

# Observed and Simulated Precipitation over Northeastern North America: How Do Daily and Subdaily Extremes Scale in Space and Time?

SILVIA INNOCENTI AND ALAIN MAILHOT

*INRS, Quebec City, Quebec, Canada*

ANNE FRIGON

*Consortium Ouranos, Montreal, Quebec, Canada*

ALEX J. CANNON

*Climate Research Division, Environment and Climate Change Canada, Victoria, British Columbia, Canada*

MARTIN LEDUC

*Consortium Ouranos, Montreal, Quebec, Canada*

(Manuscript received 10 January 2019, in final form 30 July 2019)

## ABSTRACT

The characterization of extreme precipitation at fine spatiotemporal scale represents a paramount challenge in hydroclimate sciences due to large uncertainties affecting the precipitation estimation from existing datasets. Comparing the spatiotemporal structure of precipitation extremes estimated from different datasets thus represents an essential step for climate model evaluation, as it provides insight into a model's ability to simulate atmospheric processes occurring at different scales. This study compares the probability distributions and the annual and diurnal cycles of occurrence of daily and subdaily precipitation annual maxima (AM) estimated over northeastern North America from five observed and simulated datasets: meteorological station series, the bias-corrected (CRT) satellite CMORPH, version 1.0, and the Multi-Source Weighted-Ensemble Precipitation (MSWEP), version 2, gridded datasets, various Canadian RCM, version 5 (CRCM5), simulations, and a 13-yr convection-permitting WRF, version 3.4.1, simulation. ERA-Interim-driven CRCM5 and WRF simulations well reproduced subdaily extreme quantiles and the AM annual and diurnal cycles observed at stations, while CMORPH and MSWEP displayed good performance only for daily and longer extreme statistics. The spatiotemporal statistical structure of precipitation extremes is then assessed considering the variation of AM quantiles across various spatial scales and durations. The results suggest that a two-parameter analytical relationship well describes the AM spatiotemporal structure at the regional scale, allowing us to approximate some crucial properties of point precipitation extremes from gridded datasets. Averaging the estimates from various members of the initial-condition CRCM5 Large Ensemble (CRCM5-LE) also made it possible to reduce the sampling errors and robustly estimate the AM spatiotemporal structure at the local scale of each model grid box.

---

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-19-0021.s1>.

---

Corresponding author: Silvia Innocenti, [silvia.innocenti@ete.inrs.ca](mailto:silvia.innocenti@ete.inrs.ca)

## 1. Introduction

Heavy precipitation may have severe and threatening impacts on human societies and natural ecosystems (e.g., Seneviratne et al. 2012). Moreover, as substantial changes in the frequency, duration, and spatial distribution of extreme precipitation events are expected at global and regional scales (e.g., Dai et al. 2019; Dwyer and O'Gorman 2017), the need for an accurate

DOI: 10.1175/JCLI-D-19-0021.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

characterization of extreme rainfall across various spatial and temporal scales becomes even more important. However, such an objective is still challenging, primarily because of deficiencies in available datasets (Westra et al. 2014; Herold et al. 2017). Sparse networks and short records prevent a comprehensive characterization of the complex spatiotemporal structure of extreme precipitation events (Grimaldi et al. 2015; Trenberth et al. 2017). Equally important, the biases and large uncertainties related to precipitation measurements (e.g., Sikorska and Seibert 2018), data interpolation techniques, and representativeness errors (Tustison et al. 2001), as well as the inhomogeneity of observed records (e.g., Hofstra et al. 2008), must be considered when using both observed gridded datasets (e.g., Tapiador et al. 2017).

In this context, regional climate models (RCMs) and reanalyses represent an interesting alternative to access climate information about extreme rainfall as they provide precipitation series with complete and physically coherent spatial and temporal coverage over large regions (Flato et al. 2013). Increasing RCM spatial resolution and improvements in the representation of key physical processes have led to significant advances in RCM ability at simulating key features of precipitation, such as the annual cycle and the statistics of daily extremes (e.g., Prein et al. 2015; Lucas-Picher et al. 2017). With spatial resolutions typically ranging from 12 to 50 km, state-of-the-art RCMs allow for a more realistic representation of surface forcings, such as orography and coastlines, and small-scale climate and weather processes (e.g., Rummukainen 2016; Fosser et al. 2017).

However, RCMs may still display large inaccuracies in the simulation of subdaily rainfall characteristics (e.g., diurnal cycle and hourly extremes) that have been generally attributed to the inadequate representation of important processes (e.g., cloud dynamics) at the subgrid scale (e.g., Westra et al. 2014; Cavicchia et al. 2018; Liu et al. 2017). Promising approaches for overcoming these inaccuracies rely on the use of convection-permitting models (CPMs) (e.g., Ban et al. 2015; Dai et al. 2019). Running simulations at grid resolutions of a few kilometers, CPMs more completely resolve the physical equations of deep convection (Rasmussen et al. 2019) resulting in a more realistic representation of small-scale processes generating subdaily extremes (e.g., Prein et al. 2016; Kendon et al. 2017). However, due to their high computational and storage costs, many crucial aspects of CPMs, such as the sensitivity to nesting and driving strategies, or spatial domain size, remain largely unexplored (Prein et al. 2019). For the same reasons, relatively few continuous CPM runs are currently available, and these are typically restricted

to small spatial domains and short time periods (e.g., 10 years or less; Prein et al. 2016; Mantegna et al. 2017). This limits the possibility of robustly assessing uncertainties through the use of classical multimodel and/or multimember approaches (Kendon et al. 2017; Coppola et al. 2019).

As part of the Climate Change and Hydrological Extremes (ClimEx) project, a large ensemble of high-resolution ( $0.11^\circ$  latitude–longitude) simulations has been produced for two spatial domains: one covering the northeastern part of North America (Fig. 1a) and one covering most of Europe (see Leduc et al. 2019; Fig. 2). This 50-member ensemble has been simulated over the period 1950–2100 using the fifth-generation Canadian Regional Climate Model (CRCM5; Martynov et al. 2013; Separovic et al. 2013) and is herein called the CRCM5 Large Ensemble (CRCM5-LE). Boundary conditions used to drive the CRCM5-LE were provided by the CanESM2 (von Salzen et al. 2013; Arora et al. 2011) Large Ensemble (CanESM2-LE; Sigmond and Fyfe 2016; Fyfe et al. 2017) simulated under the RCP8.5 emission scenario.

Previous investigations have demonstrated the ability of CRCM5 in reproducing annual and seasonal mean precipitation, as well as daily and multiday extreme precipitation (e.g., Diaconescu et al. 2016; Whan and Zwiers 2016). Leduc et al. (2019) evaluated CRCM5-LE monthly temperature and precipitation by comparing the large ensemble with observational gridded datasets. This study aims at broadening the evaluation of the CRCM5 by analyzing the relative change of the statistical properties of the simulated extremes (e.g., larger return period quantiles) with the spatiotemporal scale. The objective is to assess the differences between the observed and simulated spatiotemporal structure of extreme precipitation. This represents an essential step for climate model evaluation, as it provides guidance for evaluating model ability to correctly simulate atmospheric processes at different spatiotemporal scales relevant to extremes (Cortés-Hernández et al. 2016).

The present study has two specific objectives:

- (i) Assessing and comparing extreme precipitation properties (quantiles, annual and diurnal cycles, and temporal scaling) from CRCM5 simulations with the corresponding estimations obtained from meteorological stations and three gridded datasets: the Climate Prediction Center (CPC) morphing technique (CMORPH) bias-corrected satellite dataset (Xie and Xiong 2011); the Multi-Source Weighted-Ensemble Precipitation (MSWEP), version 2, dataset (Beck et al. 2019b); and one series simulated by the convection-permitting Weather

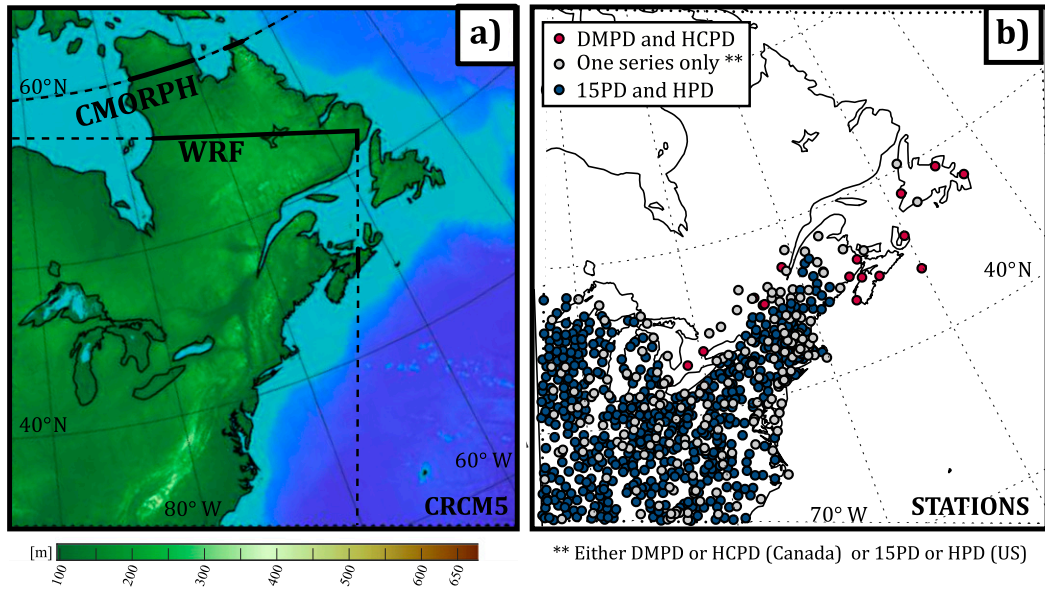


FIG. 1. Spatial domains of gridded datasets and station networks: (a) CRCM5 domain with topography and relevant boundaries of the WRF and CMORPH domains (the MSWEP grid covers the entire CRCM5 domain); (b) stations and available temporal resolution of recorded series: Daily Maxima and Hourly Canadian Precipitation Dataset (DMPD and HCPD) and 15-min and Hourly Precipitation Data (15PD and HPD).

Research and Forecasting (WRF) Model (Liu et al. 2017; Prein et al. 2019).

- (ii) Proposing a simple statistical model to characterize the variation in the quantiles of precipitation extremes across various spatial scales and durations.

The paper is structured as follows: Section 2 introduces the datasets used in this study, while section 3 describes the extraction of rainfall extremes at various spatiotemporal scales. Sections 4 and 5 define the

statistical framework for assessing and comparing extreme precipitation properties among datasets. Sections 6 and 7 present the comparison between gridded dataset and station extreme precipitation statistics. The results from the analysis of the spatiotemporal structure of precipitation extremes for the CRCM5-LE are shown and discussed in section 8. Finally, sections 9 and 10 summarize the main conclusions of the study and provide some perspectives for future work.

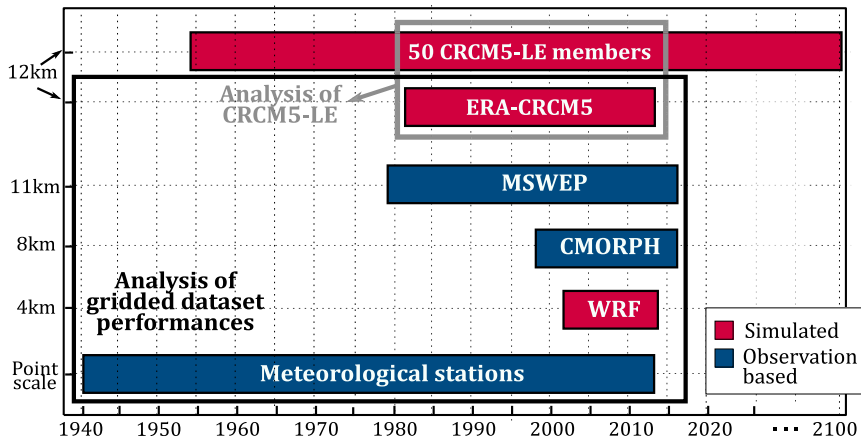


FIG. 2. Spatial resolution (y axis) and covered period (x axis) of datasets; the black rectangle defines the period (1940–2016) and identifies the datasets used for the analysis of gridded dataset performance (see section 6), while the gray rectangle defines the period (1980–2015) considered for the analysis of the CRCM5-LE (section 8).

TABLE 1. Precipitation datasets and their basic characteristics.

Dataset	Description	Period	Spatiotemporal resolution	Reference
Meteorological stations	HCPD and DMPD series	1940–2013	1 h and daily maxima <sup>a</sup>	ECCC, MDDELCC
CRCM5-LE (50 members)	15PD and HPD CanESM2-LE-driven members	1954–2099	15 min and 1 h 0.11°–1 h	NOAA (2016) <sup>b</sup> Leduc et al. (2019)
ERA-CRCM5	ERA-Interim-driven simulation	1981–2013	0.11°–1 h	Leduc et al. (2019)
WRF	ERA-Interim-driven simulation	October 2000– September 2013	4 km–1 h	Liu et al. (2017)
CMORPH (v1.0 CRT)	Bias-corrected satellite analyses	1998–2016	8 km–30 min	Xie et al. (2017)
MSWEP (v2)	Multisource gridded dataset <sup>c</sup>	1979–2016	0.1°–3 h	Beck et al. (2019b) and (2019a)

<sup>a</sup> Daily maxima depth over a 24-h window beginning at 0800 LT for durations 1, 2, 6, and 12 h.

<sup>b</sup> <http://www.ncdc.noaa.gov/data-access/land-based-station-data>.

<sup>c</sup> The CPC Unified v1.0 and real time (Xie et al. 2007; Chen et al. 2008) and GPCP Full Data Reanalysis and First Guess v7 (Schneider et al. 2014) gauge-based datasets; the CMORPH v1.0, GSMaP-MVK v5–v6 (Ushio et al. 2009), and TMPA 3B42RT v7 (Huffman et al. 2007) satellite analyses; and the ERA-Interim and JRA-55 (Kobayashi et al. 2015) reanalyses.

## 2. Data and study area

Meteorological station series are used in this study as a reference dataset and compared to CRCM5 and WRF Model simulations, and to the CMORPH satellite and MSWEP gridded datasets. The study area corresponds to the North American simulation domain of the CRCM5 model for the ClimEx experiment (Fig. 1a). Basic characteristics of the considered datasets are summarized in Table 1.

### a. Meteorological station series

Data from four rain gauge networks operating between 1900 and 2013 over the study domain were pooled for the construction of the point-scale observational dataset: the Daily Maxima Precipitation Dataset (DMPD) and the Hourly Canadian Precipitation Dataset (HCPD) [Environment and Climate Change Canada (ECCC) 2016; Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) du Québec; unpublished data obtained in 2016 from Service InfoClimat, see <http://www.environnement.gouv.qc.ca/climat/surveillance/produits.htm>] for Canada, and the Hourly Precipitation Data (HPD) and 15-min Precipitation Data (15PD) (NOAA 2016) for the United States.

The quality check of records followed the procedure used by Innocenti et al. (2017). As a result, 759 stations were selected, 640 of them having series with less than 5% of missing data. Note that 548 U.S. stations had both 15PD and HPD series, while 17 Canadian stations had both DMPD and HCPD series (blue and red locations in Fig. 1b, respectively). Selected stations are concentrated in the southern part of the study domain, with stations in northern regions having been excluded since their records usually cover the summer period (May–October).

### b. CRCM5-LE

Within the ClimEx initiative, 50 CRCM5, version 3.3.3.1, runs were produced by dynamically downscaling the CanESM2-LE (Fyfe et al. 2017) to a 0.11°-resolution grid ( $\approx 12$  km) for the period 1950–2100 (Leduc et al. 2019). The 50 independent CanESM2-LE members consider observed greenhouse gas, aerosols, and land-use emissions up to the year 2005 and radiative forcing from the RCP8.5 (Meinshausen et al. 2011) for the 2006–2100 period (Sigmond and Fyfe 2016). Hourly precipitation series were available for  $280 \times 280$  grid points over northeastern North America (Fig. 1a). The reader is referred to Martynov et al. (2013) and Separovic et al. (2013) for specific details about CRCM5, and to Leduc et al. (2019) for the CRCM5-LE simulation set-up.

An additional simulation was also produced using the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al. 2011) as atmospheric lateral CRCM5 boundaries (Leduc et al. 2019). Monthly climatologies of precipitation and temperature from the CRCM5-LE and the ERA-Interim-driven simulation (herein referred to as ERA-CRCM5) have been compared with observational gridded datasets by Leduc et al. (2019). In the present study, the ERA-CRCM5 simulation is used to evaluate the performance of the CRCM5 over the 1981–2013 period.

### c. Convection-permitting WRF simulation

The WRF Model, version 3.4.1 (Skamarock et al. 2008), was recently configured into a regional climate model with a single computational domain covering much of North America (Liu et al. 2017). For this

experiment, a 13-yr WRF simulation was performed by dynamically downscaling to 4 km the ERA-Interim reanalysis from October 2000 to September 2013 (Prein et al. 2019; Liu et al. 2017).

Previous studies showed the good performances in simulating seasonal and annual precipitation over most of the contiguous United States (Liu et al. 2017; Dai et al. 2019; Prein et al. 2019). WRF precipitation series were extracted from January 2001 to September 2013 for land grid boxes within the CRCM5 domain after the withdrawal of a 50-gridpoint buffer due to the presence of some boundary effects (e.g., Figs. S1–S4 in the online supplemental material; Leduc and Laprise 2009).

#### d. CMORPH: Bias-corrected satellite dataset

The CMORPH satellite analysis combines observations from multiple microwave satellites to produce precipitation estimates at a nearly global scale (60°N and 60°S) (Joyce et al. 2004; Xie and Xiong 2011). Recently, Xie et al. (2017) reprocessed the CMORPH series and applied a two-step bias correction procedure to obtain a homogeneous dataset at the resolution of 8 km–30 min (Xie et al. 2017). For land grid boxes, the bias correction used the CPC daily gauge analysis (Xie et al. 2010) as reference dataset. The resulting bias-corrected CMORPH v1.0 CRT dataset demonstrated good performance in representing rainfall subdaily variability during the warm season over land (Xie et al. 2017). However, the underestimation of the wintertime precipitation linked to satellite deficiencies in detecting snowfall (Xie and Joyce 2014) is a major shortcoming that must be considered when using CMORPH in high-latitude land and/or during the cold seasons (Trenberth et al. 2017).

#### e. Gridded MSWEP dataset

MSWEP, version 2 (Beck et al. 2019b, 2017), is a global precipitation dataset that blends observations from various observational and reanalysis datasets (see Table 1). In a preliminary evaluation the MSWEP dataset has been found to outperform other satellite and reanalysis datasets in terms of the 3-day and monthly precipitation patterns and for various precipitation indices (Beck et al. 2017, 2019a,b). However, considering the sparseness of ground observations and the low quality of satellite data in northern regions, MSWEP performance for precipitation extremes over our study area must also be evaluated.

### 3. Extraction of annual maxima series at various spatiotemporal scales

Let  $d_0$  represent the native temporal resolution and  $r_0$  the nominal spatial resolution of each dataset (we set

TABLE 2. Spatial scales considered for each gridded dataset.

Dataset	$k$	No. of scales	$r_k$ spatial resolution
CRCM5 <sup>a</sup>	1, 2, ..., 6	6	12, 24, ..., 72 km
WRF	1, 2, ..., 6 and 8, 10, 12, ..., 18	12	4, 8, ..., 24 km and 32, 40, ..., 72 km
CMORPH	1, 2, ..., 9	9	8, 16, ..., 72 km
MSWEP	1, 2, ..., 6	6	11, 22, ..., 66 km

<sup>a</sup> Includes CRCM5-LE and ERA-CRCM5.

$r_0 = 0$  for stations; Tustison et al. 2001; Haerter et al. 2015). From the series of precipitation depth available at  $(r_0, d_0)$ , simulated and observed extremes were assessed through the annual maxima (AM) series at various spatiotemporal scales  $(r, d)$ . In the remainder of the paper, subscripts for  $r$  and  $d$  are omitted if referring to a generic spatial and temporal scale.

The aggregated spatial scales  $r_k = kr_0$ , with  $k = 1, 2, \dots, K$ , were used and they are expressed in kilometers, for simplicity. The coarsest resolution was set to  $r_K \approx 72$  km, with the number of spatial scales considered ranging from 6 for CRCM5 simulations to 12 for WRF. The list of spatial scales considered for each dataset appears in Table 2.

For each time step, precipitation series were aggregated at scales  $r_k > r_0$  with a fixed window in space, that is, computing the average precipitation depth values for  $k \times k$  contiguous grid boxes with no overlap between coarser-scale grid boxes. Dataset grids at scales  $r_k > r_0$  were defined starting from the southwest corner of the native grid and aggregating the grid boxes toward the northeast. This aggregation procedure aims at reproducing the effects of a resolution change for gridded datasets (Eggert et al. 2015). For each aggregation level, grid boxes associated with the ocean and water bodies were removed, and spatially aggregated series were discarded if computed over less than 75% native land grid boxes. A moving window in time was then applied to temporally aggregate precipitation depth series at durations  $d = 1, 2, 3, 4, 6, 12, 18, 24, 36, 48, 60$ , and 72 h ( $d = 1, 2$ , and 4 h were not considered for MSWEP since  $d_0 = 3$  h). Finally, the AM series at each spatiotemporal scale  $(r, d)$  was extracted.

For the 565 stations having series recorded at two different temporal resolutions (blue and red points in Fig. 1b), the maximum of the two AM values extracted from the two series was retained for each common year and relevant duration  $d$ . To reduce the computational time, the WRF grid was subsampled at a 1/10 ratio, leading to the selection of  $\approx 26\,500$  randomly distributed points. All WRF grid boxes containing one or more stations were also considered.

The date and time of occurrence of the AM were extracted for each spatiotemporal scale, year, and location (i.e., a station or grid box) considering the AM starting hour (UTC – 5). Since the time of occurrence was not available for the DMPD observations, such analysis was not performed for stations or years with only DMPD records. Some problems related to the coarse measurement resolution of the tipping-bucket gauges recording 15PD and HPD series were also observed. Since their typical tip resolution (i.e., the minimum recorded nonzero value) is 2.54 mm (i.e., 1/10 in.), only a few distinct rainfall depth values can be recorded at these stations at their native temporal resolution  $d_0$  (15 min or 1 h) (NOAA data online documentation, available at [https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/PRECIP\\_15\\_documentation.pdf](https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/PRECIP_15_documentation.pdf) and [https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/PRECIP\\_HLY\\_documentation.pdf](https://www1.ncdc.noaa.gov/pub/data/cdo/documentation/PRECIP_HLY_documentation.pdf); Innocenti et al. 2017). This results in numerous ties in recorder series and therefore many occurrences of the same AM values for some specific years, therefore making it impossible to uniquely identify the date of AM occurrences for those years. Some preliminary analyses found a large number of ties in 15PD and HPD AM series for durations  $d \leq 6$  h (see Fig. S7). To cope with this issue, the date and time of occurrence of all AM ties were also extracted and included in AM occurrence analysis.

#### 4. Statistical characterization of AM series

Let  $X = (x_1, x_2, \dots, x_n)$  represent the series of AM precipitation depths (mm) at the spatiotemporal scale  $(r, d)$  for a generic location, namely, a station or grid box. AM quantiles  $\hat{x}_q$  associated with return periods  $q = 2, 10,$  and 25 years were estimated from the empirical cumulative distribution function (Hazen plotting position; Cunnane 1978). Only 2- and 10-yr return periods were considered for WRF, CMORPH, and for station series shorter than 30 years.

Note that nonstationarity and serial correlation can be expected in extreme precipitation series (Westra et al. 2014; Katz 2013). However, the statistical inference used in this study (e.g., the quantile estimation or the statistical tests presented below) implicitly assumes that data are temporally homogeneous and serially uncorrelated, since assessing the nonrandomness of the available AM series remains challenging.

##### a. Annual and diurnal cycles of AM occurrences

Many approaches have been proposed in the literature to estimate the annual and daily precipitation cycles based on precipitation timing, frequency, and/or intensity (e.g., Dai et al. 1999; Evans and Westra 2012; Cortés-Hernández et al. 2016). In the present study, the annual cycle of the extreme precipitation has been assessed at each location through the relative monthly frequencies

$f_m$  of AM occurrences. If a given AM overlapped two consecutive months, the total AM duration  $d$  was split between these two months. Similarly, the relative hourly frequencies  $f_h$  of AM occurrences were used to evaluate the diurnal cycles for durations  $1 \text{ h} < d < 6 \text{ h}$ . Each AM of duration  $d$  contributed to the frequencies of exactly  $d$  hourly bins of the 24-h histogram and was split if overlapping two consecutive days.

##### b. Spatial and temporal scaling

The temporal scaling of extreme precipitation was assessed through the change in AM quantiles with duration and estimated for each return period using the following regressions:

$$\ln(\hat{x}_{r,d}) = \alpha_r + \beta_r \ln(d) + \varepsilon_r^{(1)}, \quad (1)$$

where  $\hat{x}_{r,d}$  represents the  $q$ -yr quantile estimated at the spatiotemporal scale  $(r, d)$  and  $\varepsilon_r^{(1)}$  is the error term of the regression model. For simplicity, the index  $q$  has been omitted from Eq. (1) and in the remainder of the paper, while the indices  $r$  indicate that the model was estimated for each spatial scale  $r = r_0, r_1, \dots, r_K$ . The Theil–Sen estimator (Sen 1968) was used.

The  $K$  temporal scaling slopes  $\beta_r$  represent the average relative change of  $\hat{x}_{r,d}$  for a relative change in duration, while the intercept  $\alpha_r$  represents an estimate of  $\ln(\hat{x}_{r,1\text{h}})$ . Hence, small  $\beta_r$  values correspond to similar AM distributions over different durations, while  $\beta_r$  close to 1 implies larger variations in precipitation extremes as the duration is changed. Also, for each quantile order, station  $\beta_0$  estimates correspond to the slopes of the depth–duration–frequency (DDF) curves (Burlando and Rosso 1996; Koutsoyiannis et al. 1998), and  $\beta_r$  are analogous to the simple scaling exponents  $\beta_r^{\text{int}} = \beta_r - 1$  that have been widely used for assessing intensity–duration–frequency (IDF) curves and the temporal scaling of AM precipitation intensity distributions (e.g., Menabde et al. 1999; Blanchet et al. 2016).

Previous studies showed that temporal scaling estimates convey synthesized climatological information about the precipitation regimes of the study region (Ceresetti et al. 2010; Innocenti et al. 2017; Casas-Castillo et al. 2018). Moreover, some studies have highlighted that different synoptic regimes typically lead to distinctive scaling exponent values, with principal differences occurring between short (e.g., hourly and subhourly) and long durations (Borga et al. 2005; Eggert et al. 2015). For instance, larger  $\beta_0$  values have usually been observed for duration intervals dominated by intense convective precipitation, while extremes produced by larger-scale meteorological systems generally corresponded to weaker scaling regimes associated with

longer duration intervals (e.g., [Borga et al. 2005](#); [Panthou et al. 2014](#)). In agreement with these considerations, preliminary analyses suggested the existence of two scaling regimes for the available datasets (see Fig. S10). Therefore, durations  $1 \text{ h} \leq d < 6 \text{ h}$ , herein, short durations (SD), and  $6 \text{ h} \leq d \leq 72 \text{ h}$ , herein, long durations (LD), were analyzed separately. Only LD estimation was considered for MSWEP, since only  $d = 3 \text{ h}$  was available for SD.

Results presented in [section 7](#) also suggest that regional averages of  $\beta_r$  values vary linearly with the spatial scale  $r$ . To evaluate the validity of this linear approximation at the local scale, the following spatiotemporal scaling (STS) model was considered at each grid box:

$$\beta_r = h_0 + h_1 r + \varepsilon_r^{(2)}, \quad (2)$$

where  $\varepsilon_r^{(2)}$  is the error term of the linear regression that describes the variation of  $\beta_r$  estimates with  $r$ . The intercept  $h_0$  represents the temporal scaling slope extrapolated at the station scale (i.e.,  $r = 0$ ) from available coarser spatial scales  $r = r_0, r_1, \dots, r_K$ , while  $h_1$  expresses the average change of  $\beta_r$  associated with a unit spatial scale change (i.e., 1 km). Combining Eqs. (2) and (1), and ignoring for simplicity the residual estimate of the  $\varepsilon_r^{(1)}$  and  $\varepsilon_r^{(2)}$  error terms, the  $q$ -yr quantile at the spatiotemporal scale  $(r, d)$  can be expressed as

$$\ln(\hat{x}_{r,d}) = \ln(\hat{x}_{r,1\text{h}}) + (h_0 + h_1 r) \ln(d). \quad (3)$$

Hence, STS parameters relate AM properties across spatial scales and constitute effective metrics to compare the spatiotemporal structure of precipitation extremes reproduced by different gridded datasets. Considering that Eq. (3) reduces to  $\ln(\hat{x}_{0,d}) = \ln(\hat{x}_{0,1\text{h}}) + h_0 \ln(d)$  for  $r = 0$ , a straightforward expression for the spatiotemporal scaling of AM quantiles emerges when writing the difference  $\ln(\hat{x}_{r,d}) - \ln(\hat{x}_{0,d})$  as

$$\ln\left(\frac{\hat{x}_{r,d}}{\hat{x}_{0,d}}\right) = \ln\left(\frac{\hat{x}_{r,1\text{h}}}{\hat{x}_{0,1\text{h}}}\right) + h_1 r \ln(d). \quad (4)$$

The ratios  $(\hat{x}_{r,d}/\hat{x}_{0,d})$  and  $(\hat{x}_{r,1\text{h}}/\hat{x}_{0,1\text{h}})$  between areal and point rainfall quantiles describe the decrease in spatially averaged extremes when reducing the spatial resolution and are referred to as areal reduction factors (ARF) for durations  $d$  and 1 h ([Sivapalan and Blöschl 1998](#); [Svensson and Jones 2010](#)). According to Eq. (4),  $h_1$  thus represents the sensitivity of extreme precipitation ARF to changes in the spatiotemporal scale  $(r, d)$  with respect to the hourly scale and may serve to extrapolate the ARF beyond the empirically available range of  $r$  and  $d$ .

The correlation structure of precipitation and the mechanism driving the extremes may be expected to vary with the weather regime and the geoclimatic characteristics of the region under study, for instance, because of coastal effects or orographic forcing. Accordingly, ARF values have been found to vary with location (e.g., [Omolayo 1993](#); [Asquith and Famiglietti 2000](#)), season (e.g., [Allen and DeGaetano 2005](#)), and in some specific cases with return period (e.g., [Asquith and Famiglietti 2000](#); [Allen and DeGaetano 2005](#)). For similar reasons, ARF values are expected to increase with duration  $d$  (e.g., [NERC 1975](#); [Mineo et al. 2018](#)), as short-duration convective systems and longer duration stratiform events have different typical spatial scales ([Eggert et al. 2015](#)). However, the characteristics of the precipitation datasets (e.g., the density of the recording network or series length) may strongly affect the estimation of the spatial properties of precipitation extremes, ARFs, and their dependence on  $d$  ([Kursinski and Zeng 2006](#); [Svensson and Jones 2010](#)). It is thus crucial to assess the validity of Eqs. (2)–(4) for various datasets, as investigations of the spatiotemporal scaling of extreme precipitation remain scarce, especially for RCM simulations ([Cortés-Hernández et al. 2016](#)). To this end, the statistical significance of the  $h_0$  and  $h_1$  parameters was tested through a linear regression permutation test for Eq. (2) (see section S4.1 of the supplemental material for a detailed description of the test; [Anderson and Robinson 2001](#)).

## 5. Evaluating dataset performances

Two major issues must be considered when assessing climate model performance: the spatiotemporal resolution mismatch between observed and climate model-simulated series ([Chen and Dai 2018](#)) and the choice of adequate performance metrics (e.g., [Chardon et al. 2016](#); [Herold et al. 2017](#)). For instance, classical point-to-grid comparison of RCMs against stations makes it difficult to separate the relative contributions of spatial mismatch and structural model errors, while averaging the performance metrics over coarse-resolution grids may smooth local effects (e.g., [Tapiador et al. 2017](#)). Moreover, any regridding of point (station) or areal (grid box) precipitation series has a strong impact on the estimated extremes, which further depends on the climatological characteristics of the precipitation fields (e.g., their typical spatiotemporal correlation structure), the interpolation method, and the basic features of each specific dataset (e.g., the gauge density for station series; [Tustison et al. 2001](#)). The analysis was thus carried out considering the native grids for both observational and simulated datasets and involved the following two steps:

- (i) Comparison of grid box and station AM characteristics at native resolution: stations were first compared to the other datasets, except CRCM5-LE, using series covering the period 1940–2016 (black rectangle in Fig. 2). Each station was associated with the nearest grid box or discarded if the distance with the nearest grid box was larger than  $\sqrt{2}r$ . The selected grid box–station pairs are herein identified by the station coordinates (latitude–longitude) and referred to as L1 locations.
- (ii) AM spatiotemporal scaling for gridded datasets: the STS models were estimated and evaluated at each grid box for all gridded datasets. To this end, each grid box at native resolution  $r_0$  was associated with the overlapping grid boxes at each coarser spatial scale  $r_k$ . CRCM5-LE estimates over the 1980–2015 period (gray rectangle in Fig. 2) were also considered and compared to ERA-CRCM5 over the entire spatial domain.

Point and areal estimates of AM quantiles were first compared at the native dataset resolution using the following relative difference:

$$B_d = \frac{\hat{x}_{0,d}^{(s)} - \hat{x}_{r_0,d}^{(g)}}{\hat{x}_{0,d}^{(s)}}, \quad (5)$$

where  $\hat{x}_{0,d}^{(s)}$  and  $\hat{x}_{r_0,d}^{(g)}$  represent the  $q$ -yr return period quantiles estimated, respectively, at station  $s$  and its nearest grid box  $g$  for any relevant duration  $d$ . A permutation test based on 5000 resamples was used to estimate the statistical significance of  $B_d$  values for each station–grid box pair (Good 2013). Nonzero  $B_d$  values may result from biases in gridded estimates, station measurement errors, and representativeness errors, namely, errors that derive from the spatial resolution mismatch between the two datasets (Tustison et al. 2001).

The Perkins skill score (PSS; Perkins et al. 2007) was used to compare the annual and diurnal cycles of AM occurrences between the stations and gridded datasets. For any specific duration and location, the PSS values  $S_{M,d}$  for the annual cycles were defined as the common area between the station and gridbox histograms of the monthly frequencies of the AM occurrences ( $S_{M,d}$  close to 0 means a small overlap of the station and gridbox annual cycles, while  $S_{M,d} = 1$  indicates their perfect match). Similarly, the PSS values  $S_{H,d}$  were used to measure the overlap between station and gridded dataset diurnal cycles based on the hourly frequencies of AM occurrences. The computational details for the estimation of  $S_{M,d}$  and  $S_{H,d}$  are reported in section S3.2 of the supplemental material.

## 6. Gridded dataset and station AM statistics at native resolution

The spatial distributions of AM quantiles were found to be consistent across the datasets, with clear southwest–northeast gradients across the study domain (see Figs. S1–S4 for some examples). Local effects were observed for longer durations around the Great Lakes and the northern Atlantic coast for the two RCMs, while clusters of spuriously high AM values were observed in the northern areas for CMORPH.

### a. Relative differences in AM quantiles

Figures 3a–d shows, for each duration ( $x$  axis), the distributions over L1 locations (i.e., station–grid box pairs) of the relative differences  $B_d$  between the station and gridded dataset AM quantiles at their native resolution. For each return period, the gridded datasets underestimate station quantiles for the shortest durations, whereas they generally overestimated station extremes for durations longer than few hours, as the  $B_d < 0$  moved from the positive to negative half plane when increasing  $d$  from 1 h to longer durations.

ERA-CRCM5 and WRF show good agreement with stations for  $d = 1$  h, with  $|B_{1h}| \leq 0.05$  for more than 50% of L1 locations for 2-yr quantiles (Figs. 3a,b, first column). Besides a large spread across positive and negative values, the WRF median  $B_d$  is close to zero for all durations and return periods (solid black line in Fig. 3b). Conversely, substantial negative differences were estimated between ERA-CRCM5 and station quantiles for  $d \geq 3$  h and all return periods, with median  $B_d \leq 0.1$  and larger fractions of L1 locations displaying statistically significant differences compared to the other datasets ( $f_{H_1}$  red curve in Fig. 3).

The relative differences estimated for CMORPH and MSWEP are large for durations close to their temporal resolution but rapidly decrease close to 0 as  $d$  increases and are generally lower than ERA-CRCM5 and WRF  $B_d$  values for  $d \geq 12$  h (Figs. 3c,d). CMORPH displays substantial underestimations of the station quantiles at the hourly scale with median  $B_d \geq 0.2$  for both 2- and 10-yr return periods.

Compared to the other gridded datasets, MSWEP displays the highest relative underestimations of the station AM quantiles for  $d \leq 6$  h, whereas it presents the smallest  $B_d$  values and the lowest  $f_{H_1}$  fractions at daily and longer durations (Fig. 3d).

Figure 4 displays the spatial distributions of the relative differences estimated between station and gridbox 10-yr quantiles for  $d = 1, 3,$  and 24 h. For the two RCMs, significant positive  $B_d$  values were found in the south and west interior part of the domain for  $d \leq 3$  h



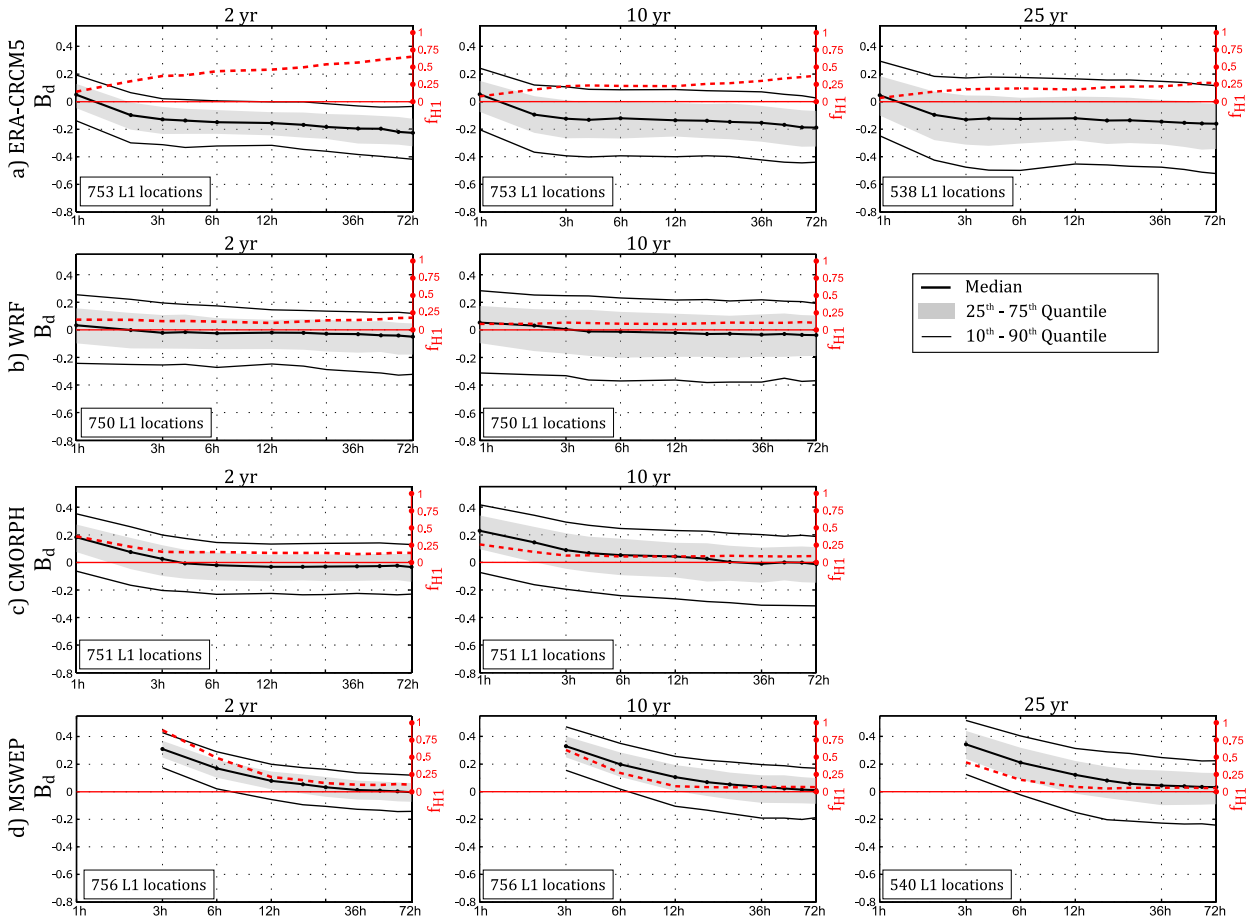


FIG. 3. Distribution over L1 locations of the relative differences  $B_d$  between station and gridbox AM quantiles at the native dataset resolution for (a) ERA-CRCM5, (b) WRF, (c) CMORPH, and (d) MSWEP. The term  $f_{H1}$  (red dashed curves and right y axis) represents the fractions of L1 locations with statistically significant  $B_d$  values. The total number of L1 available locations is indicated at the bottom-left corner of each panel.

(Figs. 4a,b). ERA-CRCM5 overestimates hourly station quantiles in eastern and central areas and negative  $B_d$  values gradually extend to the rest of the domain as  $d$  approaches 24 h (Fig. 4a), a result consistent with the wet bias found for ERA-CRCM5 mean summer precipitation by Leduc et al. (2019). WRF and CMORPH display a clear regional structure with differences between west (positive  $B_d$  values) and northeast areas (negative and generally significant  $B_d$  values) (Figs. 4b,c). For WRF, this result is consistent with the large underestimation of hourly summer extremes found in the central United States, as well as with the overrepresentation of strong mesoscale convective systems simulated over the Atlantic regions in northeastern United States (Prein et al. 2017, 2019). For  $d_0 = 3$  h, MSWEP displays positive  $B_d$  values throughout the entire domain, while at the daily scale a small number of L1 locations show the east–west pattern already observed for the other datasets. The negative  $B_d$  observed for all datasets in the

Appalachian region and eastern areas may partly result from rain gauge underestimation of daily extremes due, for instance, to wind undercatch.

Similar results were found for 2- and 25-yr AM quantiles (e.g., Figs. S5 and S6). Moreover, using spatially aggregated measures of gridded dataset performances (e.g., the normalized RMSE between station and gridbox quantiles; not shown), ERA-CRCM5 generally outperformed the other datasets at  $d_0$ , while the lowest aggregated errors were found for WRF for  $3 \text{ h} \leq d \leq 6 \text{ h}$  and for MSWEP for  $d \geq 12 \text{ h}$ .

#### b. Differences in the annual and diurnal cycles of AM occurrences

Figure 5 compares station and gridded dataset annual cycles estimated at the regional scale. As expected for midlatitudes, station AM mostly occur during summer for short durations, while daily extremes are more evenly distributed over the whole year (Figs. 5a,b).

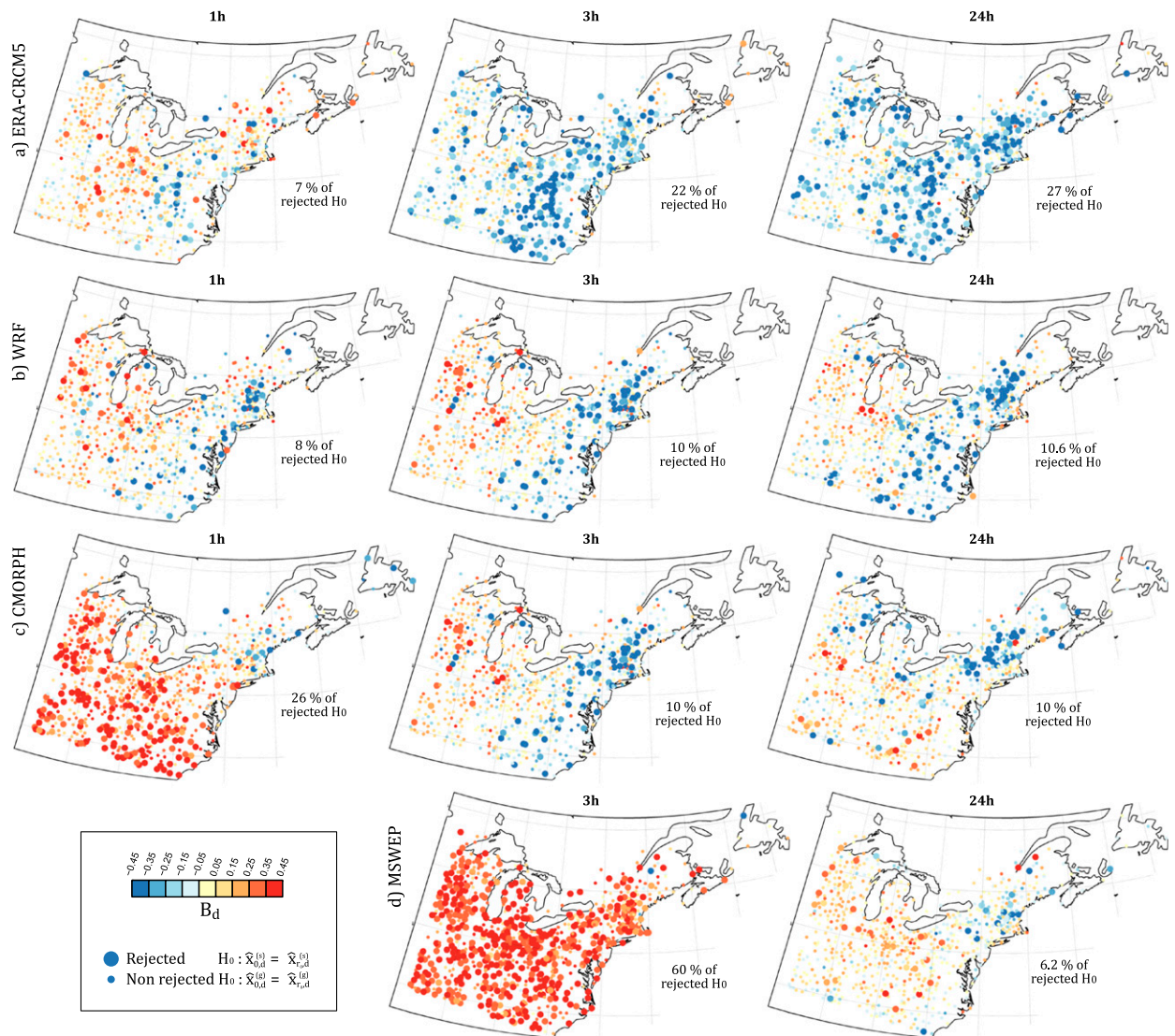


FIG. 4. Spatial distribution over L1 locations of the relative differences  $B_d$  between station and gridbox AM quantiles ( $q = 10$  years) at the native dataset resolution for (a) ERA-CRCM5, (b) WRF, (c) CMORPH, and (d) MSWEP. Smaller points represent locations where  $B_d$  was not found to be significantly different from 0 at the 0.05 significance level.

Station AM annual cycles were well reproduced by the two RCMs, although they slightly underestimate winter and fall AM frequencies. ERA-CRCM5 displays the smallest differences with stations (i.e., highest  $\overline{S_{M,d}}$  values) for  $d \leq 6$  h (Fig. 5c). Compared to other datasets, WRF simulated earlier summer peaks for all  $d$  (see the examples in Fig. S8), which is likely related to the short length of the WRF simulation, as similar results were found when restricting station annual cycle estimation to the 1995–2016 period (not shown). Low  $\overline{S_M}$  CMORPH values are due to the overestimation of AM frequencies during winter and fall and the consequent underestimation of AM occurrence frequency during summer. This result is consistent with reported

CMORPH biases and uncertainties in cold conditions (Trenberth et al. 2017). MSWEP adequately reproduced station annual cycles, resulting in the highest  $\overline{S_{M,d}}$  for  $6 \text{ h} < d \leq 24 \text{ h}$  and very good MSWEP performance at daily and longer durations.

The diurnal cycles of hourly AM are presented in Fig. 6a for all datasets except MSWEP (due to its coarse temporal resolution  $d_0 = 3$  h). Cycles from all datasets display the afternoon peak that is typical for convection-driven extremes in response to the diurnal heating of the land surface in midlatitude inland climates (Dai et al. 1999; Rasmussen et al. 2019).

ERA-CRCM5 well reproduced the shape of the observed diurnal cycle of hourly AM, although its peak was

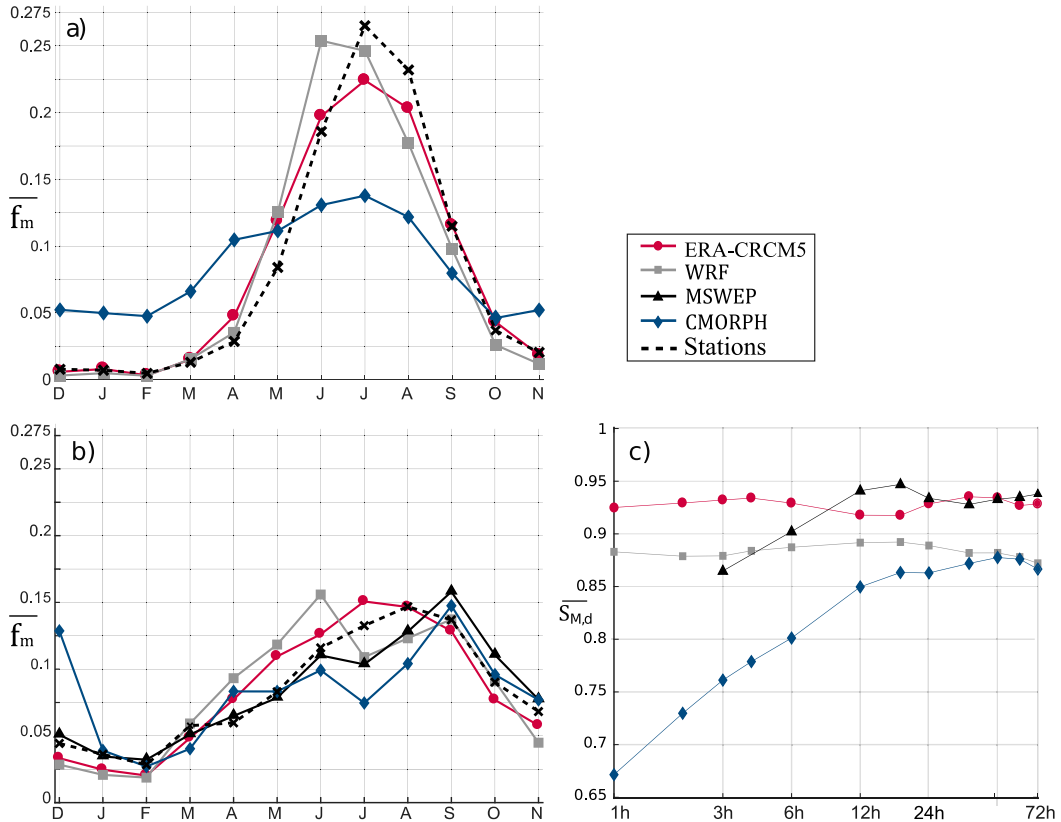


FIG. 5. Annual cycles of the date AM occurrence at the native dataset resolution: average over L1 locations of the monthly frequencies of AM occurrence for (a) 1 h and (b) 24 h; (c) Perkins Skill Score (PSS) between station and gridded dataset annual cycles averaged over L1 locations.

2 h later than the stations, contrary to what may be expected for parameterized RCMs, where the convection scheme tends to trigger precipitation too early (Prein et al. 2015). For hourly AM, WRF displays the peak 1 h later than the stations and overestimated late afternoon and early night extreme frequencies. For  $1\text{ h} < d \leq 4\text{ h}$ , station diurnal cycles were flatter than simulated ones and recorded AM more frequently occurred between midnight and 1000 (UTC - 5) (e.g., Fig. S9c). The two RCMs display comparable performance for all considered durations (Fig. 6b), suggesting that the delay in diurnal peaks observed for both RCMs cannot only be attributed to the different moist convection representation. While CMORPH slightly outperformed the RCMs in terms of  $\bar{S}_{M,d}$ , it displays noisy diurnal cycle estimates for all considered durations (e.g., Fig. S9d).

### 7. Spatiotemporal scaling of AM rainfall

Figure 7 presents for each return period and dataset scaling slopes [Eq. (1)] estimated for  $1\text{ h} \leq d < 6\text{ h}$  (SD) and  $6\text{ h} \leq d \leq 72\text{ h}$  (LD). As expected, SD extremes exhibit stronger temporal scaling than LD extremes and

gridded datasets have larger scaling slope values than stations (note the different axis scale in Figs. 7a,b). Moreover,  $\bar{\beta}_r$  values increase with the spatial scale and SD increases with  $r$  were more pronounced than for LD. Accordingly, the differences between areal and point-scale scaling slopes were more pronounced for SD (Fig. 7a) while  $\bar{\beta}_{r_0}$  values were generally within the station interquartile ranges (Fig. 7b). As an exception, MSWEP displays high LD scaling slope values, possibly due to its coarse temporal resolution. These results are consistent with the spatial smoothing expected for AM quantiles for different durations. Since short duration extremes are more likely associated with localized convective events, the spatial aggregation of precipitation series produces larger  $\beta_r$  values as  $r$  increases and the effects of this spatial aggregation on the temporal scaling is more pronounced for short duration extremes. WRF displays the largest increase of  $\bar{\beta}_r$  with  $r$  for SD (Fig. 7a), possibly due to its fine native resolution, which particularly affects the properties of finer-scale SD AM and/or more extreme events. The differences between WRF and the datasets are less pronounced for LD (Fig. 7b).

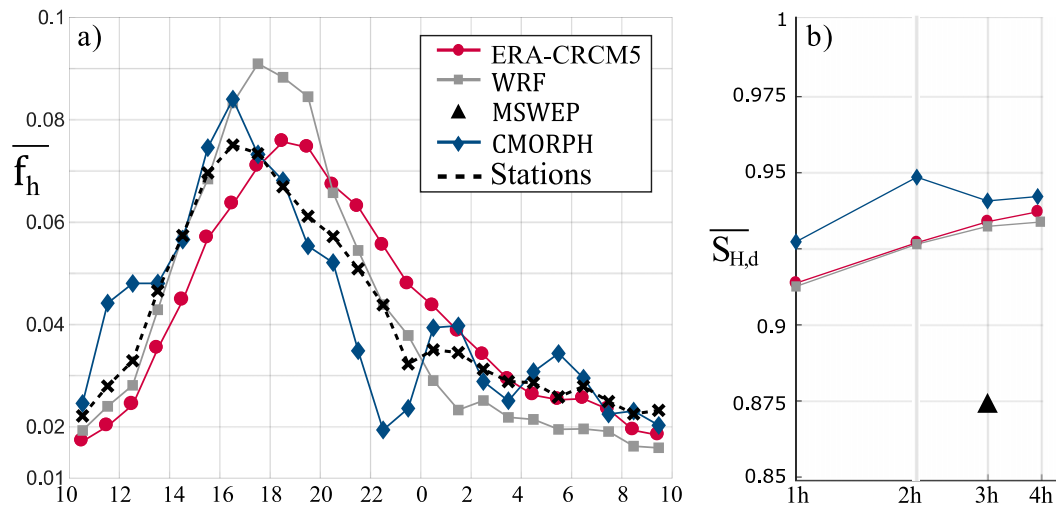


FIG. 6. Diurnal cycles of the time of AM occurrence at the native dataset resolution: (a) average over L1 locations of the hourly frequencies of hourly AM occurrence and (b) Perkins Skill Score (PSS) between station and gridded dataset diurnal cycles averaged over L1 locations. Only  $d_0 = 3$  h is considered for MSWEP.

Interestingly, the increase of the temporal scaling slopes with the spatial scale is nearly linear, as indicated in Fig. 7 by the good fit of the dotted lines to the mean scaling slope values. To further investigate this result, the  $\beta_r$  estimated at each L1 location were linearly regressed on  $r$  according to Eq. (2). The validity of this STS relationship at local scale was then assessed with a permutation test (see section S4.1 and Fig. S11). For the two RCMs, STS parameters were statistically significant for most L1 locations, despite worse linear fit and larger statistical uncertainty for the LD STS estimation (Fig. S11b). For CMORPH and MSWEP, most L1 locations rejected the use of the STS model at typical significance levels, meaning that little information about point-scale temporal scaling can be inferred from the use of multiple spatial scales. A possible explanation is that both CMORPH and MSWEP used postprocessing methods that involve the computation of precipitation statistics and the adjustment of the corresponding series at several space and time aggregations. For instance, CMORPH was bias corrected in various steps at the  $(0.25^\circ, 24\text{ h})$  and coarser resolutions (Xie et al. 2017), while the merging of multiple data sources for MSWEP necessarily involved combining data at different spatiotemporal resolutions (Beck et al. 2019b).

## 8. AM statistics for the CRCM5-LE

The proposed STS model was evaluated using the CRCM5-LE, considering ensemble mean temporal scaling slopes. The impact of lateral boundary conditions was also investigated through the comparison of

ERA-CRCM5 and CRCM5-LE estimates. The main conclusions of this analysis are summarized in Table 3.

### CRCM5-LE spatiotemporal scaling

The estimation of  $\beta_r$  was first carried out for each member and spatial scale  $r$ . The ensemble means of  $\beta_r$  parameters were then computed for subgroups of 5, 10, and 50 members and then used to estimate  $h_0$  and  $h_1$ . Subgroups of 5 and 10 members were constructed considering members 1–5, 6–10, ..., 46–50 and 1–10, 11–19, ..., 41–50, respectively.

Figure 8 shows the permutation test results for the STS regressions estimated for the first member, all members, and for the first 5- and 10-member subgroups. The fraction of grid boxes with significant  $h_1$  increases with the number of members, showing that local estimation uncertainty is considerably reduced when using larger ensembles, especially for SD and high return periods. For instance, STS parameters were significant for less than 50% of grid boxes for SD when only the first member was considered for  $q = 10$  years, while more than 90% of grid boxes had significant  $h_1$  when averaging  $\beta_r$  over 50 members before applying Eq. (2) (Fig. 8a). Moreover, the 50-member estimation of Eq. (2) generally yielded a good linear fit for all return periods (e.g.,  $R^2 > 0.9$  for more than 94% and 74% of grid boxes for SD and LD, respectively; not shown).

Interesting spatial patterns can be observed for  $h_0$  and  $h_1$  when the 50-member CRCM5 ensemble is considered (Figs. 9 and 10). For SD, lower  $h_0$  values can be observed for the interior of the domain and east of the Great Lakes, while the highest scaling regimes (e.g.,  $h_0 \geq 0.45$ )

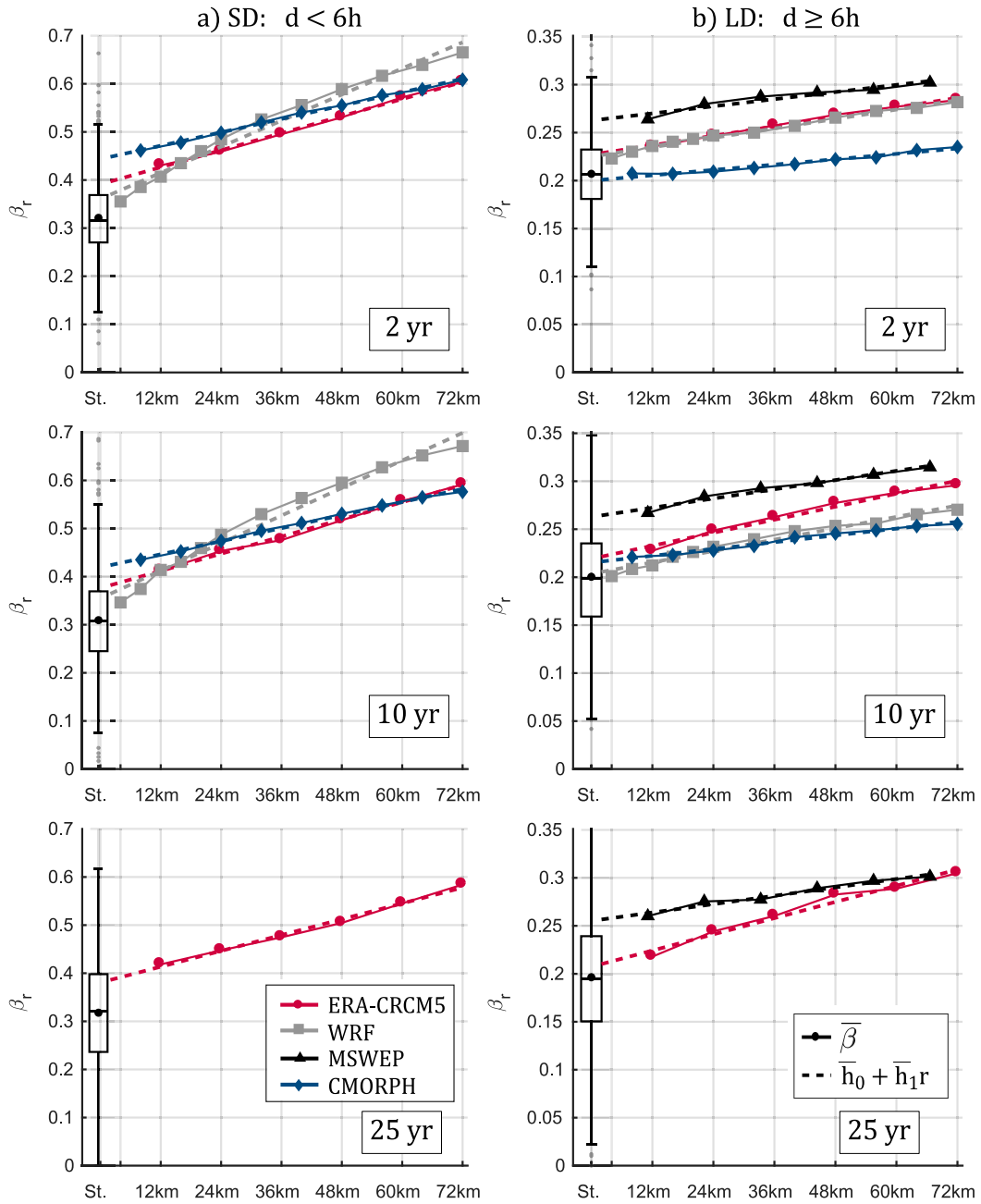


FIG. 7. Mean over L1 locations  $\bar{\beta}_r$  of the temporal scaling slopes displayed as a function of the spatial scale ( $x$  axis) for (a) SD ( $1\text{ h} \leq d < 6\text{ h}$ ) and (b) LD ( $6\text{ h} \leq d \leq 72\text{ h}$ ). Boxplots show the distribution over the L1 locations of corresponding  $\beta_0$  station values. Dotted lines represent the linear fit of  $\bar{\beta}_r$  with spatial scales  $r$ . For readability reasons some outliers are not displayed in station boxplots.

were observed at the southern and northern borders (Fig. 9a). Similar patterns were found for LD, with strong scaling regimes in southwest and northeast regions (Fig. 9b). Very small  $h_0$  values were also estimated in the northwest, between the Great Lakes and the southern coast of the Hudson’s Bay. Similar spatial patterns for stations  $h_0$  values were found by Innocenti

et al. (2017) in an extensive study of rainfall intensity scaling exponents over North America.

Evident northwest–southeast gradients can be observed for  $h_1$  values, with the highest values found over the Appalachian Mountains, especially for SD (Fig. 10a). This indicates that AM temporal scaling slopes in southern regions are more sensitive to changes in the

TABLE 3. Main results from the comparison between ERA-CRCM5 and CRCM5-LE.

Statistics	Main results
Quantiles	Important latitudinal gradient for the relative differences of the estimates: CRCM5-LE quantiles are smaller (larger) than ERA-CRCM5 in the south (northeast). Similar spatial distributions for all $d$ (e.g., Fig. S12)
Annual cycle	More AM ERA-CRCM5 occur in summer compared to CRCM5-LE for $d \leq 24$ h; for $d > 24$ h, CRCM5-LE simulated more AM than ERA-CRCM5 from November to May. Results vary according to the geographical region (e.g., Fig. S13).
Diurnal cycle	Minor differences between ERA-CRCM5 and CRCM5-LE suggesting that large-scale dynamics and boundary conditions have weak influence on diurnal cycles (the triggering of short-duration AM mainly depends on processes simulating local storm dynamics)
Temporal scaling	Minor differences for SD, while CRCM5-LE showed weaker LD scaling than ERA-CRCM5. The scaling of short duration extremes is weakly influenced by the lateral boundary conditions used to drive the simulations.

spatial scale than in northern areas, or, equivalently, that ARFs vary more strongly with the spatiotemporal scale [see Eq. (4)]. One explanation for these patterns is that the characteristics and the spatial extent of weather systems producing short duration extremes vary with latitude. Conversely, a relatively larger spatial extent is expected for weather systems generating LD extremes, as well as weaker changes in weather systems characteristics over the CRCM5 spatial domain. This

would also explain the smoother latitudinal gradients and weaker response to topography observed for LD  $h_1$  values.

For  $h_1$ , the differences between western regions and the southeast portion of the domain could also be related to the large-scale circulation patterns that strongly affect the spatiotemporal structure of precipitation extremes in these areas (Touma et al. 2018).

As a final note, although the dependence of  $h_0$  and  $h_1$  on the return period should be further investigated, some preliminary analyses showed that  $h_0$  and  $h_1$  decrease slightly with increasing return period in areas where both STS parameters are small (see Figs. S14 and S15).

## 9. Summary and discussion

An accurate characterization of the spatiotemporal structure of subdaily and daily precipitation is crucial for the development of future climate projections and to evaluate the ability of RCMs to adequately simulate intense precipitation events. The primary aim of this study was to compare the statistical properties of daily and subdaily annual maximum precipitation (quantiles, seasonal and diurnal cycles, and spatiotemporal scaling) between various observed and simulated datasets. The performance of two ERA-Interim-driven regional simulations from the CRCM5 and the convection-permitting WRF Model was evaluated against station observations and compared to the bias-corrected satellite CMORPH and the multisource MSWEP dataset.

At their native spatial resolutions both RCMs showed good agreement with station records for short duration extremes, with few statistically significant differences between RCM and station quantiles found in the eastern part of the domain. For daily and longer durations,

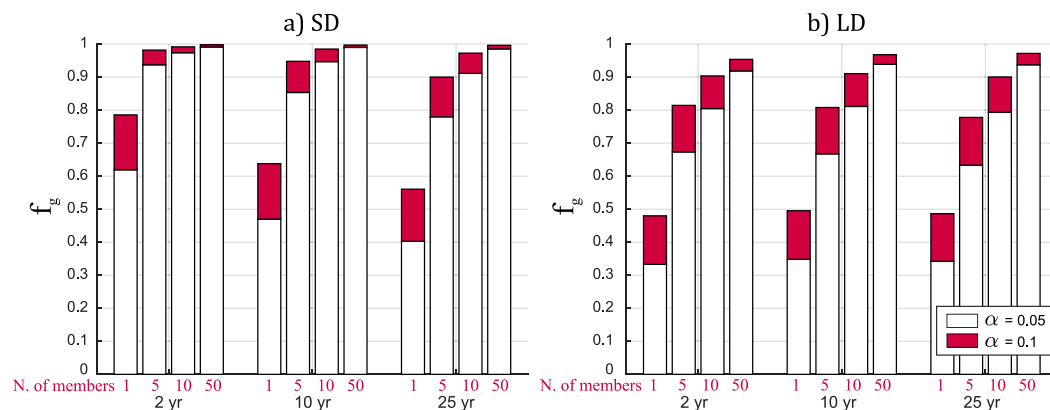


FIG. 8. Permutation test for the STS model [Eq. (2)] for subgroups of CRCM5-LE members: fraction  $f_g$  of grid boxes with significant  $h_1$  (significance levels  $\alpha = 0.05$  and  $\alpha = 0.1$ ) for each return period for (a) SD and (b) LD.

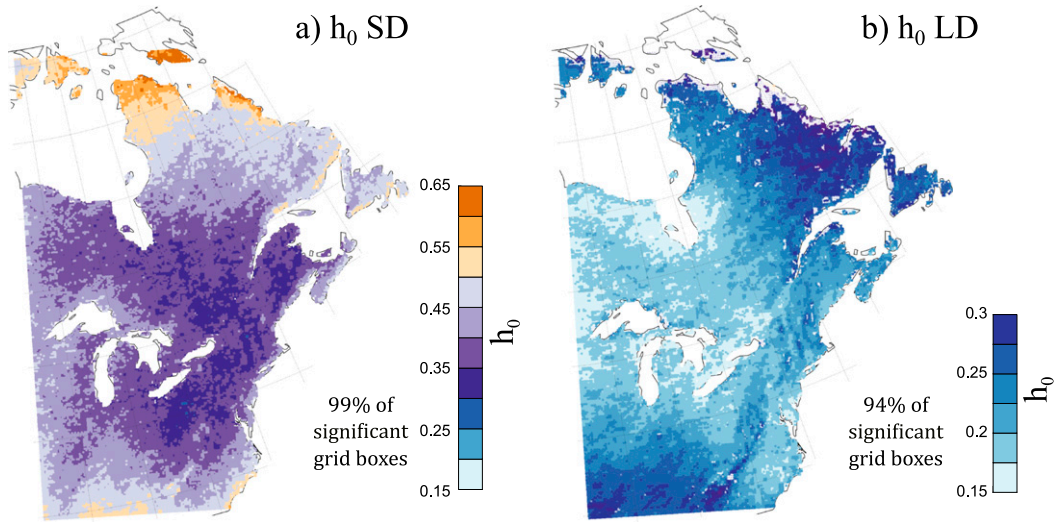


FIG. 9. Spatial distribution of the extrapolated temporal scaling slopes  $h_0$  for the AM quantiles ( $q = 10$  years) quantiles estimated using the 50 CRCM5-LE members for (a) SD and (b) LD. Only grid boxes with statistically significant STS models (0.05 significance level) are displayed.

ERA-CRCM5 significantly overestimated the extremes at many stations (e.g., for more than 50% of stations for 2-yr quantiles). Spatial patterns of the estimated relative difference are consistent with the CRCM5 and WRF bias in mean summer precipitation and hourly precipitation found in previous studies (Leduc et al. 2019; Liu et al. 2017; Prein et al. 2019).

Small differences were observed between the station and ERA-CRCM5 annual cycles for all considered durations, while WRF advanced the summer frequency

peak of subdaily extremes. Both RCMs slightly underestimated the winter frequencies of annual maxima. Interestingly, ERA-CRCM5 and WRF show comparable performances for the diurnal cycles, with later afternoon peaks for the frequency of hourly extremes compared to observations, and extreme occurrences clustered around late afternoon and early night values for longer durations.

Median underestimation up to 30% of station quantiles was observed for CMORPH and MSWEP for

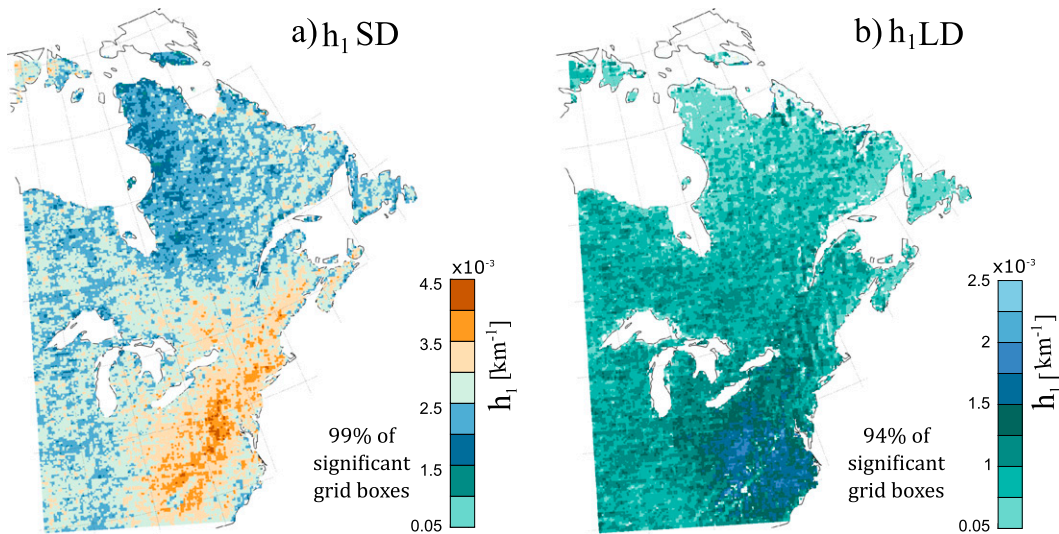


FIG. 10. Spatial distribution of the spatiotemporal scaling parameters  $h_1$  for the AM quantiles ( $q = 10$  years) estimated using the 50 CRCM5-LE members for (a) SD and (b) LD. Only grid boxes with statistically significant STS models (0.05 significance level) are displayed.

subdaily extremes, whereas MSWEP shows the smallest differences with station quantiles and adequately reproduces the observed annual cycles for  $d \geq 24$  h. However, the coarse temporal resolution of MSWEP (i.e.,  $d_0 = 3$  h) prevents any meaningful assessment of subdaily extreme characteristics.

The presence of some clusters of spuriously high quantiles in northern regions and the overestimation of extreme frequencies during fall and winter months confirms previously identified issues concerning CMORPH performance in cold regions and during cold seasons (Xie and Joyce 2014; Trenberth et al. 2017).

The spatiotemporal scaling properties of extreme precipitation quantiles were then analyzed and compared among datasets over L1 locations. The following three major results were found:

- The temporal scaling regimes are stronger for short durations (SD,  $d < 6$  h) than longer durations (LD,  $d \geq 6$  h). Hence, changes in rainfall quantiles across durations are smaller for LD, which is consistent with the fact that longer duration extremes are generated by spatially more homogeneous weather systems.
- The temporal scaling slopes increase with the spatial scale and SD increases are larger than for LD. Short duration extremes are more likely associated with intense localized convective systems and thus more sensitive to changes in the spatial resolution. Accordingly, it is interesting to note that ERA-CRCM5, WRF, and CMORPH display similar temporal scaling slopes at the smallest spatiotemporal scales but the increase of temporal scaling slopes with  $r$  is much larger for WRF, likely because of its finer native spatial resolution.
- The increase of temporal scaling slopes with the spatial scale is approximately linear for all return periods and datasets.

Accordingly, two parameters can be used at each location to describe the changes in extreme quantiles with the observational spatiotemporal scale ( $r, d$ ): the extrapolated temporal scaling slope  $h_0$  and the STS parameter  $h_1$ . The STS linear model [Eq. (3)] was found to be appropriate at most locations for SD 2-yr quantiles, suggesting the possibility of approximating the point-scale temporal scaling through the  $h_0$  values.

However, the local estimation of the STS model is highly uncertain for longer return period and longer duration extremes. For MSWEP and CMORPH, in particular, few locations adequately fit the STS relationship with significant  $h_1$  values. For these datasets, the use of postprocessing methods (e.g., bias correction applied at coarser scale than their native spatial and temporal resolution) may have altered the

spatiotemporal structure of precipitation extremes, explaining the high uncertainty of the spatiotemporal scaling estimates. However, this explanation remains a hypothesis and needs further investigation.

The second objective was to examine whether there is evidence to support the empirical linearity of the temporal scaling slopes at the local scale (e.g., for a grid box) when sampling errors associated with short series and local variability are reduced. To this end, the temporal scaling estimates from various CRCM5-LE members were averaged to robustly estimate  $h_0$  and  $h_1$  at each model grid box. Increasing the ensemble size improved the fit of STS models, and led to increasing fractions of grid boxes with significant  $h_1$  as the sample size (members) increased. While the limited length of records and historical simulations does not provide a well-constrained characterization of the extreme scaling, the CRCM5-LE allows a more robust estimation, which highlights the benefit of using a large ensemble of simulations (Thompson et al. 2017; Li et al. 2019).

Spatially, the strongest temporal scaling was observed in the north and in southwestern continental areas of the CRCM5-LE domain, while the strongest spatiotemporal scaling was found in mountainous regions. The clear spatial distribution of STS parameters can be associated with large-scale circulation patterns and topographic features that define the precipitation regimes over the study region.

## 10. Conclusions and perspectives

Although it remains difficult to definitively recommend the use of any specific dataset considered in our study, our results provide guidance on their relative performance in reproducing the daily and subdaily extreme precipitation characteristics. For short durations, both the CRCM5 and WRF well reproduced the properties of the extremes estimated at stations, while the two observation-based datasets showed good performance for approximating daily and longer duration extremes. Equally important, limited information about the spatiotemporal structure of extreme precipitation can be obtained from the CMORPH and MSWEP datasets. Their adequacy for the assessment of precipitation extremes should be particularly questioned for northern latitudes, while better performance of both gridded datasets may be expected for other regions and climate zones. Guidance on the downscaling of point-scale extreme precipitation properties from simulated series is offered by the proposed STS relationships, which can be used to consistently compare the spatiotemporal statistical structure of the precipitation extremes estimated from different datasets, regions, or periods.



Future work should consider the following issues. First, the STS model should be more thoroughly evaluated over regions with different climatological characteristics to assess the regional variability and improve the empirical evidence of the proposed relationships. Considering other climate model simulations and ensembles would also allow understanding the impacts of climate model-related uncertainty on our results. Finally, repeating the analysis for regions with very high network density would be important, as the characteristics of the available station dataset (e.g., the network density in Canada or the coarse measurement resolution of U.S. station records) may have a strong impact on AM estimates and their scaling properties (e.g., Blanchet et al. 2016).

Second, the analysis of seasonal AM should be considered to evaluate the biases and interannual variability that may affect the precipitation extreme estimation for the different datasets, and the seasonality of the STS relationships. Climate model improvements and increasing resolutions are also expected to yield a better representation of extreme properties for the range of spatiotemporal scales considered in the present study (Coppola et al. 2019). Finally, similar investigations using alternative definitions of extreme, such as event-based heavy rainfall, and projected series under future climate conditions should be carried out. In this way, the statistical scaling properties could be better related to the typical dynamics of extreme precipitation events, which can help the understanding of the main processes driving the future evolution of extreme precipitations.

*Acknowledgments.* Silvia Innocenti's scholarship was partly provided by the Consortium OURANOS. Financial support for this project was also provided by the Discovery Grants Program from the Natural Science and Engineering Research Council of Canada. The ClimEx project was financed by the Bavarian State Ministry for the Environment and Consumer Protection. CRCM5 simulations were run by SuperMUC supercomputer at Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities and funded via the Gauss Centre for Supercomputing (GCS) by the German Federal Ministry of Education and Research and the Bavarian State Ministry of Education, Science and the Arts. The CRCM5 was developed by the ESCER centre of Université du Québec à Montréal (UQAM) in collaboration with Environment and Climate Change Canada. The Canadian Centre for Climate Modelling and Analysis (CCCma) executed and made available the CanESM2-LE for the ClimEx project. We acknowledge the Canadian Sea Ice and

Snow Evolution Network for proposing the simulations. The authors also thank Jeremy Panthou and Dikra Khedhaouria for useful discussions, and all data providers who kindly give access to their datasets.

## REFERENCES

- Allen, R. J., and A. T. DeGaetano, 2005: Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts. *J. Hydrol.*, **315**, 203–219, <https://doi.org/10.1016/j.jhydrol.2005.03.028>.
- Anderson, M. J., and J. Robinson, 2001: Permutation tests for linear models. *Aust. N. Z. J. Stat.*, **43**, 75–88, <https://doi.org/10.1111/1467-842X.00156>.
- Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, <https://doi.org/10.1029/2010GL046270>.
- Asquith, W. H., and J. S. Famiglietti, 2000: Precipitation areal-reduction factor estimation using an annual-maxima centered approach. *J. Hydrol.*, **230**, 55–69, [https://doi.org/10.1016/S0022-1694\(00\)00170-0](https://doi.org/10.1016/S0022-1694(00)00170-0).
- Ban, N., J. Schmidli, and C. Schar, 2015: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophys. Res. Lett.*, **42**, 1165–1172, <https://doi.org/10.1002/2014GL062588>.
- Beck, H. E., and Coauthors, 2017: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.*, **21**, 6201–6217, <https://doi.org/10.5194/hess-21-6201-2017>.
- , and Coauthors, 2019a: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sci.*, **23**, 207–224, <https://doi.org/10.5194/hess-23-207-2019>.
- , E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler, 2019b: MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Amer. Meteor. Soc.*, **100**, 473–500, <https://doi.org/10.1175/BAMS-D-17-0138.1>.
- Blanchet, J., D. Ceresetti, G. Molinié, and J. D. Creutin, 2016: A regional GEV scale-invariant framework for intensity-duration-frequency analysis. *J. Hydrol.*, **540**, 82–95, <https://doi.org/10.1016/j.jhydrol.2016.06.007>.
- Borga, M., C. Vezzani, and G. D. Fontana, 2005: Regional rainfall depth–duration–frequency equations for an Alpine region. *Nat. Hazards*, **36**, 221–235, <https://doi.org/10.1007/s11069-004-4550-y>.
- Burlando, P., and R. Rosso, 1996: Scaling and multiscaling models of DDF for storm precipitations. *J. Hydrol.*, **187**, 45–64, [https://doi.org/10.1016/S0022-1694\(96\)03086-7](https://doi.org/10.1016/S0022-1694(96)03086-7).
- Casas-Castillo, M. C., A. Llabrés-Brustenga, A. Rius, R. Rodríguez-Solà, and X. Navarro, 2018: A single scaling parameter as a first approximation to describe the rainfall pattern of a place: Application on Catalonia. *Acta Geophys.*, **66**, 415–424, <https://doi.org/10.1007/s11600-018-0122-5>.
- Cavicchia, L., and Coauthors, 2018: Mediterranean extreme precipitation: A multi-model assessment. *Climate Dyn.*, **51**, 901–913, <https://doi.org/10.1007/s00382-016-3245-x>.
- Ceresetti, D., G. Molinié, and J. D. Creutin, 2010: Scaling properties of heavy rainfall at short duration: A regional analysis. *Water Resour. Res.*, **46**, W09531, <https://doi.org/10.1029/2009WR008603>.

- Chardon, J., A.-C. Favre, and B. Hingray, 2016: Effects of spatial aggregation on the accuracy of statistically downscaled precipitation predictions. *J. Hydrometeorol.*, **17**, 1561–1578, <https://doi.org/10.1175/JHM-D-15-0031.1>.
- Chen, D., and A. Dai, 2018: Dependence of estimated precipitation frequency and intensity on data resolution. *Climate Dyn.*, **50**, 3625–3647, <https://doi.org/10.1007/s00382-017-3830-7>.
- Chen, M., W. Shi, P. Xie, V. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, <https://doi.org/10.1029/2007JD009132>.
- Coppola, E., and Coauthors, 2019: A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. *Climate Dyn.*, <https://doi.org/10.1007/s00382-018-4521-8>, in press.
- Cortés-Hernández, V. E., F. Zheng, J. Evans, M. Lambert, A. Sharma, and S. Westra, 2016: Evaluating regional climate models for simulating sub-daily rainfall extremes. *Climate Dyn.*, **47**, 1613–1628, <https://doi.org/10.1007/s00382-015-2923-4>.
- Cunnane, C., 1978: Unbiased plotting positions—A review. *J. Hydrol.*, **37**, 205–222, [https://doi.org/10.1016/0022-1694\(78\)90017-3](https://doi.org/10.1016/0022-1694(78)90017-3).
- Dai, A., F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402, <https://doi.org/10.1029/98JD02720>.
- , R. M. Rasmussen, C. Liu, K. Ikeda, and A. F. Prein, 2019: A new mechanism for warm-season precipitation response to global warming based on convection-permitting simulations. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-3787-6>, in press.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Diaconescu, E. P., P. Gachon, R. Laprise, and J. F. Scinocca, 2016: Evaluation of precipitation indices over North America from various configurations of regional climate models. *Atmos.–Ocean*, **54**, 418–439, <https://doi.org/10.1080/07055900.2016.1185005>.
- Dwyer, J. G., and P. A. O’Gorman, 2017: Changing duration and spatial extent of midlatitude precipitation extremes across different climates. *Geophys. Res. Lett.*, **44**, 5863–5871, <https://doi.org/10.1002/2017GL072855>.
- ECDC, 2016: Hourly Canadian precipitation data (HCPD) and maximum daily precipitation data (MDPD), Historical Climate Data Canada; editing status 2016-08-09. Environment Climate Change Canada, Registry of Research Data Repositories, accessed 9 November 2016, <http://doi.org/10.17616/R3N012>.
- Eggert, B., P. Berg, J. O. Haerter, D. Jacob, and C. Moseley, 2015: Temporal and spatial scaling impacts on extreme precipitation. *Atmos. Chem. Phys.*, **15**, 5957–5971, <https://doi.org/10.5194/acp-15-5957-2015>.
- Evans, J. P., and S. Westra, 2012: Investigating the mechanisms of diurnal rainfall variability using a regional climate model. *J. Climate*, **25**, 7232–7247, <https://doi.org/10.1175/JCLI-D-11-00616.1>.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, Cambridge University Press, 741–866.
- Fosser, G., S. Khodayar, and P. Berg, 2017: Climate change in the next 30 years: What can a convection-permitting model tell us that we did not already know? *Climate Dyn.*, **48**, 1987–2003, <https://doi.org/10.1007/s00382-016-3186-4>.
- Fyfe, J. C., and Coauthors, 2017: Large near-term projected snowpack loss over the western United States. *Nat. Commun.*, **8**, 14996, <https://doi.org/10.1038/ncomms14996>.
- Good, P., 2013: *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer Science & Business Media, 228 pp., <https://doi.org/10.1007/978-1-4757-3235-1>.
- Grimaldi, S., A. Petroselli, L. Baldini, and E. Gorgucci, 2015: Description and preliminary results of a 100 square meter rain gauge. *J. Hydrol.*, **565**, 827–834, <https://doi.org/10.1016/j.jhydrol.2015.09.076>.
- Haerter, J. O., B. Eggert, C. Moseley, C. Piani, and P. Berg, 2015: Statistical precipitation bias correction of gridded model data using point measurements. *Geophys. Res. Lett.*, **42**, 1919–1929, <https://doi.org/10.1002/2015GL063188>.
- Herold, N., A. Behrangi, and L. V. Alexander, 2017: Large uncertainties in observed daily precipitation extremes over land. *J. Geophys. Res. Atmos.*, **122**, 668–681, <https://doi.org/10.1002/2016JD025842>.
- Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei, 2008: Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.*, **113**, D21110, <https://doi.org/10.1029/2008JD010100>.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- Innocenti, S., A. Mailhot, and A. Frigon, 2017: Simple scaling of extreme precipitation in North America. *Hydrol. Earth Syst. Sci.*, **21**, 5823–5846, <https://doi.org/10.5194/hess-21-5823-2017>.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Katz, R. W., 2013: Statistical methods for nonstationary extremes. *Extremes in a Changing Climate*, A. AghaKouchak et al., Eds., Water Science and Technology Library, Vol. 65, Springer, 15–37.
- Kendon, E. J., and Coauthors, 2017: Do convection-permitting regional climate models improve projections of future precipitation change? *Bull. Amer. Meteor. Soc.*, **98**, 79–93, <https://doi.org/10.1175/BAMS-D-15-0004.1>.
- Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan Ser. II*, **93**, 5–48, <https://doi.org/10.2151/jmsj.2015-001>.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas, 1998: A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J. Hydrol.*, **206**, 118–135, [https://doi.org/10.1016/S0022-1694\(98\)00097-3](https://doi.org/10.1016/S0022-1694(98)00097-3).
- Kursinski, A. L., and X. Zeng, 2006: Areal estimation of intensity and frequency of summertime precipitation over a midlatitude region. *Geophys. Res. Lett.*, **33**, L22401, <https://doi.org/10.1029/2006GL027393>.
- Leduc, M., and R. Laprise, 2009: Regional climate model sensitivity to domain size. *Climate Dyn.*, **32**, 833–854, <https://doi.org/10.1007/s00382-008-0400-z>.
- , and Coauthors, 2019: The ClimEx Project: A 50-member ensemble of climate change projections at 12-km resolution over Europe and northeastern North America with the

- Canadian Regional Climate Model (CRCM5). *J. Appl. Meteor. Climatol.*, **58**, 663–693, <https://doi.org/10.1175/JAMC-D-18-0021.1>.
- Li, C., F. Zwiers, X. Zhang, and G. Li, 2019: How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, **7**, 11–24, <https://doi.org/10.1029/2018EF001001>.
- Liu, C., and Coauthors, 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dyn.*, **49**, 71–95, <https://doi.org/10.1007/s00382-016-3327-9>.
- Lucas-Picher, P., R. Laprise, and K. Winger, 2017: Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions. *Climate Dyn.*, **48**, 2611–2633, <https://doi.org/10.1007/s00382-016-3227-z>.
- Mantegna, G. A., C. J. White, T. A. Remenyi, S. P. Corney, and P. Fox-Hughes, 2017: Simulating sub-daily intensity-frequency-duration curves in Australia using a dynamical high-resolution regional climate model. *J. Hydrol.*, **554**, 277–291, <https://doi.org/10.1016/j.jhydrol.2017.09.025>.
- Martynov, A., R. Laprise, L. Sushama, K. Winger, L. Separovic, and B. Dugas, 2013: Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: Model performance evaluation. *Climate Dyn.*, **41**, 2973–3005, <https://doi.org/10.1007/s00382-013-1778-9>.
- Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213, <https://doi.org/10.1007/s10584-011-0156-z>.
- Menabde, M., A. Seed, and G. Pegram, 1999: A simple scaling model for extreme rainfall. *Water Resour. Res.*, **35**, 335–339, <https://doi.org/10.1029/1998WR900012>.
- Mineo, C., E. Ridolfi, F. Napolitano, and F. Russo, 2018: The areal reduction factor: A new analytical expression for the Lazio Region in central Italy. *J. Hydrol.*, **560**, 471–479, <https://doi.org/10.1016/j.jhydrol.2018.03.033>.
- NERC, 1975: Meteorological Studies. Flood Studies Report, Vol. 2, UK Natural Environment Research Council, 81 pp.
- NOAA, 2016: 15 min precipitation data (15PD) and hourly precipitation data (HPD), Climate Data Online; editing status 2016-06-17. Registry of Research Data Repositories, accessed 9 November 2016, <http://doi.org/10.17616/R32059>.
- Omolayo, A. S., 1993: On the transposition of areal reduction factors for rainfall frequency estimation. *J. Hydrol.*, **145**, 191–205, [https://doi.org/10.1016/0022-1694\(93\)90227-Z](https://doi.org/10.1016/0022-1694(93)90227-Z).
- Panthou, G., T. Vischel, T. Lebel, G. Quantin, and G. Molinié, 2014: Characterising the space–time structure of rainfall in the Sahel with a view to estimating IDAF curves. *Hydrol. Earth Syst. Sci.*, **18**, 5093–5107, <https://doi.org/10.5194/hess-18-5093-2014>.
- Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate*, **20**, 4356–4376, <https://doi.org/10.1175/JCLI4253.1>.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, <https://doi.org/10.1002/2014RG000475>.
- , and Coauthors, 2016: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: High resolution, high benefits? *Climate Dyn.*, **46**, 383–412, <https://doi.org/10.1007/s00382-015-2589-y>.
- , R. M. Rasmussen, K. Ikeda, C. Liu, M. P. Clark, and G. J. Holland, 2017: The future intensification of hourly precipitation extremes. *Nat. Climate Change*, **7**, 48–52, <https://doi.org/10.1038/nclimate3168>.
- , C. Liu, K. Ikeda, R. Bullock, R. M. Rasmussen, G. J. Holland, and M. Clark, 2019: Simulating North American meso-scale convective systems with a convection-permitting climate model. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-3993-2>, in press.
- Rasmussen, K. L., A. F. Prein, R. M. Rasmussen, K. Ikeda, and C. Liu, 2019: Changes in the convective population and thermodynamic environments in convection-permitting regional climate simulations over the United States. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-4000-7>, in press.
- Rummukainen, M., 2016: Added value in regional climate modeling. *Wiley Interdiscip. Rev.: Climate Change*, **7**, 145–159, <https://doi.org/10.1002/wcc.378>.
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.*, **115**, 15–40, <https://doi.org/10.1007/s00704-013-0860-x>.
- Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Stat. Assoc.*, **63**, 1379–1389, <https://doi.org/10.1080/01621459.1968.10480934>.
- Seneviratne, S. I., and Coauthors, 2012: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field et al., Eds., Cambridge University Press, 109–230.
- Separovic, L., A. Alexandru, R. Laprise, A. Martynov, L. Sushama, K. Winger, K. Tete, and M. Valin, 2013: Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. *Climate Dyn.*, **41**, 3167–3201, <https://doi.org/10.1007/s00382-013-1737-5>.
- Sigmond, M., and J. C. Fyfe, 2016: Tropical Pacific impacts on cooling North American winters. *Nat. Climate Change*, **6**, 970–974, <https://doi.org/10.1038/nclimate3069>.
- Sikorska, A. E., and J. Seibert, 2018: Value of different precipitation data for flood prediction in an alpine catchment: A Bayesian approach. *J. Hydrol.*, **556**, 961–971, <https://doi.org/10.1016/j.jhydrol.2016.06.031>.
- Sivapalan, M., and G. Blöschl, 1998: Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves. *J. Hydrol.*, **204**, 150–167, [https://doi.org/10.1016/S0022-1694\(97\)00117-0](https://doi.org/10.1016/S0022-1694(97)00117-0).
- Skamarock, C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.5065/D68S4MVH>.
- Svensson, C., and D. A. Jones, 2010: Review of methods for deriving areal reduction factors. *J. Flood Risk Manage.*, **3**, 232–245, <https://doi.org/10.1111/j.1753-318X.2010.01075.x>.
- Tapiador, F. J., and Coauthors, 2017: Global precipitation measurements for validating climate models. *Atmos. Res.*, **197**, 1–20, <https://doi.org/10.1016/j.atmosres.2017.06.021>.
- Thompson, V., N. J. Dunstone, A. A. Scaife, D. M. Smith, J. M. Slingo, S. Brown, and S. E. Belcher, 2017: High risk of unprecedented UK rainfall in the current climate. *Nat. Commun.*, **8**, 107, <https://doi.org/10.1038/s41467-017-00275-3>.
- Touma, D., A. M. Michalak, D. L. Swain, and N. S. Diffenbaugh, 2018: Characterizing the spatial scales of extreme daily precipitation

- in the United States. *J. Climate*, **31**, 8023–8037, <https://doi.org/10.1175/JCLI-D-18-0019.1>.
- Trenberth, K. E., Y. Zhang, and M. Gehne, 2017: Intermittency in precipitation: Duration, frequency, intensity, and amounts using hourly data. *J. Hydrometeor.*, **18**, 1393–1412, <https://doi.org/10.1175/JHM-D-16-0263.1>.
- Tustison, B., D. Harris, and E. Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *J. Geophys. Res.*, **106**, 11 775–11 784, <https://doi.org/10.1029/2001JD900066>.
- Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Japan Ser. II*, **87A**, 137–151, <https://doi.org/10.2151/jmsj.87A.137>.
- von Salzen, K., and Coauthors, 2013: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of physical processes. *Atmos.–Ocean*, **51**, 104–125, <https://doi.org/10.1080/07055900.2012.755610>.
- Westra, S., and Coauthors, 2014: Future changes to the intensity and frequency of short-duration extreme rainfall. *Rev. Geophys.*, **52**, 522–555, <https://doi.org/10.1002/2014RG000464>.
- Whan, K., and F. Zwiers, 2016: Evaluation of extreme rainfall and temperature over North America in CanRCM4 and CRCM5. *Climate Dyn.*, **46**, 3821–3843, <https://doi.org/10.1007/s00382-015-2807-7>.
- Xie, P., and A.-Y. Xiong, 2011: A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res.*, **116**, D21106, <https://doi.org/10.1029/2011JD016118>.
- , and R. J. Joyce, 2014: Integrating information from satellite observations and numerical models for improved global precipitation analyses. *Remote Sensing of the Terrestrial Water Cycle*, V. Lakshmi et al., Eds., John Wiley & Sons, 43–59, <https://doi.org/10.1002/9781118872086.ch3>.
- , M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, <https://doi.org/10.1175/JHM583.1>.
- , —, and W. Shi, 2010: CPC unified gauge-based analysis of global daily precipitation. *24th Conf. on Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 2.3A, [https://ams.confex.com/ams/90annual/techprogram/paper\\_163676.htm](https://ams.confex.com/ams/90annual/techprogram/paper_163676.htm).
- , R. Joyce, S. Wu, S.-H. Yoo, Y. Yarosh, F. Sun, and R. Lin, 2017: Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. *J. Hydrometeor.*, **18**, 1617–1641, <https://doi.org/10.1175/JHM-D-16-0168.1>.