Response to the Referee's Comments

June 9, 2026

Vihotogbé Houssou & Julie Carreau

Response to Referee 2

We would like to thank Referee 2 for the careful reading of the manuscript and for the constructive and insightful comments. We appreciate the positive assessment of the relevance and clarity of the work. The suggestions provided have helped us improve the manuscript, both in terms of methodological clarification and presentation.

Below, we address each comment in detail. Referee comments are reproduced in a sans-serif font, and our responses are provided below each comment. Changes made in the manuscript are indicated where relevant.

Major comments

Comment 1 :

Fair comparison between SPR and benchmark interpolation methods : The tested approaches are not evaluated under equivalent settings. Specifically, IDW, OK, and KED are directly applied to station observations Z_d, whereas SPR is applied to anomalies relative to the RCM background field: $Z'_d = Z_d - Z_d^{\text{grid}}$. This makes the comparison difficult to interpret. It is possible that IDW, OK, or KED applied to the same anomaly fields could perform similarly or even outperform SPR. At present, the conclusions regarding the superiority of SPR remain insufficiently supported. I therefore strongly recommend evaluating all interpolation methods within the same anomaly-based framework.

Moreover, the authors should investigate the sensitivity of SPR to the definition of anomalies:

- additive anomalies: $Z'_d = Z_d - Z_d^{\text{grid}}$
- multiplicative anomalies: $Z'_d = \frac{Z_d}{Z_d^{\text{grid}}}$

The additive formulation is generally more appropriate for temperature, while multiplicative or relative anomalies are often preferable for precipitation due to heteroscedasticity.

Response

- We thank the Referee for this important comment. We clarify that SPR is applied to raw observations, but expressed in a formulation that is consistent with the PCA decomposition used to derive spatial patterns. The subtraction of the RCM mean field in Eq. (2) is not intended as an anomaly-based formulation in the climatological sense, but rather as a centering step required to ensure consistency with the PCA representation, in which spatial patterns describe variability around a mean field.
- We clarify what $\bar{Z}^{\text{grid}}$ represents in Eq. (2). It is the temporal column mean of the RCM data matrix $\mathbf{Z}_{n,p}^{\text{grid}}$ over the auxiliary period — that is, for each grid cell j, $\bar{Z}_j^{\text{grid}} = \frac{1}{n}\sum_{t=1}^{n} Z_{t,j}^{\text{grid}}$, where $n = 10{,}950$ is the number of daily time steps. This is a single constant per grid cell, not a monthly climatology, not a seasonal cycle, and not a daily evolving background. It is the standard centering constant used in SVD/PCA decomposition.

- In particular, SPR can be equivalently written as a linear model of the form:

$$Z = \bar{Z}^{\text{grid}} + V^{\text{grid}}\beta + \epsilon,$$

where the RCM-derived mean field acts as a spatially varying **intercept**, and the PCA-derived spatial patterns represent low-dimensional spatial variability. The centered formulation used in Eq. (2) is therefore a direct consequence of this decomposition, rather than a specific modelling choice.

- The uncentered form of SPR (Eq. (3), in the revised manuscript) makes explicit that all methods — SPR, IDW, OK, and KED — are evaluated on the same quantity: the full meteorological field on the original scale of the variable. The comparison is therefore consistent across methods. The difference between SPR and the baselines lies not in the scale of the data they operate on, but in the structural information they incorporate: SPR uses RCM-derived spatial patterns as a low-rank basis, while IDW and OK rely solely on spatial proximity, and KED incorporates RCM climatologies as an external drift.

- Regarding the suggestion of exploring additive versus multiplicative anomaly formulations, we note that SPR does not rely on a predefined anomaly definition but rather on the consistency between the PCA decomposition and the regression formulation. Investigating alternative formulations may represent an interesting extension, particularly for precipitation, and is left for future work.

The following equation and clarification have been added to Section 3.2.1 (lines 162–165):

> *More precisely, $\bar{Z}_p^{grid}$ is the temporal column mean of the RCM data matrix over the auxiliary period, defined for each grid cell j as $\bar{Z}_j^{grid} = \frac{1}{n}\sum_{t=1}^{n} Z_{t,j}^{grid}$. It is a single constant per grid cell, computed once over the auxiliary period, and represents the standard centering term of the SVD decomposition. It is distinct from monthly climatological means or seasonally varying backgrounds.*

The following equation and clarification have been added to Section 3.2.2 (lines 192–198):

> *The formulation in Eq. (2) can be equivalently rewritten in its uncentered form as:*
>
> $$Z_d = \bar{Z}_d^{\text{grid}} + \mathbf{V}_{d,k}^{\text{grid}}\,\beta_k + \varepsilon_d, \tag{3}$$
>
> *which highlights that SPR can be interpreted as a linear model with a spatially varying intercept term given by the RCM-derived mean field, and a low-rank representation of spatial variability defined by the PCA-derived patterns. The centered formulation used in Eq. (2) is therefore a direct consequence of the PCA decomposition, in which spatial patterns represent variability around the mean field. In particular, this uncentered form makes clear that SPR is applied to raw observations Z_d, and that its output is on the same scale as the observations used by IDW, OK, and KED, ensuring a consistent basis for comparison.*

The following text has been added to Section 5 – Discussion and Conclusions (lines 387–394):

> *Moreover, while all methods are evaluated on the same meteorological scale, they differ fundamentally in the structural information they incorporate. SPR explicitly separates a spatially varying mean structure from low-dimensional variability through the PCA-based decomposition, effectively embedding RCM-derived spatial organization into the interpolation framework. In contrast, IDW and OK rely solely on spatial proximity without incorporating any auxiliary structural information. KED occupies an intermediate position, using RCM-derived monthly climatologies as an external drift term to account for large-scale spatial variability. These differences in modeling philosophy — rather than differences in the scale or nature of the data — explain the contrasting behaviors observed across experimental configurations, particularly under sparse network conditions where the structural constraints provided by SPR become most valuable.*

Comment 2 :

Robustness to station sampling and validation strategy: For both the synthetic and real-data experiments, the results may be highly sensitive to the specific sampling of training stations, particularly under low station-density scenarios.

I recommend using a K-fold cross-validation framework (or repeated random subsampling) in which several independent training/validation splits are generated. Performance metrics (M) should then be averaged across folds K:

$$M = \frac{1}{K}\sum_{k=1}^{K} M_k$$

This would provide more robust and statistically reliable conclusions.

Response

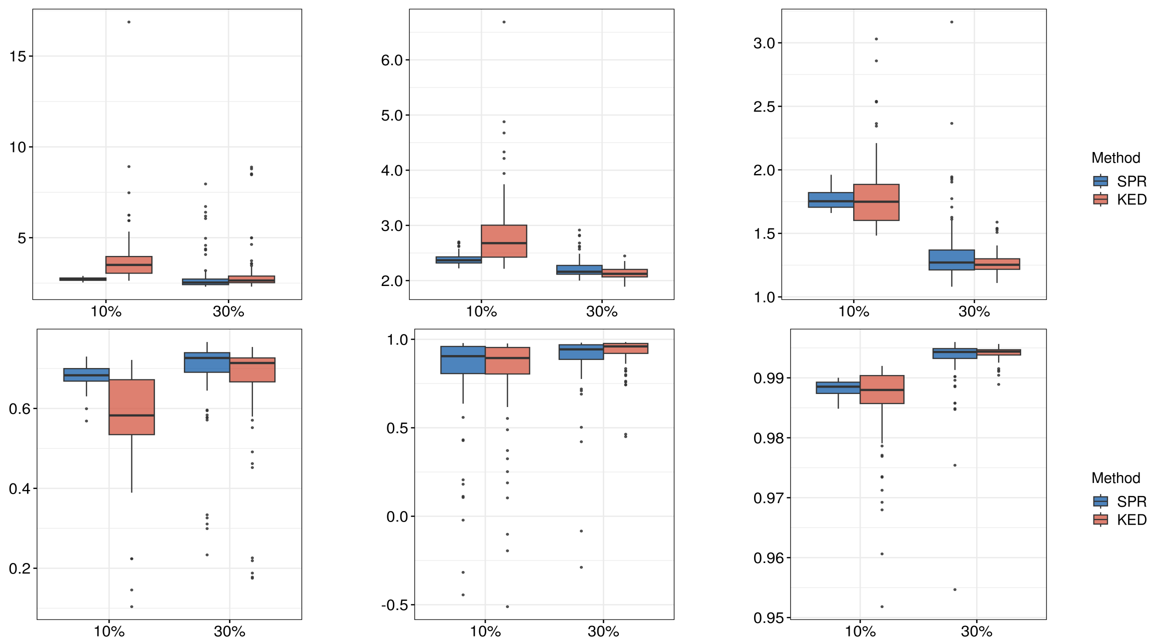
We thank the Referee for this important methodological comment. We address the synthetic and real-data experiments separately, as the robustness considerations differ substantially between the two settings.

- **Synthetic experiments.** In the synthetic setting, the ground truth is the full RCM field, available at all grid cells. The test set therefore consists of all non-training grid cells, ranging from 297 to 2,723 cells depending on the experimental configuration (see Table 1). With test sets of this size, performance metrics reflect method behavior across the full spatial domain and are not overly sensitive to the particular choice of virtual station locations. Furthermore, the factorial experimental design — combining five network density levels (10%–90%), three region sizes, and two geographic locations across three meteorological variables — provides a broad and systematic assessment of performance robustness across a wide range of conditions, which constitutes an implicit cross-validation across experimental configurations.
- **Real data experiments.** We acknowledge that the real-data experiments are more sensitive to the specific station sampling, particularly at 10% density where only 2–4 stations are available for training. To address this concern, we have repeated the real-data station selection 100 times using independent random seeds. Performance metrics (RMSE and SSIM) are now presented as boxplots (Fig. 7) instead of a table and confirm that the conclusions drawn from the original single-split analysis are robust to the choice of training stations: the ranking of methods and the relative advantage of SPR under sparse network conditions are consistent across all repetitions.

The following text has been added to Section 3.4.2 – real data section (lines 258–262):

> *To assess the sensitivity of results to the specific choice of training stations, the station selection is repeated 100 times using independent random seeds for each density level and each variable. Performance metrics are averaged across repetitions, and their standard deviation is reported to quantify sampling uncertainty. This repeated subsampling strategy ensures that the conclusions are not driven by a particular station configuration and provides a statistically robust basis for comparison across methods.*

The results shown in Table 3 are replaced by the Fig. 7 in section 4.3 (after line 354) – Results of Real data experiments: *See Fig. 7 in the revised manuscript.*



(a) Precipitation (b) Minimum temperature (c) Maximum temperature

Figure 7. Validation with real observations: distribution of mean daily RMSE (first row) and SSIM (second row) across 100 independent random station samplings for SPR and KED, for each variable. Each box shows the median, interquartile range, and whiskers extending to 1.5 times the interquartile range. Lower RMSE and higher SSIM indicate better performance. Note that 10% density corresponds to 2–4 stations in a region of 70,000 km^2.

Comment 3 :

Ground truth used for SSIM and spatial field evaluation : The manuscript should clarify what is used as the "ground truth" for the SSIM calculations and spatial field comparisons. If the reference field is the RCM itself, this may introduce an important methodological bias. It is well known that RCMs tend to smooth spatial variability, particularly for precipitation extremes. Consequently, methods reproducing smoother spatial fields (such as SPR) may be artificially favored over approaches preserving higher local variability (e.g., KED). The discussion section should explicitly acknowledge that conventional RCMs at approximately 11 km resolution strongly smooth the spatial variability of meteorological fields, especially precipitation, and that the use of convection-permitting regional climate models (CP-RCMs), typically at 1–4 km resolution, represents a promising approach for improving the realism of spatial structures used within SPR.

Response:

We thank the Referee for this important point. We address the two aspects of this comment separately.

- **Clarification of ground truth.** In the synthetic experiments, the ground truth used for both RMSE and SSIM evaluation is the full RCM field at all non-training grid cells. This is explicitly stated in the experimental design (Section 3.4.1) and is a deliberate choice: synthetic experiments are designed to provide controlled evaluation under a known reference, allowing systematic comparison across network densities, region sizes, and geographic locations. The objective of these experiments is not to assess realism with respect to observations, but to evaluate how accurately each method reconstructs the reference RCM field from a sparse subset of its grid cells under controlled conditions. A clarifying sentence has been added to Section 3.4.1 to make this explicit. In the real-data experiments, evaluation is performed against independent station observations, which provides a complementary and fully observation-based assessment.
- **RCM smoothing and potential bias.** We acknowledge the Referee's point regarding the smoothing properties of the ClimEx RCM at approximately 11 km resolution. SPR reconstructs fields as low-rank linear combinations of RCM-derived spatial patterns, which is by construction a low-pass spatial filter: truncating the PCA basis to k leading patterns retains only the dominant large-scale spatial structures while discarding finer-scale variability not captured by the leading EOFs. If the RCM ground truth is itself smooth, this truncation will result in smaller residuals for SPR. Evaluating against the RCM field itself may therefore introduce a bias in favor of SPR. However, we note that KED, through its variogram calibration, is also capable of reproducing spatially smooth fields, and the extent to which this bias specifically affects the comparison between SPR and KED may be limited. This constitutes a limitation of the synthetic evaluation framework that we have added an explicit acknowledgment of in the Discussion. As suggested by the Referee, the use of CP-RCMs at 1–4 km resolution represents a promising direction for future work.

The following sentence has been added to Section 3.4.1 (lines 241–243):

> *The reference RCM field serves as ground truth in the synthetic experiments by design: the objective is to evaluate how well each method can reconstruct a known spatially complete field from sparse observations, rather than to assess realism with respect to actual meteorological observations.*

The following text has been added to Section 5 (lines 415–424):

> *Furthermore, SPR's output may partly inherit the spatial smoothness of the RCM, since it reconstructs fields as low-rank combinations of RCM-derived spatial patterns — a consequence of the dimensionality reduction inherent in the PCA-based representation. Evaluating against the same RCM field may therefore introduce a bias in favor of SPR over methods that preserve finer local variability. However, KED, through its variogram calibration, is also capable of reproducing spatially smooth fields, and the extent to which this bias specifically affects the SPR–KED comparison may be limited. This limitation is specific to the synthetic experiments; in the real-data experiments, evaluation is performed against independent station observations, which provides a complementary assessment. More broadly, conventional RCMs at approximately 11 km resolution are known to smooth spatial variability, particularly for precipitation extremes. The use of*

convection-permitting regional climate models (CP-RCMs), typically operating at 1–4 km resolution, represents a promising direction for improving the realism of spatial structures used within SPR (Caillaud et al., 2021; Dura et al., 2025).

Minor comments

Comment 4:

No comment 4 was provided by the Referee.

Response:

The Referee's comments were numbered 1, 2, 3, and 5 through 17. No Comment 4 was included in the review. The numbering in this response letter follows the Referee's original numbering for ease of reference.

Comment 5 :

The manuscript separates "spatial heterogeneity" and "extremes", although these concepts are closely related. In this context, "variability" appears more appropriate than "heterogeneity". I suggest replacing : "spatial heterogeneity and extremes at fine spatial and temporal resolution" with "the fine spatio-temporal variability of meteorological fields, especially during extreme events".

Response:

We thank the Referee for this suggestion. We agree that the term "variability" more appropriately captures both spatial heterogeneity and extremes in this context. The sentence has been updated accordingly in the revised manuscript.

This sentence in the Introduction (lines 15–17) : *Hydrological models used for flood forecasting, water resource management, or climate impact assessments require meteorological inputs that accurately represent both spatial heterogeneity and extremes at fine spatial and temporal resolution.* has been updated to the following sentence :

> *Hydrological models used for flood forecasting, water resource management, or climate impact assessments require meteorological inputs that accurately represent the fine spatio-temporal variability of meteorological fields, especially during extreme events.*

Comment 6 :

Meteorological radar products provide gridded precipitation estimates but do not fully belong to the three categories listed by the authors. Radar datasets can be used directly, without relying on station observations or reanalysis frameworks. This category should therefore be acknowledged separately.

Response:

We thank the Referee for this observation. The text has been revised to acknowledge radar-based products as a distinct fourth category of gridded meteorological data. The corresponding reference has been added to the bibliography.

The following change has been made in the Introduction (lines 22–25):

> *In practice, gridded meteorological data used in hydrological studies are typically obtained through four main approaches: spatial interpolation of station observations (Cornes et al., 2018), physics-based numerical models (including numerical weather prediction and regional climate models) (Muñoz-Sabater et al., 2021), reanalysis products (Leduc et al., 2019; Gasset et al., 2021), and radar-based products (Vernay et al., 2025).*

Comment 7:

Lines 23–24: Several widely used gridded meteorological products should be cited when discussing interpolation, reanalysis, and climate-model-based datasets.

- Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J., and Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17), 9391–9409.
- Munoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., ... and Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383.
- Caillaud, C., Somot, S., Alias, A., Bernard-Bouissières, I., Fumière, Q., Laurantin, O., ... and Ducrocq, V. (2021). Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model. *Climate Dynamics*, 56(5), 1717–1752.

Response:

We thank the Referee for these suggestions. A reference to the E-OBS dataset (Cornes et al., 2018) and to the ERA5-Land (Muñoz-Sabater et al., 2021) have been added in the Introduction where widely used gridded meteorological products are discussed. The Caillaud et al. (2021) reference has also been added in the context of convection-permitting RCMs, as discussed in the response to Comment 3.

References to (Cornes et al., 2018) and (Muñoz-Sabater et al., 2021) have been added in the Introduction (lines 22–25) :

> *In practice, gridded meteorological data used in hydrological studies are typically obtained through four main approaches: spatial interpolation of station observations (Cornes et al., 2018), physics-based numerical models (including numerical weather prediction and regional climate models) (Muñoz-Sabater et al., 2021), reanalysis products (Leduc et al., 2019; Gasset et al., 2021), and radar-based products (Vernay et al., 2025).*

A reference to Caillaud et al. (2021) has been added to Section 5 (lines 422–424):

> *The use of convection-permitting regional climate models (CP-RCMs), typically operating at 1–4 km resolution, represents a promising direction for improving the realism of spatial structures used within SPR (Caillaud et al., 2021; Dura et al., 2025).*

The following change has also been added to Section 5 (lines 428–432):

> *Recent work has explored alternative ways of exploiting high-resolution model simulations to improve interpolation, including the use of anisotropic variograms derived from CP-RCM simulations (Dura et al., 2025), radar-based ensemble analyses (Vernay et al., 2025), and high-resolution reanalysis products (Khedhaouiria et al., 2026), which represent complementary directions for improving the spatial consistency of interpolated meteorological fields.*

Comment 8:

Lines 36–37: Climatological fields are also frequently derived directly from station observations. The article should cite datasets/methods :

- Daly, C., Taylor, G. H., and Gibson, W. P. (1997, October). The PRISM approach to mapping precipitation and temperature. In Proc., 10th AMS conf. on applied climatology (Vol. 675).

Response:

We thank the Referee for this suggestion. PRISM (Daly et al., 1997) is an interpolation method that incorporates topographic information as auxiliary covariate to produce gridded climatological fields directly from station observations. This fits naturally in the context where auxiliary gridded information such as topography and climatology is discussed. The reference has been added accordingly.

A reference to the PRISM approach is made in the introduction where topography is incorporated to spatial interpolation methods (lines 35–37):

> *To mitigate these limitations, auxiliary gridded information such as topography or climatology is commonly incorporated (Livneh et al., 2015; Werner et al., 2019; Daly et al., 1997).*

Comment 9:

Lines 38–39: If I understand correctly, the spatial structures extracted from the RCM are climatological patterns rather than daily evolving fields. This point should be clarified explicitly. Regarding Equation (1), does $Z^{\mathsf{grid}}_{n,p}$ represent all daily RCM fields? If so, the indexing notation should probably be revised to better reflect the temporal dimension.

Response:

We thank the Referee for this clarification request. The manuscript already states that $\mathbf{Z}^{\mathrm{grid}}_{n,p}$ is the RCM dataset over the auxiliary period where n is the number of daily time steps, that the PCA is computed using all daily fields over the auxiliary period, and that the spatial patterns $\mathbf{V}^{\mathrm{grid}}_{p,k}$ form a fixed basis that does not vary from day to day. To further address the Referee's concern, an explicit sentence has been added to Section 3.2.1 stating that the patterns are extracted once from the auxiliary period and do not evolve in time. The daily evolution of the reconstructed field is captured entirely through the regression coefficients $\hat{\beta}_k$, which are estimated independently for each day of the interpolation period, as described in Section 3.2.2.

The following sentence has been added to Section 3.2.1 (line 167):

> *They are extracted once from the auxiliary period and do not vary from day to day.*

Comment 10:

Lines 40–41: SPR can only be classified as a hybrid between spatial interpolation and reanalysis if the RCM predictors vary dynamically over time. If SPR only relies on climatological information extracted from the RCM, then the method is conceptually closer to a spatial interpolation framework constrained by climatological covariates. Please clarify this distinction.

Response:

We thank the Referee for this important conceptual clarification. We wish to clarify two points.

- First, while the spatial patterns $\mathbf{V}^{\mathrm{grid}}_{p,k}$ are indeed temporally fixed quantities extracted once from the RCM over the auxiliary period, the combination of these patterns varies dynamically from day to day through the regression coefficients $\hat{\beta}_k$, which are estimated independently for each day of the interpolation period. The reconstructed field therefore does evolve in time, and SPR does incorporate a dynamic component — not at the level of the patterns themselves, but through their daily amplitudes constrained by observations.
- Second, we wish to emphasize that the spatial patterns used in SPR are not climatological in the conventional sense. They are EOFs extracted from the full daily variability of the RCM over the auxiliary period, capturing recurrent modes of spatial organization that go well beyond a simple long-term mean. As discussed in Section 3.1, these patterns encode the dominant spatial structures of the meteorological field across a broad range of weather situations, making them substantially richer than climatological covariates such as monthly means or seasonal cycles.

In light of these clarifications, we have revised the characterization of SPR in the Introduction to better reflect its nature as a physically-constrained spatial reconstruction framework that bridges the gap between traditional statistical interpolation and model-based reanalyses.

The following change has been made in the Introduction (lines 41–46):

> *Specifically, we introduce Spatial Pattern Regression (SPR), a physically-constrained spatial reconstruction framework designed to bridge the gap between traditional statistical interpolation and model-based reanalyses. Unlike reanalysis systems, which assimilate observations into dynamically evolving model states, SPR relies on a fixed set of spatial patterns extracted from the RCM, whose daily amplitudes are estimated from sparse station observations. Unlike purely local interpolation methods, SPR leverages physically consistent spatial structures derived from RCM simulations to constrain the reconstruction.*

Comment 11:

Figure 1: Please add the station locations to Figure 1(b). This would greatly aid the interpretation of the geographical context.

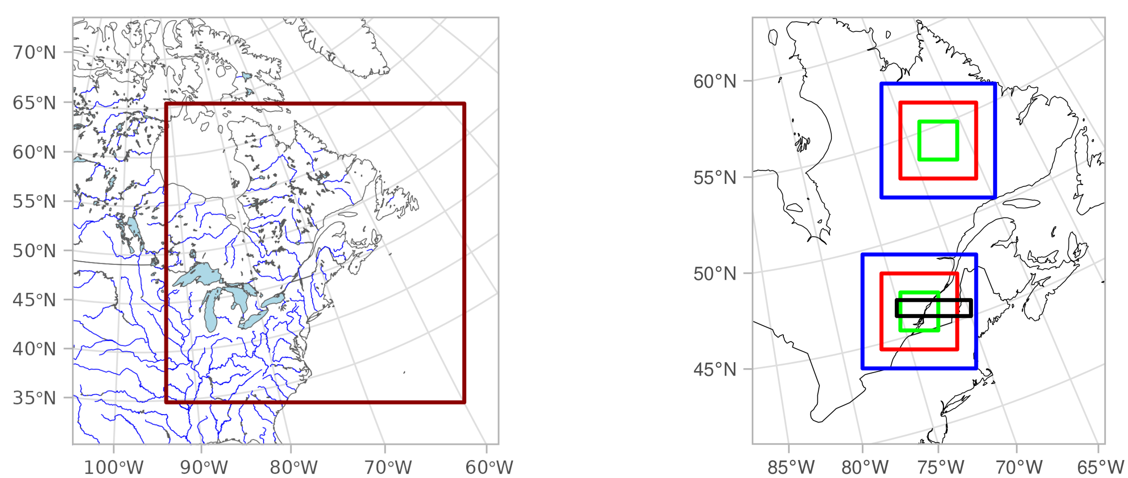
Response:

We thank the Referee for this suggestion. Rather than plotting individual station locations, which vary across variables and validation periods, we have added the spatial domain of the real-data experiments to Figure 1(b) as a dashed black rectangle, corresponding to the region defined by latitudes 45.0–47.0°N and longitudes 75.0–70.0°W (see Section 2.2). This provides clear geographical context for the real-data experiments while remaining consistent across variables. The figure caption has been updated accordingly.

The following change has been made to the caption of Figure 1(b):

> *Two study regions — referred to as the south and north regions — within the domain outlined in dark red in Fig. 1(a), shown at three different sizes (in different colours). The black rectangle indicates the spatial domain of the real-data experiments.*

Here is the update Figure 1(b) in Section 2.1 (after line 99):



(a) North American domain of the ClimEx project (dark red rectangle).

(b) Two study regions—referred to as the south and north regions— within the domain outlined in dark red in Fig. 2a, shown at three different sizes (in different colours). The black rectangle indicates the spatial domain of the real-data experiments.

Figure 1. Spatial domains used in this study.

Comment 12:

Lines 74–75: The statement that the method does not require overlap is true for precipitation but more questionable for temperature because of long-term warming trends. For temperature, inconsistencies between climatological periods may introduce systematic biases. This issue should be discussed, together with potential solutions such as detrending.

Response:

We thank the Referee for raising this important point. We agree that for temperature variables, long-term warming trends may introduce systematic biases when the auxiliary and interpolation periods are substantially separated in time, as the spatial mean field extracted from the auxiliary period may not accurately represent the mean level of the interpolation period. A clarifying sentence has been added to Section 2.1 acknowledging this limitation and suggesting detrending or temporal alignment of periods as potential mitigation strategies.

The following sentence has been added to Section 2.1 (lines 92 – 96):

> *It should be noted, however, that for temperature variables, long-term warming trends may introduce systematic biases when the auxiliary and interpolation periods are substantially separated in time. In such cases, detrending of the RCM fields prior to pattern extraction, or the selection of an auxiliary period temporally close to the interpolation period, would be advisable.*

Comment 13:

Lines 105–110: The manuscript states that SPR exploits the covariance structure. However, the method actually exploits much richer information: spatial patterns extracted from the RCM.

Using covariance structure alone would correspond to statements such as:

$$\mathrm{Corr}(Z(x_i), Z(x_j)) \to 0 \quad \text{when} \quad \|x_i - x_j\| \to \infty$$

SPR instead leverages spatial patterns derived from the RCM. This is potentially more informative, but also riskier because any regional structural bias in the RCM may directly influence the interpolation. This remark should be explicitly discussed. Some authors use RCM, CP-RCM simulations, NWP in another way, it should be mentioned in the discussion:

- Dura, V., Evin, G., Favre, A. C., and Penot, D. (2025). Improving Precipitation Interpolation Using Anisotropic Variograms Derived from Convection-Permitting Regional Climate Model Simulations. EGUsphere, 2025, 1–25.
- Khedhaouiria, D., Gasset, N., Fortin, V., Dimitrijevic, M., Bulat, M., and Wang, X. (2026). The Canadian Surface Reanalysis (CaSR) v3.2 precipitation dataset: A 45-year high-resolution analysis for North America (1980–2024). EGUsphere, 2026, 1–40.
- Vernay, M., Lafaysse, M., and Augros, C. (2025). Radar-based high-resolution ensemble precipitation analyses over the French Alps. Atmospheric Measurement Techniques, 18(8), 1731–1755.

Response:

We thank the Referee for this important distinction. We clarify that while SPR is mathematically grounded in the PCA of the RCM spatiotemporal covariance matrix, the resulting interpolation does not exploit raw covariances or distance-based correlations directly. Rather, it leverages the eigenvectors of that covariance matrix — the spatial patterns — which encode higher-order, physically consistent spatial structures that go well beyond simple pairwise correlations. It is therefore correct that SPR exploits richer information than a covariance-based representation alone.

A paragraph has been added to Section 5 acknowledging both the potential benefit of this richer representation and the associated risk of RCM structural biases propagating into interpolated fields. The three references suggested by the Referee have been added to the bibliography and cited in this new paragraph.

The following text has been added to the Introduction (lines 38–39):

More advanced approaches seek to exploit higher-order spatial structures beyond mean climatology, capturing coherent patterns of variability across space (Taylor et al., 2013; Carreau and Guinot, 2021).

The following text has been added to Section 5 (lines 425–432):

Notably, SPR leverages spatial patterns derived from RCM simulations, which encode potentially richer structural information than simple variogram-based representations of spatial dependence. However, this also introduces a dependency on the fidelity of the RCM: any systematic spatial bias may directly propagate into the interpolated fields. This sensitivity to RCM structural biases should be considered in operational applications. Recent work has explored alternative ways of exploiting high-resolution model simulations to improve interpolation, including the use of anisotropic variograms derived from CP-RCM simulations (Dura et al., 2025), radar-based ensemble analyses (Vernay et al., 2025), and high-resolution reanalysis products (Khedhaouiria et al., 2026), which represent complementary directions for improving the spatial consistency of interpolated meteorological fields.

Comment 14:

Lines 122–123: The manuscript states that SPR does not make distributional assumptions. This is not entirely correct. The regression residuals implicitly assume $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ with IID residuals. Therefore, SPR still relies on classical regression assumptions.

Response:

We thank the Referee for this correction. The original statement was imprecise. The intent was to highlight that SPR does not require constraining assumptions typically needed by geostatistical methods, such as isotropy or stationarity of the spatial field. It was not intended to claim that SPR makes no assumptions whatsoever. The manuscript has been revised to clarify this distinction. First, the opening sentence of Section 3.2 has been updated to make explicit that the assumptions being disclaimed are isotropy and stationarity specifically. Second, the assumption on the regression residuals has been added immediately after Eq. (2), where ε_d is first introduced, which is the natural location for this clarification.

The following sentence has been updated in Section 3.2 (lines 150–151):

It makes no isotropy or stationarity assumptions and relies on the representativeness of the dominant spatial patterns.

The following sentence has been added in Section 3.2.2 after Eq. (2) (line 182):

ε_d is a vector of residual errors assumed to be independent and identically distributed with zero mean

Comment 15:

Line 209: The sensitivity of the transformation step should be evaluated. In precipitation interpolation, square-root transformations are frequently used at the daily timescale to stabilize variance and reduce skewness.

- Erdin, R., Frei, C., and Künsch, H. R. (2012). Data transformation and uncertainty in geostatistical combination of radar and rain gauges. Journal of Hydrometeorology, 13(4), 1332-1346.

Response:

We appreciate the Referee's suggestion regarding the evaluation of the transformation step. To clarify, the SPR framework does not utilize a logarithm transformation, but rather applies the Softplus function ($f(x) = \ln(1+e^x)$). While square-root transformations are a well-established geostatistical tool for stabilizing variance and mitigating skewness, the Softplus function serves a slightly different, structural purpose in our framework. It acts as a smooth, differentiable link function that ensures the reconstructed precipitation field strictly respects the physical boundary of non-negativity (≥ 0). Because Softplus is integrated as a fundamental structural component of the regression framework to enforce this physical constraint—rather than an adjustable empirical parameter—a traditional sensitivity analysis is not directly applicable. However, we recognize that this choice was not sufficiently clear in the original text, and we have updated Section 3.5 to explicitly state the mathematical formulation of the Softplus transformation and its role in handling the zero-precipitation boundary.

The following clarification has been added to Section 3.5 (lines 270–279):

> *It should be noted that applying a direct inverse transformation after interpolation in transformed space may in general neglect the contribution of analysis-error variance (Fletcher and Zupanski, 2006). However, the inverse of the softplus transformation satisfies $f^{-1}(x) = \log(1 + e^x) \approx x$ for precipitation values of practical interest, implying that the back-transformation introduces negligible bias in this context. This transformation serves as a smooth link function ensuring that back-transformed interpolated values remain non-negative. Unlike variance-stabilizing transformations such as the square-root (Schiemann et al., 2010; Erdin et al., 2012), the softplus transformation does not primarily aim to reduce skewness but rather to enforce the physical boundary of non-negativity. Alternative transformations, such as the square-root transformation used by Schiemann et al. (2010), are also common for precipitation. Correction methods based on Taylor-series approximations (Fortin et al., 2015; van Hyfte et al., 2023) or quantile-based approaches (Erdin et al., 2012) are not considered here given the negligible bias of the softplus back-transformation, but represent a potential avenue for future improvements.*

Comment 16:

Line 237: I believe the Kling–Gupta Efficiency (KGE) would be more informative than RMSE alone. KGE explicitly evaluates :

- correlation,
- bias,
- variability ratio, through

$$\mathsf{KGE} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$$

  where :

- r is the correlation,
- β the mean bias ratio,
- γ the variability ratio.

In the present context, the variability component is particularly relevant because it directly evaluates biases in spatial variability.

Response:

We thank the Referee for this valuable suggestion. In the present study, interpolation performance is evaluated using two complementary metrics: RMSE, which quantifies pointwise accuracy and captures the average magnitude of errors and bias, and SSIM, which assesses the ability of each method to reproduce the spatial structure of the reference field. Together, these two metrics cover the three aspects explicitly decomposed by KGE. SSIM is defined as the product of three explicit components:

$$\mathrm{SSIM}(x,y) = \underbrace{\frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}}_{\text{luminance}} \times \underbrace{\frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}}_{\text{contrast}} \times \underbrace{\frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}}_{\text{structure}},$$

where μ_x and μ_y are local means, σ_x and σ_y are local standard deviations, σ_{xy} is the local covariance, and c_1, c_2, c_3 are stabilization constants. The contrast term compares local standard deviations and is analogous to the variability ratio γ in KGE. The structure term measures local correlation and is analogous to the correlation component r in KGE. Meanwhile, RMSE captures the average magnitude and bias of errors, analogous to the bias ratio β in KGE. The joint use of RMSE and SSIM therefore already ensures that biases in spatial variability are penalized, achieving the same evaluative goal as KGE. We acknowledge that KGE provides a complementary global diagnostic and could offer additional insights, and note its use as a direction for future work.

The following sentence has been added to Section 3.5.2 (lines 309–312) to clarify the complementarity of the two metrics:

> *While RMSE captures the average magnitude and bias of interpolation errors, SSIM explicitly accounts for spatial variance and structural correlation through its contrast and structure components respectively. Together, these two metrics provide a comprehensive evaluation covering the three aspects — bias, variability, and correlation — that are explicitly decomposed by metrics such as the Kling–Gupta Efficiency (Gupta et al., 2009).*

Comment 17:

Table 3: I recommend replacing Table 3 with boxplots (or violin plots). This would make the distribution of RMSE values much easier to interpret visually and would better highlight

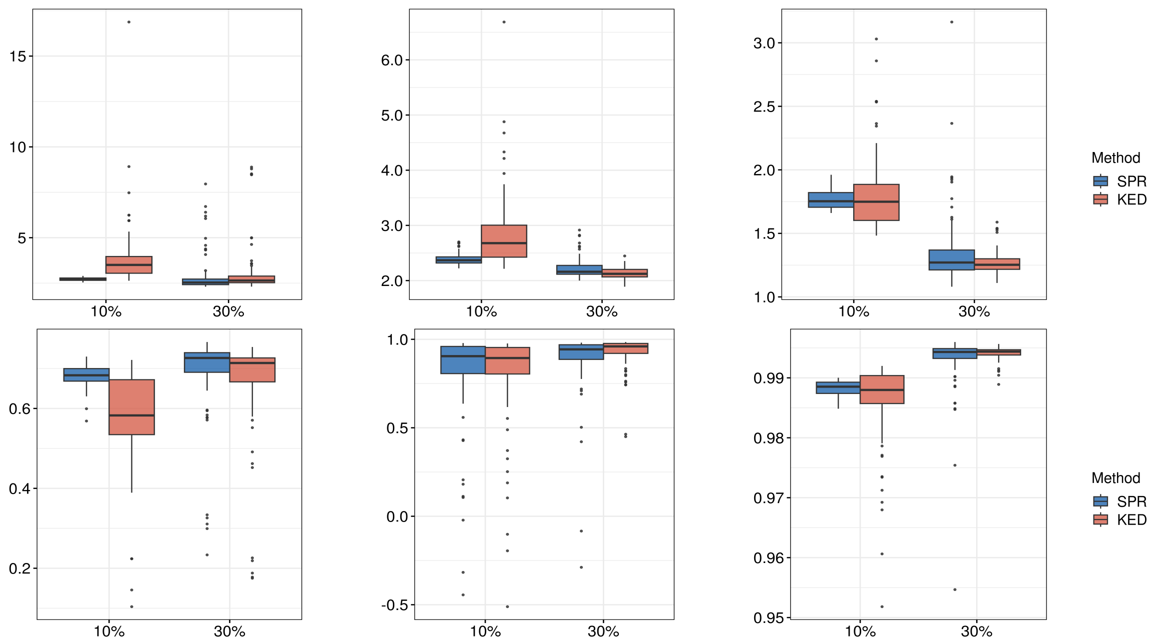
- dispersion,
- outliers,
- robustness across experiments.

Response:

We thank the Referee for this suggestion. Table 3 has been replaced with boxplots showing the distribution of daily RMSE across 100 independent random station samplings for each method, variable, and density level (Fig. 7 in the revised manuscript). This presentation directly addresses the Referee's concern by making the dispersion, outliers, and robustness of results across repetitions immediately visible. The 100 repetitions also respond to the concern raised in Comment 2 regarding sensitivity to station sampling, providing a statistically robust basis for the conclusions drawn from the real-data experiments.

The Section 4.3 has been updated to the following (lines 342–354) :

> *Figure 7 summarizes the distribution of interpolation performance across 100 independent random station samplings, for each method, variable, and density level. The consistency of results across repetitions confirms that the conclusions are robust to the specific choice of training stations.*
>
> *At the lowest density (10%), SPR systematically achieves lower median RMSE than KED for precipitation and both temperature variables, while maintaining higher SSIM values. Notably, the interquartile ranges of SPR are more compact than those of KED, indicating that SPR is not only more accurate on average but also more robust to the specific choice of training stations under very sparse network conditions.*
>
> *At 30% density, differences between SPR and KED become smaller and more variable across variables, with SPR yielding better accuracy only for precipitation. These results suggest that the primary advantage of SPR emerges in sparse-network settings, while its performance remains competitive as station density increases within the low-density regime considered here.*
>
> *These findings are consistent with the synthetic experiments, which indicate that the added value of SPR is most pronounced under sparse observational coverage.*



(a) Precipitation (b) Minimum temperature (c) Maximum temperature

Figure 7. Validation with real observations: distribution of mean daily RMSE (first row) and SSIM (second row) across 100 independent random station samplings for SPR and KED, for each variable. Each box shows the median, interquartile range, and whiskers extending to 1.5 times the interquartile range. Lower RMSE and higher SSIM indicate better performance. Note that 10% density corresponds to 2–4 stations in a region of 70,000 km^2.

References

Caillaud, C., Somot, S., Alias, A., Bernard-Bouissières, I., Fumière, Q., Laurantin, O., Seity, Y. and Ducrocq, V. (2021). Modelling Mediterranean heavy precipitation events at climate scale: an object-oriented evaluation of the CNRM-AROME convection-permitting regional climate model. *Climate Dynamics*, 56(5), 1717–1752. https://doi.org/10.1007/s00382-020-05558-y

Carreau, J., and Guinot, V.: A PCA spatial pattern based artificial neural network downscaling model for urban flood hazard assessment, Adv. Water Resour., 147, 103821, https://doi.org/10.1016/j.advwatres.2020.103821, 2021.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M. and Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123, 9391–9409. https://doi.org/10.1029/2017JD028200

Daly, C., Taylor, G. H. and Gibson, W. P. (1997). The PRISM approach to mapping precipitation and temperature. In *Proceedings of the 10th AMS Conference on Applied Climatology*, 675.

Dura, V., Evin, G., Favre, A.-C. and Penot, D. (2025). Improving precipitation interpolation using anisotropic variograms derived from convection-permitting regional climate model simulations. *EGUsphere*, 2025, 1–25. https://doi.org/10.5194/egusphere-2025-1779

Erdin, R., Frei, C., and Künsch, H. R.: Data transformation and uncertainty in geostatistical combination of radar and rain gauges, J. Hydrometeorol., 13, 1332–1346, https://doi.org/10.1175/JHM-D-11-096.1, 2012.

Fletcher, S. J. and Zupanski, M.: A data assimilation method for log-normally distributed observational errors, Q. J. R. Meteorol. Soc., 132, 2505–2519, https://doi.org/10.1256/qj.05.222, 2006.

Fortin, V., Roy, G., Donaldson, N., and Mahidjiba, A.: Assimilation of radar quantitative precipitation estimations in the Canadian Precipitation Analysis (CaPA), J. Hydrol., 531, 296–307, https://doi.org/10.1016/j.jhydrol.2015.08.003, 2015.

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873–900, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Gasset, N., Fortin, V., Dimitrijevic, M., Carrera, M., Bilodeau, B., Muncaster, R., Gaborit, É., Roy, G., Pentcheva, N., Bulat, M., Wang, X., Pavlovic, R., Lespinas, F., Khedhaouiria, D., and Mai, J.: A 10 km North American precipitation and land-surface reanalysis based on the GEM atmospheric model, Hydrol. Earth Syst. Sci., 25, 4917–4945, https://doi.org/10.5194/hess-25-4917-2021, 2021.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Khedhaouiria, D., Gasset, N., Fortin, V., Dimitrijevic, M., Bulat, M. and Wang, X. (2026). The Canadian Surface Reanalysis (CaSR) v3.2 precipitation dataset: A 45-year high-resolution analysis for North America (1980–2024). *EGUsphere* [preprint]. https://doi.org/10.5194/egusphere-2026-620

Leduc, M., Mailhot, A., Frigon, A., and others: The ClimEx Project: A 50-Member Ensemble of Climate Change Projections at 12-km Resolution over Europe and Northeastern North America with the Canadian Regional Climate Model (CRCM5), J. Appl. Meteorol. Climatol., 58, 663–693, https://doi.org/10.1175/JAMC-D-18-0021.1, 2019.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D. R., and Brekke, L.: A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013, Sci. Data, 2, 150042, https://doi.org/10.1038/sdata.2015.42, 2015.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Schiemann, R., Liniger, M. A., and Frei, C.: Reduced space optimal interpolation of daily rain gauge precipitation in Switzerland, J. Geophys. Res., 115, D14109, https://doi.org/10.1029/2009JD013047, 2010.

Taylor, M. H., Losch, M., Wenzel, M., and Schröter, J.: On the Sensitivity of Field Reconstruction and Prediction Using Empirical Orthogonal Functions Derived from Gappy Data, J. Clim., 26, 9194–9205, https://doi.org/10.1175/JCLI-D-13-00089.1, 2013.

van Hyfte, S., Le Moigne, P., Bazile, E., Verrelle, A., and Boone, A.: High-resolution reanalysis of daily precipitation using AROME model over France, Tellus A, 75, 27–49, https://doi.org/10.16993/tellusa.95, 2023.

Vernay, M., Lafaysse, M., and Augros, C.: Radar-based high-resolution ensemble precipitation analyses over the French Alps, Atmos. Meas. Tech., 18, 1731–1755, https://doi.org/10.5194/amt-18-1731-2025, 2025.

Werner, A. T., Schnorbus, M. A., Shrestha, R. R., Cannon, A. J., Zwiers, F. W., Dayon, G., and Anslow, F.: A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America, Sci. Data, 6, 1–16, https://doi.org/10.1038/sdata.2018.299, 2019.