

## Projected Changes in Precipitation Extremes for Western Canada based on High-Resolution Regional Climate Simulations

ANDRE R. ERLER AND W. RICHARD PELTIER

*Department of Physics, University of Toronto, Toronto, Ontario, Canada*

(Manuscript received 28 July 2015, in final form 4 July 2016)

### ABSTRACT

An analysis of changes in precipitation extremes in western Canada is presented, based upon an ensemble of high-resolution regional climate projections. The ensemble is composed of four independent, identically configured Community Earth System Model (CESM) integrations that were dynamically downscaled to 10-km resolution, using the WRF Model in two different configurations. Only the representative concentration pathway 8.5 (RCP8.5) scenario is considered. Changes in extremes are found to generally follow changes in the (seasonal) mean, but changes in mean and extreme precipitation differ strongly between seasons and regions (where extremes are defined as the seasonal maximum of daily precipitation). At the end of the twenty-first century, the highest projected increase in precipitation extremes is approximately 30% in winter away from the coast and in fall at the coast. Changes in winter are consistent between models; however, changes in summer are not: CESM is characterized by a decrease in summer precipitation (and extremes), while one WRF configuration shows a significant increase and another no statistically significant change. Nevertheless, the fraction of convective precipitation (extremes) in summer increases by 20%–30% in all models. There is also evidence that the climate change signal in summer is sensitive to the choice of the convection scheme. A comparison of CESM and WRF shows that higher resolution clearly improves the representation of winter precipitation (extremes), while summer precipitation does not appear to be sensitive to resolution (convection is parameterized in both models). To increase the statistical power of the extreme value analysis that has been performed, a novel method for combining data from climatologically similar stations was employed.

### 1. Introduction

There is considerable interest in the impact of anthropogenic climate change on extreme events, in particular hydroclimatic extremes such as heavy precipitation events (Herring et al. 2014; Stott 2015). Interest is especially intense in the climate change impact and adaptation community, for which changes in extreme events can be as important, or even more important, than changes in the mean climate (IPCC 2012).

In this study we focus on precipitation extremes in western Canada, which, for the purpose of this analysis, shall be taken to comprise the Canadian provinces of

British Columbia (BC) and Alberta south of 55°N. The region is characterized by extremely complex topography and a strong precipitation gradient across the Rocky Mountains. An analysis of various indices of precipitation extremes in station observations across Canada was initially conducted by Vincent and Mekis (2006). They found no evidence for an increase in extreme precipitation events in western Canada. However, several other authors have also considered projected changes in precipitation extremes based on global climate models (GCMs) and regional climate models (RCMs), and these studies typically have found increases in precipitation extremes over the area of interest.

Probably the most comprehensive such analysis to date is that of climatic extremes in the CMIP5 ensemble of GCMs discussed in Kharin et al. (2013). According to their analysis, the projected increase in the intensity of precipitation extremes in western Canada is expected to be approximately 20% by the end of the twenty-first century [under the representative concentration pathway 8.5 (RCP8.5) scenario], where precipitation extremes are defined as events with a return period of 20 yr. This

---

Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-15-0530.s1>.

---

Corresponding author address: Andre R. Erler, Department of Physics, University of Toronto, 60 St. George St., Toronto ON M5S 1A7, Canada.  
E-mail: aerler@atmos.physics.utoronto.ca

translates into an increase of roughly  $4\% \text{ }^{\circ}\text{C}^{-1}$ . The projected increase in mean precipitation under the same scenario is approximately  $10\%$  or  $2\% \text{ }^{\circ}\text{C}^{-1}$ . It is, however, important to recognize that these are changes in annual maxima and annual means, respectively. A recent assessment of changes in seasonally averaged precipitation over North America based on the CMIP5 ensemble has appeared in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Christensen et al. 2013). Therein it is reported that there is wide agreement between models that winter precipitation in western Canada will increase with global warming, although the signal in summer is much weaker and there is little agreement between models. The majority of the CMIP5 models display a tendency toward decreasing summer precipitation in the south and increasing precipitation in the north, with a transition region over western Canada (between  $50^{\circ}$  and  $55^{\circ}\text{N}$ ). It is, however, worth pointing out that the higher-resolution NARCCAP ensembles of RCMs and the high-resolution atmospheric GCM ensemble of Endo et al. (2012) both place the transition region farther south, so these models project an increase in precipitation over most of the area of interest. This may indicate that low-resolution global models are inadequate for projecting the effect of climate change in this region.

An analysis of changes in summer precipitation extremes (April–September) over Canada using a higher-resolution regional climate model [the Canadian RCM (CRCM), version 4] was presented by Mladjic et al. (2011). They find an increase in the intensity of various types of precipitation extremes of  $10\%$ – $15\%$  over western Canada by the middle of the twenty-first century. Furthermore, they report larger increases in high-latitude regions, where the projected increase in mean precipitation is also higher. Two other recent studies that have employed an ensemble of regional climate projections for the analysis of changes in precipitation extremes over North America are those by Dominguez et al. (2012) and d'Orgeville et al. (2014). However, Dominguez et al. (2012) only considered changes in winter precipitation in the western United States, while d'Orgeville et al. (2014) only considered changes in summer precipitation extremes in the Great Lakes region. Both find a substantial increase in precipitation extremes that is larger than the increase in mean precipitation [in fact, Dominguez et al. (2012) find no increase in mean precipitation]. Furthermore, analyzing simulations over the European Alps, Torma et al. (2015) have recently shown that high-resolution RCMs can add considerable value to GCM projections over complex topography and in particular to the representation of precipitation extremes. Western Canada has similarly complex topography and hence

RCMs would also be expected to add significant detail in this region.

It is worth noting that both Mladjic et al. (2011) and Dominguez et al. (2012) based their analysis on data from models with relatively low resolutions ( $30\text{--}50\text{ km}$ ; the NARCCAP ensemble and the CRCM version 4, respectively). Here we present an analysis of precipitation extremes based on an ensemble of high-resolution regional climate simulations at a resolution of  $10\text{ km}$ . The ensemble is driven by a small set of four identically configured global simulations with the Community Earth System Model, version 1 (CESM1), under the RCP8.5 forcing scenario, that differ only in their initial conditions. These simulations have been dynamically downscaled to  $10\text{-km}$  resolution using the Weather Research and Forecasting Model, version 3 (WRFV3), in two different configurations. Downscaling has been performed for a midcentury period and an end-century period over western Canada. The combination of models is the same as in d'Orgeville et al. (2014), but they employed a physics-based ensemble with only a single historical and midcentury realization for each ensemble member. To assess model uncertainty, we will also employ two different moist physics configurations in WRF. The mean climate of the CESM ensemble and one of the two WRF configurations extending to the middle of the twenty-first century was previously discussed by Erler et al. (2015) in the context of a discussion of hydrological changes in the Fraser and Athabasca River basins. For the purpose of validation of the representation of precipitation extremes, we will compare the simulations directly to daily observations from Environment Canada (EC; now Environment and Climate Change Canada) meteorological stations (Mekis and Vincent 2011). This is preferable to the use of reanalysis (as in Dominguez et al. 2012; Kharin et al. 2013), because precipitation in reanalysis is strongly affected by model physics and not well constraint by observations. Observational gridded products, on the other hand, do not have the temporal resolution necessary for our analysis (i.e., daily).

The detection of changes in precipitation extremes with decadal return periods is confounded by a high degree of natural variability and a scarcity of observational records of sufficient length. This also applies to climate model projections, albeit to a lesser extent when large ensembles are available (such as CMIP5 or NARCCAP). Simple numerical experiments show that a sample size of approximately 400 is necessary to reliably detect a  $15\%$  increase in the mean of the distribution of extremes.<sup>1</sup> To

---

<sup>1</sup> Using random samples from two generalized extreme value (GEV) distributions and the Kolmogorov–Smirnov test with  $\alpha = 0.05$ .

overcome this problem in our relatively small ensemble, we have developed a pooling technique that combines data from different stations (or grid points at station locations) based on the similarity of their climate. This is a major difference to many earlier studies, most of which estimate distribution parameters on a gridpoint-by-gridpoint basis. Other studies, such as Mladjic et al. (2011), also use regional frequency analysis (RFA; e.g., Hosking and Wallis 1997; Cooley 2009). Our approach is similar to RFA in that samples from different sites are combined (pooled). This, however, requires a degree of homogeneity between the samples, which is the main weakness of RFA: typically pooling regions have to be defined a priori, without any objective means of establishing statistical homogeneity. We solve this problem by applying a clustering algorithm based on average climatic characteristics of the stations (locations). We demonstrate that this technique yields good fits to the data and statistically meaningful results. In a similar effort, Jones et al. (2014) have recently used distribution parameters and other measures of precipitation extremes estimated at station locations to cluster station observations in the United Kingdom. However, we argue that a single station record does not provide sufficient data to adequately constrain the distribution (parameters) of precipitation extremes, and consequently the resulting clusters will be subject to higher levels of noise than clusters based on the mean climate. Furthermore, Bernard et al. (2013) argue that commonly used clustering algorithms are not suitable for data that follow a strongly nonnormal distribution (such as extremes).

For the purpose of our analysis, including the extreme value analysis, we use pooled data from station clusters for the validation period and the projection phase. d'Orgeville et al. (2014) also validated their simulations against station observations but then aggregated the data over the entire domain for the projection phase. This approach can be used for the relatively homogeneous Great Lakes region, but it is not feasible for the domain under consideration here, because of the heterogeneity of the climate. Furthermore, one can also argue that station locations are in some way representative of human habitation and activity, and thus provide a natural selection of locations with a sensible density weighting for impact assessment. We therefore choose to also employ the station clusters for the analysis of future changes. In addition and to further differentiate the present study from previous studies, we also pool data within each initial condition ensemble. This is possible because the configuration of the ensemble members is identical, so all the events can be considered to be drawn from the same parent distribution.

The same argument is usually made to pool data from long time series and can also be made for station locations that have a similar climatology.

The primary objective of this article is to document the possible impact of climate change on precipitation extremes in western Canada, based on high-resolution regional climate simulations. A secondary goal is to introduce a novel clustering and pooling technique as an alternative to regional frequency analysis. Details concerning the experimental setup and the model configurations used in the three ensembles are given in the next section (section 2). The clustering algorithm employed for pooling and the fundamentals of extreme value analysis are described in section 3. In section 4 changes in the seasonal mean and in extremes in different regions will be discussed, before results from the extreme value analysis are presented in section 5. Although the main focus in section 5 rests upon the WRF ensemble that was also employed by Erler et al. (2015), differences between WRF and CESM results and the “added value” of higher resolution will also be discussed. Section 6, on the other hand, deals primarily with differences between the two WRF ensembles and the role of convective precipitation in this context. The main results of the present study are summarized in section 7. Furthermore, it is useful to note at this point that most of the analysis in sections 4–6 will be based on three representative station clusters: one for the Pacific coast, one for the interior plateau, and one for the Prairies, which will be referred to as Pacific, plateau, and Prairies, respectively. The Prairies cluster is primarily located east of the Rocky Mountains and in their rain shadow, while the plateau cluster is located to the west of the Rocky Mountains and in the rain shadow of the Coast Mountains (see Fig. 1).

## 2. Experimental design and datasets

### a. Experimental design

The analysis presented in this article is primarily based on two small ensembles of dynamically downscaled climate projections. Both RCM ensembles downscale the same ensemble of GCM simulations, but they employ different moist physics during the dynamical downscaling phase. This makes three initial condition ensembles, each composed of four identically configured members.

The global climate projections were generated using the CESM, version 1.04, in the standard CMIP5 configuration (Gent et al. 2011); the CESM ensemble members differ only in their initial conditions (see Erler et al. 2015 for details). CESM is a fully coupled global climate model with interactive atmosphere, land, ocean, and sea ice components; the nominal resolution of

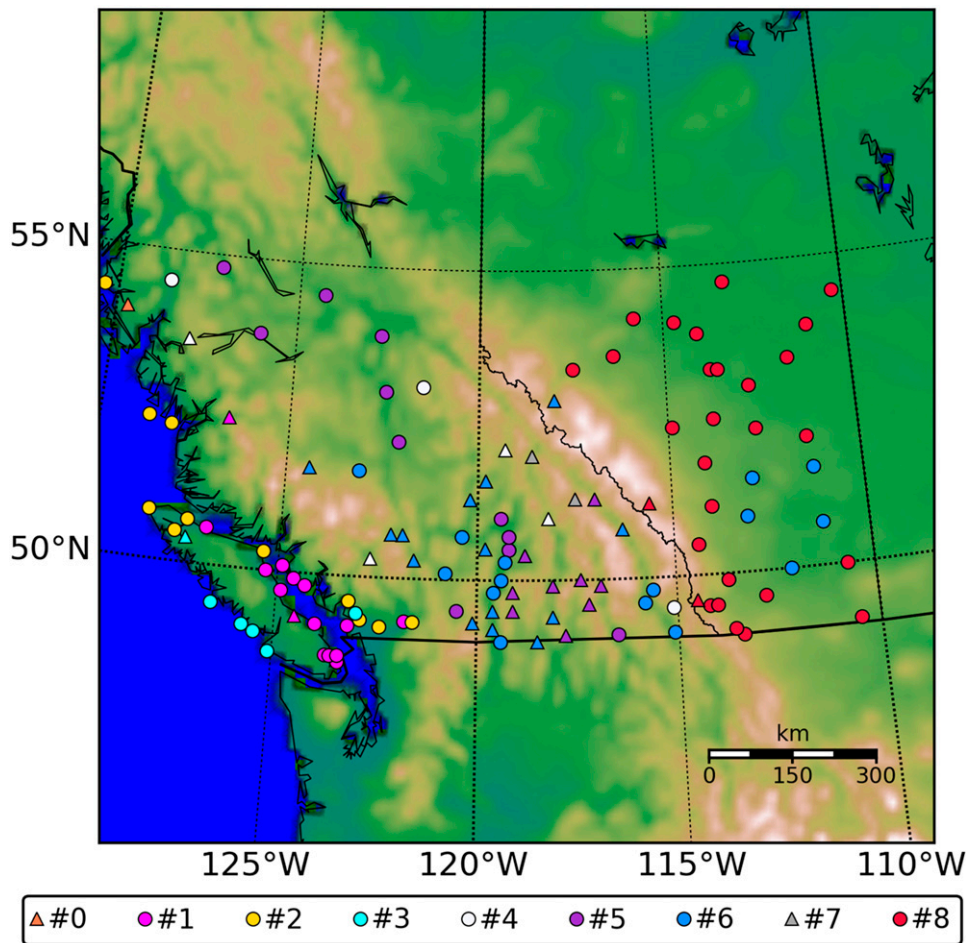


FIG. 1. Station locations and province outlines (BC and Alberta) with topography. The marker color indicates station cluster association; clusters 2, 6, and 8 were used in the analysis. Stations with more than 300-m elevation error in WRF have triangular markers and were excluded from the analysis. The topography of the inner WRF domain is shown in the background (10-km resolution).

the standard configuration is approximately  $1^\circ$  ( $\approx 100$  km at  $50^\circ\text{N}$ ).

The regional simulations differ in initial and boundary conditions (including sea surface temperatures and sea ice, which are prescribed from CESM). Each ensemble member was separately downscaled over western Canada, using a nested configuration consisting of an outer domain and an inner domain. Large-scale spectral nudging was applied to the outer domain to reduce phase drift between the global model and the regional model. The outer domain covers most of North America and the North Pacific at a resolution of 30 km, while the inner domain encompasses the Canadian provinces of British Columbia and Alberta, as well as parts of the surrounding territories, at a resolution of 10 km [see Fig. S1 in the supplementary material or Fig. 1 in [Erlor et al. \(2015\)](#)]. This resolution is much higher than most gridded datasets and all available reanalysis. The

regional simulations were performed for three 15-yr periods: a historical validation period from 1979 to 1994, a midcentury period from 2045 to 2060, and an end-century period from 2085 to 2100. Only one representative concentration pathway, the “business as usual” scenario (RCP8.5), was used. In this scenario the  $\text{CO}_2$  and  $\text{CH}_4$  concentrations are 541 ppmv and 2740 ppbv in 2050, and 936 ppmv and 3751 ppbv in 2100, respectively.

We have chosen a time-slice approach with a small initial condition ensemble in order to address the two-fold problem of high internal variability and non-stationarity. [Deser et al. \(2012\)](#) have demonstrated that natural variability can be very large at regional scales, in particular for precipitation, so many decades of data are necessary to obtain stable statistics. At the same time the process of anthropogenic global warming imposes a secular trend on the climate system with a time scale of several decades. This means that time series data of

sufficient length to obtain stable statistics (in a stationary climate) potentially stretch over several distinct climate states. By combining the time-slice approach with an initial condition (IC) ensemble, it is possible to overcome this problem, because the 15-yr time slices employed for each period are short enough that the climate statistics within them can be considered stationary and the sample size is increased through additional independent realizations from the IC ensemble. Whether 60 yr of data, with or without pooling, is sufficient to obtain a representative sampling of the distribution of any given hydroclimatic variable remains unclear at this point and in any event depends on the variable and the statistic in question as well as the required accuracy. Schindler et al. (2015) argue that only 30 years of data are sufficient to fully characterize the distribution of daily precipitation in a GCM, which would include the effect of large-scale quasi-periodic climate modes [such as ENSO and the Pacific decadal oscillation (PDO)]. However, other authors (e.g., Lucas-Picher et al. 2008) argue that a longer time series would be required to characterize mesoscale variability, and for the analysis of rare extreme events a much longer time series would certainly be desirable, which can only be effectively obtained by pooling.

The following definitions of seasons are used throughout this article; spring: March–May; summer: June–August; fall: September–November; and winter: December–February.

### *b. Model configuration*

One of the two WRF ensembles used in this study, as well as its mean climate up to midcentury, has been described previously in detail in Erler et al. (2015). In section 4 of this article, the analysis of the mean climate will be extended to the end of the twenty-first century, so as to provide context for the extreme value analysis presented later. The model release employed for downscaling is the Advanced Research version of WRF (ARW), version 3.4.1, to which a lake model (Gula and Peltier 2012; here employed in fully interactive mode rather than in offline mode) and code for time-dependent greenhouse gas (GHG) concentrations were added. The most important choices of physical parameterization schemes are as follows: the Noah land surface scheme (Tewari et al. 2004), the Morrison two-moment microphysics scheme (Morrison et al. 2009), the RRTM for GCMs (RRTMG) radiation scheme (Iacono et al. 2008), and the Grell-3 ensemble cumulus scheme (Grell and Dévényi 2002). For more information on model setup, we refer the reader to Erler et al. (2015).

To test the sensitivity to moist physics parameterizations, an alternate WRF ensemble was also integrated,

which will be discussed in section 6. The configuration is identical to the first WRF ensemble, except that the Kain–Fritsch (KF) cumulus scheme (Kain 2004) and the WRF single-moment 6-class microphysics scheme (WSM6; Hong and Lim 2006) were employed, as recommended in the WRF user guide (Wang et al. 2012) for regional climate simulations. This WRF ensemble will be referred to as the alternate ensemble (AE), while the first WRF ensemble will be referred to as the IC ensemble or simply as the first WRF ensemble.

Most of the analysis presented here will be for the IC ensemble, and differences to the CESM ensemble and the alternate ensemble will be noted where appropriate. The IC ensemble was chosen as the primary object of analysis, because it is characterized by an intermediate climate change response (between CESM and the alternate ensemble) and it achieves the best agreement with observations for precipitation extremes.

### *c. Station observations*

For the purpose of validation, we employ the EC station dataset for daily precipitation described by Mekis and Vincent (2011). The observations are quality controlled and corrected for undercatch, as well as site and instrumentation changes. In particular, records from stations separated by less than 10 km have been merged so correlations between stations should be small. For a detailed discussion of data preparation and trends in the mean climatologies, see Mekis and Vincent (2011).

The variables provided for each station are daily accumulations of total precipitation, solid precipitation, and liquid precipitation. All other variables used in this study were directly computed from these data. Because of the secular trend imposed by the global warming process, we use data only from 1950 to 2010; this choice is a compromise between data availability and stationarity of the climate. We also restrict the analysis to stations in the Canadian provinces of British Columbia and Alberta between 49° and 55°N. Furthermore, to mitigate disagreement between model simulations and station observations due to inconsistent topography, we omit all stations where the elevation error in the inner WRF domain is larger than 300 m; this excludes most alpine stations.

## **3. Analytical methods**

This section introduces the analytical framework employed in the analysis of precipitation extremes presented in this article. The same framework has also been employed by Erler and Peltier (2015) to analyze precipitation extremes in historical observations in

western Canada. Additional information is available in section S2 of the supplementary material.

### a. Station clustering

As has been outlined in section 1, it is desirable to pool station data in order to increase the amount of data from which the parameters of the probability distribution are estimated. However, it is not feasible to combine all stations, because in particular British Columbia is far too heterogeneous, so the distribution parameters vary and a reasonable fit is impossible to achieve. It should be noted, however, that the climate of Alberta is sufficiently homogeneous that pooling of all nonalpine stations within the province would be feasible (see section S2.3 in the supplementary material for further discussion).

To group stations with a similar climate and seasonal cycle, we have chosen the well-known  $k$ -means algorithm (Hartigan and Wong 1979), for which we employ the implementation from the Scikit-learn machine learning library for Python (Pedregosa et al. 2011). The  $k$ -means algorithm groups  $n$ -dimensional vectors based on their Euclidian distance into an a priori determined number of clusters. It works by iteratively adjusting the position of so-called centroid vectors so that the sum of the distances of each vector to the closest centroid is minimized. Each station was represented by a feature vector composed of the climatological seasonal cycle of monthly mean liquid and solid precipitation measured at the station ( $2 \times 12$  values–dimensions). To reduce the dimensionality of the station vectors, principal component analysis was applied prior to clustering, and only the three leading empirical orthogonal functions (EOFs) were retained, which combined explain 97% of the total variance. This reduces the dimensionality of the station vectors to three. The result of the clustering algorithm is not sensitive to the number of EOFs used, as long as the first three are included. The locations of stations and their cluster associations are shown in Fig. 1. A total of nine clusters were used in the clustering algorithm. This number was chosen so that at least one cluster with more than 500 combined years is available for each of the three major regions (Pacific coast, interior plateau, and Prairies). The number of clusters is somewhat arbitrary, but certain types of clusters occur robustly, regardless of the total number. The first division occurs between stations at the coast and stations farther land inward (which differ significantly in the amount of total winter precipitation). The next major division is between the plateau and the Prairies (which differ in the amount of solid precipitation and temperature). However, prior to the emergence of the Prairies as a separate cluster, several smaller clusters form close to the coast, where the magnitude of winter precipitation varies significantly (Vancouver Island in particular has

very high precipitation). The three main regions emerge with more than six clusters, but with a larger number, the major groups become more homogeneous, because “outliers” are moved into separate clusters.

In the design of the clustering algorithm, a conscious effort was made to identify regions that correspond to more general hydroclimatic categories, in the sense of, for example, the Köppen–Geiger climate classification system (Köppen 1936), rather than to simply cluster stations with similar precipitation extremes. There are two reasons for this choice. The first is to avoid overfitting of the observations, in order to facilitate a fair comparison between the models and observations. The second is that in this way the likelihood is higher that general precipitation characteristics, such as distribution, frequency, intermittency, and the type and process of precipitation, are similar, and not only the magnitudes; this will prove important for the analysis of cumulus precipitation in section 6. Sensitivity experiments show that the clustering results are not very sensitive to the precise choice of variables, as long as a measure of total precipitation and temperature is included. Solid precipitation was included as a measure of (winter) temperature and continentality, because temperature data were available for only about half of the stations, but using interpolated temperature data from a gridded dataset leads to very similar results. Furthermore, including measures of precipitation extremes (or variability) in the clustering algorithm does not change the results either. The reason is that the Pearson correlation coefficient of the climatological means of monthly maximum precipitation and monthly mean precipitation, computed over the seasonal cycle (monthly) and all station in BC and Alberta, is 0.97. Hence, the magnitude of extreme precipitation follows that of mean precipitation. On the other hand, using only measures of total precipitation (mean, variance, and/or maxima) leads to very small clusters at the coast (due to the strong precipitation gradient) and late separation between the plateau cluster and the Prairies cluster. Inclusion of the number of wet days also facilitates the separation of the Prairies cluster but has little effect otherwise. Also note that clustering was performed only on the observed station data; model data did not enter the algorithm. This facilitates the comparison between the three ensembles and penalizes models with a climatology that strongly deviates from the observed climatology. For the same reason it was also not possible to consider the climate change signal in the clustering algorithm, as it would make the clusters model dependent. For additional discussion of the clustering algorithm, we direct the reader to section S2.1 of the supplementary material.

Three clusters comprising 53 stations have been chosen for the purpose of our analysis: 2, 6, and 8. This choice is based on cluster size, agreement between observations and simulations, and representativeness (see below). Cluster 2 (11 stations) will be referred to as Pacific, cluster 6 (15 stations) as plateau, and cluster 8 (27 stations) as Prairies, in accord with the three regions described in [section 1](#) (note, however, that some of the stations in the plateau cluster are actually in Alberta). The Pacific cluster is sometimes loosely referred to as “the coast”; the plateau and Prairies clusters together are often referred to as “land inward.” Results will be presented separately but aggregated over each regional cluster.

### *b. Extreme value analysis*

Extreme value analysis is the study of the highest or lowest quantiles of a randomly distributed variable; these quantiles are often referred to as the tails of the distribution. There are two fundamental approaches to this analysis ([Coles 2001](#); [Katz et al. 2002](#)). One is often called the peak-over-threshold approach, which is concerned with the distribution of all samples that exceed a fixed threshold. The resulting distribution of extremes is of the generalized Pareto distribution (GPD) type. Its use makes immediate sense when a threshold can be selected a priori, but otherwise it is problematic ([Katz 2010](#)). The approach that is employed here is based on the generalized extreme value (GEV) distribution. The GEV distribution is based on a block maximum approach, where the maximum is taken over a fixed number of independent and identically distributed samples. According to the extreme value theorem, for any parent distribution of the samples, the GEV distribution is the limit distribution of such a process, when the block size tends toward infinity.

To gauge the quality of the resulting fit, we make use of the Kolmogorov–Smirnov (K–S) test ([Smirnov 1939](#); [Massey 1951](#)), which measures the maximum difference between the cumulative density function (CDF) of two distributions (discrete or analytical). We also use the K–S test to measure the agreement between the observed distribution and the (bias corrected) model distributions from the simulations, as well as to determine the statistical significance of future changes. For either comparison (samples to fit or between two samples), we report the  $p$  value associated with the K–S statistic ([Smirnov 1948](#)). The  $p$  value can be thought of as the significance level at which the null hypothesis that the two samples are drawn from the same parent distribution can be rejected ([von Storch and Zwiers 2002](#)). The  $p$  values that are reported for the quality of fit are biased upward, because the same data were used for the goodness-of-fit test as for the estimation of the fit parameters (i.e., there

is a risk of overfitting). This problem can be overcome by cross validation and is further discussed in [section S2.2](#) of the supplementary material. However, in order to make full use of the available data, we will use the entire dataset for fitting and testing, and report  $p$  values for the goodness of fit with the understanding that they are not proper hypothesis tests, but merely a measure of the quality of the fit. Further note that  $p$  values associated with the comparison of two independent samples (such as models and observations) are not biased and represent proper hypotheses tests.

In the study of climatic extremes, the canonical block size is one year ([Cooley 2009](#); [Katz 2010](#)), but we will deviate from this convention here and consider seasonal maxima instead. We limit the block size to one season, because climatic variables are typically subject to strong seasonality, so the assumption of identical distribution is questionable. Furthermore, the annual maximum is often dominated by events from only one season (e.g., the rainy season), so not much data are lost by separating seasons.

To increase the number of data points that constrain the distribution fits, we pool data from different stations within a cluster and fit a single GEV distribution to the combined data. This approach is similar to RFA ([Hosking and Wallis 1997](#); [Cooley 2009](#)), where samples from different sites across an a priori defined region are pooled. To allow a reasonable distribution fit to the combined samples, the pooling region has to be sufficiently homogeneous. In [section S2.3](#) of the supplementary material we show that a GEV fit to data pooled over the entire province of British Columbia can be clearly rejected based on the goodness-of-fit test (the K–S test). In RFA this problem is typically overcome by normalizing each sample by its site-specific sample mean, before pooling the data and fitting a distribution. However, this still requires a high degree of homogeneity in the shape of the distributions. This assumption of homogeneity is not supported by our analysis, and [Mladjic et al. \(2011\)](#) also report a failure to fit a parametric distribution to precipitation extremes pooled over western Canada. The advantage of using the clustering approach over regional frequency analysis is that the homogeneity assumption is required only within a cluster, and the station clusters have been constructed based on climatic similarity. We also find that, based on the goodness-of-fit test, the clusters are sufficiently homogeneous, so prior normalization is not necessary in our case. In principle it is possible that the pooled samples are actually not homogeneous but still allow a good distribution fit; in this case pooling can introduce biases. However, the same argument can also be made with respect to different time periods, a case where

pooling is rarely questioned. For example, in many regions the climate exhibits long-term variability on a decadal time scale, which is often associated with slowly varying climate modes, such as ENSO or the PDO in western Canada. It would thus be reasonable to expect the parent distribution of extreme events to slowly vary over time [see [Erler and Peltier \(2015\)](#) for an example]. Hence we argue that some degree of spatial variation is also acceptable. Many previous studies have avoided pooling of data altogether and instead chose to fit a distribution to every site individually. However, in section S2.4 of the supplementary material, we demonstrate that without pooling of data it is not possible to constrain the shape parameter adequately and even unphysical values cannot be rejected based on the goodness-of-fit test. Furthermore, the differences between models and the differences between validation and projection periods are often not statistically significant, if tested on a station-by-station basis (i.e., the sampling error is larger than the differences). It is of course possible to aggregate individual test results using batch-testing techniques (e.g., [Livezey and Chen 1983](#)); however, it would still be necessary to aggregate the samples into different (more homogeneous) groups for the purpose of analysis. We further argue that due to the nonlinearity of the fitting process, averaging of distribution parameters or return periods from different sites will introduce similar or even larger errors than pooling based on relatively homogeneous clusters. The key advantages of pooling are that the shape parameter can be constrained more accurately and that statistical test can be applied in a meaningful way.

For the comparison to observations and for the estimation of extremes, we employ bias correction of the distribution mean (but not of other moments). Bias correction was performed by multiplying the location and scale parameters of the modeled distribution with the ratio of the means of the observed and the modeled distribution so that the observed and the bias-corrected sample means are the same. This is equivalent to rescaling the samples, but it does not require a new fit. All computations involving the GEV distribution and related statistical tests were performed using the SciPy statistics library ([Jones et al. 2001](#)). The GEV parameters were determined using a maximum likelihood estimator.

#### 4. Climatologies of station clusters

In this section the climatological seasonal cycle and seasonal averages will be discussed; this directly extends the discussion in section 3 of [Erler et al. \(2015\)](#). We discuss changes in total and net precipitation, convective

precipitation and snowfall, and extremes of precipitation. We also discuss the spatial structure of the change in precipitation extremes. For information on temperature, snowmelt, and runoff changes, see section S1.1 of the supplementary material.

##### a. The average seasonal cycle

[Figure 2](#) shows the climatological seasonal cycle of precipitation-related variables for station clusters 2, 6, and 8 (Pacific, plateau, and Prairies). Observational data are shown as filled circles with error bars, and down-scaled IC ensemble data are shown as lines: solid lines for the historical period, dashed-dotted lines for the midcentury period, and dashed lines for the end-century period (for clarity, error bands are shown only for the historical period). The error bars and bands are based on the standard error of the mean (SEM), which is related to the sample standard deviation  $s$  by  $SEM = (1/n)^{1/2}s$ , where  $n$  is the number of samples (the number of years in this case). The error bars and bands show the 95% confidence region, corresponding to approximately  $\pm 2$  SEM. The standard deviation itself would be a measure of the interannual variability, which is to be distinguished from the uncertainty in the mean. The standard error of the mean has been estimated for each station separately and averaged subsequently. This procedure likely overestimates the true uncertainty in the aggregated station average. However, since monthly means vary on larger spatial scales, correlations between stations within the same cluster cannot be neglected, and treating all station values as independent samples would underestimate the uncertainty in the mean. The scales to the right of each panel show the annual means. The seasonal cycle of total and solid precipitation, as well as wet days of all station clusters, is shown in [Fig. S3](#) in the supplementary material. Here we note only that the number of wet days is significantly overestimated in WRF in all clusters that are not directly at the coast. This is the case for both configurations, but it is slightly more pronounced for the first (IC) ensemble. The overestimation of wet days is more pronounced for lower thresholds, suggesting that a large amount of “drizzle” is present in the simulations; this is a common problem in climate models (cf. [Stephens et al. 2010](#)).

The increase in temperature due to global warming in our simulations is approximately  $2^{\circ}\text{C}$  at the coast and  $3^{\circ}\text{C}$  land inward by midcentury; by the end of the century the increase is just over  $4^{\circ}\text{C}$  at the coast and almost  $5^{\circ}\text{C}$  land inward. The increase is generally larger in the diurnal minimum temperature and smaller in the diurnal maximum (see [Fig. S2](#) in the supplementary material). As discussed in [Erler et al. \(2015\)](#), the increase in average winter temperature is relatively



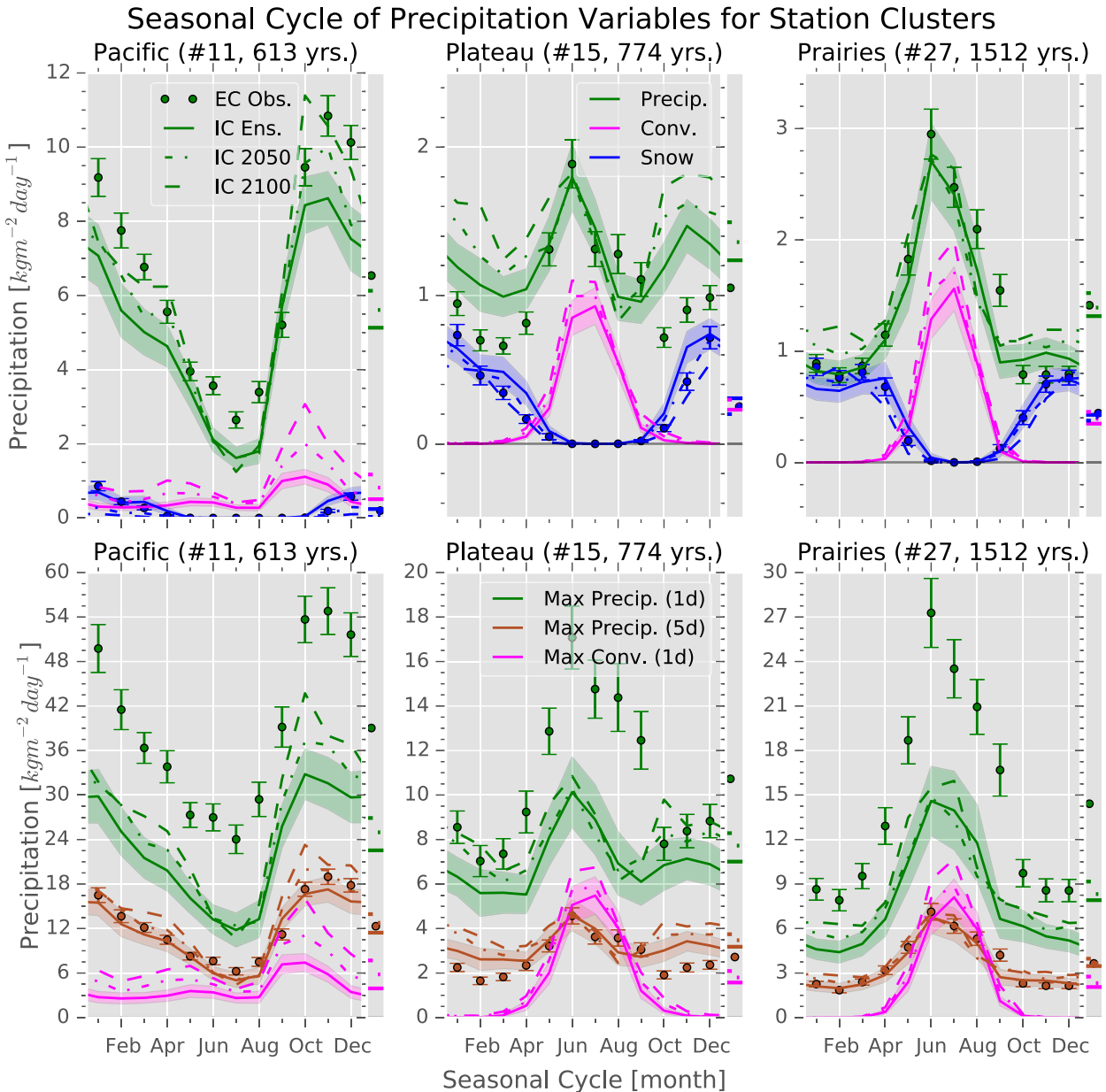


FIG. 2. Seasonal cycle of (top) precipitation types and (bottom) precipitation extremes for selected station clusters—#2, #6, and #8—for the IC ensemble. The error bars—colored bands show the standard error of the climatological mean [ $\sigma/(n)^{1/2}$ ]; the scales to the right of each panel show the value of the annual mean.

consistent between the models, while the increase in summer temperatures is larger in CESM.

Total, convective, and solid precipitation are shown in Fig. 2 (top row; monthly means). Compared to station observations, winter precipitation at the coast is significantly underestimated, while it is overestimated for the plateau and Prairies (quite substantially for the plateau). This problem was previously identified in Erler et al. (2015), who suggested that this is at least partly caused by an underestimation of the rain shadow effect of the

Coast Mountains. It is interesting to point out that at the coast, the total precipitation amount and the relative change are larger in fall than in winter, whereas both are similar in fall and winter in the Prairies; the plateau cluster has a slight asymmetry in total precipitation, but the relative change is the same. Furthermore, the peak in convective activity at the coast, as well as the largest increase therein, also occurs in fall, likely due to high frontal activity in fall combined with larger moisture availability and higher temperatures. It is not surprising

that the amount of snowfall (solid precipitation) decreases. The Pacific and plateau clusters see a decrease throughout the winter, while the Prairies, which experience colder winters due to the continental climate, see a decrease only at the beginning and the end of the snow season. The total amount of snow is somewhat overestimated toward the coast. At the end of the century, snow is projected to almost disappear completely at the coast, but it is important to note that this is the case for station locations only at lower elevations. Summer precipitation is generally represented more accurately in WRF; only the stations at the Pacific coast have a significant (low) bias. In the IC ensemble, all clusters are characterized by only a small (hardly significant) change in total precipitation, but convective precipitation increases significantly in the plateau and Prairies clusters.

Analogous figures for the CESM ensemble and the alternate WRF ensemble are shown in the supplementary material. The alternate WRF ensemble shows a very similar performance to the first WRF ensemble at the Pacific coast; in the Prairies its performance is somewhat better, with only a small low bias in fall precipitation but essentially no bias in summer and fall. In the interior plateau the seasonal cycle is much better represented in the alternate WRF ensemble, despite a high bias in summer. The climate change response is also generally similar in the alternate ensemble, except for a significant precipitation increase of approximately 10% in summer in the Prairies, which is not seen in the first WRF ensemble. In CESM, on the other hand, summer precipitation in western Canada generally decreases, which is consistent with most CMIP5 models (Christensen et al. 2013) but is at odds with the two WRF ensembles and other high-resolution climate projections (e.g., Endo et al. 2012). In fall and winter, CESM shows smaller changes than the two WRF ensembles at the coast and a significant reduction in snow during winter land inward, where WRF shows only a decrease at the beginning and the end of the snow season. The seasonal cycle is well represented at the coast, although the magnitude of precipitation is underestimated. In the interior plateau, CESM essentially simulates the same seasonal cycle and precipitation amounts as at the coast (i.e., there is no rain shadow). The seasonal cycle in the Prairies is considerably underestimated in CESM: fall, winter, and spring precipitation is overestimated by a factor of 2, while summer precipitation is simulated very accurately. All models also show a high fraction of convective precipitation in summer land inward; interestingly the fraction is the lowest in the first WRF ensemble (~55%) and the highest in the alternate ensemble (~70%). Both

WRF ensembles project a considerable increase in convective precipitation in the future, while in CESM convective precipitation decreases (see section 6b for more information). It is notable, however, that the decrease in convective precipitation is considerably less than the decrease in total precipitation.

The climate change response in winter is very robust and consistent between models, while the response in summer is inconsistent and possibly weak. The increase in winter can be understood under the assumption that the main source of moisture is the Pacific Ocean, and that winter precipitation is essentially transport limited. The ocean has a weaker seasonal cycle and is thus warmer than the air and the land in winter, so evaporation is not the limiting factor. Consequently, the air masses that bring moisture across the Rocky Mountains are close to saturation, and transport is limited only by the saturation vapor pressure, which increases by  $7\% \text{ }^{\circ}\text{C}^{-1}$  (Trenberth et al. 2003). In summer, on the other hand, precipitation is more dependent on local processes such as convection, and evapotranspiration over the continent constitutes a significant moisture source (van der Ent et al. 2010); currently both of these processes are parameterized and thus are highly model dependent.

#### *b. Changes in the mean of extreme precipitation*

Monthly maxima of daily and 5-day (pentad<sup>2</sup>) precipitation and daily convective precipitation are shown in Fig. 2 (bottom row; corresponding roughly to the 96.7th percentile). It is evident that precipitation extremes closely follow the seasonal cycle of the monthly averages. The changes in precipitation extremes in fall and winter by midcentury and the end of the century are very similar to the changes in monthly means, but in summer there is a small increase in extremes, while there is no increase in the monthly means. Note that convective precipitation, even for daily extremes, is considerably smaller than total precipitation, and the fraction is approximately the same for extremes and for monthly means. The increase in convective precipitation extremes follows the same pattern as the increase in monthly mean convective precipitation. Again, the fraction of convective precipitation extremes to total precipitation extremes is about 15% points larger in the alternate ensemble compared to the first WRF ensemble. The fraction of convective precipitation in CESM is close to that of the alternate WRF ensemble.

---

<sup>2</sup> Here pentad precipitation extremes are defined as the maxima of nonoverlapping 5-day precipitation rate averages; some studies use the 5-day accumulation instead.

The finding that across the seasonal cycle changes in extremes approximately follow changes in the mean appears to be at odds with some previous studies (Dominguez et al. 2012; Kharin et al. 2013). This may be a coincidence and only true for this region, but the effect of seasonal versus annual aggregation (or GEV block size) should also be considered: if mean and extreme precipitation increase equally during the wettest season and decrease during drier seasons, then the change in annual extremes will be close to the change during the peak season, while the increase in the annual mean will be less pronounced.

It is further evident that in most regions and seasons, the magnitude of daily precipitation extremes is approximately 40% smaller in the WRF ensemble than in observations. This is not surprising, because model values represent gridpoint averages, while station observations represent point values. The magnitude of pentad extremes, on the other hand, is actually simulated very accurately, especially in summer. This is likely due to the high degree of temporal aggregation, which compensates for the spatial aggregation effect seen in the model data (cf. Eggert et al. 2015).

The spatial distribution of the change in maxima of daily precipitation by the end of the twenty-first century is shown in Fig. 3 (corresponding roughly to the 98.9th percentile). In Fig. 3, the CESM ensemble is shown on the left, the WRF IC ensemble in the center, and the alternate WRF ensemble on the right; the top row shows the maxima for summer and the bottom row for winter. Consistent with Fig. 2, there is a wide spread increase in winter precipitation that slightly increases land inward, in particular in the lee of the Rocky Mountains (the Prairies). This is the case in all three ensembles. The picture in summer, however, is less clear: there is a suggestion of a north–south gradient in the ensemble means, but it is not consistent between individual members of any of the IC ensembles (not shown). If seasonal averages of monthly maxima of precipitation are plotted, then the north–south gradient is more easily discernible and in monthly means it is already clearly visible at midcentury (e.g., see Fig. 7 in Erler et al. 2015). It is thus likely that the same north–south gradient would emerge if more data were available to improve the signal-to-noise ratio. For comparison the changes in seasonally averaged precipitation at the end of the twenty-first century are shown in Fig. S6 in the supplementary material.

The existence of a gradient has implications for the interpretation of changes within station clusters: because station clusters are mainly oriented along the meridional direction and the southern stations are typically within the transition region, stations with different trends are being mixed; however, given the uncertainty associated with the location of the transition and the amount of

natural variability, we do not believe that a further division based on the climate change signal is warranted at this point.

## 5. Analysis of precipitation extremes

In this section, changes in the distribution of seasonal maxima and rare extreme events with decadal return periods will be discussed. The seasonal maxima roughly correspond to the 98.9th percentile of the season; the maximum of the season with the highest (extreme) precipitation also corresponds to the 99.7th annual percentile. All histograms and distribution functions shown in this section are based on seasonal maxima (i.e., one data point per season and year).

This section is concerned only with extremes of daily precipitation totals; we first show distributions for the historical and projection periods, before we discuss the behavior of the extreme tail of the distributions. The role of cumulus precipitation and important differences between the two WRF ensembles are discussed in the next section.

The probability density functions (PDFs) of fitted GEV distributions, along with histograms of the underlying data, are shown in Figs. 4, 5, 7, and 8 (corresponding PDFs and histograms are identified by color). Every panel has a reference sample (usually the observations or the historical IC ensemble), which will be indicated in the figure caption. The  $p$  values for the quality of fit and the ratios of the sample means to the reference sample are given in the upper-right corner of each panel (first and second columns, respectively). Also printed on each panel are the K–S test results from the comparison of each sample to the reference sample, as well as the ratio of the distribution means (after bias correction, if applicable). The  $p$  values associated with the quality of fit tend to be exaggerated, for the simple reason that the maximum likelihood estimator optimizes a metric that is similar to the  $p$  value; the values should simply be regarded as a guide to the quality of fit. Confidence intervals at the 95% level are shown as semitransparent bands around the PDF curves; the confidence intervals were derived from bootstrapping (random resampling with replacement; Efron and Tibshirani 1994), using 100 realizations; they indicate sampling uncertainty, that is, how sensitive the fit is to the selection of data points. Note that bootstrapping does not provide any estimate of uncertainty that has not been sampled (e.g., model uncertainty).

In Figs. 4–7 station clusters are shown in columns (from left to right: Pacific, plateau, and Prairies), as before, and seasons are organized by row, with summer in the top row and winter in the bottom row. The exception is the Pacific cluster, for which fall is shown in the upper-left panel instead of summer. The reason for

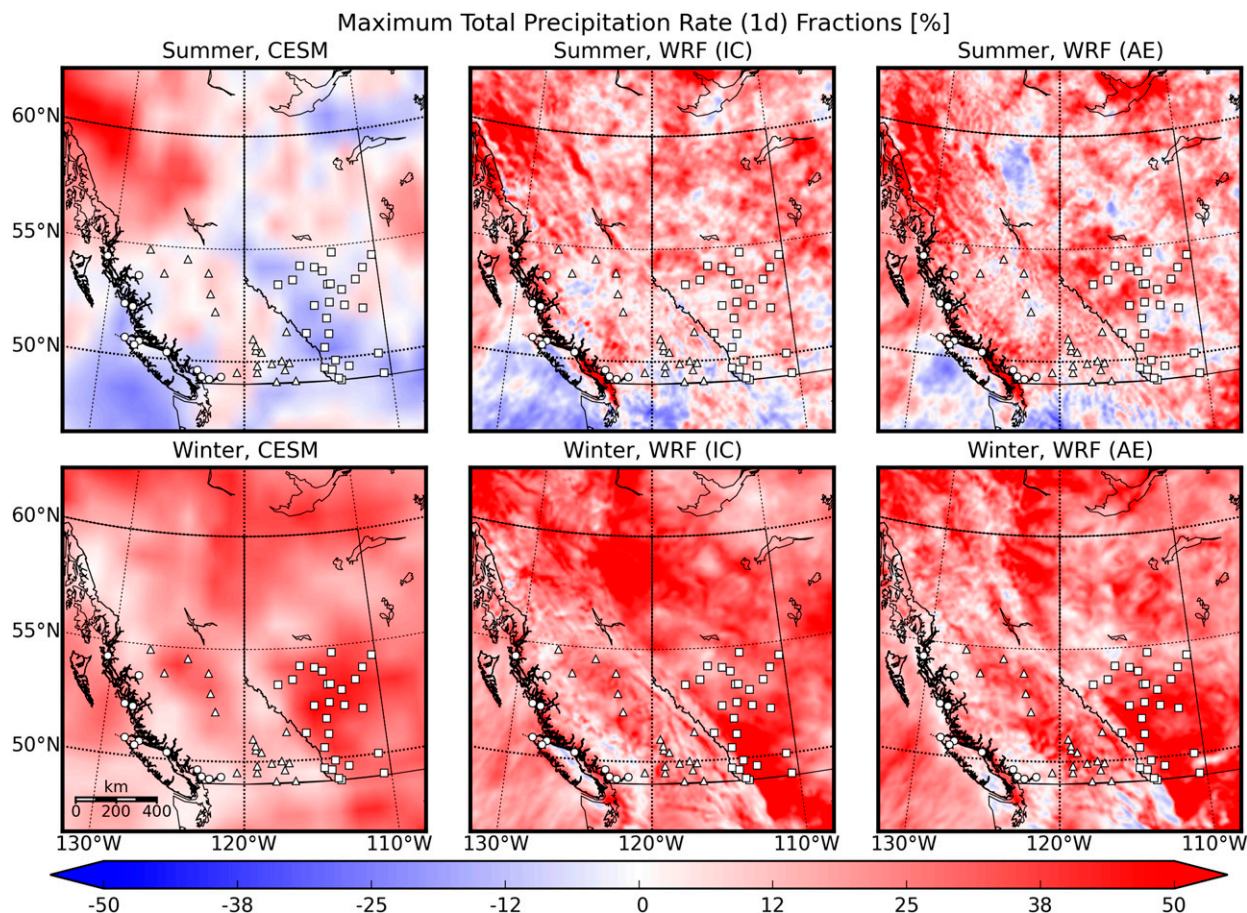


FIG. 3. Geographic distribution of changes in maximum daily precipitation for the (left) CESM ensemble, (center) the WRF IC ensemble, and (right) the WRF alternate ensemble in (top) summer and (bottom) winter. Changes are shown for the end of the twenty-first century relative to the historical period. Markers indicate station locations for the Pacific (circles), plateau (triangles), and Prairies (squares) clusters. The pattern of change is similar to but noisier than the pattern of change in the monthly means (see Fig. S6 in the supplementary material).

this choice is twofold: first, both the largest precipitation amount and the largest change in the Pacific cluster occur in fall, and summer is the time of the precipitation minimum, which does not change appreciably. Second, our confidence in the summer results for the Pacific cluster is lower, because the region could still be affected by the SST and sea ice interpolation error (which is the case for cluster 1).

The number of stations and observational data points in each cluster (for the season) are shown in the panel headers; the number of data points for the simulations is always the length of the simulation (in years, i.e., 60) multiplied by the number of stations.

#### a. The distribution of precipitation extremes

##### 1) VALIDATION

Figure 4 shows the PDFs of seasonal maxima of daily precipitation intensities for the validation period

(1979–94). PDFs are shown for the outer (green; 30-km resolution) and inner (blue; 10-km resolution) domains of the IC ensemble, as well as the driving CESM ensemble (red; ~100-km resolution). Also shown are the station observations (black), which are the reference samples in each panel. The quality of the GEV fits is generally quite good for WRF but not acceptable for CESM. The reason is that the climate of CESM significantly differs from the observed climate, and the station clusters violate the assumption of identical distribution for CESM data. For WRF, on the other hand, the station clusters appear to work well. Again, it is evident that the mean of the observed distributions is significantly larger than the mean of the model distributions; in almost all cases the magnitude of model extremes is approximately 60% of the observed extremes (except in winter in the interior plateau, where precipitation is generally overestimated). The reduction of the intensity of precipitation extremes due to temporal and spatial aggregation was

### Distribution of Maximum Total Precipitation Rate (1d)

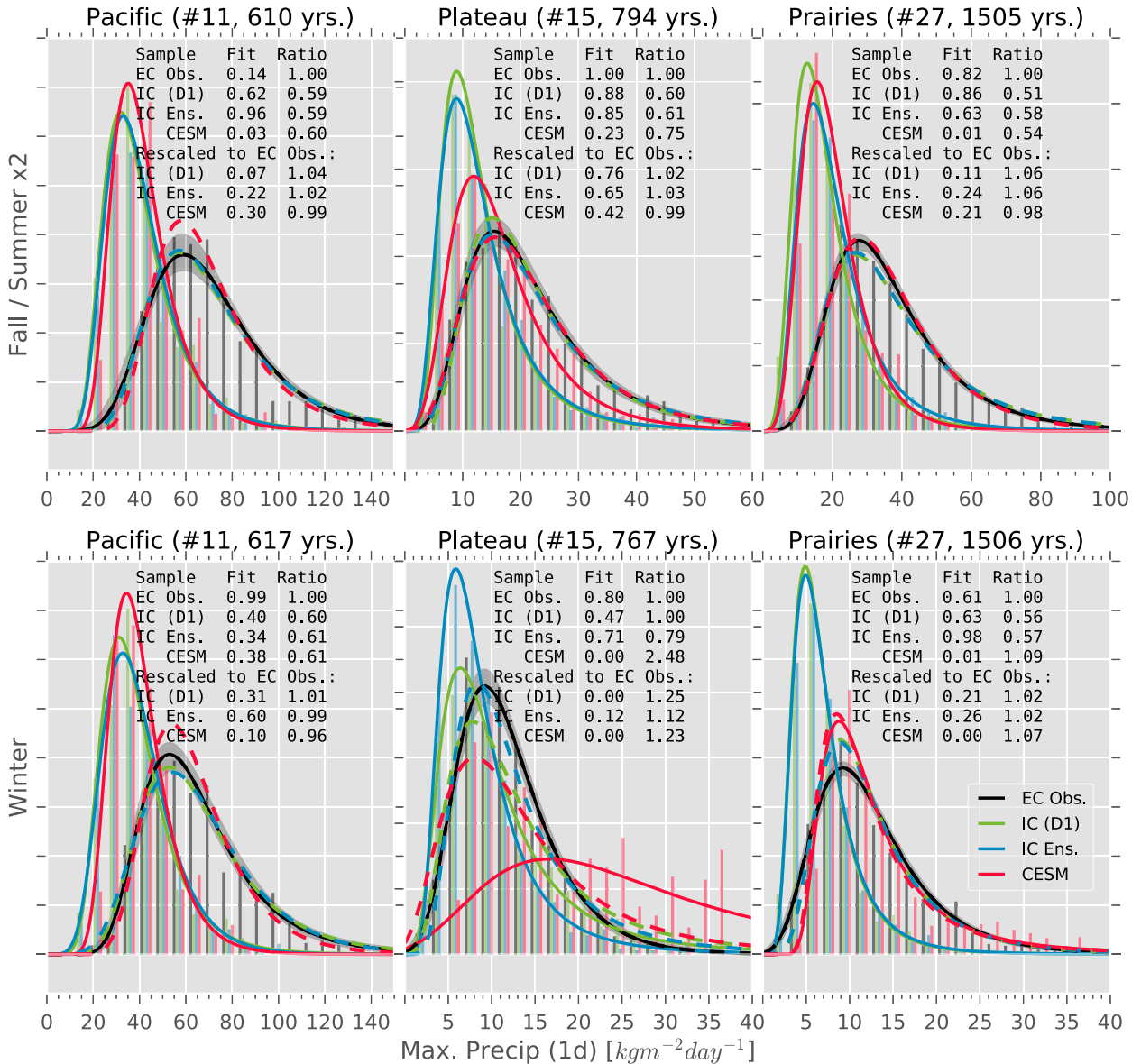


FIG. 4. Distribution of (top) fall (for only Pacific) and summer (plateau and Prairies) and (bottom) winter maxima of daily average precipitation with validation against observations in the (left) Pacific, (center) plateau, and (right) Prairies regions: observations (black), IC outer (D1; green) and inner (Ens.; blue) ensemble, and CESM (red). The dashed lines are fits that were rescaled so that the mean of the model distribution matches the distribution mean of the observed distribution. The colored bands are 95% confidence intervals, derived from bootstrapping. The station observations serve as the reference sample. The number of stations (#) and the number of samples (yrs) for each cluster and season are indicated in the panel titles.

investigated by Eggert et al. (2015), using high-resolution radar data over Germany. The area reduction factors derived by Eggert et al. (2015) can potentially explain the reduction seen in summer, when convective precipitation is dominant, but they cannot explain that the reduction factor is the same in winter, when stratiform precipitation is dominant, and also in the outer WRF domain, which has a different resolution. In this context it should also be

noted that the effective resolution of our simulations is likely a multiple of the nominal resolution (i.e., coarser) due to numerical diffusion (cf. Skamarock 2004).

Bias-corrected (rescaled) model distributions are shown as dashed lines. The *p* values in the “Rescaled Experiments” section refer to the null hypothesis that the bias-corrected model distribution is the same as the observed distribution. The apparent quality of fit is the

### Distribution of Maximum Total Precipitation Rate (1d)

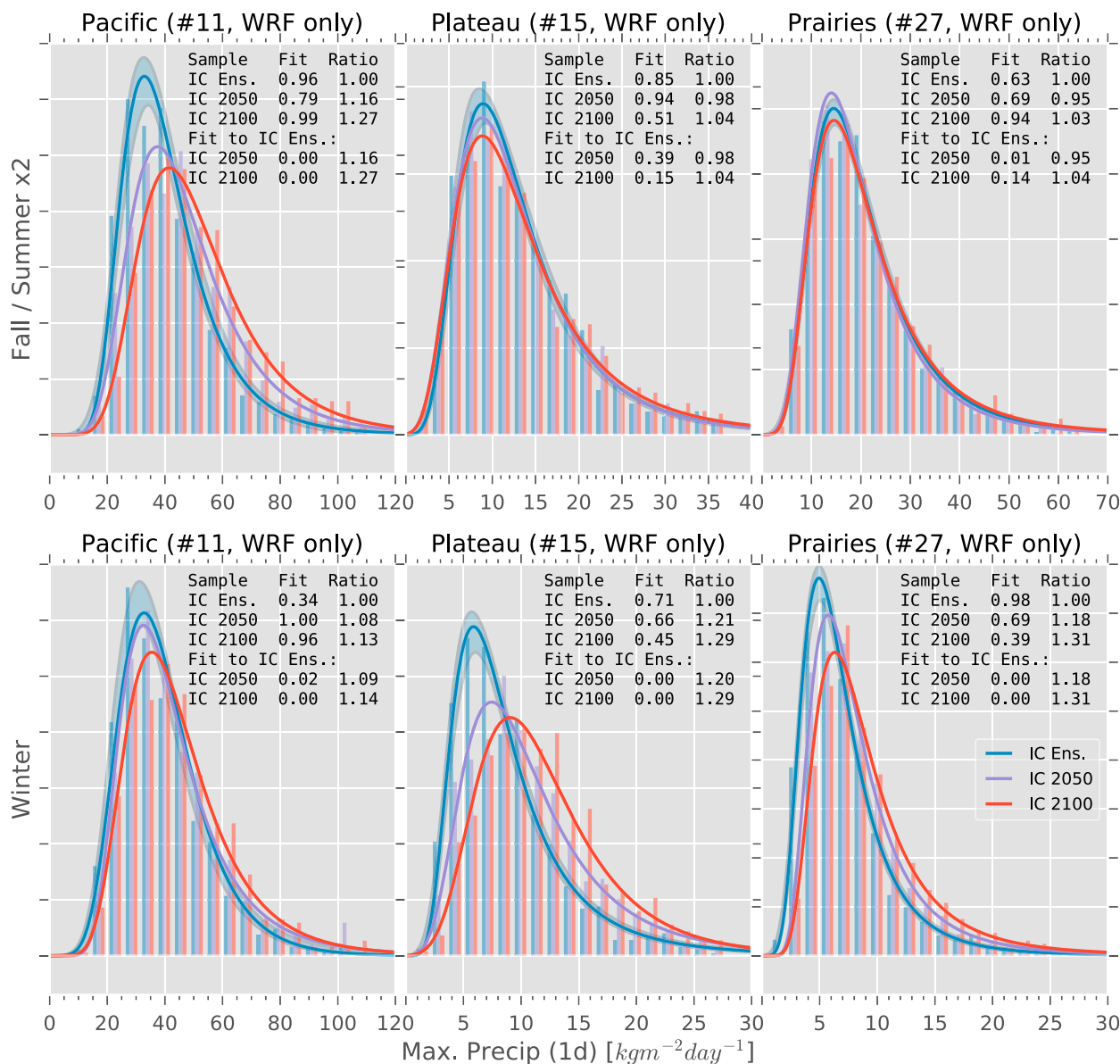


FIG. 5. Distribution of (top) fall (for only Pacific) and summer (plateau and Prairies) and (bottom) winter maxima of daily average precipitation: projection for midcentury and end of the century (no rescaling) in the (left) Pacific, (center) plateau, and (right) Prairies regions: IC ensemble (blue), IC 2050 (purple), and IC 2100 (red). The colored bands are 95% confidence intervals, derived from bootstrapping. The reference sample is the historical IC ensemble. The number of stations (#) and the number of samples (yrs) for each cluster-season are indicated in the panel titles.

highest in the Pacific cluster, but that is partly because it is the smallest, so the differences are statistically less significant. After rescaling all models/domains achieve an acceptable fit in the Pacific cluster. This is also the case for the Prairies and plateau clusters in summer because summer precipitation is in general more homogeneous, so samples can be more easily pooled (also see section S2.3 in the supplementary material). It is also

apparent that there is no relationship between resolution and the quality of fit. This is plausible because summer precipitation is largely convective, which is parameterized in all models, so summer precipitation and extremes depend more on the quality of the parameterization than the resolution of the model. The quality of fit in winter in the plateau and the Prairies is much lower, especially for CESM, and here a clear

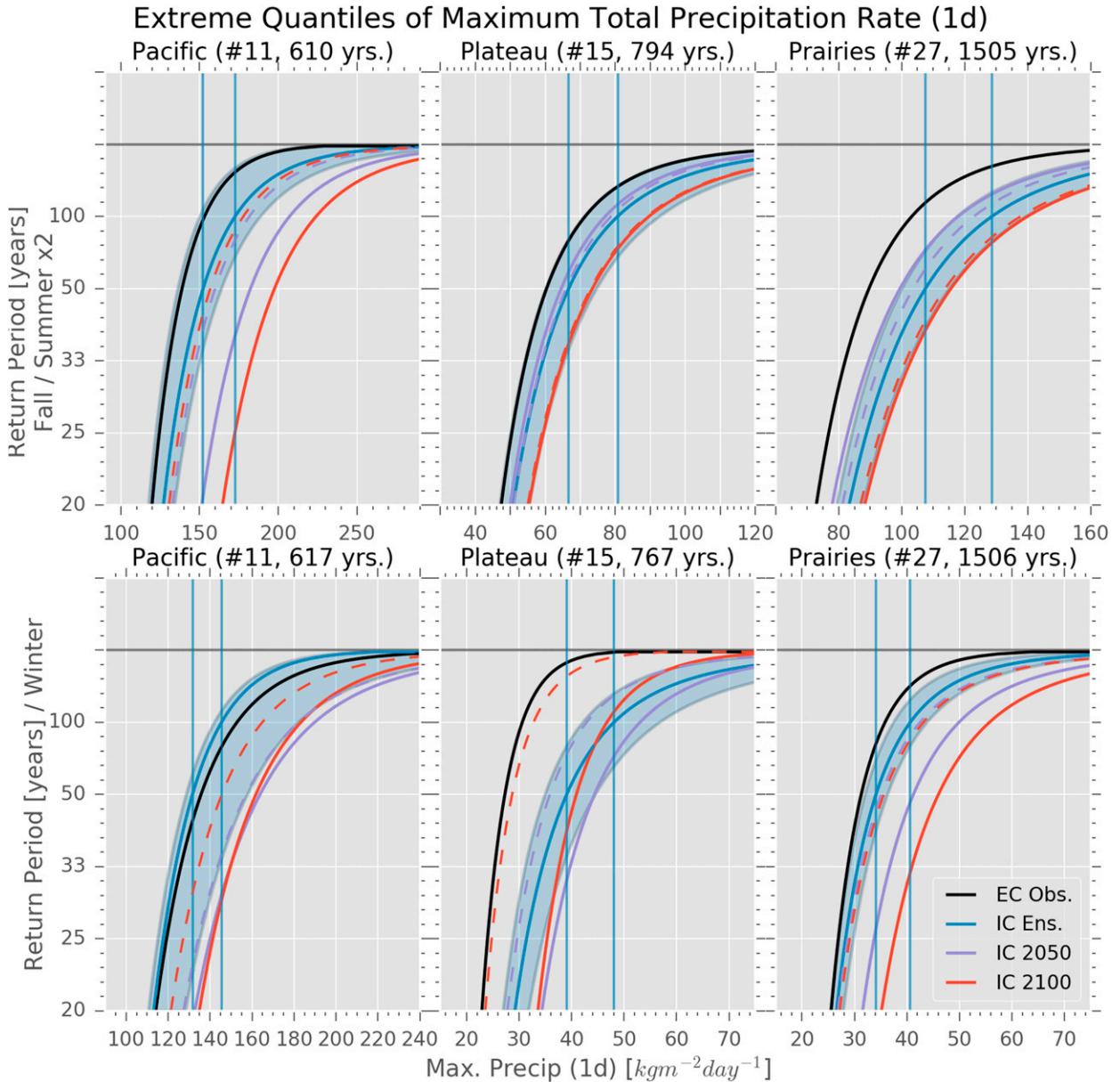


FIG. 6. Extreme quantiles of maxima of daily average precipitation (top) fall (for only Pacific) and summer (plateau and Prairies) and (bottom) winter in the (left) Pacific, (center) plateau, and (right) Prairies regions: EC observations (black), IC ensemble (blue), IC 2050 (purple), and IC 2100 (red). The model distributions have been rescaled to match the mean of the observed distribution: the curves shown in solid lines were rescaled with a single factor (equivalent to the dashed lines in Fig. 4). The curves shown in dashed lines were all rescaled separately, so their distribution means are all identical and equal to the observations in order to facilitate comparison of the shape of the distributions. Confidence intervals and annotation as in Fig. 5.

dependence on resolution is apparent. In particular in the rugged terrain of the interior plateau, very high resolution is necessary and only the inner WRF domain achieves an acceptable fit. In the Prairies, on the other hand, even the resolution of the outer domain (30 km) appears sufficient. The poor performance of CESM (even after rescaling) can probably be attributed to an inadequate representation of the topography, where the

interior plateau is essentially missing and the rain shadow in the lee of the Rocky Mountains is underestimated.

These results suggest that summer precipitation is adequately parameterized in the regional model and the global model, and the added value of higher resolution appears to be small. However, we point out that the spatial variability is much better represented in the regional model (cf. Fig. 5 in Erler et al. 2015) and that the

## Distribution of Maximum Total Precipitation Rate (1d)

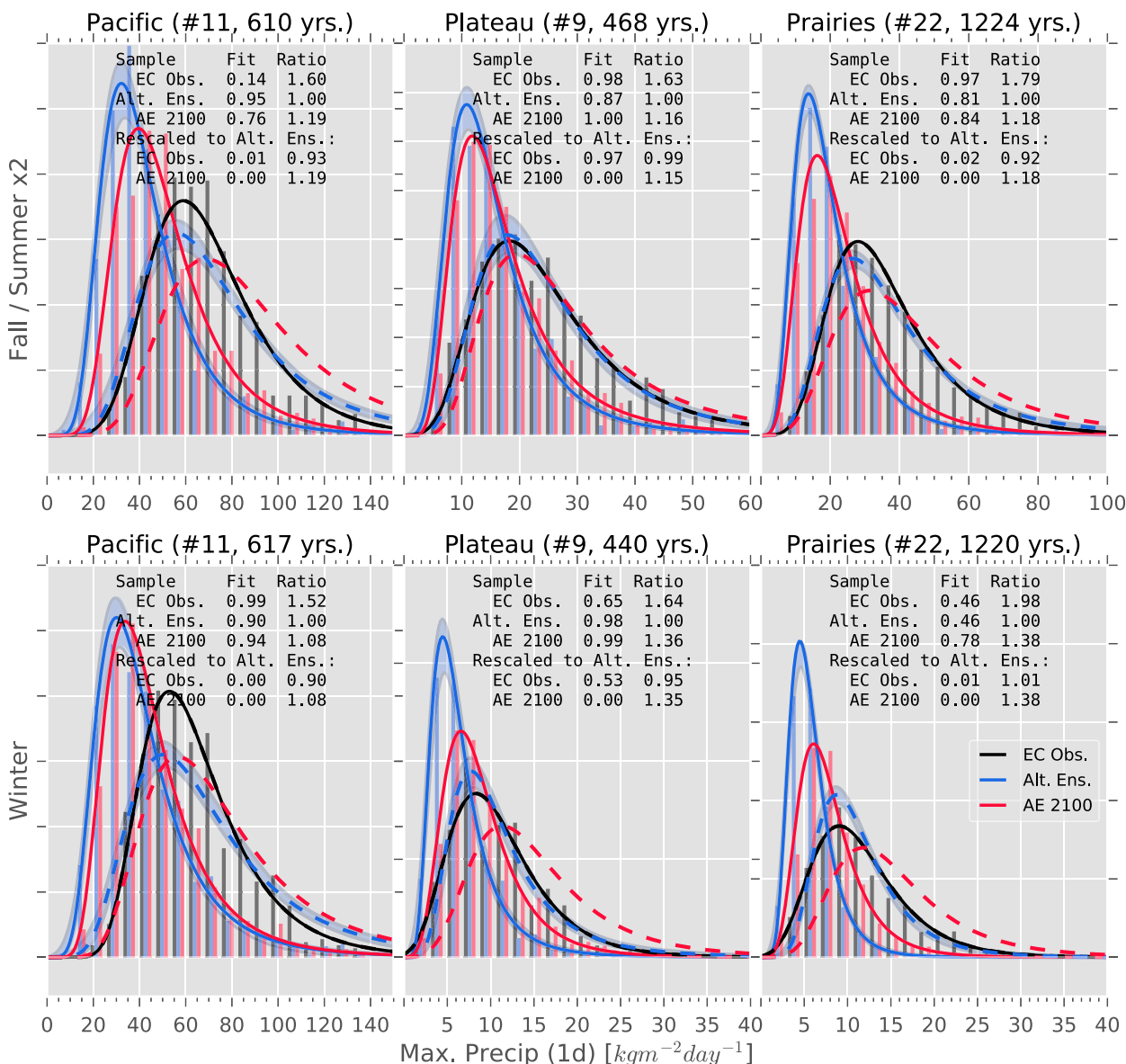


FIG. 7. Distribution (top) fall (for only Pacific) and summer (plateau and Prairies) and (bottom) winter maxima of daily average precipitation for the alternate IC ensemble (otherwise analogous to Figs. 4 and 5) in the (left) Pacific, (center) plateau, and (right) Prairies regions: EC observations (black), Alt ensemble (blue), and AE 2100 (red). The historical ensemble serves as a reference sample.

climate change signal in the global model is opposite in much of our region of interest. In winter, on the other hand, the regional model is clearly superior, and simple rescaling is not able to correct the global model. Here the added value of higher resolution is obvious and even higher resolution would likely be beneficial. In summer higher resolution might be beneficial at scales where convection is resolved explicitly. These conclusions are robust across most clusters and alternative clustering algorithms. It should be pointed out, however, that

statistically significant differences to observations also appear for the IC ensemble in the Pacific cluster in summer and in cluster 5 for most seasons (despite bias correction).

## 2) PROJECTION

The projected PDFs for the two future time periods (2045–60 and 2085–2100) are shown in Fig. 5; the model distributions for the historical period are shown for comparison; they also serve as the reference sample in this figure. Note that the  $x$ -axis scale is different from



Fig. 4, because the samples have not been bias corrected (rescaled). Low  $p$  values (e.g.,  $p < 0.05$ ) in the “Fit to IC Ens.” parts of the figure indicate that the null hypothesis that the distribution does not change can be rejected at a significance level  $\alpha = p$  (i.e., there is a detectable climate change signal if the  $p$  value is close to zero). There is a clear increase in the distribution means in all clusters in winter and fall, but no statistically significant change in summer in the plateau and Prairies. The confidence intervals (95%) confirm the K–S test results: changes are significant where the confidence bands are well separated. Again, the changes in extremes are consistent with the changes in the mean: roughly 30% increase in winter extremes in the plateau and the Prairies, and in fall at the coast; 15% increase in winter at the coast and essentially no changes in summer (by the end of the century). Considering that the temperature increase in winter is about 5°C, this implies a 6% °C<sup>-1</sup> increase in winter—very close to the Clausius–Clapeyron limit of 7% °C<sup>-1</sup> (Trenberth et al. 2003)—and is expected under the hypothesis that winter precipitation is transport limited. The climate change response in CESM is similar in winter and fall, albeit of a smaller magnitude (15%–20% increase). In summer, however, the CESM ensemble shows a strong and statistically significant decrease in precipitation extremes of 15%–20%. This is at odds with the first WRF ensemble, but again consistent with the changes in the mean (in CESM). See section 6 for further discussion of differences between the three ensembles.

The differing trends in summer and winter precipitation found in our simulations are largely consistent with scaling relationships derived from historical observations by Berg et al. (2009), who analyzed daily summer and winter precipitation intensity as a function of surface temperature for several regions in Europe and found that winter precipitation over northern Europe, a region which is climatologically similar to British Columbia, scales with surface temperature at a rate that is close to but somewhat smaller than the Clausius–Clapeyron relation, consistent with our results. For summer precipitation intensity, they do not find a consistent scaling relationship to surface temperature.

It is interesting to note that in summer in the Prairies, there is a significant change to weaker precipitation extremes at midcentury (but not at the end of the twenty-first century). It is, however, not clear whether this is a representative result, because one of the ensemble members simulates severe drought in the Prairies at midcentury (WRF-1 in Erler et al. 2015). This happens neither at the end of the century nor in any of the other ensemble members. Hence, it is possible that the ensemble mean at midcentury might be affected by an outlier. To determine the frequency of such outliers,

additional IC ensemble members would be necessary (the exceptional dryness is also evident in the driving CESM ensemble member).

Extremes in solid precipitation (snow; not shown) are less consistent between station clusters, but they are consistent with changes in temperature and precipitation. In the Pacific and plateau clusters, solid precipitation extremes do not increase, because the increase in precipitation is accompanied by a transition from solid to liquid precipitation as temperatures increase. In the Prairies, on the other hand, where winter temperatures are significantly lower, winter precipitation remains predominantly snow and the increase in solid precipitation extremes is the same as in total precipitation extremes.

### b. Extreme quantiles

Figure 6 characterizes the tail behavior of the distributions; it shows the CDFs of daily precipitation intensities of rare (decadal return period) events. The model results have been bias corrected (rescaled), so the historical model distributions have the same mean as the observations; the simulated distributions for the projection time periods have been rescaled by the same factor so that they can be usefully compared. The simulated 50- and 100-yr return period events for the historical period are marked by vertical lines; the projected future return period for such an event is given by the  $y$ -axis location where the vertical line intersects the CDF of the projection period. Note that the relative change in return periods is not affected by bias correction. As before, the semitransparent bands indicate 95% confidence intervals with respect to sampling uncertainty. Because the most extreme quantiles of a distribution are constrained only by a few extreme events, the sampling uncertainty is larger than for the bulk of the distribution. In half of the clusters/seasons, the tail behavior of the observations and the historical simulations is different, even after rescaling. However, as the otherwise acceptable fit indicates, these differences are statistically not significant (based on the K–S test; note that those with a more similar tail behavior have a higher  $p$  value in Fig. 4). The tail behavior of historical and future projections is clearly separated in fall in the Pacific cluster with a decrease in return periods by a factor of 3–4. In winter at the coast and in the Prairies, the return periods decrease by a factor of 2–3. In summer in the plateau and the Prairies, and interestingly also in the plateau in winter, the separation is weak and the change in return periods is small. Table 1 lists return periods for events that have a return period of 20, 50, and 100 yr in the (bias corrected) historical simulation.

It has been pointed out repeatedly that changes in extremes generally follow changes in the mean. To

TABLE 1. The return periods of extreme daily precipitation event, which have a return period of 20 or 50 yr in the historical simulations (bias corrected, for comparison with the observed return periods). The errors margins are based on 95% confidence intervals from bootstrapping, but in the interest of readability positive and negative errors have been averaged; typically positive errors are about twice as large as negative errors. (Obs. is observations; Ens. is ensemble.)

Experiment	Pacific Fall	Plateau Summer	Prairies Summer	Pacific Winter	Plateau Winter	Prairies Winter
Daily precipitation events with 20-yr return period						
EC Obs.	33 ± 12	26 ± 8	35 ± 6	19 ± 6	65 ± 27	23 ± 4
IC Ens.	20 ± 7	20 ± 6	20 ± 4	20 ± 6	20 ± 6	20 ± 3
IC 2050	9 ± 2	21 ± 6	25 ± 4	11 ± 2	12 ± 2	12 ± 2
IC 2100	7 ± 2	16 ± 4	17 ± 3	9 ± 2	11 ± 3	8 ± 2
Daily precipitation events with 50-yr return period						
EC Obs.	117 ± 64	75 ± 37	122 ± 38	41 ± 25	443 ± 352	70 ± 22
IC Ens.	50 ± 34	50 ± 20	50 ± 14	50 ± 20	50 ± 22	50 ± 12
IC 2050	21 ± 8	57 ± 18	68 ± 19	19 ± 4	31 ± 12	26 ± 5
IC 2100	14 ± 4	36 ± 13	39 ± 9	18 ± 6	40 ± 22	18 ± 4
Daily precipitation events with 100-yr return period						
EC Obs.	330 ± 228	169 ± 120	320 ± 140	71 ± 63	2404 ± 3256	177 ± 86
IC Ens.	100 ± 107	100 ± 53	100 ± 36	100 ± 46	100 ± 58	100 ± 29
IC 2050	38 ± 23	120 ± 48	145 ± 51	30 ± 8	68 ± 33	48 ± 12
IC 2100	25 ± 10	69 ± 30	73 ± 21	29 ± 14	117 ± 128	33 ± 8

investigate the impact of changes in the mean of the distribution as opposed to changes in the shape, we have added versions of the projected CDFs that have been rescaled to the same mean as the observations (so all periods have the same mean); these rescaled CDFs are shown as dashed lines (same color, no error bands) in Fig. 6. During fall and winter, the rescaled CDFs have considerably lower intensities than the original (bias corrected) curves; this implies that the increase in the means contributes significantly to the change in return periods. In fall, at the coast and in winter in the Prairies, the rescaled curves lie within the sampling uncertainty of the historical distribution, suggesting that essentially all change is simply due to an increase in the mean. In the plateau cluster in winter, changes in the shape appear to compensate for the increase in the mean, so there is not much change in decadal return period events. This change in shape could be associated with a change in orographic forcing or the transition from snow to rain. In summer, on the other hand, virtually all change in the most extreme quantiles is due to a fattening of the tails, without a significant shift in the mean (the dashed lines almost exactly overlap the solid lines). As will be discussed in section 6b, this is likely associated with an increase in the fraction of convective precipitation.

## 6. Sensitivity to moist physics

### a. The alternate IC ensemble

To test the robustness of the results, an alternate initial condition ensemble using different convective and

microphysics parameterizations and several sensitivity tests with different physical parameterization schemes was integrated. Figure 7 shows the distribution of seasonal maxima of daily total precipitation from the alternate WRF ensemble. The figure is analogous to Figs. 4 and 5, except that the validation and the projection periods have been combined and bias-corrected distributions are shown for all periods; for clarity, the midcentury period has been omitted. The reference sample in Fig. 7 is the historical ensemble. It is evident that significant differences from observations occur during the cold seasons (fall and winter) in all clusters. The differences occur despite bias correction and are most prominent at the coast, where the tails of the distributions are too large; in the Prairies in winter, on the other hand, the distribution is too narrow and the tail is too small. Since the driving CESM simulations are the same as in the first WRF ensemble (IC) and convective precipitation is small during the cold seasons, the differences in the distribution shape can likely be attributed to the microphysics scheme: the alternate IC ensemble uses the simpler single-moment WSM6 scheme. It should be noted here that mean winter precipitation is generally modeled more accurately by the WSM6 scheme. However, it appears plausible that a simpler but well-calibrated microphysics scheme, such as WSM6, cannot adequately capture the full distribution of precipitation, even though it reproduces the monthly climatology more accurately. The newer Morrison two-moment scheme, on the other hand, may not be sufficiently well calibrated to reproduce the mean accurately, but it also appears to have weaknesses at the

lower end of the intensity spectrum (i.e., too much drizzle; see section S1.2 in the supplementary material). Nevertheless, despite the mismatch with observations, the climate change response in winter is essentially the same in both WRF ensembles, so the climate change signal can be regarded as robust. The increase is somewhat weaker in CESM (15%–20%) but of the same sign.

The results for summer are more disconcerting. As with CESM and the first WRF ensemble, the fit to observations is quite good in summer (after bias correction), but the climate change response is very different: the alternate IC ensemble simulates a considerably larger and statistically significant increase in summer precipitation in the Plateau and Prairies clusters (15%–20%). The CESM ensemble, on the other hand, shows a statistically significant decrease of almost 20% in the plateau and Prairies clusters (cf. Fig. 3). However, because all three ensembles show good agreement with observations during the historical period, there is no immediately obvious reason to prefer one projection over any other (in summer). The reasons for the differing trends in summer extremes in the two WRF ensembles are related to the different cumulus schemes employed, and they will be discussed in section 6b. The reasons for the decrease in CESM are more complex and are beyond the scope of this analysis. We note, however, that CESM shows a similar decrease in monthly mean summer precipitation.

It is interesting to note that Mladjic et al. (2011) report a similar increase in summer precipitation extremes over the Canadian Prairies as seen in our alternate IC ensemble; the Canadian RCM in the configuration used by Mladjic et al. (2011) also employs the Kain–Fritsch cumulus scheme.

### *b. The role of convective precipitation*

Total summer precipitation extremes in the plateau and Prairies clusters do not change significantly in the first IC ensemble, but they do in the alternate IC ensemble, which utilizes a different cumulus scheme. Figure 2 reveals that the fraction of convective (or cumulus) precipitation increases significantly, for both the mean and extremes. Similarly, there is also a considerable increase in cumulus precipitation in fall in the Pacific cluster. Figure 8 shows the distribution of seasonal maxima of cumulus precipitation in the (first) IC ensemble: fall for the Pacific cluster and summer for the plateau and Prairies clusters. These are also the seasons when cumulus precipitation peaks, so they are equivalent to annual maxima. Note that cumulus or convective precipitation in this context is defined as the unresolved, parameterized component of precipitation; nonconvective or stratiform (i.e., resolved or grid

scale) precipitation is shown with dashed lines for comparison (same color, no confidence intervals). The reference sample is the historical IC ensemble. It is evident that there is a considerable increase in convective extremes in all regions and in particular at the end of the century, although it is not as high as the increase seen in winter precipitation, except at the coast, where the increase in convective precipitation is significantly larger than the already large increase in total and nonconvective precipitation. It is not clear to what extent the increase is still affected by the SST and sea ice interpolation error mentioned earlier, but similar changes are also seen in the four stations on the west side of Vancouver Island (cluster 3), where no interpolation errors have been found. On the other hand, in the alternate IC ensemble the increase at the coast is still present, but it is not as strong. A similar increase in convective precipitation with surface temperature (and compensating decrease in stratiform precipitation) was also found by Berg et al. (2009) in historical regional climate simulations over Europe. The increase in convective precipitation with temperature occurs in similar magnitude in all the cumulus schemes that were tested: Grell-3, Kain–Fritsch, and the Tiedtke scheme; the increase is always accompanied by an increase in the fraction of convective precipitation. The latter is even the case in CESM, where convective precipitation in the area of interest decreases, but the decrease is less than the decrease in total precipitation (approximately 10% points less).

However, the fraction of cumulus precipitation in summer appears to be sensitive to the choice of the convective parameterization: the fraction of convective precipitation in the first IC ensemble (using the Grell-3 scheme) is only about 55% (cf. Fig. 2), while it is approximately 70% in the alternate IC ensemble (using the Kain–Fritsch scheme) and CESM. Unfortunately, data on the type of observed precipitation events are not available for EC stations, so the fraction of cumulus precipitation cannot be validated against observations.

However, using data from weather radar and station observations in Germany, Berg et al. (2013) determined a relationship between the fraction of cumulus precipitation and surface temperature, and they find a steady increase with rising temperatures, consistent with our findings. For a typical surface air temperature of 20°C, their relationship suggests a fraction of cumulus precipitation of about 80%—closer to the alternate IC ensemble.

The higher fraction of cumulus precipitation, combined with the robust increase in convective precipitation (extremes), explains why a significant increase in summer precipitation (extremes) is seen in the alternate WRF ensemble and not in the first WRF ensemble. Given the

## Distribution of Maximum Cumulus Precipitation Rate (1d)

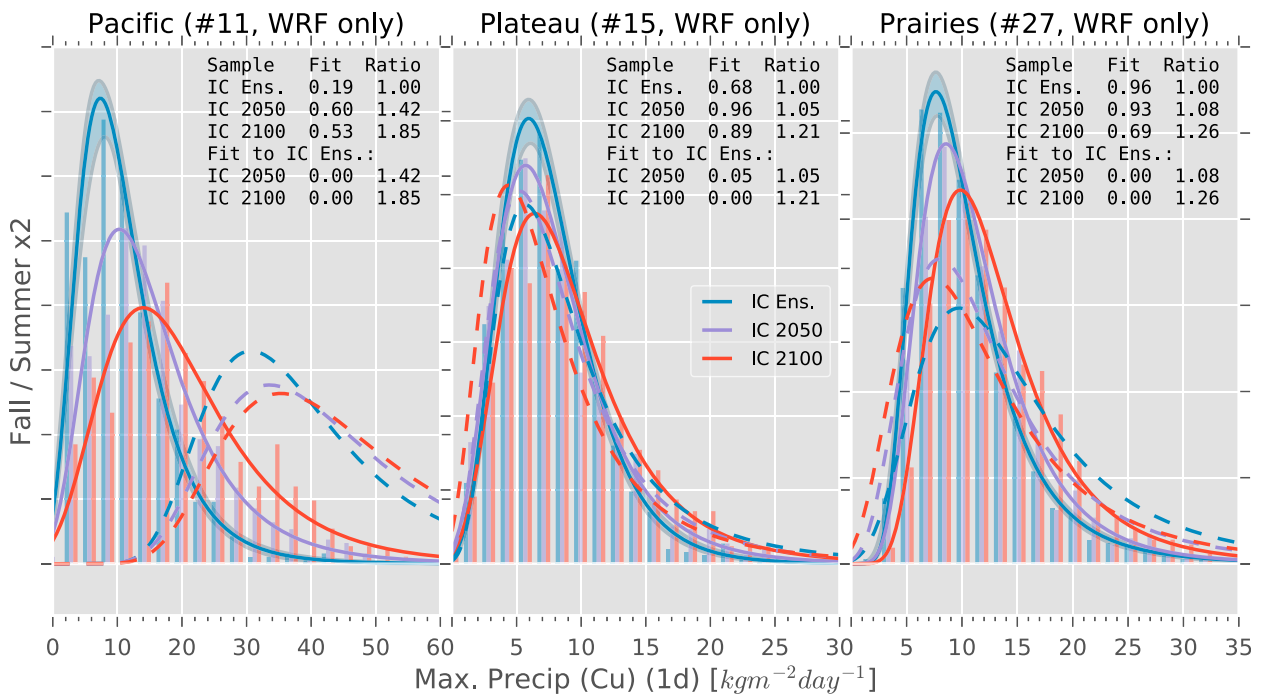


FIG. 8. Distribution of fall (for only Pacific) and summer (plateau and Prairies) maxima of daily cumulus precipitation for the IC ensemble in the (left) Pacific, (center) plateau, and (right) Prairies regions: IC ensemble (blue), IC 2050 (purple) and IC 2100 (red). Only fall–summer is shown; the dashed lines show grid-scale precipitation maxima for comparison. Stations observations are not available for convective and grid-scale precipitation separately. The reference sample is the historical IC ensemble. Confidence intervals and annotation as in Fig. 5.

cumulus fraction estimate based on the temperature dependence derived by Berg et al. (2013), the projections of the alternate IC ensemble appear to be more likely; on the other hand, the spatial pattern produced by the Grell-3 scheme is more realistic and it is an ensemble scheme, which is usually assumed to be more robust. It is clear that the climate change response in summer in midlatitudes is sensitive to the fraction of cumulus precipitation. More observational data that differentiate between stratiform and convective precipitation are needed. Simulations with convection-resolving climate models will also help to resolve this question.

This analysis is based on daily precipitation averages, but the results remain essentially unchanged, if 1-h, 6-h, or 5-day (pentad) averages are employed. This may appear surprising because shorter time scales are usually thought to be dominated by convective events in summer (e.g., Eggert et al. 2015). However, despite the increase in convective fraction, this does not appear to translate into a higher increase for extremes of a shorter time scale.

Finally we note that the Grell-3 cumulus scheme also supports spreading of subsidence to neighboring grid cells; this capability was enabled in the IC ensemble simulations. Disabling subsidence spreading results

only in a small reduction in precipitation intensity, which is not statistically significant.

## 7. Summary and conclusions

In this article we have presented an analysis of precipitation extremes, based on a small initial condition ensemble that was dynamically downscaled to 10-km resolution in two different configurations. Validation against station observations and projections until the end of the twenty-first century under the RCP8.5 scenario were discussed. A novel approach for pooling of data from stations with a similar mean climatology was introduced and employed throughout the analysis, a technique that lends a high degree of statistical power to our conclusions.

We find that in our domain of interest, western Canada, and across the seasonal cycle, changes in extremes tend to follow changes in the mean. Previous studies have typically found a larger increase in extremes than in the mean; this can be partially explained by the fact that many previous studies have considered only annual means and annual maxima. We, on the other hand, find significant differences in precipitation changes between seasons (and regions).

Changes in summer precipitation are not consistent between models and thus they are subject to significant model uncertainty, despite the large volume of data in each analysis cluster. There appears to be a transition from decreasing summer precipitation in the south to an increase in the north, for both mean and extreme precipitation, but the transition point is located significantly farther north in CESM than in WRF. As a consequence, in our region summer precipitation decreases in CESM, while the alternate WRF ensemble shows a significant increase and the first WRF ensemble shows no statistically significant change. There is, however, a major shift from nonconvective to convective precipitation, as global warming progresses; this shift appears to be robust between different models and convection schemes. However, the impact of this increase depends on the fraction of convective precipitation, which in turn depends on the convection scheme. Hence, a more active convection scheme leads to a larger increase in projected summer precipitation. Furthermore, there is evidence for a fattening of the tail of the distribution of summer precipitation extremes, even if there is no increase in the mean of the distribution. The agreement with observations is adequate in all three ensembles and no dependence on resolution is evident.

In fall and winter, on the other hand, we find a substantial and consistent increase in precipitation across the domain: a 30% increase at the coast in fall and in winter over the plateau and the Prairies, and only about a 10% increase in winter at the coast and in fall land inward. The increase in precipitation extremes is the same as the increase in the mean. The strong increase in fall precipitation is particularly relevant at the coast, where total precipitation also peaks in fall and may increase the flood risk. There is no evidence for a fattening of the tails of the distributions in winter. It is further evident that the distribution of winter precipitation extremes is sensitive to resolution and there is significant added value in higher-resolution regional simulations: CESM is unable to capture the observed distribution of winter precipitation extremes, even after linear bias correction, and the rain shadow over the interior plateau is missed entirely. In fact, in four out of six cases, CESM deviates so much from the actual climatology that it is not possible to fit a GEV distribution to the pooled data. WRF, on the other hand, is able to reproduce the observed distribution of fall and winter precipitation extremes much better, and, after linear bias correction, is statistically indistinguishable from the observations in many cases (all, for the first ensemble at 10-km resolution).

In summary, a robust and significant increase in winter precipitation is projected, while climate change projections for summer are confounded by high model

uncertainty. The added value of higher resolution is clearly evident in the distribution of winter precipitation (extremes), while there appears to be no added value of higher resolution in summer—at least as long as convection is parameterized.

*Acknowledgments.* The simulations presented in this paper were performed at the SciNet High Performance Computing facility at the University of Toronto, which is a component of the Compute Canada HPC platform. The authors thank Dr. J. Gula for assistance with the initial setup of the modeling chain, and Dr. D. Gruner and the SciNet team for assistance during the setup and operation of WRF. Furthermore, the assistance of Dr. M. d'Orgeville and Dr. G. Vettoretti in providing and processing CESM data was invaluable and is gratefully acknowledged. ARE was partially supported by an SOSCIIP Graduate Student Fellowship. The research of WRF at the University of Toronto is supported by NSERC Discovery Grant 9627.

#### REFERENCES

- Berg, P., J. Haerter, P. Thejll, C. Piani, S. Hagemann, and J. Christensen, 2009: Seasonal characteristics of the relationship between daily precipitation intensity and surface temperature. *J. Geophys. Res.*, **114**, D18102, doi:[10.1029/2009JD012008](https://doi.org/10.1029/2009JD012008).
- , C. Moseley, and J. O. Haerter, 2013: Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.*, **6**, 181–185, doi:[10.1038/ngeo1731](https://doi.org/10.1038/ngeo1731).
- Bernard, E., P. Naveau, M. Vrac, and O. Mestre, 2013: Clustering of maxima: Spatial dependencies among heavy rainfall in France. *J. Climate*, **26**, 7929–7937, doi:[10.1175/JCLI-D-12-00836.1](https://doi.org/10.1175/JCLI-D-12-00836.1).
- Christensen, J. H., and Coauthors, 2013: Climate phenomena and their relevance for future regional climate change. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1217–1308.
- Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Vol. 208, Springer, 209 pp., doi:[10.1007/978-1-4471-3675-0](https://doi.org/10.1007/978-1-4471-3675-0).
- Cooley, D., 2009: Extreme value analysis and the study of climate change. *Climatic Change*, **97**, 77–83, doi:[10.1007/s10584-009-9627-x](https://doi.org/10.1007/s10584-009-9627-x).
- Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, doi:[10.1007/s00382-010-0977-x](https://doi.org/10.1007/s00382-010-0977-x).
- Dominguez, F., E. Rivera, D. P. Lettenmaier, and C. L. Castro, 2012: Changes in winter precipitation extremes for the western United States under a warmer climate as simulated by regional climate models. *Geophys. Res. Lett.*, **39**, L05803, doi:[10.1029/2011GL050762](https://doi.org/10.1029/2011GL050762).
- d'Orgeville, M., W. R. Peltier, A. R. Erler, and J. Gula, 2014: Climate change impacts on Great Lakes Basin precipitation extremes. *J. Geophys. Atmos.*, **119**, 10799–10812, doi:[10.1002/2014JD021855](https://doi.org/10.1002/2014JD021855).
- Efron, B., and R. J. Tibshirani, 1994: *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability, Vol. 57, Chapman & Hall/CRC, 456 pp.

- Eggert, B., P. Berg, J. Haerter, D. Jacob, and C. Moseley, 2015: Temporal and spatial scaling impacts on extreme precipitation. *Atmos. Chem. Phys.*, **15**, 5957–5971, doi:10.5194/acp-15-5957-2015.
- Endo, H., A. Kitoh, T. Ose, R. Mizuta, and S. Kusunoki, 2012: Future changes and uncertainties in Asian precipitation simulated by multiphysics and multi-sea surface temperature ensemble experiments with high-resolution Meteorological Research Institute atmospheric general circulation models (MRI-AGCMS). *J. Geophys. Res.*, **117**, D16118, doi:10.1029/2012JD017874.
- Erler, A. R., and W. R. Peltier, 2015: Clustering of station observations for extreme value analysis. *Proc. Fifth Int. Workshop on Climate Informatics (CI 2015)*, Boulder, CO, NCAR, 16. [Available online at <https://www2.cisl.ucar.edu/sites/default/files/16-%20Erler.pdf>.]
- , —, and M. d'Orgeville, 2015: Dynamically downscaled high-resolution hydroclimate projections for western Canada. *J. Climate*, **28**, 423–450, doi:10.1175/JCLI-D-14-00174.1.
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.
- Grell, G. A., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1693, doi:10.1029/2002GL015311.
- Gula, J., and W. R. Peltier, 2012: Dynamical downscaling over the Great Lakes Basin of North America using the WRF Regional Climate Model: The impact of the Great Lakes system on regional greenhouse warming. *J. Climate*, **25**, 7723–7742, doi:10.1175/JCLI-D-11-00388.1.
- Hartigan, J. A., and M. A. Wong, 1979: Algorithm AS 136: A K-means clustering algorithm. *J. Roy. Stat. Soc.*, **28C**, 100–108.
- Herring, S. C., M. P. Hoerling, T. C. Peterson, and P. A. Stott, 2014: Explaining extreme events of 2013 from a climate perspective. *Bull. Amer. Meteor. Soc.*, **95** (9), S1–S104, doi:10.1175/1520-0477-95.9.S1.1.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- Hosking, J. R. M., and J. R. Wallis, 1997: *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, 240 pp.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, doi:10.1029/2008JD009944.
- IPCC, 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*. C. B. Field et al., Eds., Cambridge University Press, 582 pp.
- Jones, E., and Coauthors, 2001: SciPy: Open source scientific tools for Python. Accessed 9 June 2015. [Available online at <http://www.scipy.org/>.]
- Jones, M. R., S. Blenkinsop, H. J. Fowler, and C. G. Kilsby, 2014: Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall. *Int. J. Climatol.*, **34**, 751–765, doi:10.1002/joc.3720.
- Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.
- Katz, R. W., 2010: Statistics of extremes in climate change. *Climatic Change*, **100**, 71–76, doi:10.1007/s10584-010-9834-5.
- , M. B. Parlange, and P. Naveau, 2002: Statistics of extremes in hydrology. *Adv. Water Resour.*, **25**, 1287–1304, doi:10.1016/S0309-1708(02)00056-8.
- Kharin, V. V., F. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, doi:10.1007/s10584-013-0705-8.
- Köppen, W., 1936: *Das Geographische System der Climate*. Vol. 1, *Handbuch der Klimatologie*, Gebrüder Borntraeger, 44 pp.
- Livezey, R. E., and W. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59, doi:10.1175/1520-0493(1983)111<0046:SFSaid>2.0.CO;2.
- Lucas-Picher, P., D. Caya, R. de Elía, and R. Laprise, 2008: Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain. *Climate Dyn.*, **31**, 927–940, doi:10.1007/s00382-008-0384-8.
- Massey, F. J., Jr., 1951: The Kolmogorov–Smirnov test for goodness of fit. *J. Amer. Stat. Assoc.*, **46**, 68–78, doi:10.1080/01621459.1951.10500769.
- Mekis, É., and L. A. Vincent, 2011: An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmos.–Ocean*, **49**, 163–177, doi:10.1080/07055900.2011.583910.
- Mladjic, B., L. Sushama, M. Khaliq, R. Laprise, D. Caya, and R. Roy, 2011: Canadian RCM projected changes to extreme precipitation characteristics over Canada. *J. Climate*, **24**, 2565–2584, doi:10.1175/2010JCLI3937.1.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, doi:10.1175/2008MWR2556.1.
- Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.
- Schindler, A., A. Toreti, M. Zampieri, E. Scoccimarro, S. Gualdi, S. Fukutome, E. Xoplaki, and J. Luterbacher, 2015: On the internal variability of simulated daily precipitation. *J. Climate*, **28**, 3624–3630, doi:10.1175/JCLI-D-14-00745.1.
- Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032, doi:10.1175/MWR2830.1.
- Smirnov, N. V., 1939: On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Moscow Univ. Math. Bull.*, **2**, 3–14.
- Smirnov, N., 1948: Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Stat.*, **19**, 279–281, doi:10.1214/aoms/1177730256.
- Stephens, G. L., and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211, doi:10.1029/2010JD014532.
- Stott, P., 2015: Attribution: Weather risks in a warming world. *Nat. Climate Change*, **5**, 517–518, doi:10.1038/nclimate2640.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a. [Available online at [https://ams.confex.com/ams/84Annual/techprogram/paper\\_69061.htm](https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm).]

- Torma, C., F. Giorgi, and E. Coppola, 2015: Added value of regional climate modeling over areas characterized by complex terrain—Precipitation over the Alps. *J. Geophys. Res. Atmos.*, **120**, 3957–3972, doi:10.1002/2014JD022781.
- Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1217, doi:10.1175/BAMS-84-9-1205.
- van der Ent, R. J., H. H. Savenije, B. Schaefli, and S. C. Steele-Dunne, 2010: Origin and fate of atmospheric moisture over continents. *Water Resour. Res.*, **46**, W09525, doi:10.1029/2010WR009127.
- Vincent, L. A., and E. Mekis, 2006: Changes in daily and extreme temperature and precipitation indices for Canada over the twentieth century. *Atmos.–Ocean*, **44**, 177–193, doi:10.3137/ao.440205.
- von Storch, H., and F. W. Zwiers, 2002: *Statistical Analysis in Climate Research*. Cambridge University Press, 496 pp.
- Wang, W., and Coauthors, 2012: ARW version 3 modeling system user's guide. Version 3.3, Tech. Rep., Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, 371 pp. [Available online [http://www2.mmm.ucar.edu/wrf/users/docs/user\\_guide\\_V3.3/ARWUsersGuideV3.pdf](http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3.3/ARWUsersGuideV3.pdf).]