Simulation performance of different planetary boundary layer schemes in WRF V4.3.1 on wind field over Sichuan Basin within “Gray zone” resolution

Qin Wang¹, Bo Zeng², Gong Chen², Yaoting Li¹

¹Civil Aviation Flight University of China, Guanghan, China
²Institute of Plateau Meteorology, CMA, Chengdu/Heavy Rain and Drought-Flood Disasters in Plateau and Basin Key Laboratory of Sichuan Province, Chengdu, China

Correspondence to: Bo Zeng (bozeng126@126.com)

Abstract. The topography of Sichuan Basin is complex and unique, high-resolution wind field simulation over this region is of great significance for meteorology, air quality, and wind energy utilization. In this study, Weather Research and Forecasting (WRF) model was used to investigate the performance of different planetary boundary layer (PBL) parameterization schemes on simulating surface wind fields over Sichuan Basin at a spatial resolution of 0.33km. The experiment is based on multi-case studies, so 28 near-surface wind events from 2021 to 2022 were selected, and a total of 112 sensitivity simulations were carried out by employing four commonly used PBL schemes: YSU, MYJ, MYNN2, and QNSE, and compared to observations. The results show that the wind direction which can be well reproduced, is not very sensitive to the PBL schemes as the wind speed shows. As for wind speed, the QNSE scheme had the best performance in reproducing the temporal variation out of the four schemes, while the MYJ scheme had the smallest model bias. Further cluster analysis demonstrates that the sensitivity of the PBL schemes is affected by diurnal variation and different circulation genesis. For instance, when the surface wind event caused by the southward movement of strong cold air and occurred during 6:00 and 8:00 (UTC), the variation and speed can be well reproduced by all four PBL schemes and the differences between them are tiny. However, the simulation of surface wind events mostly occurred during midnight and early morning, showing the characteristics of poor RMSE and good COR, while the simulation results of the evening-to-evening process and southerly wind process were opposite. Overall, the four schemes are better for surface wind simulations in daytime than at night. The results show the role of PBL schemes in wind field simulation under unstable weather conditions, and provide a valuable reference for further research in the study area and surrounding areas.

1 Introduction

Wind, as the most fundamental natural phenomenon in the atmosphere, poses not only hazards to civil aviation safety and maritime transportation during severe wind events (Manasseh and Middleton, 1999; Leung et al., 2022), but also impacts the dispersion of atmospheric pollutants directly near the surface, leading to adverse
effects on public health and the environment (Liu et al., 2020; Coccia, 2020; Yang and Shao, 2021). What’s more, wind energy has attracted increasing attention because of its non-polluting and renewable nature, but due to the random nature of wind speed, wind power generation is intermittent, which poses security and stability challenges for large-scale integration of wind energy into the power network (Liu et al., 2019; Kibona, 2020; Shi et al., 2021). Therefore, the accurate prediction of near-surface wind farms has become the key to ensure traffic safety, optimize wind energy utilization and evaluate air quality, and it is also an important scientific issue for disaster prevention and mitigation, economic benefits and human life and health.

Near-surface wind fields are influenced by a combination of various factors, including atmospheric thermal and dynamic conditions, topography, and underlying surface (Zhang et al., 2021). As a state-of-the-art mesoscale weather prediction model, the Weather Research Forecast (WRF) model can predict the fine-scale structure of near-surface wind fields by simulating the evolution of various physical processes in the atmosphere, which is significantly better than the prediction model based on statistics which lacking the description of thermodynamic processes. Furthermore, there are so many researches on the prediction and simulation of the refined characteristics of local wind field by using WRF model (Prieto-Herráez et al., 2020; Salfate et al., 2020; Xu et al., 2020; Tiesi et al., 2021; Wu et al., 2022; Yan et al., 2022; Mi et al., 2023). Although the simulation of near-surface wind fields involves the nonlinear interactions of various physical processes, the physical processes in the planetary boundary layer (PBL) play a direct role in influencing near-surface wind fields. As the interaction area between the atmosphere and the ground, the thermal and dynamic structure, the turbulent motion and mixing process in the boundary layer will directly affect the distribution of the near-surface wind field, so the simulation of the boundary layer by the model can directly affect the accuracy of the near-surface wind field (Chen et al., 2020).

In the mesoscale model, since the employed grid scales and time steps cannot explicitly represent the spatiotemporal scales which turbulent eddies operate on, the PBL parameterization scheme was used to express the effects of turbulent eddies (Dudhia, 2014). The latest version of WRF model provides more than 10 kinds of PBL parameterization schemes, the differences among them are mainly due to the different methods of dealing with the turbulence closure problem, which further leads to the different simulation result. Ma et al. (2014) conducted series sensitivity simulations on spring strong wind events in Xinjiang by using the schemes of YSU, MYJ, and ACM2, the results showed more downward transport of high-level momentum in the YSU scheme. Studies by Wang et al. (2010) and Zhang and Yin (2013) indicated that the ACM2 scheme performed well in simulating winter wind conditions in Lanzhou city and Huangshan, Anhui. In addition, more studies have shown that the MYJ scheme demonstrates the best simulation of near-surface wind speeds in the coastal areas of Fujian (Yang et al., 2014), while in regions such as western Neimenggu and Jiangsu, the YSU scheme exhibits the best forecasting performance for 10-meters wind speeds (Cui et al., 2018; Li et al., 2018). In typical mountainous terrain of Guizhou, the ACM2 scheme performs better in simulating
near-surface wind speeds at 70m height compared to the MYJ and YSU schemes (Mu et al., 2017). From these studies, it is evident that WRF has obvious regional performance regarding the PBL scheme. Therefore, without considering the nonlinear amplification of initial condition errors and the inaccuracy of numerical models, the reliable wind speed prediction for specific areas is still challenging and worthy of further study.

Sichuan Basin is one of the four major basins in China, it is bordered by the Qinghai-Tibet Plateau to the west, the Daba Mountains to the north, the Wushan Mountains to the east, and the Yunnan-guizhou Plateau to the south. Because of the complex terrain of its surrounding areas, the local atmospheric circulation is also complex and unique (Yu et al., 2020), the weather here is characterized by low wind speed, low sunshine and high humidity throughout the year, therefore it is also one of the four major haze areas in China (Li et al., 2021). Under the unique terrain of the Sichuan Basin, it is difficult to determine whether cold air from mid to high latitudes can bypass the Qinghai-Tibet Plateau and then cross the Qinling Mountains to enter the basin. Besides, the basin effect makes it easier to form an inversion structure close to the surface and stabilizing the atmosphere (Gao et al., 2016; Feng et al., 2023). These factors make it one of the regions with the poorest wind forecasting performance in China (Pan et al., 2021; Xiang et al., 2023). Therefore, wind is not still as wildly studied as temperature and precipitation in Sichuan Basin, and the main focus of wind simulation is about the pollutant diffusion under stable weather conditions.

As is known, the interaction between the surface and atmosphere, as well as the characteristics of turbulent motion over the basin terrain, differ from that over plains and plateau areas. However, there is no detailed evaluation for the performance of PBL schemes in the near-surface wind field over the Sichuan Basin. Thus, the present study aims at evaluating the performance of four PBL schemes under the windy conditions over the Sichuan Basin. In the model set-up, a horizontal resolution of 0.3km was used for research, which is a major challenge in such region, because the spatial resolution is in the range of 0.1-1km, which is often referred as "gray zone" in numerical forecasting (Liu et al., 2018; Yu et al., 2022). As suggested by many studies, the spatial resolution in "gray zone", is too finely detailed with regarding to the mesoscale turbulence parameterization scheme, and too coarse for the Large Eddy Simulation (LES) scheme to analyze turbulent vortices (Shin and Hong, 2015; Honnert et al., 2016). So far, the impact of different PBL schemes under the spatial resolution of "gray zone" is still uncertain. Hence, a total of 28 wind events is simulated with a purpose of getting a reliable evaluation, and the study is based on a case study approach, rather than on continuous simulations. In general, this study not only has important significance for improving the wind field forecast in this region, but also provides a scientific basis for the further improvement and development of PBL scheme.

2 Data and Method
2.1 Data and experimental design

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. In this study, a total of 28 historical surface wind events was simulated by running WRF-ARW (version 4.3.1). We choose Guanghan Airport as the representative of Sichuan Basin, and the 28 discontinuous windy days, with a criteria of the maximum wind speed greater than 6 m s⁻¹ are simulated.

The simulation domain consists of four two-way nested domains of resolutions 9 km, 3 km, 1 km and 0.33 km, with 105*105, 103*103, 103*103 and 103*103 grids, respectively, and the vertical resolution is 45 for all domains. Figure 1 presents the domain set-up. As can be seen from Fig. 1 (a), the outermost domain (D01) covers the western Sichuan Plateau and the northern Qinlin Mountains. The surrounding mountains are mostly between 1,000 and 3,000 meters above sea level, while the basin is between 250 and 750 meters. Due to the complex topography in the upstream region, the influence of cold air on the Sichuan Basin is variable, and the wind simulation is very difficult. In the western domain 2, the elevation gradually decreases from 2000 to 500 meters, with a topography that is higher in the western and northern parts, and lower in the eastern and southern parts. In the domain 4, the transitional zone from plateau to basin is avoided. This area is located in the northern part of Chengdu Plain, and the simulation center is set at Guanghan Airport (104.32° E, 30.93° N). Additionally, Guanghan Airport is located at the western foothills of the Longquan Mountains, only 10 km away.

Given the complex terrain in study region and the high resolution of model design, the input of land surface data is particularly important, and its accuracy will directly affect the simulation of land surface processes and atmospheric boundary layer characteristics (Qi et al., 2021). Therefore, we replaced the terrain data of the 4-layer nested area with the 90 m resolution terrain data from the southwest region of Shuttle Radar Topography Mission (SRTM3).
Figure 1. Configurations of (a) four-layer nesting domains (D01-D04) in WRF and the (b) study area. The spatial resolutions are 9, 3, 1 and 0.3 km, for domains D1 to D4, respectively. The figure depicts the actual orography implemented in the experiments.

To evaluate the model’s ability in different PBL schemes, the observed wind fields at 10 meters high at Guanghan Airport station is used. The hourly reanalysis dataset ERA5 with a horizontal resolution of 0.25° and 38 vertical levels, is used to provide the initial and boundary conditions for WRF simulations, which are updated every 3 hours when input into the model. Each event is simulated using four different PBL parametrisation schemes. Thus, a total of 112 simulations are carried out. Each simulation spans 24 hours, with the corresponding high winds in the middle of the simulation, and discarding a spin-up period of 3 hours, the other model configuration is summarised in Table 1.

Table 1. Configures of the physical scheme in WRF simulation.

<table>
<thead>
<tr>
<th>Parameterizations</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro-physical scheme</td>
<td>WSM 3-class graupel scheme (same for each domain)</td>
</tr>
<tr>
<td>Longwave radiative scheme</td>
<td>RRTM shortwave (same for each domain)</td>
</tr>
<tr>
<td>Shortwave radiative scheme</td>
<td>Dudhia shortwave (same for each domain)</td>
</tr>
<tr>
<td>Cumulus convection scheme</td>
<td>Kain-Fritsch for the outermost domain, and closed in other 3-layers</td>
</tr>
</tbody>
</table>

2.2 PBL Schemes

There are more than 10 PBL parameterization schemes in WRF-V4.3.1, but four commonly used PBL schemes were selected for this study, which are YSU (Yonsei
University) scheme (Hong et. al., 2006), MYJ (Mellor-Yamada-Janjic) scheme (Janjié, 1990), MYNN2 (Mellor-Yamada-Nakanishi-Niino Level 2) scheme (Nakanishi and Niino, 2009) and QNSE (Quasi-Normal Scale Elimination) scheme (Sukoriansky and Galperin, 2006). Among them, YSU is a non-local, first-order closure scheme that represents entrainment at the top of the PBL explicitly, while the rest are local closure schemes, detail characteristics can be seen in Table 2. The surface layer scheme in the experiment is matched with each PBL scheme.

Table 2. Advantages description for the four PBL schemes used in WRF model.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>YSU</td>
<td>1st-order closure scheme that is widely utilized for its robust representation of turbulence closure processes (Hong et. al., 2006).</td>
</tr>
<tr>
<td>MYJ</td>
<td>A 1.5-order closure scheme that is known for its effectiveness in capturing vertical mixing processes (Janjié, 1990).</td>
</tr>
<tr>
<td>QNSE</td>
<td>A 1.5-order closure scheme that improves the simulation of sub-grid scale turbulence (Nakanishi and Niino, 2009).</td>
</tr>
<tr>
<td>MYNN2</td>
<td>A 1.5-order turbulence closure scheme that accounts for both turbulent and non-turbulent mixing processes in the atmosphere (Sukoriansky and Galperin, 2006).</td>
</tr>
</tbody>
</table>

2.3 Statistical metrics for validation

As suggested by Wang et al. (2017), different sky conditions and atmospheric stability will affect the simulation of wind fields. So, in order to accurately evaluate the sensitivity of four PBL schemes to the near surface wind field in the western Sichuan Basin on the east side of the Qinghai Tibet Plateau, 28 surface wind cases with an 10 minutes averaged wind speed greater than 6 m s⁻¹ from 2021 to 2022 were selected for simulation, and the result is evaluated separately through different circulation patterns and K-means clustering analysis method. The main statistical metric used includes:

Root Mean Square Error (RMSE), which is the square root of the average of the squared differences between the simulated and observed values. RMSE is a commonly used metric in model evaluation, assigning higher weight to cases with larger simulation errors:

\[
RMSE = \sqrt{\frac{\sum (O_i - S_i)^2}{N}}
\]  

where \(N\) is the total number of samples, \(O_i\) represents the observed surface wind, and \(S_i\) denotes the simulated surface wind, measured in m s⁻¹.

Correlation Coefficient (COR) is an indicator that measures the strength and direction of the linear relationship between simulation and observation. By analyzing
COR, the consistency between simulation results and observation results can be evaluated, and the corresponding PBL scheme can accurately capture the variation relationship of ground wind speed:

\[
COR = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(x_o - \bar{x})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{N} (x_o - \bar{x})^2}}
\]  

(2)

where \(N\) is the total number of samples, \(x_o\) represents the observed values, and \(x_j\) denotes the simulated values.

Mean Error (ME) refers to the average difference between simulated and observed values, reflecting the overall bias of the simulation results. If ME is close to 0, it indicates that the simulation results have good accuracy at the average level. The calculation formula is as follows:

\[
ME = \frac{1}{N} \sum_{i=1}^{N} (x_j - x_o)
\]

(3)

The Weibull distribution is a probability function used to describe the distribution of wind speed. The expression for the Weibull distribution probability density function of wind speed \(v\) is:

\[
f(v) = \frac{\kappa}{\lambda} (\frac{v}{\lambda})^{\kappa-1} \exp\left[-\left(\frac{v}{\lambda}\right)^\kappa\right]
\]

(4)

where \(\kappa\) is the shape parameter, a dimensionless parameter, and \(\lambda\) is the scale factor, measured in m s\(^{-1}\). These two parameters can be calculated using the following formulas:

\[
\kappa = \frac{\sigma}{\mu}
\]

(5)

\[
\lambda = \frac{\mu}{(0.568 + 0.434\kappa)^{1/\kappa}}
\]

(6)

where \(\sigma\) and \(\mu\) represent the standard deviation and mean value of the wind speed, respectively.

3. Overview of historical cases and evaluation of simulation results

3.1 Summary of 28 surface wind events

Since the experiment approach is concerned about multi-cases simulation in this study, it is necessary to understand the characteristics of these cases, such as the temporal variation, the peak time and dominated circulation, which can help to classify them and evaluate their simulation performance separately in the following analysis.

Therefore, Table 3 gives the detail information based on wind filed every 10 minutes. It is shown that out of the 28 surface wind events participating in the simulation, 24 were northerly events, accounting for 85% of the total. The events in
which the maximum wind is above 8 m s\(^{-1}\) accounts for 18%, and the events of 5-7 m s\(^{-1}\) accounts for 82%. Meanwhile, the wind direction corresponding to the peak time was distributed between 350° -50°, with northeasterly winds between 0-50° being the most common. Additionally, the left are 4 southerly winds cases, all of which appear to occur in summer or early autumn. As for the dominated atmospheric circulation of each event, it is shown that most of the wind events were mainly caused by cold air, only little were associated with deep convection. Influenced by this, the spring (March-May) process accounted for the most, accounting for 46%, followed by summer and autumn, both accounting for 25%. In terms of the peak time, 60% of the simulated cases appear to concentrate on 05:00 - 09:00 UTC and 10:00 - 14:00 UTC at night, then followed by 15:00 - 19:00 UTC, and there are a total of 6 events occurred at 20:00 - 23:00 UTC and 00:00 - 04:00 UTC, accounting for 21%.

Besides, the observed wind rose and time series of wind speed are presented in Fig.2. It is indicated that during these periods, the near-surface wind is mostly from a northwesterly-to-northeasterly direction.

<table>
<thead>
<tr>
<th>Event ID</th>
<th>Date yyyy-mm-dd</th>
<th>Maximum wind speed (m s(^{-1})) /direction(°)</th>
<th>Maximum wind time hh:mm</th>
<th>Circulation classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2021-03-17</td>
<td>6.0/350°</td>
<td>09:40</td>
<td>Cold air</td>
</tr>
<tr>
<td>2</td>
<td>2021-03-24</td>
<td>6.8/350°</td>
<td>08:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>3</td>
<td>2021-03-30</td>
<td>6.1/90°</td>
<td>09:50</td>
<td>Cold air</td>
</tr>
<tr>
<td>4</td>
<td>2021-03-31</td>
<td>6.4/45°</td>
<td>09:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>5</td>
<td>2021-04-23</td>
<td>6.3/47°</td>
<td>11:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>6</td>
<td>2021-04-25</td>
<td>7.0/70°</td>
<td>08:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>7</td>
<td>2021-04-27</td>
<td>8.3/18°</td>
<td>11:10</td>
<td>Cold air</td>
</tr>
<tr>
<td>8</td>
<td>2021-06-16</td>
<td>6.9/46°</td>
<td>07:40</td>
<td>Cold air</td>
</tr>
<tr>
<td>9</td>
<td>2021-07-21</td>
<td>7.1/158°</td>
<td>06:20</td>
<td>Deep convection</td>
</tr>
<tr>
<td>10</td>
<td>2021-08-22</td>
<td>8.0/47°</td>
<td>03:10</td>
<td>Cold air</td>
</tr>
<tr>
<td>11</td>
<td>2021-08-25</td>
<td>6.1/33°</td>
<td>06:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>12</td>
<td>2021-09-15</td>
<td>6.6/50°</td>
<td>15:20</td>
<td>Cold air</td>
</tr>
<tr>
<td>13</td>
<td>2021-09-19</td>
<td>6.0/183°</td>
<td>08:00</td>
<td>Deep convection</td>
</tr>
<tr>
<td>14</td>
<td>2021-09-25</td>
<td>6.1/54°</td>
<td>05:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>15</td>
<td>2021-10-01</td>
<td>6.0/332°</td>
<td>14:40</td>
<td>Cold air</td>
</tr>
<tr>
<td>16</td>
<td>2021-10-04</td>
<td>7.3/45°</td>
<td>03:30</td>
<td>Cold air</td>
</tr>
<tr>
<td>17</td>
<td>2021-11-06</td>
<td>9.6/51°</td>
<td>12:00</td>
<td>Cold air</td>
</tr>
<tr>
<td>18</td>
<td>2021-12-25</td>
<td>6.0/46°</td>
<td>20:50</td>
<td>Cold air</td>
</tr>
<tr>
<td>19</td>
<td>2022-03-19</td>
<td>7.9/10°</td>
<td>22:10</td>
<td>Cold air</td>
</tr>
<tr>
<td>20</td>
<td>2022-03-30</td>
<td>8.3/43°</td>
<td>12:20</td>
<td>Cold air</td>
</tr>
<tr>
<td>21</td>
<td>2022-04-14</td>
<td>6.0/27°</td>
<td>18:40</td>
<td>Cold air</td>
</tr>
<tr>
<td>22</td>
<td>2022-04-27</td>
<td>8.3/50°</td>
<td>17:00</td>
<td>Cold air</td>
</tr>
</tbody>
</table>
Figure 2. Observed wind rose chart (a) and time series of hourly wind speed (b) for all the 28 near-surface wind events listed in Table 3. For the wind rose, the circles represent the relative frequency (%), and the colors represent wind speed.

3.2 Overall simulation performance of 28 wind events

First, the performance of the model in different PBL schemes is assessed with
respect to wind direction. Thereby, the simulated wind rose of four PBL schemes are given in Fig. 3. By comparing with the observation (Fig. 2), it is found that four PBL schemes can reproduce the distribution of wind direction. Specifically, the simulated wind directions are basically distributed in NNW, N, NNE, NE and ENE, reproducing the characteristics of highly concentrating on NNE and NE. Besides, it is also shown that all the PBL schemes tend to overestimate the prevailing wind direction and significantly underestimate the NNW wind, indicating an clockwise bias which may be related to the plateau topography with steep terrain in the northwest and west. Therefore, it is concluded that the wind direction of the near-surface wind field in Sichuan Basin is very insensitive to the selected PBL schemes.

However, there are still some differences in wind direction simulations among four PBL schemes. In MYJ scheme, the frequency of NNE wind is higher than NE wind, which is consistent with the observations. Moreover, the frequencies of N wind and NE wind are closer to the observations. Therefore, MYJ has the best simulation of wind direction. The wind direction distribution simulated by the MYNN2 scheme is very close to QNSE scheme, but due to the worse performance in simulating NNW wind and the larger frequency of simulated NNE and NE wind, MYNN2 scheme is the worst among the four schemes. In general, for wind fields with weather processes passing through, more attention is paid to the simulation of wind speed. So, we will focus on the performance of wind speed next.
Figure 3. Same as in Fig. 2a, but for the simulated near-surface wind field corresponding to the four PBL schemes, the circles represent the relative frequency (%), and the colors represent wind speed.

In fact, by comparing Fig. 2 and Fig. 3, it seems that all the four PBL schemes exhibit obvious exaggeration of wind speed, which is also shown in other numerous studies. But, what are the specific simulation characteristics of these commonly used PBL schemes in the Sichuan Basin? To further assess the advantages and disadvantages of each scheme in simulating surface wind speed, three statistics of COR, RMSE, and ME were selected for comprehensive evaluation of the simulation results, as shown in Fig. 4. In terms of COR, the mean and median correlation coefficients between simulation of the four schemes and observation are all between 0.4-0.6, and the median is above the mean value, indicating that the correlation coefficients are all negatively skewed distribution, that is, the correlation coefficients between simulated and observed wind speed are higher than the mean value in most cases, but very poor in some certain cases. It is further illustrated by the heat map displayed in Fig. 4d, where cases No. 3, 11 and 20 demonstrate correlation coefficients below 0. In contrast, QNSE shows the best mean correlation coefficient of 0.6, suggesting the best performance in reproducing the temporal variation of observed wind speed in most cases.
Although there is little difference between the simulated and the observed wind speed in the RMSE and ME, it is ingesting that MYJ scheme has the smallest RMSE while QNSE has the largest. What’s more, in MYJ scheme not only the ME is the lowest (value is 0.96 m s\(^{-1}\)), but also the difference between the median and mean values is significant, which suggests that most of the wind speed bias produced by MYJ are actually below 0.96 m s\(^{-1}\). Therefore, it is demonstrated that the bias of near-surface wind speed produced by MYJ scheme in Sichuan Basin is the smallest based on the multiple cases simulation. The main reason for this may be associated with the basin topography, because the boundary layer is in stable condition in most time, the turbulence is mainly generated and maintained by wind shear, so that the situation showing strong locality. Therefore, the simulation error obtained by MYJ scheme is the smallest in this stable and weakly stable boundary layer, which is consistent with the research results of Zhang et al. (2012). Besides, the result that QNSE scheme has the best performance on capturing the temporal variation of wind speed, maybe because that QNSE scheme improves simulation of sub-grid scale turbulence, and considers more complex and detailed physical processes. Under stable atmospheric stratification, QNSE adopted k-ε model developed from turbulent spectral closure model, while under the unstable situation, the method of MYJ scheme is used , so QNSE scheme has more advantages in the simulation of wind speed variation trend.

![Figure 4](https://doi.org/10.5194/egusphere-2024-1532)

**Figure 4.** Different performance metrics for the comparison of observed and simulated near-surface wind speed for 28 events. Box plots shows the overall characteristics of COR, RMSE and ME, and heat-map gives details for certain case. The box represents the metrics range from first quartile to third quartile ,and the line inside the box represents the median, while the empty square represents the mean.

### 3.3 Differences of wind velocity segments and diurnal variations simulated by
four PBL schemes

Figure 5 shows the histogram of the frequency distribution of the observed and simulated wind speed at 10 meters high of the airport and the corresponding Weibull distribution fitting curve. As can be seen from the figure, the observed mode of wind speed is leftward, which is mainly due to the fact that the high wind speed sections are very concentrated and have a low frequency during 28 wind events. The corresponding Weibull fitting $\kappa$ of the observation is 1.79, and the $\kappa$ value produced by QNSE is the closest to it, while the fitting $\kappa$ values of the other three schemes are all larger. The corresponding Weibull fitting $\lambda$ values of the four parametric schemes are all larger than the observation (3.29 m s$^{-1}$).

![Weibull distribution fitting](https://doi.org/10.5194/egusphere-2024-1532)

**Figure 5.** The frequency and Weibull fitting of the observed and four PBL schemes simulated wind speed of 28 wind events.

When wind speed below 3 m s$^{-1}$, none of the PBL scheme has a good performance. Moreover, the lower the wind speed, the greater the bias. In the range of wind speed greater than 3 m s$^{-1}$ and less than 5 m s$^{-1}$, different PBL schemes show significant differences compared with observations. Specifically for wind speeds during the 3-4 m s$^{-1}$, the simulation results of the MYJ scheme are closest to the observations, followed by MYNN2. For wind speeds during the 4-5 m s$^{-1}$, YSU and MYJ simulations are closer to the observations, indicating better performance in this wind speed range. All schemes tend to overestimate when wind speed above 5 m s$^{-1}$.

Figure 6 further provides the deviations between the observed and simulated wind speed of four PBL schemes in different wind speed ranges. As can be seen, the performance of four PBL schemes differ greatly with the increase of wind speed, and
the wind speed deviation of the same PBL scheme also increases. For the wind speed below 3 m s\(^{-1}\), the simulated wind of each PBL scheme are about 1.5-2 m s\(^{-1}\) higher than the observation. In terms of mean values, the MYJ scheme exhibits relatively smaller deviations for wind speeds below 7 m s\(^{-1}\), while the MYNN2 scheme demonstrates the smallest deviation in simulation for wind speeds above 7 m s\(^{-1}\).

In general, the fitting curve of QNSE scheme is most close to the observation, and the \(\lambda\) value is slightly to the right than the mode. The mode of four schemes are to the right relative compared with the observation, tending to a normal distribution.

![Wind speed errors of four PBL schemes for 28 wind events](https://doi.org/10.5194/egusphere-2024-1532)

**Figure 6.** Wind speed errors of four PBL schemes in different wind speed segments for 28 wind events.

The variation of near-surface wind field is easily affected by surface characteristics, especially ground heating. When the weather background is fixed, the change of local thermal characteristics in a day will inevitably affect the near-surface wind field. Therefore, there will be significant differences in the wind field simulation during different time periods between different PBL schemes. According to the relationship between world time and local time, the daytime in the text corresponds to world time 00:00 - 10:00, and the nighttime refers to world time 11:00 - 23:00. Figure 7 presents the diurnal variation characteristics of wind speed deviations simulated by the four PBL schemes in the WRF model through box plots.

In terms of the mean, the performance of wind speed of each scheme is better in the daytime than in the night. The deviation is the highest at 18:00 and 19:00 UTC, which means that the strong wind occurring at this time cannot be well simulated. As for YSU scheme, the simulation ability is the best at noon, while MYJ simulated well at noon and evening, and MYNN2 simulated in the evening. QNSE has little difference in the simulation of 28 wind cases in the daytime and a large difference in...
the night, indicating that QNSE scheme is stable in the simulation of strong wind in the daytime, but unstable in the night, with a large variation of simulation performance. However, in general, QNSE scheme has the best simulation ability at noon.

**Figure 7.** Diurnal variation of wind speed errors corresponding to four PBL schemes.

### 3.4 K-means clustering analysis and performance in different types of events

From the previous analysis, it is known that as the horizontal grid spacing of 0.33 km is within the PBL gray zone resolution, QNSE scheme can better capture the trend of surface wind events over Sichuan Basin, while the bias produced by MYJ scheme is the minimum. The results also show the difference in different wind speed segment and different time in this region, but it is not significant. At the same time, Previous studies have indicated that the simulation of meteorological elements within the boundary layer is influenced by meteorological conditions such as circulation patterns. Therefore, it is necessary to further classify and analyze these 28 cases to understand the specific performance of PBL scheme in simulating surface wind events in Sichuan basin.

The K-means cluster method is used to divide the simulation results of 28 surface wind events into three categories, as presented in Fig. 8. The RMSE of the cluster center of the first class is 1.9 m s\(^{-1}\), and the COR is 0.2. A total of 10 events belong to this class, presenting the class with good RMSE but poor COR. At the cluster center of the second class, the RMSE is 2.85 m s\(^{-1}\), and the COR is 0.6. A total of 12 events belong to this class, characterized by good COR but large bias. At last, the left 6 events belong to the third category, in which both RMSE and COR are very good for simulation, and the cluster center has the RMSE of 1.25 m s\(^{-1}\), and COR of...
0.76. Furthermore, it is shown that among the three types of events, the QNSE scheme has the best simulation correlation coefficient, while the MYJ scheme has the smallest wind speed simulation error. This is consistent with the unclassified results, indicating that QNSE and MYJ schemes are relatively stable and reliable choices for the surface wind simulation in Sichuan Basin with model grid resolution of 0.3 km.

Figure 8. Scatter plot of K-means cluster analysis, the red cross symbol represents the cluster center.

According to the K-means analysis, it is found that different PBL schemes are very sensitive to the diurnal variation and circulation background of surface wind in the simulation of surface wind speed in the Basin, though there is no obvious seasonal difference. Figure 9 shows the RMSE and COR heat-maps of three types of events after cluster analysis, and peak time of gale is specially marked. It can be seen that the four PBL schemes have the least sensitivity to the event of class III. This kind of event is characterized by that the gale period basically occurs between 06:00 and 08:00 UTC, which is also the period with the highest surface temperature and the most unstable atmospheric stratification in the region. What's more, in the events of class III, except for one thunderstorm gale event, the rest are all typical strong cold air induced surface wind processes, which indicates that the four PBL schemes have the good performance in simulating the typical strong cold air wind event occurred in the afternoon. As shown in Figure 10, the RMSE ranges from 0.21 m s\(^{-1}\) to 0.96 m s\(^{-1}\), and the COR ranges from 0.05 to 0.19, with only one case having a difference of 0.3, which means that there is little difference between four PBL schemes.
Figure 9. Heat-map about the RMSE (numbers) and COR (coloring) of four PBL schemes for 28 near-surface wind simulations according to the cluster analysis. The information in the right column is gale moment (numbers) and classification label (coloring).
Figure 10. Box plots of the maximum differences during four PBL schemes in three types of events, with the green dotted line as the mean, the orange solid line as the median, and the circle as the outlier.

The most obvious differences among the four PBL schemes are mainly in the events of class I and II. Except for one southerly gale event belonging to class III, the other southerly wind events are classified into class I, indicating that the four PBL schemes often have good RMSE and poor COR for southerly wind events caused by convection in Sichuan Basin. In Figure 9, it is shown that in class I, the maximum wind speed often occurred in the two periods of 10:00 - 11:00 UTC and 15:00 - 16:00 UTC, and only two cases occurred at 06:00 - 07:00 UTC. The period of 10:00 - 16:00 UTC is the period when the atmospheric stratification in the basin changes from unstable to stable, and it is also the period when the inversion layer is established. In this kind of events, the difference between the maximum and minimum RMSE and COR obtained by different PBL schemes is as large as 1.43 m s$^{-1}$ and 0.58.

The simulation events of class II show the most significant differences among the four PBL schemes, and the characteristics such as gale occurrence time are significantly different from those in class I and class III. It is observed that the four PBL schemes often exhibit high CORR and high RMSE for surface wind events occurring in the early morning (17:00-22:00 UTC) and early afternoon (03:00-05:00 UTC), and these surface wind events are concentrated in dry and cold air scenarios. In this type the maximum difference between different PBL schemes can reach 1.49 m

435
s$^{-1}$ and 0.76. In addition, Fig. 10 shows that the differences between different PBL schemes in class I and class II events in the daytime are relatively small, while there are greater differences at night. Meanwhile, in class III, the RMSE performance at night is better than that in the daytime, but the COR is worse than that in the daytime. Therefore, it can be concluded that there are obvious and diversified differences among the simulation results shown by various PBL schemes under different types of surface wind events.

4 Summary and conclusions

In this study, a horizontal resolution of 0.33 km which is within the PBL gray zone resolution is employed to investigate the performance of four commonly used PBL schemes on near-surface wind simulation over the Sichuan Basin. In China, the near-surface wind prediction over Sichuan Basin has always a low score, and the main focus of wind simulation is about the pollutant diffusion under stable weather conditions at a horizontal resolution equals or greater than 1 km. Thus, we chose the site of Guanghan Airport as the representation, and conducted a total of 112 WRF sensitivity experiments, specifically focusing on 28 events with near-surface winds exceeding 6 m $s^{-1}$ by varying the PBL scheme, and assessed the impact of different PBL schemes on wind speed and direction simulations. Subsequent analyses considered factors such as diurnal variation of surface wind processes and circulation background to gain further understanding of their influence on model sensitivity. Therefore, the findings of our study offer the valuable insights in this region.

From our evaluation and analysis, the sensitivity of surface wind direction over Sichuan Basin to the four commonly used PBL schemes is very low, and the performance of MYNN2 is the worst when simulating the surface wind direction, while the other three schemes are generally consistent with the observations, and the MYJ scheme is the best for simulating NNE and NE winds. Our findings on wind direction is agree with the finding in many other researches (Gómez-Navarro et al., 2015; Tan et al., 2017; Shen and Du, 2023).

Generally speaking, no scheme can simulate the trend and wind speed of surface wind events well at the same time, which is also mentioned by Cohen et al. (2015). However, the 1.5-order QNSE local closure approximation scheme appears to be the best for the temporal variation, while MYJ is the scheme with smallest simulation error on wind speed. As the metrics RMSE and ME shows the similar characteristics, K-means cluster analysis is employed based on the COR and RMSE ,and the simulation results are divided into three categories. The first category of events showed poor correlation but small RMSE; the second category of events showed high correlation but large RMSE; the third category of events showed high correlation coefficient and small RMSE. Further analysis found that the four PBL schemes can simulate the ground wind events caused by the typical strong cold air (occurring at 6:00-8:00 UTC), and there is little difference between them. For the surface wind events occurring in the midnight to early morning, they are mainly concentrated in the second category; while the evening to night and the southerly wind process are
mainly concentrated in the first category. Therefore, multi-cases studies and K-means clustering analysis gives us the hint that the simulation performance of the PBL schemes mainly depends on the prevailing weather conditions of each case, such as circulation backgrounds and the time of surface wind events.

**Code and data availability.** The Weather Research and Forecasting (WRF) model version 4.3.1 used in this study is freely available online and can be downloaded from https://www2.mmm.ucar.edu/wrf/users/download/get_source.html (Skamarock et al., 2008). The ERA5 data are available from ECMWF (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, last access: 23 June 2023, DOI: https://doi.org/10.24381/cds.bd0915c6, Hersbach et al., 2018). The observations and model output upon which this work is based are available from Zenodo (https://doi.org/10.5281/zenodo.11328605, Wang et al., 2024), and the data can also be obtained from pwd@cafuc.end.cn.

**Author contributions.** QW conceptualized the study and conducted the simulations. BZ, YY and GC analyzed the model results, and QW and BZ contributed to the interpretations. The original draft of the paper was written by QW, and all the authors took part in the edition and revision of it.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Acknowledgments.** The authors acknowledge NCAR for the WRF model and ECMWF for the ERA5 reanalysis datasets.

**Financial support.** This research has been supported by the National Key Research and Development Program of China (grant no. 2022YFC3003902 and 2023YFC3007502), the National Natural Science Foundation of China (grant no. 42030611), and Sichuan Science and Technology Program (grant no. 2022NSFSC0821 and 2023NSFSC0904).

**References**


Wang, C. -G., Shen, Y. -J., Luo, F., Cao, L., Yan, J. -D., and Jiang, H. -M.: Comparison and analysis of several planetary boundary layer schemes in WRF model between clear and overcast days, Chinese J. Geophys., 60, 141-153, doi:


Zhang, L., Xin, J., Yin, Y., Chang, W., Xue, M., Jia, D., and Ma, Y.: Understanding the major impact of planetary boundary layer schemes on simulation of vertical wind