Manuscript: The Canadian Surface Reanalysis (CaSR) v3.2 precipitation dataset: A 45-year high-resolution analysis for North America (1980–2024)

Author’s response

We want to thank the Editor, and referees for their extensive reviews and for the opportunity to improve our paper. Reviewer comments are addressed in the following with a point-by-point response in italic. Added or modified sentences to the revised version of the manuscript are indicated in italic blue. We hope we have adequately answered all the reviewers’ comments.

Please note that answers to Reviewer #1 are from pages 1 to 9 and from pages 10 to 14 for reviewer #2.

= =

Reviewer #1:

This paper, submitted by Khedhaouiria et al., explains the methodology behind the recent Canadian Surface Reanalysis (CaSR) v3.2 precipitation dataset for North America and assesses its overall quality

General comment:

Developing a climate reanalysis is a task that is both meticulous and far-reaching, given the sheer volume of supporting data and the many pitfalls involved. Not to mention that Canada is a vast country, large parts of which are under-instrumented, requiring a tailored strategy. The work presented here, which focuses on total precipitation, is the result of ongoing improvements to Canadian hydroclimatic products. These products are particularly useful to the operational hydrology community.

- *Authors’ response: First, we thank the reviewer for his comments. Most of its suggestions have been incorporated in the revised paper. Responses are indicated in italic below and italic blue when it refers to revised text in the manuscript.*

Major comments:

It would have been helpful to state clearly from the outset that the database compiled relates to total precipitation, without attempts to identify the phase of the precipitation. I do not dispute the decision not to distinguish between liquid and solid precipitation, but rather the decision not to state this clearly or to justify it.

- *Authors’ response: The reviewer is right to raise this point on the phase information as it is important information for several users. Throughout the manuscript, we describe* *the precipitation reanalysis, and it was implicitly assumed that "precipitation" referred to total precipitation. This comment is even more relevant because CaSR version 3.2 also makes available four additional variables related to precipitation phase: "Quantity of liquid precipitation", "Snow amount in liquid water equivalent of falling snow", "Quantity of freezing precipitation (liquid water equivalent)" and "Quantity of ice pellets (liquid water equivalent)" (which sum-up to the total precipitation amount). These are for now forecast variables produced by the online component and are not involved in the offline precipitation reanalysis, which is the focus of this paper. We have added the "total" in line 1 of the abstract and in line 62: "The recent release of CaSR version 3.2 provides [...] total precipitation reanalysis.". We also explicitly mention the points above in the beginning of Section 4 (Lines 226-230): "The CaSR v3.2 online component distributes a broader set of hourly surface and near-surface variables relevant to hydrometeorological applications, including temperature, humidity, wind, pressure, radiation fluxes, and, since version 3.2, precipitation-phase-related fields. These variables are part of the distributed CaSR v3.2 products but fall outside the scope of the present paper, which is restricted to the offline precipitation reanalysis. Details on data access and download are provided in the Data availability section at the end of the paper".*

Section 6.1 – The manuscript draws on numerous databases, and their use is not always clear to the reader. This is particularly the case here, where numerous databases are evaluated and compared without specifying the source of the observations used as a reference. We must assume that these are the ones described in Figure 1, but they are then described as having been used for assimilation, not for evaluation. The same applies in Appendix A2: Details of assimilated data sets. In fact, the information is provided further on, in Section 6.2, line 364.

- <u>*Authors' response:*</u> *We want to thank the reviewer for this relevant comment. Operationally, CaPA is typically evaluated against independent gauge observations using a leave-one-out (LOO) framework. As part of this process, the CaPA workflow generates both the standard two-dimensional (2D) analysis and a corresponding LOO analysis, which is described in Section 5.1. In the CaPA-24h of CaSR-3.2, we added a comparison with two gridded datasets: a reanalysis, ERA5-Land, and an observation-based one, namely PRISM. We have added clarification points in several places in the article:*
    - *"This study documents the configuration of the CaSR v3.2 offline precipitation component (CaPA-24h) and evaluates its performance against CaSR v2.1*

using gridded and gauge-based observational datasets as reference, with a focus on categorical skill scores, high-impact precipitation metrics, and the representation of the diurnal cycle." (Lines 62-65)

- *The section '2.3 Reference gridded datasets for evaluation' was changed for "2.3 Evaluation datasets" and now contains a small introduction that describe the three types of datasets and have a new subsection dedicated to the surface gauges (Lines 147-156):*

  2.3 Evaluation datasets

  Three complementary references are used to evaluate the precipitation reanalysis: surface station observations in a leave-one-out framework for point-scale verification, and two gridded products, described below, for product intercomparison.

  2.3.1. Leave-one-out validation dataset

  Point-scale verification is performed using a leave-one-out validation framework applied to the same surface station observations described in Section 2.2. For each validation experiment, the target station is excluded from the assimilation, ensuring that the analyzed precipitation at that location is estimated solely from neighboring observations and the background field. This approach provides an independent reference for evaluating the precipitation analysis while maximizing the use of the available observing network. The resulting leave-one-out analyses constitute the primary reference for all station-based verification presented in this study. Details of the leave-one-out implementation are provided in Section 5.1

Moreover, the sentence "Specifically, CaPA generates LOO precipitation analyses at each station location by sequentially removing that station from the assimilation process and estimating precipitation there based solely on surrounding observations." in section 5.1 was removed to avoid redundancy.

- *Finally, we added the following sentence at the beginning of section 5.1: "Precipitation estimates from the evaluated products were verified against daily accumulations from in situ gauge observations, using the station network described in Sect 2.3.1." (Line 256-257)."*

Section 6.4 – If I understand correctly, metrics are calculated for each grid point for the 17-year period from 2002 to 2018 in the PRISM database. Since the selected indicators are annual values, the time series for each point consists of 17 values. I am not sure that it is convincing to present KGE on such short time series. In this respect, in Figure 9, I have more confidence in the assessment of bias and correlation than in that of variability and therefore the KGE. Personally, I would not present the KGE and variability. It seems to me that the issue of the length of the series deserves at least a detailed comment. What is written on line 482 seems insufficient to me.

- *<u>Authors' response</u>: We thank the reviewer for this very important comment which was also made by reviewer #2. To try to answer this comment we estimated uncertainty associated to KGEs and its components. To do so, we used a non-parametric bootstrap approach (1000 replicates of the 17 annual values, resampled in pairs [prism, CaSR /Era5-Land]) to estimate the 90% confidence interval (CI) of the KGE and its components. The results are presented as maps of the CI width in the supplementary material (Section S3 and S4) and confirm some of the reviewer's intuition:*
    - *For Rx1Day: the variability component is the least stable (median 90% CI widths across the domain of 0.56–0.69 for CaSR v3.2, depending on the season) and does influence the KGE values, whereas the bias component is considerably better constrained (0.23–0.27) and the correlation is intermediate (0.34–0.51).*
    - *For R95pTOT: the bias component is the least stable (90% CI widths across the domain of 0.39–0.55, depending on the season) probably due to a most skewed distribution than Rx1Day, while variability and correlations have small CIs.*

  Moreover, generally the CIs width for each component excludes 1 for more than 90% of grid cells, indicating that the mapped component patterns are robust to sampling variability. Therefore, the revised manuscript keeps the KGE as a qualitative indicator of spatial patterns but bases the interpretation of Fig. 9 mainly on the most robust components. We modified the manuscript consequently at several places:

  - *Section 5.2 (Lines 325-328): "To account for the short length of the time series, which can lead to substantial uncertainty in some KGE components and therefore in the KGE itself (Clark et al., 2021), a non-parametric bootstrap approach (Tibshirani and Efron, 1993) with 1000 replicates of the annual values, resampled in pairs was applied at each grid cell to estimate 90% confidence intervals."*
  - *Section 6.4 (Lines 479-480): "Given the limited number of annual samples, the corresponding bootstrap confidence intervals (Sect. S3 of the supplementary material) are used throughout this section to identify robust signals."*
  - *Section 6.4 (Lines 484-485): "Part of this noisiness reflects sampling uncertainty rather than actual disagreement, the variability ratio is the least constrained*

component of the KGE, with 90% confidence intervals typically exceeding 0.5 in width (Sect. S3)."

- *Section 6.4 (Lines 509-511): "Low KGE values are primarily driven by large biases, with both CaPA-24h v3.2 and ERA5-Land overestimating mean R95pTOT by at least a factor of two in several regions and seasons; the sign of these biases is robust to sampling variability, as the bootstrap confidence interval for β excludes 1 in 98-99% of grid cells (Sect. S3)."*
- *The last paragraph of Section 6.2 (Lines 523-527) was changed from: "KGE and its components were also computed for Rx5day and R99pTOT (not shown; see Supplement). Regional and seasonal patterns are consistent with those obtained for Rx1day and R95pTOT, although performance generally decreases for these more extreme indices, which remain more challenging to reproduce in reanalyses." for "KGE and its components were also computed for Rx5day and R99pTOT (not shown; see Supplement). Regional and seasonal patterns for Rx5day are consistent with those obtained for Rx1day, with slightly reduced sampling uncertainty due to the temporal smoothing of the 5-day accumulation. For R99pTOT, spatial patterns resemble those of R95pTOT, but sampling uncertainty becomes very large (median β CI widths exceeding 3; Sect. S4), so that grid-cell scores for this index is difficult to interpret."*

Figure 11 – The diurnal cycle is essentially a local process. Is there an issue with combining values across large areas? Furthermore, the decision to present times in UTC complicates interpretation, as not everyone is familiar with how to perform the conversion.

- <u>*Authors' response:*</u> *We thank the reviewer for this useful comment. We agree that the diurnal cycle is a local process and that averaging across regions may displace the timing and smooth the precipitation intensity. To add a bit of context, this section was added as a response to some users who wanted to look at the diurnal cycles. We found that due to the approach of stitching the different leadtimes as shown in Figure A1, some discontinuities appear in the precipitation diurnal cycle and can sometimes interfere with the real daily lows and highs. We therefore thought that it would be useful information to share with users.*
  This being said, expected diurnal cycles are still observed at the regional scale and are visible in several regions as shown in Figure 11. For example, convective precipitation tends to peak in the late afternoon or evening and a weaker diurnal cycle is observed in the winter. Some of these observations are consistent with the literature of the precipitation daily cycle across North America (Dai et al., 1999).

We also agree that presenting time only in UTC can complicate interpretation. The original choice of UTC was motivated by the structure of the product and by the 12 UTC–12 UTC accumulation window used in CaPA-24h. However, because some regions span several time zones (eg. WBoreal in light yellow of Figure B1), UTC seemed to be a good compromise. However, to better guide the readers, we added in Figure 11 the mean local solar noon, computed from the longitudes of the grid cells within each region (5th–95th percentile range). It helps prevent the use of the local time which varies according to winter or summer and regions and it permits us to see for example that few hours (3 to 6) after the mean local solar noon occurs the precipitation high in summer. Local solar time is regularly used in diurnal cycle studies (e.g. Dai et al., 2024). The caption of Figure 11 was modified accordingly: "Diurnal cycle of hourly precipitation intensity in winter (a, DJF) and summer (b, JJA) for CaSR v3.2 (blue) and ERA5-Land (gold). The thick line shows the regional median of the hourly mean precipitation over 1980–2018, and the shaded area denotes the interquartile range (25th–75th percentile). The dotted vertical line (or grey band, for regions with a large longitudinal extent) indicates mean local solar noon, computed from the longitudes of the grid cells within each region (5th--95th percentile range)."
We also added the following in the manuscript to account for this new addition:

- *Lines 535-537: "To ease the interpretation of UTC times, mean local solar noon is indicated in each panel, providing a fixed geographic reference that is independent of time-zone boundaries"*
- *Lines 539-540: "[..]where precipitation activity in both CaSR v3.2 and ERA5-Land exhibits a minimum near local solar noon and remains elevated from the evening through the early morning, [...]"*
- *Lines 547: "[...] summer precipitation generally peaks in the late afternoon, around 3 to 6 h after local solar noon (i.e. 21–00 UTC depending on the region)"*

The work presented here is partly motivated by hydrological applications. Indeed, the success of hydrological models depends largely on the quality of historical precipitation data – although this is not the only challenge they face. However, this precipitation alone is insufficient. At a minimum, an air temperature database is also required to clarify the phase of precipitation and to model the melting of snow accumulated on the ground – essential components of hydrology in a northern climate such as Canada's. Although this is not the focus of this manuscript, it would be useful for potential users if this issue were addressed in the discussion or conclusion, for example by suggesting air temperature databases (and why not other hydrometeorological variables) compatible with CaSR v3.2.

- *<u>Authors' response:</u> We think the reviewer comment is legitimate to raise. We have added sentences at two places to highlight the existence of the different other surface variables generated by the CaSR product. In the introduction, we now can read (Lines 60-62): "The recent release of CaSR version 3.2 provides approximately 40 surface variables of direct relevance to hydrological and surface models, including near-surface temperature, snow-related variables, wind, and surface fluxes, and provides the foundation for a new offline total precipitation reanalysis". We found also relevant to add the information in the section dealing with the variables distributed to users (Sect. 4), where from lines 226 to 230 we have: "The CaSR v3.2 online component distributes a broader set of hourly surface and near-surface variables relevant to hydrometeorological applications, including temperature, humidity, wind, pressure, radiation fluxes, and, since version 3.2, precipitation-phase-related fields. These variables are part of the distributed CaSR v3.2 products but fall outside the scope of the present paper, which is restricted to the offline precipitation reanalysis. Details on data access and download are provided in the Data availability section at the end of the paper."*

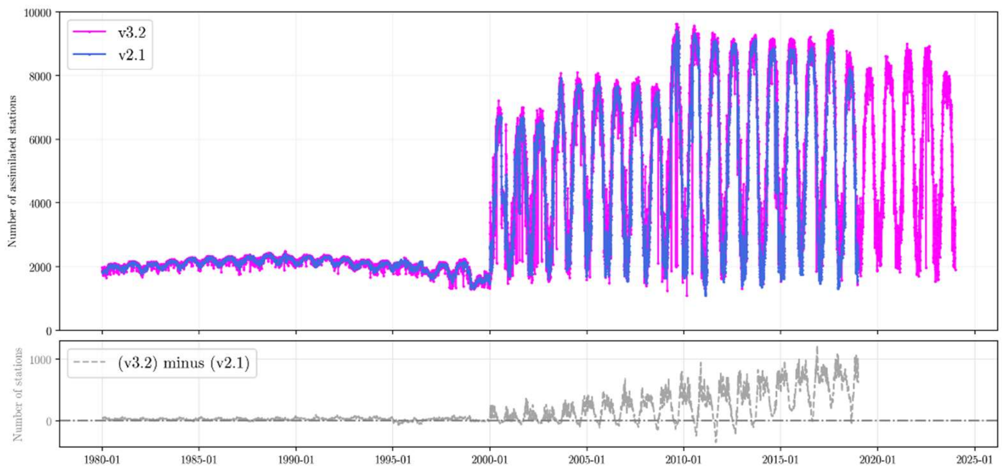
Minor comments:

Line 50 – Please write out 'NWP' in full when it first appears

- *<u>Authors' response:</u> It is already defined in line 27 in the introduction.*

Line 135 – A brief note on the annual fluctuation in Figure 1 would be helpful. The explanation comes much later in the text, at lines 599 and 617. Also, Figure 1 would be easier to interpret if the y-axis started at zero.

- *<u>Authors' response</u>: Indeed, the reviewer is right as the only sentence available in the paper in line 113: "Seasonal variability in the number of assimilated 24h accumulation observations is also evident in Figure 1." is a bit light and therefore, we have added to this sentence the following (Lines 117-120): "The lower number of assimilated gauges during winter reflects the application of more restrictive QC filters under solid precipitation conditions. Although this reduces observational coverage, its impact on the analysis is partly mitigated by the generally higher skill of the background field during winter (Khedhaouiria et al. 2020).". The suggestion for Figure 1 was also integrated as follows:*



Line 300 – The equations describing these metrics could be placed in the appendix with the others already there.

- *<u>Authors' response</u>: Done. The KGE and its component are now detailed in section D2 of Appendix D.*

Figure 14 – I am not sure that this figure is strictly necessary.

- *<u>Authors' response</u>: It is true that it is not strictly necessary, but we think it illustrates well that the operational CaPA analysis can be a good candidate to extend CaPA-24h from CaSR v3.2 beyond 2024 for selected precipitation statistics (here, seasonal accumulations). Indeed, several users expressed the need to have near real time data which is unfortunately not possible for this version of CaSR. Since operational CaPA and CaPA-24h from CaSR v3.2 differ in several respects listed in Table 1, users may legitimately wonder whether homogeneity is preserved when combining the two. Figure 14 provides evidence that it is reasonably well preserved for these statistics. We, however, agree that this comparison does not constitute a formal homogeneity assessment with statistical tests, but it offers a check for users considering such an extension.*

Ligne 645 – Is it possible to provide an appreciation of the number of stations provided to the study by these different networks?

- *<u>Authors' response</u>: As suggested by reviewer #1, we added the number of stations per network, and we take advantage of this modification to slightly re-arrange the initial version of Table A1 to add more transparency about the networks. See below the print screens of the table before and after:*

Initial version

Network	Domain	Period (Origin)	CaPA-24h
SYNOP	North America	1980–1999 (ISD)	P, T, $\|\|U\|\|$
		2000–present (ECCC)	–
SWOB	North America	1980–1999 (ISD)	P, T, $\|\|U\|\|$
		2013–present (ECCC)	–
METAR	North America	1980–1999 (ISD)	P, T, $\|\|U\|\|$
		2000–present (ECCC)	–
AdjDlyRS	Canada	1980–2023 (ECCC)	P
AdjHlyRS	Canada	2001–2023 (ECCC)	P
RMCQ	Quebec	2011–present (ECCC)	P, T, $\|\|U\|\|$
SHEF	USA	2000–present (ECCC)	P, T, $\|\|U\|\|$
CoCoRaHS	Canada–USA	(ECCC)	P

Modified version

Network	Domain	Period (Origin)	N_{sta}	Variables
SYNOP	North America	1980–1999 (ISD); 2000–present (ECCC)	2870	P, T, $\|\|U\|\|$
SWOB[b]	North America	1980–1999 (ISD); 2013–present (ECCC)	2140	P, T, $\|\|U\|\|$
METAR	North America	1980–1999 (ISD); 2000–present (ECCC)	3540	P, T, $\|\|U\|\|$
AdjDlyRS	Canada	1980–present	3730	P
AdjHlyRS	Canada	2003–present	380	P
RMCQ	Quebec	1998–present	300	P, T, $\|\|U\|\|$
SHEF[a]	USA	1980–present	21160	P, T, $\|\|U\|\|$
CoCoRaHS	Canada	2014–present (ECCC)	150	P
NoN	Canada	2000–present (ECCC)	780	P

= =

Reviewer #2:

General comment:

This paper presents an important 45-year high-resolution North American precipitation reanalysis (CaSR v3.2/CaPA-24h) that will be useful for hydrological, climatological, and impact studies. The methodology is generally well documented, and the new version demonstrates several important improvements over previous releases, particularly in reducing background-field biases and improving performance in data-sparse regions. The evaluation is comprehensive, including comparisons with CaSR v2.1, ERA5-Land, PRISM, and station-based leave-one-out verification. The authors also discuss the remaining limitations. Overall, I think the paper is well organized, and the results are convincing. I have several minor concerns regarding the presentation of the methodology, the interpretation of the results, and the wording of some conclusions.

- *Authors' response: We want to thank the reviewer for his comments as it helped to improve the first submitted version. All the comments have been incorporated in the revised paper. Responses are indicated in italic below and italic blue when it refers to revised text in the manuscript.*

Comments:

- The manuscript repeatedly refers to PRISM as an independent evaluation dataset. However, PRISM is itself a gauge-based interpolation product and likely relies on many of the same station observations assimilated by CaPA. Therefore, the comparison may not be fully independent, and part of the agreement between CaPA and PRISM could arise from their shared observational constraints. I suggest that the authors clarify the degree of independence between the two datasets and discuss this limitation.
    - *Authors' response: We thank the reviewer for this relevant comment. We have added a new paragraph in the PRISM description (Section 2.3.3, Lines 183-187): "A comprehensive description of the model algorithms and input data is provided by Daly et al. (2008). It should be noted that PRISM itself is a gauge-based product and incorporates surface observations from networks that partly overlap with those assimilated by CaPA over the CONUS. The comparison presented here therefore does not constitute a strictly*

independent validation; agreement between the two products should rather be interpreted as an indication of consistency in the representation of spatial precipitation patterns, obtained through fundamentally different methodologies." Moreover, the word 'independent' was removed from the sentence in line 626 "This study evaluated CaPA-24h v3.2 relative to its predecessor (v2.1) and to two ~~independent~~ gridded datasets, ERA5-Land" in the revised version of the manuscript.

- A related concern applies to the leave-one-out (LOO) verification. In dense gauge networks, particularly over eastern North America, spatial correlations can extend well beyond the spacing of individual stations. It would be helpful if the authors could discuss how much this residual dependence may still influence the reported verification scores.

  - *Authors' response: The reviewer is right to raise this concern. Section 5.1 already acknowledges this limitation in the context of the LOO framework: "This approach mitigates [...] independent verification." The conclusions mentioned also this point in limitations paragraph (Lines 652-654). "Although the leave-one-out framework reduces the direct impact of assimilated observations on verification, spatial dependence among validation points cannot be fully eliminated". However, this point is indeed not discussed further in the results, and the reported scores should be interpreted as an upper bound on the skill expected in truly ungauged areas.*

    A related question is whether the LOO framework nevertheless allows a correct ranking of the different products despite this potential score inflation. Two elements suggest that it does. First, the very similar scores in Figure 4 for CaPA-24h from CaSR v2.1 and v3.2 illustrate a probable limit of the LOO framework: over such a long evaluation period, and with largely similar assimilated networks (even though thinning is applied), both analyses are strongly constrained by the same observations. However, the corresponding background fields, which are not influenced by the assimilated stations, clearly show the added value of CaSR v3.2 over both CaSR v2.1 and ERA5-Land. Second, the markedly lower verification scores of ERA5-Land make it very unlikely that residual spatial dependence alone could alter the ranking CaPA-24h (CaSR v3.2) ≥ CaPA-24h (CaSR v2.1) > ERA5-Land.

    A fully independent verification (e.g., leave-region-out) would remove this residual dependence, but is difficult to implement in practice over the 45-year period: excluding entire networks or regions would either degrade the *analyses themselves or restrict the evaluation to a subdomain that is not representative of the full domain. For this reason, the LOO framework used for the operational CaPA product was retained. In addition, internal experiments using CaPA-24h from CaSR v3.2 as forcing hydrological simulations provide further, albeit indirect, support for the added value of this update.*

We have added in the manuscript the following sentences in blue in Section 6.1:

- *"Overall, differences between CaPA-24h versions 2.1 and 3.2 are relatively small. This is expected, as both versions assimilate largely similar precipitation observations and are evaluated at station locations using a leave-one-out (LOO) framework […]" (Lines 340-342)*
- *"Finally, because LOO analyses retain information from neighboring assimilated stations, verification scores in densely instrumented regions should be interpreted as an upper bound on the performance expected in truly ungauged areas. Since this residual spatial dependence is nearly identical for both CaPA-24h versions, however, it does not affect their relative comparison." (Lines 346-350)*

- The evaluation of extreme precipitation indices is based on the 2002–2018 period, which provides only 17 annual samples for metrics such as Rx1day at each grid cell. Such a limited sample size may introduce considerable sampling uncertainty into KGE and related statistics.
    - *Authors' response: This comment is indeed important and was also raised by reviewer #1. We answered this concern in the present document from page 3 to page 5. We kindly invite reviewer #2 to refer to this part of the document.*
- Since the dataset is intended as a 45-year climate record (1980–2024), I wonder whether the authors have evaluated its temporal homogeneity across the major observing-system transition around 2000. It would be helpful to discuss whether these changes could introduce artificial shifts or affect long-term trend analyses.
    - *Authors' response: We thank the reviewer for this particularly relevant comment. It is true that the use of two data sources providing significantly different numbers of stations may raise questions regarding the homogeneity of the series, especially for a reanalysis product. In the manuscript, the impact of the transition around 2000 on the verification scores was examined, and the results are discussed in Section 6. However, no trend analysis is described in the manuscript, as it would require an extension well beyond the scope of the*

present article, and could constitute a study of its own. Such work would require choosing the variables (annual maxima, annual accumulation, etc.), the period, and the geographical extent of the trend analysis, and comparing the results to other reanalyses or to trend studies based on homogenized station data (e.g., Zhang et al., 2000). However, we did the following internal tests on the winter (DJF) and summer (JJA) seasonal accumulation time series:

- *Spatial distribution of statistically significant (5% level) linear trends estimated using ordinary least squares. Significant trends were scattered across the domain, with no coherent spatial pattern.*
- *Spatial distribution of statistically significant (5% level) Pettitt change-point tests. The Pettitt test identifies a single breakpoint in each annual time series, allowing us to assess whether detected breakpoints cluster around the year 2000. As shown in Figure 1, breakpoint years are distributed throughout the study period, with no systematic concentration around 2000. A few isolated grid cells with a breakpoint in 2000 are visible in DJF, but they remain spatially limited.*

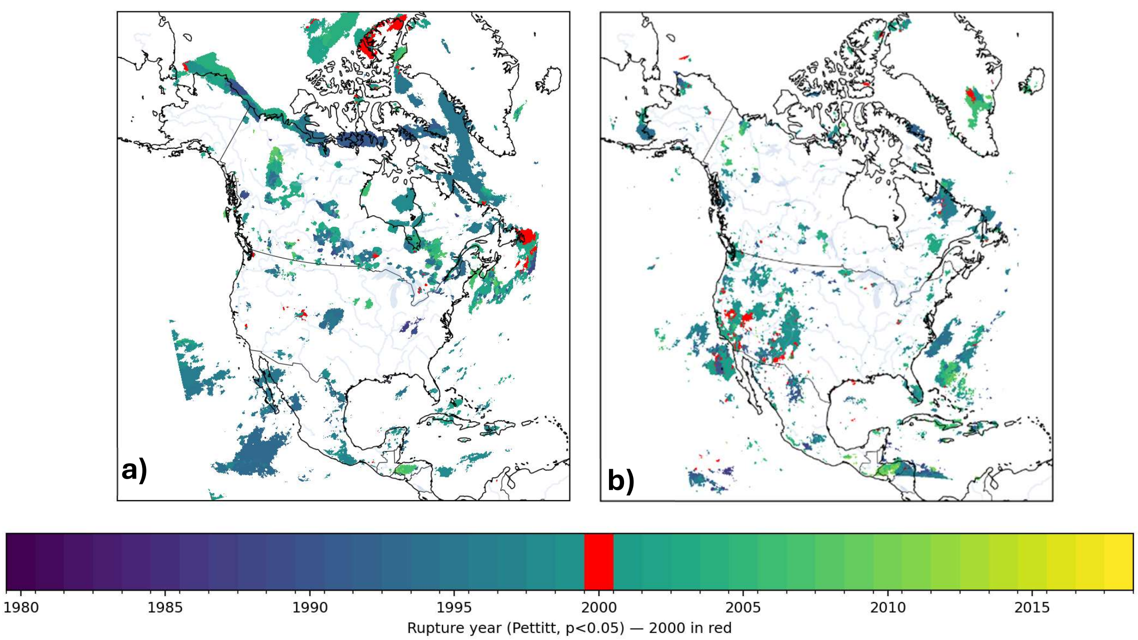


Figure 1 Significant break-point year for the DJF (a) and JJA (b) annual precipitation time series using the non-parametric Pettitt test at the 95% confidence level, the year 2000 is highlighted in red.

However, this is a preliminary work: it does not examine temporal inhomogeneity for other precipitation-based variables, nor potential breakpoints associated with other dates at which specific observing networks were added to the assimilation.

A cautionary statement regarding long-term trend analyses has been added to the limitations paragraph of the Conclusions (Lines 657-660). "The pre-2000 evaluation period is also more sensitive to sampling, owing to reduced station density and more limited wintertime gauge availability. This period also coincides with a major transition in the observing system around 2000; although preliminary evaluations did not reveal any systematic discontinuity associated with this transition, caution is advised when using the dataset for long-term trend analyses"

- In Line 42, the manuscript states that "precipitation is not directly assimilated in atmospheric reanalyses," but the following sentence mentions that systems such as CFSR ingest observation-based precipitation products. This wording may be confusing for readers who are not familiar with modern reanalysis systems. I suggest revising this paragraph.
    - Authors' response: *The reviewer is right that the original sentence could appear contradictory, as "assimilated" was used in two different contexts. The intended distinction is between the atmospheric analysis, where precipitation is not directly analyzed because it is produced by the model, and the use of precipitation observations to improve surface reanalyses. These observations can either be used to drive the land surface component (e.g., CFSR and CaSR online component) or incorporated into a separate offline precipitation analysis (e.g., CaSR). The paragraph has been revised as follows (Lines 43-47): "Unlike other atmospheric variables, precipitation is generally not directly assimilated in atmospheric reanalyses because it is not a prognostic state variable but rather the accumulation of a flux over a specific time period. Instead, observation-based precipitation information is used to constrain the land surface component. For instance, some systems ingest observation-based gridded precipitation products, such as CFSR (Saha et al 2010); while others rely on in-situ station data, as in the Canadian Surface Reanalysis (CaSR; Gasset et al. 2012)."*

References

Zhang, X., Vincent, L. A., Hogg, W. D., & Niitsoo, A. (2000). Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean*, *38*(3), 395–429. https://doi.org/10.1080/07055900.2000.9649654

Dai, A., Lin, X., & Hsu, K. L. (2007). The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low-and mid-latitudes. *Climate dynamics*, *29*(7), 727-744.