Reviewer #1

#RC1.1 This paper presents a method to interpolate daily (gauges) precipitation data using variograms derived from climate model simulations. The manuscript is well written, and results show potential in the proposed methodology to improve daily precipitation estimations, when gridded rainfall fields such radar-based rainfall estimations may not be available.

We thank the reviewer for this positive feedback.

#RC1.2 In addition of the methods of validation presented by the authors, I suggest adding a comprehensive comparison between the anisotropic variograms derived from CP-RCM (the target of this paper) with those from the radar-derived precipitation analyses. See more details below. This additional comparison targets directly the approach presented in the manuscript and may provide evidence of the advantages and limitations of the proposed technique.

Thank you very much for this suggestion. We will compare anisotropic parameters between radar and CP-RCM fields in the sub-section 4.4.

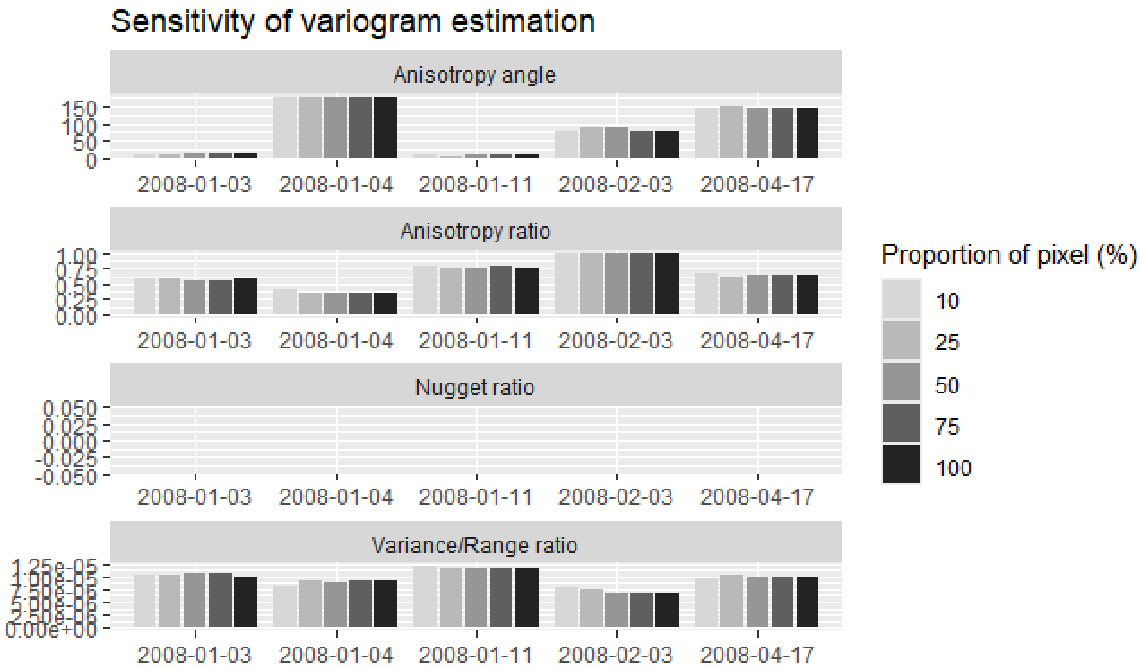
Additional comments:

#RC1.3 L134: Please describe what a Trans-Gaussian Random Field is, as I believe this is the first time the reader is introduced to this term.

Thank you for this recommendation. We introduce Trans-Gaussian Random Field L134: "To handle the strong non-Gaussianity of daily precipitation (positivity and skewness), we assume that the rainfall field can be represented as a Trans-Gaussian Random Field. That is, there exists a transformation such that the transformed field follows a second-order stationary Gaussian random field.".

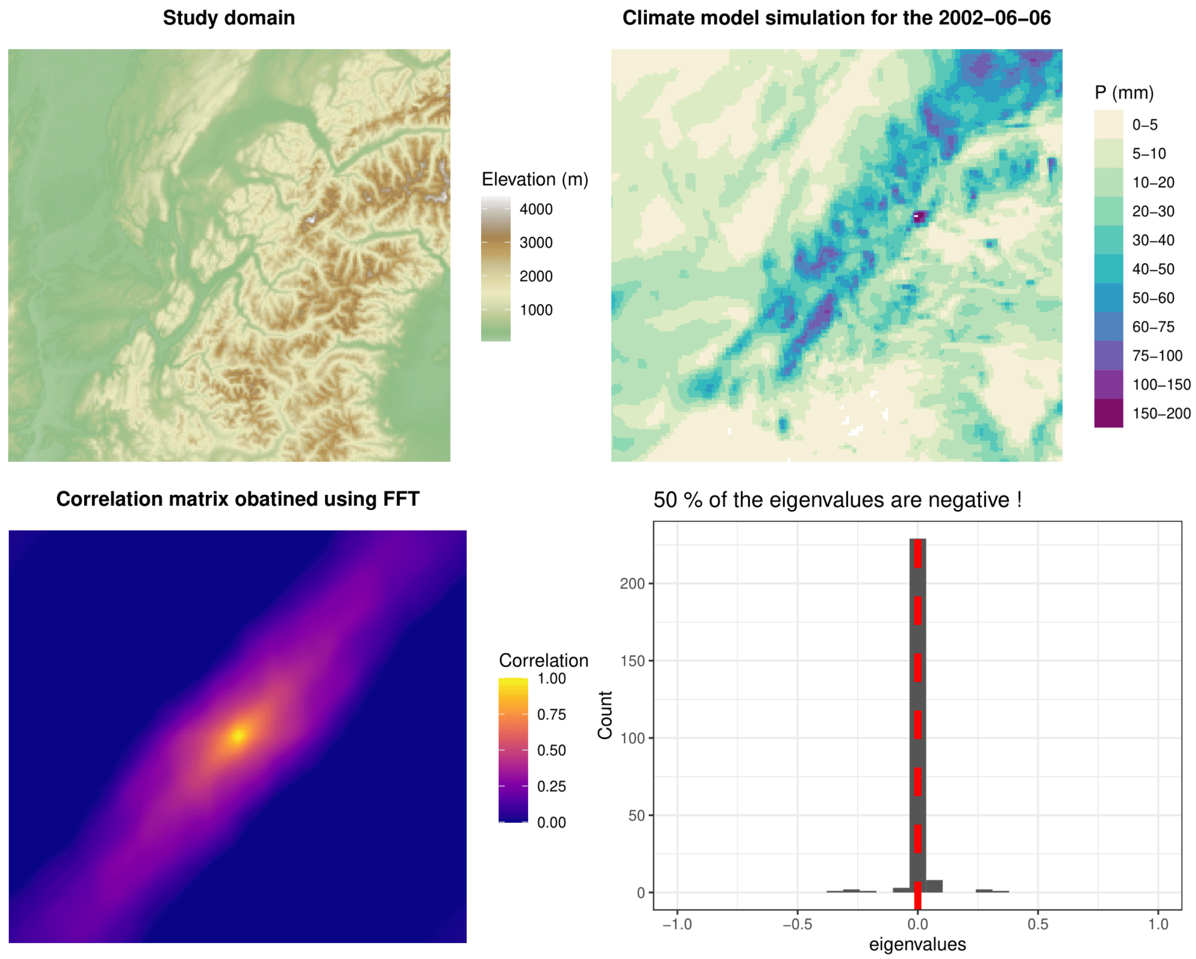
#RC1.4 L154: Please elaborate why only 25% of AROME grid cell were selected to calculate the variograms. Were the selected cells from AROME used as 'virtual' gauges to calculate the variograms? Velasco-Forero et al 2009 and other papers describe methods to estimate 2D variograms using all the grid points from radar images that could be applicable to estimate variograms from AROME and COMEPHORE datasets.

We agree with the reviewer that this part needs clarification. We selected only 25% of AROME grid cells to compute variograms in order to balance spatial representativeness and computational cost. These selected cells were treated as virtual gauges when computing empirical variograms. To confirm robustness, we performed a sensitivity test using 100% of cells for a subset of events (5 first events) , which produced similar variogram parameter estimates.



Non-parametric 2D variograms can be indeed estimated with all the grid points using FFT. We tried this approach in a first instance in another region, but the obtained covariance matrices were not semi-positive definite, which is needed for geostatistic applications.

Here are the steps we performed and the diagnostics.



Step 1 (optional): Normalizing 2D spatial data, removing the mean and dividing by the standard deviation.

Step 2 : 0-padding of the data.

Step 3: Apply FFT.

Step 4: Multiply the FFT result by its conjugate to get the spectral density matrix.

Step 5: Check that all of the values are positive. This is the case.

Step 6: Divide each value by the sum of the spectral densities to sum to 1.

Step 7: Take the inverse FFT.

Step 8: Take its real components.

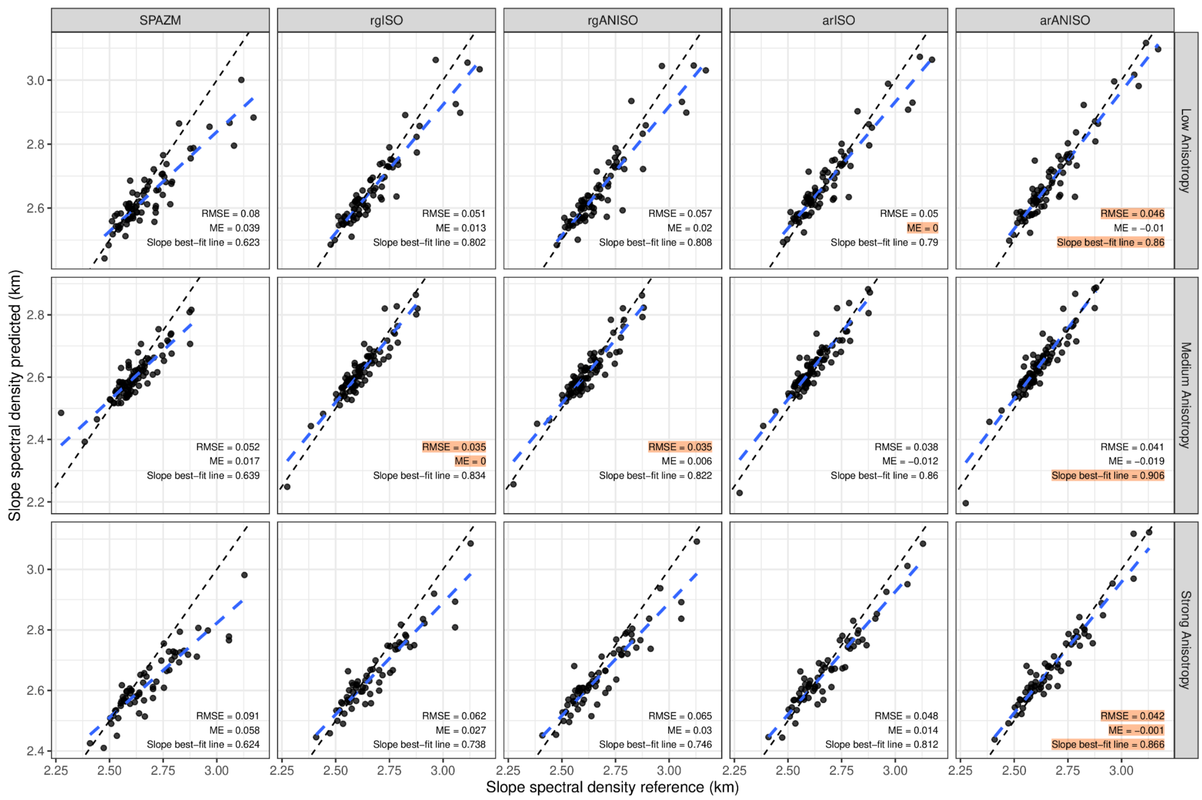
Step 8: Shift the table.

Step 9: Take the central part to remove padding.

We add L154: “We selected only 25% of AROME grid cells to compute variograms in order to balance spatial representativeness and computational cost. These selected cells were treated as virtual gauges when computing empirical variograms. To confirm robustness, we performed a sensitivity test using 100% of cells for a subset of events , which produced similar anisotropy parameter estimates (not shown).”

#RC1.5 L195: Authors are using TWS to verify the spatial structure of the precipitation fields, however spatial multi-scale dependencies are key characteristics of any rainfall fields and authors should add comparisons to account these effects. Seppo Pulkkinen et al. 2019 presents some examples on how to evaluate different rainfall fields based on their multi-scale characteristics (for example figure 8) GMD - Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0)

We thank the reviewer very much for this recommendation. We compare the slopes of the radius-averaged spectral density, using COMEPHORE data (reference) and the other gridded precipitation products.



We add L201: "Additionally, spectral analysis allows comparison of spatial variability across scales and is well suited to assess whether predictions preserve the multiscale structure of precipitation. The two-dimensional Fourier power spectrum was computed for the reference (COMEPHORE) and predicted fields and averaged radially in wave-number. For scaling processes, the radius-averaged spectrum follows a power-law relationship : $E(k) \propto k^{-\beta}$, where k is the wave number and β is the spectral slope. The slope β, estimated from a linear fit in log–log space, was used as the comparison metric. Similar values of β indicate similar spatial variability between reference and predicted fields. We compare the slopes using ME, RMSE, and slope best-fit line metrics. Whereas TWS evaluates short-range gradient image similarities, this comparison evaluates spatial multi-scale dependencies of the daily precipitation fields."

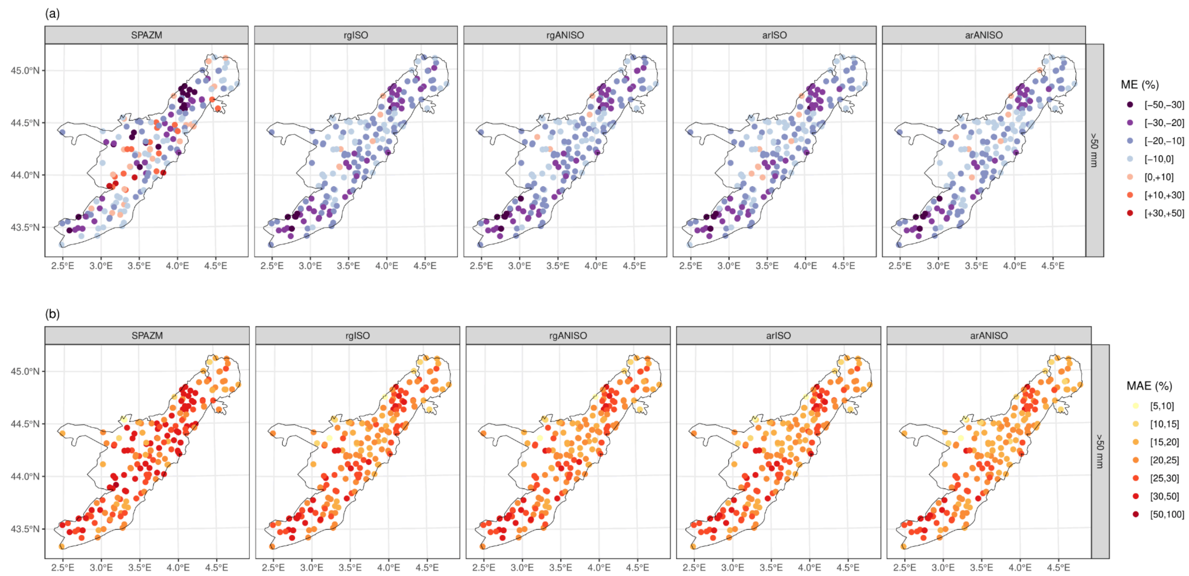
We add L44: "Figure 4 presents the spectral slopes of the reference and predicted precipitation fields. SPAZM systematically underestimates the highest slopes, indicating an overestimation of fine-scale spatial variability. Compared to rgISO, rgANISO degrades the representation of spatial variability for weakly anisotropic events and provides only limited improvement for strongly anisotropic events. In contrast, arISO and arANISO improve the representation of spatial variability, particularly for strongly anisotropic events. Among them, arANISO better reproduces spatial variability, as the highest spectral slopes are no longer underestimated."

#RC1.6 L215: Please indicate where "the ensemble means, a sample of conditional simulations, …" are show (Figure, section???)

We add L215: "In subsection 4.4, we show ensemble means … ".

#RC1.7 Figure 2: It is hard to discriminate the ME and MAE colours from the topographical background. Please try to use contours for the topography, so score colours become more visible. For the discussion of results of this figure (ME, MAE) please consider if a scatterplot between elevation and scores could help to support your conclusions. If elevation is not relevant here, then please consider removing the topography of the figure.

We thank the reviewer for this graphical advice. Elevation is not relevant here, so we removed the topography of the figure.



#RC1.8 L235: It is not true that "rgANISO does not outperform rgISO in gauge gradient similarities" as rgANISO TWS score values are mostly lower that rgISO values for the 66 events with strong anisotropy as shown in Figure 3. Also Figure 3 shows that arANISO generally outperforms rgANISO and arISO also generally outperformes rgISO, which could highlight the advantages of using AROME fields to estimate the spatial variability of the rainfall fields.

We thank the reviewer for pointing out this lack of precision. We replace L237-243 by:

“TWS scores show that SPAZM exhibits lower image gradient similarity to COMEPHORE than all geostatistical approaches, regardless of the event type, indicating a poorer representation of spatial gradients.

For both isotropic and anisotropic configurations, AROME-based methods (arISO and arANISO) consistently outperform their rain-gauge-only counterparts (rgISO and rgANISO). This systematic improvement suggests that the use of AROME fields allows a more accurate estimation of the precipitation covariance structure and, consequently, a better reproduction of spatial gradients.

When considering rain-gauge-based methods only, rgANISO outperforms rgISO for strongly anisotropic events, whereas the opposite behavior is observed for weakly anisotropic events. This indicates that accounting for anisotropy can improve gradient similarity when anisotropy is pronounced, but the limited robustness of anisotropic variogram estimation from rain gauges alone prevents drawing firm conclusions for weakly anisotropic cases.

In contrast, arANISO consistently outperforms arISO, with a marked improvement for strongly anisotropic events and a slight but systematic improvement for weakly anisotropic ones. As no robustness issues arise in the estimation of anisotropic variograms when using AROME fields, these results provide stronger evidence of the added value of explicitly accounting for anisotropy in the representation of precipitation spatial variability.

COMEPHORE relies on kriging for stratiform precipitation and radar observations for convective precipitation. While COMEPHORE uses isotropic covariance in the kriging step, anisotropy is introduced through the radar component. This suggests that anisotropic variograms, particularly when informed by AROME fields, help better reproduce the spatial structure of radar-derived precipitation patterns."

#RC1.9 L251: last sentence should indicate with dataset is used to estimate the anisotropic covariance. Is from AROME?

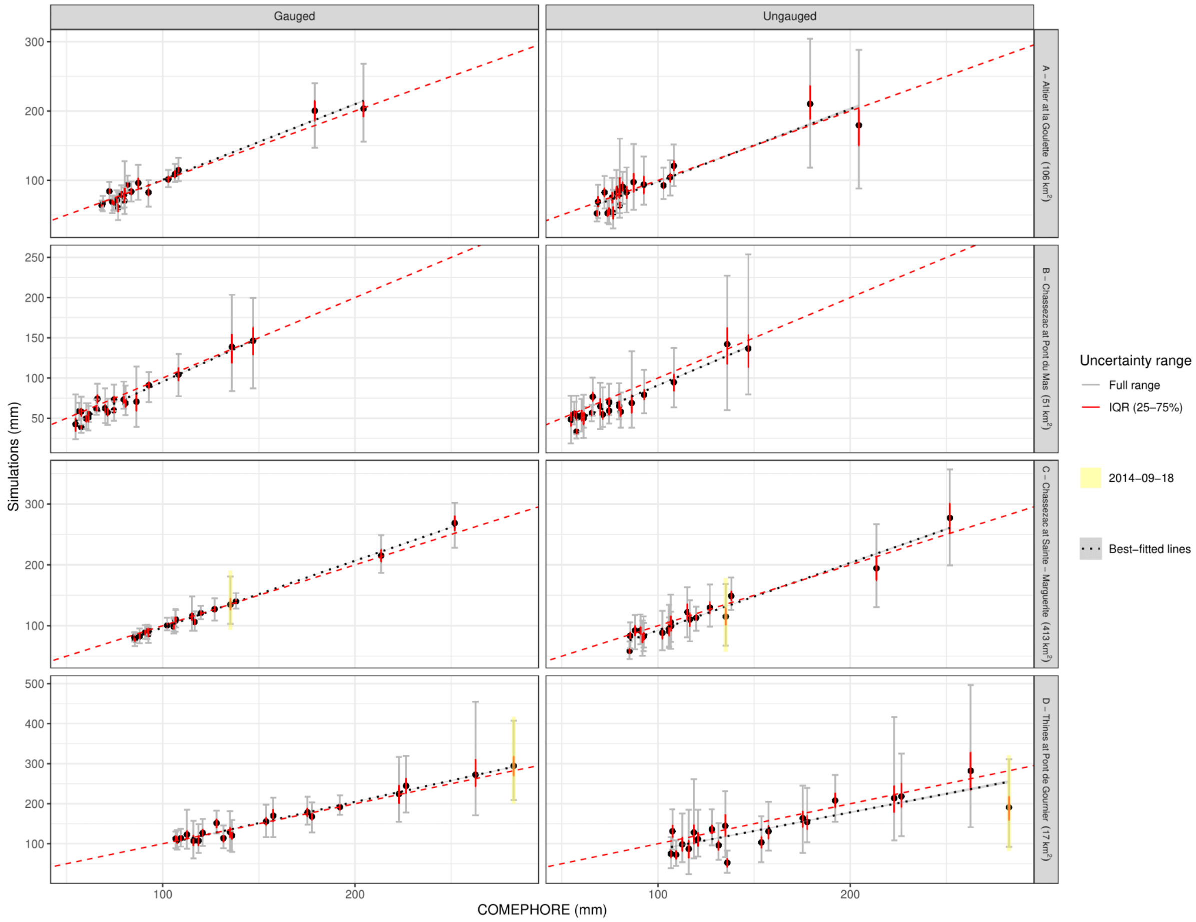
We agree with the reviewer that this information was missing. We add L251 " For this reason, we later quantify precipitation uncertainty using conditional simulations from anisotropic covariance, derived from AROME simulations."

#RC1.10 Figure 5: Please highlight the data points from the event (2014-09-18) in the scatter plots?

#RC1.11 Figure 5: Please add to the boxes with the catchments name, the code ID and areas of the catchments as described in Table 1.

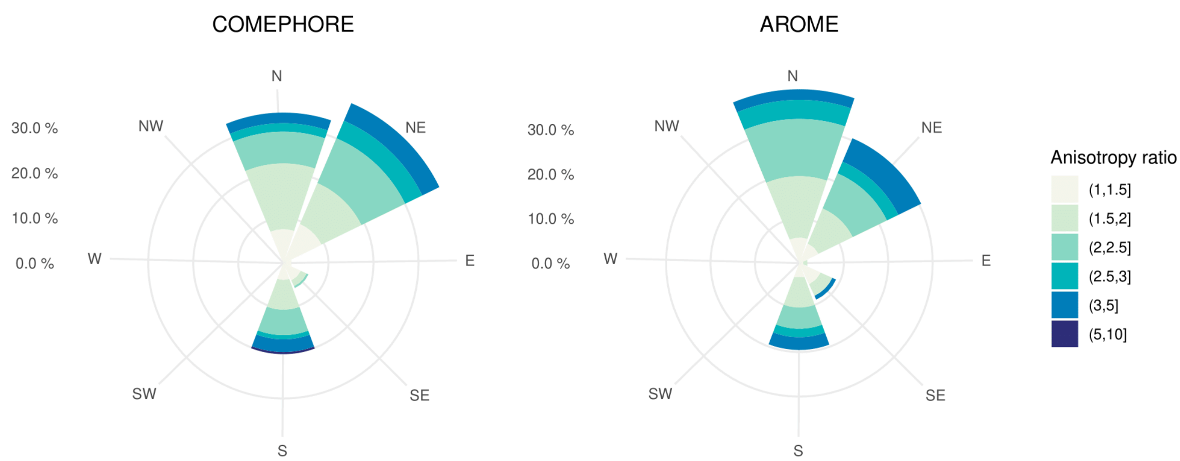
#RC1.12 Figure 5: Consider adding best-fit lines fitted across the ensemble mean points on each scatter plot as they could help to illustrate biases in the simulations for each catchment.

Thank you for these three suggestions that will improve the readiness of this figure.



#RC1.13 Table 3: it would be valuable to add the same stats for COMEPHORE in this table. This allows a direct comparison of the anisotropy parameters derived from AROME and from COMEPHORE.

Thank you very much for this suggestion. We propose to replace Table 3 by this new figure for a clearer comparison.

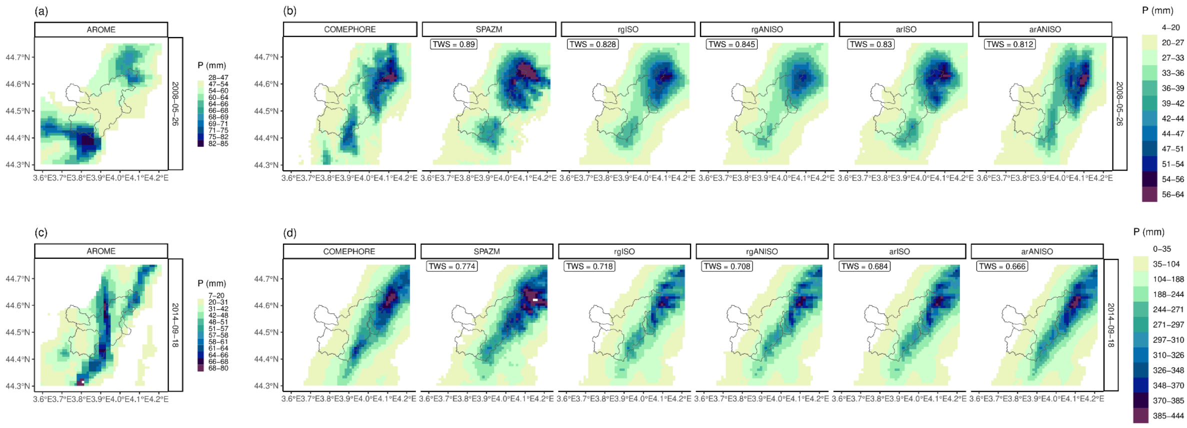


We replace L285-287 by : “Figure 8 summarizes the anisotropy parameters η (anisotropy ratio) and θ (anisotropy angle) estimated from AROME and COMEPHORE.

Both datasets show similar preferred anisotropy directions, predominantly oriented south–north (S–N) and southwest–northeast (SW–NE), with AROME exhibiting a slightly stronger S–N component. In both cases, the anisotropy is generally more pronounced along the SW–NE direction.”

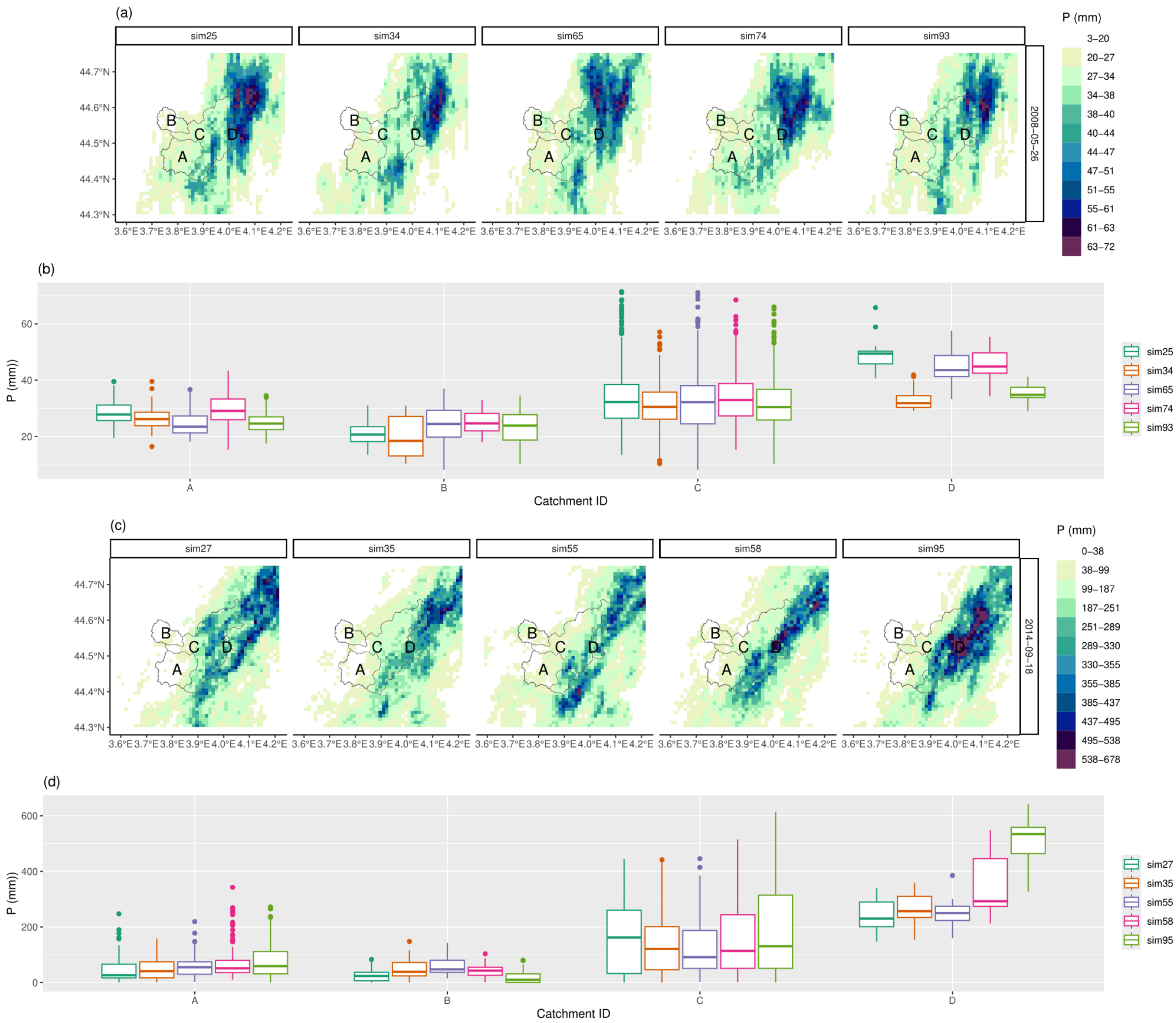
#RC1.14 Figure 6: please add AROME rainfall fields to this figure as the variogram for arISO and arANISO were derived from AROME. Please consider adding the TWS values for each field.

We thank the reviewer for this suggestion. TWS will quantify the visual differences that can be seen.



#RC1.15 Figure 7 and discussion. Given that the relatively small size of the catchments with the whole domain, would be valuable to present the distribution of the rainfall values of few (all?) members for each catchment as complement to the precipitation fields?

Thank you for this graphical suggestion. For better visualisation, we only provide the distribution of the rainfall values for the selected simulation numbers. It helps quantify the precipitation uncertainty.



We replace L280-281: “Simulation 49 gives rise to nearly 250 mm mean catchment precipitation, compared to more than 350 mm for simulation 15” by “Simulation 55 gives rise to nearly 200 mm mean catchment precipitation, compared to more than 500 mm for simulation 95”.

#RC1.16 L312: please elaborate when rgANISO improves rgISO and when does not and why?

Thank you for this suggestion. We add L314: “More specifically, for both point cross-validation and spatial variability evaluation, rgANISO outperforms rgISO during strongly anisotropic events, whereas the opposite behavior is observed for weakly anisotropic events. This indicates that anisotropy can be significant but remains difficult to robustly estimate using a rain-gauge network alone, even when the network is dense.”

#RC1.17 L334: Please elaborate what it could be needed to extend this methodology to real-time and sub-daily applications as this study only has assessed daily time scales.

“The methodology could also be extended to real-time, sub-daily interpolation. CP-RCM simulations are not continuously updated, so their replacement by hourly numerical weather forecasts should be investigated. At the daily timescale, timesteps are typically considered independent, but this assumption no longer holds at the hourly scale. To address this, temporal dependence should be incorporated into the model, as done in Sideris et al. (2014) and Frey and Frei (2025). A major limitation is the limited availability of sub-daily rain gauges. One potential solution to bypass this shortcoming is to disaggregate daily interpolated fields using radar data or numerical weather forecasts.”