnal variability in wind patterns. We plan to conduct a detailed analysis of the errors associated with these 28 cases in relation to weather conditions in a future publication. This analysis will provide valuable insights for improving wind field predictions and refining PBL schemes in the region.

Once again, thank you for your valuable feedback. We will strive to incorporate your suggestions as we move forward with our research.

Line 286: Please list some examples for the studies.

Response:

Thank you for your valuable comment. We have added the references in our manuscript at line 286.

Line 290: Assuming the mean and median were calculated over the events. Please clarify.

Response:

Thank you for your valuable comment.

In our manuscript, Figure 4 contains box plots for 28 cases, each showing the RMSE (Root Mean Square Error), COR (Correlation), and Bias of simulations from four different PBL schemes compared to observations at 10-minute intervals. So, the mean and median correlation coefficients mentioned in the sentence were calculated across 28 events for each of the four PBL schemes.

A box plot is a powerful statistical tool used to visualize the distribution and central tendencies of data. The line in the middle of the box represents the median, which gives us insight into the central tendency of the data. For instance, the median of RMSE can indicate the typical size of the model's prediction error. By observing the shape of the box and the length of the whiskers, one can determine the symmetry of the data and whether there is skewness.

As shown in the figure, both the mean and median values for all schemes fall within the range of 0.4 to 0.6. This indicates  a tendency for the coefficients to cluster around this range across the events.

In figure 4, the box plots and heat maps are both given, box plots offer an in-depth analysis of individual cases, while heatmaps provide a global overview through color gradients. By combining these two, it is possible to simultaneously analyze both local details and global trends, helping to identify specific issues in individual cases and overall trends.According to you insightful comment,, we have revised this part to make it more clear.

We hope our response addresses your concerns. Please let us know if further clarification is needed.

Line 303-304: Please clarify that the "median" of the MYJ ME is 0.96 m/s.

Response:

Thank you for your comment.

The median value of 0.96 m/s for the MYJ scheme was obtained from Figure 4c (as shown in the picture given below), which presents a box plot of the mean errors (ME) calculated from 28 cases simulated by the four PBL schemes. The box plot provides a visual summary of the distribution of these errors, including the median, quartiles.

In Figure 4c, the median line within the MYJ box corresponds to a value of 0.96 m/s, which represents the central tendency of the mean errors across the 28 cases for this specific PBL scheme. This figure was generated using the error data from each case, where the box plot effectively illustrates the spread and central values of the mean errors.

We hope this explanation clarifies how the median value was derived, and we will revised this part in our manuscript. Please let us know if further details are needed.

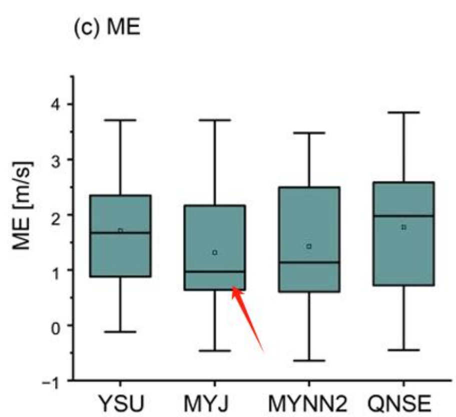


Figure1 Captured from Figure 4 c in original manuscript.

Line 307-319: This doesn't explain why MYJ is better in mean metrics while QNSE is better in variation. Since there is a suspicion that the performance of the PBL schemes differs under different stabilities, I'd suggest calculating the statistical metrics over different stabilities.

Response:

Thank you for your insightful comments and suggestions.

We agree that examining the performance of the PBL schemes under different stability conditions could provide valuable insights into their respective strengths and weaknesses. To address your suggestion, we plan to calculate the statistical metrics for four PBL schemes across different stability conditions. This analysis will help clarify the reasons behind the observed performance differences and enhance the robustness of our findings.We will incorporate this additional analysis into our revised manuscript.

Line 330: Change to “10 m”.

Response:

Suggestion taken, thank you for your comment.

Figure 8: Looks like some points belonging to Cluster 1 is more close to the centroid of Cluster 2?

Response:

Thank you for your observation regarding the K-means clustering results.

While it may visually seem that some points belong to another cluster when viewed in a simplified space, the assignment was determined by their position in the complete feature space, where the relationship between features can be more complex.

Additionally, it’s important to note that the centroids themselves are calculated based on the mean position of all points within each cluster, and slight overlaps or close proximities between clusters can occur, especially if the clusters are not well-separated.