INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2017)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.5028

RMetS
Royal Meteorological Society

An analysis of the performance of RCMs in simulating current climate over western Canada

Elaine M. Barrow* and David J. Sauchyn

Prairie Adaptation Research Collaborative, University of Regina, Canada

ABSTRACT: The performance of eight National Center for Environmental Prediction (NCEP2) reanalysis-driven regional climate models (RCMs), seven from the North American Regional Climate Change Program (NARCCAP) and one from the Coordinated Regional Downscaling Experiment (CORDEX), in simulating the 1980–2004 climate of western Canada was assessed at a number of spatial and temporal scales. Results indicated that the RCMs were more successful at capturing the seasonal spatial distribution of mean temperature than precipitation and that inaccuracies in the spatial distribution of the summer climate moisture index were likely due to the errors in precipitation distribution and amount. All RCMs performed less well in simulating summer precipitation, most likely due to continued problems with the simulation of convective precipitation.

At the grid box scale, quantile–quantile (q–q) plots for temperature indicated that all RCMs showed very similar distributions to observed but with warm or cold biases, and errors in the simulation of a number of temperature-based extremes indices were related to these biases. For precipitation, q–q plots indicated that most RCMs overestimated precipitation totals, and while tending to follow the observed quantiles at smaller precipitation amounts, they diverged at larger precipitation totals. Performance in simulating the precipitation-based extremes indices depended largely on whether or not a RCM over- or under-estimated precipitation totals – with those RCMs simulating too much precipitation underestimating the number of consecutive dry days and dry day persistence, and vice versa.

Despite improvements in RCM resolution and parameterisation schemes, this work indicates that the simulation of precipitation in particular is still problematic in western Canada. This implies that scenarios of climate change constructed from RCM output require some form of bias correction to be of most use in impacts studies.

KEY WORDS RCMs; performance; climate change; scenarios; extremes

Received 3 March 2016; Revised 11 January 2017; Accepted 17 January 2017

1. Introduction

The major impacts of climate change on most sectors (e.g. water resources, agriculture) will very likely be as a result of shifts in climate variability, exhibited through changes in the frequency and magnitude of extreme climatic events (e.g. Colombo *et al.*, 1999; Easterling *et al.*, 2000; Huntingford *et al.*, 2003; Allan and Soden, 2008; Gutowski *et al.*, 2010; IPCC, 2012; Murdock *et al.*, 2013). Studies of climate change impacts have predominantly used scenarios describing changes in mean climate (e.g. IPCC, 2007; IPCC, 2014a, 2014b; Lemmen *et al.*, 2008; and many others), although Katz and Brown (1992) and Semenov and Porter (1995) recognized over 20 years ago that shifts in climate variability have a greater effect on the frequency of extreme climatic events than do changes in mean values. Until relatively recently the main tools available for the construction of physically consistent climate change scenarios were global climate models (GCMs). The spatial scales at which these climate models operate (typically 200–300 km) are much coarser than those of the driving processes of many impacts (Giorgi *et al.*, 2001) and this has led to repeated calls for higher resolution climate information to serve the impacts community (Mearns *et al.*, 2001, 2012). Although GCMs perform well in simulating mean climate and the external forcing of global climate, their coarse spatial resolution limits their capacity to simulate climate variability and extreme events (Kharin and Zwiers, 2000; Huntingford *et al.*, 2003; Zwiers and Zhang, 2003; Kharin *et al.*, 2005; Schaeffer *et al.*, 2005).

Regional climate models (RCMs), operating typically at spatial scales of between 10 and 50 km, and utilizing reanalysis- or GCM-derived initial and lateral meteorological boundary conditions, are able to provide more reliable information on climate variability and extremes (e.g. Christensen and Christensen, 2003; Frei *et al.*, 2003). This improvement over lower resolution climate models (the 'added value'; Flato *et al.*, 2013) is due to the RCMs' better representation of finer-scale topographic features, characteristics of the land surface, and temporal detail, as well as the improved representation of smaller-scale atmospheric processes. These factors generate, or modify, atmospheric circulation at scales vital for the adequate simulation of local extremes (Mearns *et al.*, 2003). However, many of the shortcomings of RCMs (and GCMs) still stem from the lack of explicit representation

* Correspondence to: E. M. Barrow, Prairie Adaptation Research Collaborative, University of Regina, 3737 Wascana Parkway, Regina, Saskatchewan S4S 0A2, Canada. E-mail: elaine.barrow@sasktel.net

of small-scale processes, such as clouds and convection (Randall *et al.*, 2007). And, for climate extremes in particular, there are some processes, including feedbacks and land–atmosphere and ocean–atmosphere interactions, which are still poorly represented and understood (Seneviratne *et al.*, 2012).

Flato *et al.* (2013) indicated that the added value of RCMs is mostly achieved through improved simulation of extremes on small spatial and short temporal scales and of topography-influenced phenomena. For example, Huntingford *et al.* (2003) showed that while coarse-resolution simulations provided an acceptably realistic mean rainfall for a location, the high resolution of an RCM was required for the realistic simulation of extreme rainfall. Also, Rauscher *et al.* (2010) illustrated improvements in the simulation of summer precipitation over Europe when RCM resolution was increased to 25 km from 50 km. Kanada *et al.* (2008), using a 5 km resolution climate model for Japan, showed good agreement between the simulation and observations for total daily precipitation amount and frequency and the temporal and spatial characteristics of maximum daily precipitation between June and October. More recently, Kendon *et al.* (2012) demonstrated that a convection-permitting very high resolution (1.5 km) RCM simulated much more realistic rainfall over the UK when compared to a 12 km resolution RCM, with a much better representation of the duration and spatial extent of heavy rain events. Persistent light rain, a common problem in climate models, and errors in the diurnal cycle were also considerably reduced in the 1.5 km resolution RCM. This implies that for precipitation in particular, the realistic representation of extremes is very much dependent on RCM resolution.

Although the number of RCM experiments available is far fewer than their global model counterparts, they have been the focus of various international research programmes, many of which have included an assessment of RCM performance in simulating current climate. These include the North American Regional Climate Change Program (NARCCAP; Mearns *et al.*, 2007 updated 2014, 2009, 2012), the Coordinated Regional Downscaling Experiment (CORDEX; Giorgi and Gutowski, 2015), Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects (PRUDENCE; Christensen *et al.*, 2002) and Ensemble-based Prediction of Climate Changes and their Impacts (ENSEMBLES; van der Linden and Mitchell, 2009).

This article describes the first step in the process of developing RCM-derived scenarios of climate change for western Canada. Many impacts studies in this region require fine spatial- and temporal-scale climate data and projections of future climate at the same scales. In the first set of guidelines on RCM data use in impacts studies, Mearns *et al.* (2003) recommended the adoption of the delta method of scenario construction (IPCC-TGICA, 2007), which is based on 30-year mean values and combines climate model-derived changes with an observed climate record, and has been widely used with GCM output. In a more recent European study using RCM data to drive a number of impacts models (ranging from energy use to potential biomass productivity), Fronzek and Carter (2007) also cautioned against using RCM data directly due to biases in their representation of current climate. Given the systematic biases within RCMs, it is unlikely that their data will ever be sufficiently reliable to be used directly, i.e. without any further statistical manipulation, but bias correction methods (see Teutschbein and Seibert (2012) for a review and evaluation of techniques) may be applied which do not constrain the future climate information to a particular pattern of climate variability, as is the case with the delta method of scenario construction. The aim of this article, therefore, is to determine how well a number of RCMs perform in simulating current climate in western Canada in order to facilitate the selection of an appropriate bias correction method for the future construction of scenarios of climate change.

Since both GCMs and RCMs are used for a broad range of research studies, a set of measures for assessing the important aspects of climate has not yet been identified (Gleckler *et al.*, 2008). Instead, the combined use of many techniques is recommended to provide a comprehensive picture of climate model performance (Flato *et al.*, 2013). This assessment of RCM performance, therefore, examines various aspects of western Canadian climate at seasonal, monthly and daily timescales. This study uses output from reanalysis-driven RCM experiments which allows year-to-year comparisons with observed data (e.g. Roy *et al.*, 2012). In some cases, RCM control runs (driven with boundary conditions from a GCM) have been used (e.g. Blenkinsop and Fowler, 2007; van Roosmalen *et al.*, 2010), but these can only compare distributions of climate variables and not the synchronicity of the time series. Use of reanalysis boundary conditions to drive a RCM means that errors in the large-scale forcing fields are small and that the observed inter-annual and seasonal variability is incorporated into the RCM (Dulière *et al.*, 2011). For these reasons, reanalysis-driven RCMs are more likely to better simulate observed climate although, given non-linearities in the climate system and errors or parameterisations in a RCM's configuration, the simulated climate cannot ever be expected to exactly match the observed climatology. As current and future drought conditions are the driver of this research, the simulation of precipitation is the focus of this article.

2. Methodology

RCM performance in simulating current climate was assessed at various spatial and temporal scales by comparing RCM data with an observed climatology. For western Canada, seasonal and annual mean temperature and precipitation were compared, while for ten sites (see Figure 1) in southern Alberta (Carway, Gleichen, Lethbridge and Medicine Hat) and southern Saskatchewan (Kindersley, Leader, Outlook, Val Marie, Swift Current and Yellow Grass) a number of indices based on daily precipitation

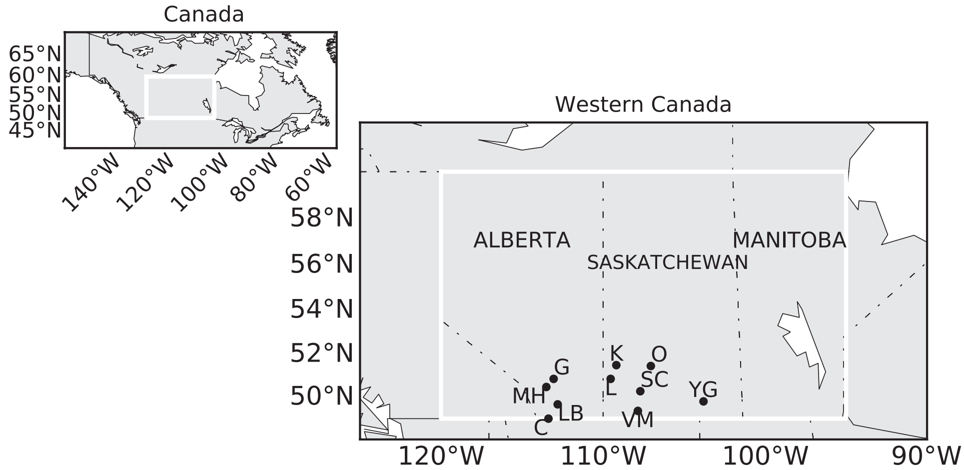

Figure 1. Location of sites used for RCM performance analysis in western Canada (area outlined in white). C, Carway; G, Gleichen; L, Leader; LB, Lethbridge; MH, Medicine Hat; O, Outlook; VM, Val Marie; SC, Swift Current; K, Kindersley; YG, Yellow Grass.

were calculated. These ten sites were chosen so as to be close to two study watersheds, Swift Current Creek, Saskatchewan, and the Oldman River, Alberta, which are the focus of current climate scenario construction work (Sauchyn *et al.*, 2016).

2.1. Data

2.1.1. NCEP-DOE AMIP-II Reanalysis (R-2) data

The National Center for Environmental Prediction-Department of Energy (NCEP-DOE) Atmospheric Model Intercomparison Project-II (AMIP-II) reanalysis (R-2; Kistler *et al.*, 2001; Kanamitsu *et al.*, 2002), henceforth NCEP2, is an updated 6-h global analysis series from 1979 to present which fixes known processing errors in the NCEP-NCAR reanalysis (R-1; Kalnay *et al.*, 1996) and includes an improved forecast model and data assimilation system. This reanalysis is a retrospective record of the atmosphere based on extensive observations, including land surface, ship, rawinsonde and satellite data, and is obtained by running a numerical weather prediction model with these observations. This data set has been used to provide boundary conditions (such as momentum, humidity and wind speed) for the RCMs involved in the NARCCAP (Mearns *et al.*, 2007 updated 2014, 2012) and also as a basic verification data set for the AMIP-II (AMIP Project Office, 1996). Here, these gridded data (at approximately 1.8° latitude/longitude resolution) were compared with output from the NCEP2-reanalysis-driven RCMs to determine the value added to their climate simulations.

2.1.2. RCM data

RCM data were obtained from the NARCCAP (Mearns *et al.*, 2007 updated 2014, 2012) and from the Canadian Centre for Climate Modelling and Analysis (Environment Canada) for the 1980–2004 time period. Details of the seven different RCMs (totalling eight experiments) run for the North American spatial domain at resolutions between 25 and 50 km are given in Table 1.

2.1.3. Observed climate data

2.1.3.1. Gridded observed data: The ANUSPLIN gridded monthly climatology developed by the Canadian Forest Service (McKenney *et al.*, 2001, 2006) is an elevation-dependent and spatially continuous data set with a spatial resolution of 300 arc s (approximately 10 km). It was constructed using ANUSPLIN software (Hutchinson, 1995, 2004) and observed data from the Meteorological Service of Canada and the US National Climatic Data Center. This monthly data set was shown to perform well in a comparative study of four gridded climate normal (1961–1990) data sets for Canada (Milewska *et al.*, 2005), with best agreement in the Prairies which are characterized by relatively flat terrain and high station density. Temperatures were shown to be within 1 °C and precipitation within several percent of station values in this region (Milewska *et al.*, 2005). Seasonal and annual fields were calculated for each year from the ANUSPLIN monthly data set and then used for comparison with each RCM at these time scales.

2.1.3.2. Observed station data: Daily observed station data from Environment Canada's second generation adjusted daily precipitation data set (Mekis and Vincent, 2011) and the second generation homogenized temperature data sets (Vincent *et al.*, 2012) were used in this analysis for the calculation of the extremes indices listed in Table 2 at the ten sites previously mentioned and for the calculation of skill scores (see Section 2.3). For the latter analysis, 105 temperature stations and 127 precipitation stations with daily data spanning the period 1980–2004 were available for western Canada.

2.2. Climate indices

2.2.1. Regional moisture deficit

Drought is a particular hazard in western Canada (e.g. Sauchyn *et al.*, 2005; Sauchyn and Kulshreshtha, 2008; Wheaton *et al.*, 2008; Sauchyn and Bonsal, 2013), so it was important to consider an index which combines

Table 1. Regional climate model details.

Model	Modelling group	Model identifier	Grid size
CRCM v4.2.0	OURANOS/UQAM (Caya and LaPrise, 1999; De Elía and Côté, 2010)	CRCM	115 × 140 (50 km)
CanRCM4	Environment Canada/UQAM (Scinocca *et al.*, 2016)	NAM-44 NAM-22	NAM-44: 130 × 155 (50 km) NAM-22: 260 × 310 (25 km)
ECPC/ECP2	UC San Diego/Scripps (Juang *et al.*, 1997)	ECP2	116 × 147 (50 km)
HRM3	UK Hadley Centre for Climate Prediction and Research (Jones *et al.*, 2004)	HRM3	130 × 155 (50 km)
MM5I	Iowa State University (Grell *et al.*, 1993)	MM5I	99 × 124 (50 km)
RCM3	UC Santa Cruz (Giorgi *et al.*, 1993a, 1993b; Pal *et al.*, 2000, 2007)	RCM3	104 × 134 (50 km)
WRFP/WRFG	Pacific Northwest National Laboratory (Skamarock *et al.*, 2005)	WRFG	109 × 134 (50 km)

Table 2. List of extremes indices calculated using observed and RCM daily temperature and precipitation data.

Index	Definition
pq90	The 90th percentile of rain day amounts
pxcdd	Maximum number of consecutive dry days (largest number of consecutive days where daily precipitation ≤1 mm)
pxcwd	Maximum number of consecutive wet days (largest number of consecutive days where daily precipitation > 1 mm)
ppww	Mean wet-day persistence (total number of consecutive wet days/total number of wet days, for the specified period)
ppdd	Mean dry-day persistence (total number of consecutive dry days/total number of dry days, for the specified period)
px3d	Greatest 3-day rainfall total

the effects of both temperature and precipitation into a meaningful measure of drought. Although there are a number of drought indices (e.g. Heim, 2002; Keyantash and Dracup, 2002), a simple moisture deficit index was used here which had minimal data requirements.

Moisture deficit, in this case defined as a measure of effective precipitation in excess of water loss by evapotranspiration (P-PET), was calculated for observed and RCM data sets. Although there are a variety of methods available to calculate potential evapotranspiration (PET; e.g. Priestley and Taylor, 1972; Jensen *et al.*, 1990; Thornthwaite, 1948; Shaw, 1994), Hogg's (1994, 1997) climate moisture index (CMI) was selected for its relative simplicity and basic climate data requirements. PET, in this case, was calculated using the simplified Penman–Monteith approach (Hogg, 1997). Hogg (1997) gave meaning to this index by illustrating that forest distribution in western Canada appears to be controlled by moisture deficit. According to Hogg (1997), a CMI value of zero (i.e. P = PET) defines the southern boundary of the boreal forest and a value of -15 mm corresponds to the aspen parkland–grassland boundary (based on 1951–1980 climate data). Moisture deficit values were calculated for each month and then accrued over the 3-month period May, June and July, to provide a measure of effective precipitation during the growing season (henceforth referred to as summer CMI).

2.2.2. *Site-specific extremes indices*

A list of standard indices related to extremes in daily temperature and precipitation was put together by Frich *et al.* (2002) and later extended by the CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices, an international effort coordinated by WMO/WCRP. These indices are used to monitor extremes in the present climate (e.g. Vincent and Mekis, 2006) and also as a result of climate change (e.g. Tebaldi *et al.*, 2006). Many of these indices were used in The Statistical and Regional dynamical Downscaling of Extremes for European Regions (STARDEX, 2005) project and MATLAB® scripts were written to calculate the indices of interest, based on the original FORTRAN routines available through STARDEX. Table 2 indicates which extremes indices were calculated from the daily observed and RCM data, the focus being on precipitation extremes.

2.3. Methods for comparing RCM and observed data

Since part of the focus of this RCM performance assessment considers their ability to simulate the frequency and intensity of extremes, this raises some unresolved issues regarding grid box data. There is some debate about comparing data from single climate model grid boxes with observed station data and about exactly how the grid box data should be interpreted (Skelly and Henderson-Sellers, 1996). Does the data value represent the centre of the grid box (e.g. Gutowski *et al.*, 2007) or is it an average value for the whole grid box (e.g. Osborn and Hulme, 1997)? Also, the climate modelling community recommends that results are averaged over multiple grid boxes since RCMs (and GCMs for that matter) cannot be expected, for numerical modelling reasons, to be skillful at their grid point scale (von Storch *et al.*, 1993; Frei *et al.*, 2003). Spatial averaging over several grid boxes smoothes out the errors at the grid point scale and potentially leads to better estimates (Herrera *et al.*, 2010; Dulière *et al.*, 2011). Teutschbein and Seibert (2010), however, found that the value of one grid cell did not differ considerably from the average over nine grid cells in their hydrological modelling study of five catchments located in a number of different climate zones across Sweden. In this case, our focus is on how well the RCMs are able to simulate extremes and averaging over several grid boxes may mask their ability to simulate extreme values. Construction of gridded observed data sets may also lead to the underestimation of extremes as a result of averaging station data to obtain gridded values (Haylock *et al.*, 2008; Hofstra *et al.*, 2010; Yin *et al.*, 2014). Being aware of these issues, it was decided, in this case, to use station data for the calculation of the observed extremes indices, rather than individual grid boxes from the daily ANUSPLIN gridded observed data set. Data from the individual RCM grid boxes closest to the station location were extracted and the extremes indices listed in Table 2 calculated. Although the RCM extreme values are likely to be underestimated, comparison with observed station data rather than gridded observed data allows us to determine the magnitude of the error and also the ability of the RCMs to simulate extremes on the scale of those observed.

The following comparisons were made:

1. Mean temperature and precipitation were averaged over western Canada (49°–60°N, 95°–120°W) and the 1980–2004 average annual cycles compared for these variables.
2. For each year, the spatial patterns of gridded seasonal fields were compared using the standard Pearson's correlation coefficient, which represents the degree of agreement between the RCM and NCEP2 climate patterns and those observed. For NCEP2 and each RCM, the original ANUSPLIN gridded observed data set was scaled up to the same resolutions by averaging the ANUSPLIN grid boxes falling within each NCEP2 or RCM grid box. As the correlation coefficient is insensitive to biases, it focuses on the ability of RCMs to simulate spatial details and contrasts (Walsh and McGregor, 1997; Herrera *et al.*, 2010). A measure of the bias between RCM, NCEP2 and observed fields was determined by calculating the mean absolute error (MAE). We used MAE, rather than the root mean square error (RMSE), a measure commonly used to describe model performance (see, e.g. Mearns *et al.*, 2012), because MAE is an unambiguous measure of average error (Willmott and Matsuura, 2005). RMSE, on the other hand, becomes increasingly larger than MAE as the distribution of error magnitudes becomes more variable, meaning that its interpretation is not clear.
3. In order to look beyond seasonal or longer averages, which can mask biases or systematic errors that are identifiable in daily data, probability density functions (pdfs) were used to permit comparisons of the entire data distribution (Perkins *et al.*, 2007). Observed and NCEP2 and RCM-derived precipitation were compared using a simple measure of similarity between two pdfs, known as the skill score (Perkins *et al.*, 2007). All observed daily station data with complete, or almost complete, records for the period 1980–2004 were used to construct the observed pdf for western Canada. In a similar manner, daily data for each grid box within the region were used to construct the RCM pdf for each RCM. Bin sizes of 1 mm day^{-1} for precipitation were used and all daily values of precipitation below 1 mm day^{-1} were omitted, since precipitation rates less than this do not contribute substantially to total daily precipitation over most regions (Dai, 2001; Sun *et al.*, 2006). The skill score, as described by Perkins *et al.* (2007), measures the common area between two pdfs by calculating the minimum value of the two distributions in each bin and then summing across the entire distribution. Values range between zero (poor performance – negligible overlap between pdfs) and one (pdfs match).
4. Daily observed station data and NCEP2 and RCM data were compared using quantile–quantile (q–q) plots at the ten study sites illustrated in Figure 1. This method compares the pdfs of two sets of data and is often used to compare empirical data with that of a known, or fitted, distribution (e.g. Gaussian or Gamma; Wilks, 2011). In this case, the plotted quantiles represent the distribution of the RCMs and NCEP2 *versus* observed data. If the points on the q–q plot closely follow the 1 : 1 line (i.e. $y = x$, and here the line upon which the observed data are plotted) then the data sets are considered to be from the same probability distribution. For precipitation, only days on which precipitation occurred were considered, with the threshold value defining a wet day being 1 mm.
5. The extremes indices listed in Table 2 were calculated at the ten study sites and compared.

3. Results

3.1. Annual cycles

All RCMs and NCEP2 capture the shape of the mean temperature and precipitation annual cycles for western Canada (Figure 2). Almost all RCMs simulate conditions that are too warm in the winter half year but in summer approximately half the RCMs are cooler than, or more closely match, observed values. HRM3 exhibits the largest warm bias, with mean temperature almost 10 °C greater than observed over western Canada in winter. CRCM has a cold bias of about 2 °C throughout most of the year in this region. NCEP2 closely matches observed conditions for the latter half of the year, but is about 1 °C cooler than observed in spring and early summer. Although all RCMs and NCEP2 capture the shape of the annual precipitation cycle in this region, the majority of RCMs are too wet throughout the cycle, particularly in winter and peak precipitation tends to occur a month earlier than observed in almost all RCMs. From October to May all RCMs simulate too much precipitation, about 20 mm (per month) greater

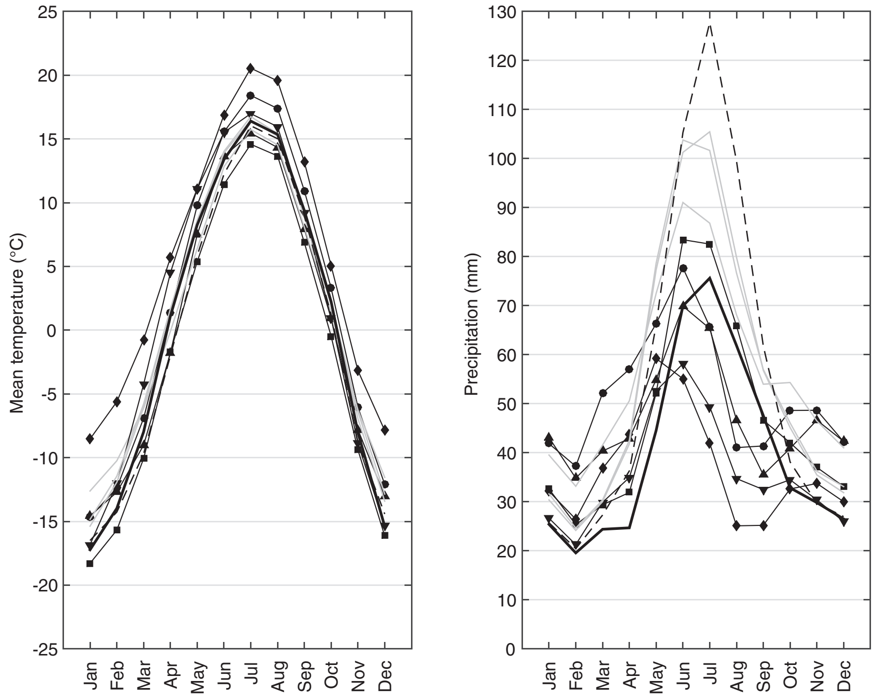

Figure 2. Average annual cycle of mean temperature (°C) and total precipitation (mm) for western Canada for 1980–2004. ANUSPLIN observed, bold black line; NCEP2, black dashed line. RCMs discussed in text are identified by the following symbols: CRCM – ■; ECP2 – ●; HRM3 – ◆; MM5I – ▲; WRFG – ▼. Other RCMs are indicated by the grey lines.

than observed values. NCEP2 closely matches observed average precipitation totals in winter, but greatly overestimates totals in summer, in particular, with values being almost double those observed. All RCMs exhibit improvements in the simulation of summer precipitation when compared with NCEP2. As RCM and NCEP2 orography is subdued and smoother than in reality, the rain-shadow effect of the Rockies is not as pronounced, thus contributing to the overestimation of precipitation in western Canada. From July to September half the RCMs (HRM3, WRFG, MM5I and ECP2) simulate too little precipitation in this region.

3.2. Comparison of spatial fields

3.2.1. Spatial pattern – Pearson correlation coefficients

Pearson correlation coefficients for mean temperature, precipitation and summer CMI are shown in Figures 3–5, by year and RCM, for western Canada. These figures allow the identification of those RCMs which may have particular problems in simulating the spatial patterns of the climate variables under study, or of particular years when all RCMs perform well or have difficulty simulating observed conditions. In general, RCMs are more successful at capturing the spatial pattern of mean temperature (Figures 3 and S1, Supporting information), compared to precipitation (Figures 4 and S2) and CMI (Figure 5), with coefficients for this variable mostly in excess of 0.9. Correlations are generally slightly lower for spring mean temperature (Figure S1) than in the other seasons, with CRCM, HRM3 and WRFG indicating lower values overall, while in summer ECP2, HRM3 and WRFG are not as successful at simulating the spatial patterns of mean temperature as the other RCMs. All RCMs are less correlated than NCEP2 with the observed climatology in spring and summer, and while values are slightly lower than NCEP2 in fall (Figure S1) and winter, most correlations are still greater than 0.9.

For precipitation over western Canada (Figures 4 and S2), the RCMs are most successful at simulating the spatial pattern of winter precipitation, with correlations generally in excess of 0.7 in all years. In summer, however, coefficients are generally less than 0.5 in all years, with some RCMs exhibiting values close to, or below, zero, in many years. In spring and fall (Figure S2), correlations are mixed, but all RCMs are generally less successful at simulating precipitation patterns in these seasons compared to winter, although more successful than in summer. In all seasons except summer, all RCMs are more closely correlated with observed conditions than is NCEP2, implying that their higher resolution has resulted in improvements to the simulation of precipitation in these seasons. For summer CMI (Figure 5), RCM3 results are least well correlated with observed conditions (1980–2004 mean correlation is 0.4), while CRCM and the two versions of CanRCM4 are more successful, with period mean correlations closer to 0.6. All RCMs, with the exception of MM5I and RCM3, are more successful than NCEP2 in simulating observed summer CMI values.

For precipitation, lower summer correlation coefficients are likely due to problems modelling convective precipitation, which is the primary source of precipitation in this region at this time of year, either through inadequate model physics or by the spatial domain not being large enough to

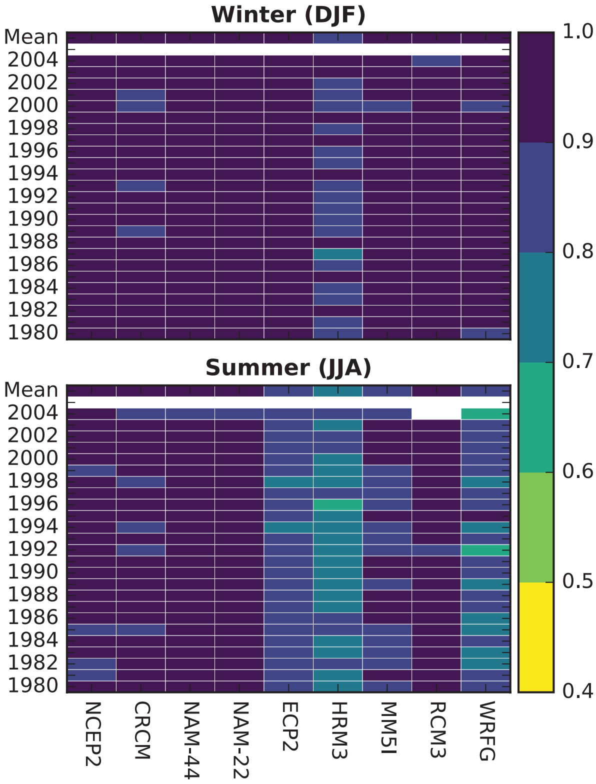

Figure 3. Spatial pattern correlations by RCM and year for mean temperature over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. (Value for RCM3 2004 is missing.) [Colour figure can be viewed at wileyonlinelibrary.com].

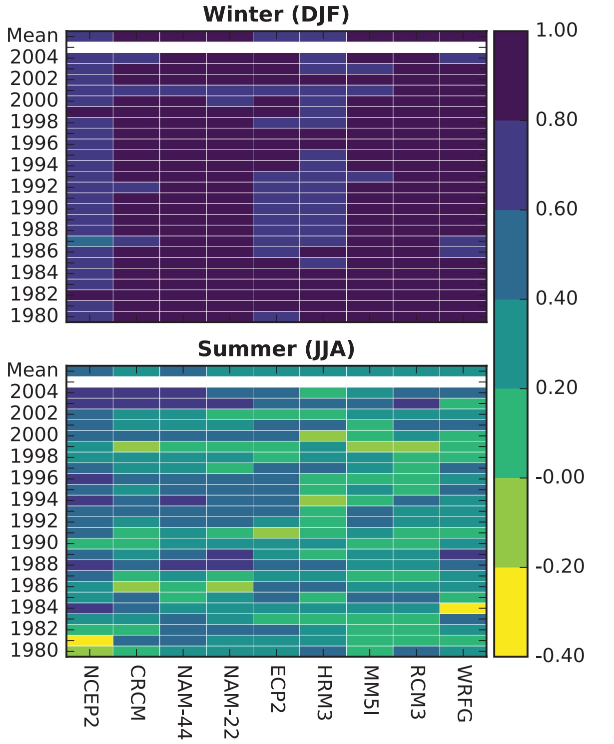

Figure 4. Spatial pattern correlations by RCM and year for total precipitation over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. [Colour figure can be viewed at wileyonlinelibrary.com].

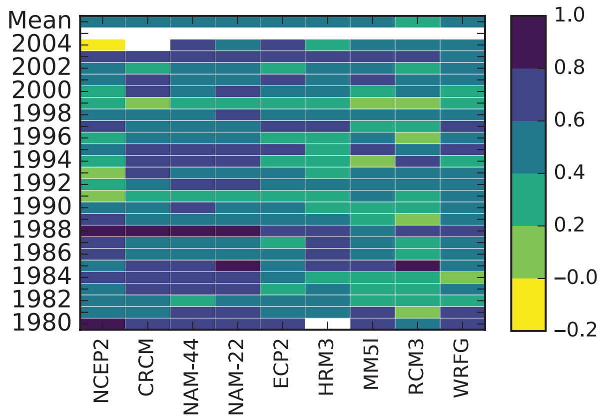

Figure 5. Spatial pattern correlations by RCM and year for summer CMI for western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of the plot. (Values for CRCM 2004 and HRM3 1980 are missing.) [Colour figure can be viewed at wileyonlinelibrary.com].

include transport of moisture from further afield (e.g. the Gulf of Mexico). At the spatial resolutions of the RCMs considered here (25–50 km), convective storms are not directly modelled, and mesoscale systems producing much of the warm-season precipitation in this region are poorly resolved (Mearns *et al.*, 2012). Convection is still modelled using sub-grid scale parameterisations, which leads to noisy patterns as the individual model grid boxes are, in effect, independent of one another. Also, gridded observed summer precipitation data sets are unrealistically smooth because of the interpolation of noisy station data and these two factors together lead to the poor correlations observed. For mean temperature, the lower correlation coefficients apparent in spring will be due to the temperature bias of the RCMs which results in either the early loss or late melt of snow cover.

3.2.2. Mean absolute error

Figures 6–8 illustrate the MAE between observed, NCEP2 and RCM-simulated values by year, model and season (winter and summer) for western Canada for mean temperature, precipitation and summer CMI, respectively. For mean temperature (Figures 6 and S3), errors are generally smaller in the summer and fall (Figure S3). HRM3 consistently exhibits the largest errors, particularly in winter (Figure 6), when the 1980–2004 mean value is between 8 and 9 °C warmer than observed. With the exception of HRM3, MAE is generally within 2 and 3 °C of NCEP2 values for all RCMs. For total precipitation (Figures 7 and S4), errors are generally similar to NCEP2 (within 0.5 mm day^{-1}) in winter (Figure 7), spring and fall (Figure S4), while in summer (Figure 7) all RCMs exhibit smaller MAE values than NCEP2. For summer CMI (Figure 8), results are mixed, but CRCM and ECP2 exhibit the smallest MAE values, unlike RCM3 and NCEP2 which indicate generally larger MAE values than the other RCMs.

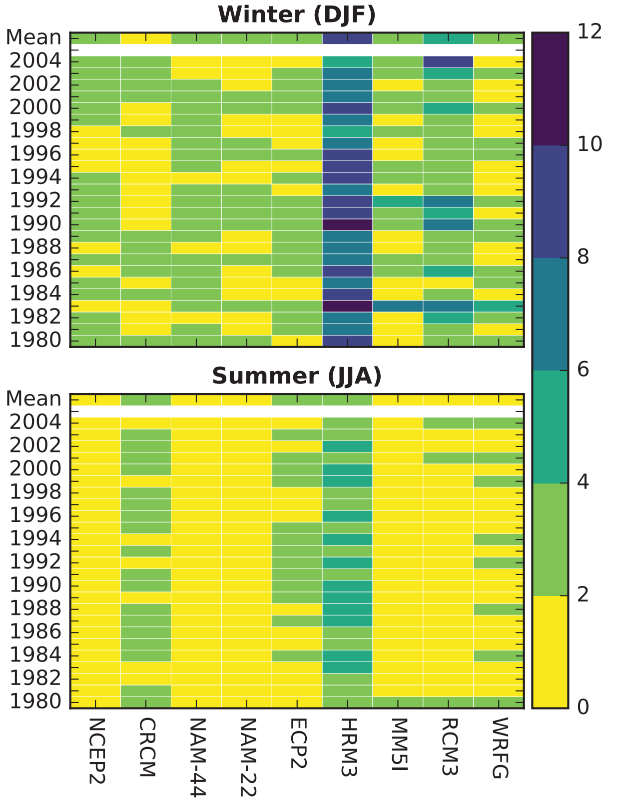

Figure 6. Winter and summer mean absolute error by RCM and year for mean temperature (°C) over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. [Colour figure can be viewed at wileyonlinelibrary.com].

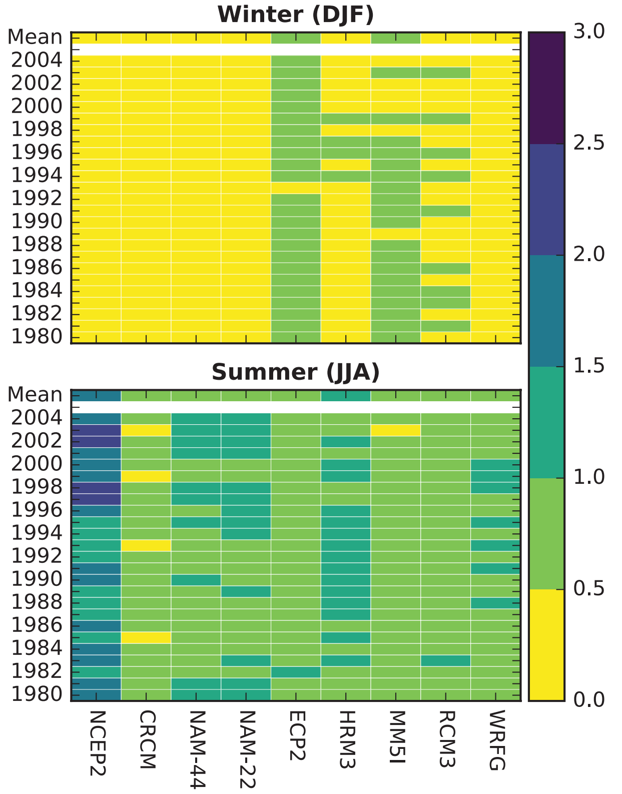

Figure 7. Winter and summer mean absolute error by RCM and year for total precipitation (mm day^{-1}) over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. [Colour figure can be viewed at wileyonlinelibrary.com].

3.2.3. *Spatial patterns associated with high and low correlations*

Since all RCMs capture the general spatial pattern of mean temperature and there is little difference between the correlation coefficients in higher and lower scoring years, and since the emphasis of this article is on precipitation simulation, no spatial patterns are illustrated for mean temperature.

For precipitation, however, the RCMs are not so successful at simulating its spatial distribution and there are more pronounced differences between seasons. Figures S5 and 9 illustrate the spatial distribution of precipitation for winter 1999 (higher correlation) and summer 1999 (lower correlation), respectively. For winter 1999 (Figure S5), where correlations are generally in excess of 0.8 (HRM3 is the exception at 0.77), all RCMs exhibit similar patterns and all overestimate the amount of precipitation in the mountains on the British Columbia – south-western Alberta border. This overestimation may be partly due to the gridded ANUSPLIN data set which exhibits its largest errors in mountainous regions (McKenney *et al.*, 2006). MAE values range between 0.2 mm day^{-1} for NCEP2 and 0.78 mm day^{-1} for ECP2. For summer 1999 (Figure 9), on the other hand, spatial correlation coefficients are less than 0.5 and negative in some instances. MAE values range between 0.46 mm day^{-1} (CRCM) and 1.56 mm day^{-1} (NCEP2). Convective precipitation contributes more to

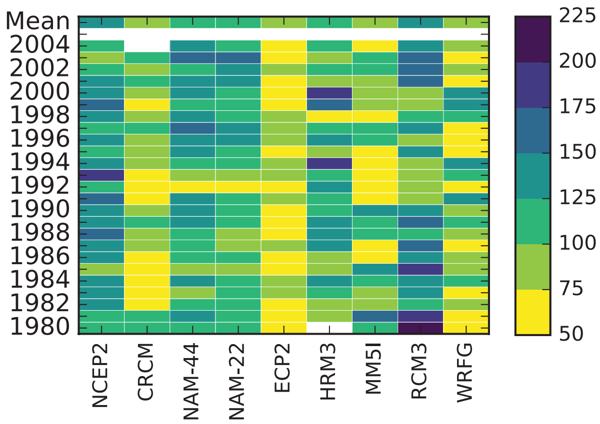

Figure 8. Mean absolute error by RCM and year for summer CMI for western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of the plot. (Values for CRCM 2004 and HRM3 1980 are missing.) [Colour figure can be viewed at wileyonlinelibrary.com].

precipitation totals during summer months, and lower correlation coefficients generally indicate the RCMs' difficulty in simulating this type of precipitation and the mesoscale systems producing precipitation in this season.

For summer CMI, two particular years stand out in Figure 5 as being consistently well-correlated (1988) and poorly correlated (1999) with observed conditions in western Canada. These years are shown in Figures S6

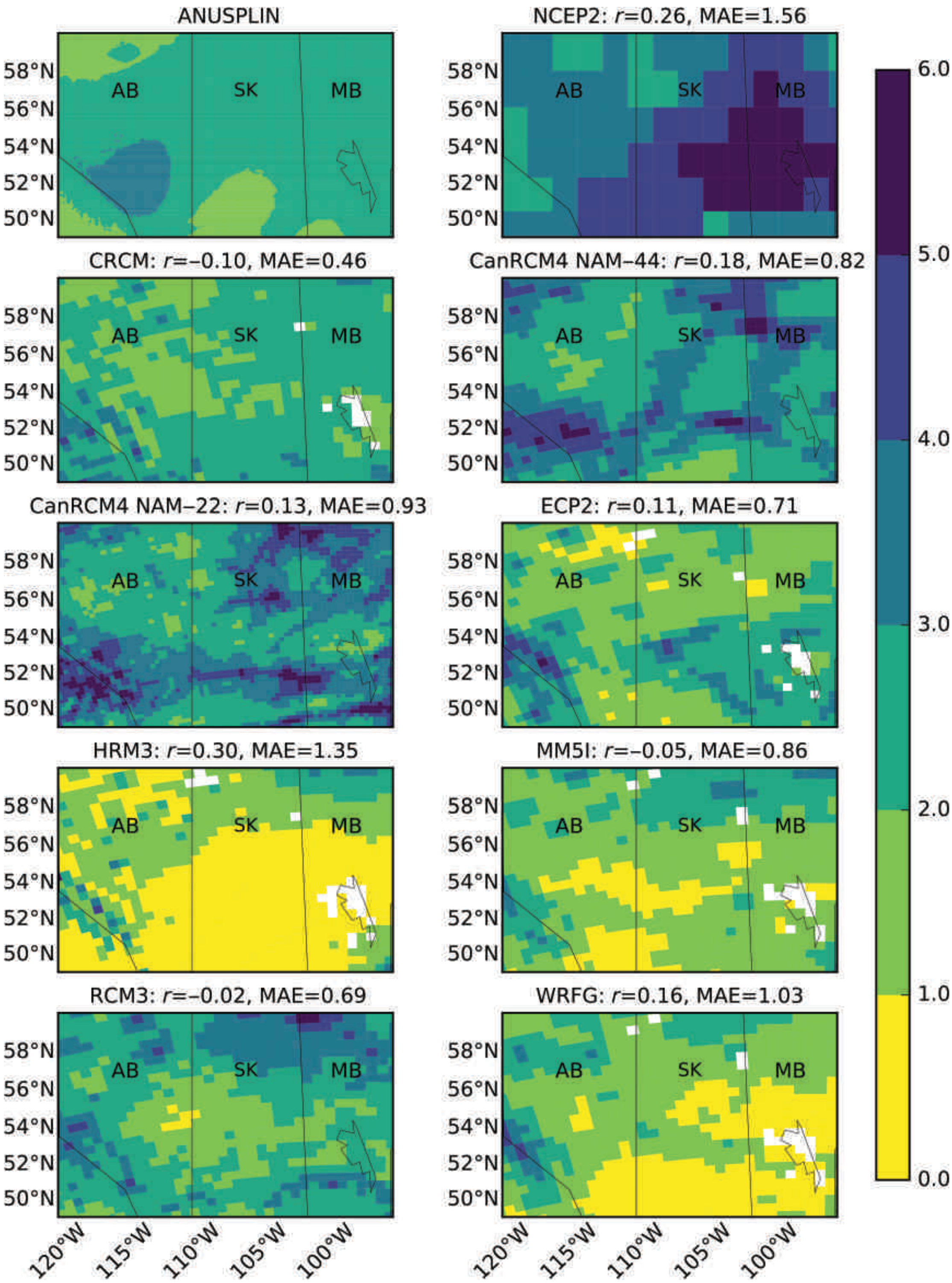

Figure 9. Total precipitation (mm day^{-1}) for summer (JJA) 1999: observed (ANUSPLIN), NCEP2, and as simulated by the eight RCMs. Correlation coefficients (r) and mean absolute errors (MAE; mm day^{-1}) are given for western Canada. [Colour figure can be viewed at wileyonlinelibrary.com].

and 10, respectively. Across the Canadian prairies, 1988 was a drought year and all RCMs and NCEP2 simulate the pattern of moisture deficit across this region with some success (Figure S6). NCEP2, CRCM and the two resolution versions of CanRCM4 (NAM-44 and NAM-22) are most successful at capturing the spatial CMI pattern. On the other hand, 1999 was the first year in a multi-year drought across the Canadian prairies (1999–2004; e.g. Chipanshi *et al.*, 2006; Bonsal *et al.*, 2013), but observed conditions were not extreme, with most moisture deficit values generally within the range ±100 mm. In this case, all RCMs struggled to simulate this pattern of moisture deficit (Figure 10), with correlations generally below 0.3. This implies that CMI is more successfully simulated

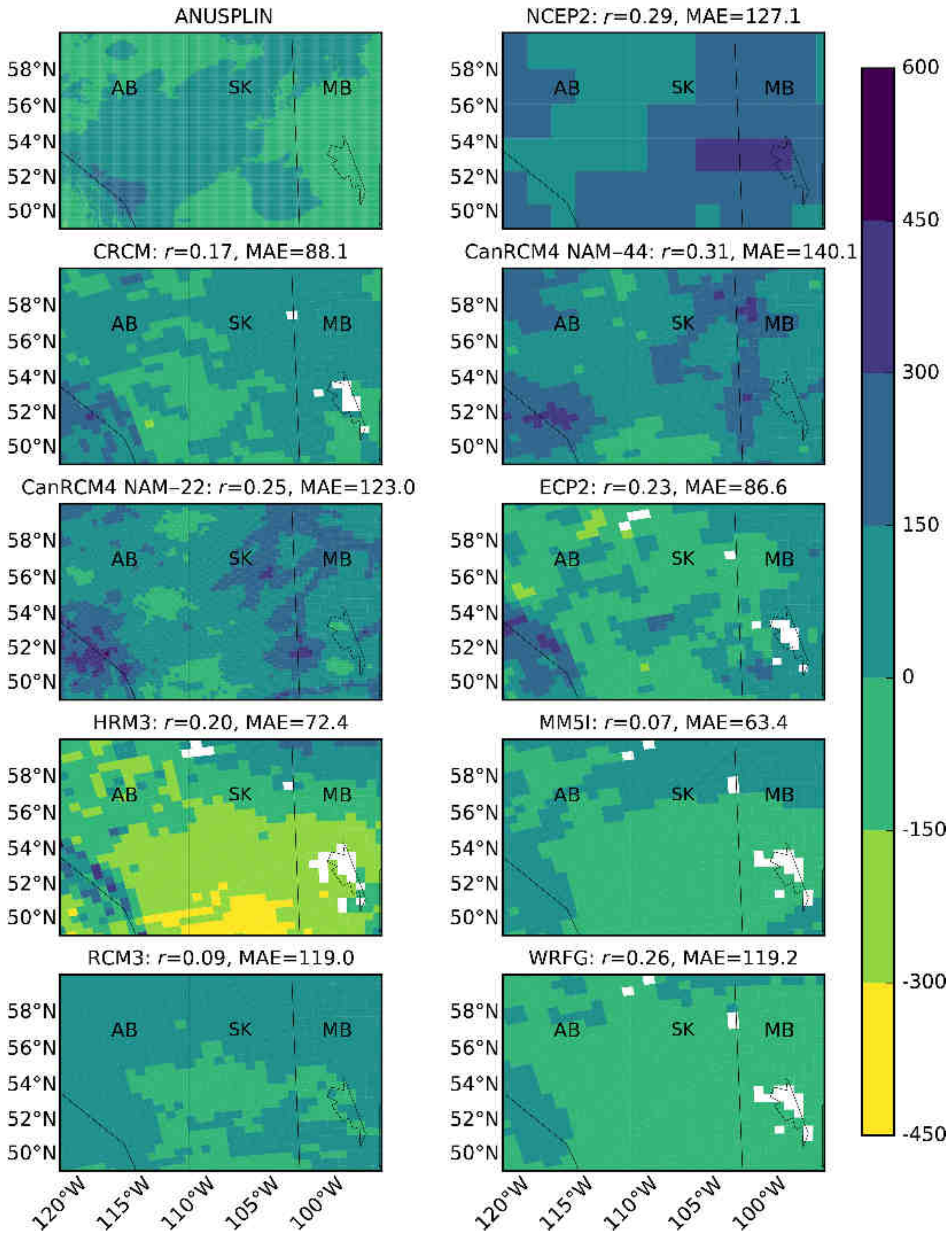

Figure 10. CMI (mm) for summer 1999: observed (ANUSPLIN), NCEP2, and as simulated by the eight RCMs. Correlation coefficients (r) and mean absolute errors (MAE; mm) are given for western Canada. [Colour figure can be viewed at wileyonlinelibrary.com].

when summer precipitation is low and less convective activity occurs.

3.3. Simulation of extremes

RCM performance in simulating extremes was evaluated at two spatial scales:

1. *At the regional scale*: skill scores (Perkins *et al.*, 2007) were calculated to determine the RCMs' ability to capture the observed probability density functions for daily mean temperature and precipitation on a seasonal basis across western Canada;
2. *At the grid box scale*: q–q plots were used to compare the observed and RCM daily data distributions for the extremes indices listed in Table 2 at the ten locations in western Canada (see Figure 1).

3.3.1. Skill scores

Tables 3 and 4 list the skill scores for daily mean temperature and precipitation by season, respectively. CRCM

Table 3. Skill scores for mean temperature over western Canada.

Model/Season	DJF	MAM	JJA	SON
NCEP2	0.85	0.68	0.83	0.86
CRCM	0.78	0.72	0.69	0.77
NAM-44	0.92	0.88	0.87	0.89
NAM-22	0.90	0.88	0.88	0.88
ECP2	0.90	0.85	0.91	0.91
HRM3	0.81	0.94	0.75	0.91
MM5I	0.90	0.73	0.80	0.84
RCM3	0.91	0.77	0.77	0.85
WRFG	0.86	0.95	0.87	0.85

DJF, winter; MAM, spring; JJA, summer; SON, fall.

Table 4. Skill scores for precipitation over western Canada.

Model/Season	DJF	MAM	JJA	SON
NCEP2	0.85	0.93	0.93	0.95
CRCM	0.86	0.83	0.87	0.84
NAM-44	0.94	0.95	0.88	0.94
NAM-22	0.94	0.97	0.90	0.96
ECP2	0.93	0.97	0.96	0.95
HRM3	0.98	0.96	0.82	0.95
MM5I	0.95	0.92	0.95	0.94
RCM3	0.91	0.91	0.95	0.90
WRFG	0.97	0.94	0.84	0.95

DJF, winter; MAM, spring; JJA, summer; SON, fall.

tends to perform less well than the other RCMs for temperature, but skill score values are generally still greater than 0.7. Most RCMs tend to indicate slightly less skill in simulating the observed mean temperature pdf in summer and this is also the case for some RCMs (NAM-44, NAM-22, HRM3 and WRFG) for precipitation. NCEP2, on the other hand, shows less skill in simulating the spring temperature distribution and winter and spring precipitation distributions. The skill score histograms (not shown) indicate that the RCMs capture the general shape of the pdfs but, for the lower skill scores, the pdfs tend to err in location and frequency when compared to observed.

3.3.2. *Quantile–quantile plots*

Quantile–quantile plots for daily mean temperature and total precipitation tend to be very similar across all ten sites, and so only a couple of examples are included here as indicative of the general results. Figure 11 illustrates seasonal q–q plots for mean temperature for Lethbridge, Alberta. For ease of interpretation, the observed daily data are represented by the $y=x$ line. In all seasons, almost all RCMs, with the exception of CRCM and RCM3 and also NCEP2, are warmer than observed. In winter, spring and fall, the RCM quantiles more closely match observed when mean temperature is positive. In summer, at mean temperatures greater than about 20 °C all RCMs and NCEP2 exhibit warmer temperatures than observed, indicating that extreme mean temperatures at this site are too high in this season.

Figure 12 shows the q–q plot for total precipitation at Swift Current, Saskatchewan. In winter, all RCMs (except NAM-22) underestimate the observed quantiles, while in the other seasons both under- and over-estimations are apparent, although no particular RCM consistently exhibits the same behaviour across all seasons. NCEP2, however, does underestimate the observed precipitation quantiles in all seasons at this site. While the RCMs and NCEP2 tend to follow the observed quantiles, discrepancies become larger at higher precipitation values.

3.3.3. *Extremes indices*

Table 2 lists the precipitation-based extremes indices calculated for observed, NCEP2 and RCM data at the ten study sites. In order to concisely present the results, mean square errors (MSE) between observed and model index values were calculated and results are given for Val Marie, Saskatchewan. This site exhibited the least error for three of the five indices considered here, while Carway, Alberta, consistently displayed the highest error values.

Although temperature-based extremes indices were calculated initially, the results are not shown because they are completely consistent with the warm and cold biases identified in the RCMs at the seasonal time scale. For example, extreme minimum and maximum temperatures are too warm, the growing season is too long and the number of growing degree days is overestimated in those RCMs with a warm bias (e.g. HRM3 and ECP2) and vice versa.

3.3.3.1. Ninetieth percentile rain day amounts: Figure 13 illustrates deviations from the observed value for 90th percentile rain day amounts. Total deviations are smaller in winter, when precipitation is low, and larger in spring and summer. Again, no single RCM or NCEP2 exhibits consistently positive or negative deviations. In summer, negative deviations occur more frequently and the positive deviations tend to be dominated by one or two RCMs, e.g. RCM3 and WRFG in 2003. Where this index value is underestimated (i.e. deviations are negative) it implies that the RCM is unable to simulate the larger precipitation amounts in the distribution. As negative deviations are generally larger and more frequent in summer, they are probably related to the poor simulation of convective precipitation in this season.

3.3.3.2. Maximum number of consecutive wet and dry days: Figure 14 shows deviations from the observed value for the maximum number of consecutive dry days. For this index, there is more consistency in the results: for example, in summer ECP2 generally overestimates the number of consecutive dry days, while NAM-44 and NAM-22 underestimate this index value. Figure 15, which gives the maximum number of consecutive wet days, indicates that most RCMs overestimate the number of consecutive wet days, particularly in spring and summer. Wet and dry day persistence emulates these results (not shown), with those RCMs which overestimate the maximum number of consecutive wet or dry days also overestimating wet or dry day persistence, respectively, and vice versa. So, while most

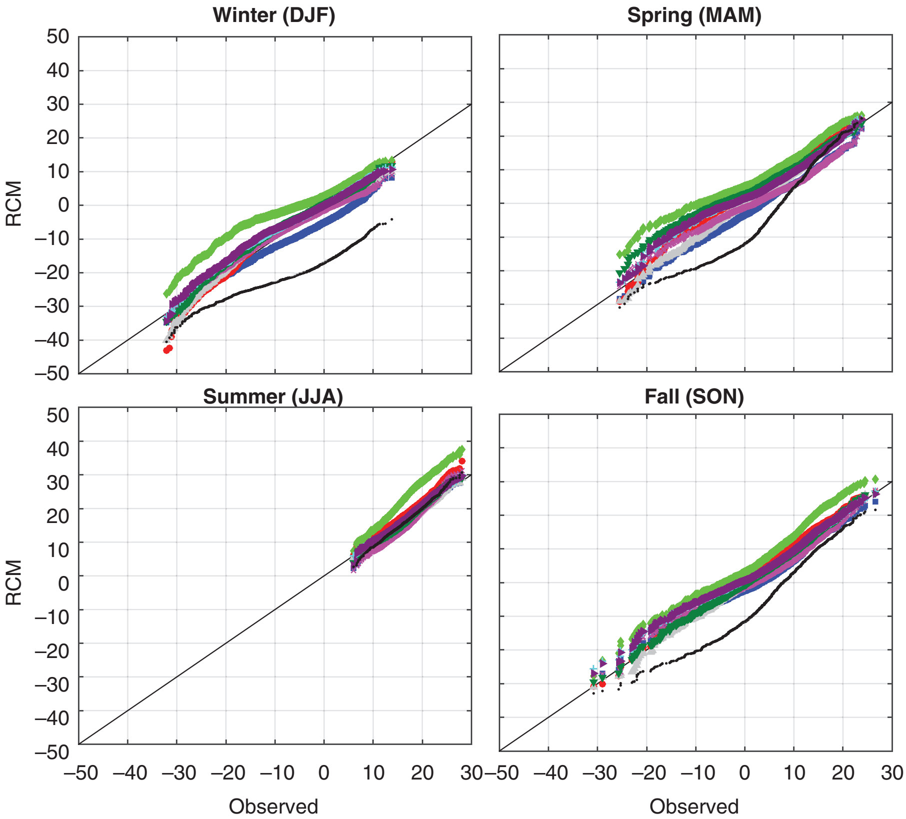

Figure 11. Quantile–quantile plots for daily mean temperature (°C) for Lethbridge, Alberta. The 1:1 line represents the observed daily data. RCMs are represented by the following symbols: CRCM – ■; NAM-22 – +; NAM-44 – ▶; ECP2 – ●; HRM3 – ◆; MM5I – ▲; RCM3 – x; WRFG – ▼; NCEP2 – •. [Colour figure can be viewed at wileyonlinelibrary.com].

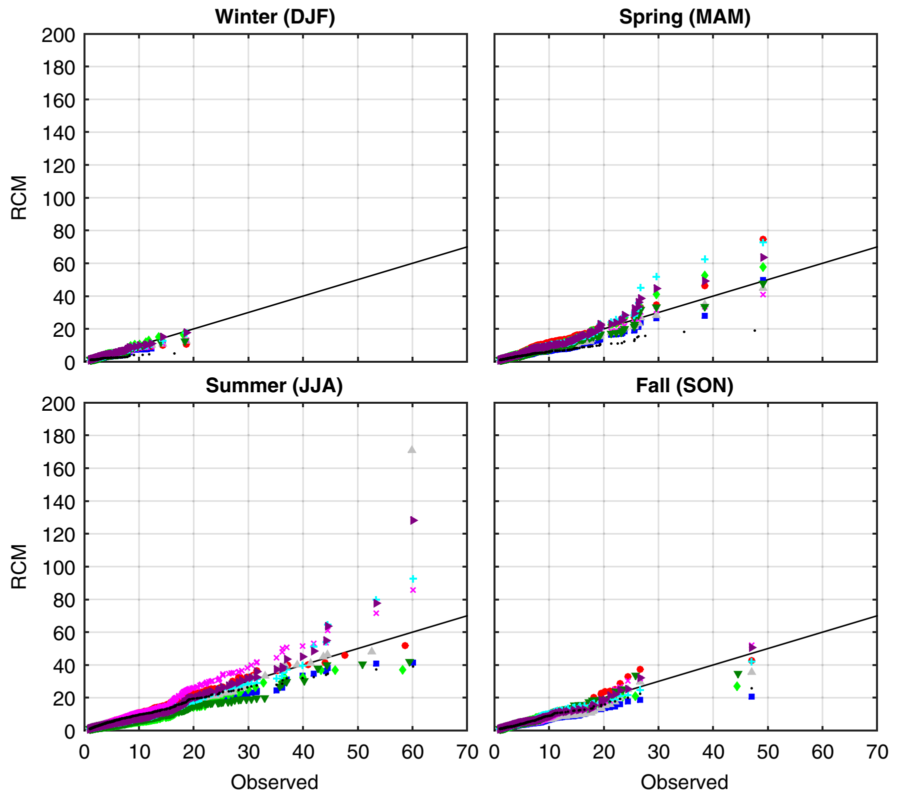

Figure 12. Quantile–quantile plots for daily precipitation (mm) for Swift Current, Saskatchewan. The 1:1 line represents the observed daily data. RCMs are represented by the following symbols: CRCM – ■; NAM-22 – +; NAM-44 – ▶; ECP2 – ●; HRM3 – ◆; MM5I – ▲; RCM3 – x; WRFG – ▼; NCEP2 – •. [Colour figure can be viewed at wileyonlinelibrary.com].

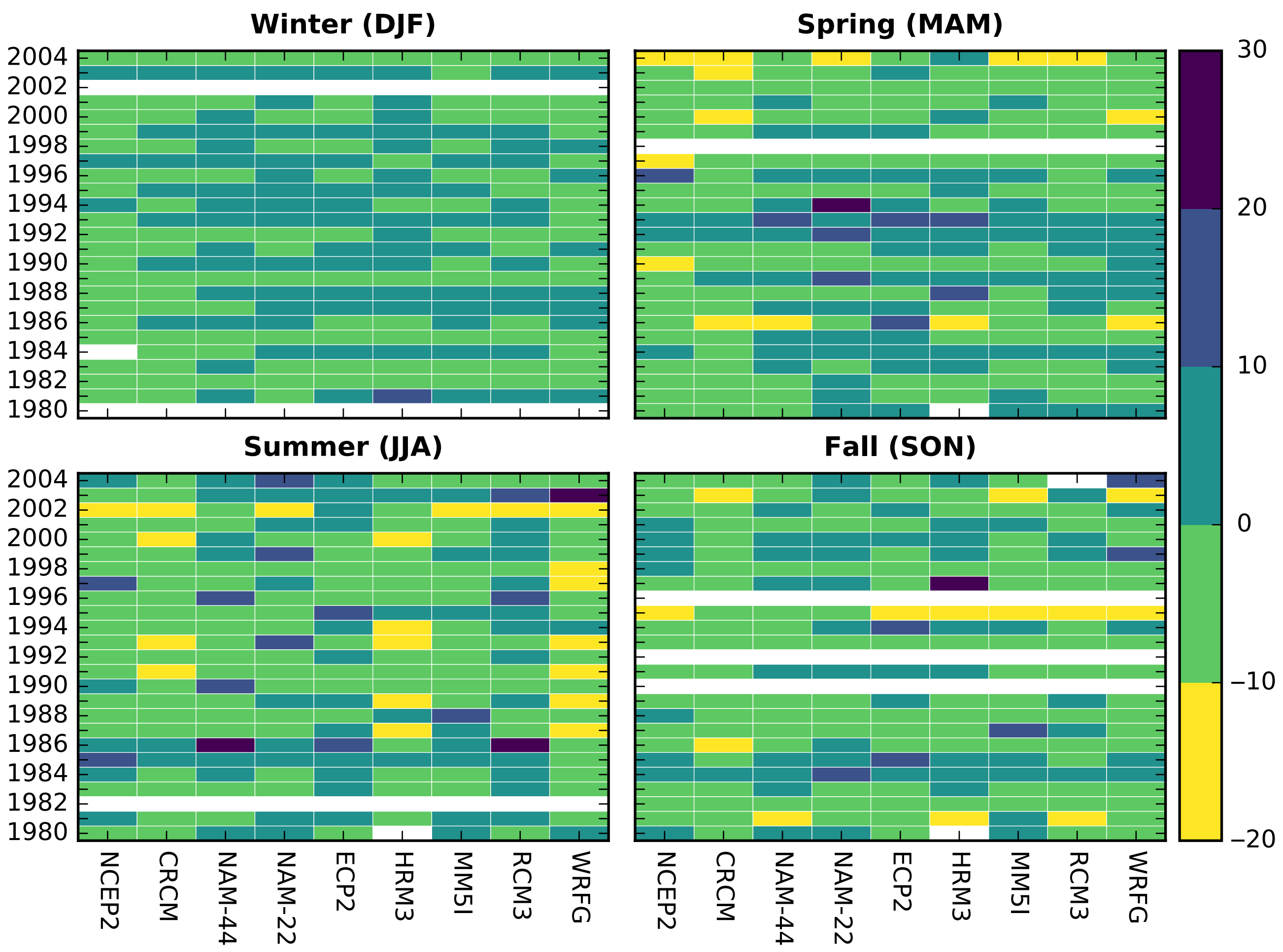

Figure 13. The 90th percentile rain day amounts (mm) for Val Marie, Saskatchewan by RCM and year for 1980–2004; deviations from the observed value. Values are missing for NCEP2 1984, all 1980 and 2002 (winter), HRM3 1980 and all 1998 (spring), HRM3 1980 and all 1982 (summer), HRM3 1980, RCM3 2004 and all 1990, 1992 and 1996 (fall). [Colour figure can be viewed at wileyonlinelibrary.com].

RCMs and NCEP2 overestimate the maximum number of consecutive wet days in spring and summer at this site, they are not able to simulate the larger precipitation values observed in these seasons (and indicated by the 90th percentile rain day amounts), thus implying that smaller precipitation amounts are occurring more frequently in the RCMs than is observed.

3.3.3.3. Greatest 3-day total precipitation: Figure 16 shows the deviations from observed values for 3-day precipitation totals. In winter deviations are smaller than in the other seasons. Largest deviations are apparent in spring and summer when much of the precipitation will be as a result of convective events. Those RCMs which overestimate this multi-day precipitation total are generally the same RCMs which overestimate the number of consecutive wet days and wet-day persistence.

4. Summary and conclusions

This article has described the analyses undertaken to assess the performance of eight RCMs (CRCM, NAM-44, NAM-22, ECP2, HRM3, MM5I, RCM3 and WRFG) in simulating the current climate of western Canada. The climate of the driving reanalysis, NCEP2, was also compared in order to determine the value added by the higher resolution climate models. The analyses were selected to gain a comprehensive picture of the RCMs' ability to simulate various aspects of current climate. As well as examining RCM performance at capturing the spatial features of current climate at the seasonal scale, their ability to simulate daily extremes of temperature and precipitation was also investigated. Results are summarized as follows:

1. RCM data generally exceed the observed mean monthly temperatures over western Canada, with the warm bias being greatest in January and December. In the summer months, at least two RCMs exhibit a cold bias and CRCM is cooler than observed throughout the year. NCEP2, provider of boundary conditions to the RCMs, is generally slightly cooler than observed, particularly in spring.
2. All RCMs capture the general shape of the annual precipitation cycle although peak values generally occur about 1 month earlier than observed. Total precipitation is overestimated in winter and early spring, but from June to October the RCMs are split

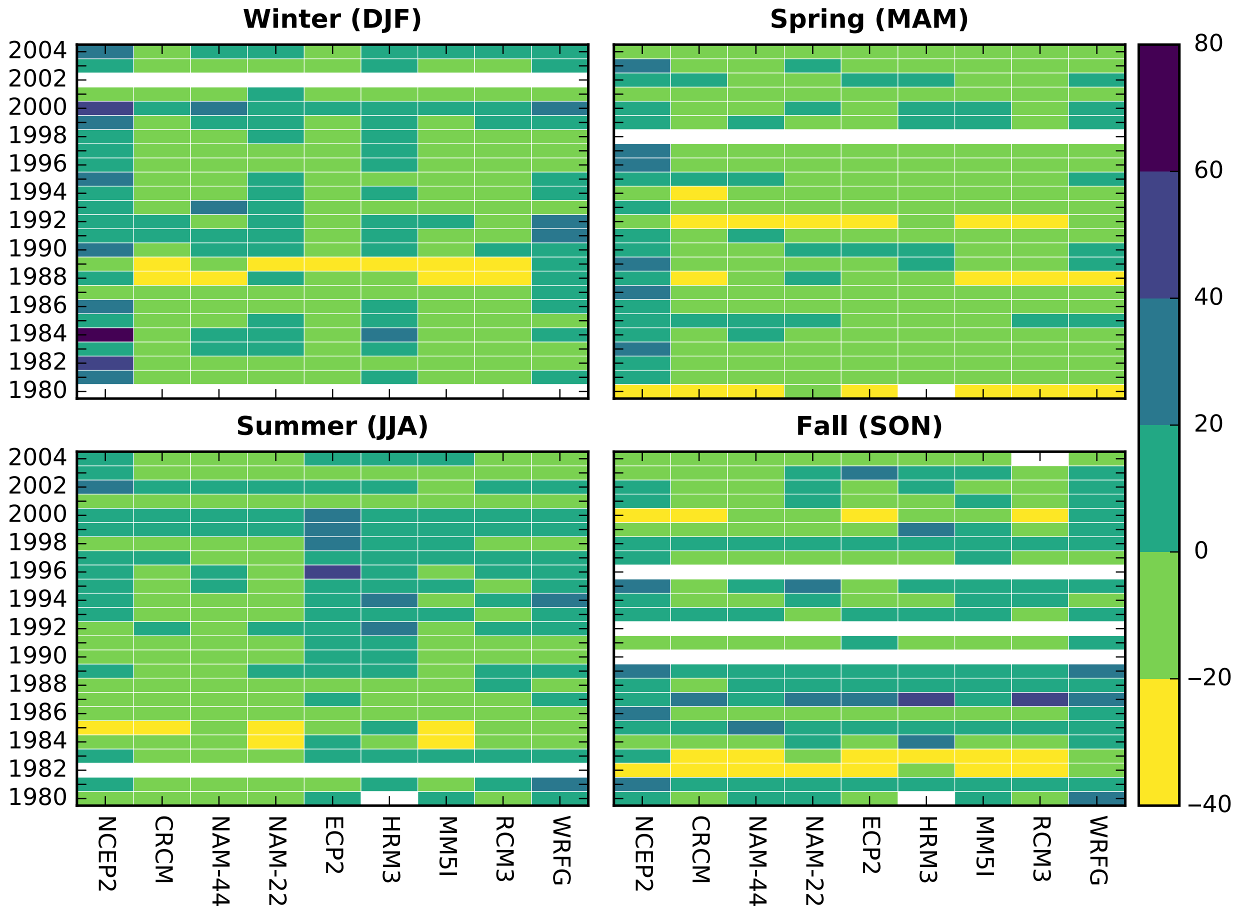

Figure 14. Maximum number of consecutive dry days for Val Marie, Saskatchewan by RCM and year for 1980–2004; deviations from the observed value. Values are missing for all 1980 and 2002 (winter), HRM3 1980 and all 1998 (spring), HRM3 1980 and all 1982 (summer), HRM3 1980, RCM3 2004 and all 1990, 1992 and 1996 (fall). [Colour figure can be viewed at wileyonlinelibrary.com].

with half simulating too much precipitation and half too little. NCEP2 precipitation totals most closely match observed values in winter, but are greatly overestimated in summer. RCM values more closely match observed in summer than do the NCEP2 data.

3. Pattern correlations indicate that all eight RCMs capture the spatial features of observed mean temperature well in all seasons, although coefficients are slightly lower in summer.
4. For total precipitation, spatial correlations are highest in winter and lowest in summer. Summer correlations range from approximately −0.3 to +0.7 and although HRM3 and MM5I exhibit the lowest 1980–2004 mean values, all RCM results include years with low correlations. Although the driving reanalysis fields are providing relatively accurate large-scale information to the RCMs, the models' ability to simulate the convective precipitation processes is still problematic at 50 km resolution.
5. MAE values for mean temperature for western Canada generally lie between 1.0 and 4.0 °C. The exception to this is HRM3 which has the largest errors in all seasons: winter temperatures are more than 8 °C warmer than observed. Errors in winter for this RCM are almost double those in the other seasons. Mearns *et al.* (2012) reported that HRM3's poor performance, in terms of seasonal temperature bias (for North America as a whole, not only Canada), does not occur elsewhere (e.g. China, South America) or to the same extent over North America when driven with the European Centre for Medium-Range Weather Forecasting Reanalysis (ERA-15). This implies that the two reanalyses are different and Mearns *et al.* (2012) suggested that the NCEP2 reanalysis data feeding into the RCM boundaries are indeed warmer and moister than ERA-15. As well as the direct increase in temperature in the boundary conditions, surface feedbacks within the RCM, such as reduced snow cover leading to enhanced absorption of solar radiation at the surface and thus further increases in surface temperature, may also be affected.
6. For precipitation, MAE values are larger in summer and generally smaller in winter, when precipitation amounts are generally low in this region. CRCM exhibits the lowest error in all seasons while NCEP2 exhibits the largest MAE values in summer.
7. Spatial correlation coefficients for summer CMI, the chosen moisture deficit index which combines the

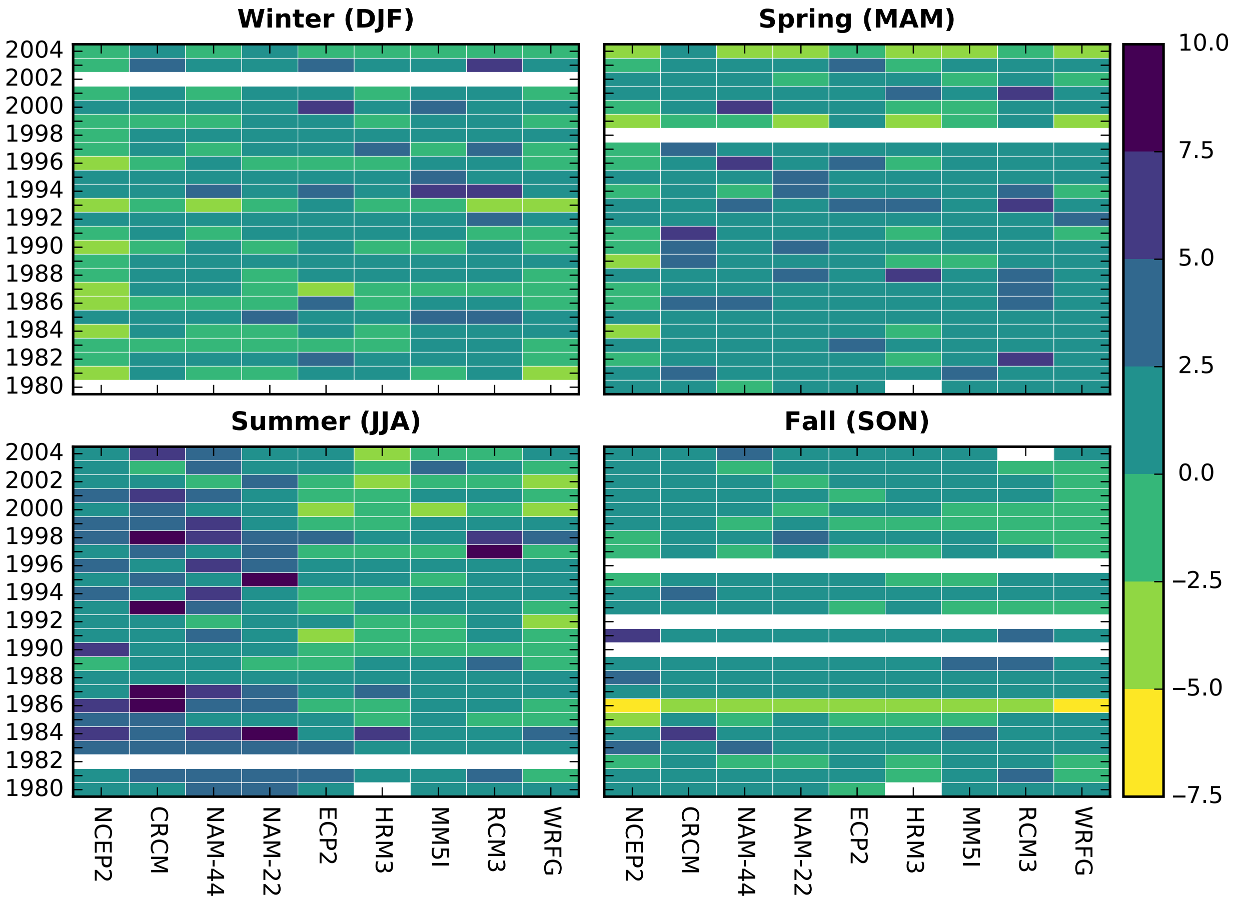

Figure 15. Maximum number of consecutive wet days for Val Marie, Saskatchewan by RCM and year for 1980–2004; deviations from the observed value. Values are missing for all 1980 and 2002 (winter), HRM3 1980 and all 1998 (spring), HRM3 1980 and all 1982 (summer), HRM3 1980, RCM3 2004 and all 1990, 1992 and 1996 (fall). [Colour figure can be viewed at wileyonlinelibrary.com].

effects of both temperature and precipitation, generally average about 0.5, although CRCM, NAM-44 and NAM-22 values are closer to 0.6. The RCMs are able to simulate the spatial features of this index with less success than for mean temperature, thus implying that errors in the precipitation fields have more influence on the spatial distribution of CMI than does mean temperature.

8. Skill scores were used at the regional scale to examine the ability of the RCMs to simulate the probability density functions of daily temperature and precipitation by season. CRCM generally scored slightly lower than the other RCMs for all variables under consideration in all seasons, but values still tended to be greater than 0.7. This indicated that all RCMs generally capture the shape of the density functions well, but where skill score values were lower there were errors in the location and frequency of the distributions when compared with observed.
9. Examination of q–q plots for the ten selected sites indicated that all RCMs simulated similar temperature distributions to those observed, but either with a cold (CRCM, RCM3 and NCEP2) or a warm bias.
10. For precipitation, the q–q plots indicate that the RCMs are not as successful at simulating distributions similar to those observed. At lower daily precipitation values the simulated distributions generally follow that of the observed distribution, but at higher values they tend to diverge.
11. Errors in the temperature-based extremes indices tend to be as a result of the warm or cold bias associated with a particular RCM. For example, HRM3 consistently simulates conditions that are warmer than observed, with fewer frost days, a larger number of growing degree days and a longer growing season. CRCM, on the other hand, simulates mean temperature conditions that are generally cooler than observed in the southern prairies, with a larger number of frost days, fewer growing degree days and a shorter growing season length.
12. Precipitation-based extremes indices present more complicated results. How well the RCMs perform in simulating these types of indices is very RCM- and site-dependent. In general, however, all RCMs tend to give results that more closely match observed index values in winter and fall. RCMs are able to simulate large precipitation amounts (e.g. multi-day totals) but most RCMs underestimate dry-day persistence

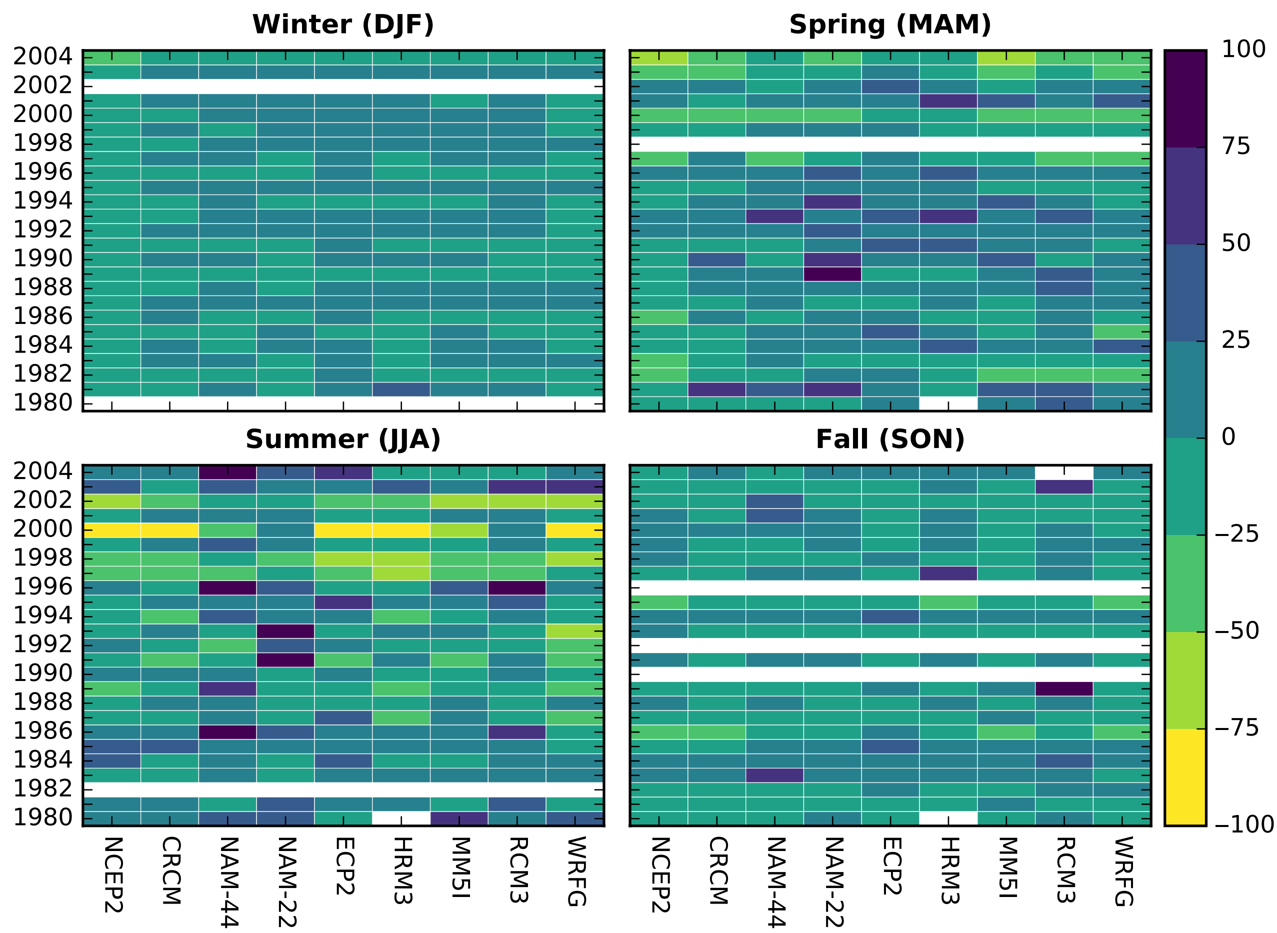

Figure 16. Greatest 3-day total precipitation (mm) for Val Marie, Saskatchewan by RCM and year for 1980–2004; deviations from the observed value. Values are missing for all 1980 and 2002 (winter), HRM3 1980 and all 1998 (spring), HRM3 1980 and all 1982 (summer), HRM3 1980, RCM3 2004 and all 1990, 1992 and 1996 (fall). [Colour figure can be viewed at wileyonlinelibrary.com].

and the number of consecutive dry days. A 1 mm threshold was used as the boundary between dry and wet days and this threshold may need adjusting for some RCMs.

The above results have determined the performance of RCMs over western Canada and paved the way for the construction of higher resolution scenarios of climate change. This work has confirmed that RCMs are able to simulate the range of observed conditions for mean temperature and, to a lesser extent, precipitation, including the representation of extremes. For this area of western Canada, the main error in NCEP2 appears to be in summer precipitation, and all RCMs improve upon the NCEP2 values. Biases apparent in the RCM simulations at the annual and seasonal scale are consistent with the errors in the simulated extremes at the daily timescale. It may be possible to improve these simulations by imposing bias corrections, which can then be applied to output from RCM future climate experiments to construct climate change scenarios. A simple linear scaling bias correction approach (e.g. Lenderink *et al.*, 2007) may be sufficient for mean temperature, as the *q*–*q* plots indicate that most of the RCM distributions appear to be similar to observed, but simply shifted to either warmer or cooler conditions. However, for precipitation, a more complex approach, such as distribution mapping (e.g. Teutschbein and Seibert, 2012), may be probably required.

The disadvantage of all bias correction methods is that they have to make the assumption of stationarity to apply the associated correction factors to data from RCM future climate change experiments. The correction methods also simply statistically manipulate the raw RCM data, i.e. none of the approaches take the physical causes of the temperature and precipitation biases into account (e.g. errors in the large-scale forcing or errors in the parameterisation of cloud processes and precipitation). Dulière *et al.* (2011) and also Kendon *et al.* (2012) have shown that finer grid spacing, which in the latter case allowed the inclusion of a convection routine, significantly improved an RCM's representation of precipitation extremes. This implies that the representation of extremes, particularly for precipitation, will only really be improved by even higher resolution modelling and subsequent improved model physics. In this work, precipitation extremes did not appear to be simulated more successfully with the higher resolution version of CanRCM4 (NAM-22). As the same physics package is used in both resolution versions of CanRCM4, it would be surprising to see large differences between NAM-22 and

NAM-44. Most of the additional value in higher resolution RCM runs comes from the higher resolution of land surface features, in particular topography and lakes. Since CanRCM4 does not have an elaborate approach to modelling lakes and there is little topographic forcing in the western Canada study region, seeing no significant differences between the NAM-44 and NAM-22 runs is reasonable (J. Scinocca, pers. comm. 2016).

Acknowledgements

RCM model output used in this study was obtained from the NARCCAP or from Yanjun Jiao at Environment Canada. NCEP_Reanalysis 2 data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/. The gridded observed monthly ANUSPLIN climate data sets were provided by Dan McKenney and Pia Papadopol at the Canadian Forest Service, Sault Ste. Marie, ON, and temperature and precipitation data from Environment Canada's daily homogenized data set were provided by Lucie Vincent and Eva Mekis, respectively. FORTRAN programs to calculate the extremes indices were obtained from the STARDEX project. The Climate Research Division of Environment Canada provided funding for the research described in this article. The authors are very grateful for the constructive and helpful comments of two anonymous reviewers.

Supporting information

The following supporting information is available as part of the online article:

Figure S1. Spatial pattern correlations by RCM and year for mean temperature over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. (White space indicates missing values.)

Figure S2. Spatial pattern correlations by RCM and year for total precipitation over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. (White space indicates missing values.)

Figure S3. Spring and fall mean absolute error by RCM and year for mean temperature (°C) over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. (White space indicates missing values.)

Figure S4. Spring and fall mean absolute error by RCM and year for total precipitation (mm day^{-1}) over western Canada (49°–60°N, 95°–120°W). The 1980–2004 mean value is given at the top of each plot. (White space indicates missing values.)

Figure S5. Total precipitation (mm day^{-1}) for winter (DJF) 1999: observed (ANUSPLIN) and as simulated by the eight RCMs. Correlation coefficients (r) and mean absolute errors (MAE; mm day^{-1}) are given for western Canada.

Figure S6. CMI (mm) for summer 1988: observed (ANUSPLIN) and as simulated by the eight RCMs. Correlation coefficients (r) and mean absolute errors (MAE; mm) are given for western Canada.

References

Allan RP, Soden BJ. 2008. Atmospheric warming and the amplification of precipitation extremes. *Science* **321**: 1481–1484.

AMIP Project Office. 1996. AMIP II guidelines. *AMIP Newsletter*, No. 8. http://www-pcmdi.llnl.gov/amip/NEWS/amipnl8.html (accessed 31 January 2017)

Blenkinsop S, Fowler HJ. 2007. Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate models. *J. Hydrol.* **342**: 50–71.

Bonsal BR, Aider R, Gachon P, Lapp S. 2013. An assessment of Canadian prairie drought: past, present and future. *Clim. Dyn.* **41**: 501–516.

Caya D, LaPrise R. 1999. A semi-Lagrangian semi-implicit regional climate model: The Canadian RCM. *Mon. Weather Rev.* **127**: 341–362.

Chipanshi AC, Findlater KM, Hadwen T, O'Brien EG. 2006. Analysis of consecutive droughts on the Canadian Prairies. *Climate Res.* **30**: 175–187.

Christensen JH, Christensen OB. 2003. Climate modelling: severe summertime flooding in Europe. *Nature* **421**: 805–806.

Christensen JH, Carter TR, Giorgi F. 2002. PRUDENCE employs new methods to assess European climate change. *EOS Trans. Am. Geophys. Union* **83**: 147.

Colombo A, Etkin D, Karney B. 1999. Climate variability and the frequency of extreme temperature events for nine sites across Canada: implications for power usage. *J. Clim.* **12**: 2490–2502.

Dai A. 2001. Global precipitation and thunderstorm frequencies. Part 1: Seasonal and interannual variations. *J. Clim.* **14**: 1092–1111.

De Elía R, Côté H. 2010. Climate and climate change sensitivity to model configuration in the Canadian RCM over North America. *Meteorol. Z.* **19**: 325–339, doi: 10.1127/094-2948/2010/0469.

Dulière V, Zhang Y, Salathé EP. 2011. Extreme precipitation and temperature over the U.S. Pacific Northwest: a comparison between observations, reanalysis data and regional models. *J. Clim.* **24**: 1950–1964.

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO. 2000. Climate extremes: observations, modelling and impacts. *Science* **289**: 2068–2074.

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M. 2013. Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds). Cambridge University Press: Cambridge, UK and New York, NY, 741–866.

Frei C, Christensen JH, Deque M, Jacob D, Jones RG, Vidale PL. 2003. Daily precipitation statistics in Regional Climate Models: evaluation and intercomparison for the European Alps. *J. Geophys. Res.* **108**: 4124–4142.

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Res.* **19**: 193–212.

Fronzek S, Carter TR. 2007. Assessing uncertainties in climate change impacts on resource potential for Europe based on projections from RCMs and GCMs. *Clim. Change* **81**: 357–371.

Giorgi F, Gutowski WJ. 2015. Regional dynamical downscaling and the CORDEX initiative. *Annu. Rev. Env. Resour.* **40**: 467–490, doi: 10.1146/annurev-environ-102014-021217.

Giorgi F, Marinucci MR, Bates GT. 1993a. Development of a second-generation regional climate model (RegCM2). Part I: boundary-layer and radiative transfer processes. *Mon. Weather Rev.* **121**: 2794–2813.

Giorgi F, Marinucci MR, de Canio G, Bates GT. 1993b. Development of a second-generation regional climate model (RegCM2). Part II: convective processes and assimilation of lateral boundary conditions. *Mon. Weather Rev.* **121**: 2814–2832.

Giorgi F, Hewitson B, Christensen J, Hulme M, Von Storch H, Whetton P, Jones R, Mearns L, Fu C. 2001. Regional climate information:

evaluation and projections (chapter 10). In *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the IPCC*, Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds). Cambridge University Press: Cambridge, UK, 739–768.

Gleckler PJ, Taylor KE, Doutriaux C. 2008. Performance metrics for climate models. *J. Geophys. Res.* **113**: D06104, doi: 10.1029/2007/D008972.

Grell G, Dudhia J, Stauffer DR. 1993. A description of the fifth generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398. NCAR: Boulder, CO, 107 pp.

Gutowski W, Kozak K, Arritt R, Christensen J, Patton J, Takle E. 2007. A possible constraint on regional precipitation intensity changes under global warming. *J. Hydrometeorol.* **8**: 1382–1396.

Gutowski WJ, Arritt RW, Kawazoe S, Flory DM, Takle ES, Biner S, Caya D, Jones RG, Laprise R, Leung LR, Mearns LO, Moufouma-Okia W, Nunes AMB, Qian Y, Roads JO, Sloan LC, Snyder MA. 2010. Regional extreme monthly precipitation simulated by NARCCAP RCMs. *J. Hydrometeorol.* **11**: 1373–1379, doi: 10.1175/2010JHM1297.1.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New MA. 2008. European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.* **113**: D20119, doi: 10.1029/2008JD010201.

Heim RR. 2002. A review of 20th century drought indices used in the United States. *Bull. Am. Meteorol. Soc.* **83**: 1149–1165.

Herrera S, Fita L, Fernandez J, Gutierrez JM. 2010. Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. *J. Geophys. Res.* **115**: D21117, doi: 10.1029/2010JD013936.

Hofstra N, New M, McSweeney C. 2010. The influence of interpolation and station network density on the distribution and extreme trends of climate variables in gridded data. *Clim. Dyn.* **35**: 841–858.

Hogg EH. 1994. Climate and the southern limit of the western Canadian boreal forest. *Can. J. For. Res.* **24**: 1835–1845.

Hogg EH. 1997. Temporal scaling of moisture and the forest-grassland boundary in western Canada. *Agric. For. Meteorol.* **84**: 115–122.

Huntingford C, Jones RG, Prudhomme C, Lamb R, Gash JH, Jones DA. 2003. Regional climate-model predictions of extreme rainfall for a changing climate. *Q. J. R. Meteorol. Soc.* **129**: 1607–1621.

Hutchinson MF. 1995. Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Syst.* **9**: 385–403.

Hutchinson MF. 2004. ANUSPLIN version 4.3. Centre for Resource and Environmental Studies, Australian National University, Canberra, Australia.

IPCC. 2007. Summary for policymakers. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, UK and New York, NY.

IPCC. 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM (eds). Cambridge University Press: Cambridge, UK, and New York, NY, 582 pp.

IPCC. 2014a. *Climate Change 2014: Impacts, Adaptation and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastandrea PR, White LL (eds). Cambridge University Press: Cambridge, UK and New York, NY, 1132 pp.

IPCC. 2014b. *Climate Change 2104: Impacts, Adaptation and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastandrea PR, White LL (eds). Cambridge University Press: Cambridge, UK and New York, NY, 688 pp.

IPCC-TGICA. 2007. *General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment*. Version 2. Prepared by TR Carter on behalf of the Intergovernmental Panel on Climate Change, Task Group on Data and Scenario Support for Impact and Climate Assessment, 66 pp. www.ipcc-data.org/guidelines/TGICA_guidance_sdciaa_v2_final.pdf (accessed 31 January 2017)

Jensen ME, Burman RD, Allen RG (eds). 1990. *Evapotranspiration and Irrigation Water Requirements*. Manuals and Reports on Engineering Practice No. 70. Amercian Society of Civil Engineers: New York, NY, 332 pp.

Jones RG, Noguer M, Hassell DC, Hudson D, Wilson SS, Jenkins GJ, Mitchell JFB. 2004. *Generating high resolution climate change scenarios using PRECIS*. Met Office Hadley Centre: Exeter, UK, 40 pp.

Juang J, Hong S, Kanamitsu M. 1997. The NMC nested regional spectral model: an update. *Bull. Am. Meteorol. Soc.* **78**: 2125–2143.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**: 437–471.

Kanada S, Nakano M, Hayashi S, Kato T, Nakamura M, Kurihara K, Kitoh A. 2008. Reproducibility of maximum daily precipitation amount over Japan by a high-resolution non-hydrostatic model. *SOLA* **4**: 105–108.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **83**: 1631–1643.

Katz RW, Brown BG. 1992. Extreme events in a changing climate: variability is more important than averages. *Clim. Change* **21**: 289–302.

Kendon EJ, Roberts NM, Senior CA, Roberts MJ. 2012. Realism of rainfall in a very high resolution regional climate model. *J. Clim.* **25**: 5791–5806, doi: 10.1175/JCLI-D-11-00562.1.

Keyantash J, Dracup JA. 2002. The quantification of drought: an evaluation of drought indices. *Bull. Am. Meteorol. Soc.* **83**: 1167–1180.

Kharin V, Zwiers F. 2000. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J. Clim.* **13**: 3760–3788.

Kharin V, Zwiers F, Zhang X. 2005. Intercomparison of near surface temperature and precipitation extremes in AMIP-2 simulations. *J. Clim.* **18**: 5201–5223.

Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Fiorino M. 2001. The NCEP/NCAR 50-year reanalysis. *Bull. Am. Meteorol. Soc.* **82**: 247–267.

Lemmen DS, Warren FJ, Lacroix J, Bush E (eds). 2008. *From Impacts to Adaptation: Canada in a Changing Climate 2007*. Government of Canada: Ottawa.

Lenderink G, Buishand A, van Deursen W. 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrol. Earth Syst. Sci.* **11**: 1145–1159.

van der Linden P, Mitchell J. 2009. *ENSEMBLES: Climate Change and Its Impacts: Summary of Research and Results from the ENSEMBLES Project*. Met Office Hadley Centre: Exeter, UK, 160 pp.

McKenney DW, Hutchinson MF, Kesteven JL, Venier LA. 2001. Canada's plant hardiness zones revisited using modern climate interpolation techniques. *Can. J. Plant Sci.* **81**: 129–143.

McKenney DW, Pedlar JH, Papadopol P, Hutchinson MF. 2006. The development of 1901–2000 historical monthly climate models for Canada and the United States. *Agric. For. Meteorol.* **138**: 69–81.

Mearns LO, Hulme M, Carter TR, Leemans R, Lal M, Whetton P. 2001. Climate scenario development (chapter 13). In *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the IPCC*, Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds). Cambridge University Press: Cambridge, UK, 583–638.

Mearns L, Giorgi F, Whetton P, Pabon D, Hulme M, Lal M. 2003. *Guidelines for use of climate scenarios developed from regional climate model experiments*. A report prepared for the IPCC-TGICA, 38 pp. http://www.ipcc-data.org/guidelines/dgm_no1_v1_10-2003.pdf (accessed 31 January 2017)

Mearns LO, McGinnis S, Arritt R, Biner S, Duffy P, Gutowski W, Held I, Jones R, Leung R, Nunes A, Snyder M, Caya D, Correia J, Flory D, Herzmann D, Laprise R, Moufouma-Okia W, Takle G, Teng H, Thompson J, Tucker S, Wyman B, Anitha A, Buja L, Macintosh C, McDaniel L, O'Brien T, Qian Y, Sloan L, Strand G, Zoellick C. 2007. (updated 2014). *The North American Regional Climate Change Assessment Program Dataset*. National Center for Atmospheric Research Earth System Grid Data Portal, Boulder, CO, doi: 10.5065/D6RN35ST. www.narccap.ucar.edu (accessed 31 January 2017)

Mearns LO, Gutowski W, Jones R, Leung R, McGinnis S, Nunes A, Qian Y. 2009. A regional climate change assessment program for North America. *EOS Trans. Am. Geophys. Union* **90**: 311–312.

Mearns LO, Arritt R, Biner S, Bukovsky MS, McGinnis S, Sain S, Caya D, Correia J Jr, Flory D, Gutowski W, Takle ES, Jones R, Leung R, Moufouma-Okia W, McDaniel L, Nunes AMB, Qian Y, Roads J, Sloan L, Snyder M. 2012. The North American Regional Climate Change Assessment Program: overview of phase I results. *Bull. Am. Meteorol. Soc.* **93**: 1337–1362, doi: 10.1175/BAMS-D-11-00223.1.

Mekis É, Vincent LA. 2011. An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmos. Ocean* **49**: 163–177.

Milewska EJ, Hopkinson RF, Niitsoo A. 2005. Evaluation of geo-referenced grids of 1961–1990 Canadian temperature and precipitation normals. *Atmos. Ocean* **43**: 49–75.

Murdock TQ, Sobie SR, Zwiers FW, Eckstrand HD. 2013. Climate change and extremes in the Canadian Columbia Basin. *Atmos. Ocean* **51**: 456–469, doi: 10.1080/07055900.2013.816932.

Osborn T, Hulme M. 1997. Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *J. Clim.* **10**: 1885–1908.

Pal JS, Small EE, Eltahir EAB. 2000. Simulation of regional-scale water and energy budgets: representation of subgrid cloud and precipitation processes within RegCM. *J. Geophys. Res.* **105**(D24): 29579–29594.

Pal JS, Giorgi F, Bi X, Elguindi N, Solmon F, Rauscher SA, Gao X, Francisco R, Zakey A, Winter J, Ashfaq M, Syed FS, Sloan LC, Bell JL, Diffenbaugh NS, Karmacharya J, Konaré A, Martinez D, da Rocha RP, Steiner AL. 2007. Regional climate modeling for the developing world: The ICTP RegCM3 and RegCNET. *Bull. Am. Meteorol. Soc.* **88**: 1395–1409.

Perkins SE, Pitman AJ, Holbrook NJ, McAneney J. 2007. Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature and precipitation over Australia using probability density functions. *J. Clim.* **20**: 4356–4376.

Priestley CHB, Taylor RJ. 1972. On the assessment of the surface heat flux and evaporation using large scale parameters. *Mon. Weather Rev.* **100**: 81–92.

Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V, Pitman A, Shukla J, Srinivasan J, Stouffer RJ, Sumi A, Taylor KE. 2007. Climate models and their evaluation. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, UK and New York, NY, 589–662.

Rauscher SA, Coppola E, Piani C, Giorgi F. 2010. Resolution effects on regional climate model simulations of seasonal precipitation over Europe. *Clim. Dyn.* **35**: 685–711.

van Roosmalen L, Christensen JH, Butts MB, Jensen KH, Refsgaard JC. 2010. An intercomparison of regional climate model data for hydrological impact studies in Denmark. *J. Hydrol.* **380**: 406–419.

Roy P, Gachon P, Laprise R. 2012. Assessment of summer extremes and climate variability over the north-east of North America as simulated by the Canadian Regional Climate Model. *Int. J. Climatol.* **32**: 1615–1627.

Sauchyn DJ, Bonsal B. 2013. Climate change and North American Great Plains' drought. In *Encyclopedia of Environmetrics*, El-Shaarawi A-H, Piegorsch W (eds). John Wiley & Sons Ltd: Chichester, UK, doi: 10.1002/9780470057339.vnn123.

Sauchyn D, Kulshreshtha S. 2008. Prairies. In *From Impacts to Adaptation: Canada in a Changing Climate 2007*, Lemmen DS, Warren FJ, Lacroix J, Bush E (eds). Government of Canada: Ottawa, 275–328.

Sauchyn DJ, Kennedy SD, Stroich J. 2005. Drought, climate change and the risk of desertification on the Canadian Plains. *Prairie Forum* **30**: 143–156.

Sauchyn DJ, Velez Upegui JJ, Masiokas M, Ocampo O, Cara L, Villalba R. 2016. Exposure of rural communities to climate variability and change: case studies from Argentina, Colombia and Canada. In *Implementing Climate Change Adaptation in Communities, Cities, Countries and via Outreach Programmes*, Filho WL, Adamson K, Dunk R (eds). Springer Science: Berlin.

Schaeffer M, Selten FM, Opsteegh JD. 2005. Shifts in means are not a proxy for changes in extreme winter temperatures in climate projections. *Clim. Dyn.* **25**: 51–63.

Scinocca JF, Kharin VV, Jiao Y, Qian MW, Lazare M, Solheim L, Flato GM, Biner S, Desgagne M, Dugas B. 2016. Coordinated global and regional climate modeling. *J. Clim.* **29**: 17–35, doi: 10.1175/JCLI-D-15-0161.1.

Semenov MA, Porter JR. 1995. Climatic variability and the modelling of crop yields. *Agric. For. Meteorol.* **73**: 265–283.

Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X. 2012. Changes in climate extremes and their impacts on the natural physical environment. In *IPCC WGI/WGII Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX)*. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM (eds). Cambridge University Press: Cambridge, UK and New York, NY, 109–230.

Shaw EM. 1994. *Hydrology in Practice*, 3rd edn. Stanley Thorne Ltd: Cheltenham, UK, 569 pp.

Skamarock WC, Klemp JB, Dudia J, Gill DO, Barker DM, Wang W, Powers JG. 2005. A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, NCAR: Boulder, CO, 88 pp.

Skelly WC, Henderson-Sellers A. 1996. Grid box or grid point: what type of data do GCMs deliver to climate impacts researchers? *Int. J. Climatol.* **16**: 1079–1086.

STARDEX. 2005. Downscaling climate extremes, 21 pp. www.cru.uea.ac.uk/projects/stardex (accessed 1 October 2014)

von Storch H, Zorita E, Cubasch U. 1993. Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J. Clim.* **6**: 1161–1171.

Sun Y, Solomon S, Dai A, Portmann RW. 2006. How often does it rain? *J. Clim.* **19**: 916–934.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. 2006. Going to the extremes: an intercomparison of model-simulated historical and future changes in extreme events. *Clim. Change* **79**: 185–211.

Teutschbein C, Seibert J. 2010. Regional climate models for hydrological impact studies at the catchment scale: a review of recent modelling strategies. *Geogr. Compass* **4**: 834–860.

Teutschbein C, Seibert J. 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. *J. Hydrol.* **456–457**: 12–29.

Thornthwaite CW. 1948. An approach towards a rational classification of climate. *Geogr. Rev.* **38**: 55–94.

Vincent LA, Mekis E. 2006. Changes in daily extreme temperature and precipitation indices for Canada over the 20th century. *Atmos. Ocean* **44**: 177–193.

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

Walsh K, McGregor J. 1997. An assessment of simulations of climate variability over Australia with a limited area model. *Int. J. Climatol.* **17**: 201–223.

Wheaton E, Kulshreshtha S, Wittrock V, Koshida G. 2008. Dry times: hard lessons from the Canadian Drought of 2001 and 2002. *Can. Geogr.* **52**: 241–262.

Wilks DS. 2011. *Statistical Methods in the Atmospheric Sciences*, 3rd edn. Academic Press, Elsevier: Oxford, UK and Amsterdam, The Netherlands, 676 pp.

Willmott CJ, Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res*. **30**: 79–82.

Yin H, Donat MG, Alexander LV, Sun Y. 2014. Multi-dataset comparison of gridded observed temperature and precipitation extremes over China. *Int. J. Climatol.* **35**: 2809–2827, doi: 10.1002/joc.4174.

Zwiers F, Zhang X. 2003. Towards regional-scale climate change detection. *J. Clim.* **16**: 793–797.

