Author response to reviewer Jenna Ritvanen

Dear Reviewer,

We thank you for the careful and constructive review of our manuscript. Your comments have helped us identify several areas that require clearer documentation, additional analysis, and more cautious interpretation. We will revise the manuscript accordingly. Our point-by-point responses are provided below.

1. Focus of the literature review

We agree that the current literature review places too much emphasis on rainfall prediction, which may obscure the main objective of the study. The purpose of the manuscript is not to develop a rainfall-prediction or nowcasting model. Rather, it introduces a continuity-equation-based diagnostic framework in which the difference between an observed rainfall field and an advection-only reference field is used to identify local rainfall evolution that cannot be explained by horizontal transport alone.

We will therefore restructure the review around methods for identifying and characterizing convective rainfall. We will distinguish among intensity- or microphysics-based convective/stratiform classification, radar- and image-based cell tracking, advection-based rainfall-field evolution, and rain-gauge-based rainfall analyses. We will also clarify that radar-object tracking and nowcasting methods generally operate on spatially continuous radar or image fields, whereas our framework operates on gauge-derived rainfall fields and diagnoses local rainfall growth relative to an advection-consistent baseline. We will add relevant studies using comparable rain-gauge data where the measurement and quality-control definitions permit a meaningful comparison.

2. Rain-gauge observations and data quality

2.1 Instrument type

We agree that the observing system requires clearer documentation. The measurements were obtained from Bangkok Metropolitan Administration (BMA) automatic tipping-bucket rain gauges. BMA confirmed that the network uses the TB3-0.5 model. The available manufacturer and BMA/DDS documentation supports a 0.5 mm bucket capacity/resolution. We will add these details and their sources to the manuscript.

We will also state the limitation of the available metadata. The analysis archive contains station identifiers, timestamps, 15 min rainfall accumulations, processing indicators, and station coordinates, but it does not contain station-level serial numbers, calibration histories, or a complete station-by-station instrument-change log.

2.2 QA/QC and outlier detection

We agree that the numerical rules and missing-data treatment should be stated explicitly. In the author-side cleaning workflow documented in the analysis notebooks, records were aligned to a common 15 min time grid and duplicate timestamps were checked. Missing or empty values were assigned an internal missing flag, negative rainfall values were assigned an error flag, and values exceeding 500 mm in one 15 min record were assigned an abnormal-value flag. Values with the error or abnormal flag were converted to missing values. Missing runs of up to six 15 min slots were time-interpolated. Runs of seven to ten slots were forward-filled from the preceding available value; backward fill from the following value was used only when no preceding value was available, such as at the start of a station series. Longer runs remained missing. We will document these rules and make clear that the 500 mm limit was a legacy operational screening bound, not a climatologically calibrated threshold for Bangkok rainfall.

We will also distinguish this original cleaning step from later diagnostic audits of the cleaned archive, which used separate range settings and must not be presented as though they were the original outlier rule. The notebooks document the author-side data screening but not every provider-side quality-control decision, so we will not present the procedure as a complete station-level BMA quality-control audit.

2.3 Spatial coverage and grid cells without gauges

We agree that a 1 km computational grid must not be interpreted as though every cell were directly observed by a gauge. We will add a spatial-support figure showing (i) the number of geolocated gauges in each 1 km grid cell and (ii) the distance from every grid-cell center to the nearest gauge. The full rectangular domain contains 71 by 57 cells, of which 1,560 cell centers are inside the Bangkok boundary. Within that boundary, 113 cells contain at least one geolocated gauge in the same 1 km cell, while 27 cells are beyond the 7.5 km smoothing-support radius used in the reconstruction. In the revised figure, cells outside this radius will be explicitly marked as having no smoothing support, and cells without a gauge in the same grid cell will be identified as having limited direct observational support. We will consistently describe the resulting fields as smoothed, gauge-derived rainfall fields rather than direct measurements at every grid cell.

2.4 Interpretation of the Table 1 data-quality categories

We agree that the original term "outliers" was misleading. The reported category combined records not retained as valid rainfall observations, including missing or unavailable records and records excluded by parsing, flags, or range checks; it did not represent physically implausible meteorological values alone. The aggregate therefore cannot be interpreted as the fraction rejected solely by the numerical bounds described above. We will relabel the Table 1 category as a historical aggregate of records not retained during cleaning, state that the archived fields do not permit an unambiguous retrospective decomposition of this aggregate, and distinguish it from the separately reported observed dry and valid non-zero rainfall categories.

For context, Ośródka et al. (2022) described a 503-station network operated by the Polish national meteorological and hydrological service, using 10 min measurements and both heated and unheated sensors at most stations. Fewer than 1.5% of observations received a reduced quality index. The fraction assigned a zero quality index—treated as erroneous and rejected—was approximately 0.4% in July and about 0.10%–0.23% in the other seasons. Those values are not directly comparable with the original Table 1 percentage: RainGaugeQC also uses neighbouring gauges, radar information, and a quantitative quality index, whereas our historical category combines missingness and several cleaning outcomes. This comparison reinforces why our percentage should not be described as a meteorological-outlier rate or classified as "normal" or "exceptional" against another network without harmonized definitions. We will add this context and state the limitation explicitly.

3. Temporal consistency of the rainfall fields

We agree that temporal consistency is an important prerequisite for applying optical flow. We computed lagged spatial correlations between rainfall fields separated by 15, 30, 45, and 60 min for all available storm-event grids. Across 1,825 usable 15 min frame pairs, the median lag-1 correlation was 0.48, with an interquartile range of 0.11–0.80. Median correlations decreased to 0.10 at 30 min, approximately 0.00 at 45 min, and -0.014 at 60 min.

We will add this analysis to the revised manuscript. The result indicates moderate short-lag spatial coherence and supports the use of consecutive 15 min fields for constructing the optical-flow reference. It does not, however, independently validate the optical-flow field as a meteorological wind product. We will make this distinction explicit and interpret the residuals as storm-window diagnostics, particularly where short-lag coherence is weak.

4. Smoothing algorithm and sensitivity

4.1 Smoothing algorithm

We will describe the field reconstruction as normalized Gaussian distance weighting of station observations on the 1 km grid. For a grid cell, the weight assigned to an observation at distance d is $w(d) = \exp[-d^2/(2\sigma^2)]$. The analysis uses $\sigma = 2.5$ km and truncates the kernel at 3σ, giving a maximum support radius of 7.5 km. The purpose is to create a spatially coherent and differentiable field for the optical-flow and residual calculations, not to claim an independent high-resolution rainfall product.

4.2 Sensitivity to the smoothing scale

We repeated the residual calculation for a three-storm subset using $\sigma = 1.0$, 1.5, 2.5, 3.5, and 5.0 km while keeping the other settings fixed. As expected, spatial-footprint metrics were sensitive to smoothing: the median peak positive-residual area increased from 466 grid cells at $\sigma = 1.0$ km to 1,166 cells at $\sigma = 5.0$ km; it was 775 cells at the default $\sigma = 2.5$ km. All three storms retained the same convective class across the five scales, and the number of strong-residual frames remained between 12 and 15. We will present these results as a limited subset sensitivity test: they demonstrate sensitivity in the magnitude and spatial extent of the residual, while suggesting qualitative classification stability for the cases tested. They should not be interpreted as a complete robustness assessment of all storms.

4.3 Smoothing-only versus optical-flow advection residuals

We also calculated a no-advection baseline by differencing consecutive smoothed rainfall fields without advecting the earlier field. For the same subset at $\sigma = 2.5$ km, the median event-integrated positive residual was larger for the no-advection baseline than for the optical-flow residual (4,311 versus 3,693 in grid-cell-time residual-sum units), and the median peak positive-residual area was also slightly larger (818 versus 775 grid cells). This indicates that optical-flow advection removes part of the frame-to-frame difference, although positive residual signatures remain. We will add this comparison and clarify the separate roles of smoothing and advection.

5. Validity and interpretation of the estimated motion fields

5.1 Low estimated speeds

We audited the unit conversion in the optical-flow and plotting pipelines. The algorithm estimates displacement between consecutive rainfall grids, which is converted from grid cells per 15 min frame to physical speed using the 1 km grid spacing and 15 min interval. One grid-cell displacement in one frame is therefore approximately 1.11 m s^{-1}. The reported p90 speed of about 0.64 m s^{-1} corresponds to a sub-grid displacement of approximately 0.6 cell over 15 min. The plotted vector scale uses the same conversion and is internally consistent.

Nevertheless, because the vectors are estimated from smoothed, gauge-derived rainfall fields, we agree that they should not be interpreted as independently observed large-scale storm-propagation speeds. We will revise the text to state this limitation clearly.

5.2 Temporal consistency of the motion fields

We computed lagged consistency statistics for the derived motion-speed fields. Across 3,150 usable 15 min frame pairs, the median lag-1 speed correlation was 0.62, with an interquartile range of 0.36–0.79. The median correlation decreased to 0.34 at 30 min, 0.19 at 45 min, and 0.12 at 60 min. Directional cosine similarity was less informative because many cells contained weak or near-zero vectors. We will report these results as an internal temporal-consistency check, not as external validation of atmospheric wind or storm motion.

5.3 Claim of spatially coherent motion

No independent radar, satellite, model-wind, or other external motion product was used for validation. We agree that the original caption overstated the evidence. We will remove the claim that the vectors have been shown to be consistent with precipitation-system advection and will instead describe them as optical-flow displacement estimates derived from consecutive gauge-based rainfall fields for constructing an internally consistent advection-only reference state.

6. Justification of the classification thresholds

We agree that the original manuscript did not document the threshold-selection process adequately. We will explain that the thresholds are operational criteria derived from exploratory analysis of storm-level residual-feature distributions, rather than universal physical constants or thresholds optimized against independent convective labels.

The analysis includes 157 storm episodes. We examined sorted metric curves, empirical percentiles, knee points, and pairwise feature-space plots. The 500-grid-cell threshold for peak positive-residual area lies close to the estimated knee at 513 cells. The event-integrated positive-residual threshold of 4,000 lies near the upper quartile (4,314) and below the estimated knee (4,703). The peak positive-residual-sum threshold of 1,000 lies between the median (766) and upper quartile (1,332), while the threshold of 10 strong 15 min frames lies slightly above the median of 9 and below the upper quartile of 13 frames.

The peak positive-residual area is the primary indicator of spatially extensive rainfall growth. The other metrics provide supporting evidence and reduce the risk of treating isolated spatial artifacts or single-frame spikes as convective storm windows. Application of these operational rules identifies 36 convective and 121 non-convective storm episodes. We will add the exploratory plots and explain both the empirical basis and the absence of independent-label calibration.

7. Statistical analysis

7.1 Units

We will add explicit units throughout the definitions, tables, captions, and axes. Storm duration will be reported in minutes; peak positive-residual area as the number of 1 km grid cells; event-integrated positive residual as a grid-cell-time residual sum; strong-residual persistence as the number or fraction of valid 15 min frames; and rainfall in its native 15 min accumulation interval, with hourly equivalents supplied only where useful for comparison.

7.2 Storm duration and within-storm persistence

The storm definition links gauge-observed rainfall episodes using an allowed gap of up to 120 min; it does not track a single cloud object. A storm window may therefore include initiation, mature, and decay phases. We agree that a storm can show a convective residual signature during only part of its duration and later evolve in a more transport-dominated manner. The storm-level label indicates sustained local rainfall-growth activity somewhere within the storm window; it does not assert that every 15 min frame is convective.

Convective storm windows have a median duration of 247.5 min (IQR 191.25–285.0 min), compared with 165.0 min (IQR 120.0–240.0 min) for non-convective windows. The median number of strong-residual frames is 12.5 (IQR 11.0–15.0) for convective windows and 8.0 (IQR 6.0–12.0) for non-convective windows. Relative to the valid residual-analysis frames, the median strong-frame fractions are 0.85 (IQR 0.81–0.91) and 0.75 (IQR 0.62–0.83), respectively. We will add these statistics and clarify the storm-window interpretation.

7.3 Rainfall-intensity distributions

We will also report rainfall-intensity summaries to facilitate comparison with intensity-threshold methods. The median event maximum 15 min accumulation is 24.25 mm per 15 min for convective storms and 13.0 mm per 15 min for non-convective storms, equivalent to simple hourly rates of 97.0 and 52.0 mm h^{-1}, respectively. The corresponding 90th percentiles are 36.5 and 26.0 mm per 15 min. We will emphasize that these values describe the groups produced by the residual-based classification; rainfall intensity alone does not define the classes.

8. Visualization improvements

We agree with all of these recommendations. We will report the full domain size (71 by 57 cells at 1 km resolution) in the relevant figure or caption. Figures showing the same event will use a common rainfall-intensity color scale wherever the plotted quantities are directly comparable, while residual fields will retain a scale appropriate to signed residual values. We will redraw the storm-count heatmap so that zero and one are visually distinguishable, replace script-variable names with reader-facing labels and titles, and remove unreferenced storm-ID point labels.

9. Repeated content

We agree that parts of the methods, operational framework, results, and conclusion repeat information unnecessarily. We will streamline Sections 4.1 and 4.2 so that they refer back to the formulation in Section 3 rather than re-deriving it, reduce repetition between Sections 7.4 and 6, remove the duplicated results around lines 587–589, and revise the conclusion to emphasize the implications and limitations rather than restating the full method.

10. Opportunistic rainfall measurements

We agree that opportunistic observations are a valuable future direction. In principle, the framework could use other dense surface rainfall observations, including personal weather stations or microwave-link-derived rainfall estimates, because it operates on rainfall-field evolution rather than on one specific gauge type. However, such observations would require additional quality control, calibration and bias correction, metadata screening, temporal synchronization, and assessment of spatial and temporal availability. Data provenance and sensor credibility may also vary substantially among providers. We will add this discussion to the conclusion and make clear that heterogeneous opportunistic measurements cannot be treated as a simple drop-in replacement for a maintained gauge network.

11. Claim concerning convective initiation from gauges alone

We agree that the original wording was too strong. Rain gauges measure surface rainfall but do not directly observe cloud-scale initiation, vertical structure, updrafts, or microphysical processes. We will therefore replace the claim that convective initiation can be analyzed using rain gauges alone with a more precise statement: dense rain-gauge observations can provide surface-based evidence of localized rainfall growth or emergence consistent with convective development. We will apply this more cautious wording consistently in the abstract, introduction, discussion, and conclusion.

We again thank the reviewer for these detailed and constructive comments. The revised manuscript will incorporate the clarifications, additional analyses, figures, and limitations described above.

Reference cited in this response

Ośródka, K., Otop, I., and Szturc, J.: Automatic quality control of telemetric rain gauge data providing quantitative quality information (RainGaugeQC), *Atmospheric Measurement Techniques*, 15, 5581–5597, https://doi.org/10.5194/amt-15-5581-2022, 2022.