

Response to Reviewer 3

General Comments

In general, I consider that the manuscript addresses a relevant problem for the WRF modeling community, since the selection of physical parameterization schemes remains one of the main sources of uncertainty in regional atmospheric simulations. The proposed weighting strategy based on the performance of multiple variables is interesting and has the potential to contribute to a more comprehensive evaluation of model configurations. However, I believe that some aspects could be strengthened to increase the scientific value of the study.

Response to General Comments:

We sincerely thank the reviewer for the positive and constructive overall assessment of our manuscript. We appreciate the reviewer's recognition that WRF physical parameterization-scheme selection is an important source of uncertainty in regional atmospheric simulations, and that the proposed performance-based multivariable weighting strategy has the potential to support more comprehensive evaluation of model configurations.

In response to the reviewer's comments, we strengthened the manuscript in three main aspects. First, we expanded the physical interpretation of the selected parameterization-scheme combinations by linking the statistical selections to the dominant regional processes in the arid Northwest and humid coastal Southeast China. Second, we clarified the scope of uncertainty addressed by the proposed framework, emphasizing that this study evaluates uncertainty associated with WRF physical parameterization-scheme selection under a fixed model configuration, rather than all sources of WRF simulation uncertainty. Third, we added a discussion of the practical meteorological meaning of the evaluated differences by connecting changes in integrated scores to variable-level RMSE changes in physically interpretable units.

These revisions were made to ensure that the methodological contribution is not presented only as a statistical ranking procedure, but is also connected to regional physical consistency, practical meteorological interpretation, and a clearly defined uncertainty scope. The above provides our overall response to the reviewer's General Comments. Detailed point-by-point responses are provided below.

In the point-by-point responses below, our responses are shown in regular black font, [whereas the corresponding revisions made in the manuscript are shown in blue regular font.](#)

Comment 3.1

The manuscript would benefit from a deeper discussion of the physical interpretation of the selected parameterization schemes. Although the methodology successfully identifies the configurations that provide the best statistical performance, the discussion remains focused mainly on the evaluation metrics. It would be valuable to explain why the selected schemes may be physically suitable for the climatic and atmospheric characteristics of the analyzed regions. Such discussion would help readers understand whether the methodology is identifying parameterizations that are not only statistically optimal but also physically consistent with the dominant atmospheric processes.

Response 3.1:

Thank you for this helpful comment. We agree that readers need to understand whether the statistically selected parameterization combinations are also physically consistent with the dominant atmospheric processes in the two study regions.

In response, we added a new Sect. 4.2, “Physical interpretation and practical meteorological relevance”. In this section, we discuss the two selected optimal configurations separately. For the arid Northwest domain, we explain why the P3, YSU, Tiedtke, GFDL, and Noah LSM combination is physically plausible in relation to intermittent precipitation, complex terrain, boundary-layer exchange, radiation, and land-surface energy-water exchanges. For the humid coastal Southeast domain, we explain why the WDM6, Shin-Hong-MM5, Grell 3, GFDL-GFDL, and Noah LSM combination is consistent with monsoon moisture transport, moist convection, cloud-precipitation processes, and coastal boundary-layer dynamics.

We also revised the tone of the discussion to avoid claiming a causal mechanism that was not directly diagnosed. The revised text therefore states that the selected schemes are physically consistent with the dominant regional processes, while acknowledging that water-vapor budget analysis, hydrometeor diagnostics, turbulent kinetic energy diagnostics, PBL-height evaluation, and surface energy-balance decomposition would be needed for formal process attribution.

Corresponding revisions:

[Section 4.2 Physical interpretation and practical meteorological relevance \(page 31, lines 673-676\):](#)

[In the arid region of Northwest China, limited water-vapor supply, complex terrain, strong land-atmosphere coupling, and large diurnal temperature variations make precipitation, near-surface wind, and temperature particularly sensitive to cloud microphysics, boundary-layer mixing, radiation, and surface energy-water exchanges. The selection of the P3 microphysics scheme is physically plausible because P3 predicts bulk ice particle properties and explicitly represents the mass and volume of rimed ice, allowing a more flexible treatment of mixed-phase hydrometeors](#)

than schemes with fixed ice categories (Morrison and Milbrandt, 2015). This is relevant for precipitation simulation over arid and topographically complex regions, where precipitation events are intermittent and strongly affected by cloud microphysical uncertainty. The selection of the YSU PBL scheme is also consistent with the importance of near-surface wind and temperature in this region, because YSU uses a non-local K-profile approach with an explicit entrainment layer to represent turbulent exchange in the unstable mixed layer (Hong et al., 2006). The Tiedtke cumulus scheme, a mass-flux scheme with shallow convection and momentum transport, provides a physically meaningful representation of sub-grid convective transport under conditions where convection is not fully resolved (Tiedtke, 1989; Zhang et al., 2011). Together with the GFDL radiation scheme, which accounts for radiative transfer with carbon dioxide, ozone, and cloud microphysical effects (Fels and Schwarzkopf, 1981), and the Noah LSM, which represents soil temperature, soil moisture, snow, and surface fluxes in the land-atmosphere system (Chen and Dudhia, 2001), this combination is consistent with the dominant controls on wind, temperature, and intermittent precipitation in the arid Northwest domain.

Section 4.2 Physical interpretation and practical meteorological relevance (page 31, lines 691-694):

In the humid coastal region of Southeast China, the dominant simulation challenges are more closely related to monsoon moisture transport, moist convection, cloud-precipitation processes, and the coastal boundary layer. The selection of WDM6 is physically interpretable because this scheme extends WSM6 by including double-moment treatment of rain and cloud water as well as prognostic cloud condensation nuclei for warm-rain processes, which is relevant for warm and moist precipitation environments (Lim and Hong, 2010). The selected Shin-Hong PBL scheme includes scale dependency for vertical transport in the convective boundary layer while retaining YSU-type treatment for stable boundary-layer and free-atmospheric mixing (Shin and Hong, 2015). This feature is consistent with the need to represent vertical exchange of heat, moisture, and momentum in coastal regions where sea-land thermal contrast and moist boundary-layer processes affect near-surface wind and humidity. The Grell 3 cumulus scheme is an improved Grell-Devenyi type mass-flux approach and can represent sub-grid convective effects through ensemble-based convective assumptions (Grell, 1993; Grell and Devenyi, 2002). Its selection in the Southeast domain is therefore consistent with the higher weight assigned to precipitation in the cumulus evaluation and with previous studies showing that convection and microphysics parameterizations strongly influence precipitation and cloud properties over Asia and East Asian summer monsoon rainfall (Liu et al., 2018; Yang et al., 2015).

Section 4.2 Physical interpretation and practical meteorological relevance (page 32, lines 707-708):

These physical interpretations should be treated as mechanism-informed explanations of the statistical selection results, not as direct process-attribution evidence.

Comment 3.2

The uncertainty analyzed in this study is associated with the selection of physical parameterization schemes. However, parameterization choice represents only one component of the overall uncertainty present in WRF simulations. Other factors, such as spatial resolution, domain configuration, land-use representation, topography, initial and boundary conditions, and spin-up time may also influence model performance. A brief discussion clarifying the specific scope of the proposed methodology and its relationship to other sources of uncertainty would strengthen the interpretation of the results.

Response 3.2:

We thank the reviewer for pointing out this important distinction. We agree that physical parameterization choice is only one component of the broader uncertainty in WRF simulations, and that the original manuscript needed to define the scope of the proposed framework more explicitly.

In response, we revised the manuscript in three connected places. First, in Sect. 2.2 and Table 1, we clarified the simulation and evaluation periods. The simulations covered May 1 to August 31, with May 1 to May 31 used as the model spin-up period and only June-August used for evaluating the candidate parameterization-scheme combinations. This clarification makes the spin-up design explicit rather than leaving it implicit in the text.

Second, we added a scope statement in Sect. 4. The revised Discussion now states that the uncertainty addressed in this study is the uncertainty associated with WRF physical parameterization-scheme selection under a fixed model configuration. In the present experimental design, all candidate simulations share the same WRF dynamical core, horizontal resolution, domain configuration, vertical setup, initial and boundary conditions, land-use dataset, spin-up design, evaluation variables, and observational reference. Therefore, the reported differences should be interpreted as differences in the evaluated multivariable performance of alternative parameterization-scheme configurations and weighting choices within this controlled framework.

Third, we explicitly acknowledge that other sources of WRF uncertainty, including horizontal and vertical resolution, domain extent and nesting strategy, terrain and land-use representation, initial and boundary forcing, and spin-up length, can also influence model performance. These factors were not independently perturbed or attributed in the present study. We therefore revised the manuscript to state that the proposed $\text{MRE} \times \text{CV}$ weighting framework is a reproducible evaluation tool for parameterization-scheme screening, rather than a complete uncertainty-decomposition framework for WRF simulations.

This revision aligns the methodological claim more closely with the actual experimental design. The role of the method is to improve the transparency of multivariable parameterization-

scheme evaluation under a fixed WRF setup and to provide additional information for variable-level diagnostic interpretation, rather than to replace a full uncertainty attribution across resolution, domains, forcing data, land-use representation, terrain, and spin-up choices. This clarification defines the interpretation range of the present results and points to future applications that combine the weighting framework with broader WRF sensitivity experiments.

Corresponding revisions:

Table 1 (page 5, lines 118-125):

The first month, from May 1 to May 31, was used as the model spin-up period, and only the June-August simulations were used to evaluate the performance of the physical parameterization scheme combinations.

Section 4.3 Robustness and uncertainty boundaries of the screening framework (page 34, lines 782-789):

The uncertainty addressed in this study should therefore be interpreted as uncertainty associated with WRF physical parameterization-scheme selection under a fixed model configuration. In the present experiments, all candidate simulations shared the same WRF dynamical core, horizontal resolution, domain configuration, vertical setup, initial and boundary conditions, land-use dataset, spin-up design, evaluation variables, and observational reference. This controlled design allows the proposed framework to diagnose how alternative physical parameterization schemes and variable-weighting choices affect the evaluated multivariable performance. It also keeps the comparison focused on the parameterization-screening problem that the method is designed to solve. Other sources of WRF uncertainty, including horizontal and vertical resolution, domain extent and nesting strategy, terrain and land-use representation, initial and boundary forcing, and spin-up length, can also influence model performance. These factors were not independently perturbed or attributed in the present study. The proposed $\text{MRE} \times \text{CV}$ weighting framework should therefore be viewed as a reproducible evaluation tool for parameterization-scheme screening, rather than as a complete uncertainty-decomposition framework for WRF simulations. Future applications could combine this weighting framework with multi-resolution, multi-domain, land-use, forcing, or spin-up sensitivity experiments to assess how parameterization-scheme uncertainty interacts with these additional uncertainty sources.

Table 1 (page 5, lines 104-105):

May 1st-Aug 31st, 2007 / May 1st-Aug 31st, 2009; Spin-up period: May 1st-May 31st; Evaluation period: June 1st-Aug 31st.

Comment 3.3

The manuscript demonstrates differences among the evaluated parameterization combinations; however, it would be useful to discuss the practical implications of these differences from a meteorological perspective. In several cases, the statistical differences between competing configurations appear relatively small. Therefore, it would be interesting to discuss whether these improvements translate into a substantially better representation of the atmospheric processes of interest or whether they mainly reflect differences in the evaluation metrics. This would help readers better assess the practical significance of the proposed methodology.

Response 3.3:

We thank the reviewer for this constructive suggestion. We agree that the manuscript should explain the practical meteorological meaning of the differences among the selected parameterization configurations, rather than presenting the differences only as changes in integrated scores.

In response, we revised Sect. 4.2 to connect the integrated evaluation results with the meteorological variables and physical-process groups that drive the scheme-selection pathway. The revised Discussion now clarifies that differences between competing configurations are not only numerical changes in the evaluation metric. They also indicate which parameterization choices better reproduce meteorological variables tied to the atmospheric processes of interest.

Specifically, the $\text{MRE} \times \text{CV}$ -selected configuration reduced RMSE for temperature-related variables in Northwest China (TEM by $0.98\text{ }^{\circ}\text{C}$ and GST by $2.47\text{ }^{\circ}\text{C}$) and for precipitation in Southeast China (PRE by 1.38 mm d^{-1}). Because RMSE retains the unit of the evaluated variable, these reductions can be interpreted in the physical units of the corresponding meteorological variables. These variable-unit diagnostics were used only to interpret the practical meteorological meaning of the selected configurations; they were not introduced as an additional scheme-ranking criterion. In other words, the scheme rankings remained based on the cMIEI-based evaluation framework, whereas the variable-unit diagnostics provided an interpretive bridge between the integrated scores and meteorological variables with physical units.

We further expanded the Discussion to translate these variable-level improvements into meteorological implications. Lower temperature-related errors imply a more reliable representation of near-surface thermal conditions and land-atmosphere heat exchange, which is relevant for regional climate simulation and heat-related extreme-event assessment in arid Northwest China. Lower precipitation errors improve the representation of water input and rainfall-forming processes, which is important for hydroclimate simulation and heavy-precipitation evaluation in humid Southeast China. Near-surface wind is also emphasized because it constrains boundary-layer mixing and moisture and heat transport.

We therefore revised the manuscript to state that the proposed weighting framework has practical value for selecting physical-process parameterizations that better represent target meteorological processes and their application-relevant variables. We also retained the appropriate scientific boundary: formal attribution of the underlying mechanisms would still require process-budget diagnostics such as water-vapor budget analysis, hydrometeor diagnostics, turbulent kinetic energy diagnostics, PBL-height evaluation, or surface energy-balance decomposition.

Corresponding revisions:

Section 4.2 Physical interpretation and practical meteorological relevance (page 32, lines 714-715):

The practical meteorological implication of the evaluated differences becomes clearer when the integrated scores are linked to the variables and physical processes that drive the scheme-selection pathway. Differences between competing configurations are not only numerical changes in the evaluation metric; they indicate which parameterization choices better reproduce meteorological variables tied to the atmospheric processes of interest. For example, the proposed $\text{MRE} \times \text{CV}$ configuration reduced RMSE for temperature-related variables in Northwest China (TEM by 0.98 °C and GST by 2.47 °C) and for precipitation in Southeast China (PRE by 1.38 mm d⁻¹). These variable-level improvements translate into practical meteorological meaning because the affected variables control different parts of the regional climate system. Lower temperature-related errors imply a more reliable representation of near-surface thermal conditions and land-atmosphere heat exchange, which is relevant for regional climate simulation and heat-related extreme-event assessment in arid Northwest China. Lower precipitation errors improve the representation of water input and rainfall-forming processes, which is important for hydroclimate simulation and heavy-precipitation evaluation in humid Southeast China. Near-surface wind is also a key constraint on boundary-layer mixing and moisture and heat transport. The proposed weighting framework therefore has practical value for selecting physical-process parameterizations that better represent target meteorological processes and their application-relevant variables, while process-budget diagnostics would still be needed for formal attribution of the underlying mechanisms.

References cited in the responses to Reviewer 3

Morrison, H. and Milbrandt, J. A.: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: Scheme description and idealized tests, *J. Atmos. Sci.*, 72, 287-311, doi:10.1175/JAS-D-14-0065.1, 2015.

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, *Mon. Weather Rev.*, 134, 2318-2341, doi:10.1175/MWR3199.1,

2006.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, 117, 1779-1800, doi:10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2, 1989.

Zhang, C., Wang, Y., and Hamilton, K.: Improved representation of boundary layer clouds over the Southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme, *Mon. Weather Rev.*, 139, 3489-3513, doi:10.1175/MWR-D-10-05091.1, 2011.

Fels, S. B. and Schwarzkopf, M. D.: An efficient, accurate algorithm for calculating CO₂ 15 μ m band cooling rates, *J. Geophys. Res.*, 86, 1205-1232, doi:10.1029/JC086iC02p01205, 1981.

Chen, F. and Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Mon. Weather Rev.*, 129, 569-585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2, 2001.

Lu, S., Guo, W., Xue, Y., Huang, F., and Ge, J.: Simulation of summer climate over Central Asia shows high sensitivity to different land surface schemes in WRF, *Clim. Dyn.*, 57, 2249-2268, doi:10.1007/s00382-021-05876-9, 2021.

Jia, W., Zhang, X., Wang, H., Wang, Y., Wang, D., Zhong, J., Zhang, W., Zhang, L., Guo, L., and Lei, Y.: Comprehensive evaluation of typical planetary boundary layer (PBL) parameterization schemes in China—Part 1: Understanding expressiveness of schemes for different regions from the mechanism perspective, *Geosci. Model Dev.*, 16, 6635-6670, doi:10.5194/gmd-16-6635-2023, 2023.

Lim, K.-S. S. and Hong, S.-Y.: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models, *Mon. Weather Rev.*, 138, 1587-1612, doi:10.1175/2009MWR2968.1, 2010.

Shin, H. H. and Hong, S.-Y.: Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions, *Mon. Weather Rev.*, 143, 250-271, doi:10.1175/MWR-D-14-00116.1, 2015.

Grell, G. A.: Prognostic evaluation of assumptions used by cumulus parameterizations, *Mon. Weather Rev.*, 121, 764-787, doi:10.1175/1520-0493(1993)121<0764:PEOAUB>2.0.CO;2, 1993.

Grell, G. A. and Devenyi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophys. Res. Lett.*, 29, 38-1-38-4, doi:10.1029/2002GL015311, 2002.

Liu, D., Yang, B., Zhang, Y., Qian, Y., Huang, A., Zhou, Y., and Zhang, L.: Combined impacts of convection and microphysics parameterizations on the simulations of precipitation and cloud properties over Asia, *Atmos. Res.*, 212, 172-185, doi:10.1016/j.atmosres.2018.05.017, 2018.

Yang, B., Zhang, Y., Qian, Y., Huang, A., and Yan, H.: Calibration of a convective parameterization scheme in the WRF model and its impact on the simulation of East Asian summer monsoon precipitation, *Clim. Dyn.*, 44, 1661-1684, doi:10.1007/s00382-014-2118-4, 2015.