Clim Dyn
DOI 10.1007/s00382-013-1717-9

Interannual variability and expected regional climate change over North America

Ramón de Elía · Sébastien Biner · Anne Frigon

Received: 25 May 2012 / Accepted: 1 March 2013

Abstract This study aims to analyse the interannual variability simulated by several regional climate models (RCMs), and its potential for disguising the effect of seasonal temperature increases due to greenhouse gases. In order to accomplish this, we used an ensemble of regional climate change projections over North America belonging to the North American Regional Climate Change Program, with an additional pair of 140-year continuous runs from the Canadian RCM. We find that RCM-simulated interannual variability shows important departures from observed one in some cases, and also from the driving models' variability, while the expected climate change signal coincides with estimations presented in previous studies. The continuous runs from the Canadian RCM were used to illustrate the effect of interannual variability in trend estimation for horizons of a decade or more. As expected, it can contribute to the existence of transitory cooling trends over a few decades, embedded within the expected long-term warming trends. A new index related to signal-to-noise ratio was developed to evaluate the expected number of years it takes for the warming trend to emerge from interannual variability. Our results suggest that detection of the climate change signal is expected to occur earlier in summer than in winter almost everywhere, despite the fact that winter temperature generally has a much stronger climate change signal. In particular, we find that the province of Quebec and northwestern Mexico may possibly feel climate change in winter earlier than elsewhere in North America. Finally, we show that the spatial and temporal scales of interest are fundamental for our capacity of discriminating climate change from interannual variability.

R. de Elía (✉) · S. Biner · A. Frigon
Consortium Ouranos, 550, Sherbrooke Street West, 19th Floor, West Tower, Montreal, QC H3A 1B9, Canada
e-mail: de_elia.ramon@ouranos.ca

R. de Elía
Centre pour l'Étude et la Simulation du Climat à l'Échelle Régionale (ESCER), Université du Québec à Montréal (UQAM), Montreal, QC, Canada

1 Introduction

Until recently, climate change research concentrated particularly on long-term changes despite user needs becoming increasingly focused on shorter terms. The reason for this dissonance between users and scientists was usually stated in the following way: in the short term, climate variability would preclude—or at least could be a serious obstacle to—any attempt to anticipate climate behavior. Over the past few years, however, things have changed. Numerous research projects on decadal prediction are currently underway (e.g., Keenlyside and Ba 2010). In addition, and perhaps related with the fact that global surface temperature has not increased markedly for almost a decade, natural climate variability has received a remarkable increase in attention, as one of the better-understood mechanisms able to temporarily slow down or revert the temperature increase associated to anthropogenic greenhouse gas (GHG) emissions (see for example, Easterling and Wehner 2009; Liebmann et al. 2010; Knight et al. 2009; Santer et al. 2011; and the Web tool described in Greene et al. 2011).

The emphasis on the part of scientists on long-term average temperature increase has perhaps produced an expectation of monotonous growth on the part of users and governments. This expectation of monotonous growth can be seen for example in reactions to short-term cooling

Published online: 19 March 2013

(see for example, Investor's Business Daily 2008; Lawson 2008; or Allegre et al. 2012 op-ed in *The Wall Street Journal* which stated "Perhaps the most inconvenient fact is the lack of global warming for well over 10 years now."). Hawkins (2011) argues that Intergovernmental Panel on Climate Change (IPCC) reports tend to underline the ensemble average and spread of Global Climate Model (GCM) projections—indicating a smooth increase in temperature over the coming century, which has the drawback of masking the role of natural climate variability. Due either to the difficulties of communicating the concept of natural climate variability to users or to the past tendency of many climate scientists to disregard the coming decade as "beyond our ability to predict", not much research has been done until recently on what we should expect in the short term, other than putting our hopes and resources in decadal prediction.

We believe that studies of natural climate variability are important for several reasons: among the non-scientific ones, users and the public have to be informed in the clearest terms that the projected increase in the mean temperature (or change in any other variable) will not be a smooth monotonous movement upwards [several publications have lately called for this point to be given particular importance, among them (Santer et al. 2011)]. This will clarify users' expectations and avoid, first, inadequate adaptation, and second, loss of faith in climate scientists when a cooling trend actually occurs. Successive events such as the extreme cold waves in Europe during the winter of 2009–2010, as well as December 2010 and February 2012, were a clear reminder of this need (see for example, the study of cold months in a warming climate in Räisänen and Ylhäisi (2011); in addition, see Cohen et al. (2010), who argue that the attribution of winter's extreme harsh conditions is critical to the debate of anthropogenic climate change; see also Hawkins (2011) effort to reach a larger audience on the importance to account for natural variability in the discussion about climate change; Deser et al. (2012) call for a dialogue between scientists and policy makers around this topic to avoid raising expectations for accurate regional predictions).

With these issues in mind, and considering that North America is one of the regions with the largest winter temperature interannual variability in the world (see Szeto 2008), we analyzed Ouranos' in-house database [constituted mainly of simulations from the Canadian Regional Climate Model (CRCM)] along with data from North American Regional Climate Change Assessment Program (NARCCAP), and studied the effect of temperature interannual variability in the regional climate change projections that we distribute to Ouranos' users (mainly government agencies). This study has some resemblances to those trying to detect a climate change signal from observations (e.g., Trenberth et al. 2007). Those studies inquire whether a certain change in temperature can be said to be significantly larger than that expected from natural climate variability. In our case, we asked ourselves how many recorded years should pass before a projected change emerges from interannual variability. In our climate models, the only source of external change is the variation in greenhouse gases and aerosols, and so attribution of changes is not in question (for issues related to climate change attribution see The International Ad Hoc Detection and Attribution Group 2005).

A few other studies have had similar goals. Among the earliest are Christensen et al. (2007) and Giorgi and Bi (2009), who investigated the moment when the climate change signal simulated by global models over regions of semi-continental scale overcomes projection uncertainty; the first study considered natural variability as noise, while in the second, noise is viewed as the combination of both model error and natural variability. They called this moment the Time of Emergence (TOE) of the climate change signal, a term that is entering common use. In our study, we concentrate on interannual variability, leaving aside model error and assuming a single GHG emissions scenario. Deser et al. (2010) asked themselves a similar question but instead of wondering about the number of years needed to obtain a detectable signal, they concentrated on the number of ensemble members that are needed for the signal to become statistically significant. Stott and Tett (1998) also presented a study in which detectability of expected temperature increases obtained from a global model are analysed at different spatial scales. Santer et al. (2011) studied the timescale dependence in the signal-to-noise ratio from the lower tropospheric temperature trends against natural variability in observed and modeled datasets. Mahlstein et al. (2011), in a similar study to Giorgi and Bi (2009), concentrate on the location of the earliest significant warming, arguing that such investigation is useful not only for understanding how humans experience climate change in different locations, but also because a range of climate impacts can generally be expected as climate changes exceed past variability. Hawkins and Sutton (2012) pursue a similar objective with a slightly different methodology, putting more emphasis in the estimation of the Time of Emergence. Deser et al. (2012) use a 40-member global model ensemble to study the impact of natural variability on the expected climate change in North America over the next decades.

Studies tackling interannual variability in the context of regional climate model simulations are few, most of them making use of the PRUDENCE (Vidale et al. 2007) and ENSEMBLES (Fischer and Schär 2009) projects which both cover European climate. In general, focus has been on the summer season, where a pronounced projected increase in

interannual surface temperature variability needed attention to improve understanding of the processes involved. Even fewer studies have been performed for North America. Duffy et al. (2006) looked into interannual variability from four RCMs focusing on the Pacific North West. The recent availability of the NARCCAP dataset will certainly facilitate further studies on aspects of interannual variability.

A fair question regarding advantages of using RCMs over GCMs may be raised, especially for surface temperature which evolves in general at scales well resolved by global models. This discussion is part of a more general debate about the value added by RCMs with respect to GCMs (see for example Di Luca et al. 2012). Here, we will circumvent this issue and justify our choice by saying that many impact studies utilise NARCCAP data and that this information will be relevant to those users.

In Sect. 2, the model data used and the methodology are presented, with special emphasis on statistical tools developed for this research. In Sect. 3, results are presented for several studies: First, the local interannual variability over North America at a seasonal scale is examined, and the expected climate change signal is discussed; then, a study of discrimination between the climate change signal and interannual variability is carried out, and conclusions are extended to different spatio-temporal scales. Results are summarized and discussed in Sect. 4.

2 Data and methodology

2.1 Observations

Two terrestrial gridded data sets were used to validate the climate models' historical simulations and to evaluate differences in their representation of past climate. Monthly mean surface temperature data were obtained from the University of East Anglia's Climatic Research Unit (hereafter CRU; Mitchell and Jones 2005) and from the University of Delaware (hereafter UDEL; Matsuura and Willmott 2009). Both data sets were provided on a global latitude-longitude grid with a $0.5° \times 0.5°$ resolution. When comparisons are performed between simulations and observations, all data are linearly interpolated to the CRCM's 45-km polar stereographic grid.

2.2 NARCCAP simulations

Some of the Regional Climate Model (RCM) simulations used in this study were provided by the North American Regional Climate Change Assessment Program (NARCCAP) described in Mearns et al. (2009). The data is referenced in (Mearns et al. 2007). Table 1 presents a list of the RCMs, along with a definition of each RCM's acronym and reference publication, as well as a summary view of the combinations of each run. In NARCCAP, a total of six RCMs were run with a horizontal grid spacing of about 50 km over similar North American domains with experiments that include simulations of climate of the recent past (driven by reanalysis) and climate change simulations (driven by global climate models). The former were driven by lateral boundary conditions derived from the National Centers for Environmental Prediction's (NCEP) Department of Energy (DOE) global reanalysis (Kanamitsu et al. 2002) for the 25-year period between 1980 and 2004, excluding one year of spin-up. The latter comprises RCM simulations, driven at their lateral and lower boundaries by outputs from AOGCM (Atmosphere–Ocean Global Climate Model) simulations, for present (1971–2000) and future climate (2041–2070) using the Special Report on Emissions Scenarios (SRES) A2 emissions scenario (Nakicenovic and Swart 2000), excluding the first 3 years of spin-up. Four AOGCMs are used to drive the RCMs: the Canadian Coupled Global Climate Model version 3 (CGCM3, Scinocca et al. 2008; Flato and Boer 2001), the NCAR (National Center for Atmospheric Research) Community Climate Model version 3 (CCSM3, Collins et al. 2006), the Geophysical Fluid Dynamics Laboratory Climate Model version 2.1 (GFDL Gamdt (The GFDL Global Atmospheric Model Development Team) 2004) and the United Kingdom's Hadley Centre Coupled Climate Model version 3 (HadCM3, Gordon et al. 2000).

From Table 1, we can see that a total of nine RCM-AOGCM pairs are used here to analyze interannual variability and the climate change signal with three RCMs (CRCM, WRFG and RCM3) driven by two AOGCMs (the former two by the CGCM3 and the CCSM, the latter by the CGCM3 and the GFDL) and three RCMs driven by only one AOGCM (ECP2 driven by GFDL, HRM3 driven by HadCM3, and MM5I driven by CCSM).

Reanalysis-driven RCM simulations use AMIP II sea-surface temperature (SST) and sea-ice (SI) concentration observations as lower boundary conditions (Kanamitsu et al. 2002). AOGCM-driven RCM simulations use SST and SI from the AOCGM data. In both reanalysis- and AOGCM-driven simulations, SST and SI surface boundary conditions are updated every 6 h using a linear interpolation between consecutive monthly-mean values. Similarly, boundary conditions are interpolated from the low resolution to the 50-km grid meshes using a linear interpolation in the horizontal.

2.3 Additional simulations

Additional analyses were performed using data from two 140-year RCM climate change simulations covering the 1961–2100 period. These projections were performed

Table 1 Acronyms, full names and modelling group of RCMs involved in the NARCCAP project, as well as the data used to drive each RCM. Of the several possible combinations of pairs of regional model-global model, those marked with an X were used in this research

RCM	RCM's full name and reference	Center	Driven by:	CGCM3	HadCM3	GFDL	CCSM
CRCM	Canadian Regional Climate Model (version 4.2.0) (Music and Caya 2007; Caya and Laprise 1999)	Ouranos/UQAM	NCEP	X			X
ECP2	Experimental Climate Prediction Center Regional Spectral Model (Juang et al. 1997)	UC San Diego/ Scripps	NCEP			X	
HRM3	Hadley Regional Model (version 3) (Jones et al. 2004)	Hadley Centre	NCEP		X		
MM5I	MM5—PSU/NCAR mesoscale model (Grell et al. 1993)	Iowa State University	NCEP				X
RCM3	Regional Climate Model (version 3) (Giorgi et al. 1993a, b)	UC Santa Cruz	NCEP	X		X	
WRFG	Weather Research and Forecasting model (Skamarock et al. 2005)	Pacific Northwest Natl Lab	NCEP	X			X

in-house by the Climate Simulation Team at Ouranos, with version 4.2 of the CRCM (Music and Caya 2007; de Elía and Côté 2010). Both simulations were configured with a regional domain similar but not identical to that of NARCCAP, running with a horizontal grid-size mesh of 45 km (polar stereographic grid true at 60°N) and covering most of North America (200 × 192 grid points). The two simulations were driven by two members of the CGCM3 (members #4 and #5 at T47, differing only in their initial conditions), allowing sampling of some of the natural climate variability. After 2000, both the driving GCM and the RCM follow the SRES A2 future emissions scenario. A spectral nudging technique was applied to large-scale winds within the interior of the regional domain, to keep the CRCM's large-scale flow close to its driving data (Riette and Caya 2002). The spectral nudging applied here is considered to be relatively weak, with horizontal winds of wavelengths greater than 1,400 km nudged with intensity varying in the vertical (starting at around 500 hPa and increasing toward the top, where only 5 % of the CRCM's large scale is replaced by that of its driving data, corresponding to a characteristic relaxation time of 10 h).

2.4 Variables and time periods

In this study, we concentrate on seasonal means of 2-m temperature. Seasons are defined as follows: Winter as December, January and February (DJF); Spring as March, April, May (MAM); Summer as June, July and August (JJA); Fall as September, October, November (SON). All seasons were analyzed, but for reasons of space, only summer and winter will be shown.

2.5 Indexes and statistical tools

Interannual standard deviations (STD) are computed with the biased estimator, which is defined as

$$STD = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2}, \quad (1)$$

where

$$\bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n \quad (2)$$

is the average of the detrended values, with x_n being the detrended value of a given season of a particular year, and N being the total number of seasons. Time series were linearly detrended before variance estimations, and this trend estimated within the time window used.

Comparisons between standard deviations were performed using a Variability Ratio Index VRI_m for each model, which is defined as

$$VRI_m = \begin{cases} \left(\frac{std_m}{std_o} - 1\right) \times 100 & if \quad std_m \geq std_o \\ -\left(\frac{std_o}{std_m} - 1\right) \times 100 & if \quad std_o > std_m \end{cases}, \quad (3)$$

where std_m is the STD value of model m, and std_o is the STD from observations. The interest of this index lies in the symmetry of its absolute value regarding excesses or deficits of variance. This index was inspired from the work of Gleckler et al. (2008) and Scherrer (2010).

Evaluation of the *VRI's* statistical significance depends on the properties of a 5 % right-sided F test (see von Storch and Zwiers 1999). For GCMs or downscaled GCMs to be compared against observations, *VRI* values below a threshold of 36 % are masked; this permits the identification of areas where the model's STD cannot be said to be statistically different from that of the observations.

In order to discriminate the influence of anthropogenic climate change from that of interannual variability, we use a measure related to signal-to-noise ratio. This last is generally used to compare the strength of a given signal to that of the background noise (in this case the signal is the

multi-year trend induced by anthropogenic GHG forcing, and the noise is interannual climate variability). This comparison can be useful to either detect a climate change signal, as is done for example with observations in Liebmann et al. (2010) or with climate model projections, to try to predict the time of emergence of the climate change signal above the noise as in Giorgi and Bi (2009) discussed in the introduction. Here, the signal-to-noise ratio will be used in a way similar to that of the latter, but with several differences that will be described in the following paragraphs. In our case, instead of using the signal-to-noise ratio directly, we will transform it into something hopefully somewhat easier to visualize. The main idea is to write the signal-to-noise ratio in the form of a statistical test and rewrite this latter as a function of the number of years of the sample (this is possible because the error associated to a trend estimation depends on the number of data points, as shown in Appendix Error in trend and variance). This permits us, with a few reasonable approximations, to use this number of years as a measure of discrimination between climate change and interannual variability. That is, given a linear trend we obtain the expected number of years that will be necessary for this discrimination to take place. The expression representing this measure will be called *Expected number of Years before Emergence* (EYE) and is written below. Its full mathematical development and resemblance with Time of Emergence (ToE), as used in other studies, is described in Appendix 1. We can write the EYE or n_α as

$$n_\alpha = \sqrt[3]{12\left(\frac{t_\alpha \sigma}{\hat{\beta}}\right)^2}, \tag{4}$$

where $\hat{\beta}$ represents the expected climate change trend, (in our case, as estimated in Sect. 3.2, by using two temporal 30-year windows), σ is the standard deviation of the interannual variability of the detrended surface temperature (in our case, unless otherwise noted, an average of two interannual variances taken from two different time windows, present and future), and t_α is the percentile of a Gaussian distribution at a level of confidence α (95 % percentile in this study).

It is important to clarify what the index EYE is and is not:

1. As can be seen from expression (4), it is another way of expressing the signal-to-noise ratio, with no additional information. It can be expressed also as the expected number of years that may have to elapse before we can detect a climate change signal from the interannual variability. This elapsed time provides a given confidence level that the expected climate change signal will emerge.
2. Since a linear approximation is used in the development of EYE, results make sense only in those regions for which projected average temperature growth with time is mostly linear. This is the case over long periods of time but obviously fails when the effect of anthropogenic greenhouse gases is felt for the first time (when the slope goes from zero to a roughly constant value during a period). Results from CMIP3 GCMs (Coupled Model Intercomparison Project, Meehl et al. 2007a; not shown) indicate that for North America, this is a good approximation for the entire XXI century with all SRES scenarios. For a given region where linear growth has existed for several years, we may consider any point within that period as the starting point. Hence, the EYE should not be considered as the number of years starting today that we may have to wait until a climate change signal is detected locally; it refers to any time window in a region of linear growth. Conversely, the "time of emergence" defined in the work of Giorgi and Bi (2009), of Mahlstein et al. (2011) and of Hawkins and Sutton (2012) specifies calendar dates.
3. It is not a prediction, but simply an expected value (as indicated by the word "expected" in EYE). In practice, even if no model error were present, actual detection could take place with more or less years of data, depending on the effect of the unexpected lows and highs. This will be further elaborated in Sect. 3.3.
4. The intuitive interpretation of this index makes it a good candidate for policy use. Still, we favor that the EYE should be understood as a rule-of-thumb for exploring geographic areas that may experience earlier than others the effects of climate change, without a precise association to the time of emergence of the signal.

The EYE index, in addition to being affected by the structural errors of models through their imperfect estimation of trends and variability, is also affected by sampling error. A theoretical estimation of this error is presented in Appendix 2.

3 Results

3.1 RCM-downscaled interannual variability

In this section, we analyze the abilities of several GCM-driven regional climate models to estimate the interannual variability of the present climate. The variability of the present climate was obtained from two sets of observations, CRU and UDEL. Comparisons between these two observational gridded datasets yielded very small differences,

although one region west of Ungava Bay in northern Quebec—an area with very few observations—displayed differences of more than 2 °C during winter and comparable differences were found on the northwestern coast of Hudson Bay during summer (not shown). Hereafter, we will refer to the UDEL dataset as the observations. Figure 1a (left panel) displays observations of winter surface temperature interannual standard deviation for the 1971–1999 period, whose main feature is a northwest-southeast axis of large variability (from Alaska to Missouri) reaching values above 4 °C. Figure 2 shows the winter *VRI* index presented in Sect. 2.5, computed against observations, for NARCCAP's regional climate models described in Table 1. The top row shows all CCSM-driven RCMs, the central row all CGCM3-driven runs, and the two lower left panels are for the two GFDL-driven RCMs. The lower-right panel shows the HadCM3-driven HRM3. As discussed in Sect. 2.5, when the two standard deviations are not statistically different, the area is masked with white. It is important to keep in mind that the masked area identifies regions of low statistical significance, however the general lack of robustness of the displayed values remains. For this reason, visual comparison between represented fields should be carried out cautiously. Large areas of variability underestimation (blue) can be seen in RCMs driven by the CGCM3 (second row) in the northwest-southeast axis, although the level of underestimation depends on the RCM. For example, in the state of Indiana, variability is severely underestimated (more than 100 %, indicating that the standard deviation of observations is double that of simulated values) by the WRFG but not by the CRCM. Remarkable overestimations (red) appear in the southeastern US (United States) for both the WRFG-CCSM and the ECP2-GFDL pairs. In these cases, red values exceed 100 %, indicating that model-simulated standard deviation doubles that of the observations. It is worth mentioning that RCMs driven by the same GCM may also show notable differences such as are seen for RCM3 and ECP2, both driven by the GFDL. These results suggest that RCMs add substantial variability information to the driving data. A general view of results is given by legends above each panel, providing the area-averaged *VRI* (indicating the general tendency to over or underestimate the variability), and the area-averaged absolute *VRI* (indicating overall departure from the observed variability).

Inspection of the temperature interannual standard deviation simulated by each RCM (not shown) indicates that some models produce a north–south temperature variability gradient (those driven by the CGCM3), while others generate fields that resemble more the observations, with an area of larger values crossing North America from the northwest to the southeast (as in Fig. 1a). All models seem to capture relatively well the location of the interannual variability maximum in the proximity of Alaska.

Figure 1b (right panel) displays observations of summer surface temperature interannual standard deviation for the 1971–1999 period. Variability is much lower in summer than in winter and dominant patterns are less easy to identify, although observations suggest higher values in central and in extreme northern Canada. Figure 3 shows the *VRI* presented in Sect. 2.5 computed against observations for the same models as described in Fig. 2, but for summer. It can be seen that CCSM-driven RCMs (top row) display variability overestimation in large parts of the continent with a clear maximum in western Canada and another one west of Illinois. Of the CGCM3-driven RCMs (second row), two (RCM3 and CRCM) display very similar patterns with underestimation in northern Canada and strong overestimation on the coast of the southeastern US. The CGCM3-driven WRFG differs considerably in the centre of the US with a strong maximum over Iowa, somehow coherent with that of the CCSM-driven WRFG. A final remarkable feature is the overestimation of the HadCM3-HRM3 in southern US that exceeds 200 % in some areas. It is worth mentioning that, although overestimations and underestimations can be large in relative terms, they are not necessarily so in absolute terms, given the fact that the summer period displays low interannual variability (see Fig. 1b).

Despite differences found among simulations from the various RCMs, we find that models simulate well the strong seasonality of the interannual variability. Winter is characterized by large values, while summer variability is much weaker (see Fig. 1 for observations). Fall and spring display values that are midway between the extreme seasons (not shown). Differences between the largest and the smallest values are remarkable in both models and observations: local values of standard deviation as high as 5 °C are found in the north during winter, while values smaller than 1 °C appear during summer.

It is beyond the scope of this research to inquire whether model variance biases can be quantitatively attributed mostly to the driving GCMs or to the RCMs [see for example Déqué et al. (2011), Vidale et al. (2007) and Duffy et al. (2006)]. However, in order to at least partially explore this issue, we present the temperature variability index for the four driving GCMs over the 1971–1999 period. Figure 4 presents the winter season *VRI* and results suggest that CGCM3 and HadCM3 display the largest biases by underestimating variability in the former and overestimating it in the latter (see legends above panels).

For the case of the CGCM3, results from Fig. 2 indicate that CGCM3-driven RCMs have the tendency to accentuate the lack of winter variability. This is especially the case for the WRFG. Another feature that seems to be

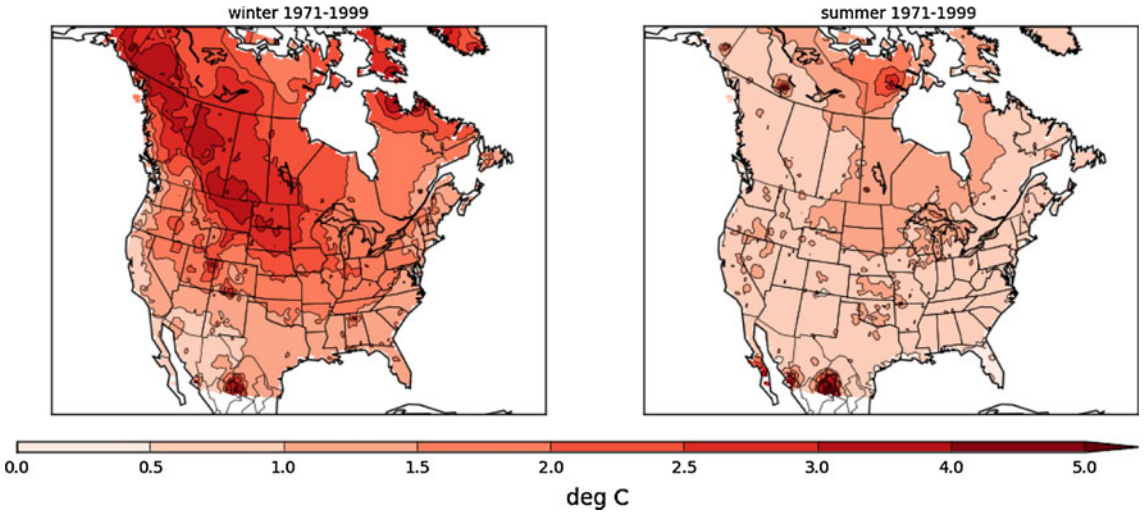

Fig. 1 Surface temperature interannual STD (in °C) over the 1971–1999 period from the UDEL observational dataset for winter (*left panel*) and summer (*right panel*)

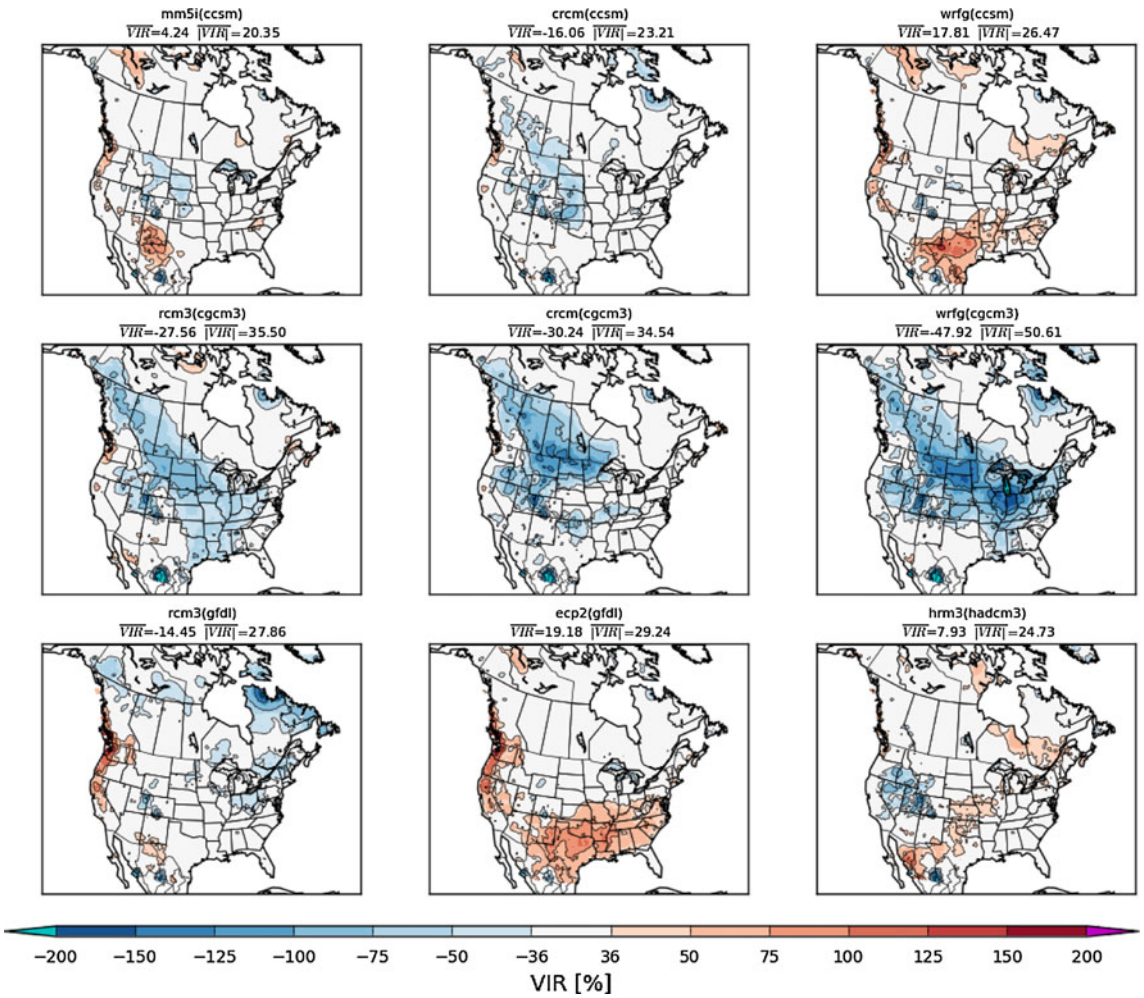

Fig. 2 Winter VRI values of surface temperature for the different RCM simulations driven by GCMs against the UDEL observational dataset over the 1971–1999 period. White masked areas with values between ± 36 %, identify regions where we cannot say that the RCM's STD are statistically different from that of the observations. Above each panel, the legend gives values of area-averaged VRI and the area-averaged absolute VRI

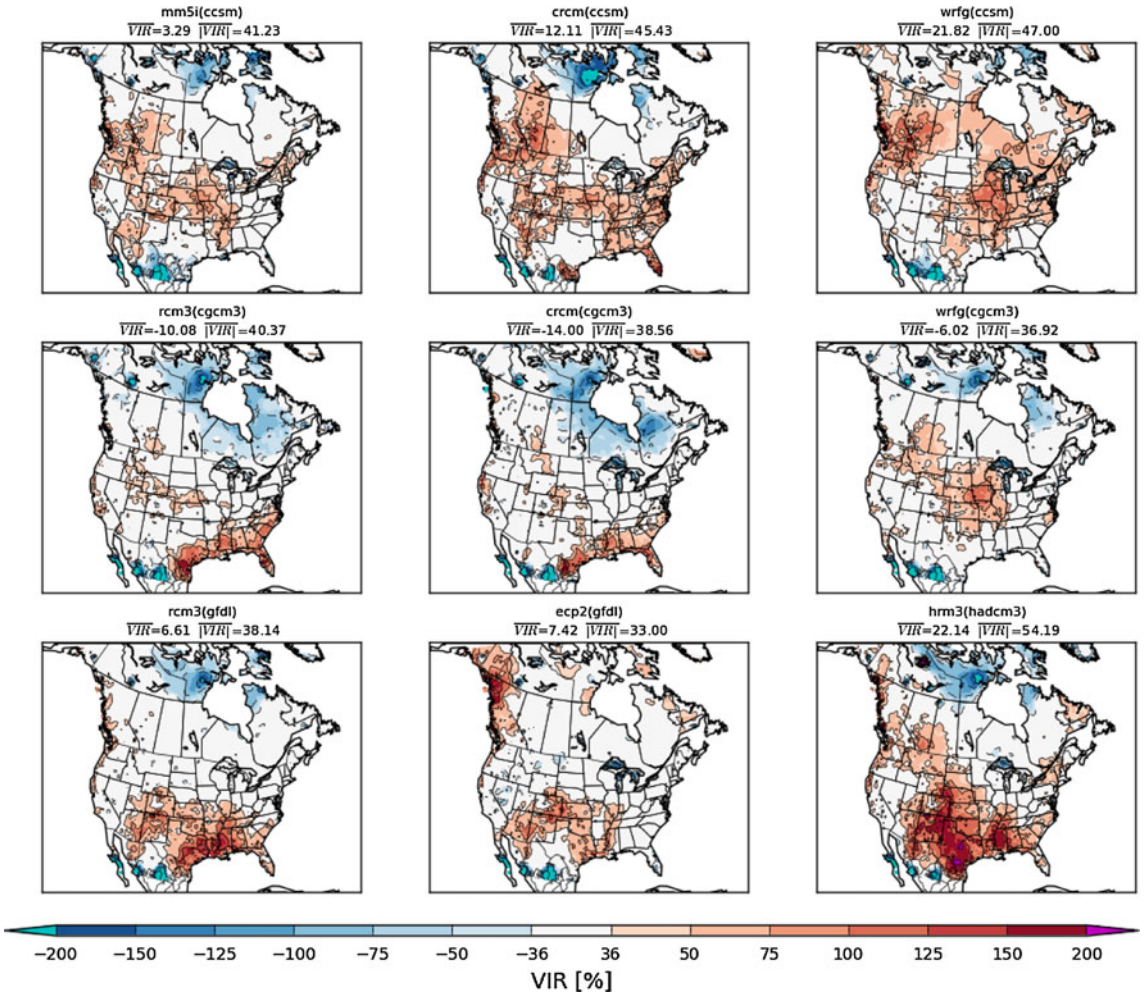

Fig. 3 Same as Fig. 2 but for summer

attributable to the driving GCM is the excess of variability in the west coast of Canada and the US present in the GFDL, and in the GFDL-downscaled data.

During summer departures from observations are proportionally larger than they are in winter, as shown in Fig. 5. The four GCMs overestimate variability in large parts of the domain, this is particularly the case for the GFDL in southern US—and underestimate it in the northern areas. This behavior is similar to that of the driven RCMs (see Fig. 3). Two remarkable facts are that: (1) the most biased model (GFDL) becomes one of the least biased after being downscaled by the ECP2; and (2) The CGCM3 displays overall excessive variance, while the three downscaled simulations all agree in a general lack of it (see legend on top of each panel for values).

The previous results suggest that interannual variability is clearly dependent on the driving model, but that RCMs introduce a large part of their own variability, particularly in summer when absolute values are lower. It is important to note that differences in variability between the driving and the driven model may have several sources, among which are modifications to the large scale by the RCM, the effect of a high-resolution topography, or the role of land-surface schemes.

Further information about attribution of an RCM's interannual variability can be gathered by studying the surface temperature *VRI* of the RCMs driven by NCEP reanalysis, against observations, over the 1980–2003 period (not shown). Unlike when driven by global models, NCEP-driven RCMs show a weaker tendency to underestimate interannual variability during winter. The downscaled climate variability from reanalysis is closer to observations even though important differences between RCM results remain. During summer, all RCMs driven by NCEP reanalysis show an inclination to overestimate the observed variance almost everywhere. The MM5I is the only one whose overestimation can be considered not too far above the threshold of statistical significance. The other models show at least one area of strong overestimation or widespread overestimation.

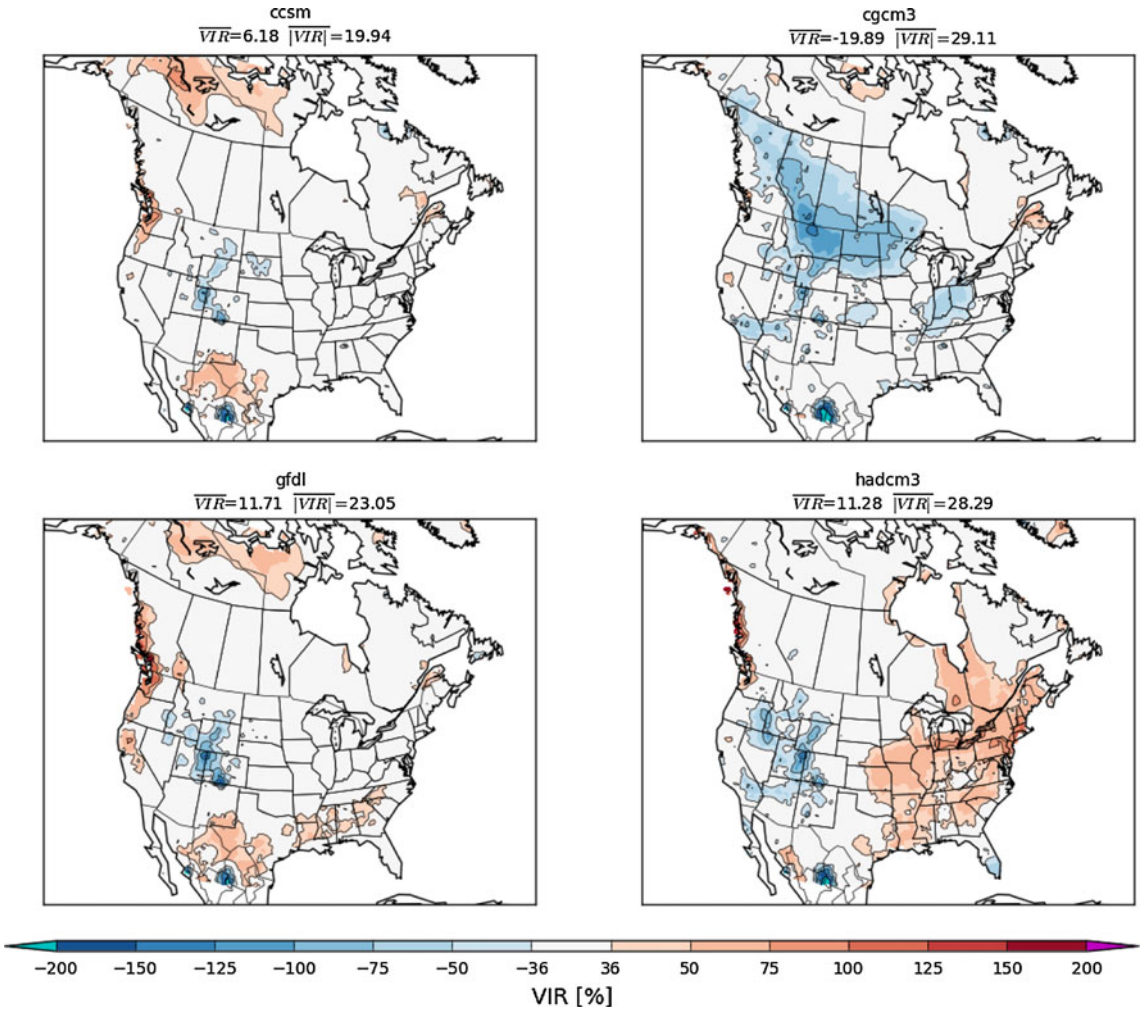

Fig. 4 Winter VRI values of surface temperature for the four driving GCMs against the UDEL observational dataset over the 1971–1999 period. Above each panel, the legend gives values of area-averaged VRI and the area-averaged absolute VRI

3.2 Expected climate change signal

Figures 6 and 7 present, for winter and summer respectively, the projected increase in temperature due to anthropogenic greenhouse gas emissions for the same regional models as in previous figures. Estimations from NARCCAP have been discussed before, for example in Di Luca et al. (2012) and in Sobolowski and Pavelsky (2012), who focus on the southeastern United States. In this work, trends were computed by making a difference between the averages of two time windows: future (2041–2070) minus present (1971–2000), and dividing the result by the intervening years (70). Afterwards, values are normalized to °C/century. In order to simplify the interpretation, in what follows we will consider these trends as good estimations of the expected trends, in the sense that they are affected by a small sampling error, originating from interannual and interdecadal variability. This is equivalent to saying that these estimations are not far from the ensemble mean estimation of an RCM (using an ensemble of a given RCM driven by many members generated by the same GCM). This assumption is necessary, given that a single simulation member is available for each model setup in NARCCAP. As shown in Appendix 2—and in the discussion about Fig. 10 in Sect. 3.3, a typical sampling error is of approximately 1 °C/century, but it will be disregarded in this study.

The winter temperature climate change signal (Fig. 6) displays a clear north–south warming gradient in all models, except in the RCM3-GFDL simulation (lower left panel), where a maximum on a northwest-southeast axis is the main feature (note: inspection of the GFDL driving simulation suggests that this particular pattern is not attributable to the driving data), and in the WRFG-CGCM3, where, with the exception of Hudson Bay's temperature increase, is rather homogeneous. Local maxima of warming are located in Hudson Bay and the Labrador Sea for all models (except for the RCM3–GFDL),

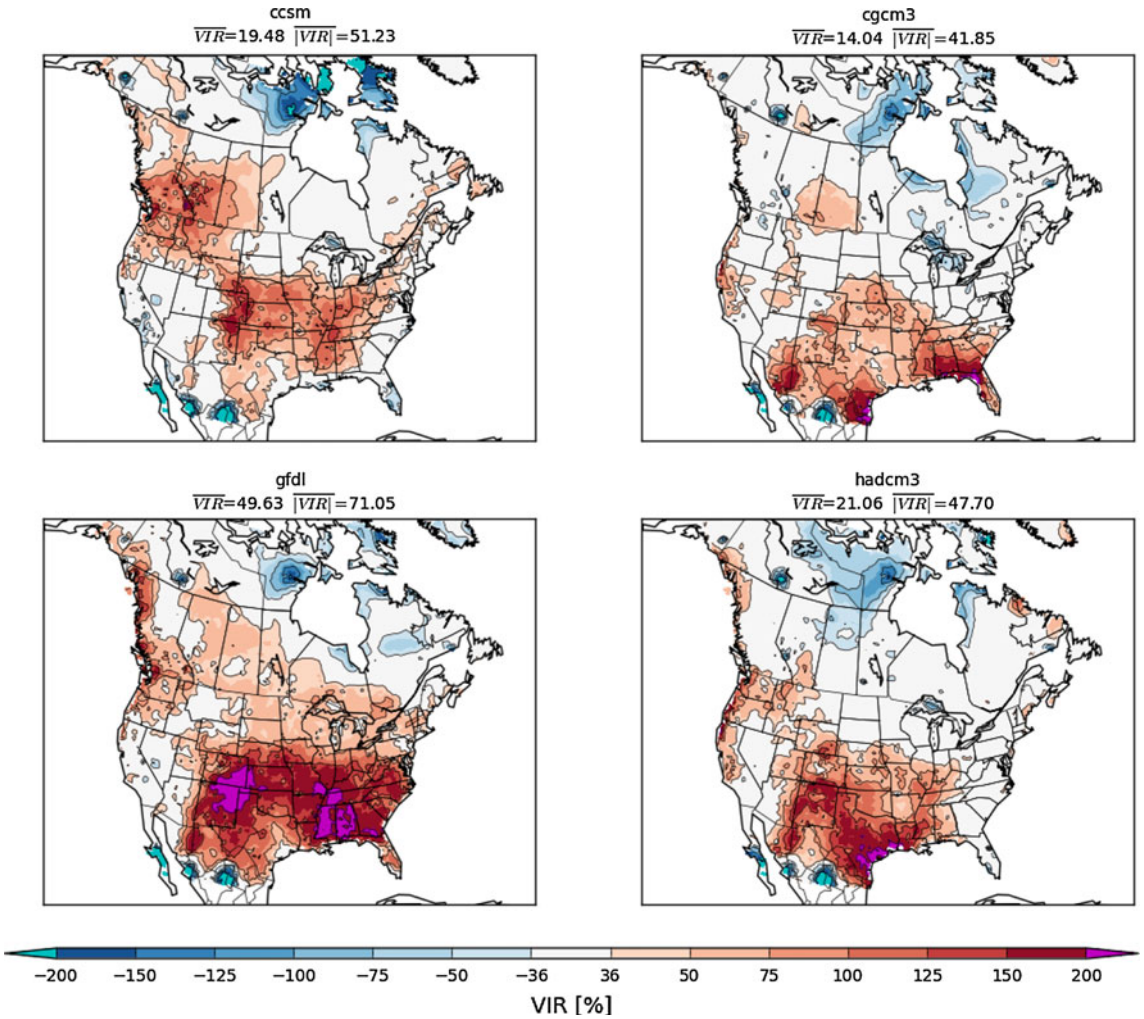

Fig. 5 Same as Fig. 4 but for summer

where a decrease in sea-ice cover, depth and fraction further accentuates temperature increases.

It is interesting to note pattern similarities between results from independent model combinations such as WRFG–CCSM, HRM3–HadCM3 and CRCM–CGCM3, as well as discrepancies between different RCMs driven by the same global model (CGCM3). Even results in Hudson Bay, which depend strongly on ocean data, seem to be affected differently in the driven RCMs (see differences between RCM3–CGCM3 and CRCM–CGCM3, as well as RCM3–GFDL and ECP2–GFDL). In addition, it is worth mentioning that RCMs that project mostly weak temperature trend fields for a driving GCM (WRFG–CGCM3 and RCM3–GFDL) respond differently with different driving models (WRFG–CCSM and RCM3–CGCM3).

The summer temperature climate change signal (Fig. 7) differs considerably from that of winter. Most models display maxima in central United States, except for WRFG, independently of the driving model. These maxima are in part related to positive feedback induced by a decrease of latent-heat fluxes and an increase in sensible fluxes due to a drop in surface soil moisture (Seneviratne et al. 2010).

Regional models show different intensity and locations for these maxima, such as the extension through British Columbia and Alaska in the case of HRM3–HadCM3. It is worth noting the presence of slight cooling in the Northwest Territories in northern Canada in the GFDL-driven ECP2 simulation, as well as only slight warming in western Canada for both WRFG simulations.

3.3 Examples of realistic possible trends at different horizons

In order to have a visual sense of what a *possible* trend—rather than an *expected* trend, as discussed in the previous section—may look like in the next few decades, Figs. 8 and 9 display trends for winter and summer computed over periods of 30, 50, and 70 years starting in the year 2000. The trends are presented for two simulations obtained from continuous 140 year CRCM integrations driven by two

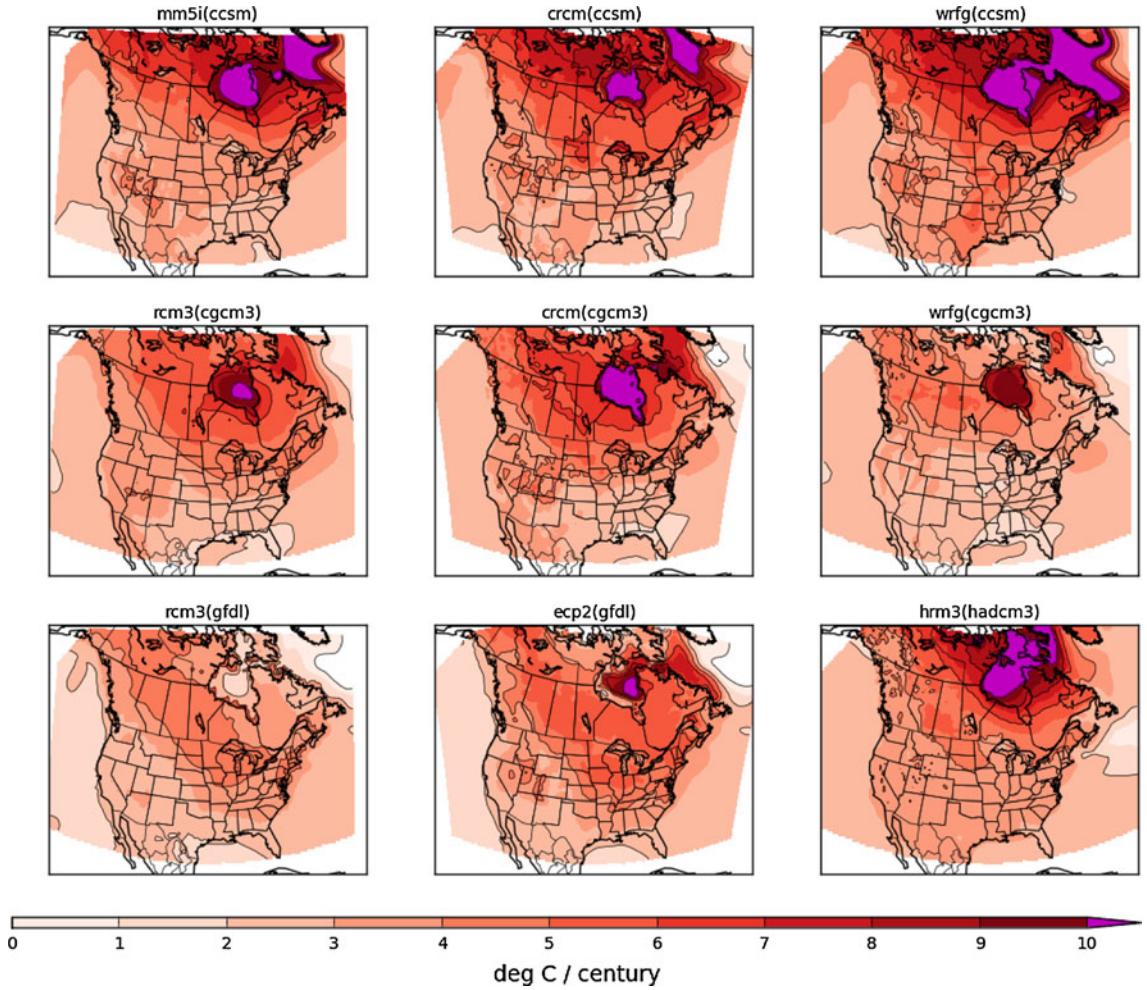

Fig. 6 Projected change in winter surface temperature between the 2041–2070 and 1971–2000 periods for each NARCCAP simulation, normalized to °C/century

different CGCM3 members as described in Sect. 2.3, allowing sampling of the effect of natural climate variability. Note that trends are normalized in °C/century to ease comparison among panels. Figure 8 shows that it may take many years before positive trends in winter affect all regions simultaneously. For example, 30-year trends (upper panels) show a slight negative trend in the southeastern US (left panel), and in northern Canada (right panel). The case displayed also shows that two possible projected trends can be quite different from each other during the first decades, as well as being different from the expected trend (represented by the 70-year trend in the lower panels, where both members start to converge towards similar local trends). For example, the 30-year trend (upper panels) shows two quite distinct values in British Columbia, one of the order of +8 °C/century (left panel), while the other is around +2 °C/century (right panel). Similar results are also reported in Knight et al. (2009). The shaded area on the figure identifies the regions where statistically significant changes are detected with a two-sided Student test at 95 % confidence [see Appendix 1 and Eq. (12)]. In the upper panels of Fig. 8 (30-year trends), we can see that areas where trends are detected depend strongly on the simulation, as can be seen over the Rockies. It is noteworthy that on the upper left panel—that in which the emergence of the climate change signal is present in a large percentage of the western part of the continent—the Pacific Ocean also has an emerging signal over the coast. Also interesting is the difference in behavior between the two simulations regarding Hudson Bay; one becomes ice-free earlier (on the left), showing a strong positive trend, but still not strong enough to emerge from the interannual noise, while the other (on the right) displays a very weak, non-significant trend.

It is interesting to see that simulations on the southeastern US and in particular over Florida have yet to have a detected trend after 50 and even 70 years (see lower right panel). Note that 70 years is the distance between present and future windows used in Sect. 3.2.

Figure 9 displays the summer trends under the same conditions discussed for Fig. 8. Negative trends are in this

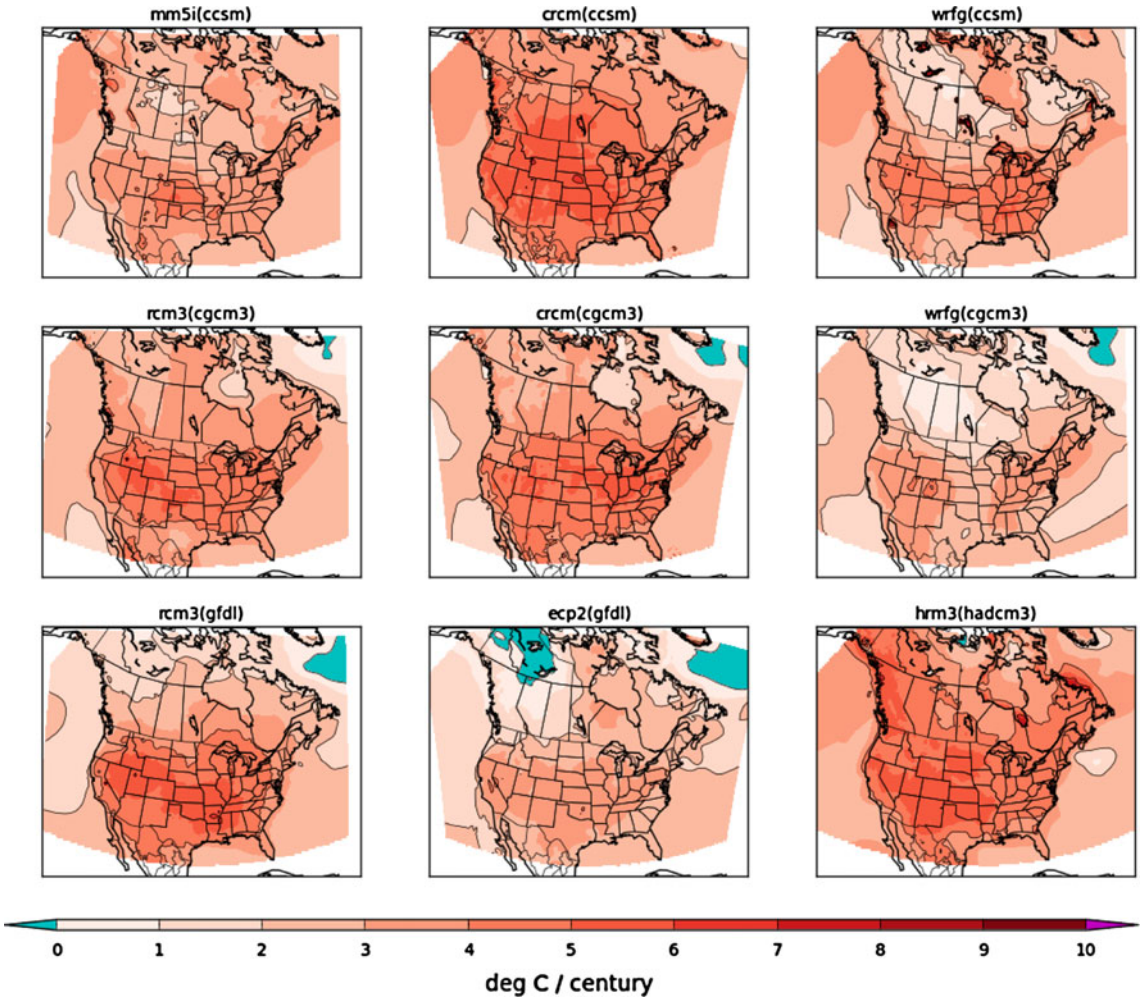

Fig. 7 Same as Fig. 6 but for summer

case non-existent on the continent, except for a small region on the boundary between northern British Columbia and northern Alberta for one of the members (upper left panel). Trend differences in the central US between both members (upper left and right panels), as well as differences between trends obtained with different period lengths (upper right panel 30-year trend, and lower right panel 70-year trend), are quite remarkable. This illustrates that early detection of an almost violent warming trend does not necessarily foretell worse news in the long term than would a low pace increase (see for example the evolution of the trend in South Dakota in both upper panels of Fig. 9).

It is worth noting that statistical significance arrives sooner in summer than it does in winter (compare upper panels of Figs. 9 and 8), although convergence into a similar pattern by the two ensemble members is less fulfilled in summer than in winter (compare lower panels of Figs. 9, 8). The cases here highlight a lack of certainty regarding trends in the next few decades, even when no model error is taken into account, and when no climate prediction—as in Keenlyside and Ba (2010)—is attempted. This predominant role of natural variability as a source of uncertainty at this timescale, was already discussed by Hawkins and Sutton (2009).

A general feature of the other seasons is the presence of large areas of warming as well as smaller but still considerably large regions of cooling. The distribution of these regions is of course dependent on the driving-GCM member chosen, and as illustrated in the previous figures, the chances of finding a negative trend diminishes by increasing the length of the period used to estimate the trend.

An examination of the statistical significance of these trends shows that a region with a given long-term trend may have different realized trends over the shorter term (as in the Rocky Mountain area in the upper panels of Fig. 8 for 30-year trends). This may lead to a detection of climate change in one case (left panel) and not in the other (right panel). The same can be said of two regions that have similar long-term trends (as for example central Quebec and the Northwest Territories, west of Hudson Bay) but

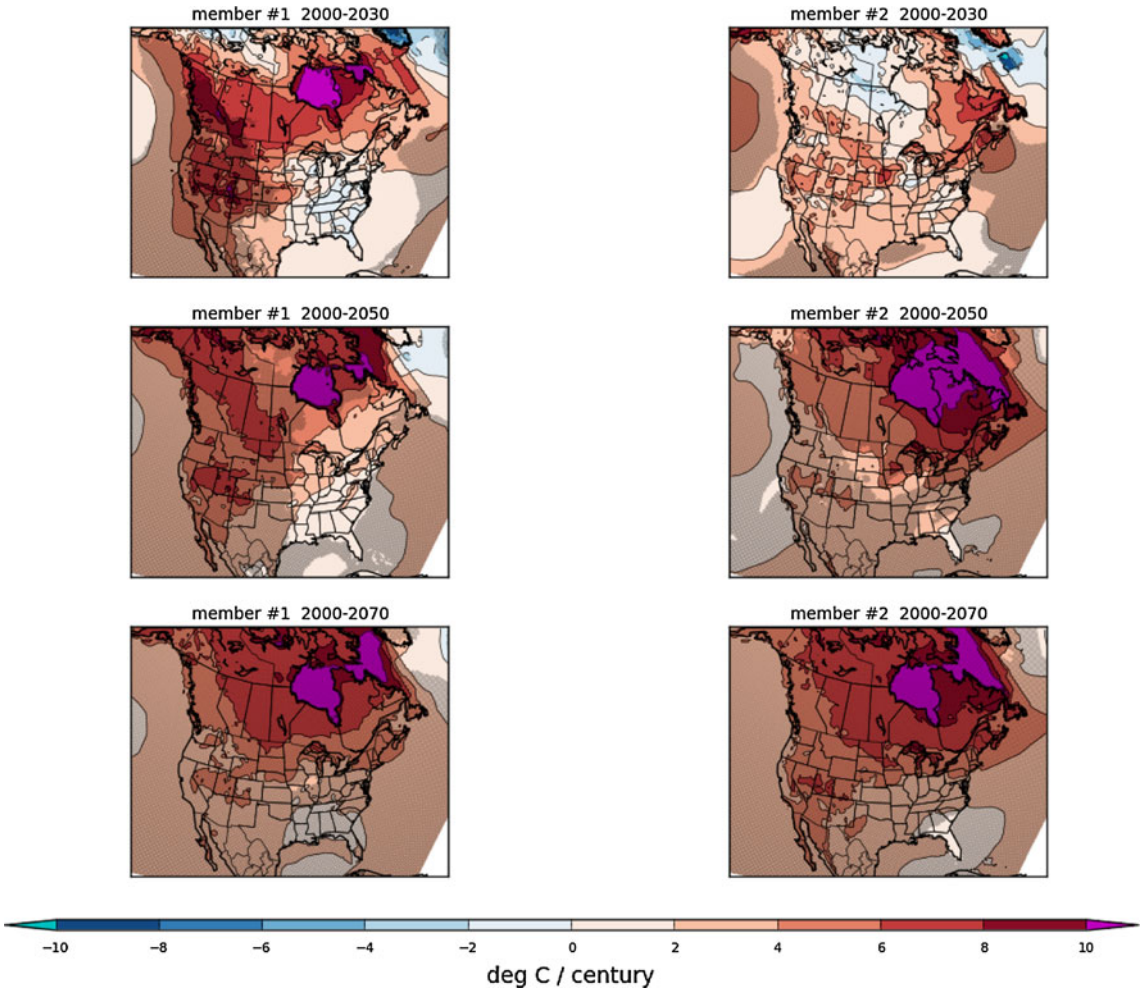

Fig. 8 Trends computed in °C/century for winter surface temperature, from two CRCM 140-year continuous simulations (driven by two members of CGCM3, shown over each column), using periods of 30 (*top*), 50 (*middle*) and 70 years (*bottom*), starting from the year 2000. Statistical significance of the trends using a Student's test (as described in Appendix 1) is displayed by the grey shading

that in the short term may present very different behavior (see Fig. 8, upper right panel, where there is cooling for a 30-year period west of Hudson Bay). This clearly illustrates the fact that interannual variability contributes to the existence of cooling trends embedded in expected long-term warming trends, and to either early or late detection of temperature warming trends.

Figure 10 illustrates how the root mean-squared difference (RMSD) over North America between trend patterns of these two CRCM simulations decreases with period length, and how values depend on the season. It is clear from this image that for periods shorter than 30 years, the variety of possible trends becomes considerably large. The convergence with an average difference less than 1 °C/century is equivalent to that theoretically estimated in Appendix 2. We can see that the largest RMS differences in the short term occur in winter, while the smallest are found in summer, the transition seasons (fall and spring) generally give intermediate values. Note that annual and summer trends have very similar values.

3.4 Discriminating anthropogenic climate change from interannual variability

In Sect. 3.2, we discussed the estimation of the expected climate change trend from several projections, and in the previous section we illustrated that the expected climate change trend may differ substantially from the trend one may in fact observe in the short term. Still, the concept of expected climate change signal is very valuable. In order to quantify our expectations about the potential for each region of North America to be affected by climate change in the shorter term, we will use here the statistical tool presented in Sect. 2.5.

Figure 11 displays the Expected number of Years before Emergence (EYE) for winter temperature for all

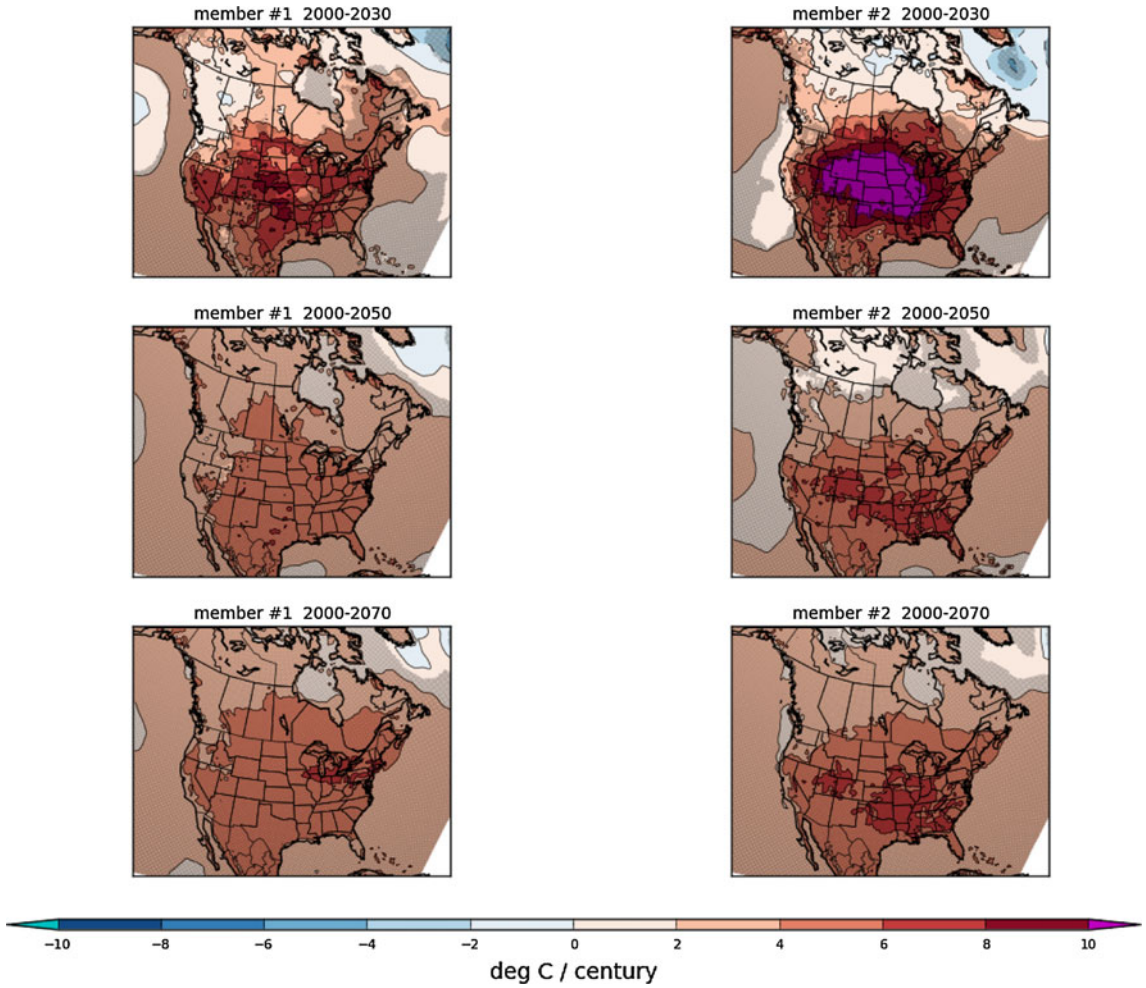

Fig. 9 Same as Fig. 8 but for summer

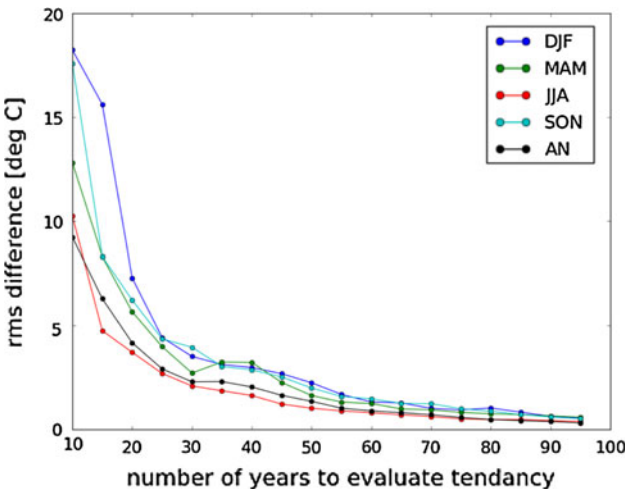

Fig. 10 RMSD over North America between trends (in °C/century) of the two 140-year continuous CRCM simulations when using different length periods (years), for each season (*in color*), and for annual trends (*in black*)

NARCCAP simulations. Let us recall that an EYE value of a large number of years implies a climate change signal that is more difficult to disentangle from interannual variability. In general, models show different behavior, although several common features can be found. A northwest-southeast axis of relatively high values is a common feature. We find that the EYE's spatial pattern for winter is predominantly influenced by that of the interannual standard deviation (see Fig. 1a for observations from the present climate) and not much by the north–south warming gradient of expected change (Fig. 6). Most models agree that for Southern Alaska, British Columbia, central United States, and Florida, many years may be required—in some cases well over sixty—for distinguishing climate change from interannual variability. There is also considerable agreement that the province of Quebec and northwestern Mexico could feel the reality of climate change earlier than other places in North America—within less than 30 years of data according to some models. It is important to recall

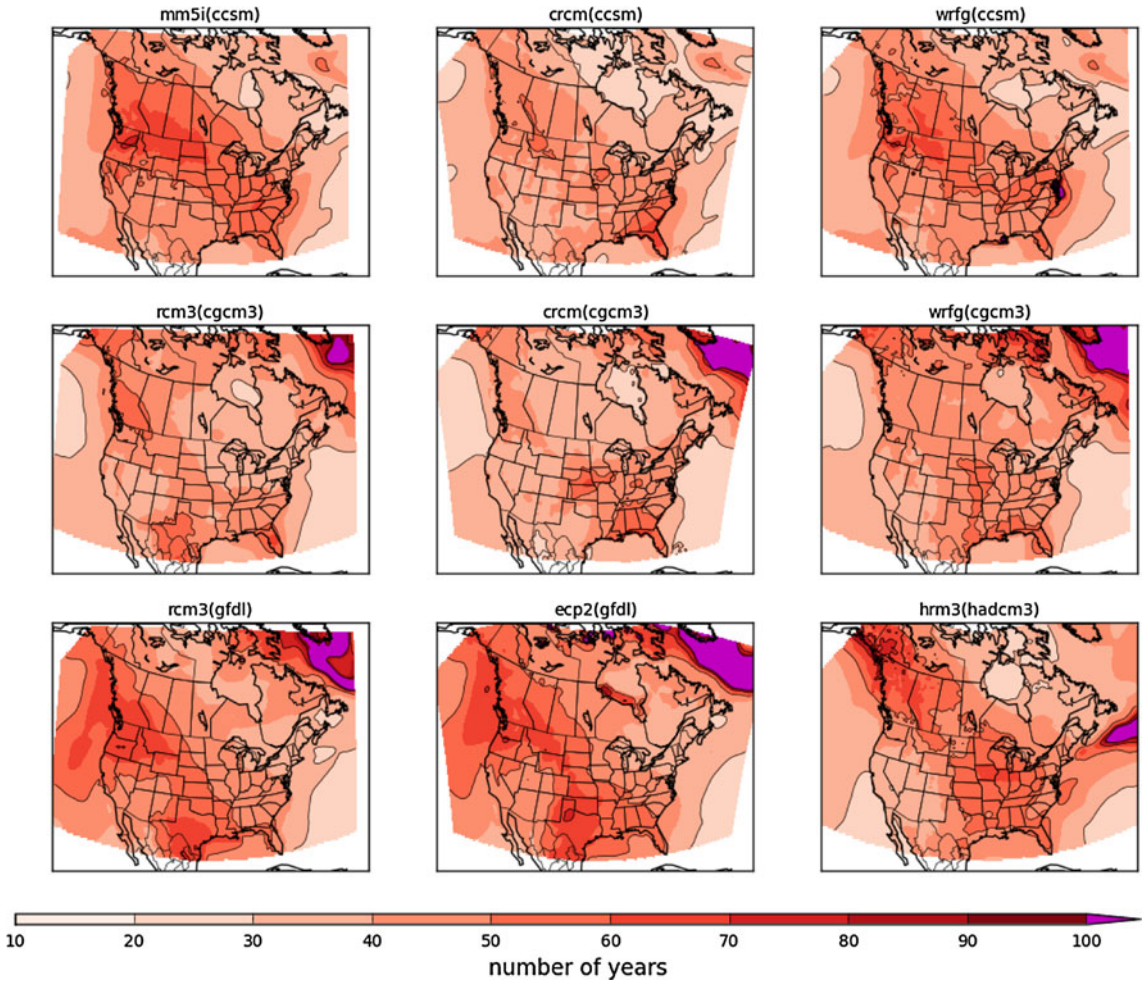

Fig. 11 Expected number of Years before Emergence (EYE) for winter surface temperature change for each NARCCAP simulation

that the trend used in expression (4) is the expected one; so, as discussed in Sect. 3.3, the actual trend may be discriminated earlier or later.

The upper left panel of Fig. 13 shows the median winter EYE value at each gridpoint from the nine simulations depicted in Fig. 11 and confirms what has been discussed above.

Summer temperature EYE presented in Fig. 12 tells a more complicated story. Models disagree substantially over northern areas, a disagreement that can be traced back to climate change trend differences discussed in Sect. 3.2 where it was found that the WRFG (with both driving models) and the ECP2 showed a particularly weak climate change trend. The EYE index further exacerbates these differences since it is a function of the inverse of the trend (see Eq. 4). With a slight maximum of EYE in a north–south axis in the center of North America, it would seem that the interannual standard deviation's pattern (see Fig. 1b for observations from the present climate) has generally a greater influence than the maxima of expected warming in central US (Fig. 7), but this relation is not as strong as was the case in winter. However, the regions of EYE values exceeding 100 years are dominated by the relatively low warming for the different combinations of RCMs and GCMs. The median field of summer EYE is presented in the upper right panel of Fig. 13 and gives an overview of the main features discussed above.

The lower panels of Fig. 13 depict the EYE as in the upper panels but consider the observed variability in the present climate time window instead of the modeled one (cf. Eq. 4). These results allow us to analyse the effect of assuming observed variability as a better approximation than modeled variability. It can be seen that EYE values are sensitive to this assumption but that several features remain unchanged. Notable modifications are the larger values of EYE during winter in the northwest southeast axis (later detection), and lower values for summer in central US (earlier detection). Random error of the EYE fields presented in Figs. 11 and 12 are discussed in Appendix 2.

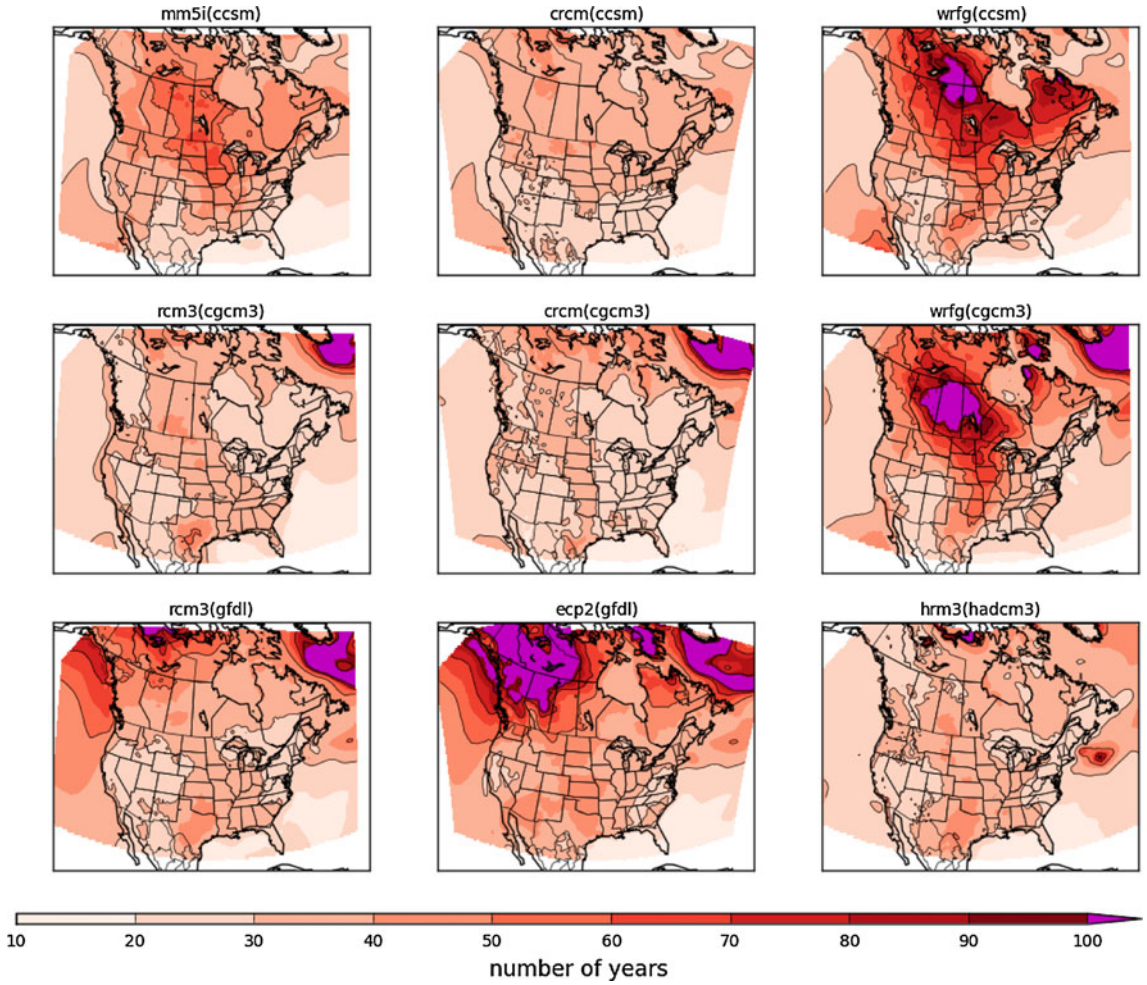

Fig. 12 Same as Fig. 11 but for summer

In several publications it has been mentioned that despite the fact that winter temperature has a much stronger climate change signal, summer temperature changes could be detected or felt earlier, as a result of winter's stronger interannual variability (e.g. Hawkins and Sutton 2012). This is partially true in our case too, especially for median values, but as shown above, very large values of EYE are present during summer in Canada in three of the simulations.

3.5 Importance of spatio-temporal scale

Until now we have discussed interannual climate variability in the context of gridpoint seasonal values. However, many important applications are more concerned with average values at larger scales (see for example the needs of insurance industries in Mills 2005). In this section, we will discuss separately the effect of time and space averaging on reducing the interannual variability, and also its impact on our ability to discriminate climate change.

3.5.1 Area averaging

Figure 14a displays the annual cycle of temperature interannual variability from a CGCM3#4-driven CRCM 140-year continuous simulation (for simulation details see Sect. 2.3). Here, the interannual variability is taken as the standard deviation of the monthly time series detrended with a linear regression. Two different quantities are presented; a black cross indicates the continental North American average of interannual standard deviation at each grid point (henceforth AG, for average gridpoint variability), for monthly values. In mathematical terms, we can write this as

$$\sigma_{AG} = \sqrt{\sum_{i=1}^{N} \sigma_i^2}, \tag{5}$$

where σ_i^2 represents the interannual variance of the detrended monthly time series at each land grid point i, and N is the total number of grid points in continental North America. A black diamond denotes the interannual

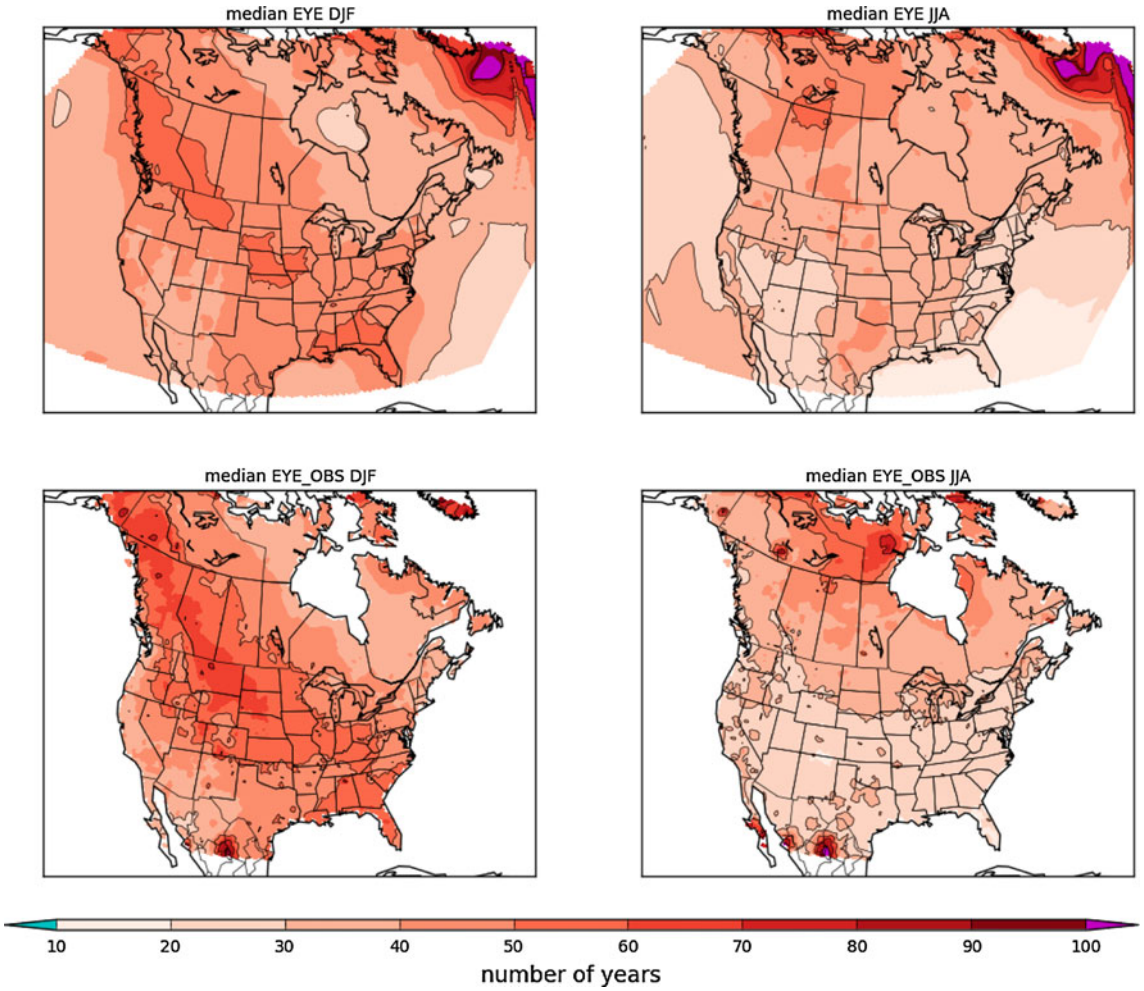

Fig. 13 Median Expected number of Years before Emergence (EYE) for winter (*left*) and summer (*right*). The computation of EYE values use the nine NARCCAP simulated (*upper panels*) and UDEL observed (*lower panels*) interannual variability for the present climate

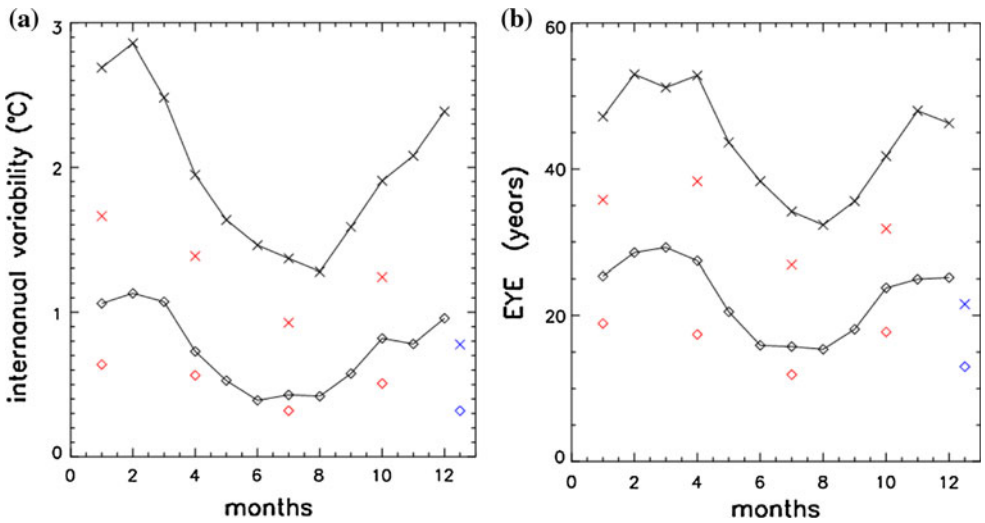

Fig. 14 **a** Temperature interannual variability (standard deviation in °C) from gridpoints in North America (continental averages, *crosses*) and for the entire North American continent (*diamonds*). The colors stand for: black for monthly values, red for seasonal, and blue for yearly, **b** the same as Fig. 13a but for Expected number of Years before Emergence (EYE in years). Values are computed using the CGCM3#4-driven CRCM 140-year simulation described in Sect. 2.3

standard deviation based on a single monthly value representing all of continental North America. The detrended North American time series is obtained by area averaging monthly temperature values for all land gridpoints *i* within the North American continent (henceforth referred as CN, for continental variability).

We can see from the figure that continental averaging of the time series reduces the standard deviation by a factor of around three, when passing from individual gridpoints to the entire continent (from black crosses to black diamonds). In terms of the Expected number of Years before Emergence (EYE) we can see in Fig. 14b with the same symbols, that area averaging corresponds to a reduction of about a factor of two. Stott et al. (2011) illustrate a similar situation with an example of the increase in the temperature's signal-to-noise ratio as estimated over Europe and over Paris.

It is interesting to estimate what the expected reduction by averaging would have been, had the gridpoints been spatially independent. Using the well-known formula for the error associated to the sample mean (e.g., Wilks 2006)

$$\sigma_{\overline{x}} = \frac{\sigma_x}{\sqrt{N}}, \tag{6}$$

we obtain that reduction in interannual standard deviation should have been of a factor of 100 (if the 12,500 grid points in the entire land area of North America had been independent). Hence, the factor three found with RCM data is equivalent to having only about nine independent points (or regions) for the entire North American continent. Identical computations performed with the CRCM driven by the other ensemble member of the CGCM3 yielded very similar results (not shown). This simple estimation yields comparable results to those of more sophisticated computations of spatial dependence of observed surface temperature [see for example Hansen and Lebedeff (1987) and Giorgi (2002)].

The high correlation between grid point values for monthly surface temperature indicates that not much noise reduction is gained by spatial averaging, if this is restricted to small regions. At the same time, it also indicates that point estimations, although suffering from higher noise than regional estimations, are still highly informative. Similar results in terms of histogram spread for future temperatures at different spatial scales can also be found in Hawkins (2011).

3.5.2 *Time averaging*

Figure 14a also illustrates the filtering effects of considering different season lengths on temperature interannual variability, in particular for monthly, seasonal and yearly periods. As in the previous subsection, values are also estimated for AG variability, as given by Eq. 5, and are represented by the same cross symbols, this time in red for the four seasons and in blue for yearly values. The figure also shows the continental (CN) interannual variability of temperature with a diamond, in red for the four seasons and in blue for yearly values. Since successive months and seasons are weakly correlated in time at gridpoint scale, time averaging is quite effective in reducing interannual variability (compare black crosses with red and blue crosses). A similar effect can be seen in the EYE illustrated in Fig. 14b. Reduction of variability by averaging in time seems less effective at the continental scale: notice for example reductions of around 30 % for all seasons for AG variability (compare black and red crosses in Fig. 14a), and that of a much smaller factor for CN variability (compare black and red diamonds in Fig. 14a). Strong seasonality in interannual variability results, nevertheless, in yearly values being quite similar to those for summer.

In terms of both EYE and interannual variability, it can be seen that averaging monthly temperature values on a continental scale (black diamonds) is as effective in reducing noise as averaging gridpoint monthly values into yearly values (blue crosses).

4 Summary and discussion

The aim of this study was to analyse the role of interannual variability in concealing the effect of temperature increases due to GHGs on a local scale. As discussed in the introduction, a number of studies published recently have analysed similar issues, however, very few of these studies have been performed with RCMs, and even fewer over North America. In order to accomplish this, we used an ensemble of regional climate change projections over North America, performed at a nominal resolution of 50 km, most of them belonging to the NARCCAP project.

We first analysed the ability of the different models (i.e. RCMs and GCMs) to reproduce present interannual variability. We found that they simulate well the strong seasonality of temperature interannual variability, but show large departures from observed variability in some cases, and also from that of the driving models.

The expected climate change signal for winter displays a clear north–south warming gradient over North America in the majority of simulations, while the summer warming pattern is not as strong, and shows its maximum in central United States. These general results coincide with those presented in several previous publications, but some models depart from this behaviour. An additional pair of 140-year continuous runs from the Canadian RCM, driven by two different members of the Canadian CGCM3, was used to illustrate the effect of natural climate variability in decadal and multi-decadal scales. The predominance of natural variability at short time scales can hence contribute to either earlier or later detection of temperature warming trends.

A new index EYE, related to signal-to-noise ratio, was developed to evaluate the expected number of years before the warming trend emerges from interannual variability. One of the advantages of this index is the clarity of its

meaning, which may appeal to policy makers. Still, due to the number of approximations and assumptions, we prefer to interpret these results in a qualitative way. Our results suggest that detection of the climate change signal is expected to occur earlier in summer than in winter almost everywhere. Hence, despite the fact that winter temperature has a much stronger climate change signal, summer temperature changes could be detected or felt earlier. This shows, as previous research has also shown, the importance of considering interannual variability, since early detection of climate change can happen in regions different from those for which a stronger climate change signal is expected. In particular, we find that the province of Quebec and northwestern Mexico may possibly feel climate change earlier than elsewhere in North America during winter, with less than 30 years of data according to some models. Santer et al. (2011) show that temperature records of at least 17 years may be required for identifying human effects on global-mean tropospheric temperature, which is surprisingly close to our local results.

Considering the differences found among RCM simulations, these results should inspire caution with respect to any practical application. In addition, we may also wonder whether the benefits of using high-resolution RCM instead of GCM data—privileged by most studies described in the introduction—are not counterbalanced by the necessarily smaller size of the ensemble available. This is a question that deserves further attention especially given the results of Bukovsky (2012) that illustrate that at least short-term temperature trends in RCMs may be quite different from those in the driving data.

Since both time and spatial averaging can be used for noise reduction, the length of the time period as well as the size of the region of interest become fundamental for our capacity to discriminate climate change from natural variability.

It is worth mentioning that in many instances, users do not have the freedom of choosing a preferred spatial or time scale to study climate change. In general, they are confronted with a specific problem with well-defined temporal and spatial scales.

From our results we can see that industries sensitive to weather phenomena that extend over a large part of the year are generally less subject to interannual variability than other industries whose interests are concentrated in a short period. An example of the former is hydroelectric power generation that accumulates water during the entire year, while an example of the latter is tourism for a particular event—like a city summer festival—that may last for less than a week.

Similarly, in the spatial scale, small farm producers are affected by occurrences in their local area, while large agro-industrial corporations or insurance companies distribute weather-related risk over vast regions, sometimes located on different continents. Therefore, a business affected by a given climate vulnerability operating over large geographical regions undoubtedly dampens the effects of climate-related interannual variability, but simultaneously increases the chances of being affected by climate change earlier. In a strategic business report for the insurance industries, Ernst and Young (2008) rank climate change as the top risk factor for the insurance industry.

Acknowledgments The authors wish to thank the North American Regional Climate Change Assessment Program (NARCCAP) for providing the data used in this paper. NARCCAP is funded by the National Science Foundation (NSF), the US Department of Energy (DoE), the National Oceanic and Atmospheric Administration (NOAA), and the US Environmental Protection Agency Office of research and development (EPA). We would also like to acknowledge the Ouranos climate simulation and analysis team for generating and supplying output from the continuous 140-year runs from the Canadian RCM, and the Canadian Center for climate modelling and analysis (CCCma) for kindly providing the CGCM3 driving data. The authors also want to thank the editor and the reviewers, especially one of them who contributed considerably to the improvement of the manuscript.

Appendix 1: Expected number of years for statistically significant climate change

There are several hypotheses that are considered in the definition of the EYE. First of all that the trend is linear; this seems a reasonable approximation as discussed in Sect. 2.5 and as shown for example by Santer et al. (2011) (who considered that "to first order, the signal is timescale invariant") and Mahlstein et al. (2011). This is also implicitly assumed when a trend is obtained from the averages of two time windows. However, although linearity is a reasonable approximation for the present century, the choice of the beginning of this trend causes some difficulties. In this work, we do not concern ourselves with this issue, since we are only interested in this measure as a qualitative indicator. Hence, the view taken here is that climate has been stationary until a given time, after which temperature grows linearly (see for example Figs. 10.4 and 10.5 in Meehl et al. 2007b, where results suggest that linear growth is a good approximation after the 1970s and during several decades). In addition, inferences regarding the discrimination of climate change from natural variability will be taken only considering model data and from an arbitrarily determined starting point. This will provide information regarding the expected length of the time series needed to discriminate climate change, but not its actual place in time.

Two possible misunderstandings should be avoided. First, if the chosen starting point for the analysis of the slope were 1990, the number of years N obtained by this method should not be interpreted as meaning that climate change will *only* be detected in the year $1990 + N$. One

can choose any starting point as long as the linearity approximation stands. Second, the EYE is by definition an expected value; hence, it is not a prediction and cannot be applied to a single realization. It does not indicate the time that emergence from natural variability will happen, but when it is expected to occur.

Another hypothesis is that seasonal values are interannually decorrelated. Our studies suggest (not shown) that the hypothesis of total decorrelation cannot be rejected with local (grid point) time series of a 100 years.

It is important to keep in mind that the data analyzed from the simulated 140-year time series in Sect. 3.3 may lack a realistic decadal variability as is the case of many climate simulations [see validation analyses of multidecadal variability among CMIP3 models discussed by Kravtsov and Spannagle (2008) and in Santer et al. (2011)]. This is one further reason why these estimations should be considered carefully if applied for policy use.

The test for the slope of a cloud of points based on the Student distribution will be developed. For its comparable performance with respect to other tests in temperature trends see Liebmann et al. (2010) and Mahlstein et al. (2011).

The standard definition of a Student test can be written as

$$t = \frac{\hat{\beta}}{s_\beta}, \tag{7}$$

where the numerator is the estimated statistic and the denominator is the estimated standard deviation of the statistic. In the case of the trend of a time series, this expression can be rewritten as

$$t = \frac{\hat{\beta}}{s/\sqrt{SS}}, \tag{8}$$

where

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y})^2, \tag{9}$$

with i being the year of each value y_i, with a total of n years, and $\hat{y}$ is the ordinate value obtained by linear regression. This last expression is the detrended variance. The term SS can be expressed as

$$SS = \sum_{i=1}^{n} i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} i\right)^2. \tag{10}$$

For a thorough derivation of these expressions see Scheaffer and McClave (1990). The term SS can be expressed in a simpler way by rewriting the summations using the properties of finite series of integers as

$$SS = \frac{n}{12}\left(n^2 - 1\right). \tag{11}$$

Using (11), we can now rewrite expression (8) as

$$t = \frac{\hat{\beta}}{s}\sqrt{\frac{n}{12}(n^2-1)}. \tag{12}$$

This expression of the statistic t is now function of the estimated slope, the estimated detrended variance of the time series, and the number of years in the sample.

For n reasonably large, t can be described by a Gaussian distribution. By assuming $n^2 - 1 \approx n^2$, we can isolate n for a given level of significance α, yielding

$$n_\alpha = \sqrt[3]{12\left(\frac{t_\alpha \sigma}{\hat{\beta}}\right)^2}, \tag{13}$$

where σ now represents the detrended variance and the trend $\hat{\beta}$ is in degrees per year.

This expression can be interpreted in several ways. In the present work, $\hat{\beta}$ is going to be the *expected* trend (obtained from a long term climate projection of a single simulation or an ensemble as in Sect. 3.2). At the same time σ, the estimated interannual variability, will be the average detrended standard deviation (average between past and future) when the climate change trend is estimated by a simple difference between future minus present temperature.

An alternative formulation using two time windows instead of a linear regression was also developed (not shown). It produced a similar functional form.

Appendix 2: Error estimations

Error in trend and variance

The error in the estimation of the detrended variance can be obtained using the expression from von Storch and Zwiers (1999), but taking into consideration that there are $n-2$ degrees of freedom,

$$\sigma^2_{S^2_T} = 2\frac{\sigma_T^4}{(n-2)}, \tag{14}$$

where the distribution is assumed to be symmetrical (zero kurtosis), S_T^2 is the estimator of the temperature variance and σ_T^2 the population temperature variance. In our case with two time windows of $n = 30$, the variance considered in Sect. 3.3 is

$$\sigma_T^2 = \frac{1}{2}\left(\sigma^2_{T_p} + \sigma^2_{T_f}\right), \tag{15}$$

where T_p and T_f refer to present and future temperature respectively.

Using properties of the Chi square distribution, and assuming similar variances in both periods, the error of S_T^2 may be written as

$$\sigma^2_{S^2_T} = \frac{\sigma^4_T}{(n-2)}, \tag{16}$$

where the variance over T now considers both present and future temporal windows (the use of a double window explains the decrease in variance estimation error from 14 to 16).

The slope of temperature trends are estimated by using

$$\beta = \frac{1}{N_y}\left(\overline{T}_f - \overline{T}_p\right), \tag{17}$$

where N_y is the distance in number of years between the centers of the time windows.

The error on beta can be estimated using the standard formula for error in the mean, as well as properties of the variance operator (see von Storch and Zwiers 1999) as

$$\sigma^2_\beta = \frac{1}{N_y^2}\left(\frac{\sigma^2_{T_p}}{n} + \frac{\sigma^2_{T_f}}{n}\right). \tag{18}$$

For most gridpoints and NARCCAP models, temperature interannual variances in future and present cannot be said to statistically differ. The exception is the HRM3-HadCM3, which shows a loss of variance in the future during winter over most of Canada, and the CCSM-driven models in the northern tip of Canada during summer, which show an increase of variance (not shown). With this information, the previous expression can confidently be approximated by

$$\sigma^2_\beta \approx \frac{2}{N_y^2}\frac{\sigma^2_T}{n}, \tag{19}$$

where σ^2_T represents the mean variance between the present and future climate. This assumption is taken only for the sake of obtaining an estimate of the error bar, and is consistent with results and assumptions used elsewhere (e.g., Hawkins and Sutton 2009).

Under the conditions of the experiment discussed in this paper ($n = 30$, $N_y = 70$, and S of around 3 °C; see Fig. 1), it can be seen that the sampling error in the estimation of the trend in climate change is of around 1 °C/century.

Error in expected number of years before emergence (EYE)

The other quantity whose error needs to be estimated is the EYE, presented in Sect. 2.5 and derived in Appendix 1. For the sake of completeness, we rewrite below the final expression (13).

$$n_\alpha = \sqrt[3]{12\left(\frac{t_\alpha \sigma}{\beta}\right)^2}. \tag{20}$$

It can be shown that for a function f of two independent variables x and y of the form

$$f = a\left(\frac{x}{y^2}\right)^{1/3}, \tag{21}$$

its error can be written as a function of those of the independent variables by error propagation as

$$\sigma^2_f \approx a^2 \frac{1}{3^2}\left(\frac{x}{y^2}\right)^{2/3}\left(\frac{\sigma^2_x}{x^2} + 4\frac{\sigma^2_y}{y^2}\right), \tag{22}$$

where σ^2_f is the error in function f, and σ^2_x and σ^2_y the error associated to variables x and y. The normalized error can be expressed as

$$\left(\frac{\sigma_f}{f}\right)^2 \approx \frac{1}{9}\left(\frac{\sigma^2_x}{x^2} + 4\frac{\sigma^2_y}{y^2}\right). \tag{23}$$

Using the definition of the EYE from (20) and taking σ^2as x (notice that σ^2 will be estimated by S^2_T), $\hat{\beta}$ as y, $\sqrt[3]{12(t_\alpha)^2}$as a, and n_α as f, it is easy to see that (23) becomes

$$\left(\frac{\sigma_{n_\alpha}}{n_\alpha}\right)^2 \approx \frac{1}{9}\left(\frac{\sigma^2_{S^2_T}}{S^4_T} + 4\frac{\sigma^2_\beta}{\beta^2}\right), \tag{24}$$

Taking advantage of the estimation of errors for each of these variables discussed in (16) and (19), we get

$$\left(\frac{\sigma_{n_\alpha}}{n_\alpha}\right)^2 \approx \frac{1}{9n}\left(1 + 8\frac{S^2_T}{N_y^2\beta^2}\right). \tag{25}$$

In most cases this expression can be approximated by neglecting the left hand side of the addition. In per cent units, we obtain

$$\frac{\sigma_{n_\alpha}}{n_\alpha} \approx 90\,\%\,\frac{1}{\sqrt{n}}\frac{1}{N_y}\frac{S_T}{\beta}. \tag{26}$$

For example, for the typical conditions proposed in this research, with $n = 30$ and $N_y = 70$, a trend of 4 °C/century and an interannual standard deviation of 3 °C gives a relative error of around 20 %. This error increases with larger variability and smaller trend.

References

Allegre C, and others (2012) No need to panic about global warming. Wall Street J, OP-ED, Jan 27

Bukovsky MS (2012) Temperature trends in the NARCCAP regional climate models. J Clim 25:3985–3991. doi:10.1175/JCLI-D-11-00588.1

Caya D, Laprise R (1999) A semi-Lagrangian semi-implicit regional climate model: the Canadian RCM. Mon Weather Rev 127(3): 341–362

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V,

Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional climate projections. In: Solomon S, Qin D, Manning M, Chen Z (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change

Cohen J, Foster J, Barlow M, Saito K, Jones J (2010) Winter 2009–2010: a case study of an extreme Arctic oscillation event. Geophys Res Lett 37:1–6. doi:10.1029/2010GL044256

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The community climate system model version 3 (ccsm3). J Clim 19:2122–2143

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:10.1127/0941-2948/2010/0469

Déqué M, Somot S, Sanchez-Gomez E, Goodess CM, Jacob D, Lenderink G, Christensen OB (2011) The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Clim Dyn. doi:10.1007/s00382-011-1053-x

Deser C, Phillips AS, Bourdette V, Teng H (2010) Uncertainty in climate change projections: the role of internal variability. Clim Dyn. doi:10.1007/s00382-010-0977-x

Deser C, Knutti R, Solomon S, Phillips AS (2012) Communication of the role of natural variability in future North American climate. Nature Clim Change 2:775–780

Di Luca A, de Elía R, Laprise R (2012) Potential for small scale added value of RCM's downscaled climate change signal. Clim Dyn. doi:10.1007/s00382-012-1415-z

Duffy PB, Arritt RW, Coquard J, Gutowski W, Han J, Iorio J, Kim J, Leung L-R, Roads J, Zeledon E (2006) Simulations of present and future climates in the Western United States with four nested regional climate models. J Clim 19:873–895

Easterling DR, Wehner MF (2009) Is the climate warming or cooling? Geophys Res Lett 36:4–6. doi:10.1029/2009GL037810

Ernst & Young (2008): Strategic business risk 2008: insurance. EYG no. EG0015

Fischer EM, Schär C (2009) Future changes in daily summer temperature variability: driving processes and role for temperature extremes. Clim Dyn 33:910–935. doi:10.1007/s00382-008-0473-8

Flato G, Boer G (2001) Warming asymmetry in climate change simulations. Geophys Res Lett 28:195–198

GFDL Gamdt (The GFDL Global Atmospheric Model Development Team) (2004) The new GFDL global atmosphere and land model AM2-LM2: evaluation with prescribed SST simulations. J Clim 17:4641–4673

Giorgi F (2002) Dependence of the surface climate interannual variability on spatial scale. Geophys Res Lett 29:1–4. doi:10.1029/2002GL016175

Giorgi F, Bi X (2009) Time of emergence (TOE) of GHG-forced precipitation change hot-spots. Geophys Res Lett 36:L06709. doi:10.1029/2009GL037593

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

Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. J Geophys Res 113:1–20. doi:10.1029/2007JD008972

Gordon C, Cooper C, Senior C, Banks H, Gregory J, Johns T, Mitchell J, Wood R (2000) The simulation of sst, sea ice extents and ocean heat transports in a version of the hadley centre coupled model without flux adjustments. Clim Dyn 16:147–168

Greene A, Goddard L, Cousin R (2011) Web tool deconstructs variability in twentieth-century climate. EOS Trans Am Geophys Union 92(45):397–398

Grell G, Dudhia J, Stauffer DR (1993) A description of the fifth generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech Note NCAR/TN-398-1A, National Center for Atmospheric Res, p 107

Hansen J, Lebedeff S (1987) Global trends of measured surface temperature. J Geophys Res 92:13345–13372

Hawkins E (2011) Our evolving climate: communicating the effects of climate variability. Weather 66:175–179

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc 90:1095–1107. doi:10.1175/2009BAMS2607.1

Hawkins E, Sutton RT (2012) Time of emergence of climate signals. Geophys Res Lett 39:1–6. doi:10.1029/2011GL050087

IADG (International ad hoc Detection and Attribution Group) (2005) Detecting and attributing external influences on the climate system: a review of recent advances. J Clim 18:1291–1314. doi:10.1175/JCLI3329.1

Investor's Business Daily (2008) Alarmists still heated even as world cools, 4 Nov

Jones RG, Hassell DC, Hudson D, Wilson SS, Jenkins GJ, Mitchell JFB (2004) Workbook on generating high resolution climate change scenarios using PRECIS. UNDP, New York

Juang HMH, Hong SY, Kanamitsu M (1997) The NCEP regional spectral model: an update. Bull Am Meteorol Soc 78:2125–2143

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo J, Fiorino M, Potter G (2002) NCEP-DOE AMIP-II reanalysis (R-2). Bull Am Meteorol Soc 83(11):1631–1643

Keenlyside NS, Ba J (2010) Prospects for decadal climate prediction. Wiley Interdiscip Rev Clim Change 1:627–635

Knight PJ, Kennedy JJ, Folland C, Harris G, Jones GS, Palmer M, Parker D, Scaife A, Stott P (2009) Do global temperature trends over the last decade falsify climate. Bull Am Meteorol Soc 3:22–23

Kravtsov S, Spannagle C (2008) Multidecadal climate variability in observed and modeled surface temperatures. J Clim 21:1104–1121. doi:10.1175/2007JCLI1874.1

Lawson N (2008) An appeal to reason: a cool look at global warming. Gerald Duckworth and Co, London, p 144

Liebmann B, Dole RM, Jones C, Bladé I, Allured D (2010) Influence of choice of time period on global surface temperature trend estimates. Bull Am Meteorol Soc 91:1485–1491. doi:10.1175/2010BAMS3030.1

Mahlstein I, Knutti R, Solomon S, Portmann RW (2011). Early onset of significant local warming in low latitude countries. Environ Res Lett. doi:10.1088/1748-9326/6/3/034009

Matsuura K, Willmott CJ (2009) Terrestrial air temperature and precipitation monthly time series (1900–2008) version 2.01 released June 2009. Available online at http://climate.geog.udel.edu/~climate/

Mearns, L.O. et al. (2007), updated 2011. The North American regional climate change assessment program dataset, National Center for atmospheric research earth system grid data portal, Boulder, CO. Data downloaded 2011-12-01. http://www.earthsystemgrid.org/project/NARCCAP.html

Mearns L, Gutowski W, Jones R, Leung R, McGinnis S, Nuñes A, Qian Y (2009) A regional climate change assessment program for North America. Eos Trans AGU 90(36):311–312

Meehl GA, Covey C, Delworth T, Latif M, McAvaney B, Mitchell JFB, Stouffer RJ, Taylor K (2007a) The WCRP CMIP3

multi-model dataset: a new era in climate change research. Bull Am Meteorol Soc 88:1383–1394

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C (2007b) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Mills E (2005) Insurance in a climate of change. Sci Mag 309(5737):1040–1044

Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int J Climatol 25:693–712. doi:10.1002/joc.1181

Music B, Caya D (2007) Evaluation of the hydrological cycle over the Mississippi river basin as simulated by the Canadian regional climate model (CRCM). J Hydrometeor 8(5):969–988. doi:10.1175/JHM627.1

Nakicenovic N, Swart R (eds) (2000) Special report on emissions scenarios. Cambridge University Press, Cambridge, p 599

Räisänen J, Ylhäisi JS (2011) Cold months in a warming climate. Geophys Res Lett 38:1–6. doi:10.1029/2011GL049758

Riette S, Caya D (2002) Sensitivity of short simulations to the various parameters in the new CRCM spectral nudging. In: Ritchie H (Ed.): Research activities in atmospheric and oceanic modeling, WMO/TD No. 1105, report no. 32: 7.39–7.40

Santer BD, Mears C, Doutriaux C, Caldwell P, Gleckler PJ, Wigley TML, Solomon S, Gillett NP, Ivanova D, Karl TR, Lanzante JR, Meehl G, Stott P, Taylor KE, Thorne PW, Wehner MF, Wentz FJ (2011) Separating signal and noise in atmospheric temperature changes: the importance of timescale. J Geophys Res 116:1–19. doi:10.1029/2011JD016263

Scheaffer R, McClave JT (1990) Probability and statistics for engineers. PWS-Kent, Boston

Scherrer SC (2010) Present-day interannual variability of surface climate in CMIP3 models and its relation to future warming. Int J Clim. doi:10.1002/joc.2170

Scinocca JF, McFarlane NA, Lazare M, Li J, Plummer D (2008) Technical note: the CCCma third generation AGCM and its extension into the middle atmosphere. Atmos Chem Phys 8:7055–7074. www.atmos-chem-phys.net/8/7055/2008/

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ (2010) Investigating soil moisture–climate interactions in a changing climate: a review. Earth Sci Rev 99:125–161. doi:10.1016/j.earscirev.2010.02.004

Skamarock WC, Klemp JB, Dudhia J, Gill JO, Barker DM, Wang W, Powers JG (2005) A description of the advanced research WRF version 2. NCAR Tech Note NCAR/TN-468 + STR, National Center for Atmospheric Research, p 88

Sobolowski S, Pavelsky T (2012) Evaluation of present and future North American Regional Climate Change Assessment Program (NARCCAP) regional climate simulations over the southeast United States. J Geophys Res 117. doi:10.1029/2011JD016430

Stott P, Tett SFB (1998) Scale-dependent detection of climate change. J Clim 11:3282–3294

Stott PA, Christidis N, Betts RA (2011) Changing return periods of weather-related impacts: the attribution challenge. Clim Change 109:263–268. doi:10.1007/s10584-011-0265-8

Szeto KK (2008) On the extreme variability and change of cold-season temperatures in Northwest Canada. J Clim 21:94–113. doi:10.1175/2007JCLI1583.1

Trenberth KE, Jones PD, Ambenje P, Bojariu R, Easterling D, Klein Tank A, Parker D, Rahimzadeh F, Renwick JA, Rusticucci M, Soden B, Zhai P (2007) Observations: surface and atmospheric climate change. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, AverytKB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Vidale PL, Lüthi D, Wegmann R, Schär C (2007) European summer climate variability in a heterogeneous multi-model ensemble. Clim Change 81:209–232. doi:10.1007/s10584-006-9218-z

Von Storch H, Zwiers F (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge

Wilks DS (2006) Statistical methods in the atmospheric sciences, 2nd edn. International geophysics series, vol 59, Academic Press, p 627