# Observed Precipitation Trends Inferred from Canada's Homogenized Monthly Precipitation Dataset

XIAOLAN L. WANG,[a] YANG FENG,[a] VINCENT Y. S. CHENG,[a] AND HONG XU[b]

[a] *Climate Research Division, Science and Technology Branch, Environment and Climate Change Canada, Toronto, Ontario, Canada*
[b] *Ottawa, Ontario, Canada*

(Manuscript received 4 April 2023, in final form 10 August 2023, accepted 10 August 2023)

ABSTRACT: This study first developed a comprehensive semiautomatic data homogenization procedure to produce gap-infilled and homogenized monthly precipitation data series for 425 long-term/critical stations in Canada, which were then used to assess Canadian historical precipitation trends. Data gaps in the 425 series were infilled by advanced spatial interpolation of a much larger dataset. The homogenization procedure repeatedly used multiple homogeneity tests without and with reference series to identify changepoints/inhomogeneities, the results from which were finalized by manual analysis using metadata and visual inspection of the multiphase regression fits. As a result, 298 out of the 425 data series were found to be inhomogeneous. These series were homogenized using quantile matching adjustments. The homogenized dataset shows better spatial consistency of trends than does the raw dataset. The improved gridding and regional mean trend estimation methods also provide more realistic trend estimates. With these improvements, Canadian historical precipitation trends were found to be dominantly positive and significant, except in central-south Canada where the trends are generally insignificant and small with mixed directions. For annual precipitation, the largest increases are seen in southeastern Canada and along the Pacific coast; however, the largest relative increases (in percent of the 1961–90 mean) are seen in northern Canada. The largest trend difference between northern and southern Canada is seen in winter, in which significant increases in the north were matched with significant decreases in the south.

SIGNIFICANCE STATEMENT: This study aims to produce a homogenized long-term monthly precipitation dataset for Canada, which is then used to assess Canadian historical precipitation trends. The work is important because it developed a comprehensive algorithm for homogenization of precipitation data, and the results provide better representation of precipitation climate and more robust estimates of precipitation trends. It also identified the causes for large biases in the published estimates of precipitation trends over Canada.

KEYWORDS: Precipitation; Trends; Changepoint analysis; Quality assurance/control, Time series; Climate change

## 1. Introduction

Historical records of observations of various key climate variables, such as precipitation and surface air temperature, are required for climate research and applications, and for climate services. Climate observations and research in turn inform mitigation and adaptation decisions in Canada and internationally.

Since precipitation occurs in different forms (e.g., rainfall, snowfall, ice pellets) and is highly variable both temporally and spatially (i.e., from location to location), it is very challenging to measure and quantify its temporal and spatial distributions, especially for northern countries, such as Canada, with a lot of snowfall or other types of solid precipitation.

Since the early 1990s, Environment and Climate Change Canada (ECCC; previously known as Environment Canada) has developed an Adjusted and/or Homogenized Canadian Climate Dataset (AHCCD) for trend and climate change detection analysis. This includes two generations of the adjusted precipitation dataset (Mekis and Hogg 1999; Mekis and Vincent 2011). The adjustments include 1) conversion of ruler measurements of snowfall to its water equivalent and 2) corrections for gauge-related issues (undercatch and evaporation due to wind effects, gauge-specific wetting loss) and for trace precipitation amounts. The adjusted precipitation dataset consists of monthly total precipitation series for 463 manual observation stations. These were chosen by Mekis and Hogg (1999) and refined by Mekis and Vincent (2011), primarily based on whether the stations (after some joining of stations) have data records for the baseline period 1961–90 that are needed to define the baseline climate normal for computing anomalies.

The AHCCD precipitation dataset had been used to produce the Canadian Gridded precipitation relative anomalies (i.e., the CanGRD precipitation dataset). The gridded dataset is then used to derive the regional mean precipitation relative anomalies and trends for the Climate Trend and Variation Bulletin (CTVB), which is published routinely every season by ECCC. Although the AHCCD precipitation data were not

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-23-0193.s1.

*Corresponding author*: Xiaolan L. Wang, xiaolan.wang@ec.gc.ca

DOI: 10.1175/JCLI-D-23-0193.1

homogenized at the station level, at the time it was the best precipitation dataset available for trend analysis.

With the evolution of precipitation monitoring technology, Canada's monitoring network has transitioned from manual observations to using automatic gauges. Due to station automation and closure of volunteer stations, this 463-station network has declined drastically in the recent decade. Only 71 out of the 463 manual stations remain active in 2016 (with data throughout the whole year), and none of the active stations are located north of 60°N. It had become impossible to update and use the AHCCD precipitation dataset to produce reliable precipitation trends. Thus, the update and subsequent products (CanGRD, CTVB) were temporarily halted in 2017, pending extensive data reconciliation to link the current automated observations to the historical manual observations by performing data homogenization.

Climate data homogenization is a procedure to detect and eliminate nonclimatic changes induced in climate data time series by changes in site condition/environment, location, instrumentation, etc. (Wang and Feng 2013). This procedure is necessary and critical, since results of climate research, especially trend analysis, could be largely biased, as will be shown later, if the data series contains nonclimatic changes.

The ultimate objective of this study is to assess Canada's historical precipitation trends using a homogenized high-quality monthly precipitation dataset, which is however not yet available. Thus, the first objective here is to develop a homogenized historical monthly precipitation dataset.

The remainder of this article is arranged as follows. The selection and formation of the long-term stations network for this study, and preparation of the joined precipitation data series are detailed in section 2. The methods for, and results of, data homogenization are described in section 3. Historical trends and the effects of data inhomogeneity thereon are presented and discussed in section 4. This article is completed with a summary and discussion in section 5.

## 2. Data sources and preparation

The following precipitation datasets were used in this study to form and homogenize the long-term monthly precipitation data series for assessing Canada's historical precipitation trends:

- The quality-controlled version 2020 of the Adjusted Daily Rainfall and Snowfall (AdjDlyRS) dataset (Wang et al. 2017), which was developed by Wang et al. (2017) and has undergone an in-depth quality control (QC) procedure (Cheng et al. 2023). The AdjDlyRS dataset includes 3346 stations across Canada and was produced with the following processing (Wang et al. 2017): 1) conversion of snowfall ruler measurements to their water equivalents, 2) corrections for gauge undercatch and evaporation due to wind effect, for gauge specific wetting loss, and for trace precipitation amount, and 3) treatment of flags (e.g., accumulation flags); also, long runs of miscoded missing values were also identified and set to missing.
- Daily total precipitation data from automated gauge stations (including the Belfort, Fisher and Porter, Nipher, Geonor, and Pluvio gauges), taken from the ECCC digital Climate Archive. We also processed and adjusted these data for use in this study.
- Version 1 of ANUSPLIN surfaces of the adjusted monthly precipitation, which was developed by MacDonald et al. (2021) using version 2016 of the AdjDlyRS dataset (Wang et al. 2017). The ANUSPLIN surfaces are thin plate smoothing splines fitted to climate data recorded at in situ stations (more details in section 2c below).
- The Twentieth Century Reanalysis version 3 (20CRv3) ensemble-mean monthly precipitation data for the period from 1840 or later to 2015 (Slivinski et al. 2019) were also used as reference series but without assuming temporal homogeneity (see section 3 below).

In the subsections below, we described the procedures to select long-term/critical stations of precipitation data and to form and derive the long-term monthly total precipitation data series, the method to infill the temporal data gaps in the data series, and the quality control results.

### *a. Selection and formation of core stations*

This study aims to homogenize historical monthly precipitation data for a set of long-term and/or critical stations (referred to as core stations) in Canada, for assessing precipitation trend and variability for Canada. Since it is difficult and time-consuming to develop a homogenized precipitation dataset, it is important to ensure that the stations that have precipitation data to be homogenized will have good chance to be operated and well maintained into the distant future (i.e., so-called future-proof), producing complete data (with few missing values) of good quality for decades to come, to ensure the longevity and temporal consistency of observations.

With this in mind, we selected stations from the currently active automated gauge stations that are producing good precipitation data, and from the 200+ Reference Climate Stations, which are also automated and supposed to be continued into the distant future because it is the only Canadian observing network that was officially designed for climate purposes. From this set of stations, we choose all stations that have collocated/nearby manual stations of precipitation data to form a long-term precipitation data series. The nearby manual stations of adjusted precipitation data were extracted from the quality-controlled AdjDlyRS version 2020 dataset (Cheng et al. 2023). When choosing the most suitable nearby stations to join, we searched for the best nearby stations within a 300-km radius and considered the similarity between the stations (correlation between the first-order difference series), as well as the length and completeness and quality of data record. As a result, 94.6% and 99.6% of the joined stations are within 100- and 200-km radius, respectively. For data-sparse areas, we also allowed stations of shorter and/or poorer data record to be retained as core stations to infill critical spatial data gaps. In the meantime, we also tried our best to have a higher station density for areas of complex topography (e.g., the Rocky Mountain areas) or of low spatial correlation.

Using the above selection criteria, we have chosen the 425 stations shown in Fig. 1 as core stations and formed the joined daily precipitation data series for these stations,

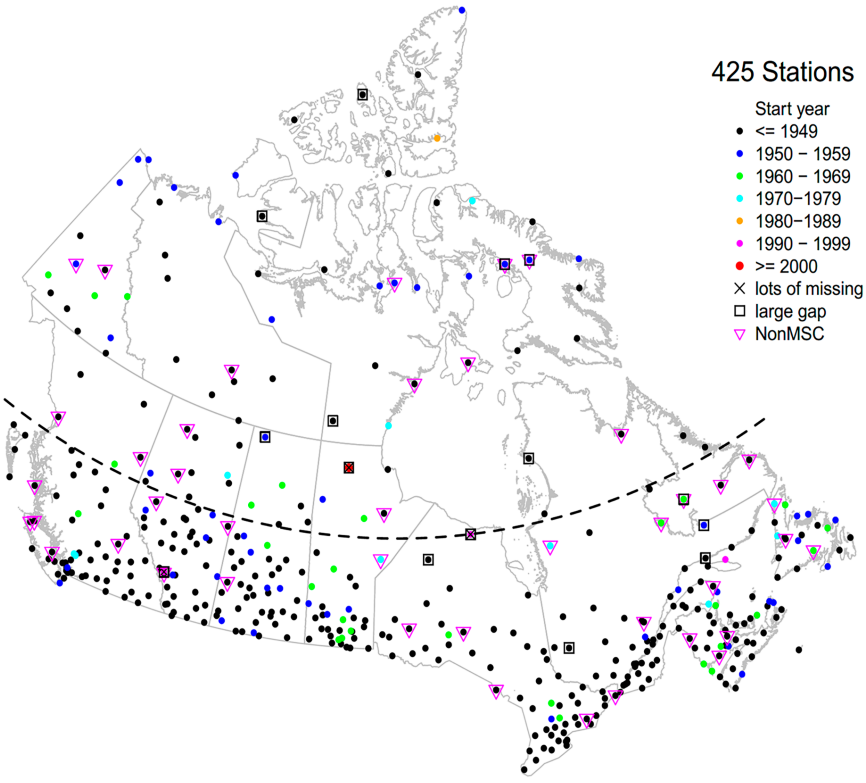


FIG. 1. The 425 long-term and/or critical stations chosen as core stations. The dashed line marks the 55°N parallel.

finishing the procedure I in Fig. 2. Clearly, the station density is much lower in the region north of 55°N latitude. With the joining of stations, these include 75 stations of data since 1890 or earlier, 321 stations of data since 1949 or earlier, and 3 stations of very short record (starting after 1990; see magenta and red dots in Fig. 1). Note that these also include 40 stations that are not owned by ECCC (i.e., NonMSC in Fig. 1); and many of these stations are in the data-sparse areas or have very long data record. It is of critical importance for ECCC to find a way to ensure the continuation of quality observations at these sites.

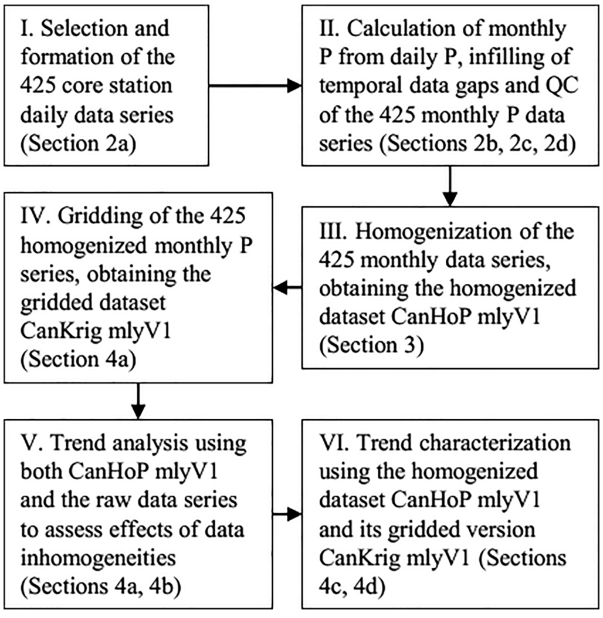


FIG. 2. The flowchart of the procedures of analysis.

### b. *Calculation of monthly total precipitation*

Here we used zero tolerance to do so, namely, a monthly total precipitation ($P$) value is set to missing if there are one or more days of missing values in the month. However, the nonzero monthly total precipitation calculated by allowing an unlimited number of missing daily values is kept as an estimated value for the month (flagged with OE) if it is greater than the ANUSPLIN estimate (see section 2c below). In other words, the total precipitation amounts calculated with missing daily values are not discarded, unless they are less than the corresponding ANUSPLIN estimates. It is also worth noting that data from automated stations are usually of lower quality, having higher percentages of missing data.

Note that the daily precipitation data might contain accumulated amounts, which are flagged with A or F and preceded with a sequence of the days of zero amount with C or L flag (indicating precipitation occurred but amounts not recorded). The accumulation period includes the day with an accumulated amount plus the preceded sequence of consecutive days with

C or L flag. In the AdjDlyRS dataset, an accumulated amount was either distributed to the days in the accumulation period using neighbor stations amounts for these days or set to missing when no qualified neighbor stations data are available for such distribution (Wang et al. 2017). In the calculation of monthly $P$ here (Fig. 2, procedure II), we recovered the set-to-missing accumulated amounts, for such an amount is valid for inclusion in the monthly total $P$, although it is not a valid daily amount (i.e., it occurred in the sequence of 2 or more days) and hence was set to missing in the daily data series. Note that this recovery procedure will make some monthly totals different from those calculated directly from the corresponding quality-controlled daily precipitation (QCdP) data, but it makes the monthly total precipitation amounts more accurate.

### c. *Data and method for infilling data gaps*

We used the quality-controlled ANUSPLIN estimates of monthly precipitation to infill temporal data gaps in the joined monthly precipitation data series. The ANUSPLIN-AdjMlyV1 surfaces were developed by MacDonald et al. (2021) using the ANUSPLIN software package (Hutchinson and Xu 2013) and the AdjDlyRS dataset version 2016 (Wang et al. 2017), which was the best precipitation dataset available then for producing gridded precipitation for Canada. (Note that version 2016 has now been superseded by version 2020.) From these surfaces, the ANUSPLIN estimates of monthly precipitation were extracted for 3881 stations, including the 3346 AdjDlyRS stations and all stations used in developing the 425 joined precipitation data series. Before being used to infill temporal data gaps, these estimates were first quality-controlled as described below.

The ANUSPLIN estimates are one type of optimal spatial interpolation based on high-quality adjusted precipitation data for a large set of stations (3346) across Canada. However, the number of observations is very limited in the pre-1900 period (MacDonald et al. 2021; Wang et al. 2017). This makes the ANUSPLIN estimates for the pre-1900 period much less reliable. To exclude unreliable/bad estimates, we set to missing all ANUSPLIN estimates that deviate more than six standard deviations (SDs) from their respective long-term monthly means, with the SDs and long-term means being calculated from the data over the period of 1900–2015. This produced a quality-controlled version of the ANUSPLIN estimates, which were then used to infill temporal data gaps in the joined monthly precipitation series, and to extend all the 425 joined data series back to at least January 1948 (so the data series cover the period 1948–2015 or longer). The infilled values are flagged with AE (for ANUSPLIN estimate). Further, the monthly total precipitation calculated with one or more missing daily values in the month is used to infill the gap (with flag OE) when it is greater than the ANUSPLIN estimate for that month. In other words, for a month of one or more days of missing observations, the value used to infill the missing monthly precipitation is either an ANUSPLIN estimated monthly value (flagged with AE) or the sum of all nonmissing daily values in the month (flagged with OE), whichever is greater (unless the OE value is identified as an erroneous value). All the AE or OE flagged values that deviate more than six SD from the mean are discarded (recorded as missing with a bAE or bOE flag).

The ANUSPLIN estimates are very good for most cases (e.g., Stn155 in Fig. S1 in the online supplemental material) but could be systematically biased for data-sparse areas and/periods (e.g., Stn401 in Fig. S1). Nevertheless, the infilling of the temporal data gaps increases the sample size and thus improves the data homogenization results, as long as the infilled data segments were tested for homogeneity.

### d. *Data quality control*

Recently, Cheng et al. (2023) developed a comprehensive QC system for in situ precipitation data records, with application to the 425 joined monthly precipitation data series for the period up to September 2019. This system includes a procedure to identify erroneous data of unusually large values, and another one to identify erroneous zero or near-zero monthly total precipitation amount. It uses four nearest stations' data for the same month to determine if a suspect value is a real extreme value or an erroneous value (Cheng et al. 2023). As a result, 351 erroneous values were identified in the 425 joined monthly precipitation series, all of which were set to missing before the data series was subjected to homogeneity testing.

Like infilling other missing values, we also used the quality-controlled ANUSPLIN estimates (described in section 2c above) to infill the months of set-to-missing erroneous values. Since the ANUSPLIN-AdjMlyV1 surfaces were fitted to version 2016 of the AdjDlyRS dataset, which did not undergo the systematic QC of Cheng et al. (2023), all ANUSPLIN estimates that differ less than 2.0 mm from the original erroneous values are deemed unacceptable and hence not used to infill the missing values (i.e., remain missing with flag bAE). Further, set-to-missing erroneous values remain missing if an ANUSPLIN estimate is not available (mostly for the post-2015 period).

## 3. Data homogenization

The quality-controlled and gap-infilled version of the 425 joined monthly precipitation data series (i.e., the end results of procedure II in Fig. 2) are the raw data series to be tested for homogeneity below. Considering the nonnormality of monthly precipitation data and that log-transform is sufficient to diminish the nonnormality of such data, we applied homogeneity tests to log-transformed monthly precipitation (in mm) series, $\log(P_t + 0.1)$. The addition of the 0.1 mm in the transformation is to enable dealing with valid 0 values in the data series; this has no effect on homogeneity of the data series.

### a. *Changepoint detection procedure*

Changepoints are times of discontinuities in a data time series. Nonclimatic changepoints are those caused by nonclimatic factors, such as changes in observing techniques/equipment, locations, environment, and so on. Some of these changes are documented, while others are not. Thus, there are documented and undocumented changepoints, which are also called type-0 and type-1 changepoints. Nonclimatic changepoints must be eliminated from the data series, so that the data series can better present the true climate and changes therein. Nonclimatic changes can affect the mean or variance or the whole

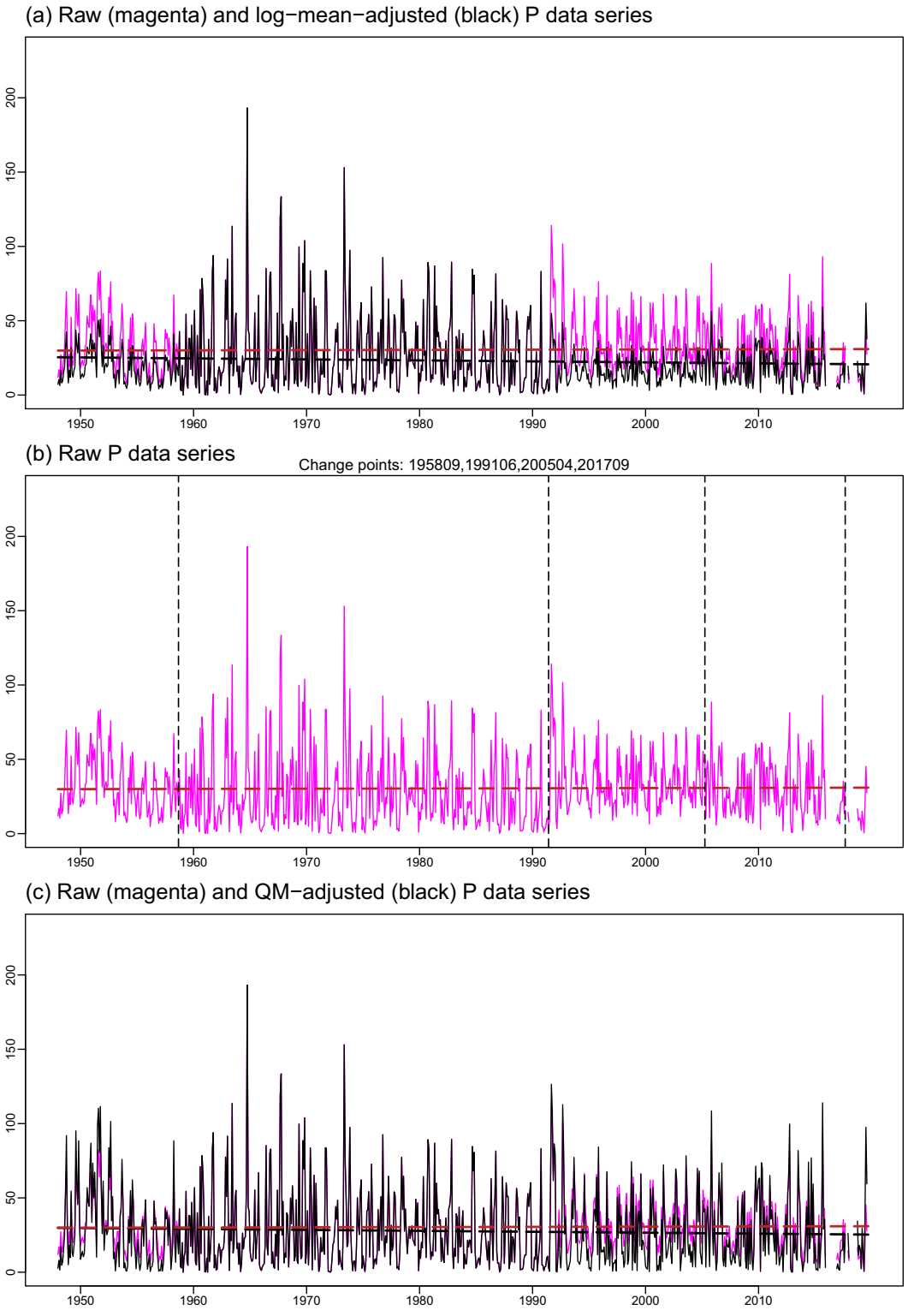


FIG. 3. The raw and log mean-adjusted or QM-adjusted monthly precipitation (*P*; in mm) data series for Stn361 (ID: 2400573). The vertical dashed lines mark the changepoints.

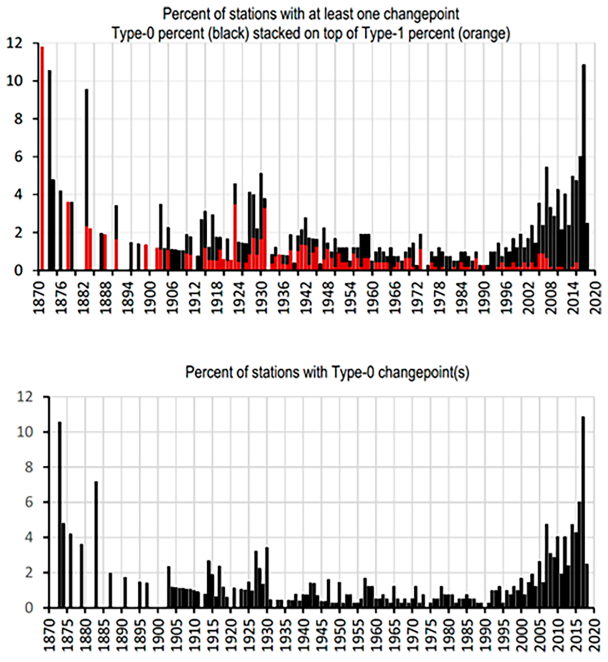


FIG. 4. The annual series of the percentage of stations (i.e., data series) with at least one changepoint of the indicated types (type 0 in orange, type 1 in black).

distribution. Most of homogeneity tests focus on testing sudden changes in the mean (e.g., Reeves et al. 2007; Wang et al. 2007; Wang 2008a,b; Wang and Feng 2013); only a few test for changes in variance (e.g., Dai et al. 2011). Nevertheless, if accompanied by a change in the mean, changes in the variance or distribution can also be eliminated by quantile matching adjustments (see section 3d below).

For a type-0 changepoint, since the time of change is known (documented), the regular $t$ or $F$ test can be used to determine the statistical significance of a sudden change in the mean (i.e., mean-shift). For a type-1 changepoint, the time of change is unknown and thus the most probably time of change needs to be identified using a maximal type of test, such as maximal $t$ or maximal $F$ tests (e.g., Wang 2008a). The RHtestsV4 package (Wang and Feng 2013) is the only software package that allows users to test for both types of changepoints in tandem, and thus was used in this study. This package includes the FindU.wRef and FindU functions for detecting type-1 changepoints with and without using a reference series, respectively, obtaining a list of type-1 changepoints. Then, a list of potential type-0 changepoints can be obtained by metadata investigation. The statistical significance of each changepoints in the combined list of type-1 and type-0 changepoints can be determined using the StepSize.wRef or StepSize function (depending on whether a reference series is used), which also estimates the magnitude of the mean shifts.

More specifically, we developed a semiautomatic procedure to detect significant changepoints by applying multiple homogeneity tests without a reference (noRef) and with reference series (wRef). We used the FindU.wRef, FindU, StepSize.wRef, and StepSize functions of the RHtestsV4 package (Wang and Feng 2013) to do the various homogeneity tests needed, with the level of significance being set to 1% (this high level was chosen to avoid overadjustments to the data). For each base series, we used up to six reference series in total: the ANUSPLIN estimates of monthly precipitation for the station in question, the 20CRv3 ensemble-mean monthly precipitation data series for the grid point closest to the station, as well as up to four best significantly-correlated (at 5% level) neighbor stations' data series (there could be fewer than four neighbor stations within 200-km radius that are qualified as reference; the correlation is calculated using the first difference series). This semiautomatic procedure is detailed in the online supplemental material SM1, with a shortened description below.

First, for each base series, we conducted homogeneity tests with each of the chosen reference series, separately, as well as a test without any reference series, to identify all possible type-1 changepoints (p1Cs; these are statistically significant without metadata support, see Wang and Feng 2013). We then investigated all available metadata, including but not limited to periods of missing observations and times of station joining, to produce a list of all potential type-0 changepoints (p0Cs; these are deemed documented changepoints). All the possible type-1 and potential type-0 changepoints were then pooled together and sorted, obtaining the list Pooled_p1Cs+p0Cs, which was then analyzed to produce the merged list Merged_p1Cs+p0Cs. The analysis includes defining initial groups of nearby changepoints (including finding the most likely one to represent the group), as well as replacing a type-1 changepoint with its nearest type-0 changepoint if they are no more than 5 months apart. For more details, readers are referred to the first four steps in SM1.

Second, we obtained a list of not insignificant changepoints by eliminating, one by one, all changepoints in the list Merged_p1Cs+p0Cs that are not significant in the base data series. For each not insignificant changepoint $c$, we redefined its group of nearby changepoints using StepSize and StepSize.wRef, subsequently, and the list Pooled_p1Cs+p0Cs, thereby obtaining the lists betaFCs.noRef and betaFCs.wRef, respectively, as detailed in steps 5 and 6 in SM1. The redefining of groups is necessary, because the initial definition could be inaccurate due to the existence of many other insignificant changepoints that were in the list for testing back then (which could notably bias the multiphase regression fit).

Third, we excluded all changepoints in the list betaFCs.noRef that do not meet the following duo-significance requirement. We fit the base series with this list of changepoints by calling StepSize and StepSize.wRef, subsequently, thereby obtaining one fit without a reference, and another fit with the reference series 20CRv3est (20CRv3est was chosen because it covers the whole data period of every base station). A changepoint could be significant in one fit but insignificant in the other fit, due to imperfectness of the reference/base series. Here, we retained only those changepoints that are significant in both fits (with a few exceptions), obtaining the new list of changepoints, 2sig_betaFCs.noRef. We repeated this procedure for the list betaFCs.wRef to obtain the new list 2sig_betaFCs.wRef, as detailed in step 7 in SM1. Since the 20CRv3est reference series ends at the end of 2015, any

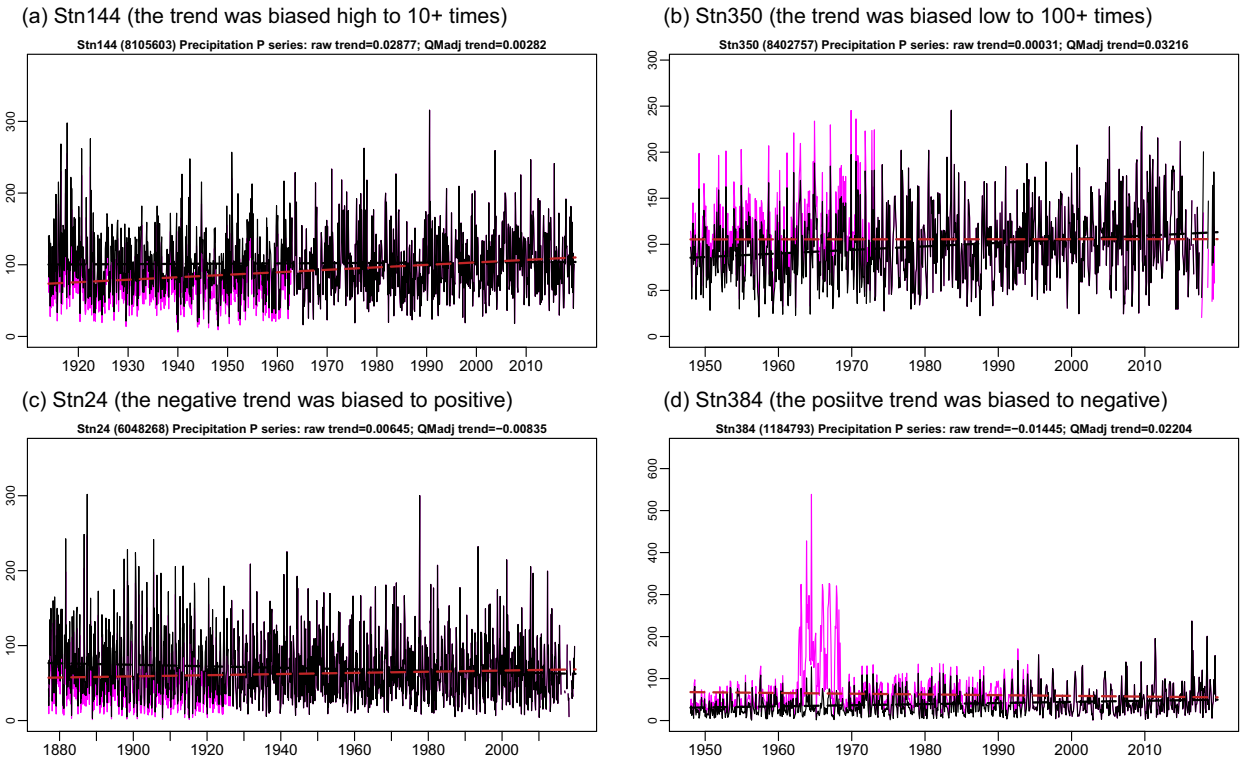


FIG. 5. Examples of inhomogeneity effects on trends of individual monthly precipitation data series. Raw data series are shown in magenta, with its trend in red. Adjusted data series and its trend are shown in black. The trend unit is mm month$^{-1}$. The 10+ and 100+ times are the ratios between the raw trend and the QMadj trend.

changepoints after year 2015 are retained without the duo-significance requirement (these are significant only in the noRef test). The duo-significance requirement would delete any changepoint that exists only in the reference series, not in the base series, because such changepoints will not be found significant in the fit without using the reference series. This is crucial here because the reference series 20CRv3est could also be inhomogeneous.

Fourth, as the final results from the automatic procedure, we chose from the two sets of changepoints, 2sig_betaFCs.noRef and 2sig_betaFCs.wRef, the set that fits the base series with the smaller sum of squared errors, obtaining the list AutoFinalCs.

We finalize the results by manual analysis using metadata and visually inspection of the regression fits with all changepoints in AutoFinalCs and of the associated trend maps, obtaining the final list of significant changepoints (ManualFinalCs) for each base data series. Whenever necessary, we revisited the list Pooled_p1Cs+p0Cs to see if the automated procedure above failed to retain the correct changepoint(s) or to find it at its physically most reasonable point of time. Some expert judgement was needed and used in this step.

Manual investigation is necessary for a few reasons. First, some changepoints in the list AutoFinalCs have "?" as significance (i.e., might be significant) because the significance is estimated to be within the 95% uncertainty range (Wang and Feng 2013). In this case, expert judgement based on investigation of the data series with the multiphase regression fit and trend map is needed to determine whether to keep these as significant changepoints. Wrong inclusion or exclusion of one or more changepoints could invalidate the results due to the nature of the multiphase regression fit (see SM1 and Fig. S1 for examples). Second, there is uncertainty in the metadata information used in this study. The exact dates of some of the changes were not recorded. Also, we treated the last valid data point before a data gap and the last point of a segment of ANUSPLIN estimates as potential type-0 changepoints, suspecting that some changes could have happened when the observations were resumed. Third, the statistical results (especially the significance estimate) are less reliable for very short data segments (usually less than 12 data points) because of the very small sample size; thus, some of such changepoints were excluded or included wrongly, which should be corrected by the manual investigation.

Moreover, it is important to point out that some station joinings should be included for adjustments even if it caused a mean shift that is of low or marginal significance, since this is a case of reliable cause for inhomogeneity; otherwise the trend of the data series could be biased to become inconsistent with its surrounding stations, since the effect is just enough to bias the trend but the mean shift does not stand out significantly (see SM1 for an example). Note that the majority of the station joinings in the last two decades are joinings of manual and automated stations, which are more likely to have systematic differences.

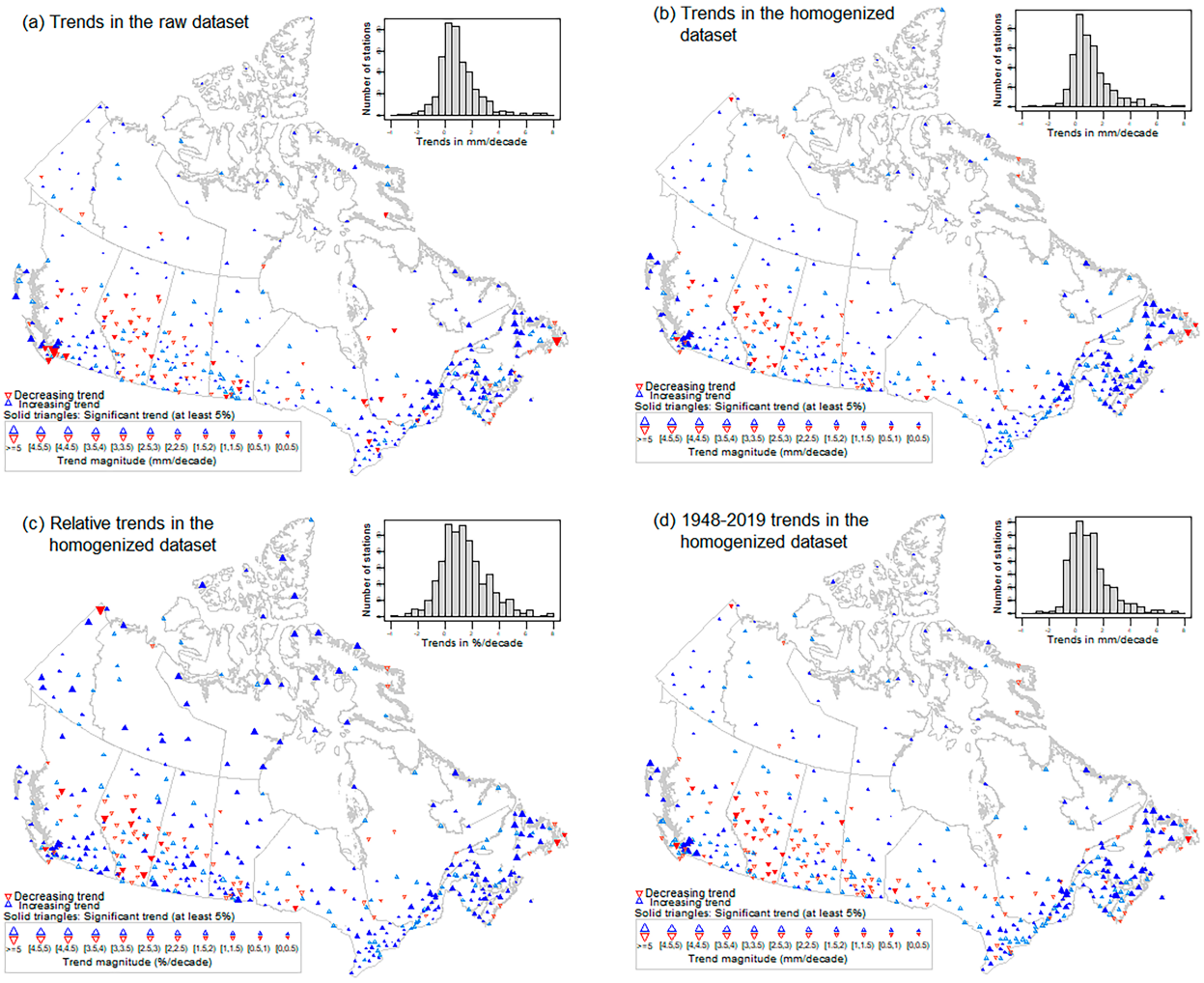


FIG. 6. Trends and relative trends in the indicated monthly precipitation dataset for (a)–(c) the entire data record period and (d) the 1948–2019 period.

However, as shown in supplemental Table S2, most of the modifications made during the manual analysis are due to having "?" as significance or questionable type-0 error (just because of missing data) or of too short segment (too small sample size to give a reliable estimate of significance); many of them are minor changes that improve the accuracy; the automated procedure gave unreasonable results only for a few station series, as would be expected as type-I and type-II errors of statistical tests at 1% significance level.

### b. *Adjustment procedure for eliminating data inhomogeneity*

First, it is of critical importance to point out that one shall not apply mean adjustments to the log-transformed data series and then convert the homogenized log-transformed data series back to obtain homogenized precipitation data series. This is because a mean adjustment ($\Delta$) to a log-transformed data series (referred to as a log mean adjustment) is a multiplier adjustment to the precipitation data series $P_t : P_t^a = e^{\log(P_t)+\Delta} = e^{\Delta}P_t$. This means that the variance of the adjusted precipitation series $P_t^a$ is changed by a factor of $e^{2\Delta}$, regardless of whether there is any variance change accompanying the mean shift. More specifically, the variance is reduced when $\Delta < 0$, and increased when $\Delta > 0$. Such a log-mean adjustment will make the adjusted precipitation data series much worse than the raw data series, creating artificial variance change(s), when the mean needs to be adjusted down while the variance needs to be adjusted up (or vice versa). For example, the negative log-mean adjustments notably reduced the variance of the adjusted precipitation series after June 1991, causing a notable variance decrease in the adjusted precipitation series (see the black line in Fig. 3a).

Second, although mean adjustments can be derived from the untransformed precipitation data series after changepoints are identified, these adjustments cannot be easily applied to the precipitation data series, unless you want to adjust the series to the segment of the highest long-term mean value, because subtracting an amount from the precipitation data could make the precipitation amount negative, which is unphysical.

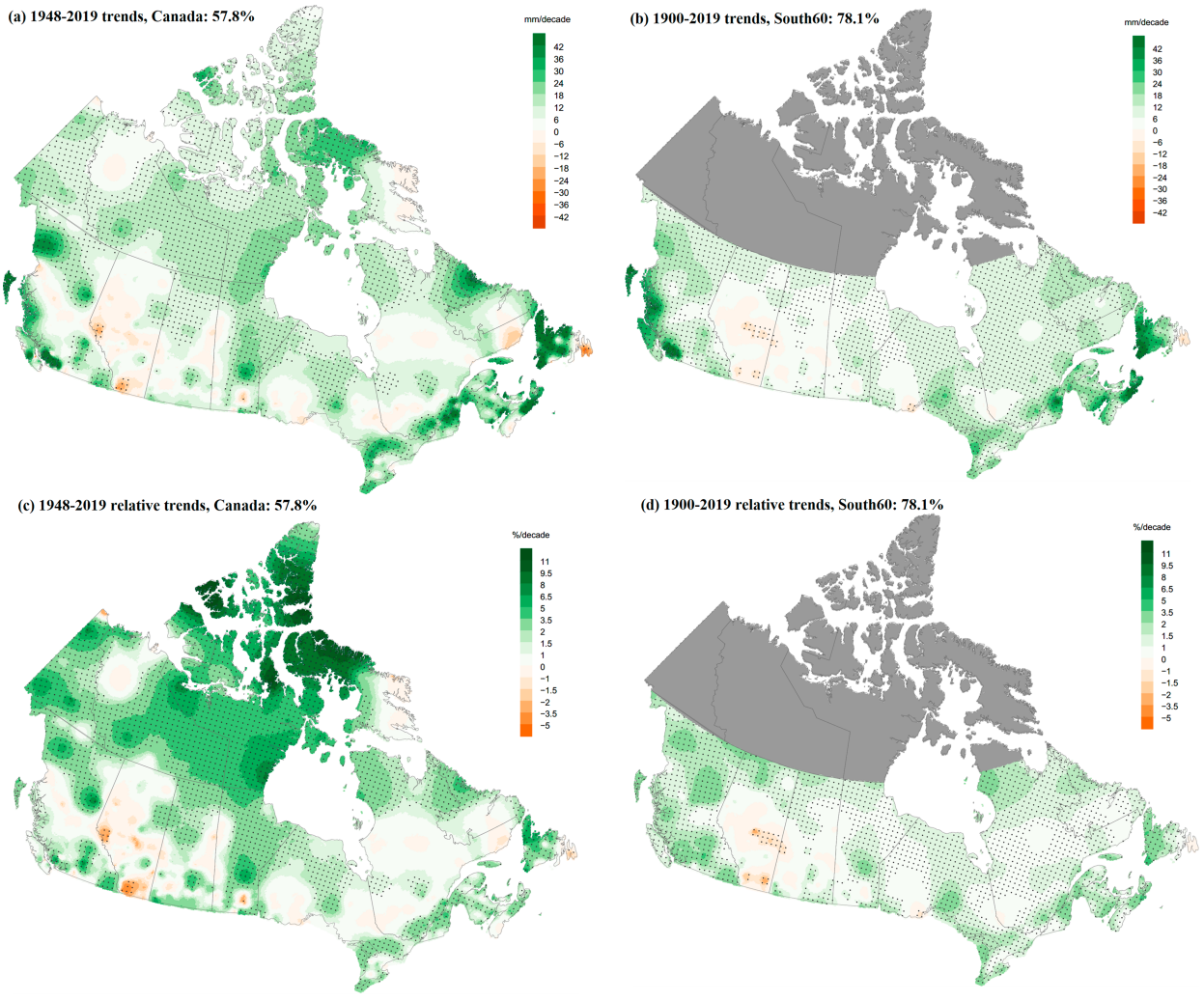


FIG. 7. Maps of trends in annual total precipitation. Areas of a significant trend (at the 5% level) are marked with black dots. The relative trends are the trends expressed in percentage of the corresponding 1961–90 mean values at the grid points. The percentage of grid points with a significant trend is also given in the title line.

Therefore, in this study, we applied quantile-matching (QM) adjustments (Wang et al. 2010; Vincent et al. 2012) to the log-transformed data series to homogenize it. Then, the homogenized log-transformed data series was converted back to obtain the homogenized precipitation data series. The QM adjustments aim to adjust the whole distribution, including but not limited to the mean and variance, as shown in Fig. 3c (black line).

We developed two procedures to estimate the QM adjustments. One is to use the StepSize function, which does not use any reference series (referred to as noRefQMadj). The other is to use StepSize.wRef with the Best-correlated Reference Homogeneous segment (BRHseg) if available, or to use StepSize if BRHseg is not available, for the changepoint in question. The second procedure is a dynamic QM-adjustment procedure (DynQMadj), in which some changepoints will be adjusted using BRHseg as reference, and others will be adjusted without using any reference; the BRHseg, if available, could be different for different changepoints in the same data series (see SM2 in the online supplemental material for the details of DynQMadj).

It is common practice to adjust the data to the most recent segment because this makes ongoing updating of the dataset easier. However, the precipitation data from the most recent period are of much lower quality and have more missing values due to the use of automated gauges with insufficient regular maintenance, which is a bigger disadvantage than making the ongoing updating of the dataset a little more complicated. Thus, we adjusted the data series to the longest observed data segment, i.e., the segment (i.e., homogeneous subperiod) that has the highest valid observation count (excluding estimated values), because this is the period of the most consistent observations (usually manual observations) of higher quality. If there are two or more such homogeneous segments, the most recent one is chosen. The number of categories used to estimate the cumulative distribution function is $Mq = \min[12, C_{\min}/5]$, where $C_{\min} \geq 5$ is the count of valid data in the shorter of the

TABLE 1. The top half shows trends in the regional mean series of seasonal or annual total precipitation over the indicated periods, as estimated from the CanKrig mlyPv1 dataset. The bottom half shows regional mean relative trends as derived by expressing the trends from the top portion in percentage of the 1961–90 mean of the corresponding data series. The 95% confidence interval of trend is shown in parentheses. Insignificant trends are shown in italic.

| Trends in physical units (mm decade$^{-1}$) | | | | | | |
|---|---|---|---|---|---|---|
| Period | Region | DJF total $P$ | MAM total $P$ | JJA total $P$ | SON total $P$ | Annual total $P$ |
| 1948–2019 | Canada | *0.70 (−0.60, 2.02)* | 3.18 (1.46, 4.89) | 3.86 (2.16, 5.49) | 4.20 (2.79, 5.72) | 11.26 (8.58, 14.08) |
| | North55 | 1.90 (0.85, 2.89) | 2.67 (1.11, 4.29) | 5.00 (2.75, 7.04) | 3.35 (1.76, 4.55) | 11.02 (3.49, 18.07) |
| | South55 | *−0.91 (−3.01, 1.21)* | 3.37 (1.15, 5.61) | 2.95 (0.24, 5.52) | 5.30 (2.18, 8.65) | 9.98 (4.25, 14.98) |
| 1900–2019 | South55 | 1.78 (0.68, 2.85) | 1.77 (0.86, 2.66) | 2.66 (1.41, 3.97) | 4.32 (2.97, 5.65) | 10.48 (8.01, 12.69) |

| Regional mean relative trends (% decade$^{-1}$) | | | | | | |
|---|---|---|---|---|---|---|
| Period | Region | DJF | MAM | JJA | SON | Annual |
| 1948–2019 | Canada | *0.54 (−0.46, 1.55)* | 2.61 (1.2, 4.01) | 1.94 (1.09, 2.76) | 2.34 (1.56, 3.2) | 1.79 (1.36, 2.23) |
| | North55 | 2.54 (1.13, 3.85) | 3.59 (1.49, 5.77) | 3.35 (1.84, 4.72) | 2.60 (1.36, 3.53) | 2.58 (0.82, 4.22) |
| | South55 | *−0.46 (−1.5, 0.6)* | 1.86 (0.63, 3.09) | 1.13 (0.09, 2.12) | 2.19 (0.9, 3.58) | 1.13 (0.48, 1.69) |
| 1900–2019 | South55 | 0.89 (0.34, 1.42) | 0.98 (0.48, 1.47) | 1.02 (0.54, 1.52) | 1.79 (1.23, 2.34) | 1.18 (0.91, 1.43) |

two data segments surrounding the changepoint in question. That is, Mq = 12 if $C_{min} \geq 60$; otherwise, Mq = $C_{min}/5 < 12$ (note that $C_{min} \geq 5$).

We carried out both noRefQMadj and DynQMadj and found that the results are mostly similar, with the DynQMadj being problematic for a few stations due to the relatively low station density and thus imperfect reference series. Thus, we chose to use the noRefQMadj results in this study, to avoid possible adverse effects of any imperfection in the reference series on the adjusted base data series. The 425 homogenized data series are assigned to the station ID of this longest observed data segment (i.e., the base segment). This dataset is referred to as version 1 of the Canadian Homogenized monthly Precipitation dataset (CanHoP mlyV1).

### c. Changepoint statistics and characteristics

While 127 out of the 425 monthly precipitation series were found to be homogenous, we have identified 647 artificial changepoints in the remaining 298 precipitation series, including 465 type-0 and 182 type-1 changepoints. Among the inhomogeneous series, 121 are single-changepoint series, and 84, 49, 21, and 15 are 2-, 3-, 4-, and 5-changepoint series, respectively, with 8 series having more than 5 (6–8) changepoints each. Among the 465 type-0 changepoints, 72.3% are due to data gaps, 23.4% are due to station joining, and 4.3% are due to changes in gauge type and/or observing frequency.

Figure 4 shows the annual series of the percentage of stations (i.e., data series) with at least one changepoint (regardless of types). The percentage is the lowest in the period from the 1960s to the end of the 1990s (mostly below 1%), and higher in the early period and recent decade (Fig. 4a). This is also true for the type-0 changepoints (Fig. 4b), for which the percentage is the highest (10.8%) in 2017, due to station automation and more missing values in the automated gauge data records. However, the percentage of stations with type-1 changepoints (orange bars in Fig. 4a) decreases over time, thanks to improvement in metadata collection and more missing data in the recent decade (missing data information was used as metadata in this study; see SM1 in the supplemental material). Note that the relatively higher numbers of type-0 changepoints in the 1915–30 arose from relatively more data gaps in this period: 93.3% of the type-0 changepoints (42 out of 45) are due to data gaps, and the remaining three are due to station joining.

## 4. Historical trends and the effects of data inhomogeneity on trend estimates

### a. Trend analysis method

After the precipitation data series is homogenized, we used the Mann–Kendall trend estimator (Mann 1945; Kendall 1955) to estimate trends. This estimator is a nonparametric estimator and is also less vulnerable to gross errors than the least squares estimator. Specifically, we used the Mann–Kendall trend estimation method of Wang and Swail (2001) to estimate linear trends in the consecutive monthly precipitation data series. This trend estimator accounts for the effects of lag-1 autocorrelation in the data series and was found to perform best in comparison with other trend calculation methods (Hartmann et al. 2013, p. 2SM-12). For each station separately, we applied this method to the deseasonalized series of raw and homogenized monthly precipitation data, that is, to the anomalies relative to the long-term mean annual cycle (one value for each calendar month). The deseasonalization makes the data closer to a normal distribution and lowers the autocorrelation, so that a better trend estimate can be obtained. The results are also compared with each other to show the effects of data inhomogeneity on trend results (i.e., procedure V in Fig. 2), as presented in the next subsection.

To estimate regional mean precipitation trends, we also used the kriging-based gridding scheme KGNGrA of Abbasnezhadi and Wang (2023, unpublished manuscript) to grid the homogenized precipitation dataset CanHoP mlyV1 (i.e., procedure IV in Fig. 2). KGNGrA is based on ordinary kriging, a

geostatistical spatial interpolation model, which uses a weighted average of a subset of neighboring point values to estimate a particular interpolation point value [see Wackernagel (1995) for a detailed description]. More specifically, ordinary kriging was used to grid the climate normal values (i.e., the 1961–90 means) and the corresponding relative anomalies, separately. Then, the gridded normal values and gridded relative anomalies were used to derive gridded total precipitation dataset, which is named as version 1 of Canadian Kriged monthly Precipitation (CanKrig mlyPv1) and was used to estimate precipitation trends at grid points and to derive regional mean precipitation series for estimating regional mean trends (presented in section 4c below).

For calculation of regional mean trends, we used the 55°N parallel to separate the north and south (North55 and South55), because this is the parallel of most clear-cut difference both between north and south station density (Fig. 1) and between north and south historical trends (shown later). However, for comparison purpose, we also derived the regional mean data series and trends for the regions north and south of 60°N parallel (North60 and South60).

### b. *Effects of data inhomogeneity on trend estimates*

Figure 5 shows a few examples of how data inhomogeneity affects the trend estimation for individual data series, which are results of procedure V in Fig. 2. The trend could be biased high or low (e.g., Figs. 5a,b). Data inhomogeneity could also result in wrong conclusions about the trend direction: it could turn a negative trend into a positive one, and vice versa (e.g., Figs. 5c,d).

Precipitation trends estimated from the raw and homogenized monthly datasets are compared in Figs. 6a and 6b. Clearly, we see better spatial consistency of trends in the homogenized dataset (CanHoP mlyv1) than in the raw dataset, especially in southern British Columbia and southeastern Canada (from southern Ontario to Newfoundland and Labrador). Note that, in Fig. 6, two hollow triangles of different directions do not indicate a significant inconsistency, because both insignificant trends are not statistically different from 0 (no trend).

### c. *Historical precipitation trends at individual stations*

Based on the homogenized dataset, as shown in Fig. 6b, a significant upward trend prevails in the Pacific Coast and northern and southeastern Canada, while insignificant trends of mixed directions prevail in central-south Canada (from Alberta to central Quebec); there are only a few significant downward trends, most of which are seen in Alberta and the easternmost corner of Canada. The largest increases (in physical units) are seen in southeastern Canada and the Pacific Coast (Fig. 6b). However, the largest relative increases are seen in the north (Fig. 6c). More detailed trend results are presented below.

### d. *Historical trends inferred from the new gridded precipitation data*

Figure 7 shows the maps of trends in annual total precipitation for the indicated periods and regions, as derived from the gridded homogenized dataset CanKrig mlyPv1. Not surprisingly, the trend patterns resemble the corresponding counterpart shown

TABLE 2. As in the bottom half of Table 1, but for the relative trends (% decade$^{-1}$) in the regional mean series of seasonal or annual total precipitation for the indicated periods and climate regions (as estimated from the CanKrig mlyPv1 dataset). B.C. is British Columbia.

| Period | Climate region | DJF | MAM | JJA | SON | ANN |
|---|---|---|---|---|---|---|
| 1948–2019 | Arctic mountains and fiords | *2.85 (−1.47, 6.88)* | 4.95 (1.25, 8.1) | 4.08 (1.54, 6.19) | *1.08 (−0.81, 3)* | 2.7 (1.24, 4.08) |
| | Arctic tundra | 6.08 (3.48, 8.96) | 6.89 (5.18, 8.69) | 3.7 (1.55, 5.65) | 3.78 (2.46, 5.13) | 3.99 (1.41, 6.68) |
| | Mackenzie district | 3.62 (1.13, 6.11) | 4.11 (1.82, 6.56) | 3.86 (1.93, 6.04) | 2.54 (0.91, 4.14) | 3.34 (2.22, 4.39) |
| | Yukon/North B.C. mountains | *1.81 (−0.3, 3.92)* | 2.45 (0.52, 4.82) | 3.56 (1.63, 5.25) | 2.93 (1.15, 5.06) | 2.76 (1.19, 4.07) |
| | Pacific coast | *−0.46 (−2.6, 1.47)* | 3.58 (1.05, 6.12) | *1.02 (−1.95, 3.76)* | *1.98 (−0.05, 3.74)* | 1.39 (0.32, 2.49) |
| | South B.C. mountains | *−1.79 (−4.09, 0.22)* | 4.72 (1.88, 7.76) | *1.12 (−1.77, 3.76)* | 2.85 (0.74, 5.1) | 1.35 (0.29, 2.47) |
| | Northwestern forest | *−0.21 (−1.96, 1.36)* | *1.92 (−0.05, 4.08)* | 2.03 (0.55, 3.53) | *1.6 (−0.51, 3.6)* | 1.37 (0.28, 2.39) |
| | Prairies | *−3.41 (−5.62, −1.03)* | *1.33 (−1.31, 3.89)* | *1.13 (−1.21, 3.41)* | *2.56 (−0.67, 5.99)* | *0.82 (−0.73, 2.43)* |
| | Northeastern forest | *−0.14 (−1.49, 1.37)* | 1.45 (0.19, 2.51) | *0.99 (−0.18, 2.1)* | 2.29 (0.94, 3.58) | 0.92 (0.2, 1.7) |
| | Great Lakes/St. Lawrence | *0.59 (−1.45, 2.45)* | *2 (−0.41, 4.59)* | 1.78 (0.19, 3.31) | 2.65 (0.42, 4.77) | 1.72 (0.64, 2.79) |
| | Atlantic Canada | *0.71 (−0.9, 2.15)* | *1.24 (−0.49, 2.97)* | 2.47 (0.39, 4.47) | 3.54 (1.17, 5.72) | 1.91 (0.88, 3.01) |
| 1900–2019 | Pacific Coast | 1.63 (0.5, 2.69) | 2.3 (1.13, 3.45) | *1.16 (−0.12, 2.39)* | 1.79 (0.86, 2.79) | 1.77 (1.08, 2.46) |
| | South B.C. mountains | 1.14 (0, 2.38) | 2.35 (1.06, 3.68) | 1.4 (0.29, 2.51) | 2.01 (0.88, 3.08) | 1.62 (0.97, 2.32) |
| | Northwestern forest | *0.74 (−0.08, 1.53)* | *0.54 (−0.38, 1.4)* | 1.3 (0.65, 1.97) | 1.09 (0.21, 2.05) | 0.99 (0.47, 1.47) |
| | Prairies | *−0.24 (−1.37, 0.78)* | *0.29 (−0.95, 1.52)* | *0.13 (−0.93, 1.38)* | *0.48 (−1.1, 2.06)* | *0.3 (−0.39, 0.98)* |
| | Northeastern forest | *0.62 (−0.05, 1.3)* | 0.77 (0.21, 1.33) | 1.02 (0.48, 1.51) | 1.88 (1.01, 2.72) | 1.05 (0.74, 1.35) |
| | Great Lakes/St. Lawrence | 1.03 (0.23, 1.88) | 1.3 (0.19, 2.48) | 1.33 (0.51, 2.1) | 1.95 (0.89, 3.01) | 1.34 (0.78, 1.93) |
| | Atlantic Canada | 1.35 (0.61, 2.07) | 1.14 (0.44, 1.86) | 1.9 (0.92, 2.82) | 2.58 (1.7, 3.44) | 1.81 (0.98, 2.61) |

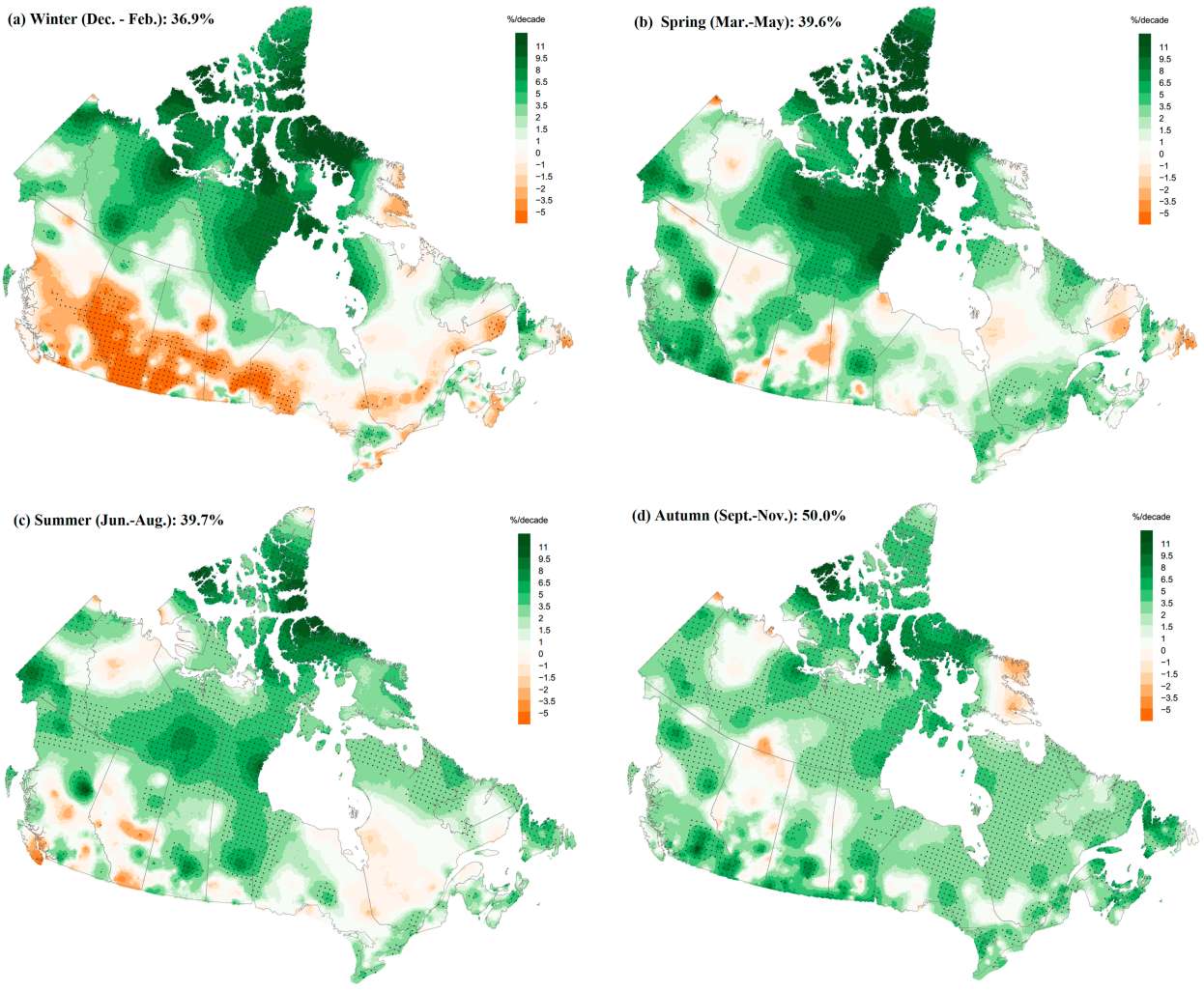


FIG. 8. As in Figs. 7c and 7d, but for the 1948–2019 relative trends in seasonal total precipitation in the indicated seasons.

in Fig. 6. For the period of 1948–2019 (Fig. 7c), 57.8% of the grid points were found to have a significant trend; 90.8% of the grid points have a positive trend, with largest relative increases being seen in the central Canadian Arctic. For the period of 1900–2019 (Fig. 7d), 78.1% of the grid points in South60 were found to have a significant trend; 93.6% of the grid points have a positive trend, with 84.3% of the increases in the South60 are less than 2% decade$^{-1}$. Note that the magnitudes of relative trends in Figs. 7c and 7d are much smaller than those shown in Vincent et al. (2015), although both studies show similar trend patterns (more comparison below).

Linear trends in the regional mean annual total precipitation series are presented in Table 1 (top). These trends (in physical units) are also expressed in percentage of the 1961–90 mean value of the data series in question, obtaining the relative trends presented in Table 2. Over the period 1948–2019, the relative annual trend is estimated to be 2.58%, 1.13%, and 1.79% decade$^{-1}$ for the North55, South55, and Canada, respectively; and over the period 1900–2019, it is 1.18% decade$^{-1}$ for the South55 (Table 1, bottom).

The patterns of seasonal relative trends are shown in Fig. 8 (see Fig. S3 for the trends in mm decade$^{-1}$). The winter trend pattern (Fig. 8a) is characterized by significant increases in the central Canadian Arctic, matched with extensive significant decreases in central-south Canada. The winter precipitation trend pattern agrees with the northward shifting of the storm track reported by Wang et al. (2006a) and Douville et al. (2021), and with the winter cyclone activity trend inferred from station pressure observations (Wang et al. 2006b). In the central Canadian Arctic, the relative increases are statistically significant in all four seasons but are largest in spring and winter, while the decreases in central-south Canada are statistically significant only in winter (Fig. 8). Extensively significant increases are also seen in BC and southern Ontario and Quebec in spring (Fig. 8b), and in southeastern Canada and British Columbia in autumn (Fig. 8d).

The regional mean trends also show some seasonality. In the north, spring has the largest relative increase among the four seasons, but in the south autumn has the largest relative increases. In both regions, winter has the smallest trend among

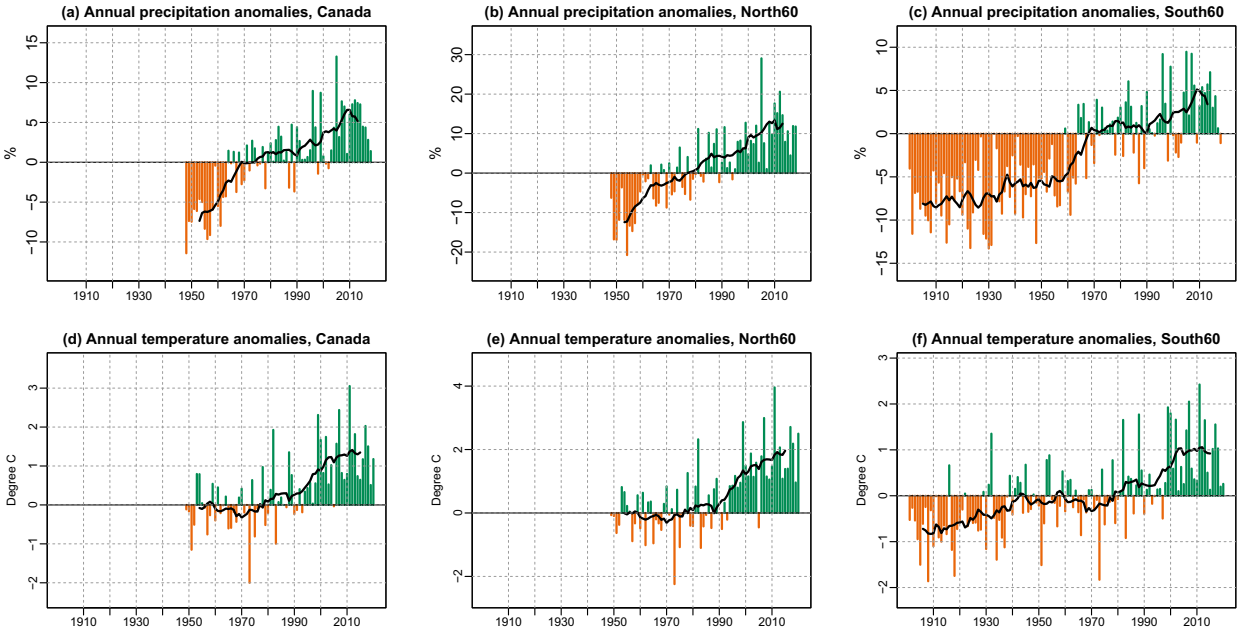


FIG. 9. Time series of anomalies of regional mean (a)–(c) annual total precipitation and (d)–(f) annual mean surface air temperature for the indicated regions. The anomalies are relative to the climate normal (1961–90 mean) of the regional mean data series; those in (a)–(c) are expressed in percent of the climate normal. The thick black line is an 11-yr running mean.

the four seasons. Over the period 1948–2019, the winter trend is statistically insignificant in the south and averaged over Canada (Table 1).

The regional mean trends for the 11 climate regions used in CTVB are also presented in Table 2, and the provincial/territorial mean trends in Table S3. A significant increase is seen in Arctic tundra and the Mackenzie district in all four seasons, with the largest relative increases in the Arctic tundra in winter and spring, while the Prairies experienced a significant decrease in winter but insignificant trends in all the other seasons (Table 2).

For comparison with the related previous studies (e.g., Vincent et al. 2015), time series of anomalies of regional mean annual total precipitation are shown for North60, South60, and Canada in Figs. 9a–c (and for those of seasonal total precipitation in Fig. S4). The anomalies are relative to, and expressed in percentage of, the climate normal (i.e., 1961–90 mean) of the regional mean precipitation. Averaged over Canada, there is a steady rapid increasing trend over the period from 1948 to early 2010s, with a higher rate of increase before 1970 (Fig. 9a, thick line). For the South60, the increasing trend is smaller before 1960 and after 1970, with a rapid increase in the 1960s (Fig. 9c, thick line). As shown in Fig. S6, a similar rapid increase is also seen in the corresponding anomaly series derived from the 20CRv3 ensemble-mean precipitation, and from the ANUSPLIN-gridded monthly precipitation data for North America (ANUSPLIN_NA) that are based on unadjusted and unhomogenized precipitation data from a mixture of manual and automated stations in Canada, the United States, and Mexico (MacDonald et al. 2020). The rapid increase in the 1960s (Fig. 9c) is also seen when the region is limited to the 40°–50°N latitude zone, which is dominated by stations in the United States (Fig. S6d). However, the large negative anomalies before 1960 and the rapid increase in the 1960s are not seen in the corresponding regional mean temperature anomalies (Figs. 9d–f).

As shown in Table 3, the old CanGRD-based estimate of national mean precipitation trend is about 64% and 45% higher than the corresponding new CanKrig-based estimate for the period of 1970–2012 and 1948–2012, respectively. Causes for the unphysically high values of the old estimates will be detailed later in the discussion in section 5.

Given the corresponding observed (surface air) temperature trend for Canada (Table 3), the new CanKrig-based estimate of the 1970–2012 and 1948–2012 precipitation scaling rate is 3.16% and 7.73% per 1°C of warming, respectively. While the 1970–2012 rate is close to the corresponding global rate (2%–3% increase per 1°C of warming; Douville et al. 2021), the 1948–2012 rate is more than double the global rate, which is due to the rapid precipitation increase in the 1960s that did not have a corresponding increase in temperature (see Fig. 9). The old CanGRD-based estimates are an 11.2% and 5.2% increase in annual precipitation per 1°C of warming for the period of 1970–2012 and 1948–2012, respectively; both are unphysically high due to data inhomogeneity issues and problems in the method used to estimate regional mean trends, as detailed below.

## 5. Summary and discussion

As the first objective of this study, a comprehensive semiautomatic data homogenization procedure was developed and used to produce a homogenized monthly precipitation dataset CanHoP mlyV1. First, data gaps in the 425 data series were

TABLE 3. Comparison of annual precipitation trends with corresponding annual temperature trends for the indicated regions over the indicated periods. The 95% confidence interval of the trend is shown in parentheses.

| Annual trends | Canada, period: 1970–2012 | Data source | Canada, period: 1948–2012 | Data source |
|---|---|---|---|---|
| Temperature trend (°C decade$^{-1}$) | 0.43 (0.14, 0.75) | CanGRD data from Vincent et al. (2015) | 0.26 (0.17, 0.35) | Vincent et al. (2015) (CanGRD) |
| New $P$ trend (% decade$^{-1}$) | 1.36 (0.17, 2.33) 3.16% per 1°C of warming | CanKrig mlyPv1 | 2.01 (1.53, 2.48) 7.73% per 1°C of warming | CanKrig mlyPv1 |
| Old $P$ trend (% decade$^{-1}$) | 2.23 (0.65, 3.79) 5.2% per 1°C of warming | CanGRD data from Vincent et al. (2015) | 2.92 (2.31, 3.38) 11.2% per 1°C of warming | Vincent et al. (2015) (CanGRD) |

infilled by ANUSPLIN estimates of monthly precipitation for the station in question; all the infilled data segments of 5 months or longer were tested to see if they are significantly different from the observed data segments (namely, if the estimates induced any temporal data inhomogeneity). After the gap infilling, the 425 data series have few missing values in the period of 1948–2015; many stations have nearly complete data records for the period from 1900 to the late 2010s, with some going back as far as the 1840s. Then, multiple homogeneity tests with and without a reference series were used repeatedly to identify changepoints, the results from which were then finalized by manual analysis using metadata and visual inspection of the multiphase regression fits. As a result, 298 out of the 425 data series were found to be inhomogeneous. These series were then homogenized using quantile matching adjustments, which adjusted the whole distribution of the data. The homogenized dataset shows better spatial consistency of trends than does the raw dataset, and thus was used to assess Canadian historical precipitation trends.

The homogenized monthly precipitation dataset shows a significant upward trend prevailing in the Pacific Coast and northern and southeastern Canada, with only a few significant downward trends in Alberta and the easternmost corner of Canada. The largest increases in physical units are seen in southeastern Canada and the Pacific Coast. However, the largest relative increases (in percent of the climate normal value) are seen in the north (Fig. 6). The largest trend difference between northern and southern Canada is seen in winter, in which significant increases in the north were matched with significant decreases in central-south Canada (Fig. 8a). The winter precipitation trend pattern agrees with the reported northward shifting of the storm track (Wang et al. 2006a; Douville et al. 2021) and the winter cyclone activity trend inferred from station pressure observations (Wang et al. 2006b).

The scaling rate for annual precipitation over Canada is estimated to be 3.16% increase per 1°C of warming for the period 1970–2012, and 7.73% for the period 1948–2012. While the 1970–2012 rate is close to the global rate (2%–3% per 1°C of warming; Douville et al. 2022), the 1948–2012 rate is higher than the rate expected for extreme daily precipitation (7% increase per 1°C of warming; Douville et al. 2022), although a lower rate is expected for annual precipitation. The high rate is due at least in part to the rapid precipitation increase in the 1960s that was accompanied by little change in temperature (see Fig. 9).

The trend magnitudes presented in this study are much smaller and physically more reasonable than those reported in previous studies (Vincent et al. 2015; Zhang et al. 2000). As shown in Table 3, the old estimates of annual precipitation scaling rate as derived from Vincent et al. (2015) are 11.2% and 5.2% per 1°C of warming for the period of 1970–2012 and 1948–2012, respectively. The large overestimation of trend was mainly due to two reasons: the problem in estimating relative trends from regional mean series of relative anomalies (Abbasnezhadi and Wang 2023, unpublished manuscript, hereafter AW23) and the use of unhomogenized precipitation data series. The use of regional mean series of gridded relative anomalies to estimate regional mean relative trends (i.e., the CanGRD approach) resulted in as much as 38% overestimation of the regional mean relative trend for northern Canada, in comparison to the relative trend estimated from the regional mean total precipitation series (AW23). Moreover, due to the low station density in the north, a bias in one single station's trend would spread out to affect a large area by the interpolation between stations far away from each other (examples shown in Fig. S5).

It is worth pointing out that it is difficult and time-consuming to develop a homogenized precipitation dataset suitable for climate research, application, and services. This implies the critical importance of the longevity of observing stations and of the temporal consistency of observations. As revealed in AW23, missing data, especially a temporally changing percentage of missing data, could significantly bias the trend estimates.

*Acknowledgments.* The authors wish to thank Tommy Jang and Rodney Chan for their help in data/metadata extraction, Kevin Lawrence (Natural Resources Canada) for extracting the ANUSPLIN monthly precipitation estimates from the ANUPLIN surfaces for us, and Kian Abbasnezhadi for his guidance on using his gridding codes. The authors are also grateful to Alex Cannon and Hui Wan for their helpful internal review comments on an earlier version of this paper. Note that Hong Xu is an independent contractor who had a contract with Environment and Climate Change Canada to work, under Dr. X. L. Wang's direction, on this research project, which was proposed and designed by X. L. Wang (the scientific authority of the contract).

*Data availability statement.* The source datasets used in this study are available in the Digital Archive of Canadian Climatological Data (https://climate.weather.gc.ca) and the Canada Open Data Portal (http://open.canada.ca/data/en/dataset/

d8616c52-a812-44ad-8754-7bcc0d8de305) for the Adjusted Daily Rain and Snowfall dataset version 2020. The new datasets produced in this study, CanHoP mlyV1 and CanKrig mlyPv1, can be found by clicking on Precipitation on this web page: https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data.html.

## REFERENCES


Cheng, V. Y. S., X. L. Wang, and Y. Feng, 2023: A quality control system for historical in situ precipitation data. *Atmos.–Ocean*, in press.

Dai, A., J. Wang, P. W. Thorne, D. E. Parker, L. Haimberger, and X. L. Wang, 2011: A new approach to homogenize daily radiosonde humidity data. *J. Climate*, **24**, 965–991, https://doi.org/10.1175/2010JCLI3816.1.

Douville, H., and Coauthors, 2021: Water cycle changes. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 1055–1210, https://doi.org/10.1017/9781009157896.010.

——, S. Qasmi, A. Ribes, and O. Bock, 2022: Global warming at near-constant tropospheric relative humidity is supported by observations. *Commun. Earth Environ.*, **3**, 237, https://doi.org/10.1038/s43247-022-00561-z.

Hartmann, D. L., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 159–254.

Hutchinson, M. F., and T. Xu, 2013: ANUSPLIN version 4.4 user guide. Australian National University, Fenner School of Environment and Society, 55 pp., https://fennerschool.anu.edu.au/files/anusplin44.pdf.

Kendall, M. G., 1955: *Rank Correlation Methods*. Charles Griffin, 196 pp.

MacDonald, H., D. W. McKenney, P. Papadopol, K. Lawrence, J. Pedlar, and M. F. Hutchinson, 2020: North American historical monthly spatial climate dataset, 1901–2016. *Sci. Data*, **7**, 411, https://doi.org/10.1038/s41597-020-00737-2.

——, ——, X. L. Wang, J. Pedlar, P. Papadopol, K. Lawrence, Y. Feng, and M. F. Hutchinson, 2021: Spatial models of adjusted precipitation for Canada at varying time scales. *J. Appl. Meteor. Climatol.*, **60**, 291–304, https://doi.org/10.1175/JAMC-D-20-0041.1.

Mann, H. B., 1945: Non-parametric tests against trend. *Econometrica*, **13**, 245–259, https://doi.org/10.2307/1907187.

Mekis, E., and W. D. Hogg, 1999: Rehabilitation and analysis of Canadian daily precipitation time series. *Atmos.-Ocean*, **37**, 53–85, https://doi.org/10.1080/07055900.1999.9649621.

——, and L. A. Vincent, 2011: An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmos.-Ocean*, **49**, 163–177, https://doi.org/10.1080/07055900.2011.583910.

Reeves, J., J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, 2007: A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteor. Climatol.*, **46**, 900–915, https://doi.org/10.1175/JAM2493.1.

Slivinski, L. C., and Coauthors, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the twentieth century reanalysis system. *Quart. J. Roy. Meteor. Soc.*, **145**, 2876–2908, https://doi.org/10.1002/qj.3598.

Vincent, L. A., X. L. Wang, E. J. Milewska, H. Wan, F. Yang, and V. Swail, 2012: A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis. *J. Geophys. Res.*, **117**, D18110, https://doi.org/10.1029/2012JD017859.

——, X. Zhang, R. D. Brown, Y. Feng, E. Mekis, E. J. Milewska, H. Wan, and X. L. Wang, 2015: Observed trends in Canada's climate and influence of low-frequency variability modes. *J. Climate*, **28**, 4545–4560, https://doi.org/10.1175/JCLI-D-14-00697.1.

Wackernagel, H., 1995: Ordinary kriging. *Multivariate Geostatistics: An Introduction with Applications*, Springer, 74–81, https://doi.org/10.1007/978-3-662-03098-1.

Wang, X. L., 2008a: Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal *t* or *F* test. *J. Appl. Meteor. Climatol.*, **47**, 2423–2444, https://doi.org/10.1175/2008JAMC1741.1.

——, 2008b: Penalized maximal *F* test for detecting undocumented mean-shift without trend change. *J. Atmos. Oceanic Technol.*, **25**, 368–384, https://doi.org/10.1175/2007JTECHA982.1.

——, and V. R. Swail, 2001: Changes of extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes. *J. Climate*, **14**, 2204–2221, https://doi.org/10.1175/1520-0442(2001)014<2204:COEWHI>2.0.CO;2.

——, and Y. Feng, 2013: RHtestsV4 user manual. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada, 28 pp., https://github.com/ECCC-CDAS.

——, V. R. Swail, and F. W. Zwiers, 2006a: Climatology and changes of extra-tropical cyclone activity: Comparison of ERA-40 with NCEP/NCAR reanalysis for 1958–2001. *J. Climate*, **19**, 3145–3166, https://doi.org/10.1175/JCLI3781.1.

——, H. Wan, and V. R. Swail, 2006b: Observed changes in cyclone activity in Canada and their relationships to major circulation regimes. *J. Climate*, **19**, 896–915, https://doi.org/10.1175/JCLI3664.1.

——, Q. H. Wen, and Y. Wu, 2007: Penalized maximal *t* test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol.*, **46**, 916–931, https://doi.org/10.1175/JAM2504.1.

——, H. Chen, Y. Wu, Y. Feng, and Q. Pu, 2010: New techniques for detection and adjustment of shifts in daily precipitation data series. *J. Appl. Meteor. Climatol.*, **49**, 2416–2436, https://doi.org/10.1175/2010JAMC2376.1.

——, H. Xu, B. Qian, Y. Feng, and E. Mekis, 2017: Adjusted daily rainfall and snowfall data for Canada. *Atmos.–Ocean*, **55**, 155–168, https://doi.org/10.1080/07055900.2017.1342163.

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