Chief editor:

Dear authors,

in my role as Executive editor of GMD, I would like to bring to your attention our Editorial version 1.2: https://www.geosci-model-dev.net/12/2215/2019/

This highlights some requirements of papers published in GMD, which is also available on the GMD website in the 'Manuscript Types' section: http://www.geoscientific-model-development.net/submission/manuscript_types.html

In particular, please note that for your paper, the following requirements have not been met in the Discussions paper:

"The main paper must give the model name and version number (or other unique identifier) in the title."

"If the model development relates to a single model then the model name and the version number must be included in the title of the paper. If the main intention of an article is to make a general (i.e. model independent) statement about the usefulness of a new development, but the usefulness is shown with the help of one specific model, the model name and version number must be stated in the title. The title could have a form such as, "Title outlining amazing generic advance: a case study with Model XXX (version Y)"."

As you use WRF-Fire for your analysis, please add something like "a case study using WRF-FIRE version x.y" to the title of your manuscript upon submission of the revised version to GMD.

Thanks for your kindly suggestion, the new title is "Dynamical Linkages Between Planetary Boundary Layer Schemes and Wildfire Spread Processes: a case study using WRF-Fire version 4.6".

Reviewer 1:

This manuscript evaluates the performance of five PBL schemes within the WRF-Fire coupled atmosphere-fire model, utilizing a case study simulation of a forest fire in a mountainous region of China. The study assesses simulation results through near-surface variables (2-m temperature and 10-m wind speed) and vertical turbulence structures (TKE and PBL height). The manuscript is well-written with high-quality figures, and the analysis holds scientific value. However, the manuscript appears to lack content regarding model development, suggesting it might be more suitable for submission to other journals.

We thank the reviewer for the suggestion. The primary objective of this study is to evaluate the applicability of the coupled atmosphere-fire model in complex forested terrain. This evaluation relies heavily on hourly data from automatic weather stations located in the immediate vicinity of the fire, including one emergency mobile station (covering two positions) and five automatic observation stations (comprising two upwind stations, two downwind stations, and one high-altitude station). Notably, the mobile station operated within a 3 km range; the nearest automatic station (Chaoyang Primary School) is less than 3 km in linear distance from the mobile station, and the farthest station (Batang) is less than 9 km away. Consequently, these observational data provide a robust basis for evaluating the forest fire simulation results. The value of this study lies in the application of rare observational data acquired in close proximity to the fire ground and the assessment of how heat released by the fire impacts local circulation, providing significant scientific value for the future predictive application of coupled models.

Issues to be Addressed

1. It is currently unclear what the role of including the fire module in these simulations is, as fire spread is not compared or discussed. The main goal of the paper appears to

be evaluating the performance of PBL schemes under "perturbed" conditions (i.e., induced by fire heat perturbations). The manuscript does not compare the effects of different PBL schemes on the final fire spread, nor does it quantify the degree of fire perturbation on the PBL (e.g., TKE from coupled simulation minus TKE from uncoupled simulation). Including results related to simulated fire behavior would make a significant contribution to the existing literature on coupled fire modeling.

We thank the reviewer for pointing out the important issue regarding the role of the fire module and the depth of the fire–boundary layer coupling analysis. We have supplemented and clarified this issue in the revised manuscript as follows:

First, the purpose of this study is to evaluate the applicability and usability of the coupled atmosphere-fire model WRF-Fire v4.6 in complex terrain regions. Currently, evaluations of this coupled model in complex mountainous areas are scarce. We utilize the strong, spatially heterogeneous thermal forcing generated by a real forest fire event as the physical background to test the differences in performance of different PBL schemes under non-typical, strongly perturbed conditions. Compared to idealized heating or no-fire scenarios, WRF-Fire provides a more realistic spatiotemporal distribution of heat release, making the response differences of various PBL schemes under the same external fire forcing comparable and practically significant.

Regarding the issue that fire behavior was not directly shown in the original manuscript, the revised manuscript has added the burned area results simulated by WRF-Fire and performed a quantitative comparison with the MODIS burned area product. This analysis shows that the spatial scale of the simulated fire ground is consistent with observations, thereby verifying the rationality of the fire–atmosphere coupled simulation at the regional scale and avoiding the treatment of the fire module merely as an "idealized heat source." The simulated burned area varied considerably (47.38%–92.82%) across the five boundary layer schemes tested (MYJ, MYNN2, MYNN3, BouLac, and UW). The MYNN3 scheme demonstrated superior performance, providing the most accurate estimate.

Regarding quantifying the degree of fire perturbation on the PBL (such as the difference between fire-on and fire-off simulations), we agree that this analysis is physically meaningful. In fact, the authors have systematically carried out fire-on/fire-off control experiments in previous research targeting fires in the forest-grassland transition zone (see: https://www.mdpi.com/2571-6255/6/11/443), evaluating the impact of atmosphere-fire interactions on forest fire behavior under both coupled and uncoupled experimental settings. The focus of the current study is specifically to compare the differences in response modes of various PBL schemes under identical fire thermal forcing conditions, rather than quantifying the absolute intensity of the fire perturbation. Therefore, this paper adopts a unified fire source configuration and places the analytical emphasis on the comparison between schemes. Relevant assumptions and limitations have been clearly stated in the discussion section of the revised manuscript.

The new sentences have been added to Lines 60-71 and 129-145 in the revised manuscript.

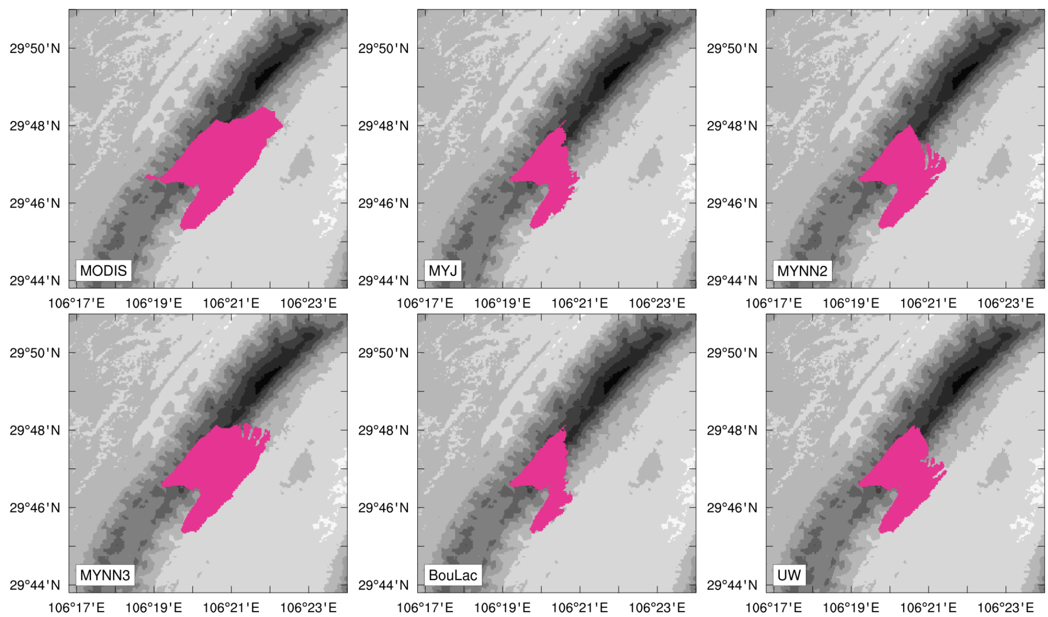

Figure MODIS and simulated burned areas from different boundary layer schemes, with gray shading area indicating terrain elevation.

2. Comparing PBL schemes and their performance during the fire process can help understand the characteristics of these schemes under non-typical conditions. However, fire-induced processes occur at the microscale, where turbulent eddies propagate in 3D space, and these processes are not represented by PBL schemes. Most (if not all) PBL schemes are 1D, with only vertical fluxes at each grid point. Therefore, poor performance of the atmospheric state near the fire is expected, especially at spatial resolutions that should ideally be resolved by LES. This is also true in complex terrain regions.

We thank the reviewer for this critical comment regarding fire-induced microscale processes and the applicability of PBL parameterizations. We agree that fire-induced turbulence essentially possesses distinct 3D characteristics and occurs within the microscale range; such processes are not the design target of traditional PBL schemes. Therefore, near the fire ground, especially under complex terrain conditions, PBL schemes have inherent limitations in characterizing the local atmospheric state.

It should be noted that although the PBL schemes used here are theoretically 1D vertical parameterizations, we enabled diff_opt=2 and km_opt=4 in the model configuration. This allows turbulent diffusion in both horizontal and vertical directions to participate in the numerical evolution. This setting enhances the model's ability to dissipate and adjust sub-grid scale perturbations to some extent, ensuring that fire-induced thermal perturbations are not entirely confined to a single-column structure but can propagate limitedly at the grid scale. We have added an explanation of this configuration in the Methods section of the revised manuscript.

Using the default model configuration with 45 vertical layers and three nested domains (7500 m, 1500 m, and 300 m; Rothermel fire grid: 30 m), we tested time steps of 45 s, 15 s, and 10 s, finding that only the 10 s time step allowed the simulation to complete the full 145-hour duration; following the reviewer's suggestion, we subsequently attempted a three-domain nested simulation using WRF-LES for the innermost domain (increasing vertical layers to 65 to meet the minimum requirement of 60 layers for WRF-LES) with the same horizontal resolutions and tested time steps of 10 s, 5 s, 2 s, and 1 s, but the results showed that the WRF-LES configuration was unstable at every time step for this 145-hour event, and since the instability in the coupled atmosphere-fire model cannot be resolved in the short term, the PBL scheme remains the currently feasible approach despite its limitations. Furthermore, at a 2 s time step, the ratio of computation time to simulation time exceeded 5:1 (5 minutes of runtime for 1 minute of simulation), which is unacceptable for operational fire spread forecasting; in future work, the authors plan to conduct WRF-LES simulations in flat terrain to analyze microscale fire-induced processes and the impact of 3D turbulence propagation on fire spread.

Based on the reasons above, this study focuses on comparing the relative response characteristics of multiple PBL schemes under identical fire-induced thermal forcing conditions, provided that stable operation is ensured. This choice does not imply that PBL schemes can fully describe fire-induced microscale turbulence; rather, it is a compromise solution that is feasible and reproducible under current resolution and computational conditions. We have explicitly pointed out this limitation in the Discussion section of the revised manuscript and emphasized that the conclusions should be understood as a comparison of PBL parameterization behaviors, not a direct resolution of the fire ground's microscale turbulent structure.

3. A 300 m grid spacing is too fine for using PBL schemes, especially in the presence of fire-induced perturbations. A more reasonable approach would be to use LES for the innermost domain (d03), with the PBL scheme from d02 providing boundary conditions for d03—this is a typical configuration for WRF-Fire. Vertical resolution is also critical:

this study uses the WRF default (i.e., CONUS) settings, where the first layer is located from the ground to a height of approximately 50 m. However, the heat source of the fire is confined to within a few meters of the surface. A 50 m thick vertical layer cannot reasonably represent the actual heat intensity and the magnitude of the resulting perturbations.

We thank the reviewer for the suggestion. As mentioned previously, when boundary layer schemes are enabled for d01 and d02 while d03 adopts the LES scheme, the model requires a vertical layer count of 60 or greater. In our tests, we employed vertical layer counts of 75 and 65, with the first atmospheric layer set within 10 meters of the ground; however, the coupled model could not run normally. When the vertical layer count was set to 60 with d03 using a PBL scheme, the simulation was also unstable. Currently, the atmosphere-fire coupled model does not support restart simulations, meaning we cannot resume from a breakpoint.

The height of the fire source is closely related to the height of the fuel and the thickness of the first atmospheric layer. In this study, forest trees account for over 98% of the fuel. In WRF-Fire, the Rothermel forest fire spread model utilizes the synthetic wind method proposed by Albini et al., which synthesizes wind and slope effects into an "effective wind." This effective wind is then substituted into the Rothermel formula using a wind speed height of 6.096 meters. This approach maintains the form of the original formula while allowing application in complex terrain and non-heading wind situations. Considering the actual circumstances of forest fire occurrence, a reasonable layer thickness would be 10–20 meters. However, due to the excessive heat released by the fire source, the calculation of the coupled model becomes unstable. Therefore, the coupled model cannot reproduce this forest fire spread process when the model layer height is set too low.

4. The manuscript currently does not discuss these key limitations. These limitations are critical aspects of simulating fire-induced atmospheric processes and the fire spread

they drive. From the perspective of studying fire-atmosphere coupling processes and their impact on the chaotic nature of fire spread, the manuscript's model configuration is not convincing. However, LES (including nested LES) is computationally expensive, and there may be application scenarios that do not require focusing on microscale processes. Since this manuscript does not focus on the typical challenges of WRF-Fire, it is suggested that the authors strengthen the explanation of the research motivation and potential application scenarios, and clarify the limitations of the model configuration.

We thank the reviewer for the in-depth comments regarding the innermost grid resolution, the applicability of PBL versus LES, and the vertical resolution issues. We agree that this issue is one of the key technical challenges when simulating fire–atmosphere coupling processes. In response to the previous and above comments, our response is as follows:

First, we acknowledge the reviewer's judgment: the 300 m horizontal resolution lies within the "gray zone" (terra incognita) between traditional PBL parameterization and LES. In the presence of strong fire-induced thermal perturbations, employing PBL schemes inevitably has physical limitations. At the same time, the first layer thickness of approximately 50 m is indeed difficult to resolve the strong heat release and vertical gradients occurring within the few meters near the ground where the fire source actually exists. These issues have been explicitly discussed as important limitations of the model configuration in the revised manuscript.

Regarding the choice of methodology, we attempted to use a nested LES for the innermost domain following the typical WRF-Fire configuration (with the PBL scheme of d02 providing boundary conditions for d03). However, under the combined effects of complex mountainous terrain, non-stationary fire heat release, and the continuous multi-day simulation period involved in this study, the LES experiments faced significant difficulties in terms of numerical stability and computational cost. Actual tests showed that multiple sets of LES simulations struggled to run stably for more than about one day, and the computational cost made it unfeasible to conduct a systematic comparison of multiple PBL/LES schemes under current computing conditions. Therefore, LES was not selected as the primary simulation framework for this paper.

Based on the aforementioned practical constraints, this paper adopts an innermost resolution of approximately 300 m combined with PBL parameterization schemes, aiming to achieve a compromise configuration that is numerically stable, computationally reproducible, and capable of covering the complete fire process. It needs to be emphasized that this study does not attempt to resolve fire-induced microscale turbulent structures or directly characterize the chaotic nature of fire spread, but rather focuses on the relative response differences of different PBL schemes regarding near-surface wind fields, turbulence intensity, and boundary layer structures under the same fire heat forcing background. Under this research objective, PBL schemes still provide a comparable diagnostic framework.

Furthermore, although the first layer thickness is about 50 m, the 2-m temperature and 10-m wind speed are not simple averages of this layer, but are diagnosed by the surface layer parameterization scheme based on stability and surface fluxes. Therefore, the response differences of different PBL–surface layer combinations to observed variables under fire conditions can still reflect the handling characteristics of parameterization schemes towards fire-induced perturbations. Relevant physical assumptions and their limitations have been explained in the revised manuscript.

Finally, we have strengthened the elaboration of the research motivation and potential application scenarios in the revised manuscript, explicitly pointing out that the results of this paper are mainly applicable to the following scenario: evaluating and comparing the behavior of PBL parameterization schemes in strong, heterogeneous thermal perturbation environments under conditions where computational resources are limited and long-duration nested LES cannot be conducted. This paper does not aim to replace high-resolution LES or microscale fire behavior models, but to provide a reference for understanding and improving the applicability boundaries of PBL schemes in coupled models.

The new sentences have been added to Lines 358-370 in the revised manuscript.

5. The authors must clearly explain the motivation for using PBL schemes given the 300 m grid spacing and a first layer height of 50 m. It is strongly suggested to include a set of simulations using the WRF LES option, with the first layer set closer to the ground. This would help in understanding the impact of using PBL instead of LES and assessing whether this causes critical biases. If conditions permit, adding a 3-D PBL scheme (configured consistently with other PBL schemes) would also provide valuable (though not mandatory) insights.

We thank the reviewer for this critical suggestion. We have explicitly supplemented the motivation for adopting PBL schemes in the revised manuscript. Although the horizontal resolution of the innermost atmospheric grid reaches 300 m, this scale is still insufficient to meet the basic premise of LES, which is to explicitly resolve the energy-containing turbulent eddies. At this resolution, the majority of turbulent kinetic energy remains within the sub-grid scale range; thus, employing LES is not physically self-consistent and makes it difficult to obtain stable, reproducible multi-day simulation results numerically. In contrast, PBL parameterization schemes remain a more robust and widely applied choice within this scale range, especially suitable for the research objective of conducting multi-scheme comparisons and sensitivity experiments in complex terrain regions.

We also agree with the reviewer regarding the importance of vertical resolution and have explicitly pointed out in the revised manuscript that the default WRF vertical stratification (with the first layer at approximately 50 m) indeed cannot finely represent the strong thermal gradients and turbulent structures within the few meters near the fire source. However, this study focuses on the integrated response of near-surface meteorological elements (such as 2-m temperature and 10-m wind speed) under fire conditions, and the relative performance differences of different PBL schemes to fire-induced perturbations, rather than the microscale structure of the fire plume itself. Within the framework of the current multi-day continuous simulation and multiple experiments, increasing the near-surface vertical resolution would significantly increase computational costs and introduce numerical instability. Therefore, this study adopted a vertical stratification setting that represents a compromise between feasibility and physical consistency.

We acknowledge that LES is of great value for understanding fire–atmosphere interactions and have tested multiple LES configurations in our preliminary work, including sensitivity experiments with reduced first-layer heights. However, under complex terrain conditions, LES coupled with the interactive fire model exhibits poor numerical stability; the simulation duration typically struggles to exceed 24 hours, and the computational cost increases significantly, making it difficult to support the multi-day simulations and systematic multi-scheme comparisons required for this paper. Consequently, LES results were not included as the main analytical content of this article.

It should be noted that the vast majority of PBL schemes in WRF (including MYJ, MYNN, BouLac, and UW) are inherently based on the one-dimensional vertical flux assumption; their differences mainly lie in turbulence closure forms, mixing lengths, and stability functions. In this study, by enabling 3D turbulence diffusion options (diff_opt = 2, km_opt = 4), we allowed sub-grid turbulent diffusion in both horizontal and vertical directions, thereby mitigating the limitations of the pure 1D assumption to a certain extent. Due to limitations in computational resources and the research focus, we did not further introduce experimental 3D PBL configurations, but we agree that this direction holds potential value and have acknowledged it as a possible extension for future work in the revised manuscript.

Under the conditions of 300 m horizontal resolution and 50 m first-layer height, the amplitude of fire-induced microscale turbulence and strong thermal plumes might be underestimated, which could affect the absolute intensity of fire–atmosphere interactions. However, since all experiments employ the same dynamic framework and vertical stratification, this bias is consistent across different PBL schemes and thus does not affect the conclusions regarding the relative performance and sensitivity of the different PBL schemes presented in this paper.

The new sentences have been added to Lines 60-71 and 358-370 in the revised manuscript.

6. The observation comparison uses 2-m temperature and 10-m wind speed, but the temperature and wind in the simulation represent an average layer of 0–50 m. The atmosphere and land surface in WRF are coupled, but fire heat couples only to the atmosphere, not directly to the soil. Therefore, soil temperature increases come only from air temperature, not the fire, leading to excessive thermal conduction inertia and underestimated thermal gradients. Since T2 is diagnosed from T (representing a 50 m thick layer), the fire feedback is significantly weakened. The same applies to 10-m wind; they are interpolations rather than truly resolved. It is suggested to add a control experiment with fire feedback turned off (or completely no fire) to demonstrate the magnitude of fire perturbation on each PBL (which might be very small). Furthermore, this would allow comparison of how each scheme transports fire heat, i.e., comparing the temperature or energy differences between fire/no-fire simulations for each PBL.

We thank the reviewer for the in-depth discussion regarding the diagnostic mechanism of near-surface variables and the fire–land–atmosphere coupling issue. Our understanding and response to this issue are as follows.

First, 2-m temperature (T2) and 10-m wind speed (U10) are not simply equivalent to the layer average of the first model level (0–50 m). In WRF, these two variables are diagnosed by the surface layer parameterization scheme based on the atmospheric state of the first layer, surface fluxes, and Monin–Obukhov similarity theory. Although the thickness of the first layer is indeed approximately 50 m, the calculation processes for T2 and U10 explicitly introduce stability corrections, friction velocity, and scaling variables. Their response to fire-induced perturbations depends to a certain extent on the surface layer scheme itself, rather than being a simple interpolation of the layer-averaged temperature or wind speed. Therefore, the response differences of different PBL–surface layer combinations to T2 and U10 under fire conditions still possess physical significance.

Second, we agree with the important fact pointed out by the reviewer: in WRF-Fire, fire heat is only coupled to the atmosphere and does not directly act on the soil. Consequently, changes in soil temperature mainly originate from the adjustment of sensible heat flux after the air is heated by the fire, rather than direct fire–soil heat conduction. This mechanism indeed leads to greater thermal inertia in land–atmosphere heat exchange and may underestimate the near-surface thermal gradient. We have explicitly discussed this physical limitation in the revised manuscript and emphasized that the results of this paper should be understood as the response of fire-induced atmospheric perturbations at the model-resolvable scale, rather than a precise characterization of the actual thermal structure at the meter scale near the ground.

Regarding the suggestion to add a "fire-off" (closed fire feedback) or no-fire control experiment, we consider this suggestion physically reasonable. However, given that the authors have conducted similar research previously, it was not adopted within the framework of this study, primarily based on the following considerations: the objective of this study is to compare the relative performance differences of different PBL schemes under the same fire forcing condition, rather than to quantify the absolute intensity of the fire perturbation. Since all experiments used a completely identical fire source configuration, the fire heat input serves as the same external forcing across various PBL schemes; thus, this paper focuses on the differences in how each scheme responds to the same fire heat input.

Finally, regarding the LES issue, we agree that adopting LES and setting the first layer height lower is theoretically more beneficial for resolving fire-induced turbulent structures, which is also one of the main advantages of LES or ultra-high-resolution simulations. However, in the simulation of complex mountain terrain and long-duration fire processes involved in this study, nested LES faces significant challenges in terms of numerical stability and computational cost. Therefore, we adopted a first layer thickness of approximately 50 m as a compromise solution, allowing the simulation to cover the complete fire process and perform multi-scheme comparisons while ensuring numerical stability. The revised manuscript has explicitly stated that this configuration is more suitable for comparative studies of model behavior and parameterization schemes, rather than for the direct resolution of microscale fire–turbulence structures.

7. Finally, the CONUS namelist does not include the fire configuration. Please include the fire-related configuration, including the ignition location and all &fire options that differ from the defaults.

We thank the reviewer for the suggestion. The namelist settings have been added as follows:

```
&domains
sr_x = 0,        0,       10,
sr_y = 0,        0,       10,
&fire
ifire =    0,    0,    2,
fire_num_ignitions          = 0,    0,    1,
fire_ignition_start_lon1    = 0,    0,    106.3428,
fire_ignition_start_lat1    = 0,    0,    29.7705,
fire_ignition_end_lon1      = 0,    0,    106.3428,
fire_ignition_end_lat1      = 0,    0,    29.7705,
```

```
fire_ignition_radius1          = 0,    0,    35,
fire_ignition_start_time1      = 0,    0,    52200,
fire_ignition_end_time1      = 0,    0,    52204,
fire_print_msg      = 0,    0,    0,
fire_print_file      = 0,    0,    0,
fmoist_run          = .false., .false., .true.
fmoist_interp       = .false., .false., .true.,
fmoist_only         = .false., .fasle., .false.,
fmoist_freq         = 0,      0,      0,
fmoist_dt           = 10,    10,    10,
fire_fmc_read    = 0,      0,      0,
fire_boundary_guard      = -1,    -1,    -1,
fire_fuel_left_method      =    1,      1,      1,
fire_fuel_left_irl         =    2,      2,      2,
fire_fuel_left_jrl         =    2,      2,      2,
fire_atm_feedback     =    1.,    1.,    1.,
fire_grows_only        =    1,      1,      1,
fire_viscosity               =    0.4, 0.4, 0.4,
fire_upwinding            =    3,      3,      3,
fire_lfn_ext_up              =    1.0, 1.0, 1.0,
fire_test_steps              =    0,      0,      0,
fire_topo_from_atm        =    1,      1,      0,
```

It should be noted that the officially released WRF version only permits the use of YSU for meso-LES configurations. However, with very minor modifications to the registry, other PBL schemes can also be made to work: the key lies in ensuring that the variables in the pbl=0 package include those required by the corresponding PBL scheme.

We appreciate the reviewer's suggestion and will strengthen our work in this direction in future research.

Reviewer 2:

The authors attempt to discuss the dynamic relationship between wildfire processes and model schemes within WRF. However, the article does not address discussions regarding wildfire ignition and spread. Furthermore, WRF-Fire has limitations in modeling and practical application within the Wildland-Urban Interface (WUI). The authors need to explain how the information from these simulation results can be used for decision-making in the built environment. The reference to "wildfire" in the title is misleading and requires modification. When the wildfire aspect is removed, the article reads very similarly to other papers comparing PBL schemes at a single station. To distinguish it from existing research, the authors need to clearly point out the unique contribution of this study.

We thank the reviewer for the constructive comments. In the revised manuscript, we have further clarified the research objectives, the core role of observational data, and the unique contributions of this paper relative to existing work. Specific explanations are as follows.

Although this fire occurred near a city, over 98% of the burned area consisted of forest and natural vegetation. The fire ground was located within a scenic area covered by continuous woodland and did not involve the burning of buildings or typical Wildland-Urban Interface (WUI) fire behavior processes. Therefore, this study remains focused on the issue of atmospheric response under wildfire conditions, rather than urban or structural fire simulation. We have explicitly clarified the study area and fuel types in the revised manuscript.

The focus of this study is not to finely characterize fire ignition mechanisms or provide engineering-grade predictions of fire spread, but rather to utilize the strong thermal perturbations generated by a real wildfire event as a "stress test" to examine the performance differences of different boundary layer schemes under extreme near-surface conditions. To avoid remaining solely at a theoretical or idealized comparison level, this paper conducts evaluations based on station observational data near the fire ground (distances ranging from 1 to 9 km).

Specifically, this study employed hourly data from 5 fixed automatic weather stations and 1 mobile observation station surrounding the fire ground to systematically compare the 2 m temperature, 10 m wind speed, and their evolution characteristics simulated by WRF-Fire. This type of observational data possesses irreplaceable importance: the observation locations were directly affected by the fire field, capable of truthfully reflecting the near-surface response under conditions of fire-induced heating and enhanced turbulence; the distribution of multiple stations (upwind, fire edge, and downwind) allows the evaluation to go beyond a single point and reflect spatial differences under complex terrain conditions; the mobile station provided supplementary information at close range to the fire, enhancing the constraints on model results. It is precisely based on these observational data that this study is able to distinguish the response characteristics of different PBL schemes regarding near-surface wind fields, turbulence intensity, and boundary layer structures in a fire environment.

Furthermore, a comparative analysis between the WRF-Fire simulated burned area and the MODIS burned area product has been added to the revised manuscript to verify the rationality of the fire–atmosphere coupled simulation on a spatial scale, thereby avoiding the simplification of the research into a "fire-module-removed PBL comparison experiment". The five boundary layer schemes (MYJ, MYNN2, MYNN3, BouLac, and UW) produced burned area estimates ranging from 47.38% to 92.82%, the MYNN3 scheme had the best result.

In summary, the unique contribution of this study lies in the systematic evaluation of the performance of multiple PBL schemes under strong fire-induced thermal forcing conditions, within the context of a real wildfire in complex terrain, utilizing near-field station observational data. The research results provide valuable references for model development regarding the applicability and limitations of PBL parameterization schemes under non-typical, forced conditions, which is highly aligned with GMD's focus on model behavior and applicability boundaries.

The new sentences have been added to Lines 60-71 and 129-145 in the revised manuscript.

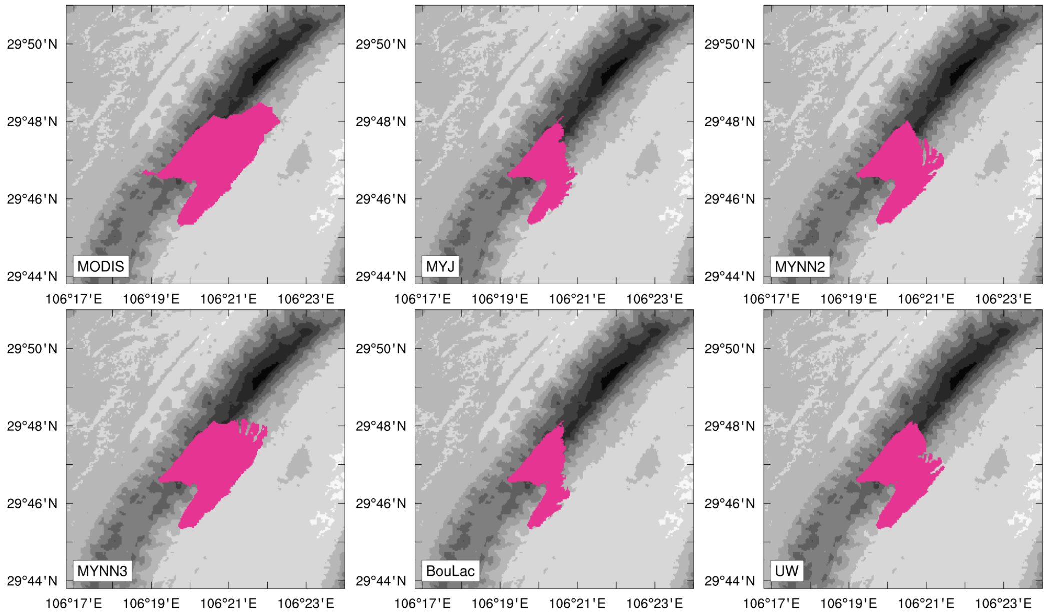

Figure MODIS and simulated burned areas from different boundary layer schemes, with gray shading area indicating terrain elevation.

Other Comments:

1. The authors need to provide a justification for the choice of PBL schemes. Why were these five chosen from the many WRF PBL schemes? Also, why were more recent "Hybrid PBL–LES" schemes not considered?

We thank the reviewer for this important question. In this study, we selected the MYJ, MYNN2.5, MYNN3, BouLac, and UW boundary layer schemes, aiming to represent different turbulence closure concepts within WRF, and because they are widely used in wildfire–atmosphere coupling research.

The objective of this study is not to be exhaustive of all PBL schemes, but to compare the schemes that are most representative, exhibit good stability, and hold practical application value in regional-scale wildfire simulations.

We did not adopt Hybrid PBL–LES schemes because this paper focuses on practical application scenarios at a 300 m resolution where reliance on PBL parameterization is still necessary. Hybrid schemes typically require finer vertical resolution and additional configuration; their introduction would increase uncertainty and exceed the scope of this study regarding PBL sensitivity comparison. Relevant explanations have been added to the revised manuscript.

Although Hybrid PBL–LES schemes have been introduced into WRF in recent years to smooth the transition of turbulence description between RANS and LES under high-resolution conditions, we did not employ such schemes in this study based primarily on the following considerations:

First, Hybrid PBL–LES schemes are typically designed for the "turbulence gray zone" scale. Their applicability premise is that the atmospheric grid resolution is sufficiently close to the dominant turbulence scales, accompanied by high vertical resolution and idealized terrain conditions. In complex mountainous environments, the forced circulations introduced by terrain undulation, shear discontinuities, and strong fire-induced thermal perturbations would significantly magnify gray zone uncertainties, making it difficult for the physical advantages of Hybrid schemes to be clearly manifested.

Second, the current systematic evaluation of Hybrid PBL–LES schemes within the WRF-Fire coupled system remains limited, especially regarding their turbulence closure and energy partition mechanisms under the action of strong local heat sources (such as firelines), which lack sufficient verification. Considering GMD's requirements for model reproducibility and physical consistency, this study prioritized traditional PBL parameterization schemes that are more mature and have more existing application instances in wildfire simulation.

Finally, from the perspective of experimental design, a core objective of this study is to perform a comparative analysis of the performance of different PBL schemes under identical resolution and physical frameworks. Introducing Hybrid PBL–LES would introduce additional degrees of freedom at the turbulence closure level, which is not conducive to explicitly distinguishing the impact of the boundary layer parameterization itself on the simulation results. Therefore, under the current spatial resolution and research objectives, Hybrid PBL–LES was not included in the model configuration of this study.

2. Why do all PBL schemes use the same surface layer scheme? For example, MYNN has its own specific surface layer scheme.

In WRF, the surface layer scheme acts as a critical interface between the land surface model and the boundary layer scheme. It does not directly predict atmospheric state variables but diagnoses momentum, heat, and moisture fluxes based on Monin–Obukhov similarity theory, utilizing the lowest-level atmospheric variables and surface characteristics. The surface layer scheme provides the boundary layer scheme with friction velocity, surface fluxes, and stability parameters; these quantities serve as the lower boundary conditions for the boundary layer turbulence equations, directly regulating turbulence production and vertical mixing intensity.

Different surface layer schemes employ distinct variants in calculating stability functions and roughness lengths, which affect the computation of turbulent exchange. The purpose of using a single, unified surface layer scheme across all experiments is to maximize the isolation of the effects of the boundary layer schemes themselves. If different PBL schemes were paired with different surface layer schemes, it would introduce additional coupling effects, making it difficult to distinguish whether the differences in results originate from the PBL scheme or the flux calculations.

Although the MYNN scheme comes with a dedicated surface layer option, a unified surface layer setting helps ensure that all schemes are compared under identical surface forcing conditions.

Therefore, the sensitivity of wildfire–atmosphere coupling to PBL schemes partly originates from how surface-layer fluxes are ingested and redistributed within the boundary layer, rather than from differences in the surface layer formulation itself.

The new sentences have been added to Lines 76-81 in the revised manuscript.

3. WRF ideally uses a 1:3 nesting ratio to reduce numerical dispersion. Higher ratios are typically used when downscaling to LES to avoid the "terra incognita" zone. Is there a specific reason for using a 1:5 nesting ratio in this paper?

The adoption of a 1:5 nesting ratio in this study (with an innermost atmospheric model resolution of 300 m and a fire spread model resolution of 30 m) represents a trade-off between computational efficiency, applicability in the gray zone, and regional representativeness.

Fire–atmosphere coupling is highly non-linear; even a time step difference of 0.2 s can alter the turbulent structure and the local circulation around the fire source, leading to differences in wind speed and temperature feedback. Adopting a 1:5 ratio not only reduces the number of nesting levels but also allows the time step of the innermost domain to be minimized.

Although a 1:3 nesting ratio helps reduce numerical dispersion, a 1:5 ratio is not uncommon in existing WRF-Fire research when downscaling from synoptic scales to convection-resolving or fire scales. The outer grids primarily provide large-scale background circulation constraints, while the innermost 300 m grid focuses on the relative differences between different PBL schemes. Since all schemes utilize the same model configuration, this setup is suitable for comparing simulation differences arising from different physical schemes, ensuring the scientific validity and practicality of the conclusions.

4. The 300 m resolution lies within the "gray zone" of applicability for PBL schemes. Why not further refine the grid and use LES instead of PBL?

We agree that the horizontal resolution of 300 m lies within the "gray zone" (or terra incognita) between traditional PBL parameterization schemes and explicit LES. Under ideal conditions, further refining the grid and adopting LES would allow for a more direct resolution of turbulent structures. However, in the context of regional-scale, multi-day continuous coupled wildfire simulations, LES still faces a series of practical and unresolved challenges; therefore, it was not selected as the primary configuration for this paper.

First, the demand for computational resources by LES increases non-linearly. Refining the horizontal resolution from 300 m to 100 m or even 50 m requires not only a significant reduction in the time step but also a synchronous refinement of vertical stratification (particularly requiring a layer thickness of <10 m in the surface layer). This typically increases the total computational cost by more than an order of magnitude. Given the complex terrain and the multi-day wildfire process investigated in this study, such a configuration is difficult to support systematic sensitivity experiments under currently available standard computational resources.

Second, the numerical stability of LES under strong wildfire–atmosphere coupling conditions remains an open question. The high spatiotemporal heterogeneity of fire-induced heat fluxes significantly amplifies the buoyancy production term in the surface layer, rendering LES highly sensitive to sub-grid scale (SGS) schemes, filter scales, and dissipation parameters. In complex terrain, this instability often leads to: abnormal sensitivity of simulation results to the choice of SGS schemes; the appearance of non-physical oscillations or energy accumulation near the fireline; and difficulties in maintaining reproducibility across different cases or regions.

Third, the compatibility of LES with existing operational wildfire forecasting frameworks remains limited. Currently, most regional-scale wildfire forecasting systems (including operational WRF-Fire applications) still rely on PBL parameterization as the core configuration. The objective of this study is precisely to evaluate the performance differences and applicability boundaries of different PBL schemes under strong fire-induced thermal perturbations within this realistic and widely used model configuration, rather than to explore the optimal simulation scheme under idealized conditions.

Therefore, although the 300 m resolution lies within the transition zone for PBL schemes, it still possesses clear practical significance for regional-scale wildfire simulations. The results of this paper can provide a benchmark reference for future research utilizing higher resolutions or hybrid PBL–LES frameworks.

5. For wildfire research, 300 m resolution is too coarse, especially since wind speed accuracy in complex terrain is the most critical factor for studying ignition and spread.

We acknowledge that a horizontal resolution of 300 m cannot resolve fire-line scale turbulence structures and local wind field details. Especially under complex terrain conditions, this resolution is insufficient to finely characterize the ignition process and the transient dynamics of the fire front. However, it is important to emphasize that the objective of this study is not to finely reproduce the fire spread process itself, but to focus on the following two issues of greater regional-scale significance:

First, this paper focuses on the modulation effect of fire-induced heat fluxes on the overall boundary layer structure. Although the internal structure of the fire line cannot be resolved at 300 m resolution, the total heat released by the fire and its impacts on near-surface stability, turbulence intensity, and wind speed profiles can still be effectively characterized. Simulation results indicate that at this resolution, the model possesses a good capability to reproduce the burned area and the overall evolution of the fire zone, which provides a reasonable basis for analyzing the differences in the fire–boundary layer coupling response among different PBL schemes.

Second, the importance of wind speed accuracy in wildfire research is not limited to the fire-line scale. In regional-scale simulations, wind fields within the surface layer and boundary layer are equally crucial for predicting fire spread direction, expansion trends, and potential risk areas. By comparing different PBL schemes, this paper reveals their different response mechanisms regarding wind speed, turbulence, and stability under strong thermal forcing. These differences have direct significance for understanding and improving future wildfire forecasts.

Finally, from an application perspective, the 300 m resolution represents a realistic trade-off between computational cost and physical representation capability. At the current stage, relying entirely on LES for regional-scale, complex terrain, multi-day wildfire simulations still face significant computational and technical barriers. In comparison, systematically evaluating the applicability of PBL schemes at this resolution helps to clarify the strengths and limitations of different schemes; Guide scheme selection in future higher-resolution or hybrid frameworks; Provide physical and engineering references for the gradual transition to finer simulations.

Therefore, although the 300 m resolution cannot resolve fire-line scale processes, its scientific value and practical significance in regional-scale wildfire simulation and boundary layer response research remain clear.

The new sentences have been added to Lines 358-370 in the revised manuscript.

6. In this study, how many vertical layers are located within the boundary layer? 2-m and 10-m winds are diagnostic variables, and their accuracy in complex terrain depends on the height of the first grid level.

Under the default 45-layer configuration, the daytime planetary boundary layer typically contains approximately 4–6 model layers (depending on the specific PBL height). In WRF, the surface layer scheme does not explicitly resolve a finite layer of fixed depth. Instead, it applies Monin–Obukhov similarity theory between the surface and the first atmospheric model level, assuming constant fluxes. Diagnostic 2-m temperature/humidity and 10-m winds are obtained through similarity-based interpolation or extrapolation, regardless of whether the first model level is located below or above 10 or 50 meters.

Under complex terrain conditions, the diagnostic accuracy of 2-m temperature and 10-m wind speed does not simply depend on the height of the lowest model level; rather, it relies more on the combined interaction of the lowest model level height, surface roughness characteristics, and the validity of the surface layer similarity theory assumptions.

7. What land use data and terrain data were used for simulation initialization?

For the initialization of the atmospheric model in this study, the land use data adopted the default land use classification scheme of the WRF/WPS system to maintain consistency in surface parameterization across all experiments.

Given the sensitivity of complex terrain and wildfire processes to high-resolution terrain and fuel information, this study made targeted replacements for terrain height and fuel data related to fire behavior. Specifically, the terrain height data utilized 30 m resolution DEM data to more accurately characterize slope, aspect, and local topographic relief features within the study area. The fuel/vegetation type data required by the fire behavior model employed the FROM-GLC 30 m global land cover dataset, which was remapped according to the fuel classification requirements of WRF-Fire.

This configuration improves the representation accuracy of the spatial distribution of terrain and fuels without altering the atmospheric physical parameterization settings, thereby facilitating a better characterization of the fire spread process under complex terrain conditions and its perturbation to the near-surface atmosphere.

8. Figures 5 and 6 do not clearly demonstrate that MYNN3 is the optimal scheme, particularly given the absence of observational uncertainty (error bars) in the analysis.

We thank the reviewer for highlighting this important issue. We agree that relying solely on the individual statistical metrics presented in Figures 5 and 6 makes it difficult to definitively conclude that MYNN3 is "optimal" across all meteorological variables and aspects. Consequently, we have modified the relevant terminology in the revised manuscript and supplemented the analysis with additional evidence to support the assessment of MYNN3's comprehensive performance.

First, it should be clarified that the definition of "optimal" in this study is not based solely on a single meteorological variable or statistical metric, but rather on a comprehensive consideration of the consistency in simulating near-surface thermal states, wind field structures, and fire behavior. In the revised manuscript, we have introduced a comparative analysis between the simulated burned area and the MODIS satellite-retrieved burned area product. The results indicate that among all PBL schemes, MYNN3 shows the closest agreement with the satellite retrievals regarding the spatial distribution and overall magnitude of the burned area. From the perspective of fire behavior response, this result provides independent evidence for the comprehensive applicability of MYNN3 in the context of fire–atmosphere coupling, extending beyond reliance solely on station statistics of near-surface meteorological elements.

Second, regarding the observation comparison, MYNN3 exhibits the smallest mean bias for 2-m temperature, suggesting a relatively more reasonable characterization of near-surface thermal conditions. It is important to note that both 2-m temperature and 10-m wind speed are variables diagnosed via surface layer similarity theory. Their errors are associated not only with the PBL scheme but are also collectively influenced by the surface layer scheme, terrain representativeness errors, and the spatial heterogeneity of fire-induced perturbations. Under conditions of complex terrain and strong heterogeneous heat sources, the sensitivity of different variables to model errors is inherently inconsistent.

Regarding 10-m wind speed, MYNN3 exhibits relatively larger mean bias and RMSE; this finding has been truthfully retained and discussed in the revised manuscript. It remains difficult to unequivocally attribute this bias to a single mechanism; it is likely related to the combined action of the following factors:

（1）10-m wind speed is highly sensitive to sub-grid terrain effects and local channeling effects, which are difficult to fully resolve at 300 m resolution;

（2）The impact of fire-induced thermal perturbations on the wind field possesses significant spatiotemporal heterogeneity, making wind speed errors at the station scale more prone to amplification;

（3）Since 10-m wind speed is a variable diagnosed via surface layer similarity theory rather than explicitly resolved by the PBL scheme, different PBL schemes indirectly influence friction velocity and stability parameters by altering the wind speed profile, turbulence mixing intensity, and thermal stability in the first model layer. Under strong fire-induced perturbations and complex terrain, this indirect influence may be amplified, thereby increasing the uncertainty of the 10-m wind diagnosis.

Given the lack of wind field observational data with higher spatiotemporal resolution, this paper does not perform a further quantitative distinction of the aforementioned mechanisms but identifies this as an issue requiring in-depth investigation in future work.

Finally, regarding the issue of observational uncertainty (error bars), we agree that this is an important aspect, albeit one with practical limitations. Systematic errors and representativeness errors of automatic weather stations are difficult to accurately quantify in complex terrain and fire environments; therefore, we did not perform a unified estimation of error bars in this paper. To avoid over-interpretation, we have softened the absolute claim of an "optimal scheme" in the revised manuscript, shifting the emphasis to the relative consistency advantage of MYNN3 across multi-metric and multi-perspective evaluations.

The new sentences have been added to Lines 371-381 in the revised manuscript.