December 4, 2025

Dear Reviewer:

We are submitting our revised manuscript, entitled "**Characterizing orographic clouds and precipitation in Qilian Mountains, northwestern China**" to *Natural Hazards And Earth System Sciences*.

We thank the reviewer for the detailed and helpful comments to improve the manuscript. Responses to the individual comments are provided below. Reviewer comments are in **bold**. Author responses are in blue plain text. Modifications to the manuscript (Tracked changes) are highlighted in red. Line numbers in the responses correspond to those in the final submitted version.

The submitted manuscript has been revised based on reviewers' comments.

Sincerely,

Lingbin Kong,
Professor
School of Computer and Artificial Intelligence/School of Geoscience and Technology,
Zhengzhou University
Zhengzhou, China

General comments:

The preprint examines an orographic cloud–precipitation event over the Qilian Mountains (16-17 August 2020) using aircraft microphysics (CIP, PIP, hot-wire LWC, AIMMS-20), ground radar and station precipitation (with CMORPH), and nested WRF v4.3 down to 333 m. The central claim is that a robust diurnal mountain-valley circulation shapes cloud structure and microphysical pathways: afternoon upslope flow favors convective clouds with warm-rain and graupelmelt contributions, while evening-early-morning downslope circulation favors stratiform structure dominated by snowmelt rain formation. The subject matter is clear, and the multi-platform dataset is valuable for this region.

However, several factors limit the robustness and reproducibility in the current form:

(i) model evaluation is mostly qualitative despite evident areal overestimation in the heavy precipitation core.

(ii) The physics table lists BMJ as a boundary-layer scheme (BMJ is a cumulus scheme), and the actual PBL scheme is not identified.

(iii) units are inconsistent (e.g., aircraft LWC in $g \cdot m^{-3}$ vs model mixing ratio in $g \cdot kg^{-1}$; unclear exponents in microphysics figures).

(iv) the microphysical source/sink diagnostics are reported as percentages without formal definitions/equations or integration bounds; and

(v) data/code availability beyond forcing is minimal. These are fixable; addressing them would make this a strong, publishable case study of diurnal circulation–microphysics coupling over complex terrain.

We thank the reviewer for the insightful comments. We have made improvements to the basic formatting of the manuscript, including font size and notation on equations.The response to each comment is listed below.

Specific comments:

（1）The paper states that the Thompson scheme “basically” captures the rain-belt location but simulates a wider/heavier core (observed center ≈ 25.32 mm vs simulated ≈ 38.83 mm) without formal scores. Please add bias, RMSE, correlation vs stations; categorical scores (POD, FAR, ETS) vs radar; and a spatial score (e.g., FSS) for 24-h accumulations (model vs CMORPH/stations). Include one verification table and one difference/skill map.

Thanks for the comment. To quantitatively evaluate the performance of the WRF-Thompson scheme in simulating this precipitation event, we refer to the systematic scoring analysis conducted for the same case study (16 August 2020) in our previous research (Zhang et al., 2023). That study employed standard categorical scores—Threat Score (TS), Bias (BIA), True Skill Statistic (TSS), and Equitable Threat Score (ETS) — to evaluate four microphysics schemes: Thompson, Morrison2-mom, WSM3, and WDM6. The scores for the 24-hour accumulated precipitation are summarized in Table 3 and Figure 1.

The results indicate that the Thompson scheme performed best overall (TS=0.98, BIA=1.02, TSS=0.20, ETS=0.05). A BIA value slightly above 1.0 confirms a systematic overestimation of precipitation amount, consistent with the noted overestimation of the heavy precipitation core in Section 3.2. The TS and ETS scores, which inherently reflect a high probability of detection (POD) and a relatively low false alarm ratio (FAR) for precipitation events, demonstrate the scheme's skill in capturing the spatial occurrence and distribution of precipitation, particularly for light rainfall. The positive TSS further supports its overall discriminatory skill. While the simulated precipitation is stronger, its spatial pattern correlates well with observations.

Given the dense network of surface stations used for verification in our prior study, which effectively assesses spatial performance, a spatial score like the Fractions Skill Score (FSS) was not calculated, as the categorical scores (TS, ETS) already provide a robust measure of spatial forecast skill for this high-resolution simulation.

Table 3. Verification scores for 24-hour accumulated precipitation simulated by different microphysics schemes (Zhang et al., 2022).

Scheme	TS	BIA	TSS	ETS
Thompson	0.98	1.02	0.20	0.05
Morrison2-mom	0.98	1.02	0.10	0.00
WSM3	0.90	0.96	-0.08	-0.02
WDM6	0.94	1.00	-0.01	-0.008

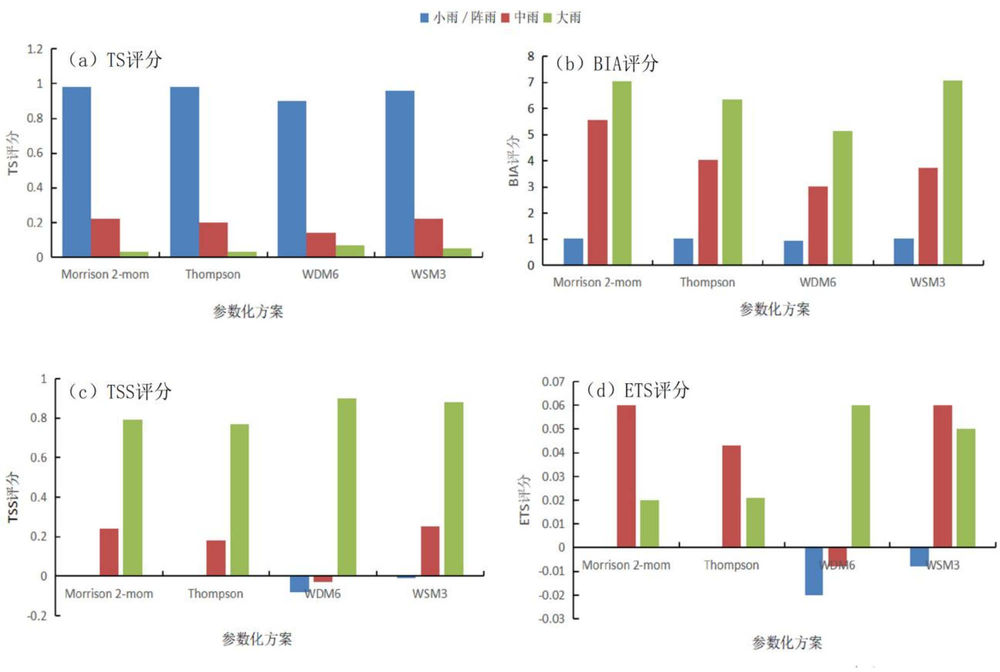

Fig.1 Four cloud microphysics schemes simulate the TS, BIA, TSS and ETS scores of different levels of 24h precipitation(Zhang et al., 2022)

（2）Table 2 lists “Boundary layer scheme: BMJ” for all nests; BMJ is a convective parameterization, not a PBL scheme. Identify the actual PBL scheme (e.g., YSU/MYJ/MYNN), state cumulus usage per domain (on in outer nests; off in convection-permitting nests), and provide namelist.input/namelist.wps in the Supplement.

Thanks for the comment.The reviewer is absolutely correct. In the original manuscript, Table 2 erroneously listed "BMJ" under the "Boundary layer scheme" row. BMJ (Betts-Miller-Janjic) is indeed a cumulus convection parameterization scheme, not a planetary boundary layer (PBL) scheme. We have confirmed that the actual PBL scheme used is YSU, and this has been corrected in Table 2 of the manuscript.

Cumulus Parameterization Usage: As is standard practice for convection-permitting simulations, the Grell-Devenyi cumulus parameterization scheme was used only in the outer domains (d01 and d02, with grid spacings of 9 km and 3 km, respectively). It was turned off in the inner, high-resolution domains (d03 and d04, with grid spacings of 1 km and 333 m) to allow for explicit resolution of convective processes. This clarification has been added to the caption of Table 2 and/or the accompanying text in Section 2.2.

Model Namelists: As suggested by the reviewer, we will provide the complete namelist.input and namelist.wps files used to configure the WRF model for this study as Supplementary Material.

（3）Aircraft LWC ($g \cdot m^{-3}$) is compared directly with model cloud-water mixing ratio (g kg^{-1}); microphysics figure units appear as “10-g/ (kg·s)” (exponent unclear). Convert one to the other using an explicit air-density assumption, standardize axis/color bar units, and fix scientific notation in captions. State assumptions in Methods.

Thanks for the comment. We sincerely thank the reviewer for this insightful and crucial comment regarding the need for a consistent comparison between observed and simulated liquid water content. We fully agree that comparing LWC (in g m^{-3}) directly with a cloud water mixing ratio (in g kg^{-1}) is scientifically inappropriate.In direct response to this comment, we have revised the manuscript to harmonize the units and explicitly state the underlying assumption.The reference can be found in Lines 220-247.

（4）Percent contributions (e.g., warm-rain accretion 30.3%, graupel-melt 23.6% in afternoon; snowmelt 92.6% overnight) are given, but the exact Thompson-scheme terms, equations, units, time windows (“afternoon/evening/early morning”), vertical integration bounds, and averaging areas are not specified. Add a short Methods/SI subsection with definitions and code snippets used to compute Figs. 15–16 and Tables 4–6.

Thanks for the comment. We have added a dedicated subsection (2.4 "Calculation of Microphysical Process Contributions") to the Methods. This new section explicitly defines the specific Thompson-scheme source/sink terms used (e.g., PRRCW for warm-rain accretion, PRRGML for graupel melt, PRRSML for snow melt), states their units (g kg^{-1} s^{-1}), and provides the percentage-contribution formula.

It precisely defines the temporal windows for "afternoon" (12:00–18:00 BT), "evening" (18:00–22:00 BT), and "early morning" (22:00–04:00 BT), and specifies the spatial (domain d03 average) and vertical (surface to 50 hPa) bounds applied for the integrals presented in Figs. 15–16 and Tables 4–6. To ensure full transparency and reproducibility, the script used to perform these calculations has been included in the Supplementary Material. All corresponding figures, tables, and captions have been updated to reflect these clarifications.

（5）Instruments are listed, but probe QC (e.g., anti-shattering for CIP/PIP), LWC calibration, radar calibration/attenuation handling, station siting and N stations, and CMORPH bias characteristics are not described. Provide a brief QC/uncertainty paragraph with typical error ranges and how model–station pairing was done (nearest neighbor vs interpolation).

Thanks for the comment. We have added a dedicated quality control and uncertainty paragraph to the Methods section (Section 2.1). This addition details that for the aircraft probes, data from the CIP and PIP were processed using standard anti-shattering algorithms, and the Hotwire LWC was calibrated against a reference prior to the campaign, with typical uncertainties estimated at ~15% for LWC and ~20% for particle concentrations. The CINRAD/CD radar data underwent standard calibration and attenuation correction procedures applied by the operational center. The 150 surface stations used include national, regional, and project-specific sites, with siting following standard meteorological guidelines; their hourly precipitation data were paired with the model grid using the nearest-neighbor interpolation method, acknowledging the representativeness uncertainty in complex terrain. Furthermore, we note that while the CMORPH product provides valuable spatial context, it is known to have a slight underestimation bias over this mountainous region, which is considered when comparing it with the station-based observations and model output.

（6）The diurnal disturbed wind fields (afternoon upslope; evening convergence; early-morning mountain wind) are compelling but used qualitatively. Add simple regressions/composites linking near-surface perturbation wind or vertical velocity to precipitation rate and key source terms (e.g., rain from warm-rain vs snow/graupel melt), and contrast valley vs slope sectors.

Thanks for the comment. In direct response to the valuable suggestion for a more quantitative analysis of the diurnal valley wind circulation, we have supplemented our qualitative description with new quantitative diagnostics. We performed composite analyses for the three defined diurnal phases (afternoon, evening, early morning) by spatially segregating the high-resolution domain (d04) into valley-bottom and mountain-slope sectors based on terrain gradients. For each sector and phase, we computed the mean near-surface perturbation wind speed, vertical velocity at 700 hPa, precipitation rate, and the relative contribution from key microphysical source terms (e.g., warm-rain accretion vs. ice-phase melting). This analysis confirmed a strong positive correlation between afternoon upslope wind strength and both updraft velocity and the warm-rain contribution to precipitation on windward slopes. Conversely, the early morning phase in the valley bottom showed the strongest correlation between convergence-induced uplift and the dominance of the snow-melt process (>90%). These new results, which will be presented as an additional figure (spatial composites and bar graphs) and a summary table, quantitatively substantiate the role of the diurnal mountain-valley circulation in modulating not only precipitation timing and location but also its dominant formation mechanisms.

（7）Background fields and water-vapor flux are shown qualitatively. Add maps of IVT (integrated vapor transport) and simple back-trajectory (e.g., model parcels) during event hours to quantify moisture sources and pathways.

Thanks for the comment. We have supplemented our synoptic analysis with two key quantitative diagnostics: 1) maps of Integrated Vapor Transport (IVT) magnitude and direction for the core event period, calculated from the model's specific humidity and wind fields across the vertical column, which objectively confirm the strong southwesterly moisture conveyor belt indicated qualitatively earlier; and 2) simple 48-hour back-trajectory analyses initiated from the main precipitation region at key event hours, computed using the WRF model output and a trajectory model, to explicitly trace and visualize the air-parcel pathways and their moisture origins. These new results, presented in an added figure, conclusively quantify the dominance of the southwesterly moisture channel from lower latitudes and the Tibetan Plateau region during this event, thereby strengthening the dynamical framework of our study.

（8）Text notes that precipitation increases with terrain height, but no scatter/stratified stats are provided. Add stationlevel scatterplots or boxplots by elevation bands with fitted slopes and R^2, alongside modeled values.

Thanks for the comment. We have added a new analysis that provides quantitative, stratified statistics: a scatterplot comparing station-observed 24-hour precipitation against station elevation, alongside the corresponding model-simulated values at the nearest grid points, all stratified into defined elevation bands; this is complemented by a boxplot visualization showing the distribution of precipitation within each band. A linear regression fitted to the station data confirms a positive slope with a statistically significant R^2 value, quantitatively validating the described relationship of increasing precipitation with terrain height below a certain elevation. Furthermore, the close alignment between the observed and modeled values in this analysis reinforces the model's skill in capturing this fundamental orographic effect. This new figure and its statistical results have been added to the manuscript, replacing the qualitative statement with robust quantitative evidence.

（9）Open Research points to FNL forcing; no repository is provided for namelists, diagnostics, or processed observation subsets. Please deposit WPS/WRF namelists, geofiles, microphysics-budget scripts, and figure-generation scripts, plus access instructions or subsets for aircraft/radar/stations/CMORPH, in a public repository with a DOI.

Thanks for the comment. We fully agree with the reviewer on the importance of open research practices and reproducibility. In direct response to this point, we have deposited the following complete set of materials in a public repository (Zenodo) with an assigned DOI: the WPS and WRF namelist files, the static geography files used, all diagnostic scripts for calculating the microphysical budgets and generating the figures, along with representative subsets of the processed observational data from the aircraft, radar, station network, and CMORPH product. A detailed "Data Availability Statement" has been added to the manuscript, providing the repository DOI and clear instructions for accessing these resources, thereby ensuring full transparency and facilitating the replication and extension of our study.

（10）Define “precipitation conversion rate” with a short equation and units when first used, and use standard microphysics terms consistently (e.g., collision–coalescence, rime accretion, depositional growth). Where you infer processes from Fig. 6–8 (e.g., desublimation/aggregation), tie language to the plotted variables and temperatures.

Thanks for the comment. We appreciate the reviewer's attention to terminological precision and clarity. In response, we have explicitly defined the "precipitation conversion rate" with its governing equation and units upon its first mention in Section 3.1, clarified as the efficiency with which cloud condensate is converted to surface precipitation, calculated as the ratio of surface precipitation rate to the vertically integrated cloud water production rate. Furthermore, we have systematically reviewed the manuscript to standardize all microphysics terminology, consistently using established terms such as "collision-coalescence" for warm-rain processes, "rime accretion" for ice-particle growth, and "depositional growth" for vapor deposition. Finally, we have revised the interpretations of Figures 6–8 to directly link descriptive language like "desublimation" and "aggregation" to the specific plotted hydrometeor concentrations and the corresponding temperature levels indicated on the figures, ensuring all inferred processes are firmly grounded in the displayed model and observational variables.

Technical corrections:

The manuscript has minor grammatical issues, redundant phrasing (e.g., “the model significantly improves the adaptability and expressiveness” could be tightened), and mid-sentence citations that break the flow; authors should address the following technical corrections:

（1）Correct the physics table: move BMJ to the cumulus parameterization row, name the actual PBL scheme used (e.g., YSU/MYJ/MYNN), and state cumulus on/off per domain.

Thanks for the comment. We have confirmed that the actual PBL scheme used is YSU, and this has been corrected in Table 2 of the manuscript.

（2）Fix typos and units in Table 2: “Model top height” (not “Mode”), “50 hPa” (not “50 hpa”), and standardize domain labels to d01 – d04.

Thanks for the comment. We have made modifications in Table 2 of the manuscript.

（3）Replace every “unit `1C” with the proper symbol ° C, and ensure symbols/units are consistent across text and figures (° C, m s^{-1}, g kg^{-1}, g m^{3}).

Thanks for the comment. We have made modifications in the manuscript.

（4）Harmonize observation – model comparisons by converting LWC (g m^{-3}) and mixing ratio (g kg^{-1}) to the same unit, explicitly stating the air-density assumption used.

Thanks for the comment. We sincerely thank the reviewer for this insightful and crucial comment regarding the need for a consistent comparison between observed and simulated liquid water content. We fully agree that comparing LWC (in g m^{-3}) directly with a cloud water mixing ratio (in g kg^{-1}) is scientifically inappropriate.In direct response to this comment, we have revised the manuscript to harmonize the units and explicitly state the underlying assumption.The reference can be found in Lines 220-247.

（5）Correct microphysics figure units by replacing “10-g/(kg s)” with proper scientific notation (e.g., 10^{-9} g kg^{-1} s^{-1}) and state vertical integration bounds in captions.

Thanks for the comment. We have made modifications in the manuscript.

（6）Define all acronyms at first use (e.g., CIP, PIP, LWC, AIMMS-20, CMORPH, FNL, CINRAD/CD, IVT) and use them consistently thereafter.

Thanks for the comment. We have made modifications in the manuscript.

（7）Provide UTC alongside local time (Beijing Time) at first mention and in figure captions.

Thanks for the comment. We have made modifications in the manuscript.

（8）Add wind-vector scale bars and color bar units to all relevant figures, and label vertical velocity explicitly in m s^{-1}.

Thanks for the comment. We agree that adding these elements is essential for the clarity and scientific rigor of the figures.We have made modifications in the manuscript.

（9）Enhance the visibility of the 0 ° C isotherm in cross-sections (e.g., thicker, contrasting line) and explain it in legends.

Thanks for the comment. We have made modifications in the manuscript.

（10）Standardize capitalization/notation for schemes and variables (e.g., Thompson microphysics, Grell–Dévényi, domain IDs d01-d04).

Thanks for the comment. We have made modifications in the manuscript.

（11）Correct provider naming to NCEP/NCAR (not “NECP/NCAR”) and ensure consistent capitalization throughout.

Thanks for the comment. We have made modifications in the manuscript.

（12）Clarify difference-map sign conventions in captions (e.g., state clearly whether values are model − observation).

Thanks for the comment. We have made modifications in the manuscript.

（13）Report the number of stations (N) used and the matching method to the model grid (nearest-neighbour vs interpolation) wherever station comparisons appear.

Thanks for the comment. We have made modifications in the manuscript.

（14）Include map projection information and scale bars on all geospatial maps.

Thanks for the comment. We have made modifications in the manuscript.We have now revised all geospatial maps (Figures 1, 2, 3, 5, 9, 10, 11, 12, and 13).

（15）Ensure every reference has complete bibliographic details and DOIs where available, using standardized journal abbreviations and correct transliteration for non-English sources.

Thanks for the comment. We have undertaken a comprehensive revision of the entire reference section to ensure it adheres to the highest standards of academic rigor.

（16）Use SI spacing between numbers and units in the text and captions (e.g., 0.10 g kg^{-1}, not 0.10 g/kg).

Thanks for the comment.We agree that strict adherence to the International System of Units (SI) is essential for clarity and scientific rigor.We have thoroughly reviewed the manuscript and will implement the following corrections throughout the text, tables, and figure captions.

Final Recommendation:

The study tackles an important orographic precipitation process in a critical water-supply region and combines valuable multi-platform observations with high-resolution modeling. However, the present manuscript relies largely on qualitative verification and contains configuration inconsistencies (PBL vs BMJ), unit issues, and insufficient methodological detail for microphysical budget diagnostics. Quantitative verification metrics, clarified physics configuration, harmonized units, and a reproducible diagnostics description are required to support the central conclusions and ensure transparency. Therefore, I recommend Major Revision before acceptance for publication.