Improving the Representation of Historical Climate Precipitation Indices Using Optimal Interpolation Methods

Alexis Pérez Bello & Alain Mailhot

Submit your article to this journal ⤤

View related articles ⤤

View Crossmark data ⤤

Improving the Representation of Historical Climate Precipitation Indices Using Optimal Interpolation Methods

Alexis Pérez Bello* and Alain Mailhot

Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ETE), Québec, Quebec, Canada

ABSTRACT *Defining a reference climate for precipitation is an important requirement in the development of climate change scenarios to support climate adaptation strategies. It is also important for many hydrological and water resource applications. This, however, remains a challenge in regions that are poorly covered by meteorological stations, such as northern Canada or mountainous regions. Reanalyses may represent an interesting option to define a reference climate in such regions. However, these need to be validated and corrected for bias before they can be used. In this paper, two data assimilation methods, Optimal Interpolation (OI) and Ensemble Optimal interpolation (EnOI), were used to combine four reanalysis datasets with observations in order to improve the representation of various precipitation indices across Canada. A total of 986 meteorological stations with minimally 20-year precipitation records over the 30-year reference period (1980–2009) were used. Annual values of ten Climate Precipitations Indices (CPIs) were estimated for each available dataset and were then combined (reanalysis plus observations) using OI and EnOI. A cross-validation strategy was finally applied to assess the relative performance of these datasets. Results suggest that combining reanalysis and observations through OI or EnOI improves CPI estimates at sites where no recorded precipitation is available. The EnOI dataset outperformed OI applied to each reanalysis independently. An evaluation of the gridded interpolated observational dataset from Natural Resources Canada showed it should be used with considerable caution for extreme CPIs because it can underestimate annual maximum 1-day precipitation, as well as overestimate the annual number of wet days.*

RÉSUMÉ *[Traduit par la rédaction] La définition d'un climat de référence pour les précipitations est un élément important dans l'élaboration de scénarios de changements climatiques sur lesquels s'appuieront les stratégies d'adaptation à de tels changements. Cette référence est également importante pour de nombreuses applications en hydrologie et en gestion des eaux. Cependant, la référence est difficile à établir dans des régions où la couverture des stations météorologiques est lacunaire, comme dans le nord du Canada ou dans les régions montagneuses. Les réanalyses semblent être une option intéressante pour y parvenir. Elles doivent toutefois être validées et corrigées des biais avant de pouvoir être utilisées. Dans cet article, deux méthodes d'assimilation de données, l'interpolation optimale (IO) et l'interpolation optimale d'ensemble (IOE), ont servi à combiner quatre jeux de données de réanalyses avec observations afin d'améliorer la représentation de divers indices de précipitations au Canada. Au total, 986 stations météorologiques ayant des relevés de précipitations couvrant au moins 20 des 30 années de la période de référence (1980–2009) ont été utilisées. Les valeurs annuelles de dix indices de précipitations ont été estimées pour chaque jeu de données disponibles puis combinées (réanalyses plus observations) par IO et IOE. Enfin, une stratégie de validation croisée a été appliquée pour évaluer la performance de chacun des jeux de données en question. Selon les résultats, la combinaison de réanalyses et d'observations par IO ou IOE améliore les estimations d'indices de précipitations des sites où les précipitations ne sont pas enregistrées. Le jeu de données de l'IOE a surpassé l'IO appliquée à chaque réanalyse séparément. Une évaluation du jeu de données d'observation interpolées aux points de grille de Ressources naturelles Canada montre qu'il doit être utilisé avec circonspection en ce qui a trait aux indices de précipitations extrêmes, car les valeurs annuelles de précipitations maximales en un jour peuvent être sous-estimées et le nombre de jours de pluie par année surestimé.*

KEYWORDS optimal interpolation; ensemble optimal interpolation; climate precipitations indices; data assimilation; northern Canada

1 Introduction

Flooding, landslides, soil erosion, crop damages, and degradation of water quality are just a few examples of how extreme precipitation events may affect ecosystems, the economy, and human activities. These events have become more frequent and intense over many regions of the world during recent decades (e.g., North America and Europe) (Hartmann et al., 2013) and are expected to become even more frequent and intense in the future as the climate warms (Stocker et al., 2013). Characterizing baseline climatologies of precipitation-related variables, such as extreme events, is essential for many applications and sectors, for example, the design of water infrastructure and water resource applications (Fletcher et al., 2019). These baselines or reference climatologies are also essential for developing climate change scenarios (Hatzaki et al., 2010; Tebaldi et al., 2006; Westra et al., 2014).

Historical climate can be characterized through various indices, such as the Climate Precipitations Indices (CPIs) proposed by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Peterson et al., 2001; Zhang et al., 2011). These indices have been estimated directly from available observational records around the world (e.g., Frich et al., 2002). Gridded datasets of these indices have also been produced through interpolation of observational datasets (e.g., Alexander et al., 2006; Donat et al., 2013). This process is relatively straightforward in regions with temporally consistent, dense meteorological networks but represents a challenge in remote regions with low-density networks and/or short periods of record as is the case for mountainous and northern regions of Canada (Mekis & Vincent, 2011).

Reanalyses represent an interesting complement to conventional observational datasets because they provide a comprehensive and, in general, good spatiotemporal representation of many variables. Reanalyses consist of a background forecast model and data assimilation routine that combines available observational datasets with a model forecast to generate gridded datasets (Bosilovich et al., 2008). They have been widely used in different climate research applications (e.g., Alexeev et al., 2012; Lindsay et al., 2014; Zou et al., 2014). Extreme precipitation and temperature indices from reanalyses have also been used as references for the evaluation of climate models (e.g., Sillmann et al., 2013). Their consistency with global gridded observational datasets has been verified (Donat et al., 2014), showing improvements in the representation of the spatial patterns after 1979 when satellite data were included in the assimilation process. Despite their overall good spatiotemporal representation of many variables, some variables, for instance precipitation, are still not well represented in some regions, mainly because observational datasets for these variables have not been explicitly used in the assimilation process (Kobayashi et al., 2015; Rienecker et al., 2011). For instance, in evaluating daily climate indices over the northern parts of Canada, Diaconescu et al. (2018) recommended caution when using extreme daily precipitation indices from reanalyses as reference datasets for these regions because their performance was, in some cases, lower than Regional Climate Model simulations. Reanalyses, therefore, need to be "corrected" before they can be used. Their overall performance depends on the variable considered, the data assimilation routine, the assimilated observational datasets, and the forecast model (Lorenz & Kunstmann, 2012).

One option for improving the CPI representation in the historical climate is to use bias correction methods (e.g., linear-scaling factor methods and quantile mapping) (Alidoost et al., 2019). Another option is to combine reanalysis datasets with in situ observations not considered in the reanalysis assimilation process through data assimilation methods. These methods combine model fields with observational datasets to create an improved model dataset (Bertino et al., 2003; Kalnay, 2002). They can be incorporated into the modelling process through two different modes: an online mode, which increases the quality of the initial conditions sequentially for new model simulations, and an off-line mode, which is used to create the best estimation, combining the final outputs of the model with observational data (Candiani et al., 2013; Matsikaris et al., 2015).

In this study, the off-line mode was applied using the Optimal Interpolation (OI) (Daley, 1991; Kalnay, 2002; Sen, 2009) and Ensemble Optimal Interpolation (EnOI) methods (Evensen, 2003; Oke et al., 2010). The OI method was applied to each reanalysis independently and EnOI to all reanalyses combined. Both methods are quite similar, their main difference being in how the background error covariance is defined and estimated (Ren & Hartnett, 2017). These methods were used because they are relatively straightforward, computationally inexpensive, and take the uncertainty of the various datasets into account. The OI method has been used to merge precipitation datasets in various regions of the world (e.g., Häggmark et al., 2000; Mahfouf et al., 2007; Soci et al., 2016). Furthermore, significant improvements were obtained when merging satellite information with observations and model predictions at a global scale (Nie et al., 2016). This method has also been used for daily retrospective estimation of streamflow at ungauged river sections (Lachance-Cloutier et al., 2017).

The objective of the paper is to improve historical CPI datasets for Canada combining observations and reanalyses using the OI and EnOI methods. The paper is organized as follows. Section 2 describes the observational datasets and reanalyses considered in this study. The CPI, the data assimilation techniques (OI and EnOI), and the cross-validation procedure are briefly described in Sections 3, 4, and 5, respectively. Section 6 presents the results followed by a summary and discussion (Section 7).

2 Datasets

Five gridded datasets were considered in this study (Table 1). The first four datasets are widely used reanalyses while the fifth dataset is the interpolated daily precipitation dataset from Natural Resources Canada (NRCan) generated from

TABLE 1. Gridded datasets used in this study (acronyms appearing in parentheses are used when referring to these datasets).

Gridded datasets	Period	Spatial Resolution	Temporal Resolution	Reference	Precipitation Assimilated
Climate Forecast System Reanalysis (CFSR)	1979–2009	0.312º × 0.312º	Hourly	Saha et al. (2010)	2.5º pentad dataset of CMAP and 0.5º land-base of CPC
ERA5 reanalysis (ERA5)	1979 onwards	0.25º × 0.25º	Hourly	Hersbach et al. (2019)	Satellites and ground-based radar (2009 onwards)
Japanese 55-year Reanalysis (JRA55)	1958–2013	0.56º × 0.56º	3 h	Kobayashi et al. (2015)	Not assimilated
Modern-Era Retrospective Analysis for Research and Applications (MERRA)	1979–2012	0.50º × 0.66º	Hourly	Rienecker et al. (2011)	Not assimilated
Precipitation data from Natural Resources Canada (NRCan)	1950–2012	0.083º × 0.083º	Daily	McKenney et al. (2011)	Observations

station records using a thin plate spline smoothing algorithm (ANUSPLIN) (Hutchinson et al., 2009; McKenney et al., 2011). It was used for comparison purposes and to estimate missing values in station records (further details are provided in Section 4).

Reanalyses differ in spatiotemporal resolutions, period covered, assimilation method, and assimilated datasets. Of the four reanalyses, the Climate Forecast System Reanalysis (CFSR) is the only reanalysis that directly assimilates precipitation datasets derived from ground-based meteorological records. Two sets of global precipitation data are indeed assimilated in CFSR. The first is the pentad dataset of the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP), which is a 5-day mean precipitation global dataset with a spatial resolution of 2.5 degrees and combines rain-gauge and satellite observations. The second dataset is the CPC global land-based daily dataset with a spatial resolution of 0.5 degrees obtained by interpolating approximately 30,000 rain-gauge stations worldwide (Saha et al., 2010). Precipitation from satellite data and rates from ground-based radar observations (2009 onwards) are assimilated by the fifth major global reanalysis produced by the European Centre for Medium-range Weather Forecasts (ERA5) (Hersbach et al., 2019). Finally, The Japanese 55-year Reanalysis (JRA-55)

and the Modern-Era Retrospective Analysis for Research and Applications (MERRA) do not assimilate precipitation-related variables (Kobayashi et al., 2015; Rienecker et al., 2011).

The NRCan dataset has fairly good spatial resolution (10 km grid) and covers Canada (the study domain) over a relatively long period (1950–2012). Previous studies showed an overall underestimation of precipitation by this dataset (e.g., Bajamgnigni Gbambie et al., 2017; Diaconescu et al., 2018; Hutchinson et al., 2009). The observations used in NRCan spatial models were not corrected for evaporation loss or undercatch (Hutchinson et al., 2009), which may partly explain these underestimations. To our knowledge, it is the only available dataset with such spatial and temporal resolutions that entirely covers Canada. Some other gridded datasets exist but they cover short time periods (e.g., Canadian Precipitation Analysis (CaPA); Lespinas et al., 2015; Mahfouf et al., 2007) or parts of Canada (e.g., dataset from the Ministère de l'Environnement et de la Lutte Contre les Changements Climatiques (MELCC) covering the province of Quebec; see Bajamgnigni Gbambie et al., 2017). Because each dataset covers a different period, a common period, 1980–2009, was considered in the following.

Daily precipitation records from the Adjusted and Homogenized Canadian Climate Data (AHCCD) developed by Environment and Climate Change Canada (Mekis & Vincent, 2011) were used, as well as from the stations operated by the MELCC in Quebec. The records from the AHCCD were corrected for evaporation loss and wind undercatch (Mekis & Vincent, 2011). Only stations with at least 20 valid years during the 30-year period 1980–2009 were considered for a total of 986 stations across Canada. A valid year was defined as a year when less than 10% of the daily values were missing. This criterion was applied as a compromise between more strict conditions (e.g., a smaller percentage of missing data for valid years or a larger number of valid years over the 1980–2009 period) that would result in a much smaller number of stations and less strict conditions that would raise concerns about data representativity. Figure 1 shows the map locating these stations. It clearly shows that stations are unevenly distributed and mainly concentrated in the southern parts of the country.

It is important to note that not all the stations considered in this study were used to create the NRCan dataset (for instance, some stations operated by the MELCC in Quebec were not
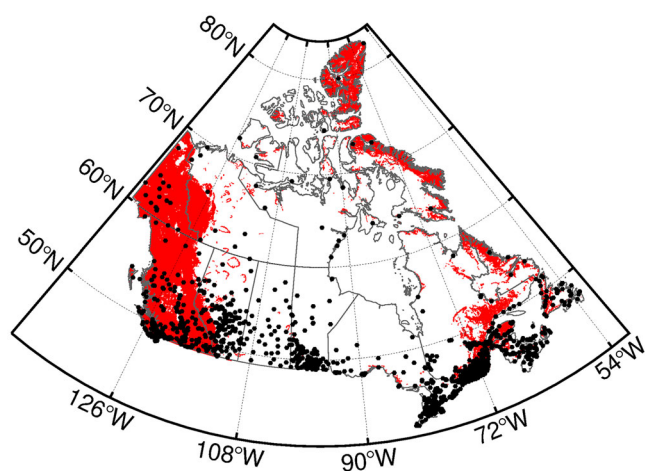

Fig. 1 Map of the 986 stations considered in this study. Regions in red correspond to mountainous regions according to the K3 classification of Karagulle et al. (2017).

used). Also, it is well known that interpolated precipitation, especially extreme values, at daily time scales can be highly uncertain in regions with sparse network coverage (Gervais et al., 2014).

3 Climate precipitation indices and super-observations

The CPI defined by the ETCCDI, as well as mean annual precipitation (PRmean), have been considered in this study (the list of indices and their meanings can be found in Table 2). Annual CPI values were first estimated, as well as the climatological mean, over the 1980–2009 period. The CPI were computed at each native grid point of each dataset (reanalysis or NRCan dataset). A second-order conservative remapping (Jones, 1999) was then applied to each CPI gridded dataset created from CFSR, JRA55, and MERRA using the ERA5 grid as the reference grid. The OI and EnOI methods were applied once the CPI gridded datasets from each reanalysis were created on the common grid.

As shown in Fig. 1, the number of stations in southern Canada (below 60°N) is relatively high (938 stations for 95% of the total number of stations) compared with northern regions. Station density in southern Canada is approximately one station per 80 km x 80 km, while for northern regions it is one station per 290 km x 290 km. Therefore, in southern regions, observations from neighbouring stations can be highly correlated at daily time scales. Furthermore, reanalyses may not properly represent small-scale processes (Vihma et al., 2014). Super-observations (Daley, 1991), also called upscaled observations (Cherubini et al., 2002), were, therefore, created using the reanalysis grid with the highest resolution as a reference grid (i.e., ERA5) in order to guarantee the numerical stability of the OI and EnOI methods. Super-observation values were estimated by combining station values within the same grid cell using the following expression:

$$\bar{x} = \frac{\sum_{i=1}^{n} D_i \gamma_i}{\sum_{i=1}^{n} D_i}, \qquad (1)$$

where n is the total number of stations inside a given grid cell, D_i is the number of valid days for station i for a specific year, and γ_i

TABLE 2. Climate Precipitation Indices (adapted from Zhang et al., 2011).

| Index	Description	Units
PRmean	Annual mean of daily precipitation	mm
PR1mm	Number of wet days (daily precipitation ≥ 1 mm)	days
RX1day	Annual maximum 1-day precipitation amount	mm
RX5day	Annual maximum 5-day precipitation amount	mm
SDII	Simple daily intensity index (ratio of annual total precipitation to the number of wet days ≥ 1 mm)	mm d^{-1}
R10mm	Annual number of days with precipitation ≥ 10 mm	days
R20mm	Annual number of days with precipitation ≥ 20 mm	days
CDD	Annual maximum number of consecutive days with precipitation below 1 mm	days
CWD	Annual maximum number of consecutive days with precipitation ≥ 1 mm	days
PRCPTOT	Annual total precipitation from days ≥ 1 mm	mm
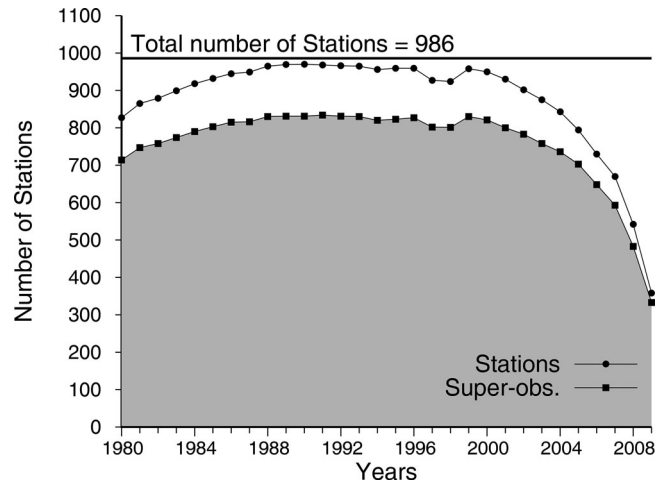
Fig. 2 Total number of stations and super-observations available over the 1980–2009 period.

represents the annual CPI value for that year. Longitudes and latitudes for the super-observations were also estimated using a similar equation. The super-station is superposed on an available station when there is only one station within a grid cell. In the construction of super-station values, no correction was added to account for possible differences in elevation among stations located within a grid cell.

Figure 2 shows the number of stations and corresponding super-observations with valid records for each year of the 1980–2009 period. It shows that the number of stations stay almost constant over the first 20 years but quickly decreased from 2000 to 2009. It should be noted that the spatial distribution of the super-observations changes from year to year as the number and location of available stations changes. These super-observation networks were used to apply OI and EnOI for each year independently.

4 Optimal interpolation methods

a Optimal Interpolation

This section provides a brief description of the OI method and how it was applied in this paper. A more general and comprehensive description of this method can be found in Daley (1991), Kalnay (2002), and Sen (2009).

The OI method, also known as statistical interpolation, is widely used (e.g., Lespinas et al., 2015; Mahfouf et al., 2007) to create an analysis field by combining a background state, or model forecast, with an observation network (Daley, 1991). The analysis value $\psi_A^{(j)}$ at grid point j is estimated by the background field $\psi_B^{(j)}$, the linear combination of the differences between the observations $\psi_O^{(i)}$ ($i = 1, \ldots, n$), where n is the number of observations close to this grid point and the background values at the observation site $\psi_B^{(i)}$. The so-called analysis equation is defined as,

$$\psi_A^{(j)} = \psi_B^{(j)} + \sum_{i=1}^{n} \mathbf{W}^{(j,i)}(\psi_O^{(i)} - \psi_B^{(i)}). \qquad (2)$$

Assuming that the background error depends only on distance (so-called homogeneous conditions), is isotropic, and observation errors are uncorrelated, the interpolation weights matrix, $\mathbf{W}$, can then be estimated using the following equation (Daley, 1991):

$$[\mathbf{B}_{EC} + \mu\,\mathbf{I}]\,\mathbf{W} = \mathbf{b}_{EC}, \tag{3}$$

in which $\mathbf{B}_{EC}$ is the background error correlation matrix between the station values, $\mathbf{b}_{EC}$ is the background error correlation vector between the observations and the background field at the corresponding grid point, $\mathbf{I}$ is the identity matrix and $\mu = O_{EV}/B_{EV}$, where O_{EV} and B_{EV}, are the observation and background error variances respectively. Parameter μ can be adjusted such that the weights granted to observations can be increased or reduced (Kalnay, 2002). For each index and reanalysis, O_{EV} and B_{EV}, respectively, were estimated according to the methodology described in Section 4.3 of Daley (1991).

The available CPI annual series over the 30-year reference period were used to estimate the spatial dependence of the correlation function $\mathbf{B}_{EC}(x)$ in Eq. (3). The values at each observation point from the background field were obtained for each year by means of bilinear interpolation, and the missing values in the observational dataset were filled out with the CPI calculated from the NRCan dataset. Then, correlations $\mathbf{R}_{ij}$ of the differences between the observations and background field at sites i and j were calculated using the following equation (Daley, 1991; Sen, 2009):

$$\mathbf{R}_{ij} = \frac{\overline{(\mathbf{O}_i - \mathbf{B}_i)(\mathbf{O}_j - \mathbf{B}_j)}}{\sqrt{\overline{(\mathbf{O}_i - \mathbf{B}_i)^2}\ \ \overline{(\mathbf{O}_j - \mathbf{B}_j)^2}}}, \tag{4}$$

where the overlined variables correspond to the average values over the reference period. The mean correlations over each 30 km distance interval were then computed. Many different spatial correlation functions $R(x)$ were tested and the following expression was finally selected and adjusted to the estimated values (Sen, 2009):

$$R(x) = a + b\ e^{(-x/c)}, \tag{5}$$

where x is the distance between stations and a, b, and c are the fitting parameters of the correlation functions. These parameters were estimated for each CPI and each reanalysis (not shown) and PR1mm was the most correlated CPI (highest values of c for all datasets) while RX5day and RX1day are the least correlated indices.

Figure 3 shows an example of the spatial correlation for the RX1day index estimated from CFSR with the corresponding fit of Eq. (5). Because of the averaging process over 30 km intervals, Figure 3 shows that at zero distance correlation has to be estimated. Using Eq. (5) and setting $x = 0$, the isotropic background error correlation $\mathbf{B}_{EC}$ used
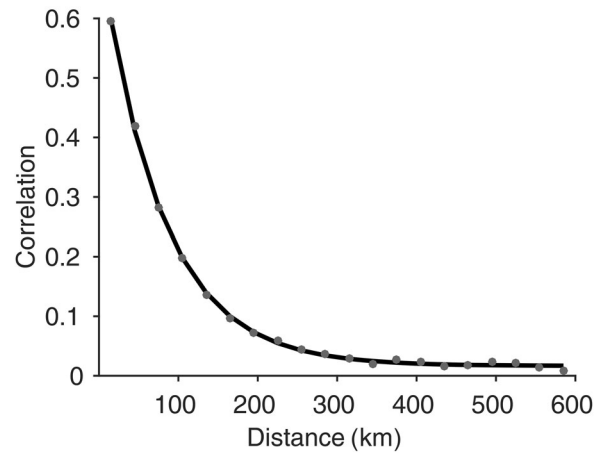

Fig. 3 Estimated correlation function (continuous curve) with corresponding discrete values for RX1day estimated from CFSR.

in Eq. (3) is given by (Sen, 2009):

$$\mathbf{B}_{EC}(x) = \frac{R(x)}{R(0)} \quad \text{or} \quad \mathbf{B}_{EC}(x) = \frac{a + b\,e^{(-x/c)}}{a + b}. \tag{6}$$

The relative expected analysis error ε for each grid point can be estimated by the following equation (Daley, 1991; Kalnay, 2002; Sen, 2009):

$$\varepsilon = 1 - \sum_{i=1}^{n} \mathbf{W}^{(i)}\,\mathbf{b}_{EC}^{(i)}, \tag{7}$$

where values of ε close to 0 means there is a larger contribution from neighbouring observations in the estimation of the CPI value at this specific grid point, while values of ε close to 1 means there is a larger contribution from the background field.

b *Ensemble Optimal Interpolation*

This section provides a brief description of the EnOI method and how it was applied. A more general and comprehensive description can be found in Evensen (2003) and Oke et al. (2010).

The EnOI is an approximation of the ensemble Kalman filter (EnKF) in which an ensemble of model states in long time integration (stationary) is used to compute the analysis in the space spanned and only one model state is used as a background field (Oke et al., 2007). Defining a matrix $\mathbf{E}$ with ensemble members:

$$\mathbf{E} = [M_1, M_2, \ldots, M_n], \tag{8}$$

where n is the number of ensemble members that are stored (vector form) in each column of $\mathbf{E}$. For each CPI, the ensemble members are the 30 years of each reanalysis combined ($n = 120$ members). The anomaly or ensemble perturbation matrix $\mathbf{E}_n$ can be defined by:

$$\mathbf{E}_n = \mathbf{E} - \mathbf{E}_m, \tag{9}$$

where $\mathbf{E}_m$ is the ensemble mean matrix in which each column has the same ensemble mean vector calculated from all the ensemble members stored in $\mathbf{E}$. Then, the ensemble covariance matrix $\mathbf{E}_{CM}$ can be estimated by

$$\mathbf{E}_{CM} = \frac{\mathbf{C}_F \circ \alpha\, \mathbf{E}_n\, \mathbf{E}_n^T}{n-1}, \qquad (10)$$

where the superscript T means a matrix transpose and the parameter $\alpha \in (0,1]$ is a scaling parameter of the model anomalies. This parameter is introduced to reduce the variance produced by the ensemble over a long time period that can overestimate the instantaneous background field error variance (Evensen, 2003; Fu et al., 2011). In our case, the value $\alpha = 1$ was used. The correlation function matrix $\mathbf{C}_F$ is used to implement the localization (each analysis grid point is calculated using only the nearest stations; Evensen, 2003), and "$\circ$" represents the Hadamard product that multiplies two matrices of the same dimensions element by element (Oke et al., 2007).

The EnOI analysis field was estimated using Eq. (2) described in Section 4.a and the following weight matrix $\mathbf{W}$ (Oke et al., 2007):

$$\mathbf{W} = \mathbf{E}_{CM}\, \mathbf{H}^T [\mathbf{H}\, \mathbf{E}_{CM}\, \mathbf{H}^T + \mathbf{R}]^{-1}, \qquad (11)$$

where $\mathbf{H}$ is an operator that interpolates model grid points at observation sites (bilinear interpolation), and $\mathbf{R}$ is the observation error covariance matrix. Combining Eqs (10) and (11), $\mathbf{W}$ can finally be expressed as

$$\begin{aligned} \mathbf{W} = \mathbf{C}_F \\ \circ \mathbf{E}_n\, \alpha\, (\mathbf{HE}_n)^T [\mathbf{C}_F \circ \mathbf{HE}_n\, \alpha\, (\mathbf{HE}_n)^T + (n-1)\, \mathbf{R}]^{-1}. \end{aligned} \qquad (12)$$

The background field used in the EnOI was the ensemble mean over the four reanalyses for each year. The observation errors were assumed to be uncorrelated as in Section 4.a; therefore, the observation error covariance matrix $\mathbf{R}$ was diagonal with values equal to the observation error variance O_{EV}. The correlation function was the same as Eq. (6), and the parameters a, b, and c for this equation were the mean parameters over the four reanalyses. The ensemble O_{EV} (defined as EO_{EV}) was estimated using the following equation:

$$EO_{EV} = \frac{1}{4} \sum_{i=1}^{k} \frac{O_{EV(i)}}{k}, \qquad (13)$$

where k is the number of reanalyses (in our case four). The "four factor" preceding the sum is used for convenience to provide more weight to the observations in the analysis equation. The O_{EV} differs from one reanalysis to another because it was estimated from the relationship between the observations and the background values at the observation sites (see Section 4 and Section 4.3 of Daley (1991). Elevation was not used in the analysis equation to estimate the CPI.

Regions with abrupt elevation changes could, therefore, be incorrectly represented. A possible solution could be to use only observational sites with absolute differences to grid average elevation less than 400 m (Brasnett, 1999). However, it would considerably reduce the number of observation sites used in this study.

5 Cross-validation procedure

In order to evaluate the performance of the OI and EnOI methods and check how the quality of the CPI improved at sites without observations, the sub-network cross-validation approach proposed by Panthou et al. (2012) was used. Available stations are first subdivided into two sub-networks: the calibration and validation sub-networks.

Validation sub-networks were generated by eliminating stations at varying distances in order to control the average distance between sites of the calibration and validation sub-networks. Two types of calibration and validation sub-networks were defined. Near sub-networks were designed to assess the performance close to the calibration points and were created according to the following steps: (1) the distances between all pairs of stations are first calculated; (2) the pair of closest stations is selected; (3) considering these two stations, the one with the smallest distance to the remaining stations is put in the validation sub-network; and (4) steps (1), (2), and (3) were then repeated using the remaining stations until the percentage of stations in the validation sub-network is reached. Far sub-networks were designed to assess the performance of the datasets at distant locations from the calibration stations. Far sub-network creation starts with step (1) as previously explained. Closest neighbouring stations are then paired. The station with the largest distance from the neighbouring station is then selected and removed from the calibration sub-network and put in the validation sub-network. Step 4 (mentioned above) was then applied. Different validation sub-networks including 10%, 30%, and 50% of the stations were considered for the Near and Far configurations.

Figure 4 presents violin plots (Hintze & Nelson, 1998) and boxplots of the distributions of distances between all pairs of validation sites and their closest calibration sites for Near and Far sub-networks respectively. The 10% Near sub-network corresponds to the configuration with the smallest distances between calibration and validation sites while the Far configuration has the largest distances. These distances increase as the fraction of validation sites increases for the Near sub-networks, while they decrease for the Far sub-networks. Furthermore, these distances decrease rapidly for Far configurations as the fraction of validation sites increases, which is consistent with the spatial distribution of stations shown in Fig. 1, where the network density is much higher in the southern part of Canada, and only a small fraction of stations covers the northern regions.

The distance between each land-based grid point over the study area and the closest station is presented in Fig. 5a. All
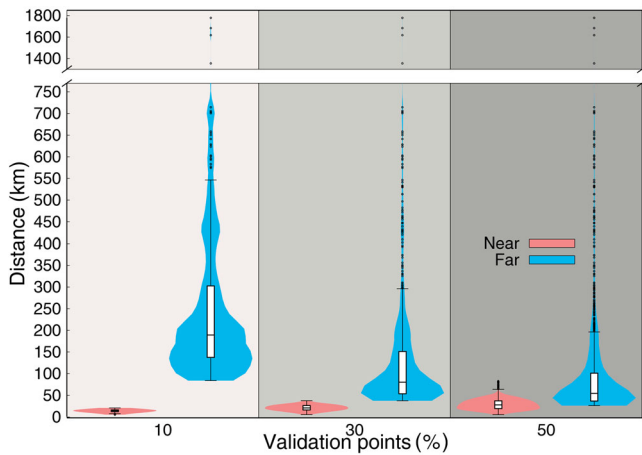
Fig. 4 Violin and box plots showing the distances between all pairs of validation sites and closest calibration sites for Near (red) and Far (blue) sub-networks over the study period. Boxes represent the inter-quartile range (first quartile Q1 and third quartile Q3); line inside the box is the median value; and the black dots are outliers. Outliers are defined as values 1.5 times the inter-quartile distance (IQR = Q3–Q1, corresponding to the whiskers) above or below Q3 or Q1, respectively.

stations with records over the 30-year study period were considered despite their number and spatial distribution changing from year to year as a result of missing data. In fact, no year in the 1980–2009 period includes all stations as shown in Fig. 2. Figure 5a shows the areas where the estimated CPIs were more influenced by nearby station values (areas in blue) or by the initial field (reanalysis; areas in red) values after the application of OI and EnOI. Figure 5b provides the relative expected analysis error values (Daley, 1991; Kalnay, 2002; Sen, 2009) (in this case the expected relative average CPI error). Values close to zero indicate that the neighbouring observations have the largest influence on estimated CPI fields, while values close to one indicate that the background or initial field has the largest impact on the estimated CPI fields. Despite CPI values at each grid point being estimated using neighbouring stations, not only the closest station, Fig. 5b provides valuable information about the relationship between the expected CPI error and the distance to the station having the most significant impact on the estimated values. It is possible to see that the contribution from observations to the analysis field decreases more quickly with distance for some indices (e.g., RX5day, RX1day, R20mm, CDD, and CWD). Indices with the most substantial contribution from neighbouring observations for a given distance are PRCPTOT, PRmean, and PR1mm. It can be seen that for some remote regions, the main contribution is from the reanalysis field, and the overall quality of the analysis field highly relies on the quality of the reanalysis field.

6 Results and discussion

a Comparison between Reanalysis, OI, and EnOI Datasets

Taylor diagrams (Taylor, 2001) were used to compare CPI annual series at stations with (i) reanalysis CPI values; (ii)
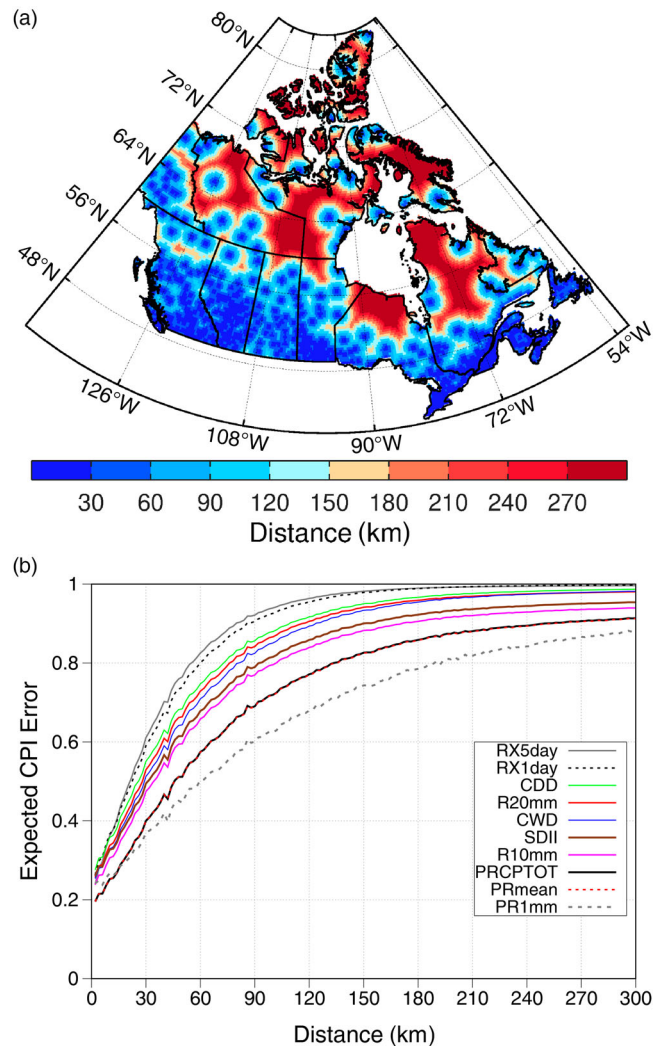

Fig. 5 (a) Distance between each grid point and the closest station. (b) Expected relative CPI error as a function of the distance to the closest station.

reanalysis CPI values after the application of OI; (iii) ensemble CPI values before the application of EnOI; and (iv) CPI values after the application of EnOI (Fig. 6). The statistics displayed in the Taylor diagrams were calculated for each CPI by combining the annual CPI values from all stations.

Reanalysis performance depends on considered CPI, the largest normalized root mean square difference (NRMSD) value being obtained for the CFSR CWD index and the smallest one for the ERA5 PRmean, ERA5 R10mm, and ERA5 PRCPTOT indices. The reanalysis with the best overall performance is ERA5, while CFSR and MERRA displayed the poorest performance. Correlation coefficients for all indices and all reanalyses are larger than 0.5. After applying OI, most indices display correlation values larger than 0.9, NRMSD smaller than 0.5, and normalized standard deviation (NSD) close to 1. The ensemble mean before applying EnOI (Fig. 6c) confirms the CWD index with the largest NRMSD
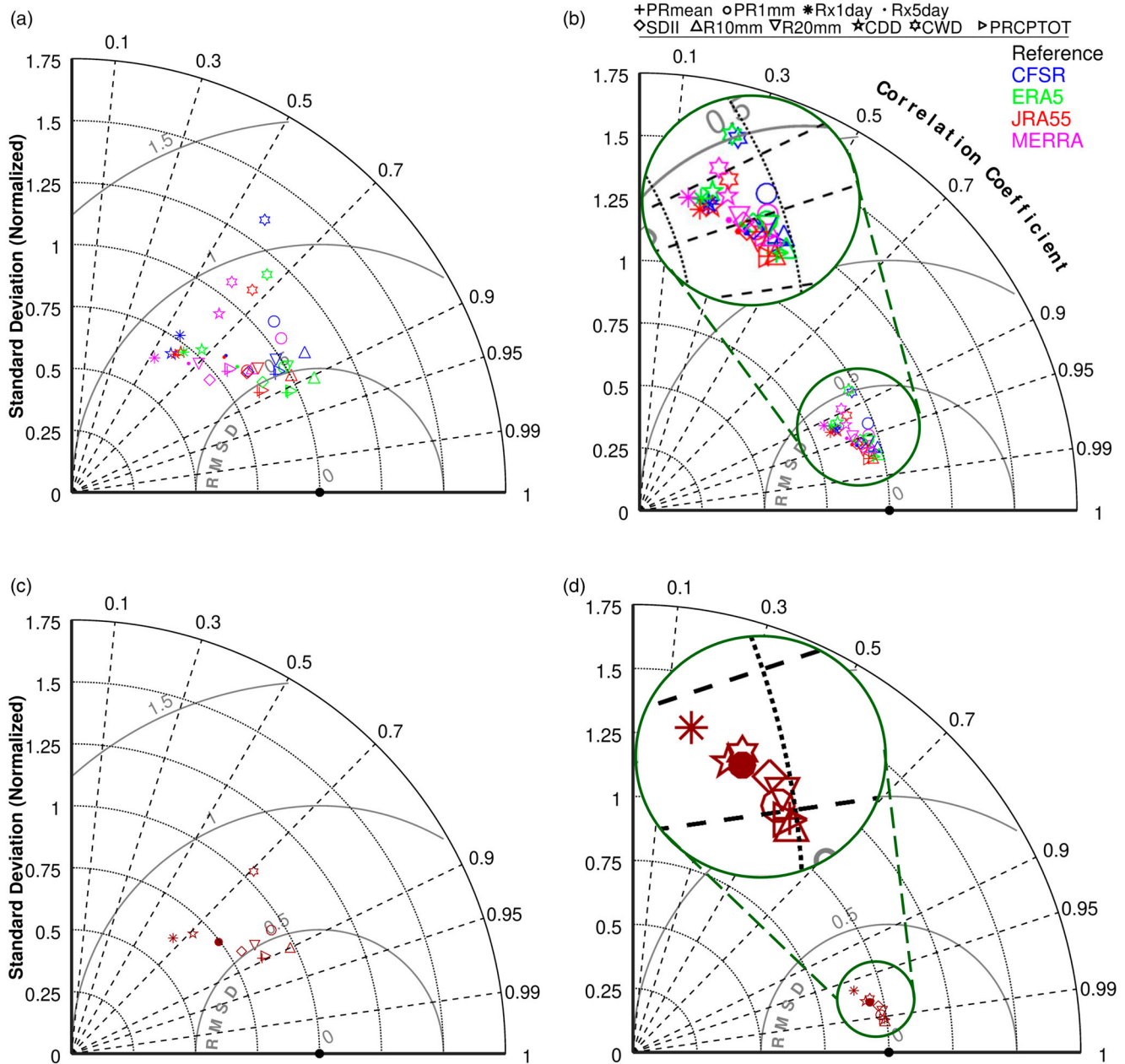
Fig. 6 Taylor diagrams comparing the observed and estimated annual CPI values over the 1980–2009 period at station sites: (a) before applying OI; (b) after applying OI; (c) before applying EnOI, and (d) after applying EnOI. Enlargements for the more dense areas are provided for (b) and (d).

and the smallest for PRmean, R10mm, and PRCPTOT indices. Combining the various reanalyses through EnOI (Fig. 6d) still improved the overall performance because the NRMSD for all indices is smaller than 0.3, and the correlation coefficient is larger than 0.95. The indices with the largest NRMSD are RX1day, CDD, CWD, and RX5day.

Considering that the climatology of northern and mountainous regions are generally poorly represented, a similar analysis was carried out for northern regions (48 stations located north of 60°N) and mountainous regions (246 stations according to the K3 classification of Karagulle et al. (2017); regions in red in Fig. 1) before and after applying EnOI. Figure 7 shows the

Taylor diagrams for northern regions (Figs 7a and 7b), and mountainous regions (Figs 7c and 7d). It is possible to see that applying EnOI improves the CPI estimates over these specific regions. The largest NRMSD values for northern regions were obtained after applying EnOI for the R20mm index, while for mountainous regions the largest NRMSD values were estimated for the CDD index. Small differences are observed in the global performance for these regions compared with the entire territory. These results are not really surprising because performance is assessed at the same sites as those used for the data-assimilation process. They, however, confirm that applying OI or EnOI improves the CPI estimates at station sites.
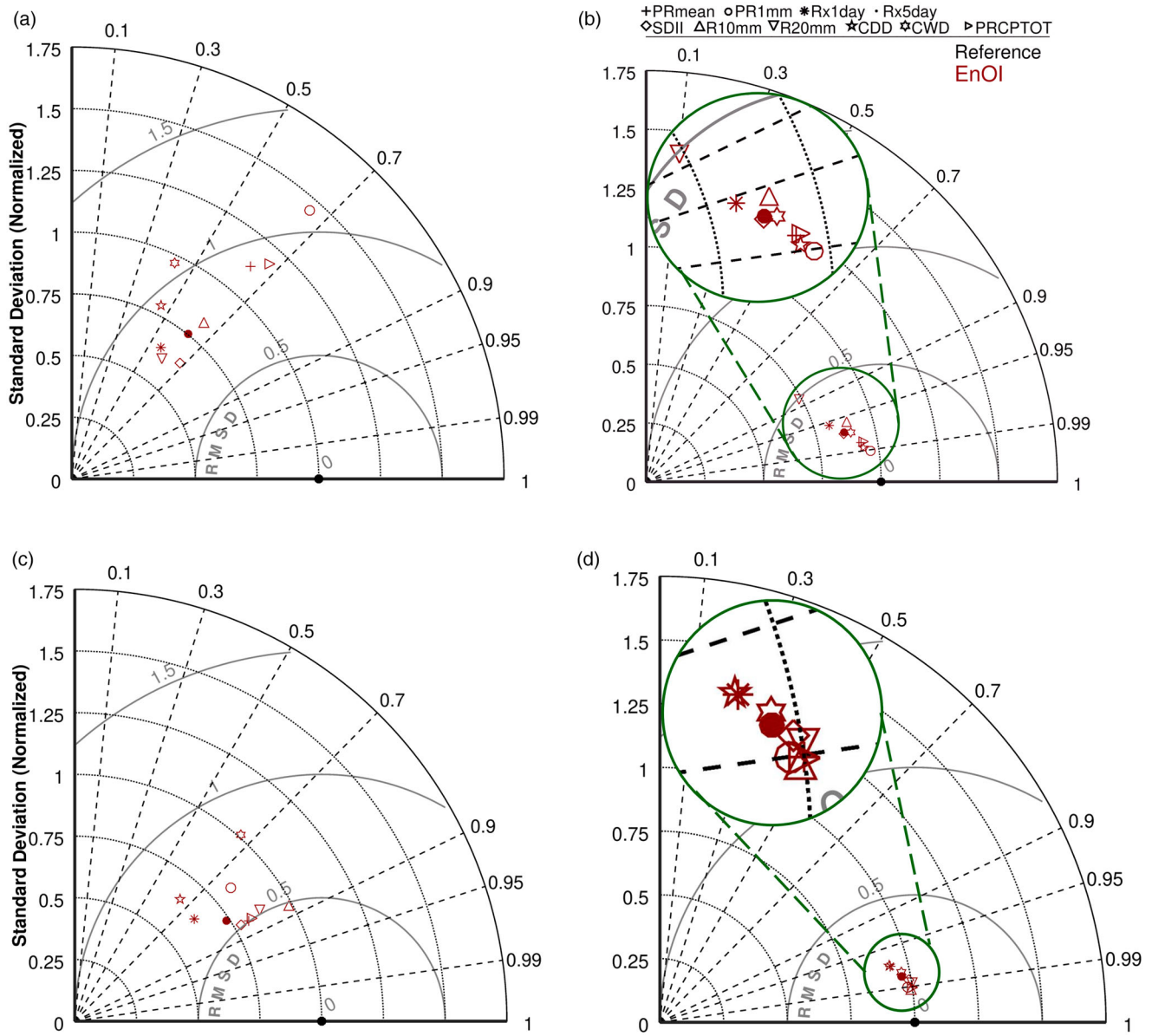
Fig. 7 Taylor diagrams comparing the observed and estimated annual CPI values over the 1980–2009 period at station sites in northern regions (48 stations located north of 60°N): (a) before applying EnOI and (b) after applying EnOI; and in mountainous regions (246 stations according to the K3 classification of Kar-agulle et al., 2017): (c) before applying EnOI and (d) after applying EnOI.

b *Cross-Validation*

Cross-validation was applied to assess the performance of OI and EnOI at points that were not used to upgrade the initial field. Figure 8 shows a heatmap with the values of $N_{(MAE)}$ defined by

$$N_{(MAE)} = \frac{MAE_A}{MAE_B}, \qquad (14)$$

where MAE_A is the mean absolute error of the annual CPI values at validation sites of the analysis field (reanalysis after applying OI or EnOI), and MAE_B is the mean absolute error of the annual CPI values of the initial or background field (each reanalysis value or average reanalysis value).

It can be seen that for each CPI and cross-validation network Near or Far (and, therefore, even for points far from the calibration network), estimated CPI fields were improved after applying OI or EnOI compared with reanalysis values (all $N_{(MAE)}$ values are smaller than one). The largest $N_{(MAE)}$ values (smallest improvement compared with the background field) are obtained, in general, for the Far configuration with 10% validation sites, which represents the configuration with the largest average distance between calibration and validation networks (see Fig. 4). When increasing the percentage of validation points to 30% or 50%, $N_{(MAE)}$ decreases because validation points are closer to calibration points (see Fig. 4). Reanalyses cannot be compared with each other in Fig. 8
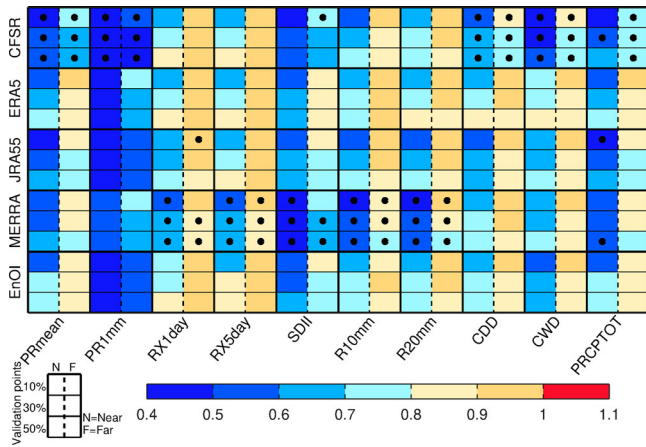
Fig. 8 Mean Absolute Error of the analysis field (MAE$_A$) normalized by the Mean Absolute Error of the background field (MAE$_B$) over all validation sites for the entire period (1980–2009) (Eq. (14)) and each CPI (x-axis). Near (N) and Far (F) configurations with 10%, 30%, and 50% validation points after applying OI to the four reanalyses (top four rows) and after applying EnOI (bottom row) are presented. Black dots identify, for each CPI, the datasets with the lowest $N_{(MAE)}$ values for a given sub-network (N or F with 10%, 30%, or 50% of validation points).

Fig. 9 RMSD of the analysis field (RMSD$_A$) normalized by the standard deviation for each annual CPI series at all validation sites over the 1980–2009 period ($N_{(RMSD)}$ values; (Eq. (15)) and each CPI (x-axis). Near (N) and Far (F) configurations with 10%, 30%, and 50% validation points after applying OI to the four reanalyses (top four rows) and after applying EnOI (bottom row) are presented. Black dots identify the datasets with the best performance (smallest $N_{(RMSD)}$ values) for a given sub-network (N or F with 10%, 30%, or 50% validation points) and CPI.

because reference datasets are different in each case (denominator of Eq. (14)). For instance, one dataset can display a large improvement, but the initial performance could have been poor, while another dataset can display a small improvement but perform well initially.

The $N_{(MAE)}$ values for the various datasets were compared for each sub-network (Near or Far with 10%, 30%, or 50% validation points) and each CPI. The black dots on Fig. 8 identify the dataset with the largest improvement (smallest $N_{(MAE)}$) for each sub-network and each CPI. The CFSR and MERRA reanalyses showed the largest improvements for most sub-networks. The CPIs with the smallest improvements, especially for Far sub-networks, were RX1day and RX5day, while PR1mm stands out as the CPI with the largest improvement. These results are consistent with Fig. 5b in which contributions from observations to RX1day and RX5day decreased as distance to the closest station increased, while PR1mm is the CPI with the largest contribution from observations for the closest station.

A similar heatmap is presented in Fig. 9 for $N_{(RMSD)}$ values defined by

$$N_{(RMSD)} = \frac{RMSD_A}{\sigma_o}, \qquad (15)$$

where RMSD$_A$ is the root mean square difference of the analysis field (reanalysis after applying OI or EnOI), and σ_o is the standard deviation of annual CPI values for all validation sites over the 1980–2009 period. Each dataset is compared with the same reference dataset; therefore, it is possible to select the datasets displaying the best performance for each CPI. Figure 9 shows that, for most CPIs and sub-networks,
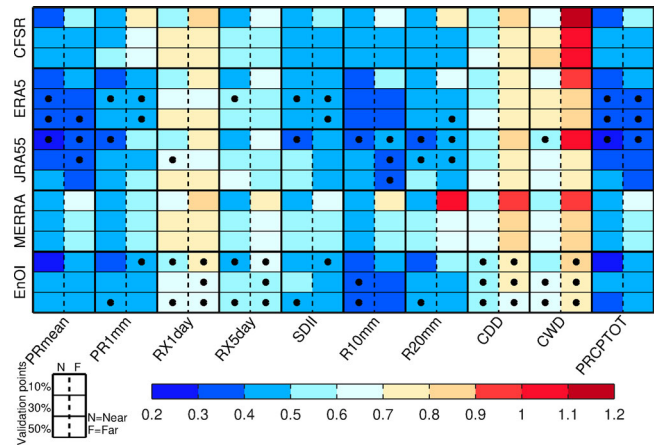
EnOI outperformed the OI datasets (28 CPI and sub-network combinations representing 46.7% of all cases), followed by the analysis field based on OI applied to JRA55 (17 CPI and sub-network combinations for 28.3% of all cases) and finally by the analysis field based on OI applied to ERA5 (15 CPI and sub-network combinations for 25% of all cases). It can also be seen that EnOI outperformed the other datasets for PR1mm, RX1day, RX5day, SDII, CDD, and CWD for the Far configuration with 10% of validation points corresponding to the case with the largest average distance between calibration and validation sites (Fig. 4).

These results show that applying OI and EnOI methods improved the representation of the CPI fields across Canada compared with the initial field (single reanalysis) for all CPIs and even at ungauged sites far from the sites that were used for OI and EnOI calibration. Also, it shows that the EnOI in most cases outperformed OI. This last result is interesting because it shows that combining all the available information (all reanalyses) through EnOI is the best strategy.

c Comparison between EnOI and NRCan Gridded Interpolated Datasets

The CPI dataset created from EnOI was compared with the CPI values estimated from the NRCan dataset. Although the NRCan dataset is constructed by interpolating station records and, therefore, should adequately reproduce CPI values at sites where stations are located, two major issues must be considered. First, station density has an effect on estimated interpolated precipitation values suggesting possible large errors for remote regions with low station density (Gervais et al., 2014). Second, representativeness errors defined by Tustison et al. (2001) as "the errors in representing data (i.e., either model

output or observations) at a scale other than their own inherent scale" may also be influenced by the interpolating procedure. Besides, as was pointed out in Section 2, some stations used in this study were not considered when creating the NRCan dataset. Because the NRCan dataset grid (0.083° × 0.083°) has a finer spatial resolution than EnOI (0.25° × 0.25°), it was necessary to remap the NRCan dataset to the EnOI grid resolution in order to do a grid-to-grid comparison. A conservative remapping method (Jones, 1999) was used.

Figure 10 shows maps of EnOI and NRCan mean CPI values over the 1980–2009 period and the corresponding grid-to-grid differences between the EnOI and NRCan values for three specific indices (PRmean, PR1mm, and RX1day). The smallest differences between the EnOI and NRCan datasets were observed for PRmean with generally smaller values for the NRCan dataset. The PR1mm values for EnOI are globally lower than the NRCan values (especially

in the centre and southern part of the study area). In the northern part of the study area, the NRCan dataset also displays a small area with lower values for this index. Lower RX1day values are observed for the NRCan dataset for almost the entire domain, and this could probably be attributed to the interpolation process used to create this dataset, which smooths out extreme values or the fact that corrections (e.g., evaporation loss and wind undercatch) were not applied to the observations used to build this dataset.

Annual CPI values over the reference period estimated from the EnOI and NRCan datasets were then compared with corresponding station values. Figure 11 shows the corresponding scatterplots for these three indices. It shows that the NRCan dataset slightly underestimates the observed PRmean values, but underestimations are much more important for RX1day, especially at sites with large observed RX1day values. Finally, PR1mm estimated from NRCan overestimates
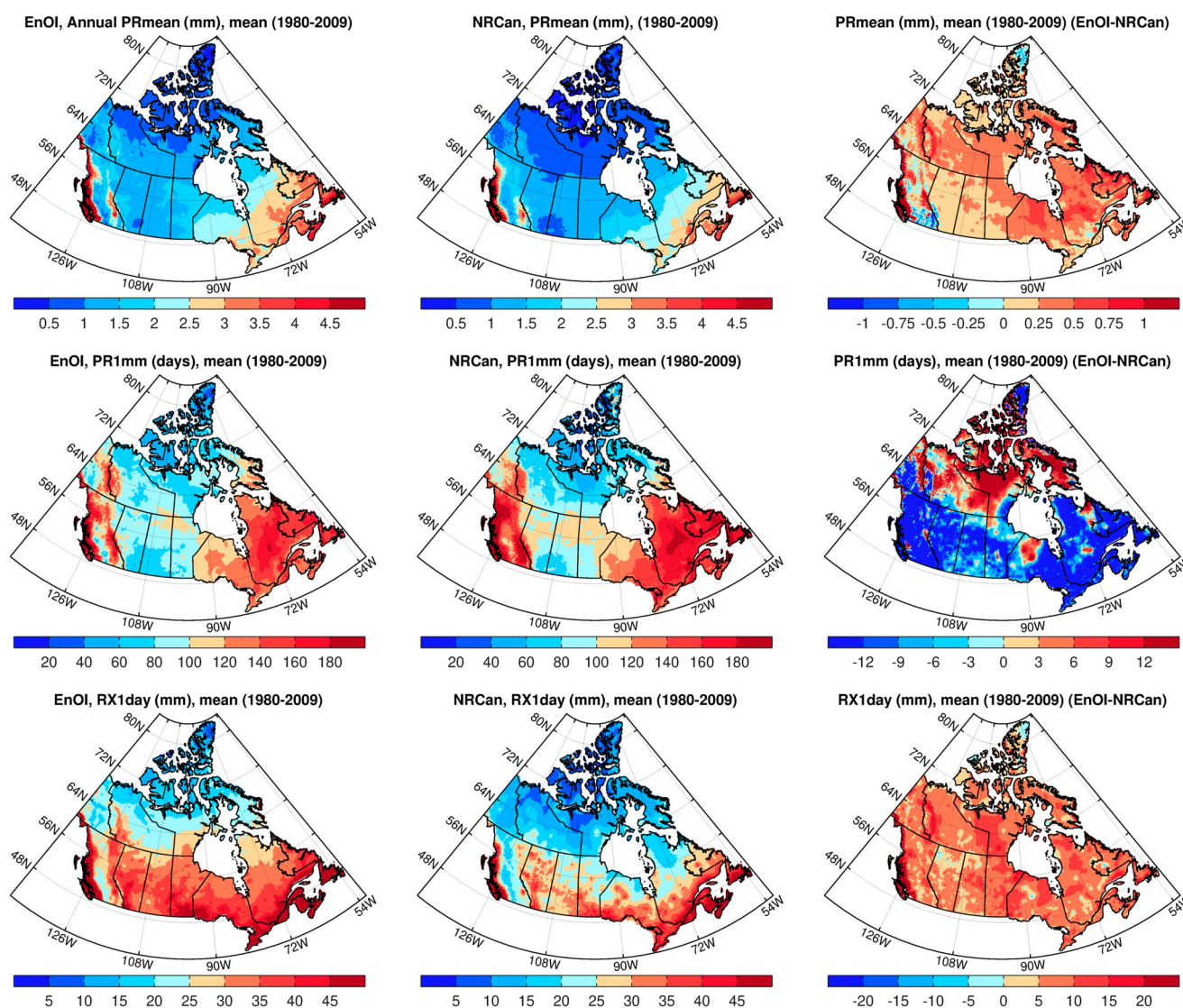

Fig. 10   Mean PRmean (first row), PR1mm (second row), and RX1day (third row) values over the 1980–2009 period as estimated from the EnOI (left column), NRCan datasets (middle column), and corresponding grid-to-grid differences between EnOI and NRCan (right column).
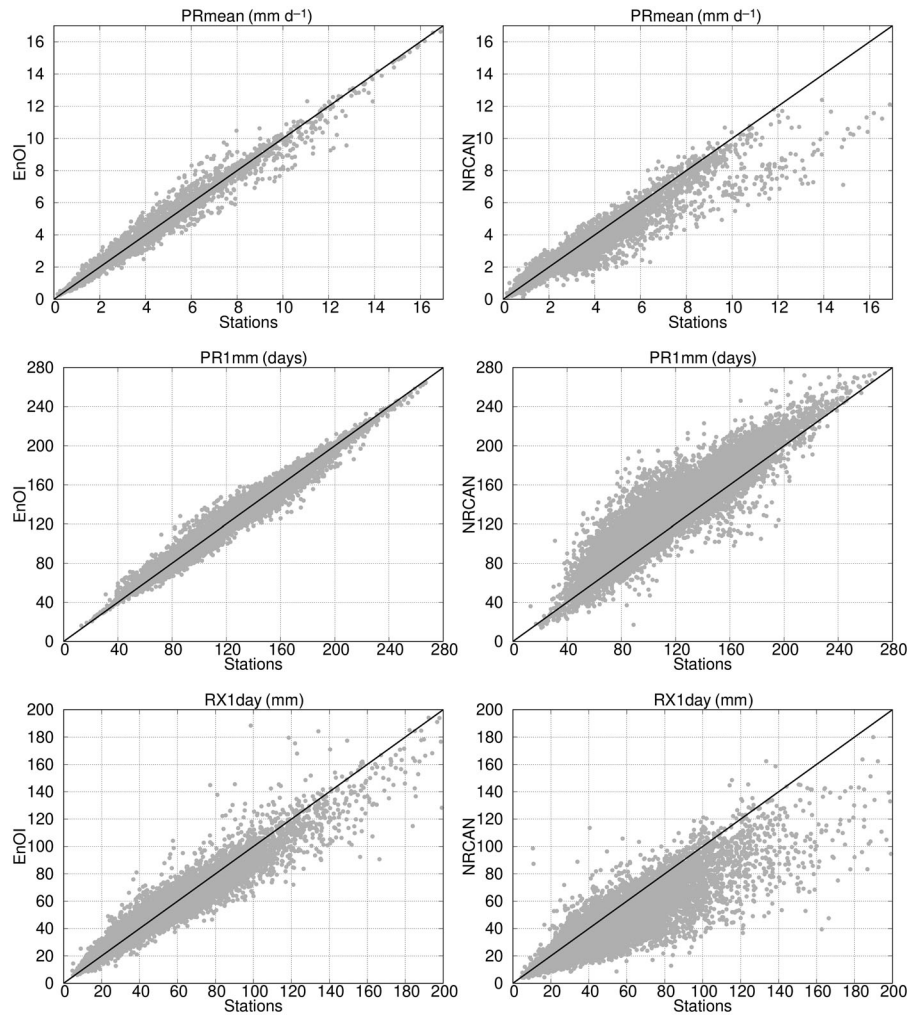
Fig. 11   Scatterplots of PRmean (first row), PR1mm (second row), and RX1day (third row) values estimated from EnOI (left column) and NRCan (right column) dataset against the corresponding observed values.

corresponding station values. Biases are much smaller for the EnOI ensembles for all three indices.

These results show that the CPI estimated from the EnOI dataset, obtained through the combination of the four reanalyses with the observational dataset, outperformed the NRCan dataset over Canada. Also, these results show that the NRCan dataset must be used with caution for extreme indices in Canada as previously reported (Diaconescu et al., 2018).

7   Summary and conclusions

Defining a reference climate for precipitation is an important issue in many hydrological and water resource applications. It is also essential when assessing the impact of climate change and comparing various adaptation strategies. However, it can be challenging in remote regions that are historically poorly covered by station networks. Reanalyses represent an interesting option for such regions but can contain large biases because many do not assimilate surface precipitation observations. Combining reanalysis and station records could, therefore,

provide an improvement in the spatial representation of precipitation. In this study, the off-line data assimilation techniques (OI and EnOI) were applied to define a historical climate for Canada through the combination of observational and reanalysis datasets. The OI method combined each reanalysis independently with the available station records, while the EnOI method combined station records with the information provided by all reanalyses. A period of 30 years (1980–2009) was considered. Annual series for ten CPIs were estimated for each OI and EnOI dataset.

In order to assess the performance of these datasets, preliminary analyses were first carried out comparing the estimated annual precipitation indices over the 1980–2009 period at each observational site with the corresponding nearest gridpoint values estimated through OI or EnOI. Results showed that reanalysis performance strongly depends on the precipitation indices considered. For instance, indices displaying the best overall performance were annual mean precipitation (PRmean) and the annual total precipitation from days with precipitation greater than or equal to 1 mm (PRCPTOT),

while the index showing the poorest performance was the annual maximum number of consecutive days with precipitation greater than or equal to 1 mm (CWD). The ERA5 reanalysis globally outperformed the other reanalyses, while CFSR and MERRA were the reanalyses with the poorest global performance. As expected, OI and EnOI led to substantial reductions in the differences between observed and estimated values at station sites.

Cross-validation was then used to evaluate the performance at sites not included in the calibration process. For this purpose, stations were classified into two sub-networks, one used for calibration and the other for validation. Two different sub-networks, Near and Far, were also defined based on the average distance between calibration and validation sites. Validation networks including 10%, 30%, and 50% of the original stations were considered. A total of 60 sub-networks per dataset (six for each of the ten CPIs, three Near, and three Far) were considered. Improvements in the estimated OI CPI annual series compared with their respective reanalysis were observed for all sub-networks. These improvements were larger for CPIs estimated from CFSR and MERRA than CPIs estimated from other reanalyses after applying OI. Among the datasets obtained after applying OI, CPIs estimated from JRA55 and ERA5 displayed the overall best performance, while those estimated from MERRA and CFSR had the poorest performance. Finally, the EnOI dataset also outperformed OI applied to each reanalysis.

Finally, the CPI annual series estimated from the EnOI dataset were compared with the corresponding values estimated from the NRCan dataset. A grid-to-grid comparison was carried out for three specific indices: PRmean, PR1mm, and RX1day. Differences between each of these datasets and observed values remained small for PRmean, while the NRCan dataset globally overestimated PR1mm values and underestimated RX1day values. The smoothing effect from the interpolation method may partly explain these biases. These results show that the NRCan dataset should be used with caution for extreme indices such as RX1day.

Precipitation observations are notoriously difficult to measure in northern environments and may contain significant site- and instrument-related biases and uncertainties. Assimilating observational datasets with reanalyses as proposed in this study is based on the assumption that station records represent the most valuable information to characterize the past climatology of these regions despite the possible biases and uncertainties. Corrections of possible biases caused by, for instance, evaporation loss, wind undercatch, and trace precipitation are, therefore, important because they may help improve the quality of observational datasets and all datasets derived from this primary source of data.

The proposed dataset provides a unique reference climate dataset for ten precipitation indices covering Canada for the 1980–2009 period. It combines the main available reanalyses that were bias corrected with available station records. Such a gridded dataset increases our ability to characterize the past climate over regions that are poorly covered with station networks, and it is an important contribution to climate change impact studies, as well as the development of adaptation strategies. Future work will look at the extension of the period presented in this study to include more recent years, as well as to incorporate other datasets in the assimilation process (e.g., satellite data). Different comparison analyses will be carry out with other available datasets for recent years (e.g., the Canadian Precipitation Analysis). In addition, the application of these methods to daily precipitation series is planned, enabling a possible application to hydrology.

References

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., … Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, *111*(D5), D05109. https://doi.org/10.1029/2005JD006290

Alexeev, V. A., Esau, I., Polyakov, I. V., Byam, S. J., & Sorokina, S. (2012). Vertical structure of recent Arctic warming from observed data and reanalysis products. *Climatic Change*, *111*(2), 215–239. https://doi.org/10.1007/s10584-011-0192-8

Alidoost, F., Stein, A., Su, Z., & Sharifi, A. (2019). Multivariate copula quantile mapping for bias correction of reanalysis air temperature data. *Journal of Spatial Science*, 1–17. https://doi.org/10.1080/14498596.2019.1601138

Bajamgnigni Gbambie, A. S., Poulin, A., Boucher, M.-A., & Arsenault, R. (2017). Added value of alternative information in interpolated precipitation datasets for hydrology. *Journal of Hydrometeorology*, *18*(1), 247–264. https://doi.org/10.1175/JHM-D-16-0032.1

Bertino, L., Evensen, G., & Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography. *International Statistical Review*, *71*(2), 223–241. https://doi.org/10.1111/j.1751-5823.2003.tb00194.x

Bosilovich, M. G., Chen, J., Robertson, F. R., & Adler, R. F. (2008). Evaluation of global precipitation in reanalyses. *Journal of Applied*

Meteorology and Climatology, *47*(9), 2279–2299. https://doi.org/10.1175/2008JAMC1921.1

Brasnett, B. (1999). A global analysis of snow depth for numerical weather prediction. *Journal of Applied Meteorology*, *38*(6), 726–740. https://doi.org/10.1175/1520-0450(1999)038<0726:AGAOSD>2.0.CO;2

Candiani, G., Carnevale, C., Finzi, G., Pisoni, E., & Volta, M. (2013). A comparison of reanalysis techniques: Applying optimal interpolation and ensemble Kalman filtering to improve air quality monitoring at mesoscale. *Science of the Total Environment*, *458-460*, 7–14. https://doi.org/10.1016/j.scitotenv.2013.03.089

Cherubini, T., Ghelli, A., & Lalaurette, F. (2002). Verification of precipitation forecasts over the alpine region using a high-density observing network. *Weather and Forecasting*, *17*(2), 238–249. https://doi.org/10.1175/1520-0434(2002)017<0238:VOPFOT>2.0.CO;2

Daley, R. (1991). *Atmospheric data analysis*. Cambridge University Press.

Diaconescu, E. P., Mailhot, A., Brown, R., & Chaumont, D. (2018). Evaluation of CORDEX-Arctic daily precipitation and temperature-based climate indices over Canadian Arctic land areas. *Climate Dynamics*, *50*(5), 2061–2085. https://doi.org/10.1007/s00382-017-3736-4

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J. H., Willett, K. M., Aguilar, E., Brunet, M., Caesar, J., Hewitson, B., Jack, C., Klein Tank, A. M. G., Kruger, A. C., Marengo, J., Peterson, T. C., Renom, M., Oria Rojas, C., Rusticucci, M., … Kitching, S. (2013). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, *118*(5), 2098–2118. https://doi.org/10.1002/jgrd.50150

Donat, M. G., Sillmann, J., Wild, S., Alexander, L. V., Lippmann, T., & Zwiers, F. W. (2014). Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets. *Journal of Climate*, *27*(13), 5019–5035. https://doi.org/10.1175/JCLI-D-13-00405.1

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Fletcher, S., Lickley, M., & Strzepek, K. (2019). Learning about climate change uncertainty enables flexible water infrastructure planning. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-018-07882-8

Frich, P., Alexander, L. V., DellaMarta, P., Gleason, B., Haylock, M., Tank, A. M. G. K., & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, *19*(3), 193–212. https://doi.org/10.3354/cr019193

Fu, W., She, J., & Zhuang, S. (2011). Application of an ensemble optimal interpolation in a North/Baltic Sea model: Assimilating temperature and salinity profiles. *Ocean Modelling*, *40*(3–4), 227–245. https://doi.org/10.1016/j.ocemod.2011.09.004

Gervais, M., Tremblay, L. B., Gyakum, J. R., & Atallah, E. (2014). Representing extremes in a daily gridded precipitation analysis over the United States: Impacts of station density, resolution, and gridding methods. *Journal of Climate*, *27*(14), 5201–5218. https://doi.org/10.1175/JCLI-D-13-00319.1

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Olugokencky, E. J., Easterling, J. R., Kaplan, A., Soden, B. J., Thome, P. W., Wild, M., & Zhai, P. M. (2013). Observations: Atmosphere and surface. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Hatzaki, M., Flocas, H., Oikonomou, C., & Giannakopoulos, C. (2010). Future changes in the relationship of precipitation intensity in Eastern Mediterranean with large scale circulation. *Advances in Geosciences*, *23*, 31–36. https://doi.org/10.5194/adgeo-23-31-2010

Häggmark, L., Ivarsson, K.-I., Gollvik, S., & Olofsson, P.-O. (2000). Mesan, an operational mesoscale analysis system. *Tellus A: Dynamic Meteorology and Oceanography*, *52*(1), 2–20. https://doi.org/10.3402/tellusa.v52i1.12250

Hersbach, H., Bell, W., Berrisford, P., Horányi, A,J,M-S, Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., & Dee, D. (2019, Spring). Global reanalysis: Goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter*, *159*, 17–24. https://doi.org/10.21957/vf291hehd7

Hintze, J. L., & Nelson, R. D. (1998). Violin plots: A box plot-density trace synergism. *The American Statistician*, *52*(2), 181–184. https://doi.org/10.1080/00031305.1998.10480559

Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E., & Papadopol, P. (2009). Development and testing of Canada-wide interpolated spatial models of daily minimum–maximum temperature and precipitation for 1961–2003. *Journal of Applied Meteorology and Climatology*, *48*(4), 725–741. https://doi.org/10.1175/2008JAMC1979.1

Jones, P. W. (1999). First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, *127*(9), 2204–2210. https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2

Kalnay, E. (2002). *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press.

Karagulle, D., Frye, C., Sayre, R., Breyer, S., Aniello, P., Vaughan, R., & Wright, D. (2017). Modeling global Hammond landform regions from 250-m elevation data. *Transactions in GIS*, *21*(5), 1040–1060. https://doi.org/10.1111/tgis.12265

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., & Takahashi, K. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Lachance-Cloutier, S., Turcotte, R., & Cyr, J.-F. (2017). Combining streamflow observations and hydrologic simulations for the retrospective estimation of daily streamflow for ungauged rivers in southern Quebec (Canada). *Journal of Hydrology*, *550*, 294–306. https://doi.org/10.1016/j.jhydrol.2017.05.011

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., & Stadnyk, T. (2015). Performance evaluation of the Canadian precipitation analysis (CaPA). *Journal of Hydrometeorology*, *16*(5), 2045–2064. https://doi.org/10.1175/JHM-D-14-0191.1

Lindsay, R., Wensnahan, M., Schweiger, A., & Zhang, J. (2014). Evaluation of seven different atmospheric reanalysis products in the Arctic. *Journal of Climate*, *27*(7), 2588–2606. https://doi.org/10.1175/JCLI-D-13-00014.1

Lorenz, C., & Kunstmann, H. (2012). The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis. *Journal of Hydrometeorology*, *13*(5), 1397–1420. https://doi.org/10.1175/JHM-D-11-088.1

Mahfouf, J.-F., Brasnett, B., & Gagnon, S. (2007). A Canadian precipitation analysis (CaPA) project: Description and preliminary results. *Atmosphere-Ocean*, *45*(1), 1–17. https://doi.org/10.3137/ao.v450101

Matsikaris, A., Widmann, M., & Jungclaus, J. (2015). On-line and off-line data assimilation in palaeoclimatology: A case study. *Climate of the Past*, *11*(1), 81–93. https://doi.org/10.5194/cp-11-81-2015

McKenney, D. W., Hutchinson, M. F., Papadopol, P., Lawrence, K., Pedlar, J., Campbell, K., Milewska, E., Hopkinson, R. F., Price, D., & Owen, T. (2011). Customized spatial climate models for North America. *Bulletin of the American Meteorological Society*, *92*(12), 1611–1622. https://doi.org/10.1175/2011BAMS3132.1

Mekis, É, & Vincent, L. A. (2011). An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmosphere-Ocean*, *49*(2), 163–177. https://doi.org/10.1080/07055900.2011.583910

Nie, S., Wu, T., Luo, Y., Deng, X., Shi, X., Wang, Z., Liu, X., & Huang, J. (2016). A strategy for merging objective estimates of global daily precipitation from gauge observations, satellite estimates, and numerical predictions. *Advances in Atmospheric Sciences*, *33*(7), 889–904. https://doi.org/10.1007/s00376-016-5223-y

Oke, P. R., Brassington, G. B., Griffin, D. A., & Schiller, A. (2010). Ocean data assimilation: A case for ensemble optimal interpolation. *Australian Meteorological and Oceanographic Journal*, *59*(1SP), 67–76. https://doi.org/10.22499/2.5901.008

Oke, P. R., Sakov, P., & Corney, S. P. (2007). Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dynamics*, *57* (1), 32–45. https://doi.org/10.1007/s10236-006-0088-8

Panthou, G., Vischel, T., Lebel, T., Blanchet, J., Quantin, G., & Ali, A. (2012). Extreme rainfall in West Africa: A regional modeling. *Water Resources Research*, *48*(8), W08501. https://doi.org/10.1029/2012WR012052

Peterson, T. C., Folland, C., Gruza, G., Hogg, W., Mokssit, A., & Plummer, N. (2001). *Report on the activities of the working group on climate change detection and related rapporteurs 1998-2001*. WMO, Rep. WCDMP-47, WMO-TD 1071, Geneve, Switzerland, http://etccdi.pacificclimate.org/docs/wgccd.2001.pdf

Ren, L., & Hartnett, M. (2017). Sensitivity analysis of a data assimilation technique for hindcasting and forecasting hydrodynamics of a complex coastal water body. *Computers & Geosciences*, *99*(C), 81–90. https://doi.org/10.1016/j.cageo.2016.10.012

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., … Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., … Goldberg, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058. https://doi.org/10.1175/2010BAMS3001.1

Sen, Z. (2009). *Spatial modeling principles in earth sciences*. Springer Netherlands.

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, *118*(4), 1716–1733. https://doi.org/10.1002/jgrd.50203

Soci, C., Bazile, E., Besson, F., & Landelius, T. (2016). High-resolution precipitation re-analysis system for climatological purposes. *Tellus A: Dynamic Meteorology and Oceanography*, *68*(1), 29879. https://doi.org/10.3402/tellusa.v68.29879

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P. M. (eds.) (2013). IPCC, 2013: Summary for Policymakers. In *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Tebaldi, C., Hayhoe, K., Arblaster, J. M., & Meehl, G. A. (2006). Going to the extremes. *Climatic Change*, *79*(3–4), 185–211. https://doi.org/10.1007/s10584-006-9051-4

Tustison, B., Harris, D., & Foufoula-Georgiou, E. (2001). Scale issues in verification of precipitation forecasts. *Journal of Geophysical Research: Atmospheres*, *106*(D11), 11775–11784. https://doi.org/10.1029/2001JD900066

Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjernström, M., Lüpkes, C., Nygård, T., Notz, D., Weiss, J., Marsan, D., Cheng, B., Birnbaum, G., Gerland, S., Chechin, D., & Gascard, J. C. (2014). Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: A review. *Atmospheric Chemistry and Physics*, *14*(17), 9403–9450. https://doi.org/10.5194/acp-14-9403-2014

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kandon, E. J., Lenderink, G., & Roberts, N. M. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, *52*(3), 522–555. https://doi.org/10.1002/2014RG000464

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., & Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(6), 851–870. https://doi.org/10.1002/wcc.147

Zou, H., Zhu, J., Zhou, L., Li, P., & Ma, S. (2014). Validation and application of reanalysis temperature data over the Tibetan Plateau. *Acta Meteorologica Sinica*, *28*(1), 139–149. https://doi.org/10.1007/s13351-014-3027-5