The DOI for this manuscript is doi: 10.1175/JAMC-D-18-0021.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Leduc, M., A. Mailhot, A. Frigon, J. Martel, R. Ludwig, G. Brietzke, M. Giguére, F. Brissette, R. Turcotte, M. Braun, and J. Scinocca, 2019: ClimEx project: a 50-member ensemble of climate change projections at 12-km resolution over Europe and northeastern North America with the Canadian Regional Climate Model (CRCM5). J. Appl. Meteor. Climatol. doi:10.1175/JAMC-D-18-0021.1, in press.

ClimEx project: a 50-member ensemble of climate change projections at 12-km resolution over Europe and northeastern North America with the Canadian Regional Climate Model (CRCM5)

Martin Leduc*

Ouranos and Centre ESCER, Université du Québec à Montréal, Montréal, Québec, Canada

Alain Mailhot

INRS - Eau, Terre et Environnement, Québec, Canada

Anne Frigon

Ouranos, Montréal, Québec, Canada

Jean-Luc Martel

École de Technologie Supérieure, Montréal, Québec, Canada

Ralf Ludwig

Ludwig Maximilians University of Munich, Munich, Germany

Gilbert B. Brietzke

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching, Germany

Michel Giguère

Ouranos, Montréal, Québec, Canada

François Brissette

École de Technologie Supérieure, Montréal, Québec, Canada

Richard Turcotte

Ministère du Développement durable, Environnement et Lutte contre les changements climatiques (MDDELCC)

Marco Braun

Ouranos, Montréal, Québec, Canada

John Scinocca

Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada

**Corresponding author address:* Ouranos, 550 Sherbrooke West, West Tower, 19th floor Montreal, Quebec, Canada, H3A 1B9

E-mail: leduc.martin@ouranos.ca

ABSTRACT

The Canadian Regional Climate Model (CRCM5) Large Ensemble (CRCM5-LE) consists of a dynamically downscaled version of the CanESM2 50-member initial-conditions ensemble (CanESM2-LE). The downscaling was performed at 12-km resolution over two domains, Europe (EU) and north-eastern North America (NNA), and the simulations extend from 1950 to 2099, following the RCP8.5 scenario. In terms of validation, warm biases are found over the EU and NNA domains during summer, while during winter, cold and warm biases appear over EU and NNA respectively. For precipitation, simulations are generally wetter than the observations but slight dry biases also occur in summer. Climate-change projections for 2080-2099 (relative to 2000-2019) show temperature changes reaching 8°C in summer over some parts of Europe, while exceeding 12°C in northern Québec during winter. For precipitation, central Europe will become much dryer during summer (-2 mm/day) and wetter during winter (>1.2 mm/day). Similar changes are observed over NNA although summer drying is not as prominent. Projected changes in temperature interannual variability were also investigated, generally showing increasing and decreasing variability during summer and winter respectively. Temperature variability is found to increase by more than 70% in some parts of central Europe during summer, and to increase by 80% in the northernmost part of Québec during the month of May as the snow cover becomes subject to high year-to-year variability in the future. Finally, CanESM2-LE and CRCM5-LE are compared with respect to extreme precipitation, showing evidence that the higher resolution of CRCM5-LE allows a more realistic representation of local extremes, and especially over coastal and mountainous regions.

1. Introduction

As the latest phase of the Bavaria-Québec long-term collaboration on climate change, the ClimEx (Climate change and hydrological Extremes) project investigates the implications of extreme hydrometeorological events on water management in Bavaria and Québec. In order to assess future hydrological impacts from climate change, a complex chain of interlinked processes needs to be taken into account, from how anthropogenic greenhouse gases and aerosols emissions affect the global climate, to the local impacts of climate variability on hydrological processes.

In practice, local hydrological impacts of climate change are studied using a variety of impact models, which use state-of-the-art climate model simulations for inputs. For instance, Global Climate Models (GCMs) (Earth System Models in their current generation) are commonly used to generate large scale climate-change projections over periods from decades to centuries (Collins et al. 2013). However, since GCMs are computationally expensive to run due to their high complexity, they typically use rather coarse spatial resolutions –ranging from 100 to 450 km in the Coupled Model Intercomparison Project Phase 5 (CMIP5) ensemble. These resolutions are often too coarse for hydrological applications (Fatichi et al. 2014; Fowler et al. 2007; Wigley et al. 1990). To fill the gap between GCMs and local scales, downscaling methods have been developed to refine GCM output before driving the hydrological model over a region of interest (Xu 1999; Fowler et al. 2007).

Regional Climate Models (RCMs) offer a convenient approach to downscale GCM output at sufficiently high resolutions for impact modeling. RCMs represent an intermediate step that enables the concentration of computational power on a limited area (rather than on the entire globe as with a GCM) to obtain downscaled climate projections at spatial resolutions typically ranging from 12 to 50 km (Giorgi and Gutowski 2015). RCMs are essentially built as GCMs in terms of dynamical core and parameterizations of sub-grid processes, but must be driven by either GCMs or reanalyses through their lateral and surface boundaries. With their higher resolution, RCMs provide a much better representation of the heterogeneity in surface forcings (e.g., land-sea contrasts, orography, distribution of lakes and rivers, canopy types from vegetation to urban surfaces and soil properties), and an extended range of resolved atmospheric spatio-temporal scales toward finer processes (Lucas-Picher et al. 2016). For all these reasons, RCMs are excellent candidates for driving hydrological models since, compared to coarse-resolution GCMs, they can better account for processes relevant to the scale of many hydrological applications.

Since they provide hydrologically relevant output variables such as precipitation, runoff and evapotranspiration, RCMs can already be used to assess some hydrological impacts from climate change without the need to run a hydrological model (e.g., Music et al. 2012). At the basin scale, however –where complex topography and heterogeneity in soil characteristics are important factors– applications using RCM-driven hydrological models are increasingly popular in the assessment of the hydrological impacts of climate change. It is a common practice to bias correct RCM data to ensure that calibrated hydrological models are driven by realistic meteorological conditions (Muerth et al. 2013). However, there is some debate as to whether an RCM output should be, or not be, bias-corrected prior to drive a hydrological model, as bias correction may introduce further uncertainty into future hydrological simulations (Chen et al. 2017; Clark et al. 2016). Therefore, raw RCM outputs may be preferred to drive hydrological models for some applications, as when Lucas-Picher et al. (2015) reconstructed the Richelieu River flooding of spring 2011, one of the most important flood that occurred in Québec over the last years.

The use of a hydro-modelling chain including a GCM, an RCM and a hydrological model appears to be necessary for the proper assessment of hydrological impacts driven by climate change. This approach, however, requires the various sources of uncertainty that may affect climate-change projections be considered. The World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project (CMIP) multi-model datasets CMIP3 (Meehl et al. 2007), CMIP5 (Taylor et al. 2012) and the upcoming CMIP6 (O'Neill et al. 2016) are vast multi-model ensembles that allow to sample the three main sources of uncertainty: 1) future pathway (scenario) of greenhouse-gas and aerosol (GHGA) emissions; 2) climate sensitivity (structural uncertainty) to fixed GHGA emissions scenario; 3) natural climate variability. These uncertainties are sampled using an "ensemble of opportunity" framework: modelling centres around the world voluntarily generate simulations (based on their own resources and interests) using different GHGA-emission scenarios and GCM models. Some modelling centres also generate multiple realizations of the same experiment (i.e. a specific GCM model driven by a specific GHGA scenario), by adding slight perturbations to the model's initial conditions to sample the effect of natural climate variability (Deser et al. 2014) –an approach that reflects the intrinsic chaotic nature of the climate system. Ensembles involving multiple RCMs are also increasingly common, as they are built upon CMIP-like ensembles of GCMs, such as the CORDEX-coordinated project (Giorgi and Gutowski 2015), which consists of a multi-scenario, multi-GCM, multi-RCM ensemble.

Given the large amount of resources involved in the production of climate model simulations, the multi-model ensemble framework does not generally provide every possible combination of scenarios and models. In addition, models are often represented by a single realization, leading to a weak sampling of natural climate variability. In this sense, it is important to note that, for short-term climate projections, the contribution from natural climate variability to uncertainties is often more important than the contributions from the other factors (Hawkins and Sutton 2009, 2011). Moreover, as extreme events are by definition rare, multiple realizations from one model are important to more robustly assess how climate change may affect their occurrence and intensity.

For extremes floods, for instance, short-term data records translate into large uncertainties for 100-year return-level estimates (Schulz and Bernhardt 2016).

To better understand the role of natural variability and extreme events in current climate projections, it has become increasingly popular in recent years to use the large-ensemble framework, consisting of using a single GCM to generate several realizations of a same experiment. Recent examples are the Community Earth System Model Large Ensemble (CESM-LE) (Kay et al. 2015), which now contains at least 40 members of transient climate-change projections under the RCP8.5 emissions scenario, or its 15-member RCP4.5 version (Sanderson et al. 2015). Similarly, the CanESM2 Large Ensemble (CanESM2-LE) (Fyfe et al. 2017) was produced by the Canadian Centre for Climate Modelling and Analysis (CCCma) at Environment and Climate Change Canada (ECCC), and consists of 50 members under the RCP8.5 scenario. Two 40-member ensembles use the CESM model driven by historical radiative forcing, one using a dynamical ocean model, and the other one observed sea-surface temperatures (Mudryk et al. 2013). The Dutch Challenge Project produced another ensemble, consisting of 62 members from the Community Climate System Model (CCSM1) driven by a business-as-usual scenario (Selten et al. 2004). Also worth noting is the "Essence" project (Sterl et al. 2008), a 17-member ensemble of climate-change simulations using the ECHAM5/MPI-OM climate model forced by the "Special report on Emissions Scenarios" (SRES) A1B pathway. All of these large ensemble projects use many initial-condition members to filter the effects of internal variability to better detect the climate-change signal related to a phenomenon of interest and to estimate the ranges of natural variability, an important information for impacts and adaptations studies.

As natural climate variability can highly depend on the spatial scale under consideration (Giorgi 2002), a better assessment of local climate-change impacts from natural variability and extreme events implies that the regional climate modelling community also began to follow the large-ensemble framework (Deser et al. 2014). An important recent example is "Database for Policy Decision-Making for Future Climate Change" (d4PDF) (Mizuta et al. 2016), which involved the dynamical downscaling of a GCM large ensemble at a spatial resolution of 20 km over Japan. Also, the Canadian Regional Climate Model version 4 (CanRCM4) was used to perform a 35-member ensemble over North America on a 50-km grid mesh (Fyfe et al. 2017). Another example is the 16-member ensemble performed over western Europe and the Alps using the Royal Netherlands Meteorological Institute's regional model KNMI-RACMO2 at 12-km resolution driven by the EC-EARTH global model (Aalbers et al. 2017).

In the scope of the ClimEx project, a 50-member ensemble of climate-change projections at 12-km resolution was produced to assess hydrological impacts from climate change in Bavaria and Québec. This paper presents initial results from this new dataset –the Canadian Regional Climate Model (CRCM5) Large Ensemble (CRCM5-LE; Ouranos 2017, unpublished data)– which is characterized by continuous simulations from 1950 to 2099 under the RCP8.5 GHGA emission scenario and was produced over two domains, Europe and northeastern North America. CRCM5-LE consists of a dynamically downscaled version of CanESM2-LE, which was used to drive the CRCM5 through its boundary conditions.

This paper is organized as follows. Section 2 describes the experimental framework of CRCM5-LE, which builds on CanESM2-LE. In section 3a, a preliminary analysis of the CanESM2-LE initialization is proposed. Sections 3b to e present the main results for CRCM5-LE as follows: model validation with observations (section 3b) and climate-change projections of mean climate (section 3c) and natural variability (section 3d). In section 3e, CRCM5-LE is compared with its driving ensemble (CanESM2-LE) regarding the representation of precipitation extremes. Finally, Section 4 provides a discussion and conclusions.

2. The ClimEx experimental framework

Figure 1 shows the general framework of the ClimEx experiment. The Canadian Earth System Model version 2 (CanESM2; Arora et al. 2011), developed at the CCCma, was used to generate a large initial-condition ensemble of climate-change projections at 2.8° resolution. This dataset, namely the CanESM2 Large Ensemble (CanESM2-LE; Sigmond et al. 2018; Fyfe et al. 2017), is based on a 1,000 years equilibrium simulation (CMIP5 piControl run) forced by pre-industrial conditions (i.e. constant 284.7 ppm atmospheric CO2 concentration). Random atmospheric perturbations (in the cloud-overlap value) were then applied to this simulation to obtain five historical runs starting on 1 January 1850. Applying new random atmospheric perturbations in 1950, each historical run was used to generate ten members, resulting into 50 members from five "families", which differ by a 100-year spin-up from 1850 to 1949. All members were forced with observed emissions (CO2 and non-CO2 GHGs, aerosols and land use) including observed explosive volcanoes and solar-cycle forcings during the historical period up to year 2005, while simulations were extended from 2006 to 2099 following radiative forcing from the representative concentration pathway RCP8.5. From 2006, simulations are forced by a repetition of roughly the last observed solar cycle (prior to 2006) without volcanic aerosol forcing. As will be shown in section 3a, this approach leads to 50 simulations that can be assumed as independent realizations of the modelled climate system after a few years from their initialization in 1950.

The Canadian Regional Climate Model version 5 (CRCM5; Martynov et al. 2013; Separovic et al. 2013) is developed by the ESCER Centre (*Centre pour l'Étude et la Simulation du Climat à l'Échelle Régionale*) of l'Université du Québec à Montréal in collaboration with Environment and Climate Change Canada. This RCM was used by the Ouranos Consortium on Regional Climatology and Adaptation to Climate Change to dynamically downscale CanESM2-LE from 2.8° (≈310 km) to 0.11° (≈12 km) resolution over the 1950-2099 period. The downscaling experiment was performed for two domains, Europe (EU) and northeastern North America (NNA), both using an integration domain of 380 x 380 grid points (Figure 2). In order to validate the performance of the CanESM2 driven CRCM5, ERA-Interim driven runs covering the period from 1979 to 2013 were also performed over both domains and at the same resolution (12 km).

CRCM5 lateral boundary conditions are updated every six hours and linearly interpolated to the five-minute time step of the model. GCM output fields of temperature, surface pressure, specific humidity and horizontal wind components are used to drive the RCM with a one-way nesting procedure over a 10 grid points surrounding blending zone (Davies 1976). A smooth spectral nudging of large scales (Riette and Caya 2002; Separovic et al. 2012) was applied to the horizontal wind component within the RCM domain interior. The spectral nudging configuration consists of large-scale features being defined with a half-response wavelength of 3,113 km and a relaxation time of 13.34 hours. These large scales are imposed inside the RCM domain and vary along the vertical: the nudging strength is set to zero from the surface to a height of 500 hPa and increases linearly onward to the top of the model's simulated atmosphere (10 hPa). In the ERA-Interim driven run, the cut-off length was set slightly shorter due to the higher spatial resolution of ERA-Interim compared to CanESM2. In comparison, the current spectral nudging configuration was much weaker than that used in Liu et al. (2016), where the nudging was applied to all geopotential, horizontal wind, and temperature fields, with shorter relaxation time, and linearly increasing from the top of the planetary boundary layer to a full strength fifth level above. At the bottom boundary, the sea surface temperature and sea ice fraction are prescribed from the driving dataset (CanESM2 or ERA-Interim).

Removing both the 10 grid point wide Davies' blending zone and the 10 point halo (which provides upstream data in the semi-Lagrangian interpolation) included in the periphery of the integration domain results into a 340 x 340 “free domain”, where the model is technically free from direct imposition of lateral boundary conditions. However, RCM applications are known to suffer from boundary effects inside their free domain because small-scale features –which are absent from the lateral boundary conditions– need space (Leduc and Laprise 2009; Leduc et al. 2011; Matte et al. 2016) and time (de Elía et al. 2002) to develop from the coarse-resolution boundary conditions. For this reason, an additional 30 grid point wide security zone was removed within the free domain to favour the development of fine-scale features over the region of interest, corresponding to a 280 x 280 grid points analysis domain (Figure 2) over which all CRCM5 outputs were archived.

The CRCM5 Large Ensemble (CRCM5-LE) dataset will be made available to the scientific community. More information about data access and the complete list of archived variables with corresponding time frequencies (e.g., one hour for precipitation, three hours for surface-air temperature) are posted at `www.climex-project.org`.

3. Results

a. Spin-up time from initial conditions in CanESM2-LE

The fact that large ensembles allow to thoroughly quantify natural climate variability relies on the assumption that the ensemble members consist of independent realizations of the model’s climate system. While climate models are expected to forget their initial conditions after some spin-up time from the beginning of a simulation, it is not clear how much time is required before all members from the five families (see Figure 1) become completely independent. This question is important since a longer spin-up time means a shorter simulated period available for climate analysis. In addition, a lack of independence between ensemble members could undermine further statistical assessments (e.g., extreme values) from both CanESM2-LE and CRCM5-LE by reducing the effective number of independent members.

In order to assess the length of the spin-up time in the current experiment, the time evolution of the inter-member spread is analyzed using various five-member ensemble combinations that may belong to one of the following two categories: 1) runs sharing the Same Ocean Initial Conditions (SOIC) in 1950 (i.e. five members from a same ocean family); 2) runs with Mixed Ocean IC (MOIC) (i.e. five members, one from each ocean family). Ten five-run ensembles were constructed for each category.

The Inter-Member Standard Deviation (IMSD) was calculated for each five-member ensembles and averaged over either land or ocean grid points for various time period. Figure 3a presents the ranges of land-averaged IMSD obtained for the SOIC and MOIC categories respectively during the first year of simulation. It can be seen that after about 100 days, the surface-air temperature over land appears to become independent from its initial conditions in the SOIC ensembles, as seen by the overlap with the MOIC distribution. However, over ocean (Figure 3b, first 1100 days shown), the SOIC ensembles completely overlap the MOIC distribution after a much longer period of time, namely around 800 days of simulation. In comparison, the spin-up period obtained for precipitation (Figure 3c and d) is around 25 and 150 days over land and ocean grid points respectively. It is clear that for slowly evolving processes such as the deep-ocean circulation, the spin-up period would range from hundreds to thousands years (Stouffer 2004) although these time scales are beyond the scope of the ClimEx ensemble framework. For the time scales, regions and variables of interest here, it is reasonable to assume that the CanESM2-LE members are independent a few years after initialization, and therefore that they consist of independent boundary conditions for driving CRCM5-LE.

b. Validation of the historical climate

In this section, the CRCM5 is evaluated in terms of its performance to reproduce the historical climatology. Since biases in the output of an RCM can originate both from inaccurate driving data, or due to the RCM itself, the performance of the ERA-Interim driven run is first compared with that driven by the first CanESM2 member to investigate the possible sources of bias. Here, only one member of the large ensemble (e.g., rather than the ensemble mean) is used to make a proper comparison with the single realization of the ERA-Interim run. Using a 32-year climate period for validation, the climates of the different members slightly differ due to internal variability, but the general conclusions drawn from this validation hold across the ensemble. While the following discussion focusses on the differences between CRCM5 output and the observed climatology, the simulated climatology of the different variables and domains can be found in Supplementary Figures S1 to S4.

Figure 4 presents the seasonal mean surface-air temperature averaged over the 1980-2012 period from the E-OBS observational gridded dataset (0.22° resolution; Haylock et al. 2008) for the EU domain (first column), the difference between the ERA-Interim driven CRCM5 and E-OBS (second column) and the difference between the CanESM2 driven CRCM5 and E-OBS (third column). All data are linearly interpolated onto the CRCM5 grid for comparison purpose. It can first be seen that CRCM5 bias depends on geographical location and season, but systematic warm biases (especially in winter) appear over mountainous regions such as the Alps, Pyrenees, Balkans and the Carpathians (see also the CRCM5 topography in Figure 2). During winter, the reanalysis driven run (second column) shows a systematic cold bias larger than -1°C over most regions and exceeding -3°C in central Europe, while for the CanESM2 driven run, the bias is not systematically negative (generally between -1°C and 1°C). The fact that the CRCM5 winter bias is larger when driven by ERA-Interim may appear counterintuitive, as a reanalysis is expected to provide a better representation of the observed climate than a GCM. While the cold bias is likely partly attributable to the CRCM5 itself, the improvement observed when the CRCM5 is driven by the GCM may be due to some sort of bias cancellations between these two models. For the other seasons, biases are relatively insensitive to the nature of the driving data, although as expected, the CanESM2 driven run always shows a slightly higher RMSD than the ERA-Interim driven run. The generalized cold bias also appears during fall and spring, although with about half of the magnitude of the winter bias obtained from the CanESM2 driven run. During summer, a warm bias exceeding 2°C is observed for the eastern part of the domain.

Figure 5 shows corresponding results for precipitation over the EU domain. Throughout the year, there is a wet bias appearing over most parts of Europe. During winter, the bias is relatively large for the CanESM2 driven run, exceeding 3 mm/day in western Europe. In comparison, bias from the ERA-Interim run are generally smaller than 2 mm/day over the same region. The wet biases during spring and fall are as well less important for the ERA-Interim driven run. The CanESM2 driven run shows a marked dry bias exceeding -1 mm/day in the eastern part of the domain during summer.

Figure 6 presents the CRCM5 evaluation for surface-air temperature over the NNA domain using the Climatic Research Unit dataset (CRU; 0.5° resolution; Harris et al. 2013). The bias obtained for the ERA-Interim driven run generally ranges between -2°C and 2°C. RMSD values are approximately two times larger for the CanESM2 driven run than for ERA-Interim driven run. This is especially due to the important warm bias detected over most parts of the domain throughout the year for the CanESM2 driven run, which exceeds 4°C in western regions during summer and in the central part of the domain during winter. The cold bias occurring during winter and spring over northern Québec persists independently of the lateral boundary conditions, which suggests that the bias may originate from the CRCM5 itself.

Figure 7 shows the same analysis for precipitation over the NNA domain. A systematic wet bias around 1-2 mm/day exists for most parts of the domain and through the year for the ERA-Interim driven run. Biases are quite similar to those detected from the CanESM2 driven run for winter and spring, but for summer and fall, the CanESM2 driven run is characterized by a dry bias in the western (-2 mm/day) and southern (-1 mm/day) parts of the domain respectively.

Finally, to place these results into a more general context, it is worth recalling that the performance of the CRCM5 in terms of reproducing the current climate when driven by the ERA-Interim reanalysis is comparable to other state-of-the-arts RCMs over Europe and North America (Kotlarski et al. 2014; Martynov et al. 2013; Diaconescu et al. 2016).

c. Projected changes in climatological means

Figure 8 presents the short-term projected changes (2020-2039 versus 2000-2019) in precipitation for December estimated from ensemble members 1 to 24 over both domains. Reminding that the ensemble members differ only by slight random perturbations in their initial conditions, these results clearly show how natural variability can lead to very different projections. Some regions with strong precipitation changes may even show opposite signs for different members (e.g., members 4 and 6 over both domains). This also demonstrates how the practical use of single-member ensembles of regional climate projections may lead to misleading recommendations for planning short-term adaptation strategies to climate change. To focus on climate-change features that are robust across the ensemble, the ensemble mean signal is analyzed in the following. The statistical significance of the signal will be quantified by applying a Student's t test on the difference between future and historical ensemble-mean climates, and the dependence of this measure to the time horizon and the ensemble size will be assessed.

Ensemble mean climate-change signal between the 2000-2019 and 2080-2099 periods for the monthly mean surface-air temperature over the EU domain is first analyzed (Figure 9). The signal is stronger from June to September, with August showing temperature increases exceeding 8°C in western and southeastern Europe. There is also an enhanced warming in the northeastern part of the domain during winter, partly attributable to the decreasing snow cover-albedo feedback (Fischer et al. 2010).

Figure 10 shows the ensemble mean climate-change signal for monthly mean precipitation over the EU domain (2080-2099 versus 2000-2019). These simulations show that the climate in Europe will become dryer in summer and wetter in winter. Precipitation increase in December is as large as 2 mm/day on the west side of the Alps and along the west coast of the Balkan Peninsula. A large decrease of 2 mm/day in summer precipitation is detected during July and August on both the north and south sides of the Alps. However, the projected changes in precipitation are not significant everywhere, even for such a far horizon, as can be seen from the hatched regions, where the signal is not statistically significant. Notably, precipitation changes in winter over the Mediterranean Sea and the Iberian Peninsula are too weak to emerge from the noise of natural climate variability.

In order to investigate the relative contribution of natural variability and climate-change signal, changes in temperature and precipitation over different future periods were estimated and compared to the ensemble mean of the 50 members and to the first five members ensemble mean. Figures 11a, b and c show the 50-member ensemble mean temperature change (for December only) for three different time horizons; 2020-2039 (short term), 2040-2059 (mid-term) and 2080-2099 (long term; as in Figure 9) respectively. Similarly, Figure 11d, e, and f show the five-member ensemble mean temperature over the same three future periods. The five-member mean results are very similar to those of the full ensemble and the signal remains statistically significant everywhere in the domain for both mid-term (2040-2059) and long-term (2080-2099) projections. However, when considering short-term projections (2020-2039), the 50-member ensemble still shows statistically significant changes (Figure 11a), while the signal has not emerged from natural variability over most land areas for the five-member ensemble (Figure 11d). Similar conclusions hold for other months (see Supplementary Figures S5, S9 and S10).

Comparing the 50- and five-member ensemble mean precipitation change for July (Figure 11)g to l), the general features seen for the 50-member ensemble are still present for the five-member ensemble. Particularly, for long-term projections, the decrease in precipitation is statistically significant, although the intensity of the change is greater for this particular five-member ensemble. For short-term projections (2020-2039), the 50-member ensemble allows to detect small significant decreases in precipitation for western and southwestern Europe (Figure 11g), while the five-member ensemble mean displays practically no region with statistical significance changes in the short term, and very few statistically significant areas in the mid-term (Figure 11j). It is interesting to note that larger parts of the domain with statistically significant changes for the short-term period are reported for the 50-member ensemble than for the mid-term period for the five-member ensemble. These conclusions generally hold for the other months (see also Supplementary Figures S6, S11 and S12), and in several cases, even the long-term projections show very low statistical significance for the five-member ensemble while the 50-member ensemble generally allows to detect a signal over an appreciable fraction of the domain.

Repeating the previous analysis for the NNA domain, the climate-change signal in 2080-2099 for the monthly mean temperature is shown in Figure 12 based on the 50-member ensemble. A prominent maximum increase of temperature appears over the Hudson Bay. It exceeds 14°C from January through March and attenuates in April. It is worth noting that this regional feature is mostly inherited from the CanESM2 driving model, because its sea-surface temperature and sea-ice values are prescribed to the CRCM5. The positive ice-albedo feedback occurs as Hudson Bay becomes partially covered, instead of completely covered, by sea ice during winter by the end of the 21st century in the CanESM2 simulations (not shown). The important temperature change in winter extends into northern Québec and is influenced by the feedback from Hudson Bay sea ice, and by snow-albedo feedback as snow cover decreases.

Figure 13 shows the projected changes in precipitation over the NNA domain. From November through May, precipitation increases over land regions (exceeding 0.8 mm/day in northern Québec), Hudson Bay and Atlantic Ocean. In June, precipitation decreases by more than 0.4 mm/day over most land regions with the exception of northern Québec, and this drying pattern slowly decays until August, when only a small drying area remains over Ontario. Over the Atlantic, minimal change is observed during December, while precipitation decreases slightly during April/May, to reach values exceeding -1.8 mm/day in July/August. The important decrease in summer precipitation occurs in the area of the North Atlantic storm track and might be related to the poleward shift of mid-latitude storm tracks (Woollings et al. 2012), as well as to the weakening of the North Atlantic Meridional Overturning Circulation (Brayshaw et al. 2009) in CanESM2 simulations.

As for the EU domain, reducing the ensemble from 50 to 5 members does not significantly modify the patterns in temperature change (Figures 14a to f, results shown for December only). Short-term projections are also statistically significant for the 50-member ensemble (Figure 14a) while for the five-member ensemble (Figure14d) the southern half of the domain shows practically no statistically significant change during winter. Similar conclusions are obtained for the other months, namely that statistically significant changes are observed everywhere with the exception of some regions in the short-term projection for the five-member ensemble (see also Supplementary Figures S7, S13 and S14).

Comparing the 50-member ensemble with a five-member ensemble for precipitation over the NNA domain (for July only), Figures 14j to l show that the fraction of the domain with statistically significant changes is very small for the five-member ensemble. For short-term projections, however, the 50-member ensemble (Figure 14g) already shows a significant, though small, decrease in precipitation in the western part of the domain, which progressively extends in size for the mid-term and long term projections. Similar results are obtained for the other months, that is, no statistically significant changes over the largest fraction of the domain for the five-member ensemble, even in long-term projections are observed, while the 50-member ensemble generally allows to detect such changes (see also Supplementary Figures S8, S15 and S16). But it is also important to note that precipitation change remains a challenging variable even with the full ensemble, as the signal is generally weak while the variability is high.

d. Projected changes in temperature interannual variability

Here the large ensemble is used to assess the effect of climate change on temperature interannual variability, which can be defined as follows. Given a time window extending from year a to b inclusively, the overall variance calculated over this period of $P = b - a + 1$ years at a given gridpoint can be written as

$$\sigma_{a,b}^2 = \frac{1}{P(N-1)} \sum_{t=a}^{b} \sum_{i=1}^{N} (X_{it} - \bar{X}_{ot})^2, \tag{1}$$

where N is the ensemble size ($N = 50$), X_{it} the monthly mean temperature over the given time period for member i and year t, and $\bar{X}_{ot}$ the ensemble mean (average over all members) at year t. Assuming ergodicity between temporal and inter-member variances (Nikiéma et al. 2017), $\sigma_{a,b}$ (i.e. the square root of equation 1) can be interpreted as an estimate of the interannual variability for this specific time period. In the case of a climate system under transient forcing, the use of equation 1 to assess temporal variability using the inter-member spread involves weaker assumptions than calculating the residual temporal variability from detrended time series. The latter approach is nevertheless popular when assessing natural variability using small ensembles (Hawkins and Sutton 2009, 2011; Leduc et al. 2016a,b; Räisänen 2002).

Figure 15 shows the monthly patterns of interannual variability of surface-air temperature calculated over the 2000-2019 period for the EU domain. These patterns show a marked annual cycle reaching a maximum of around 4°C during winter in the northern regions, while the variability generally remains below 2.5°C for the rest of the year. The relative changes in interannual variability from 2000-2019 to 2080-2099 are presented in Figure 16, where the statistical significance is assessed using the F-test with a 99% confidence level. A large increase in interannual variability occurs from May through September over most of western and central Europe, and extending into the Scandinavian Peninsula. The maximum change is reached in August, when interannual variability increases by more than 70% (approximately 1°C), compared to the 2000-2019 period for which interannual variability is around 1.5°C (Figure 15). In addition to the mean surface-air temperature increase of around 7°C over this area and month in 2080-2099 (Figure 9), this highlights the importance of considering the effect of climate change on both mean climate and interannual variability when investigating the effect of climate change on heat waves, for instance (Schär et al. 2004).

The important projected decrease in mean precipitation during summer (see Figure 10) leads to a decrease in soil-moisture content (not shown) over a large part of Europe. The heat capacity of the land surface thus decreases, strengthening land-atmosphere coupling. As described in Seneviratne et al. (2006), the enhancement of the land-atmosphere coupling over Europe is an important contributor to the projected increase in temperature interannual variability. For instance, the surface-air temperature becomes more strongly influenced by variations in incident solar radiation, which is converted into sensible rather than latent heat flux (Brown et al. 2017). This suggests that local temperature variability could highly depend on geophysical characteristics in this case. It is also worth noting that the increase in summer temperature interannual variability is known to relate to both land-atmosphere interactions and projected changes in global atmospheric circulation patterns (e.g., Meehl and Tebaldi 2004).

For the rest of the year (i.e. October through April), Figure 16 shows that interannual variability tends to decrease throughout the 21st century. Several physical mechanisms support this result. Sea-ice retreat in the North Atlantic plays a role as westerly circulation becomes less affected by sea-ice albedo variability, but also as the atmosphere is no more isolated from the ocean which has a much greater heat capacity (Stouffer and Wetherald 2007). As another key physical mechanism that could explain this decreasing variability, it is known that sub-seasonal temperature variability is strongly affected by Arctic amplification. As shown by Screen (2014), rapid warming in the Arctic translates into a warming of cold air advected by northerly winds, which decreases sub-seasonal variability of surface-air temperature.

Figure 17 shows the annual cycle of interannual variability over the NNA domain for the period 2000-2019. Variability is much larger during the cold season in the northern part of the domain, which is in general agreement with observations (see Figure 1 in de Elía et al. 2013). From January through March, interannual variability exceeds 3°C for Hudson Bay and most of Québec. High values persist into April and May in a narrow region of maximum temperature variability that extends from the south shore of Hudson Bay and across Québec. It is worth noting that these regions are also characterized by a high level of interannual variability in snow-cover fraction (not shown). This corresponds with the transition zone separating permanent snow cover in the north and rare spring snow in the lower latitudes (Krasting et al. 2013). This link between high temperature variability and the edges of snow-covered regions is consistent with the results of Fischer et al. (2010), and as well as with Lehner et al. (2017) who showed the evidence of an existing thermodynamical link between snow cover and surface air temperature variability.

Figure 18 shows changes in monthly mean temperature interannual variability over the NNA domain from 2000-2019 to 2080-2099. There is a systematic decrease in interannual variability during winter over a dominant fraction of the domain and an increase during summer for the southern regions. This is in agreement with the relationship between temperature variability and thermal advection (Holmes et al. 2016), based on the fact that land-sea temperature contrasts will tend to increase during summer and decrease during winter, while the temperature gradient from pole to equator decreases mostly during winter due to Arctic amplification.

The northernmost part of Québec experiences a 80% increase (corresponding to about 1°C) in interannual temperature variability in May. This can be partly explained by the northward migration of the snow transition zone, which is located in the northernmost part of Québec in 2080-2090 while being around 10° further south in the reference period. In other words, the snow cover in a specific year may completely disappear in May in the northernmost region for some ensemble members while persisting in others. So interannual variability increases in a region when persistent snow cover transforms into a new transition region (northernmost region of Québec), while inversely, a transition region that becomes permanently without snow will rather experience a decrease in interannual variability. This may also explains the narrow east-west band in northern Québec where variability decreases by 30% during May.

While a rich literature describes the physical mechanisms underlying changes in temperature variability, the patterns of these changes are often difficult to assess with a high degree of confidence when using smaller ensembles. Similarly to what was done in Section c, it can be shown that using only the first five members of the ensemble leads to much less regions where changes in temperature interannual variability are statistically significant. Nevertheless, it is worth noting that some general features can still be detected with the smaller ensemble, such as the general decrease in variability over the northern regions during winter, or the increasing variability that is specific to central Europe during summer. More details about these results can be found in Supplementary Figures S17 and S18.

e. CRCM5-LE added value for extreme precipitations

A fundamental reason for producing large initial-condition ensembles is to obtain a satisfactory sampling of extreme events, these being poorly characterized in a single-member framework. In addition, it has been widely shown in the literature that RCMs have potential to add value compared to GCMs due to their higher spatial resolution, and especially over regions with specific heterogeneous features that can have an impact on surface forcings such as vegetation, lakes, orography, land-sea contrasts (e.g., Lucas-Picher et al. 2016; Prein et al. 2015; Di Luca et al. 2011; Kanamitsu and DeHaan 2011). To extend the concept of RCM added value to the case of large ensembles, the two large ensemble involved in the ClimEx project (CanESM-LE and CRCM5-LE) are compared in terms of the 20-year daily Annual Maximum Precipitation (AMP). For both ensembles, this climate extreme index was calculated by first extracting the daily annual maxima precipitation series at each grid point over the 2000-2019 period for each member (20 years x 50 members), from which the 95th percentile empirical quantile (20-year return level) was estimated.

Figure 19a and b show the daily AMP over the EU domain as calculated from CanESM2-LE and CRCM5-LE respectively. The largest fraction of grid points have daily AMP values ranging between 20-60 mm/day for CanESM2-LE while corresponding values for CRCM5-LE are mostly around 40-80 mm/day. In terms of the spatial distribution of the daily AMP, it is clear that the effect of orography on extreme precipitation patterns is more realistic for CRCM5-LE than CanESM2. Maximum values of about 60-80 mm/day occur over a few grid points in central Europe for CanESM2-LE, which correspond to the Alps region as seen from the CanESM2 topography (Figure 2d). Due to its coarse resolution, CanESM2 topography barely represents the Alps, as compared with CRCM5 topography (Figure 2b) where they are more realistically represented in terms of both height and spatial extent. This necessarily has an effect on the spatial structure of the AMP maximum over this region in CanESM2. For CRCM5-LE, coastal regions and areas with complex orography such as the southwest part of Scandinavian Mountains, the Atlantic coast of the Iberian Peninsula, the Alps and Dinaric Alps, the Pyrenees, are characterized by high precipitation extremes that are often around 120 mm/day, and even exceed 200 mm/day in some localized areas. Similar features were also detected from observations by Nikulin et al. (2011), although the reported AMP values were generally smaller.

For the NNA domain (Figure 19b), there is a north-south gradient from 30 mm/day in northern Québec to values around 100 mm/day in the southern part of the domain for CanESM2-LE. For CRCM5, this gradient ranges from about 40 mm/day in the north to about 160 mm/day in the south. This gradient, as well as the area of higher values detected along the east coast of United-States, is better represented in CRCM5-LE in terms of spatial distribution as compared with Gervais et al. (2014b,a) who have analyzed the 97th percentile of the observed daily precipitation.

As for the mean precipitation climatology (Section b), CRCM5-LE likely has some biases in extreme values. Nevertheless, this analysis shows that CRCM5-LE provides a much better representation of local extremes as compared with its driving model. In addition to its more detailed representation of surface forcings, a 12-km resolution model is generally more suitable for resolving extreme values at short time scales such as the daily AMP, as also shown by Innocenti et al. (2018).

4. Discussion and conclusions

The series of extreme flood events that occurred in Bavaria and Québec in recent decades has been of great concern to local governments, and has led to the development of the ClimEx project, which builds on the longstanding collaboration between Bavaria and Québec. The main goal of ClimEx is to help decision makers to implement robust climate-change adaptation strategies regarding flood risk, and more particularly, to better understand the role of natural climate variability and extreme meteorological events in the quantification of risk. This project is structured as a hydro-modelling chain: a Global Climate Model (GCM) large ensemble is dynamically downscaled with a Regional Climate Model (RCM), whose outputs will serve as input to hydrological model simulations over Bavaria and Québec. In this context, the current paper introduced the dynamical downscaling phase of ClimEx (i.e. the CRCM5 Large Ensemble) to the scientific community and was framed with the aim of facilitating the use of this unique dataset in future climate applications and research. The CRCM5 Large Ensemble (CRCM5-LE) consists in the dynamically downscaled version of the CanESM2 large initial-conditions ensemble from 2.8° ($\approx$ 310 km) to 0.11° ($\approx$ 12 km) resolution using the CRCM5 regional model over two regions of interest: Europe (EU) and northeastern North America (NNA).

In a preliminary analysis, the initial spin-up period of CanESM2-LE was analyzed in order to assess the time from which CRCM5-LE is driven by independent climate realizations, and therefore to ensure that the simulated natural variability can be assumed as physically consistent in future applications. For surface-air temperature, spin-up times of 100 and 800 days were found over land and ocean regions respectively, while for precipitation much shorter periods were found (25 and 150 days respectively). Therefore, an 800-day spin-up is the characteristic time after which the boundary conditions of CRCM5-LE can be assumed as independent realizations from

CanESM2, given the time scales of interest in the ClimEx project. In the light of these results, and since the CRCM5 also needs some time to become independent from its own initial conditions (not shown), it is reasonable to define the 1955-2099 period as the one where climate analysis could be performed.

A climatological validation of CRCM5-LE was performed for monthly mean surface-air temperature and precipitation. As for other climate models, CRCM5 reproduces the historical climate with biases that can be related to two main sources: the RCM model itself (e.g., domain configuration, spatial resolution, parameterization packages, land-surface scheme) and the nature of the boundary conditions (e.g., GCMs or reanalyses). For the analyzed variables, it was shown that biases of CanESM2 driven simulations are generally larger than those from the reanalysis-driven run, with the exception of a cold bias occurring during winter over Europe. These results suggest that a significant part of the total bias in CRCM5-LE may originate from both the CanESM2 and CRCM5 models. This climatological validation step should provide guidance to future users to select the most suitable bias-correction methods when using CRCM5-LE as an input for impact models (e.g., Muerth et al. 2013).

Climate-change projections of the monthly mean variables were next analyzed. The added-value of the large ensemble was investigated by comparing two ensemble sizes (5 vs 50 members) and three time horizons for the projections (short term 2020-2039, mid-term 2040-2059 and long-term 2080-2099 relative to 2000-2019) with regard to the spatial extent of the statistically significant climate-change signal. As expected, the highest extent of statistical significance was obtained using the full ensemble, and for long-term projections when the signal is large relative to the noise. While for temperature, a five-member ensemble was generally enough to detect short-term signals, for precipitation the 50-member short-term projection was often needed for long-term projection of the fraction of the domain with statistically significant signal. An interesting finding was that the 5-member ensemble displayed large scale patterns of the climate response often very similar to the 50-member ensemble, although the local climatic response –investigated through grid-point series– was generally not statistically significant. This suggests, as previously reported for instance by Deser et al. (2012), that natural variability plays a major role at local scales. Averaging over a larger ensemble improves our ability to detect local climatic response changes by ‘filtering out the local internal variability noise’, but it is worth noting that the actual future local response could be very different from the ensemble mean estimate because of internal variability.

Similarly, the projected changes in interannual variability of monthly mean surface-air temperature were investigated. Such analysis is possible when using a large ensemble while it remains very difficult to assess changes in interannual variability based on a single or few simulations. The patterns of change in temperature variability generally showed an increase during summer and a decrease during winter, which is in agreement with previous studies using GCM initital-conditions ensembles (e.g., Holmes et al. 2016). The current results however provided a more detailed characterization of temperature variability at the regional scale, as compared with the previous studies based on GCMs. A striking result is the dipole of decreasing/increasing variability that was found in northern part Québec during May, which was mostly attributable to the northward progression of the transition zone in the snow cover as the mean surface-air temperature increases.

Finally, the potential added-value of CRCM5-LE compared to CanESM2-LE was investigated by comparing 20-year daily AMP. While both ensembles allow empirical estimations of high AMP quantiles because to the large number of members –hence bypassing assumptions made in a parametric analysis–, CRCM5-LE allowed a much more realistic representation of important regional features regarding extreme precipitation over both domains, and especially over regions characterized by contrasting land-sea interfaces and complex topography such as in the southwest part of

Scandinavian, the Iberian Peninsula, the Alps and Dinaric Alps, the Pyrenees and along the east coast of United-States.

It is worth reminding that the CRCM5-LE framework does not address neither the model nor the scenario uncertainties, since it uses only one combination of global (CanESM2) and regional climate models (CRCM5), along with a single future pathway of GHGA emissions (RCP8.5). CRCM5-LE rather samples the internal variability of the CanESM2 model, which is downscaled at the regional scale using the CRCM5 that also adds its own internal variability (although generally smaller than that of a GCM). But despite not spanning the full range of uncertainty, the natural climate variability of this high-resolution regional climate system was assessed at a degree of detail never reached before.

In this context, an important strength of CRCM5-LE resides in short-term climate-change projections, which is supported by the previous conclusion that a large number of members is necessary to obtain statistically significant signals for short-term projections. This is also in agreement with Hawkins and Sutton (2009, 2011) who have shown that natural climate variability is a major contributor (especially for precipitation) to the total uncertainty of climate-change projections on short lead times at the regional scale. This important characteristic of single-model large ensembles should always be taken into account through the diversity of new applications that could emerge from CRCM5-LE, including the analysis of extreme compound events (e.g. heat waves, floods, droughts, forest fires), or the development of innovative techniques involving machine-learning algorithms to link meteorological patterns with high-impact events, among others. For long-term projections toward the end of the 21st century, CRCM5-LE results become increasingly dependent on the CRCM5 and CanESM2 models and the RCP8.5 scenario, which implies either to assume a storyline approach, or the include other models/ensembles in the analysis.

From this wider perspective, as more single-RCM large ensembles will become available in the future using other models and scenarios, inter-comparison of these datasets will be critical to better cope with the uncertainty related to future GHGA emissions, climate sensitivity (i.e. structural uncertainty) and natural variability within a common framework, at spatial and temporal scales suitable for climate-change impact applications. It is therefore necessary that future single-GCM large ensemble projects plan to provide all the necessary fields to drive RCMs. For instance, in the current experiment, CanESM2-LE was the only GCM allowing to drive an RCM with 50 continuous climate simulations from 1950 to 2099, whereas the CESM large ensemble (Kay et al. 2015) was also providing the necessary output but for a limited number of 10-year periods.

Acknowledgments. The authors would first like to thank Michel Valin (UQAM), Mourad Labassi (Ouranos) and Jens Weismüller (LRZ) for their technical support with the SuperMUC and Ouranos computational infrastructures; René Laprise (UQAM) for scientific discussions; and James Anstey (Environment and Climate Change Canada) for reviewing the first draft of the paper. We would also like to thank Laxmi Sushama and Katja Winger (UQAM) for supporting the development of the CRCM5. The ClimEx project is a new achievement for the Quebec-Bavaria International Collaboration on Climate Change (QBic3), which would have been impossible without the contribution of several important actors, notably Dieter Kranzlmüller (LRZ), Diane Chaumont (Ouranos) and Simon Ricard (Quebec Environment Ministry).

The ClimEx project is funded by the Bavarian State Ministry for the Environment and Consumer Protection. The CRCM5 is developed by the ESCER centre of Université du Québec à Montréal (UQAM; `www.escer.uqam.ca`) in collaboration with Environment and Climate Change Canada. We acknowledge Environment and Climate Change Canada's Canadian Centre for Climate Modelling and Analysis for executing and making available the CanESM2 Large Ensemble simulations used in this study, and the Canadian Sea Ice and Snow Evolution Network for proposing the simulations. Computations with the CRCM5 for the ClimEx project were made on the SuperMUC supercomputer at Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities. The operation of this supercomputer is funded via the Gauss Centre for Supercomputing (GCS) by the German Federal Ministry of Education and Research and the Bavarian State Ministry of Education, Science and the Arts.

References

Aalbers, E. E., G. Lenderink, E. van Meijgaard, and B. J. J. M. van den Hurk, 2017: Local-Scale Changes in Mean and Heavy Precipitation in Western Europe, Climate Change Or Internal Variability? *Climate Dynamics*, doi:10.1007/s00382-017-3901-9.

Arora, V. K., and Coauthors, 2011: Carbon Emission Limits Required To Satisfy Future Representative Concentration Pathways of Greenhouse Gases. *Geophys. Res. Lett.*, **38 (5)**, doi: 10.1029/2010gl046270.

Brayshaw, D. J., T. Woollings, and M. Vellinga, 2009: Tropical and Extratropical Responses of the North Atlantic Atmospheric Circulation To a Sustained Weakening of the Moc. *Journal of Climate*, **22 (11)**, 3146–3155, doi:10.1175/2008jcli2594.1.

Brown, P. T., Y. Ming, W. Li, and S. A. Hill, 2017: Change in the Magnitude and Mechanisms of Global Temperature Variability With Warming. *Nature Climate Change*, doi:10.1038/nclimate3381.

Chen, J., F. P. Brissette, P. Liu, and J. Xia, 2017: Using Raw Regional Climate Model Outputs for Quantifying Climate Change Impacts on Hydrology. *Hydrological Processes*, **31 (24)**, 4398–4413, doi:10.1002/hyp.11368.

Clark, M. P., and Coauthors, 2016: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports*, **2 (2)**, 55–64, doi:10.1007/s40641-016-0034-x.

Collins, M., and Coauthors, 2013: *Long-term Climate Change: Projections, Commitments and Irreversibility*, book section 12, 1029–1136. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.024, URL www.climatechange2013.org.

Davies, H. C., 1976: A Lateral Boundary Formulation for Multi-Level Prediction Models. *Quarterly Journal of the Royal Meteorological Society*, **102 (432)**, 405–418, doi:10.1256/smsqj.43209.

de Elía, R., S. Biner, and A. Frigon, 2013: Interannual Variability and Expected Regional Climate Change Over North America. *Climate Dynamics*, **41 (5-6)**, 1245–1267, doi:10.1007/s00382-013-1717-9.

de Elía, R., R. Laprise, and B. Denis, 2002: Forecasting Skill Limits of Nested, Limited-Area Models: A Perfect-Model Approach. *Monthly Weather Review*, **130 (8)**, 2006–2023, doi:10.1175/1520-0493(2002)130$⟨$2006:fslonl⟩2.0.co;2.

Deser, C., R. Knutti, S. Solomon, and A. S. Phillips, 2012: Communication of the Role of Natural Variability in Future North American Climate. *Nature Climate Change*, **2 (11)**, 775–779, doi:10.1038/nclimate1562.

Deser, C., A. S. Phillips, M. A. Alexander, and B. V. Smoliak, 2014: Projecting North American Climate Over the Next 50 Years: Uncertainty Due To Internal Variability. *Journal of Climate*, **27 (6)**, 2271–2296, doi:10.1175/JCLI-D-13-00451.1.

Di Luca, A., R. de Elía, and R. Laprise, 2011: Potential for Added Value in Precipitation Simulated By High-Resolution Nested Regional Climate Models and Observations. *Climate Dynamics*, **38 (5-6)**, 1229–1247, doi:10.1007/s00382-011-1068-3.

Diaconescu, E. P., P. Gachon, R. Laprise, and J. F. Scinocca, 2016: Evaluation of Precipitation Indices Over North America From Various Configurations of Regional Climate Models. *Atmosphere-Ocean*, **54 (4)**, 418–439, doi:10.1080/07055900.2016.1185005.

Fatichi, S., S. Rimkus, P. Burlando, and R. Bordoy, 2014: Does Internal Climate Variability Overwhelm Climate Change Signals in Streamflow? The Upper Po and Rhone Basin Case Studies. *Science of The Total Environment*, **493**, 1171–1182, doi:10.1016/j.scitotenv.2013.12.014.

Fischer, E. M., D. M. Lawrence, and B. M. Sanderson, 2010: Quantifying Uncertainties in Projections of Extremes-A Perturbed Land Surface Parameter Experiment. *Climate Dynamics*, **37 (7-8)**, 1381–1398, doi:10.1007/s00382-010-0915-y.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking Climate Change Modelling To Impacts Studies: Recent Advances in Downscaling Techniques for Hydrological Modelling. *International Journal of Climatology*, **27 (12)**, 1547–1578, doi:10.1002/joc.1556.

Fyfe, J. C., and Coauthors, 2017: Large Near-Term Projected Snowpack Loss Over the Western United States. *Nature Communications*, **8**, 14 996, doi:10.1038/ncomms14996.

Gervais, M., J. R. Gyakum, E. Atallah, L. B. Tremblay, and R. B. Neale, 2014a: How Well Are the Distribution and Extreme Values of Daily Precipitation Over North America Represented in the Community Climate System Model? A Comparison To Reanalysis, Satellite, and Gridded Station Data. *Journal of Climate*, **27 (14)**, 5219–5239, doi:10.1175/jcli-d-13-00320.1.

Gervais, M., L. B. Tremblay, J. R. Gyakum, and E. Atallah, 2014b: Representing Extremes in a Daily Gridded Precipitation Analysis Over the United States: Impacts of Station Density, Resolution, and Gridding Methods. *Journal of Climate*, **27 (14)**, 5201–5218, doi:10.1175/jcli-d-13-00319.1.

Giorgi, F., 2002: Dependence of the Surface Climate Interannual Variability on Spatial Scale. *Geophysical Research Letters*, **29 (23)**, 16–1–16–4, doi:10.1029/2002gl016175.

Giorgi, F., and W. J. Gutowski, 2015: Regional Dynamical Downscaling and the Cordex Initiative. *Annual Review of Environment and Resources*, **40 (1)**, 467–490, doi:10.1146/annurev-environ-102014-021217.

Harris, I., P. Jones, T. Osborn, and D. Lister, 2013: Updated High-Resolution Grids of Monthly Climatic Observations - the Cru Ts3.10 Dataset. *International Journal of Climatology*, **34 (3)**, 623–642, doi:10.1002/joc.3711.

Hawkins, E., and R. Sutton, 2009: The Potential To Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, **90 (8)**, 1095–1107, doi:10.1175/2009bams2607.1.

Hawkins, E., and R. Sutton, 2011: The Potential To Narrow Uncertainty in Projections of Regional Precipitation Change. *Climate Dynamics*, **37 (1-2)**, 407–418, doi:10.1007/s00382-010-0810-6.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European Daily High-Resolution Gridded Data Set of Surface Temperature and Precipitation for 1950-2006. *Journal of Geophysical Research*, **113 (D20)**, D20 119, doi:10.1029/2008jd010201.

Holmes, C. R., T. Woollings, E. Hawkins, and H. de Vries, 2016: Robust Future Changes in Temperature Variability Under Greenhouse Gas Forcing and the Relationship With Thermal Advection. *Journal of Climate*, **29 (6)**, 2221–2236, doi:10.1175/jcli-d-14-00735.1.

Innocenti, S., A. Mailhot, A. Frigon, A. J. Cannon, and M. Leduc, 2018: Observed and simulated precipitation over North East North-America: how do sub-daily extremes scale in space and time? *Submitted to Journal of Climate, January 2018*.

Kanamitsu, M., and L. DeHaan, 2011: The Added Value Index: A New Metric To Quantify the Added Value of Regional Models. *Journal of Geophysical Research*, **116 (D11)**, D11 106, doi:10.1029/2011jd015597.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society*, **96 (8)**, 1333–1349, doi:10.1175/BAMS-D-13-00255.1.

Kotlarski, S., and Coauthors, 2014: Regional Climate Modeling on European Scales: a Joint Standard Evaluation of the Euro-Cordex RCM Ensemble. *Geoscientific Model Development*, **7 (4)**, 1297–1333, doi:10.5194/gmd-7-1297-2014.

Krasting, J. P., A. J. Broccoli, K. W. Dixon, and J. R. Lanzante, 2013: Future Changes in Northern Hemisphere Snowfall. *Journal of Climate*, **26 (20)**, 7813–7828, doi:10.1175/jcli-d-12-00832.1.

Leduc, M., and R. Laprise, 2009: Regional Climate Model Sensitivity Domain Size. *Clim. Dyn.*, **32 (6)**, 833–854, doi:10.1007/s00382-008-0400-z.

Leduc, M., R. Laprise, R. de Elía, and L. Separovic, 2016a: Is Institutional Democracy a Good Proxy for Model Independence? *Journal of Climate*, doi:10.1175/JCLI-D-15-0761.1.

Leduc, M., R. Laprise, M. Moretti-Poisson, and J.-P. Morin, 2011: Sensitivity To Domain Size of Mid-Latitude Summer Simulations With a Regional Climate Model. *Climate Dynamics*, **37 (1-2)**, 343–356, doi:10.1007/s00382-011-1008-2.

Leduc, M., H. D. Matthews, and R. de Elia, 2016b: Regional Estimates of the Transient Climate Response To Cumulative CO2 Emissions. *Nature Clim. Change*, **6 (5)**, 474–478, letter.

Lehner, F., C. Deser, and L. Terray, 2017: Toward a New Estimate of ”Time of Emergence” of Anthropogenic Warming: Insights From Dynamical Adjustment and a Large Initial-Condition Model Ensemble. *Journal of Climate*, **30 (19)**, 7739–7756, doi:10.1175/jcli-d-16-0792.1.

Liu, C., and Coauthors, 2016: Continental-Scale Convection-Permitting Modeling of the Current and Future Climate of North America. *Climate Dynamics*, **49 (1-2)**, 71–95, doi:10.1007/s00382-016-3327-9.

Lucas-Picher, P., R. Laprise, and K. Winger, 2016: Evidence of Added Value in North American Regional Climate Model Hindcast Simulations Using Ever-Increasing Horizontal Resolutions. *Clim Dyn*, doi:10.1007/s00382-016-3227-z.

Lucas-Picher, P., P. Riboust, S. Somot, and R. Laprise, 2015: Reconstruction of the Spring 2011 Richelieu River Flood By Two Regional Climate Models and a Hydrological Model. *J. Hydrometeor*, **16 (1)**, 36–54, doi:10.1175/jhm-d-14-0116.1.

Martynov, A., R. Laprise, L. Sushama, K. Winger, L. Separovic, and B. Dugas, 2013: Reanalysis-Driven Climate Simulation Over Cordex North America Domain Using the Canadian Regional Climate Model, Version 5: Model Performance Evaluation. *Clim Dyn*, **41 (11-12)**, 2973–3005, doi:10.1007/s00382-013-1778-9.

Matte, D., R. Laprise, J. M. Thériault, and P. Lucas-Picher, 2016: Spatial Spin-Up of Fine Scales in a Regional Climate Model Simulation Driven By Low-Resolution Boundary Conditions. *Climate Dynamics*, doi:10.1007/s00382-016-3358-2.

Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell, 2007: The WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bulletin of the American Meteorological Society*, **88 (9)**, 1383–1394, doi:doi:10.1175/BAMS-88-9-1383.

Meehl, G. A., and C. Tebaldi, 2004: More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century. *Science*, **305 (5686)**, 994–997, doi:10.1126/science.1098704.

Mizuta, R., and Coauthors, 2016: Over 5000 Years of Ensemble Future Climate Simulations By 60 Km Global and 20 Km Regional Atmospheric Models. *Bulletin of the American Meteorological Society*, BAMS–D–16–0099.1, doi:10.1175/bams-d-16-0099.1.

Mudryk, L. R., P. J. Kushner, and C. Derksen, 2013: Interpreting Observed Northern Hemisphere Snow Trends With Large Ensembles of Climate Simulations. *Clim Dyn*, **43 (1-2)**, 345–359, doi:10.1007/s00382-013-1954-y.

Muerth, M. J., and Coauthors, 2013: On the Need for Bias Correction in Regional Climate Scenarios To Assess Climate Change Impacts on River Runoff. *Hydrology and Earth System Sciences*, **17 (3)**, 1189–1204, doi:10.5194/hess-17-1189-2013.

Music, B., D. Caya, A. Frigon, A. Musy, R. Roy, and D. Rodenhuis, 2012: *Canadian Regional Climate Model as a Tool for Assessing Hydrological Impacts of Climate Change at the Watershed Scale*, 157–165. Climate Change, Springer Science + Business Media, doi:10.1007/978-3-7091-0973-1_12.

Nikiéma, O., R. Laprise, and B. Dugas, 2017: Energetics of Transient-Eddy and Inter-Member Variabilities in Global and Regional Climate Model Simulations. *Climate Dynamics*, doi:10.1007/s00382-017-3918-0.

Nikulin, G., E. Kjellström, U. Hansson, G. Strandberg, and A. Ullerstig, 2011: Evaluation and Future Projections of Temperature, Precipitation and Wind Extremes Over Europe in an Ensemble of Regional Climate Simulations. *Tellus A*, doi:10.3402/tellusa.v63i1.15764.

O'Neill, B. C., and Coauthors, 2016: The Scenario Model Intercomparison Project (ScenariOMIP) for CMIP6. *Geoscientific Model Development Discussions*, **2016**, 1–35, doi:10.5194/gmd-2016-84.

Prein, A., and Coauthors, 2015: Precipitation in the Euro-Cordex 0.11° and 0.44° Simulations: High Resolution, High Benefits? *Climate Dynamics*, 1–30, doi:10.1007/s00382-015-2589-y.

Räisänen, J., 2002: CO2-Induced Changes in Interannual Temperature and Precipitation Variability in 19 CMIP2 Experiments. *Journal of Climate*, **15 (17)**, 2395–2411, doi:10.1175/1520-0442(2002)015$⟨$2395:ciciit⟩2.0.co;2.

Riette, S., and D. Caya, 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. Tech. rep.

Sanderson, B. M., K. W. Oleson, W. G. Strand, F. Lehner, and B. C. O'Neill, 2015: A New Ensemble of Gcm Simulations To Assess Avoided Impacts in a Climate Mitigation Scenario. *Climatic Change*, doi:10.1007/s10584-015-1567-z.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The Role of Increasing Temperature Variability in European Summer Heatwaves. *Nature*, **427 (6972)**, 332–336, doi:10.1038/nature02300.

Schulz, K., and M. Bernhardt, 2016: The End of Trend Estimation for Extreme Floods Under Climate Change? *Hydrological Processes*, **30 (11)**, 1804–1808, doi:10.1002/hyp.10816.

Screen, J. A., 2014: Arctic Amplification Decreases Temperature Variance in Northern Mid- To High-Latitudes. *Nature Climate Change*, **4 (7)**, 577–582, doi:10.1038/nclimate2268.

Selten, F. M., G. W. Branstator, H. A. Dijkstra, and M. Kliphuis, 2004: Tropical Origins for Recent and Future Northern Hemisphere Climate Change. *Geophys. Res. Lett.*, **31 (21)**, doi:10.1029/2004gl020739.

Seneviratne, S. I., D. Lüthi, M. Litschi, and C. Schär, 2006: Land-Atmosphere Coupling and Climate Change in Europe. *Nature*, **443 (7108)**, 205–209, doi:10.1038/nature05095.

Separovic, L., A. Alexandru, R. Laprise, A. Martynov, L. Sushama, K. Winger, K. Tete, and M. Valin, 2013: Present Climate and Climate Change Over North America As Simulated By the Fifth-Generation Canadian Regional Climate Model. *Clim Dyn*, **41 (11-12)**, 3167–3201, doi:10.1007/s00382-013-1737-5.

Separovic, L., R. de Elía, and R. Laprise, 2012: Impact of Spectral Nudging and Domain Size in Studies of Rcm Response To Parameter Modification. *Climate Dynamics*, **38 (7-8)**, 1325–1343, doi:10.1007/s00382-011-1072-7.

Sigmond, M., J. C. Fyfe, and N. C. Swart, 2018: Ice-free arctic projections under the paris agreement. *Nature Climate Change*, doi:10.1038/s41558-018-0124-y.

Sterl, A., and Coauthors, 2008: When Can We Expect Extremely High Surface Temperatures? *Geophys. Res. Lett.*, **35 (14)**, L14 703, doi:10.1029/2008gl034071.

Stouffer, R. J., 2004: Time Scales of Climate Response. *Journal of Climate*, **17 (1)**, 209–217, doi:10.1175/1520-0442(2004)017$⟨$0209:tsocr⟩2.0.co;2.

Stouffer, R. J., and R. T. Wetherald, 2007: Changes of Variability in Response To Increasing Greenhouse Gases. Part I: Temperature. *Journal of Climate*, **20 (21)**, 5455–5467, doi:10.1175/2007jcli1384.1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93 (4)**, 485–498, doi:10.1175/bams-d-11-00094.1.

Wigley, T. M. L., P. D. Jones, K. R. Briffa, and G. Smith, 1990: Obtaining Sub-Grid-Scale Information From Coarse-Resolution General Circulation Model Output. *Journal of Geophysical Research*, **95 (D2)**, 1943, doi:10.1029/jd095id02p01943.

Woollings, T., J. M. Gregory, J. G. Pinto, M. Reyers, and D. J. Brayshaw, 2012: Response of the North Atlantic Storm Track To Climate Change Shaped By Ocean-Atmosphere Coupling. *Nature Geoscience*, **5 (5)**, 313–317, doi:10.1038/ngeo1438.

Xu, C.-y., 1999: From GCMs To River Flow: a Review of Downscaling Methods and Hydrologic Modelling Approaches. *Progress in Physical Geography*, **23 (2)**, 229–249, doi:10.1177/030913339902300204.

LIST OF FIGURES

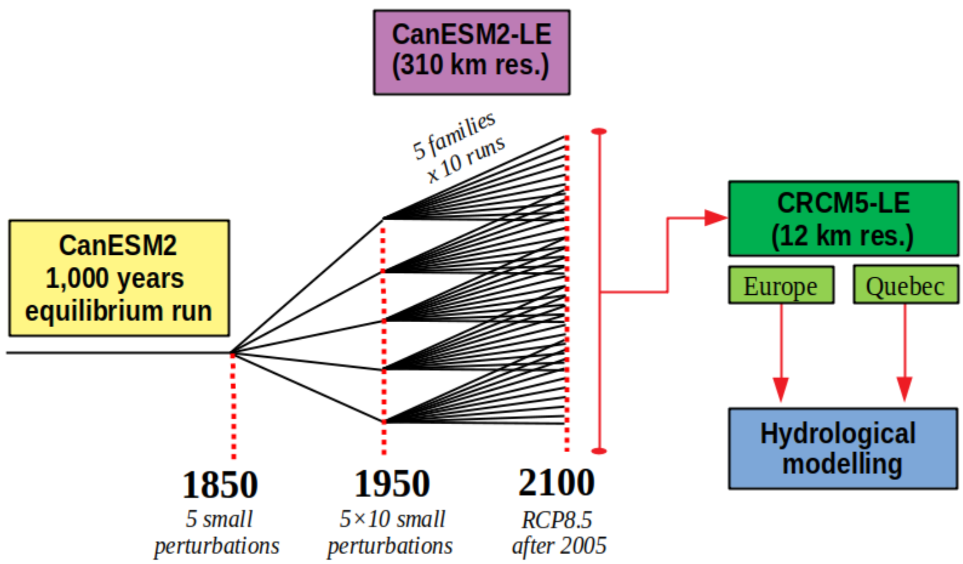
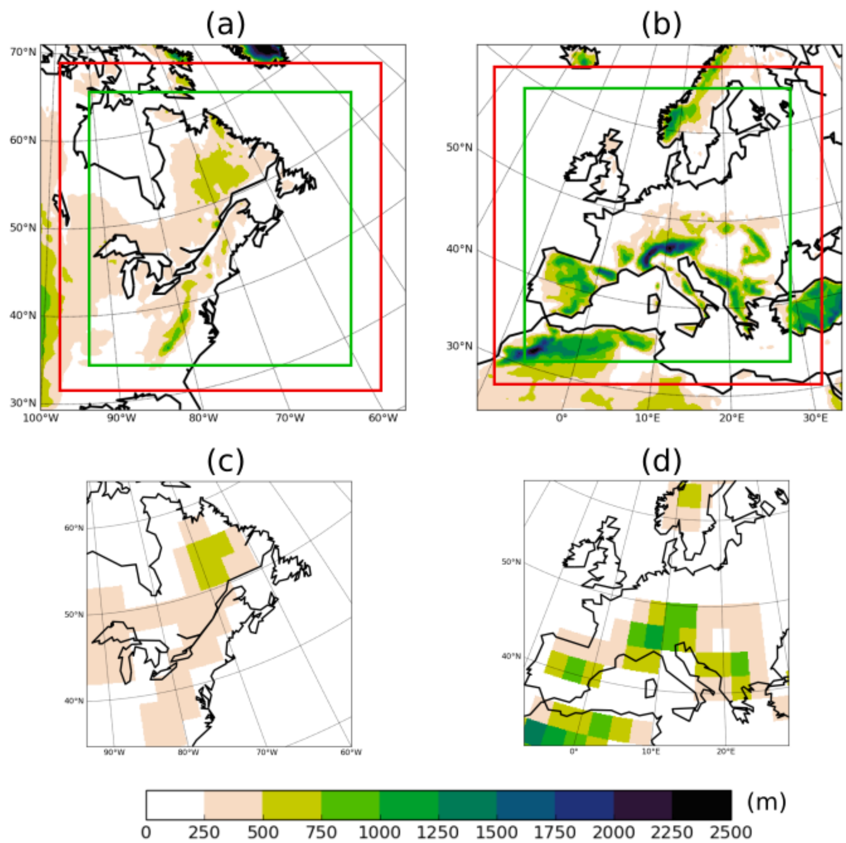


FIG. 1. Schematic representation of the ClimEx modelling chain where the CanESM2 members are used to drive the CRCM5 and hydrological models.



961 FIG. 2. Topography used by CRCM5 to produce the ClimEx large ensemble over the (a) northeastern North America (NNA; left panel) and (b) Europe (EU; right panel) domains. Integration domain is shown in each case (380x380 grid points; full map), as well as the "free domain" (340x340 grid points; in red) where the model is technically free from direct imposition of lateral boundary conditions, and the "analysis domain" (280x280; in green), that is the region where output fields were archived; (c) and (d) Topography used by CanESM2 as seen from the perspective of the NNA and EU analysis domains respectively.

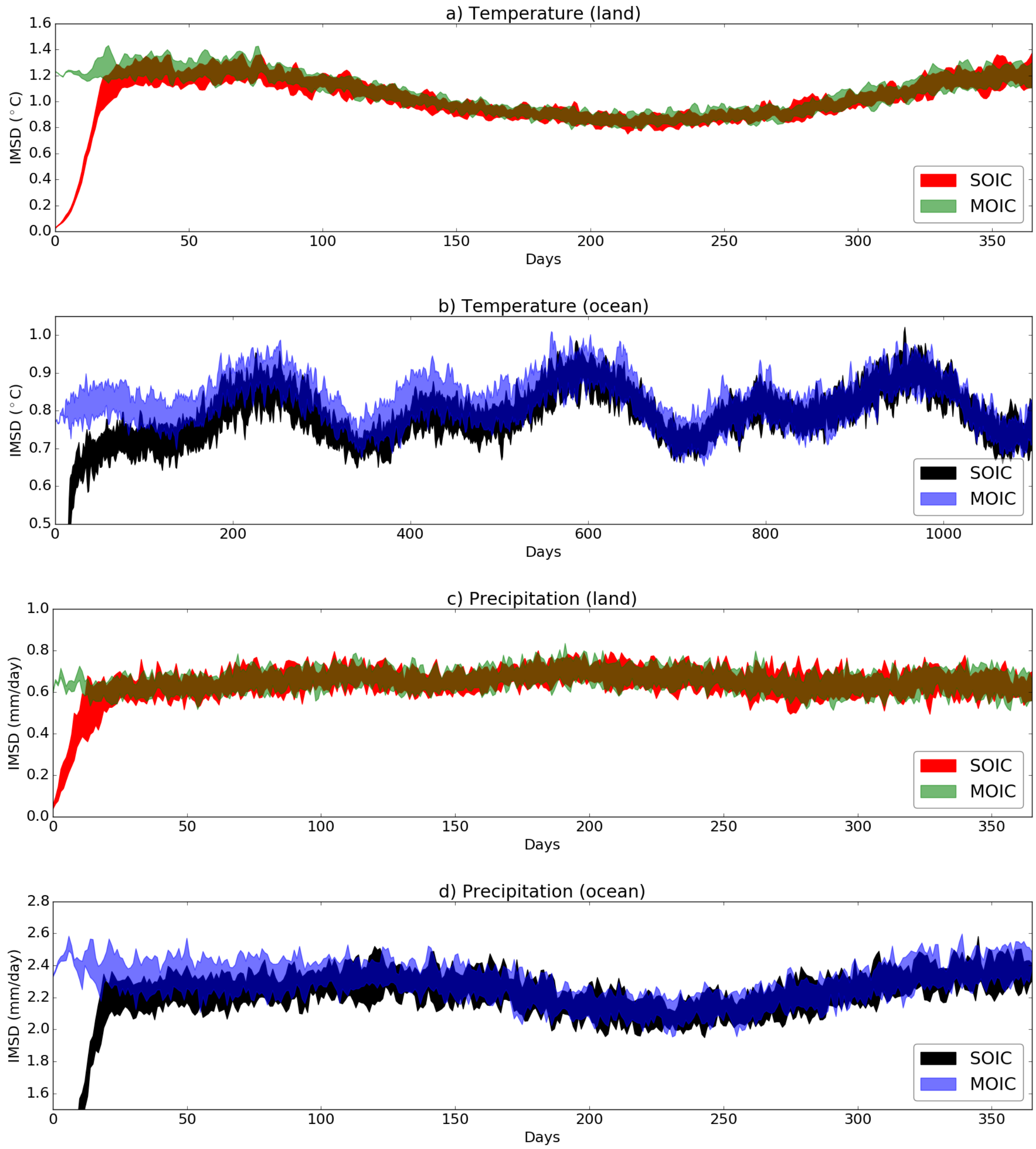


FIG. 3. Spatially averaged inter-member standard deviation (IMSD) as a function of time (in days) from the beginning of the simulations (starting on 1 January 1950) for the CanESM2 large ensemble. The IMSD was calculated for twenty groups of 5 runs, and these groups divide in two categories: 1) same ocean initial conditions (SOIC) and 2) mixed ocean initial conditions (MOIC). The SOIC groups (red and black curves) correspond to members 1 to 5, 6 to 10, and so on up to members 45 to 50 (see families in Figure 1. The ten MOIC groups (green and blue curves) correspond to members (1,11,21,31,41) up to (10,20,30,40,50). In panels (a) and (b), the IMSD calculated for surface-air temperature was averaged over land and ocean grid points respectively while (c) and (d) present the corresponding results for precipitation. Note the different scales for axes.

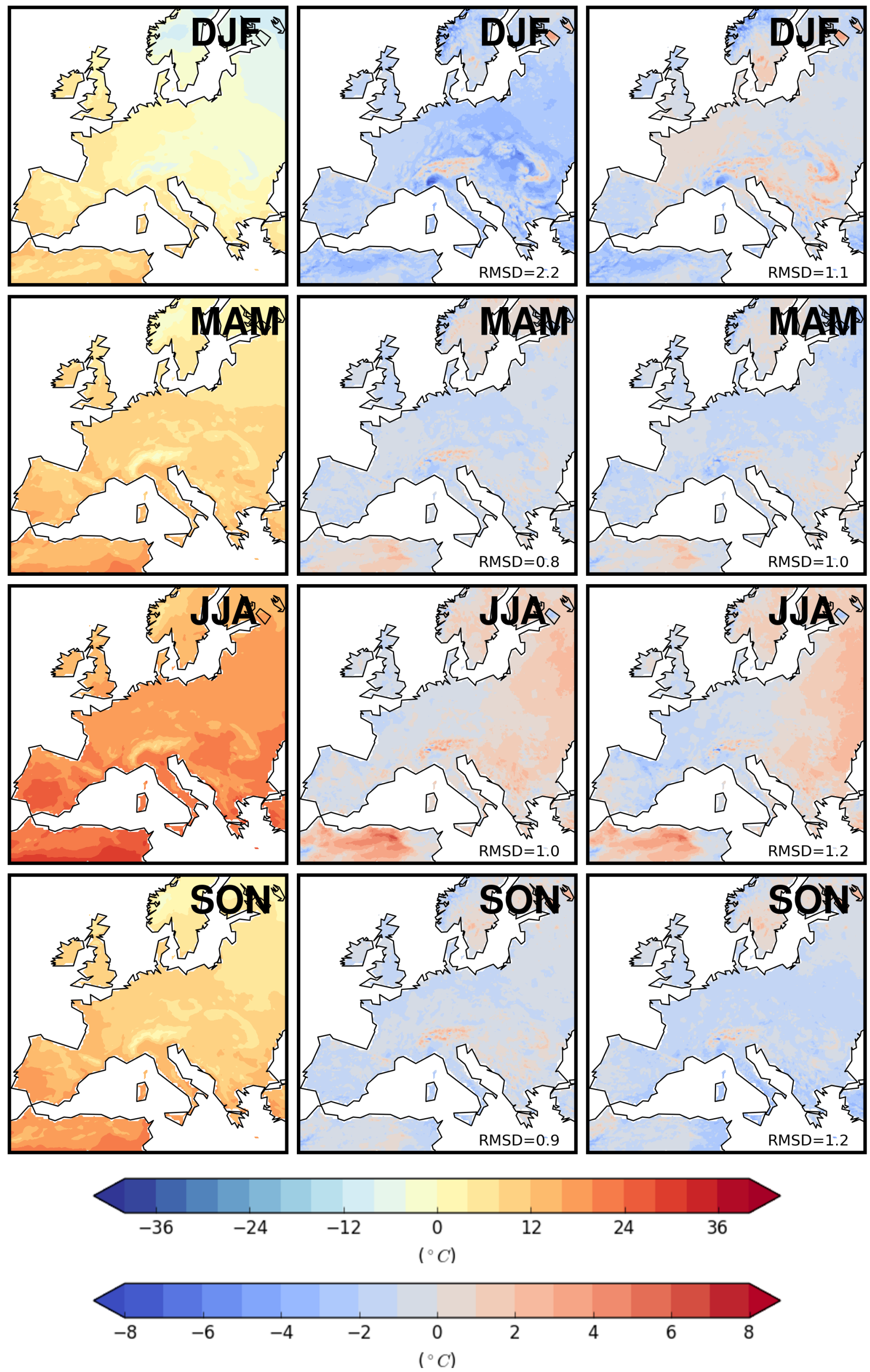


FIG. 4. Mean seasonal surface-air temperature over the 1980-2012 period for E-OBS (first column), differences with the CRCM5 run driven by ERA-Interim (second column), and with the first member of the CanESM2-LE (third column) for the EU domain. A positive difference corresponds to an overestimation of the observed values by the simulations. The top color bar applies to the first column while the bottom one applies to second and third columns. The root mean square differences (RMSD) over land grid-points of the domain is indicated at the lower-right corner of each panel in central and right columns.

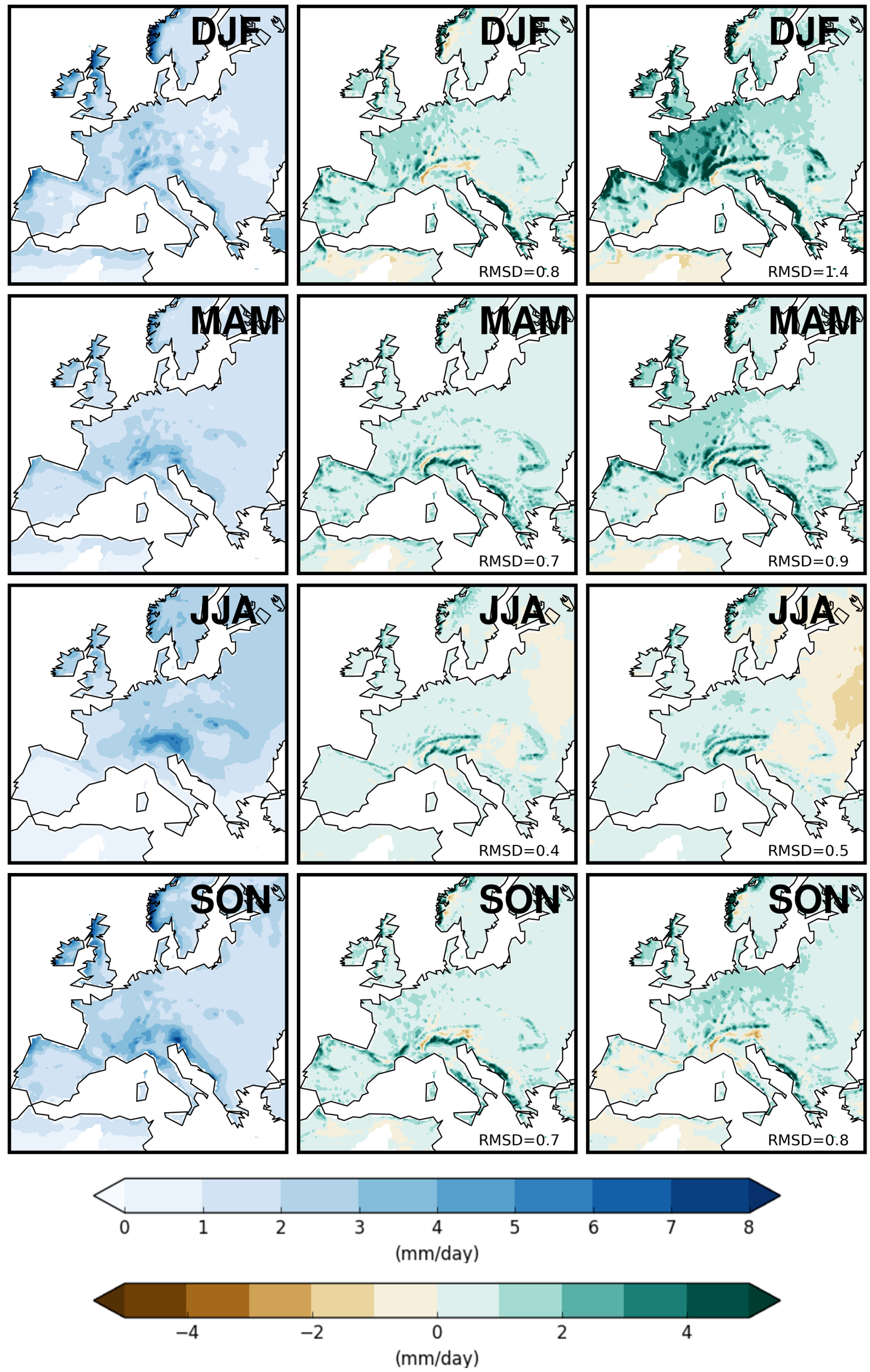


FIG. 5. Mean seasonal precipitation over the 1980-2012 period for E-OBS (first column) and its difference from the CRCM5 run driven by ERA-Interim (second column) and from the first member of CanESM2 (third column) for the EU domain. A positive difference corresponds to an overestimation of the observed values. The top colorbar applies to the first column. The root mean square difference (RMSD) is provided for each difference.

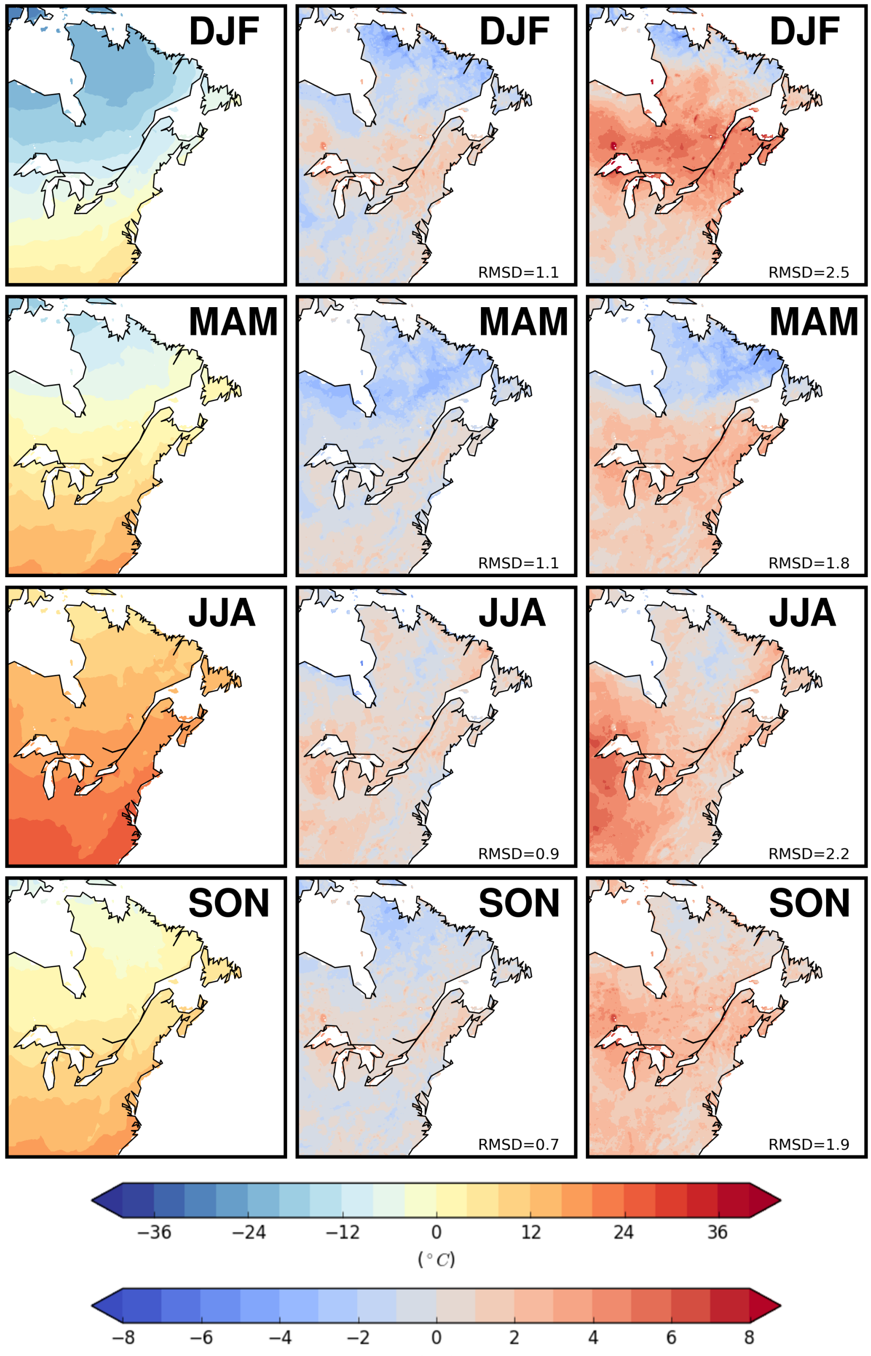
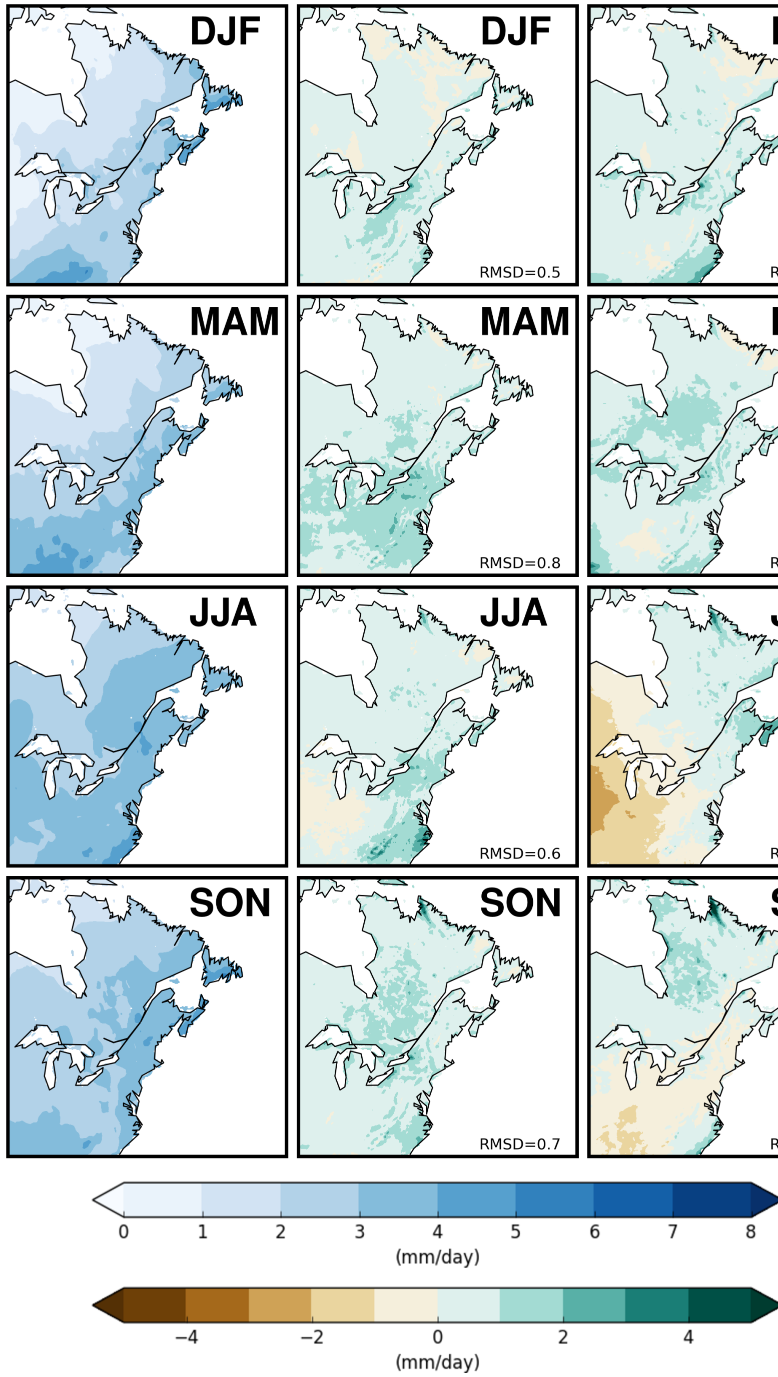
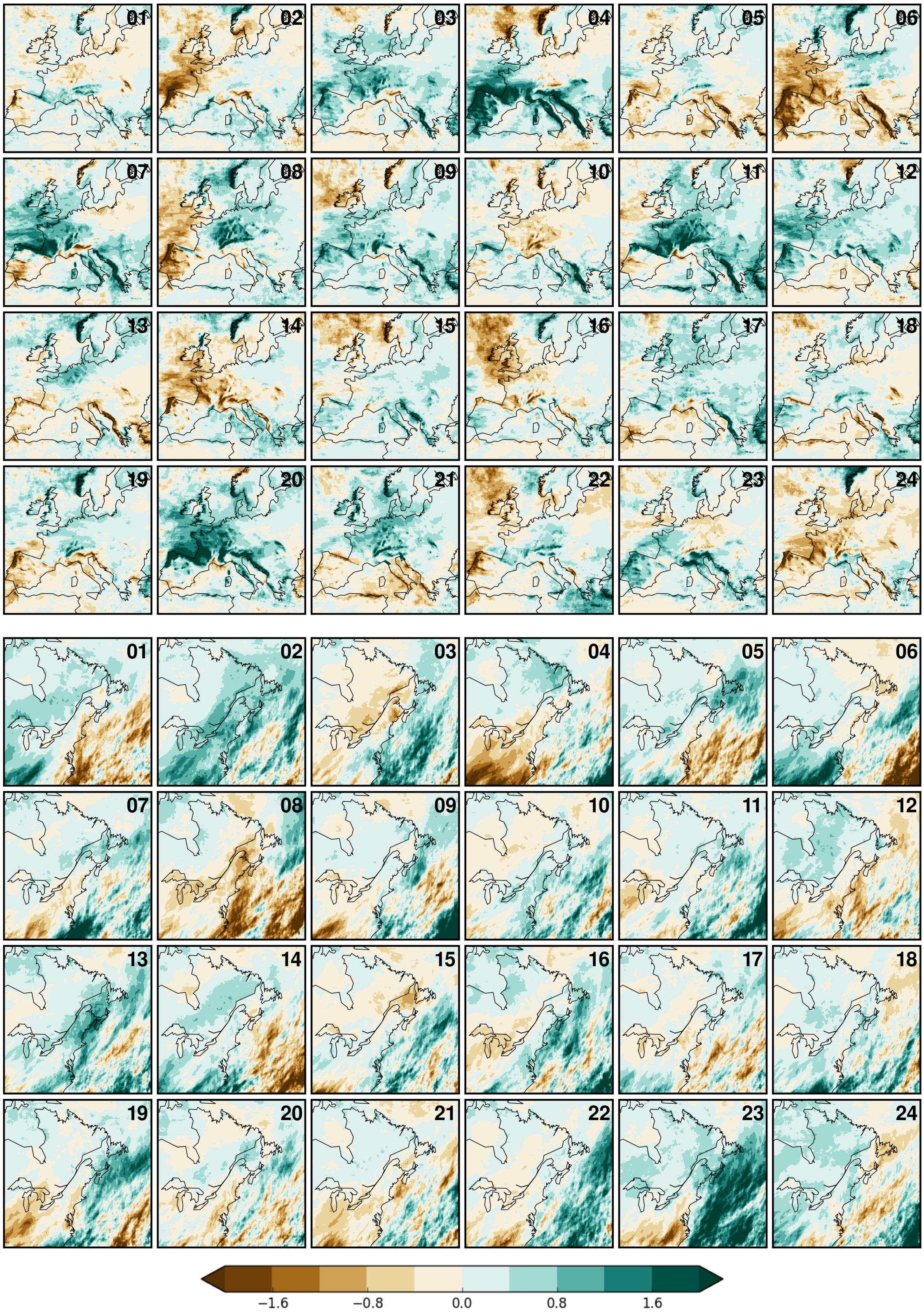


FIG. 6. Mean seasonal surface-air temperature over the 1980-2012 period for CRU (first column) and its difference from the CRCM5 run driven by ERA-Interim (second column) and from the first member of CanESM2 (third column) for the NNA domain. A positive difference corresponds to an overestimation of the observed values. The top colorbar applies to the first column. The root mean square difference (RMSD) is provided for each difference.

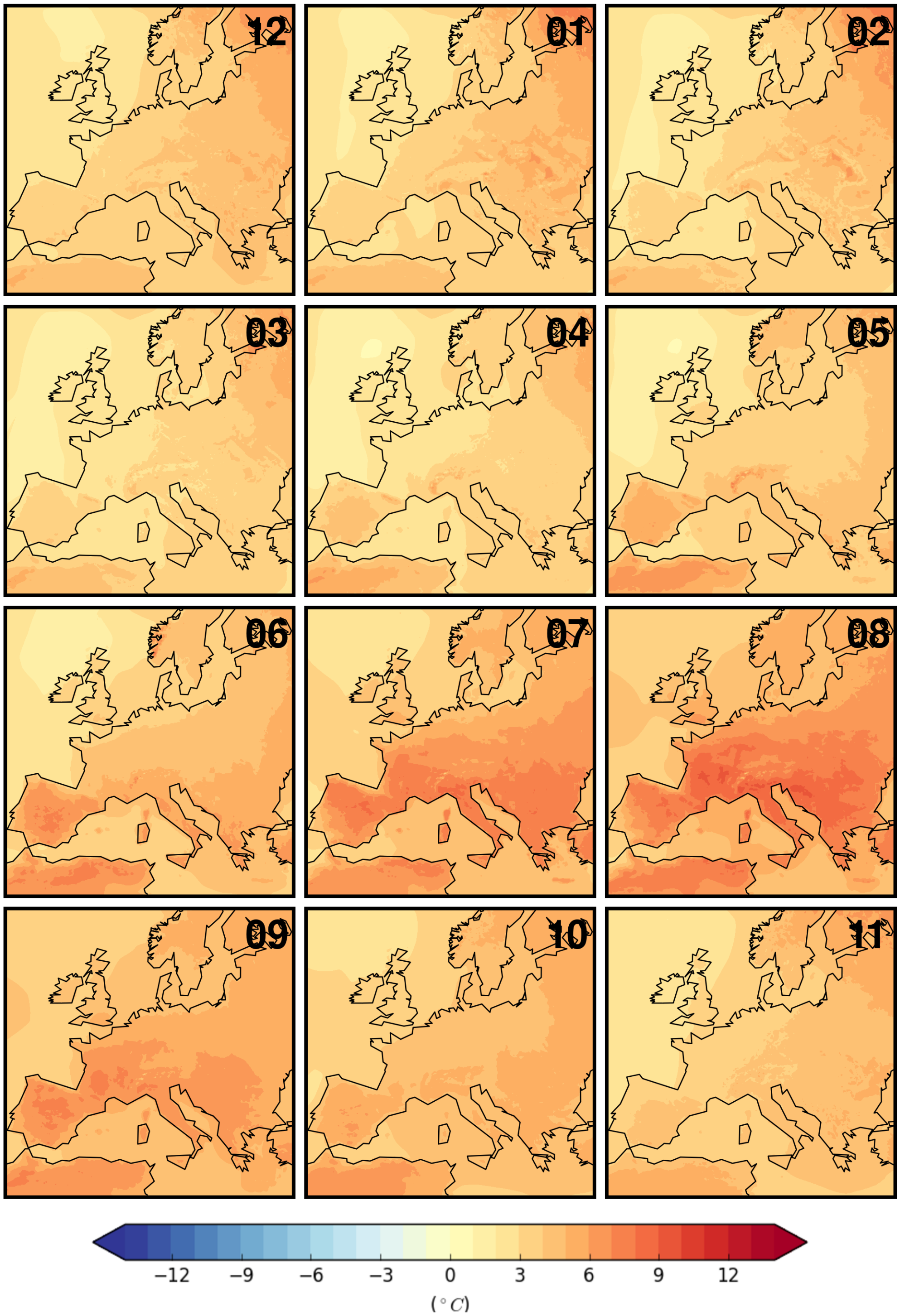


991 FIG. 7. Mean seasonal precipitation over the 1980-2012 period for CRU (first column) and its difference from
992 the CRCM5 run driven by ERA-Interim (second column) and from the first member of CanESM2 (third column)
993 for the NNA domain. A positive difference corresponds to an overestimation of the observed values. The top
994 colorbar applies to the first column. The root mean square difference (RMSD) is provided for each difference.



995 FIG. 8. Short-term climate-change projections (2020-2039 vs 2000-2019) for mean December precipitation

996 from the ensemble members 1 to 24 over the EU (top rows) and NNA (bottom rows) domains.



997 FIG. 9. The CRCM5 50-member ensemble mean climate-change signal for surface-air temperature computed
998 as the difference between the 2080-2099 and 2000-2019 monthly climate means for the EU domain. All reported
999 changes are statistically significant at the 99% confidence level (Student's t test with unequal variances). Months
1000 are labeled from 1 to 12.

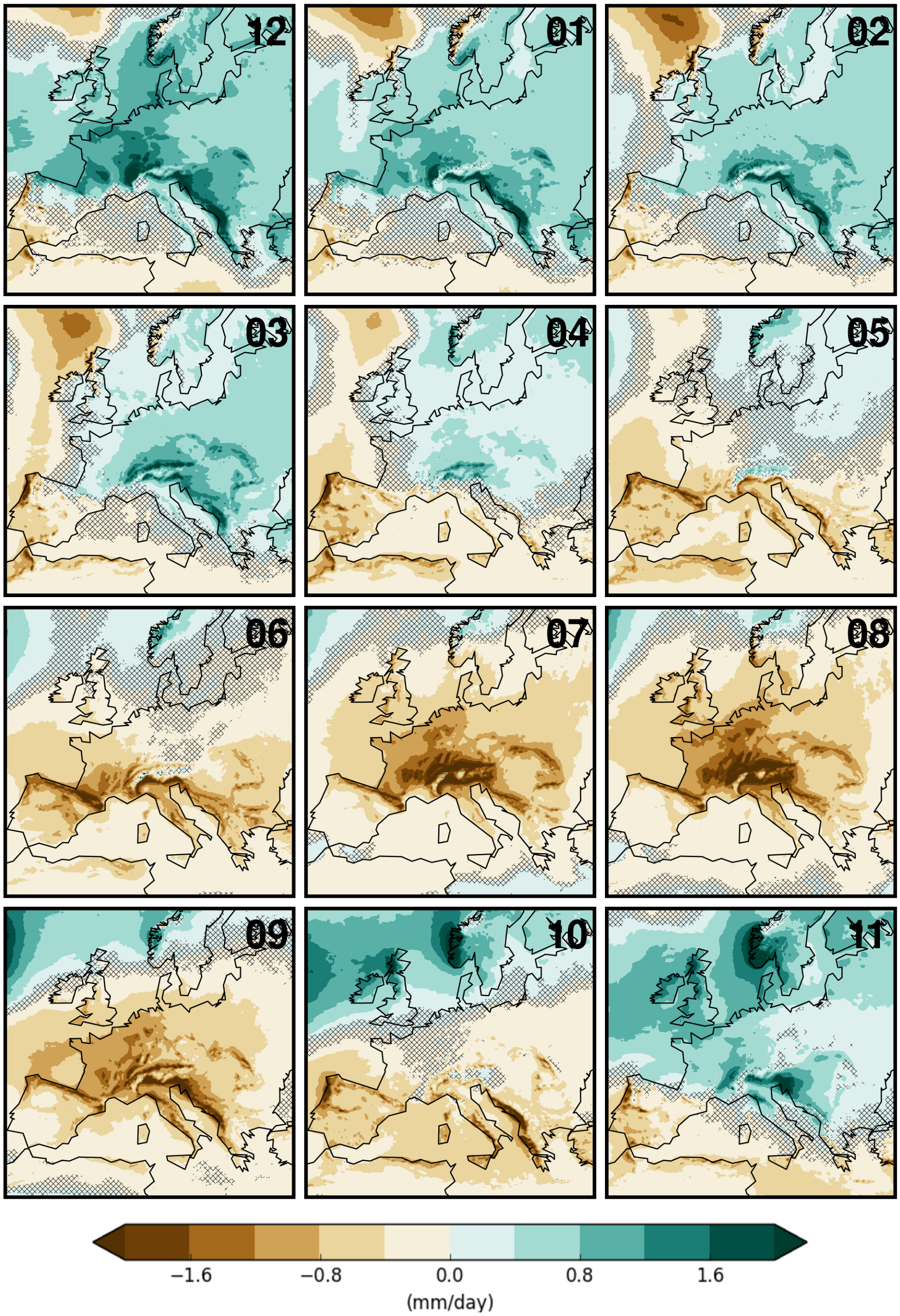


FIG. 10. Same as Figure 9 for precipitation during the 2080-2099 period over the EU domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances).

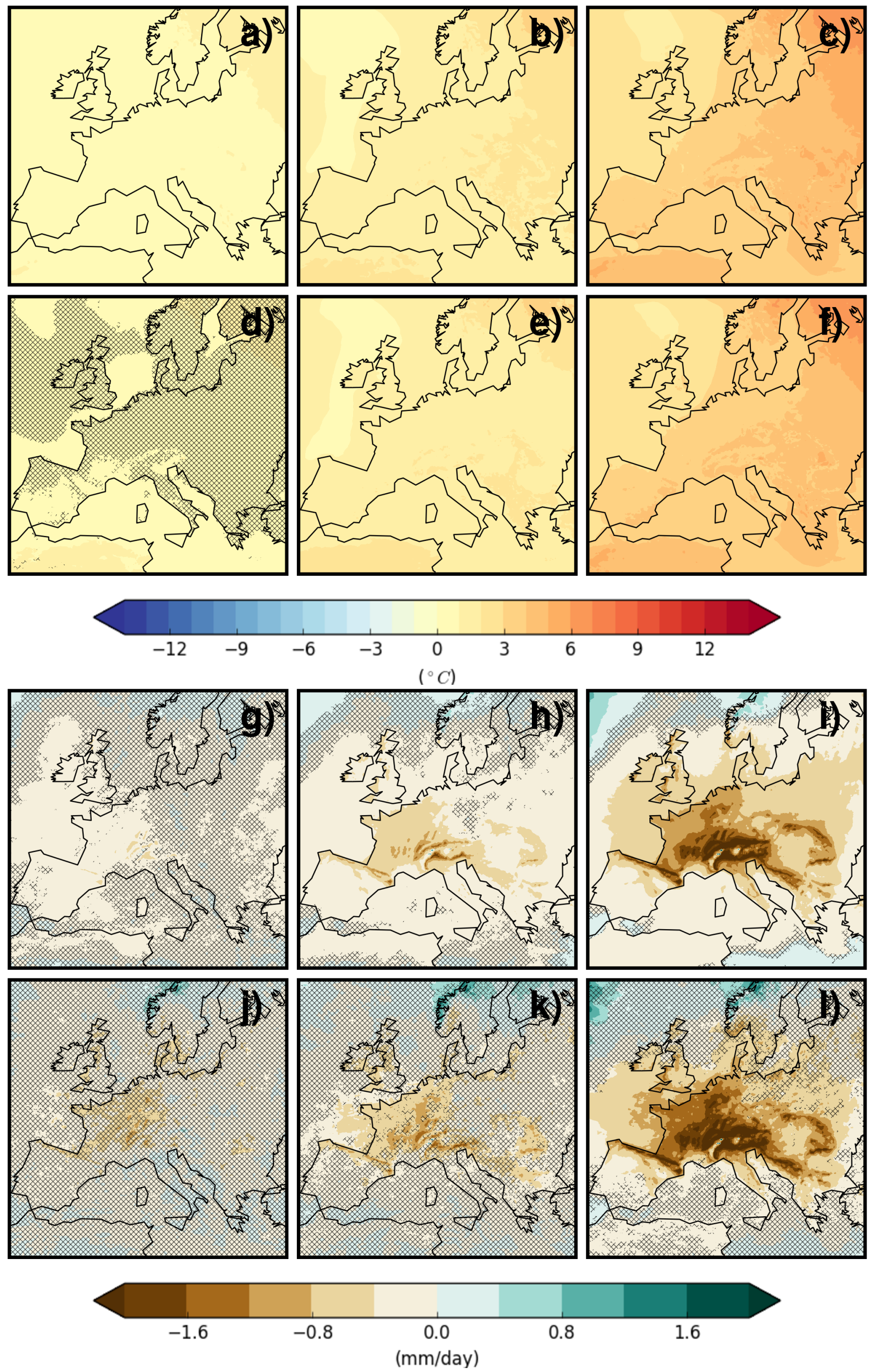


FIG. 11. (a) to (c): CRCM5 50-member ensemble mean climate-change signal for surface-air temperature during December over the EU domain computed for the (a) 2020-2039, (b) 2040-2059, and (c) 2080-2099 periods relative to 2000-2019; (d) to (f): Same as (a) to (c) for the first five members of the ensemble; (g) to (i) and (j) to (l): Same as (a) to (c) and (d) to (f)for precipitation during July. Panels (c) and (i) are reproduced from Figures 9 and 10 for clarity. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances).

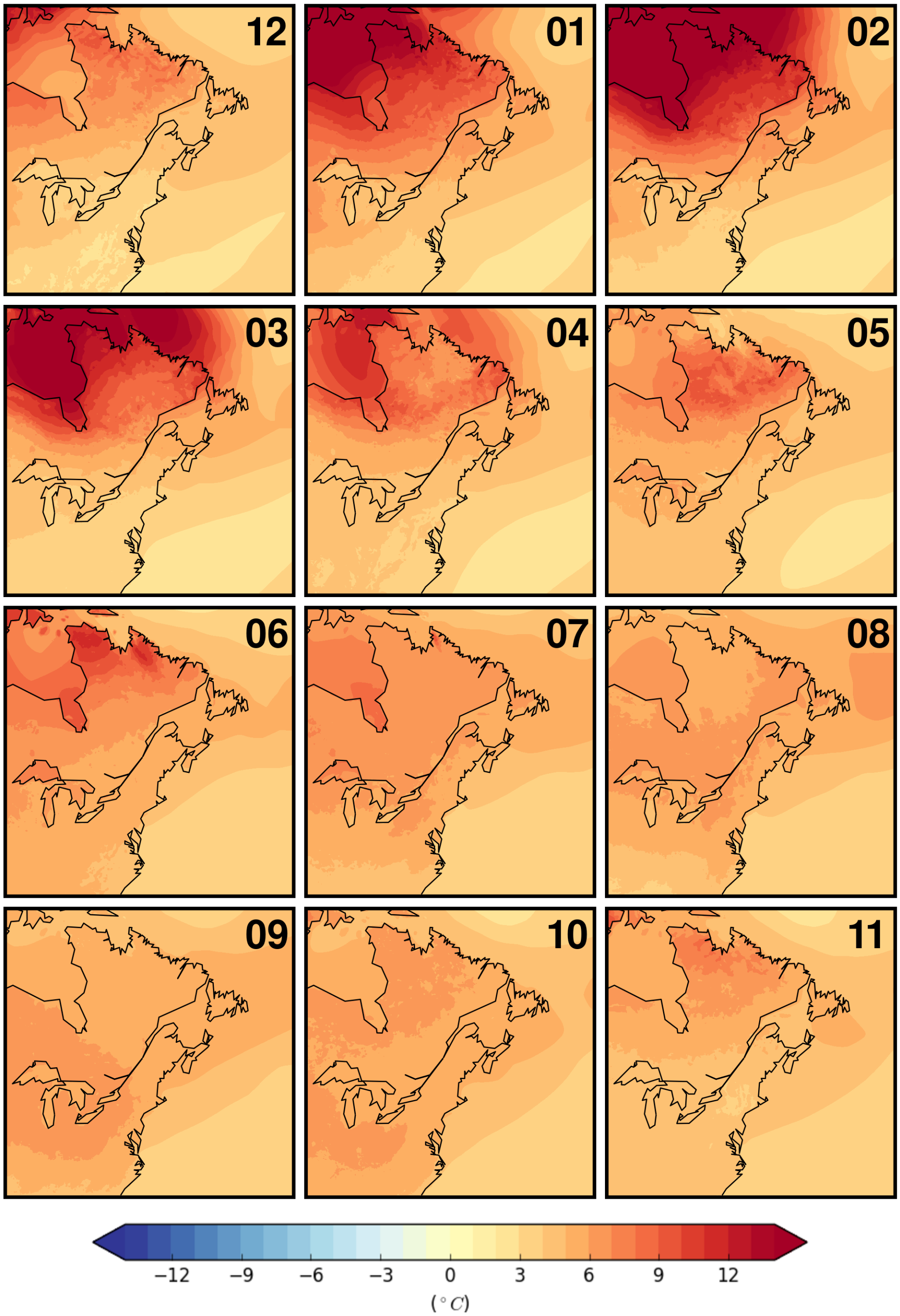


FIG. 12. Same as Figure 9 for surface-air temperature during the 2080-2099 period over the NNA domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances).

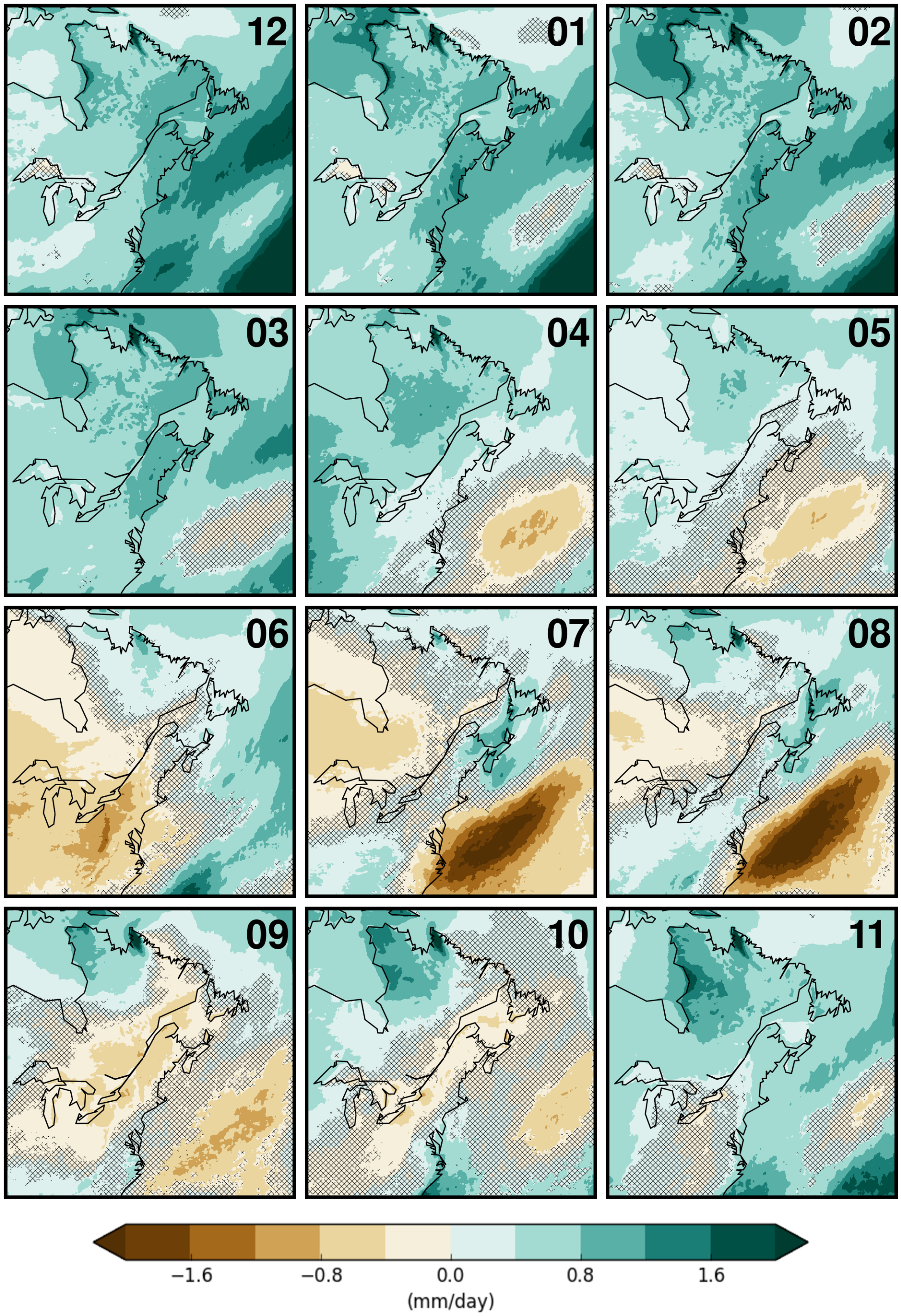


FIG. 13. Same as Figure 9 for precipitation during the 2080-2099 period over the NNA domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances).

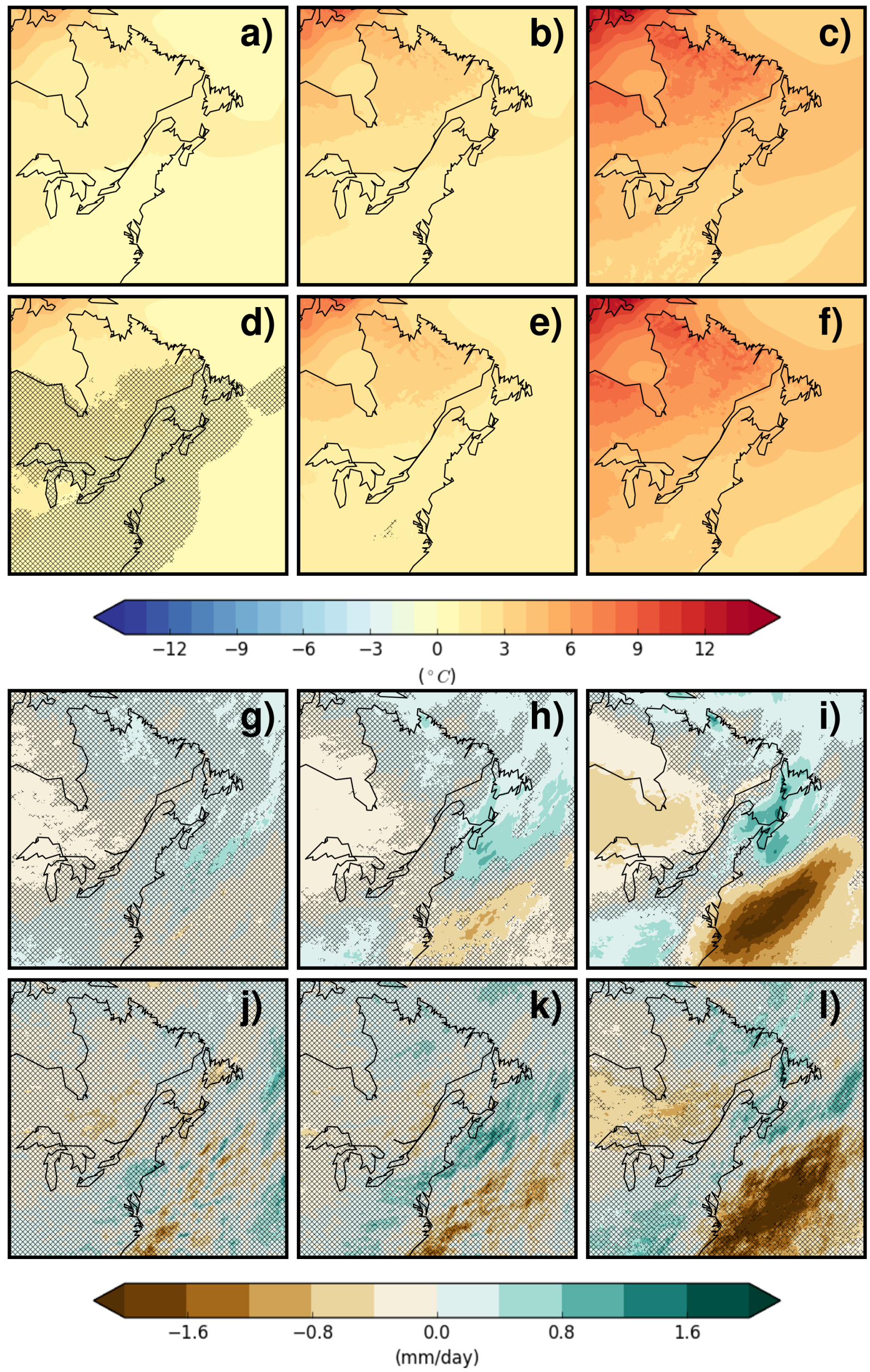


FIG. 14. (a) to (c): CRCM5 50-member ensemble mean climate-change signal for surface-air temperature during December over the NNA domain computed for the (a) 2020-2039,(b) 2040-2059, and (c) 2080-2099 periods relative to 2000-2019; (d) to (f): Same as (a) to (c) for the first five members of the ensemble; (g) to (i) and (j) to (l): Same as (a) to (c) and (d) to (f) but for precipitation during July. Panels (c) and (i) are reproduced from Figures 12 and 13 for clarity. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances).

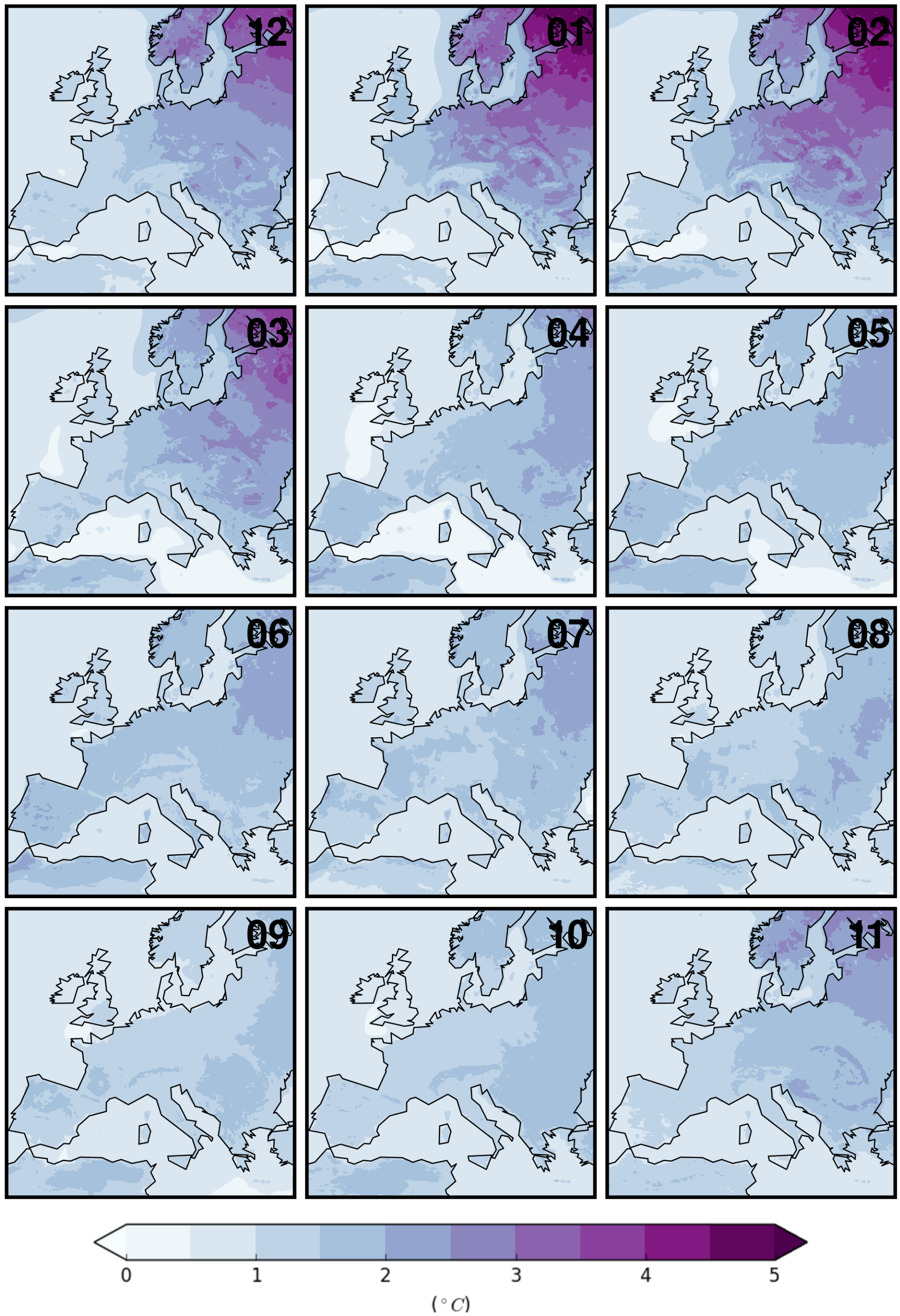
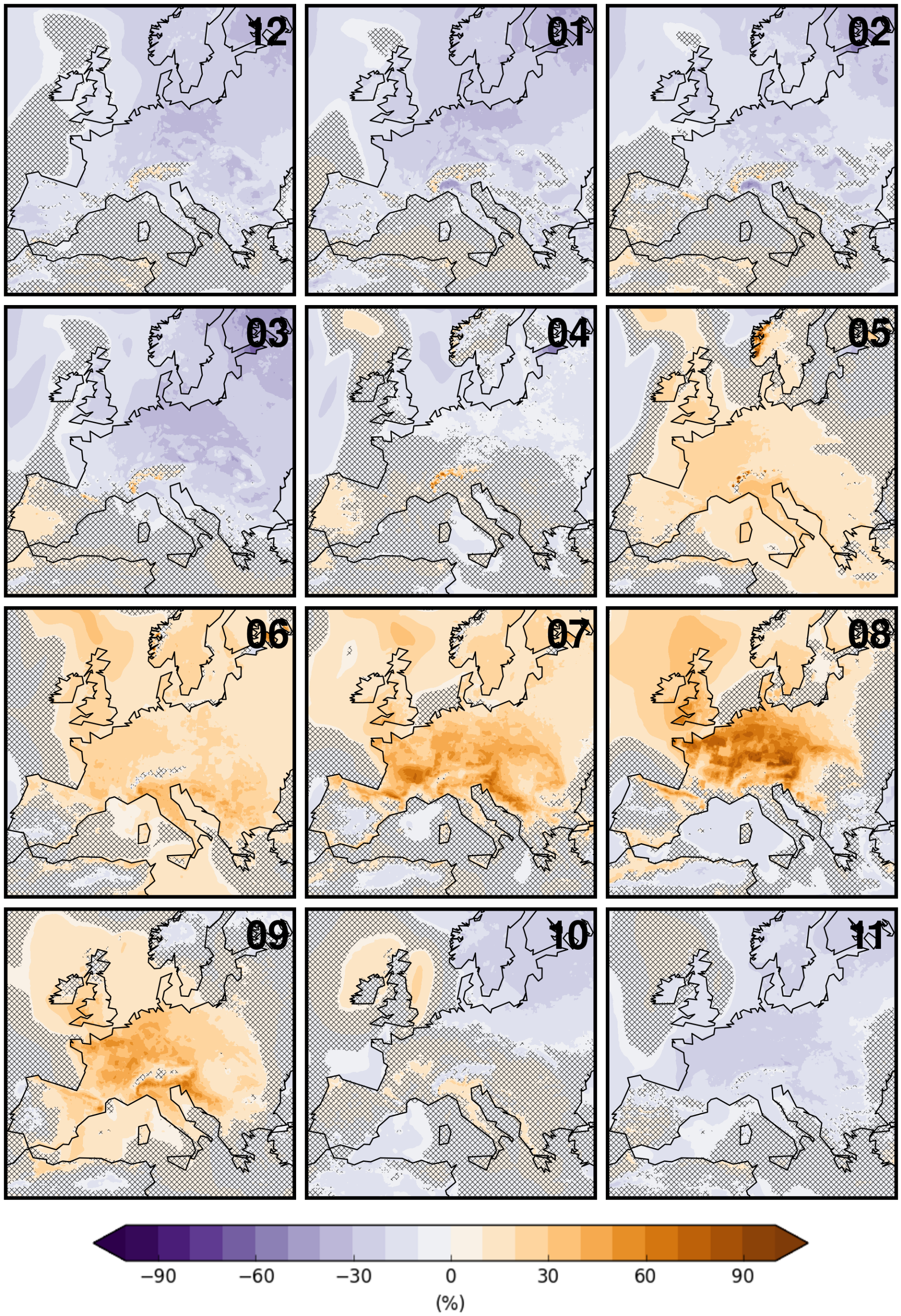
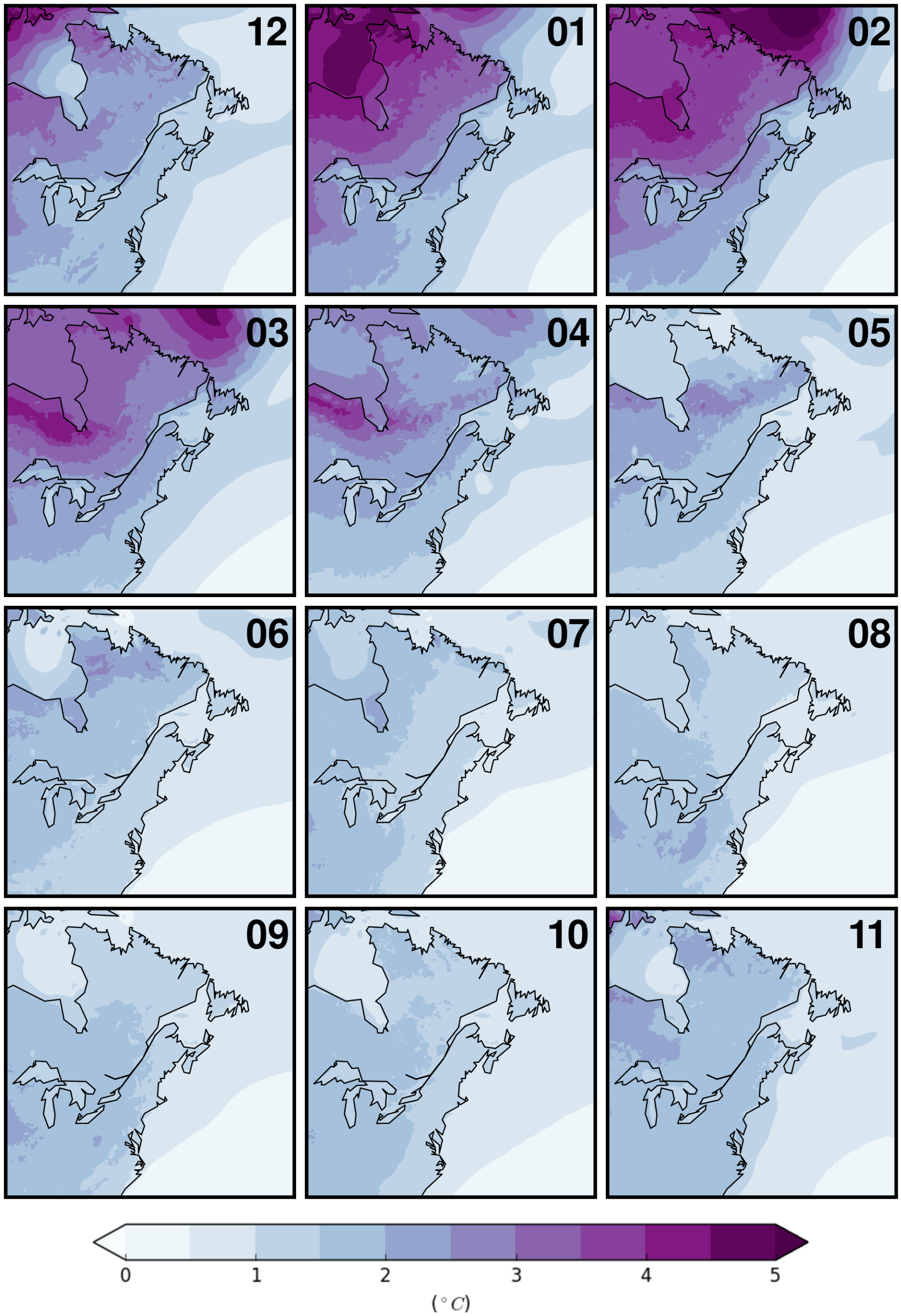


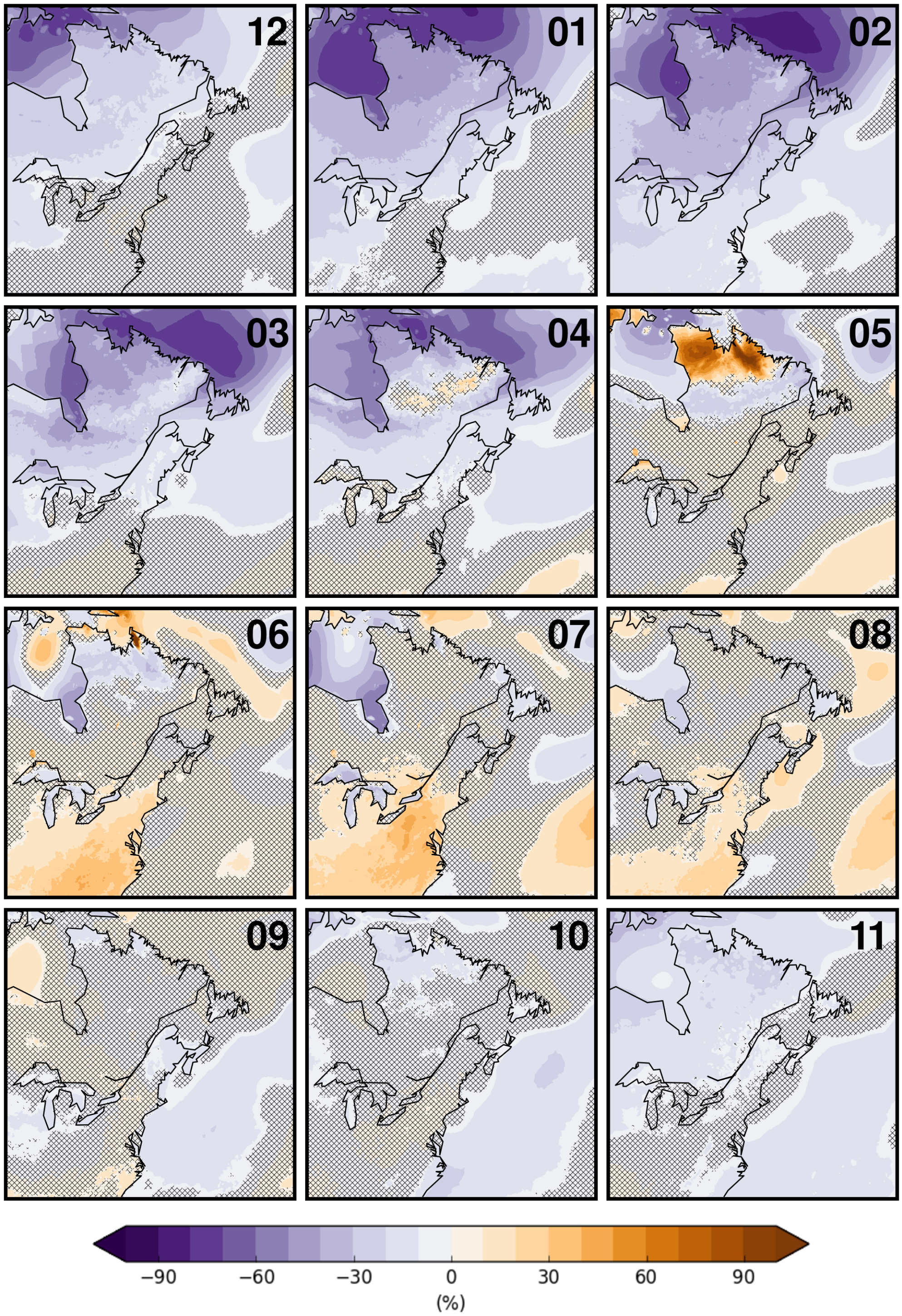
FIG. 15. Interannual variability of monthly mean surface-air temperature over the EU domain calculated as the yearly inter-member spread averaged during the 2000-2019 period. Months are labeled from 1 to 12.



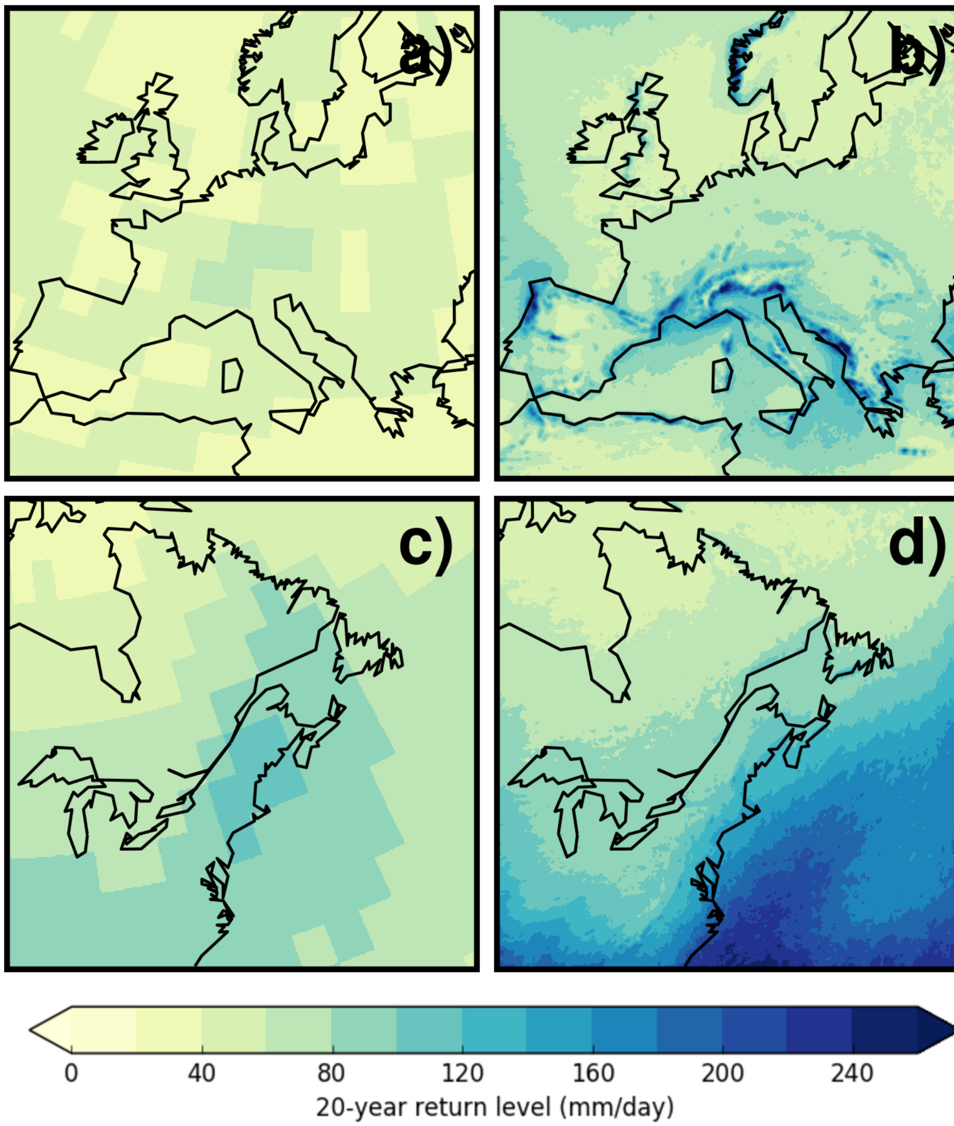
1024 FIG. 16. Relative change in interannual variability for the monthly mean surface-air temperature (2080-2099
1025 vs 2000-2019) over the EU domain. Hatched regions identify where changes are not statistically significant at
1026 the 99% confidence level (F-test). Months are labeled from 1 to 12.



1027 FIG. 17. Interannual variability of monthly mean surface-air temperature over the NNA domain calculated as
1028 the yearly inter-member spread averaged during the 2000-2019 period. Months are labeled from 1 to 12.



1029 FIG. 18. Relative change in interannual variability for the monthly mean surface-air temperature (2080-2099
1030 vs 2000-2019) over the NNA domain. Hatched regions identify where the changes are not statistically significant
1031 at the 99% confidence level (F-test). Months are labeled from 1 to 12.



1032 FIG. 19. (a) and (b): The 20-year return period values of the daily annual maximum precipitation during 2000-2019 over the EU domain as calculated from CanESM2-LE and CRCM5-LE respectively. (c) and (d): Same as (a) and (b) over the NNA domain.