



Land Surface Model influence on the simulated climatologies of temperature and precipitation extremes in the WRF v.3.9 model over North America

Almudena García-García^{1,2}, Francisco José Cuesta-Valero^{1,2}, Hugo Beltrami¹, Fidel González-Rouco³, Elena García-Bustamante⁴, and Joel Finnis⁵

¹Climate & Atmospheric Sciences Institute, St. Francis Xavier University, Antigonish, Nova Scotia, Canada.

²Environmental Sciences Program, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

³Physics of the Earth and Astrophysics Department, IGEO (UCM-CSIC), Universidad Complutense de Madrid, Spain.

⁴Research Center for Energy, Environment and Technology (CIEMAT), Madrid, Spain.

⁵Department of Geography, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

Correspondence: Hugo Beltrami (hugo@stfx.ca)

Abstract. The representation and projection of extreme temperature and precipitation events in regional and global climate models are of major importance for the study of climate change impacts. However, state-of-the-art global and regional climate model simulations yield a broad inter-model range of intensity, duration and frequency of these extremes. Here, we present a modeling experiment using the Weather Research and Forecasting (WRF) model to determine the influence of the land surface model (LSM) component on uncertainties associated with extreme events. First, we evaluate land-atmosphere interactions within four simulations performed by the WRF model using three different LSMs from 1980 to 2012 over North America. Results show LSM-dependent differences at regional scales in the frequency of occurrence of events when surface conditions are altered by atmospheric forcing or land processes. The inter-model range of extreme statistics across the WRF simulations is large, particularly for indices related to the intensity and duration of temperature and precipitation extremes. Areas showing large uncertainty in WRF simulated extreme events are also identified in a model ensemble from three different Regional Climate Model (RCM) simulations participating in the Coordinated Regional Climate Downscaling Experiment (CORDEX) project, revealing the implications of these results for other model ensembles. This study illustrates the importance of the LSM choice in climate simulations, supporting the development of new modeling studies using different LSM components to understand inter-model differences in simulating temperature and precipitation extreme events, which in turn will help to reduce uncertainties in climate model projections.



1 Introduction

General Circulation Models (GCMs) and Regional Climate Models (RCMs) are currently the most useful tools for the study of processes affecting the frequency, duration and intensity of extreme temperature and precipitation events, as well as projecting their evolution under different emission scenarios at global, regional and local scales. Both observational data and climate model simulations confirm all of these statistics respond to climate change (Seneviratne et al., 2012; Orłowsky and Seneviratne, 2012; Jeong et al., 2016). However, state-of-the-art global and regional climate models differ substantially in their interpretation of the climatology and response to warming of various indices of temperature and precipitation extremes (Sillmann et al., 2013a, b). Climate information provided by models is currently employed by public and private institutions dedicated to the evaluation and management of risks from extreme events and associated disasters (IPCC, 2013; Arneth, 2019). It is, therefore, essential that climate models represent extreme events and their evolution as realistically as possible to aid in the design of appropriate policies to mitigate climate change and build resilience. In this study, we evaluated the representation of a set of extreme indices, previously included in international reports such as IPCC (2013) and Seneviratne et al. (2012), as simulated by the Weather Research and Forecasting (WRF) model with different land surface model (LSM) components.

Land-atmosphere interactions have been identified as a key factor in the simulation of extreme events (e.g. Lorenz et al., 2016; Vogel et al., 2017). Soil conditions affect and are affected by near-surface atmospheric phenomena, through energy and water exchanges at the ground surface. For example, previous observational studies have shown the impact of soil moisture deficits on hot extreme temperatures through changes in evapotranspiration over southeastern and western Europe and Russia (Hirschi et al., 2011; Miralles et al., 2012; Hauser et al., 2016). Additionally, soil moisture regimes have been found to alter the energy and water exchanges at the surface, influencing inter-annual summer temperature variability in central parts of North America (Donat et al., 2016), and precipitation events in western North America (Diro et al., 2014). Land-Atmosphere interactions, and consequently near-surface conditions, are influenced by vegetation and snow covers (Stieglitz and Smerdon, 2007; Diro et al., 2018). For example, Diro et al. (2018) showed that interactions between snow cover and atmospheric processes influence extreme events, increasing the frequency of cold events over western North America and affecting the variability in warm events over northeast Canada and the Rocky mountains.

Metrics built on the representation of land-atmosphere interactions have been employed as a basis for evaluating extreme temperature and precipitation events in climate model simulations (Knist et al., 2016; Davin et al., 2016; Lorenz et al., 2016; Sippel et al., 2017; Gevaert et al., 2018; García-García et al., 2019). For example, Lorenz et al. (2016) evaluated outputs from six GCMs participating in the Global Land-Atmosphere Coupling Experiment of the Coupled Model Intercomparison Project, Phase 5 (GLACE-CMIP5) and concluded that ranges of intensity, frequency and duration of extreme events among climate projections are strongly related to inter-model differences in the representation land-atmosphere interactions. Gevaert et al. (2018) evaluated the representation of land-atmosphere interactions within a set of off-line LSM simulations, finding similar spatial patterns of soil moisture-temperature coupling among LSM simulations, but large variability in the degree and local patterns of land-atmosphere coupling. García-García et al. (2019) employed a simple metric derived from soil and air temperatures to evaluate outputs from the CMIP5 models and the North American Regional Reanalysis (NARR) against



observations over North America, suggesting a strong dependency of the simulated land-atmosphere interactions on the LSM component employed.

In contrast with the variety of LSM components employed in the new generation of GCMs, reanalyses use simplified versions of LSM components, typically included as part of the atmospheric model component. For example, all reanalysis products produced by the European Centre for Medium-range Weather Forecasts (ECMWF) model (CERA-20C, ERA-15, ERA20C, ERA-Interim and ERA-40 products) employed the same LSM component included in the code of the ECMWF atmospheric model. The two Modern-Era Retrospective analysis for Research and Applications (MERRA) global products employed the GEOS-5 Catchment land surface model (Reichle et al., 2011). The Japanese Reanalysis (JRA) products employed a modified version of the Simple Biosphere (SiB) LSM (Onogi et al., 2007), while most of National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) products employed the NOAH LSM (Tewari et al., 2004). The complexity and variety of these LSM components are limited in order to reduce computational costs, affecting the quality of the represented land surface processes. This has already been noted by the scientific community, and some have attempted to address the issue by incorporating updated versions of LSMs in new land reanalysis products though offline LSM simulations forced by observational data products (LDAS, MERRA-land, ERA-Interim/Land, Rodell et al., 2004; Reichle et al., 2011; Balsamo et al., 2015). Although these new products can be useful for LSM development and provide data about the soil states and fluxes (Balsamo et al., 2015), the offline character of the new land products inhibits the representation of land-atmosphere feedbacks.

Here, we perform a set of modeling experiments to evaluate for the first time the influence of the LSM component on the simulation of key extreme indices and land-atmosphere interactions within land-atmosphere coupled climate simulations at continental scales. For this purpose, four regional simulations are performed over North America (1979-2012) using the WRF model including three different LSM components widely employed in model simulations and reanalysis products, as described in Section 2. The methodology for the analysis of land-atmosphere interactions and the representation of extreme events is described in Section 3. Section 4 presents the evaluation of land-atmosphere interactions, the analysis of LSM differences in the representation of temperature and precipitation extremes, and the comparison between the WRF simulations and three Coordinated Regional Climate Downscaling Experiment (CORDEX) Evaluation simulations. A discussion about previous results and the main conclusions and implications of this study are presented in Section 5 and 6, respectively.

2 Description of the modeling experiment

We performed four regional simulations over North America (NA) using the version 3.9 of the Advanced Research WRF (ARW-WRF) model (Michalakes et al., 2001) including three different land surface models: the NOAH LSM (NOAH, Tewari et al., 2004), the NOAH LSM with multiparameterizations options (NOAH-MP, Niu et al., 2011), and the Community Land Model version 4 LSM (CLM4, Oleson et al., 2010). Vegetation cover was prescribed in these three simulations (NOAH, NOAH-MP and CLM4); an additional simulation was conducted with dynamic vegetation cover in the NOAH-MP LSM (NOAH-MP-DV), allowing for the evaluation of the influence of dynamic vegetation on extremes (NOAH-MP-DV). The use



Table 1. Characteristics of the LSM components employed for the WRF simulations performed in this analysis.

LSM	Vegetation Types	Vegetation Mode	Soil Layers	Soil Depth	Snow	Reference
NOAH	Dominant vegetation type in one grid cell	Prescribed	4	2 m	Single Layer	Tewari et al. (2004)
NOAH-MP	Dominant vegetation type in one grid cell	Prescribed	4	2 m	Up to 3 Layers	Niu et al. (2011)
NOAH-MP-DV	Dominant vegetation type in one grid cell	Dynamic	4	2 m	Up to 3 Layers	Niu et al. (2011)
CLM4	Up to 10 vegetation types in one grid cell	Prescribed	10	4.32 m	Up to 5 Layers	Oleson et al. (2010)

of different LSM components in a RCM permits the study of the influence of surface and soil processes on the simulated
 85 climate system in contrast to LSM offline simulations (Laguë et al., 2019).

The LSM components employed have been previously included in climate model studies or in reanalysis products. The
 CLM4 LSM component has been coupled to several GCMs participating in the CMIP5 project (Collins et al., 2006; Vertenstein
 et al., 2012). The NOAH LSM has been extensively used for reanalysis products, as well as for RCM simulations as those
 participating in the CORDEX project (Mesinger et al., 2006; Katragkou et al., 2015). The NOAH-MP LSM has been selected
 90 for current studies using WRF (e.g. Liu et al., 2017). The NOAH LSM is a rather basic LSM developed by the National Center
 for Atmospheric Research (NCAR) and the National Centers for Environmental Prediction (NCEP), based on the Oregon State
 University (OSU) LSM. This LSM component describes soils using 4 layers with thickness 10, 30, 60 and 100 cm, using a
 zero-flux bottom boundary condition at a depth of 2 m. The NOAH LSM estimates soil moisture and temperature at the node
 of each soil layer, taking into account snow cover, canopy moisture, and soil ice. The NOAH-MP LSM is based on the NOAH
 95 LSM, introducing relevant improvements, such as a dynamic vegetation option; a new separated vegetation canopy cover that
 improves the computation of energy, water and carbon fluxes at the surface; a separate scheme for computing energy fluxes
 over vegetated surfaces and bare soils; a new 3-layer snow model; a more permeable frozen soil; and an improved description
 of runoff and soil moisture. Although the NOAH-MP LSM is the updated version of the NOAH LSM and has been shown to
 improve the simulation of surface processes in comparison to the NOAH LSM (e.g. Niu et al., 2011; Yang et al., 2011), the
 100 NOAH-MP LSM has not yet been implemented in any reanalysis product. The CLM4 represents one of the most advanced
 LSM components, incorporating a detailed description of biogeophysics, hydrology and biogeochemistry. The CLM4 classifies
 vegetation cover according to 4 different plant functional types, considering the physiology and structure of different plants.
 The soil vertical structure is divided into a layer for the vegetation canopy, 5 layers for snow cover, and 10 soil layers, placing
 the zero-flux bottom boundary condition at approximately 4.32 m. The main characteristics of the employed LSM components
 105 are summarized in Table 1.

Beyond the structural differences among LSM components, the remaining options and parameters are identical for the four
 WRF simulations. Boundary conditions for the WRF experiments are provided by the North American Regional Reanalysis
 (NARR) product, which is formed by the NCEP Eta atmospheric model, the NOAH LSM and the Regional Data Assimilation
 System (RDAS); (Mesinger et al., 2006). NARR data are provided with a 32 km grid and three-hourly temporal resolution,
 110 available at the National Center for Environmental Information (NOAA) archive. The domain set for the WRF simulations



has 50 km horizontal resolution and 27 atmospheric levels, covering North America in a Lambert projection. The land use categories employed for the four simulations are derived from the Moderate Resolution Imaging Spectroradiometer (MODIS, Barlage et al., 2005). The four WRF simulations start in January 1st 1979, which is the first year of the NARR product, and end in December 31st 2012, using a time-step of 300 seconds for the model integrations. We use the first year of each simulation
115 as spin-up and the other 33 years for the analysis. The employed physics parameterizations include the WSM 6-class graupel scheme for the microphysics (Hong and Lim, 2006), the Grell-Freitas ensemble scheme for cumulus description (Grell and Freitas, 2014), the Yonsei University scheme as planetary boundary layer scheme (YSU, Hong et al., 2006), the revised MM5 monin-Obukhov scheme for the surface layer (Jiménez et al., 2012), and the CAM scheme for the integration of radiation physics each 20 min intervals (Collins et al., 2004).

120 The gap in resolution from the employed boundary conditions (32 km) to the final simulations (50 km) can be counter-intuitive for a RCM experiment ; indeed. The rationale for this decrease in resolution is that this set of simulations constitutes an ensemble of WRF sensitivity experiments to using different LSM components. The computational resources saved with this coarse resolution allow us to perform simulations long enough for the evaluation of land-atmosphere interactions and extreme events at climatological scales and yet similar horizontal resolution and domain to those employed in the North American
125 component of the CORDEX project (Giorgi and Gutowski Jr., 2015) can be attained. Additionally, we do not apply any nudging technique, ensuring that the RCM evolves freely according to each LSM component and its representation of land-atmosphere interactions.

3 Methodology

Different metrics have been employed in the literature for the evaluation of land-atmosphere interactions within climate model
130 simulations and observations. Among these metrics, we selected the Vegetation-Atmosphere Coupling (VAC) index (Zscheischler et al., 2015) as our evaluation metric for the representation of land-atmosphere interactions at monthly scales. This index has been previously employed in the literature to identify regions with episodes of strong land-atmosphere coupling within climate model simulations and observational data (Zscheischler et al., 2015; Gevaert et al., 2018; Sippel et al., 2017; Li et al., 2017; Philip et al., 2018). The VAC index is segregated in four categories based on the simultaneous occurrence of some given
135 extreme percentile ranges of Surface Air Temperature (SAT) and latent heat flux (LH, Philip et al., 2018):

$$\begin{aligned} VAC_a & \text{ if } SAT < 30^{th} Pctl. \text{ and } LH < 30^{th} Pctl. \rightarrow Atmos. \text{ Control} \\ VAC_b & \text{ if } SAT > 70^{th} Pctl. \text{ and } LH > 70^{th} Pctl. \rightarrow Atmos. \text{ Control} \\ VAC_c & \text{ if } SAT > 70^{th} Pctl. \text{ and } LH < 30^{th} Pctl. \rightarrow Land \text{ Control} \\ VAC_d & \text{ if } SAT < 30^{th} Pctl. \text{ and } LH > 70^{th} Pctl. \rightarrow Land \text{ Control} \\ 0 & \text{ otherwise} \end{aligned} \tag{1}$$

Extremes of SAT and LH are defined as values exceeding (below) the 70th (30th) percentile, relative to a 20-year period (1980-2000) (Eq. 1). The 30th and 70th percentile thresholds were also employed in previous studies based on monthly data



(Sippel et al., 2017). The VAC index classifies areas depending on the soil moisture regime into energy-limited areas, where
140 atmospheric forcing controls processes at the land surface (VAC_a and VAC_b), and transitional areas, where land surface pro-
cesses are driven by soil moisture deficits (VAC_c and VAC_d). As explained in Zscheischler et al. (2015), the VAC_a category
is associated with low SAT caused by the presence of clouds and precipitation, which leads to low vegetation activity likely
rising soil moisture. The VAC_b category is frequent in wet areas with high SAT, usually related to clear sky and high radiation,
which is associated with the increase in vegetation activity inducing the depletion of soil moisture. During VAC_c episodes, the
145 combination of high SAT and soil moisture deficits leads to diminished vegetation activity, followed by low precipitation and
consequently reduced soil moisture and high SAT, promoting heat waves and droughts. The VAC_d category is associated with
high precipitation over dry soils which stimulates vegetation activity, increases soil moisture and decreases SAT.

We calculate the frequency of occurrence for each VAC category using deseasonalized and detrended monthly SAT and
LH series following the typical methodology (Sippel et al., 2017) at each grid cell from 1980 to 2012, hereafter the analysis
150 period. The frequency of occurrence for each VAC category is calculated by counting the VAC events for the analysis period
seasonally; in boreal winter (December, January, and February; DJF), in spring (March, April, and May; MAM), in summer
(June, July, and August; JJA), and in fall (September, October, and November; SON). The VAC frequencies of occurrence
for each category are considered significant when higher than the 95th percentile of the population obtained by 100 randomly
sorted 34-year time series of SAT and LH. For the study of land-atmosphere coupling within each simulation, we represent the
155 averaged frequency of events under atmospheric control (VAC_a and VAC_b) and under land control (VAC_c and VAC_d) at grid
cells with significant frequency of occurrence for at least one of the two VAC categories.

After the evaluation of land-atmosphere interactions in our set of simulations, we assess the representation of extreme
events across the WRF simulations coupled to different LSM components. There are several definitions of indices related
to temperature and precipitation extremes, mainly using thresholds based on absolute values or statistical percentiles (e.g.
160 Sillmann et al., 2013a). The evaluation of model simulations in representing indices based on absolute values could include
model-specific biases, that can be corrected by bias removal techniques. However, the advantage of applying bias removal
techniques techniques is not clear for the study of future climate trends and climate variability, since they have been proven
to modify the spatiotemporal consistency of climate models as well as internal feedback mechanisms and conservation terms
(Ehret et al., 2012; Cannon et al., 2015). Additionally, the simulation of absolute temperatures are of central importance
165 for temperature dependent processes that may have important consequences for society and ecosystems, such as soil carbon
processes (Hicks Pries et al., 2017). Studies based on statistical percentiles improve the comparison among models but hamper
the interpretation of results by losing the physical meaning of the variable (temperature or precipitation). Since extreme indices
based on both absolute values and statistical thresholds present advantages and disadvantages, we selected a set of indices
including both categories from the list of 27 indices recommended by the Expert Team on Climate Change Detection and
170 Indices (ETCCDI, Karl et al., 1999, Table 2). Since we are interested in the climatology of extreme events, temporal averages
of each annual index are computed for the analysis period at each grid cell for each WRF experiment. Then, we compute the
inter-model range of each index across the WRF simulations (i.e., the difference between the maximum and minimum values at



Table 2. List of extreme indices used in this study defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Karl et al., 1999). Percentiles are calculated over the period 1980-2000.

Index	Definition	Unit
Cold Event		
Intensity		
TX _x DJF	Maximum value of daily maximum temperature (hottest day) in winter	°C
TN _n DJF	Minimum value of daily minimum temperature (coldest night) in winter	°C
Frequency		
TN10p	Percentage of days in a year when daily minimum temperature < the calendar day 10th percentile centered on a 5-day window	%
TX10p	Percentage of days in a year when daily maximum temperature < the calendar day 10th percentile centered on a 5-day window	%
Duration		
CSDI	Cold Spell Duration Index: annual count of days with at least 6 consecutive days when daily minimum temperature < the calendar day 10th percentile centred on a 5-day window	Days
Warm Event		
Intensity		
TX _x JJA	Maximum value of daily maximum temperature (hottest day) in summer	°C
TN _n JJA	Minimum value of daily minimum temperature (coldest night) in summer	°C
Frequency		
TN90p	Percentage of days in a year when daily minimum temperature > the calendar day 90th percentile centered on a 5-day window	%
TX90p	Percentage of days in a year when daily maximum temperature > the calendar day 90th percentile centered on a 5-day window	%
Duration		
WSDI	Warm Spell Duration Index: annual count of days with at least 6 consecutive days when daily maximum temperature > the calendar day 90th percentile centred on a 5-day window	Days
Precipitation Event		
Intensity		
R95p	Annual total precipitation when daily accumulated precipitation on a wet day > 95th percentile of precipitation on wet days	mm
Frequency		
R10mm	Annual count of days when daily accumulated precipitation ≥ 10mm	Days
Duration		
CDD	Maximum length of dry spell: maximum annual number of consecutive days with daily accumulated precipitation < 1mm	Days
CWD	Maximum length of wet spell: maximum annual number of consecutive days with daily accumulated precipitation ≥ 1mm	Days



each grid cell considering the four WRF simulations), using it as metric for the uncertainty in the WRF simulation of extreme events arising from the LSM component.

175 The LSM effect on the WRF simulation of extreme temperature and precipitation events was also compared with the representation of extreme events by three different RCMs participating in the North America CORDEX (NA CORDEX) program, using daily data from three Evaluation simulations (Table S1). These CORDEX simulations were performed by the WRF model (Skamarock et al., 2008), the RCA4 model (Samuelsson et al., 2011), and the CRCM-UQAM model (Martynov et al., 2013), using boundary conditions from the ERA-Interim reanalysis (Dee et al., 2011). The spatial domain and resolution of
180 the NA CORDEX simulations are similar to that of the WRF simulations, as indicated in Section 2. Refer to Table S2 for information about the availability of the data employed in this work.

4 Results

4.1 Evaluation of land-atmosphere interactions in WRF simulations

All WRF simulations with different LSM components display similar spatial patterns for VAC categories, agreeing in the
185 seasonality and broadly in the areas with high probability of episodes when atmospheric forcing or soil conditions control processes at the land surface (Figures 1 and 2). Atmospheric forcing controls surface processes at middle and high latitudes in MAM, JJA and SON, moving southward in DJF (Figure 1). Areas frequently driven by soil processes are displayed over the western Mexican coast in DJF, spreading across low and middle latitudes in MAM, JJA and SON (Figure 2). Despite the broad agreement between LSM simulations in the spatial distribution of the VAC categories, there are regional differences in
190 their representation of land-atmosphere coupling. These regional differences allow us to identify the NOAH LSM as the one simulating the weakest annual land control on processes at the surface, mainly due to a relatively weak land control during MAM and JJA (Figure 2).

The areas where LSM simulations differ in the probability of episodes under atmospheric control (VAC_a and VAC_b) vary with the season; for example the NOAH-MP LSM simulates a large area under atmospheric control over the southeastern US
195 in DJF, while the CLM4 and NOAH LSMs identify atmospheric control areas below the Great Lakes following a northwestern direction (Figure 1). In MAM, the NOAH-MP LSM represents higher probability of atmospheric control episodes over the northern US in comparison with the CLM4 and NOAH simulations (Figure 1). The NOAH simulation shows the strongest atmospheric control in JJA as compared with the remaining simulations, particularly over eastern and western regions of Hudson Bay, the southeastern US and small areas in Mexico (Figure 1). During SON, the NOAH-MP LSM reaches the highest
200 probability of episodes under atmospheric control at middle and high latitudes. The contribution of the VAC_a and VAC_b categories to these episodes is broadly similar across LSMs, with slightly higher VAC_a in all seasons; modest LSM-specific differences include a tendency for the NOAH simulation to show slightly higher VAC_a probabilities across all seasons (but especially DJF) (Figures S1 and S2). LSM differences in the representation of VAC_a and VAC_b probabilities suggest the LSM influence on the evolution of atmospheric conditions.

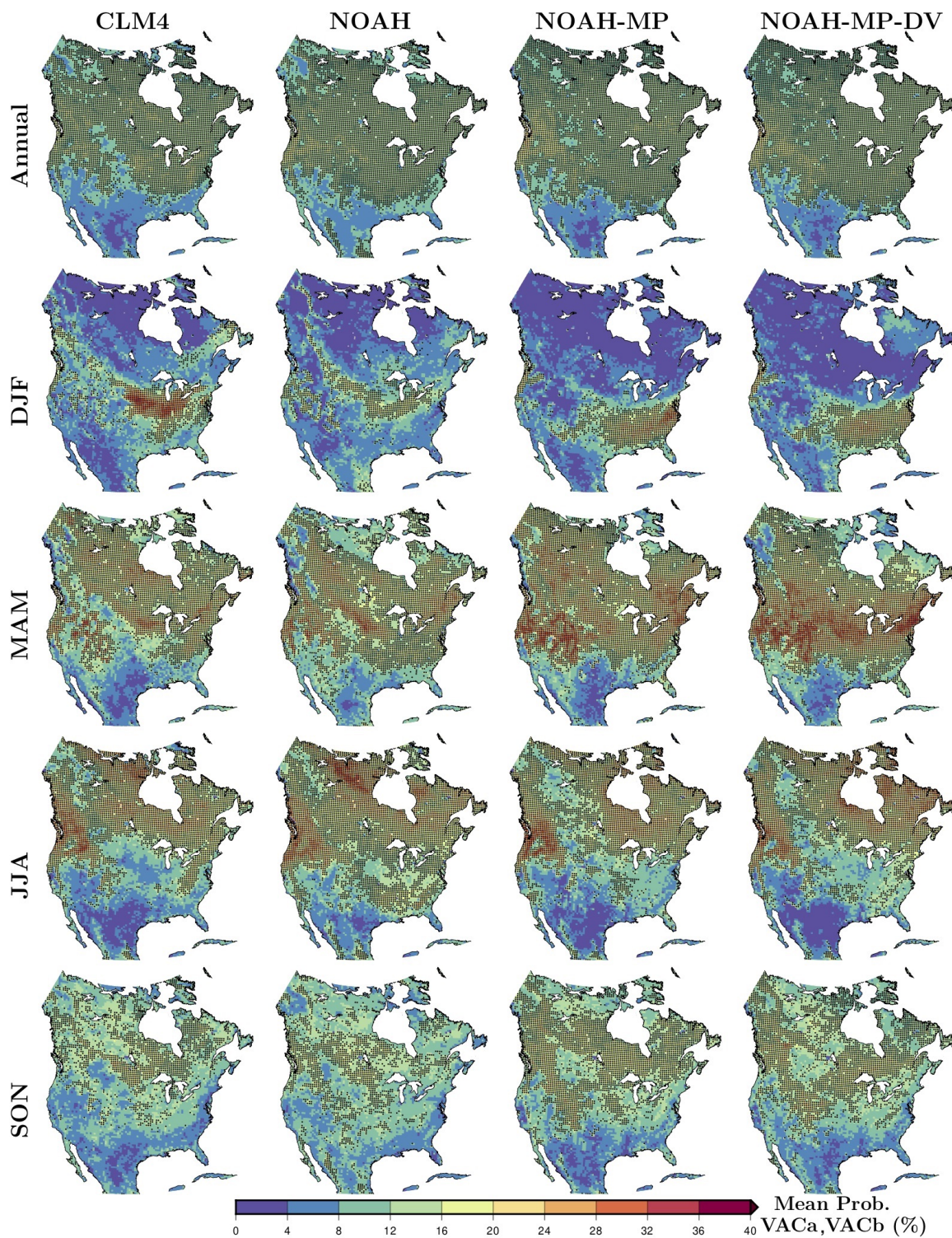


Figure 1. Mean frequency of occurrence for VAC categories associated with atmospheric control (VAC_a and VAC_b) for each simulation annually and seasonally; DJF, MAM, JJA and SON. Black dots in the maps indicate VAC values of at least one category larger than the 95th percentile of the randomly generated series.

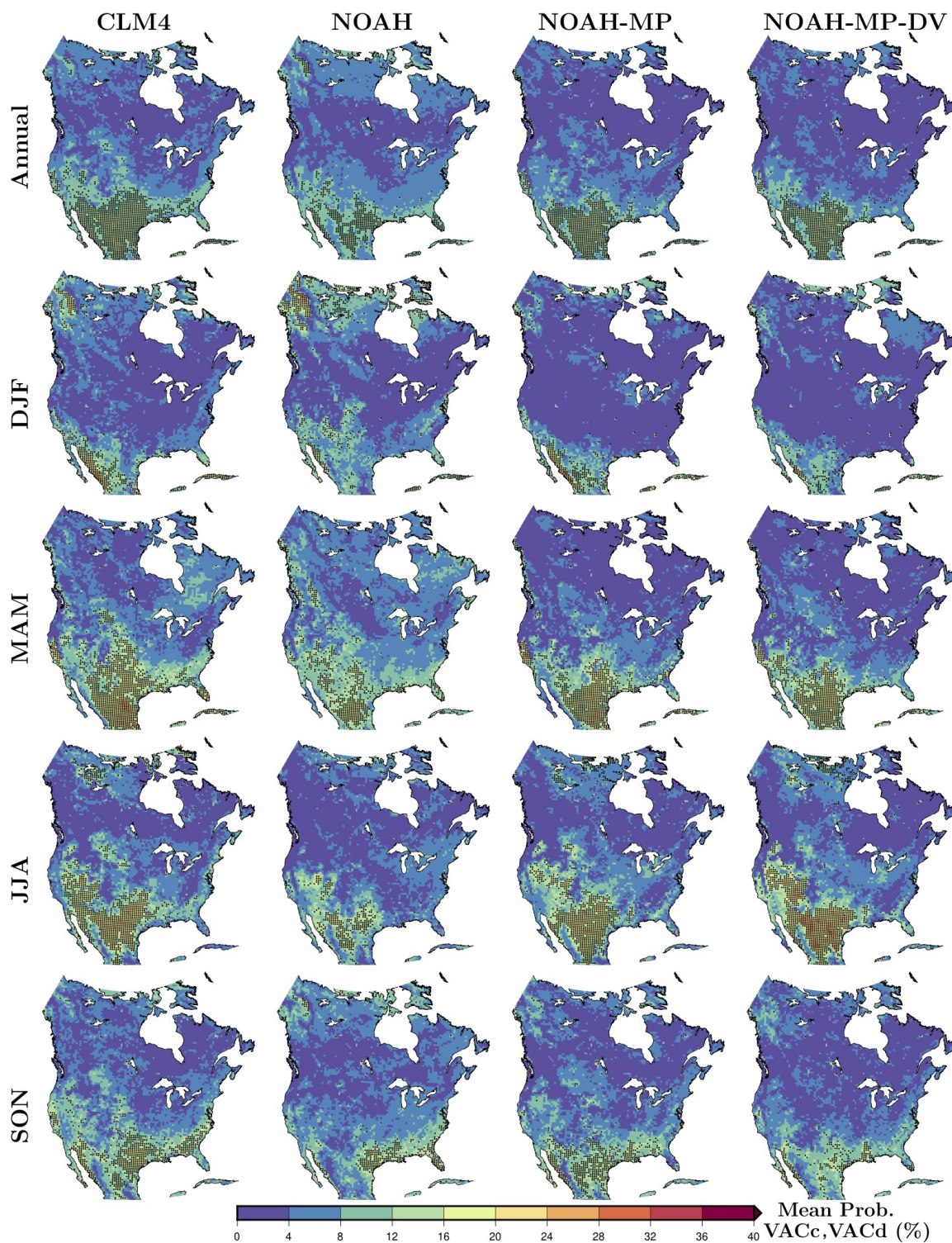


Figure 2. Mean frequency of occurrence for VAC categories associated with land control (VAC_c and VAC_d) for each simulation annually and seasonally; DJF, MAM, JJA and SON. Black dots in the maps indicate VAC values of at least one category larger than the 95th percentile of the randomly generated series.



205 Although the NOAH simulation displays the weakest land control for all seasons, it shows regions under land control over
northwestern North America in DJF also indicated by the CLM4 simulation, but absent in the NOAH-MP and NOAH-MP-
DV simulations (Figure 2). The probability of land control episodes over the western Mexican coast is higher in the CLM4
and NOAH-MP simulations than in the NOAH and NOAH-MP-DV simulations in DJF. In JJA, however, the NOAH-MP-DV
simulation presents a stronger land control at low and middle latitudes than the NOAH-MP simulation (Figure 2). There are
210 also regional differences between LSM simulations in SON, particularly over the southeastern US coast where the CLM4
shows the strongest land control, followed by the NOAH-MP simulation (Figure 2). Exploring the contribution of respective
 VAC_c and VAC_d separately, it is shown they present small differences; for example, the VAC_c probability in DJF is slightly
higher than the VAC_d probability for all simulations, showing the opposite behavior in JJA for the NOAH-MP and the NOAH-
MP-DV simulations (Figures S3 and S4). The LSM differences shown in the representation of land control VAC categories
215 likely imply LSM differences in the simulated statistic of extreme events because of the relationship between VAC_c episodes
and heat waves and droughts (Zscheischler et al., 2015).

4.2 Climatologies of temperature and precipitation extremes in WRF simulations

The climatologies of temperature and precipitation extremes for the analysis period, represented by their means, show similar
spatial patterns across all WRF simulations with different LSM component (Figures S5, S6 and S7). Figure 3 represents the
220 simulated climatologies of all extreme indices for the ensemble mean, formed by all WRF simulations. The WRF ensemble
mean shows the most intense cold events at high latitudes and high elevations, with cold events being more frequent and longer
over northwestern North America and over Mexico (Figure 3a). The simulation of warm events is more intense in coastal
areas of the US and Mexico and over the central US, being more frequent and longer over southern North America with a
high percentage of hot nights over northeastern NA (Figure 3b). Precipitation events are heavier and more frequent at higher
225 elevations and over southeastern NA (Figure 3c). The longest dry periods are simulated over the western Mexican and US
coasts, reaching more than 80 consecutive dry days, while the longest wet periods are represented over the Rockies and the
northwestern Mexican coast (Figure 3c).

Figure 4 summarizes the averaged climatology of each extreme index for each simulation. Averages are computed over six
regions adapted from Giorgi and Francisco (2000): Central America, CAM; Western North America, WNA; Central North
230 America, CNA; Eastern North America, ENA; Alaska, ALA; and Greenland, GRL. Colors in the figure correspond to the
hottest (red) and coldest (blue) index values among the WRF simulations for the representation of cold and warm temperature
extremes, and to the driest (brown) and wettest (green) index values for the representation of precipitation extremes over each
region. This approach helps us to identify the CLM4 simulation as that with the weakest and shortest cold extreme events,
although simulating more frequent cold events than the rest of LSM components (Figure 4a). Meanwhile, the NOAH-MP-
235 DV simulation shows more intense cold extremes during shorter periods over most of the regions (CAM, CNA, ENA and
ALA) in comparison with the NOAH-MP simulation which uses prescribed vegetation (Figure 4a). The CLM4 simulation
also corresponds to the most intense representation of warm extremes for the index based on maximum temperatures, while
the intensity index based on minimum temperatures shows higher values in the NOAH-MP simulation, except for the CAM



CLIMATOLOGY OF EXTREMES

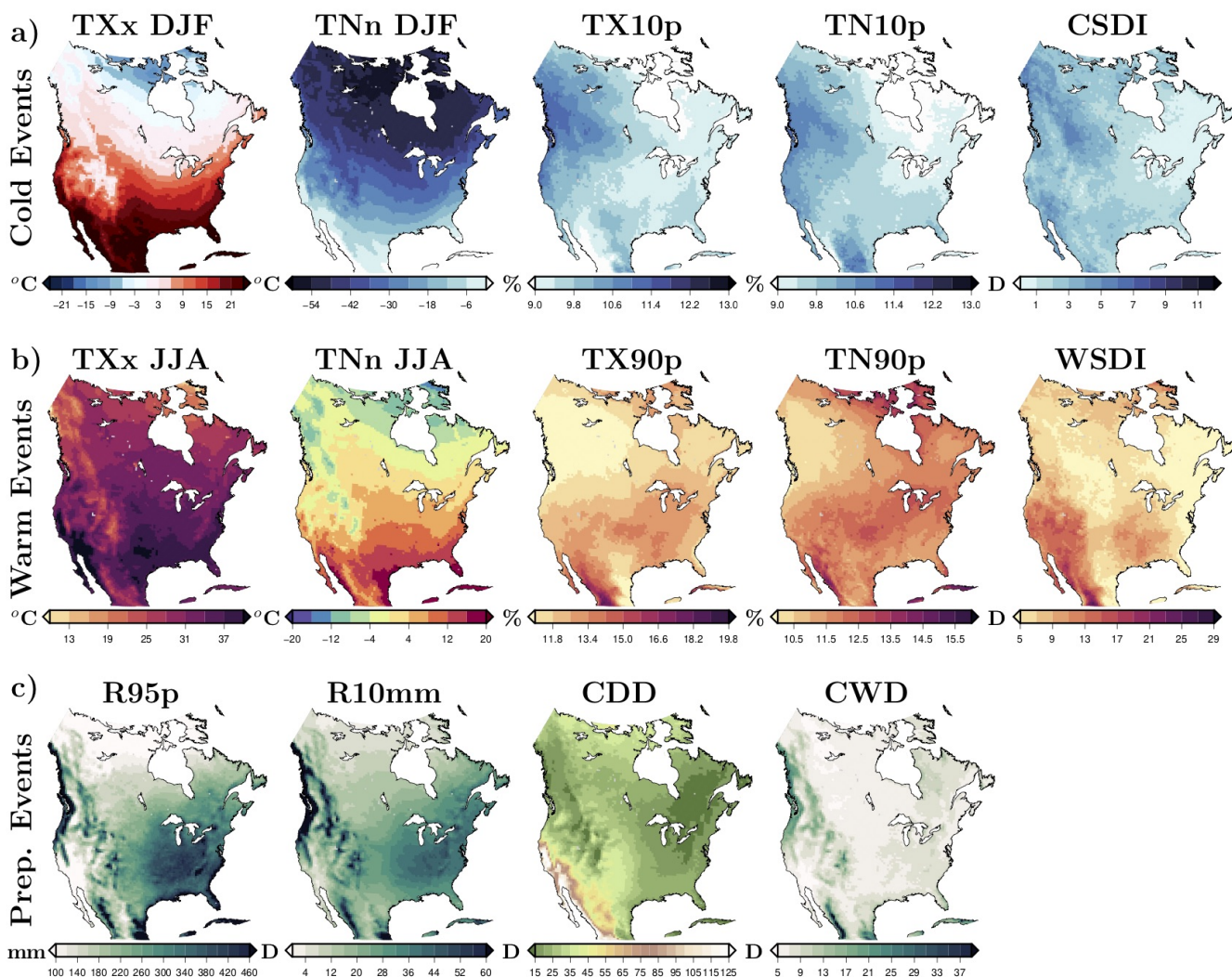


Figure 3. Climatology of extreme indices associated with cold temperature events (a), warm temperature events (b), and precipitation events (c) for the WRF ensemble mean (Table 2). The climatology of each index is estimated as the mean of each extreme index at each grid cell for the analysis period (1980-2012).

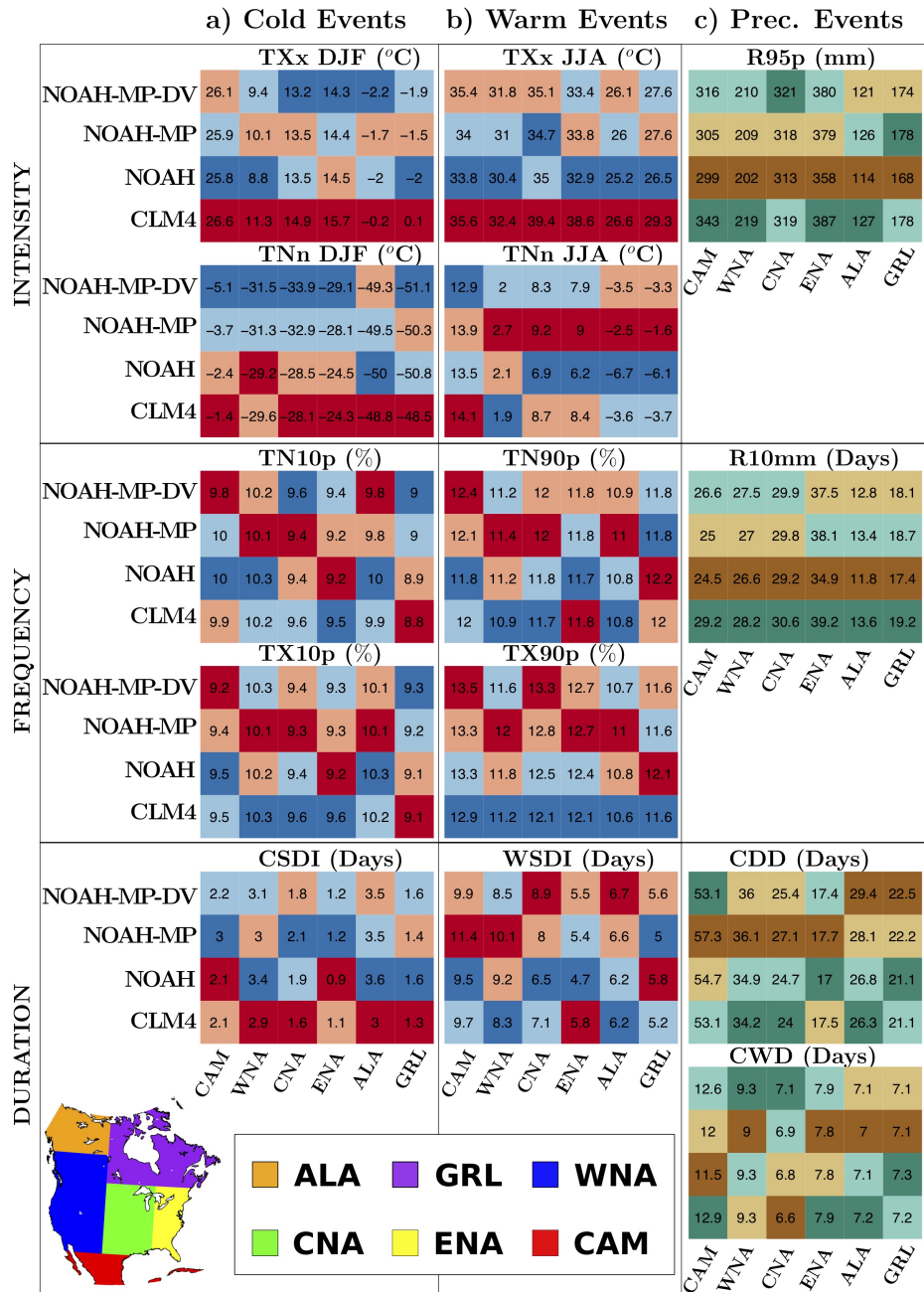


Figure 4. Comparison of the simulated climatologies of temperature and precipitation extreme indices included in Table 2 among the WRF simulations averaging over six land North American regions adapted from Giorgi and Francisco (2000) (Central America, CAM; Western North America, WNA; Central North America, CNA; Eastern North America, ENA; Alaska, ALA; and Greenland, GRL). Mean values of each index over each region are sorted, identifying the warmest (red) and coldest (blue) simulations for the indices associated with cold (a) and warm (b) temperature events as well as identifying the driest (brown) and wettest (green) simulations for the indices associated with precipitation events (c).



region (Figure 4b). The NOAH simulation is associated with the weakest and shortest warm extremes over most areas, and
240 the NOAH-MP and NOAH-MP-DV simulations with the most frequent and longest events. The effect of dynamic vegetation
seems to weaken hot extremes at nights over all regions, making them longer at middle and high latitudes (CNA, ENA,
ALA and GRL), except in the western US (Figure 4b). For precipitation extreme events, the CLM4 simulation shows the most
intense and frequent precipitation events over most areas, while the NOAH simulation shows the weakest and the least frequent
precipitation events (Figure 4c). The NOAH-MP simulation produces the longest dry periods over all regions except at high
245 latitudes, where the NOAH-MP-DV simulation yields a higher number of consecutive dry days (Figure 4c). The simulation
with dynamic vegetation yields wetter results than the simulation with prescribed vegetation at middle and low latitudes, while
at high latitudes the NOAH-MP-DV simulation is generally drier than the NOAH-MP simulation (Figure 4c).

In summary, although all simulations represent a similar spatial pattern of the climatology of extreme indices, each LSM
simulation produces different values for temperature and precipitation extreme events. Thus, the CLM4 LSM is identified
250 as the component yielding the highest temperatures during cold and warm events over most of North America as well as
the heaviest and most frequent precipitation extremes over most locations. Meanwhile, the NOAH LSM produces one of the
weakest climatologies for all temperature and precipitation extreme indices over most regions.

4.3 LSM uncertainty in the simulation of temperature and precipitation extremes

Although all WRF simulations show similar spatial patterns for temperature and precipitation extreme indices (Figures S5,
255 S6 and S7), there are large uncertainties in the climatology of each extreme index associated with the use of different LSM
components. For the simulation of the intensity of cold events, the multi-model range across the WRF simulations for the hottest
day in DJF (TXx DJF) shows large values over the boreal forest and the Rockies, where the index climatology is close to 0°C
(Figures 3 and 5a). The representation of the coldest night in DJF (TNn DJF) shows large LSM dependency, yielding ranges
up to 12 °C over the US and a spatial average of 4 °C, displaying large uncertainties over areas where the index climatology
260 approaches to 0°C (Figures 3 and 5a). The simulated intensity of warm temperature events, measured by the temporal average
of the hottest day in summer (TXx JJA), differs up to 10 °C among simulations over eastern North America, with a spatial
average of 3.5 °C (Figure 5a). The simulation of the mean coldest night in summer (TNn JJA) varies across simulations from 2
to 3 °C over the whole domain, except in the Arctic where the range across simulations reaches approximately 15 °C and the
index value yields negative temperatures for some simulations (Figure 3 and 5a). The frequency of warm extreme temperature
265 events varies among simulations; the range for the number of hot days (TX90p, based on maximum temperatures) is up to 4.2%
over the US with a spatial average of 0.97% over the whole domain, and the range for the number of hot nights (TN90p, based
on minimum temperatures) reaches values up to 3.8% at low latitudes with a spatial average of approx. 0.7% (Figure 5b). Large
values of the multi-model range for the TX90p index approximately coincide with the largest index values (Figures 3 and 5b).
Note that ranges of more than 2% in the number of hot days and nights correspond to differences of more than 7 days per year
270 in the index climatology simulated by different LSMs. Ranges of indices related to the frequency of cold events show smaller
values than those for warm temperature events, displaying no clear spatial pattern with averages of ~ 0.5% (i.e. 1.8 days per
year) for the number of cold days and nights (TX10p and TN10p; Figure 5b). The duration of warm spells is greatly affected



by the choice of the LSM component, while its effect is weaker on the simulated duration of cold events (Figure 5c). The range of the duration of warm spells across simulations yields values of more than 10 days over Mexico and over broad areas of the central and southern US, with a spatial average of 2.8 days (Figure 5c). Otherwise, the LSM effect on the simulated duration of cold spells is weaker, reaching differences of about 6 days among simulations in central Canada with a spatial average of 1.3 days (Figure 5c). For both indices, the LSM differences are larger where the duration indices display larger values (Figure 3 and 5c).

The simulated climatology of the intensity of extreme precipitation events is also strongly affected by the choice of LSM, with the R95p index reaching LSM differences larger than 100 mm at low latitudes and over the eastern US with a spatial average of 39 mm (Figure 6a). The frequency of heavy precipitation events varies among simulations in about 35 days per year at some locations in Mexico and the US, with a spatially averaged range of 3.5 days per year (Figure 6b). The areas with the largest inter-model range of the precipitation frequency index across simulations are located in Mexico, the Rockies and at some grid cells over the eastern US coast (Figure 6b). The simulation of the number of consecutive dry and wet days also depends on the choice of the LSM component, presenting larger differences among simulations in the climatology of the consecutive dry days index than in the climatology of the consecutive wet days index (Figure 6c). The inter-model range across LSM simulations reaches 37 days for the number of consecutive dry days over central and southwestern North America, with a spatial average of 4 days per year (Figure 6c). Meanwhile, the simulated number of consecutive wet days also shows LSM differences of more than 20 days at a few grid cells, but lower values over most of the domain, yielding a spatial average of ~ 1.2 days (Figure 6c). Large inter-model ranges of precipitation indices across WRF simulations coincide with areas where each index reaches the maximum values (Figure 3 and 6).

Results for the VAC metric present some similarities with the spatial pattern of uncertainties in the WRF simulation of temperature and precipitation extreme events. We estimate the seasonal component of each index to compare with the metrics of land-atmosphere coupling seasonally, except for the temperature intensity indices which are defined in JJA and DJF (Figures S8, S9, and S10). The areas showing large uncertainty in the simulation of the intensity indices of cold extremes coincide with areas where LSM simulations differ in the representation of DJF atmospheric control VAC categories (VAC_a and VAC_b ; Figures 1 and 5). The seasonal components of the inter-model range for the simulated percentage of cold days and nights show small uncertainty in winter with very noisy fields (Figure S8). However, the seasonal decomposition of the range for the duration index of cold events shows a region with large uncertainty over western NA in MAM, corresponding to an area with marked differences between LSM simulations in the MAM atmospheric control VAC categories (Figures S8c and 1). For the simulation of warm extremes, large LSM differences in the intensity indices correspond to LSM differences in the JJA VAC categories associated with atmospheric control episodes (Figures 1 and 5). Areas with large uncertainty in the JJA simulation of warm frequency indices coincide with areas showing strong land control on surface processes as well as regional differences between LSM simulations (Figures S9ab and 2). The duration index of warm extremes also shows large inter-model range in JJA over regions under land control (Figures S9c and 2). The range of the intensity index of precipitation extremes displays a large JJA component over areas under land control at low latitudes and under atmospheric control at middle and high latitudes (Figures S10a, 1 and 2). The MAM and SON components of the range for the precipitation intensity index also show large values over

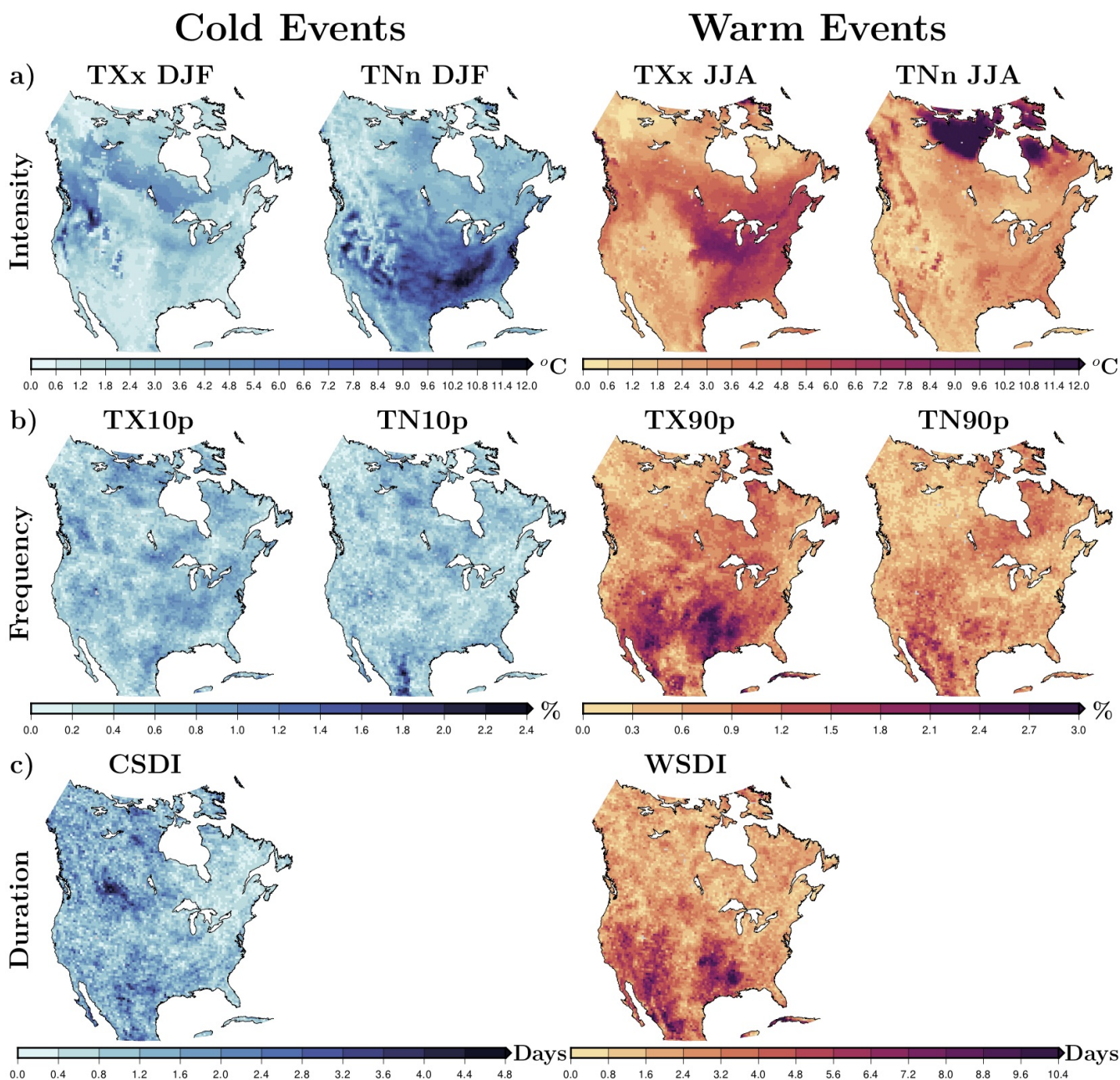


Figure 5. Multi-model ranges across the WRF simulations (i.e., difference between the highest value and the lowest value of the simulation ensemble at each grid cell) of extreme indices associated with the intensity (a), frequency (b), and duration (c) of cold (left) and warm (right) extreme temperature events (Table 2). The range among simulations is computed using the mean of each index from 1980 to 2012 for each simulation.

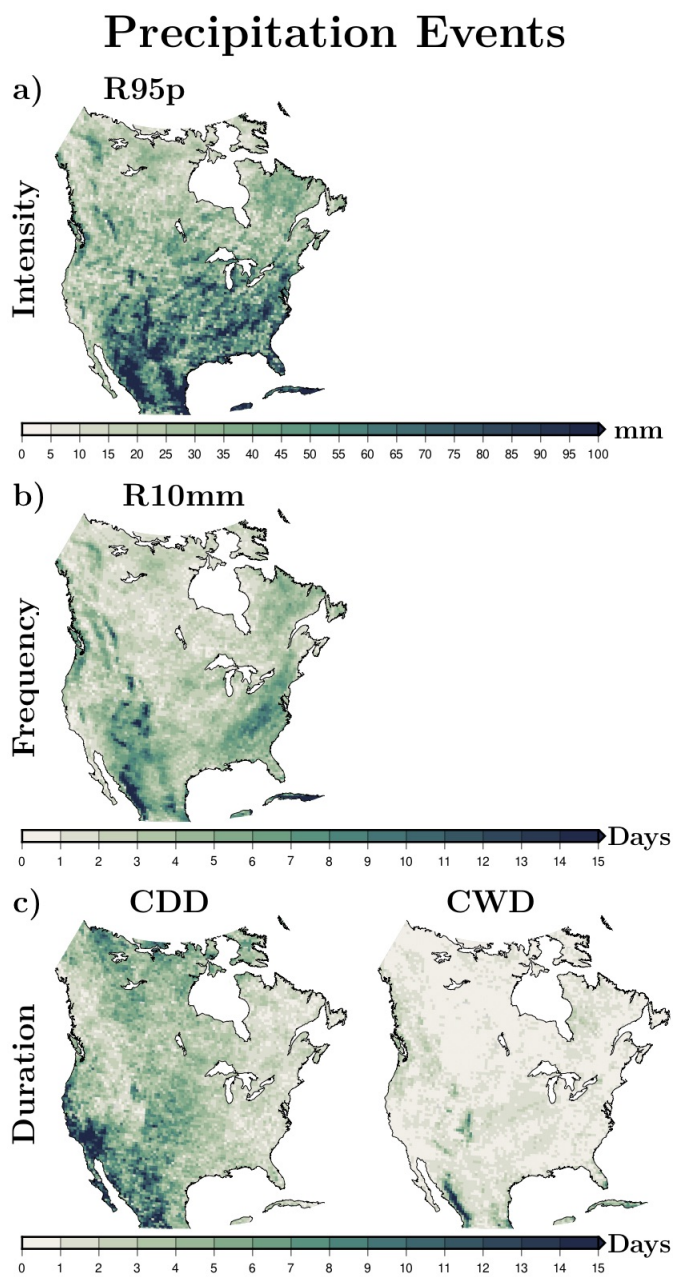


Figure 6. As in Figure 5 but for extreme precipitation events.



310 areas with atmospheric control in MAM and over areas with land control in SON (Figures S10a, 1 and 2). The frequency index of precipitation events presents large inter-model range over small regions in JJA, coinciding with areas under atmospheric control (Figures S10b and 1). The inter-model range of dry periods coincides with land control areas at low latitudes in all seasons and with atmospheric control areas at high latitudes in MAM (Figures S10c 1 and 2). The inter-model range in the simulation of consecutive wet days is large in JJA over a small Mexican region classified under atmospheric control with different degree of coupling between simulations (Figures S10d and 1).

315 In order to address the LSM influence on the simulation of extreme events, we compute the ranges among WRF simulations using the 95th percentile of the analysis period for each extreme index. The uncertainty in the WRF simulations due to the LSM component when using the 95th percentile for each extreme index leads to similar conclusions (Figures S11 and S12). The LSM differences using the 95th percentile of the analysis period are larger for all extreme temperature and precipitation indices than using the period mean as expected, but the marked areas are analogous (Figures 5, 6, S11 and S12). The agreement in the representation of areas with large uncertainty in extreme indices between results using mean and extreme climatologies 320 suggests the LSM influence on extreme events at climatological and shorter time scales.

4.4 Comparison between WRF simulations and three CORDEX Evaluation simulations

The climatologies of temperature and precipitation extreme statistics as simulated by the RCMs participating in the NA-CORDEX project (Table S1) show similar spatial patterns to those from the WRF ensemble (Figures S5-S7 and S13-S15). Although spatial patterns are similar in both ensembles, the WRF simulations yield colder minimum temperatures in DJF 325 (TNn DJF) and less frequent cold nights (TX10p) than the CORDEX simulations (Figures S5 and S13). The percentage of hot days, however, is higher and warm spells are longer in the WRF simulations than in the CORDEX simulations, particularly over southwestern NA (Figures S6 and S14). The intensity of heavy precipitation extremes is generally higher within the WRF ensemble than in the CORDEX ensemble, while dry periods are longer in the CORDEX simulations (Figures S7 and S15).

The uncertainties in the simulation of extreme statistics within the CORDEX ensemble show some similarities with the 330 WRF uncertainties arising from the LSM component. For example, the simulated climatology of DJF coldest night (TNn DJF) shows large uncertainties over the US for both ensembles, particularly over the eastern US (Figures 5a and 7a). The climatologies of DJF hottest day (TXx DJF) display large inter-model range within the WRF ensemble over areas where temperatures approximate to $0^{\circ}C$, expanding southward for the CORDEX ensemble. The CORDEX inter-model ranges of the frequency indices for cold extremes do not show a clear spatial pattern in agreement with the WRF ensemble. There is, 335 however, a region over the central US with slightly larger ranges among the CORDEX simulations than among the WRF simulations (Figures 5b, 7b, and S16b). The duration of cold spells presents large uncertainties in the CORDEX ensemble over the eastern US/Mexican border and over western Canada, coinciding with a small region with large inter-model range among the WRF simulations (Figures 5c and 7c). For the simulation of warm temperature extremes, the uncertainties in the intensity indices among the CORDEX simulations show large ranges over the eastern US for the JJA hottest day (TXx JJA) 340 in agreement with the WRF simulations, and at high latitudes for the coldest night (TNn JJA), including the eastern region of Hudson Bay also marked by the WRF ensemble (Figures 5a and 7a). The frequency indices of warm events show large

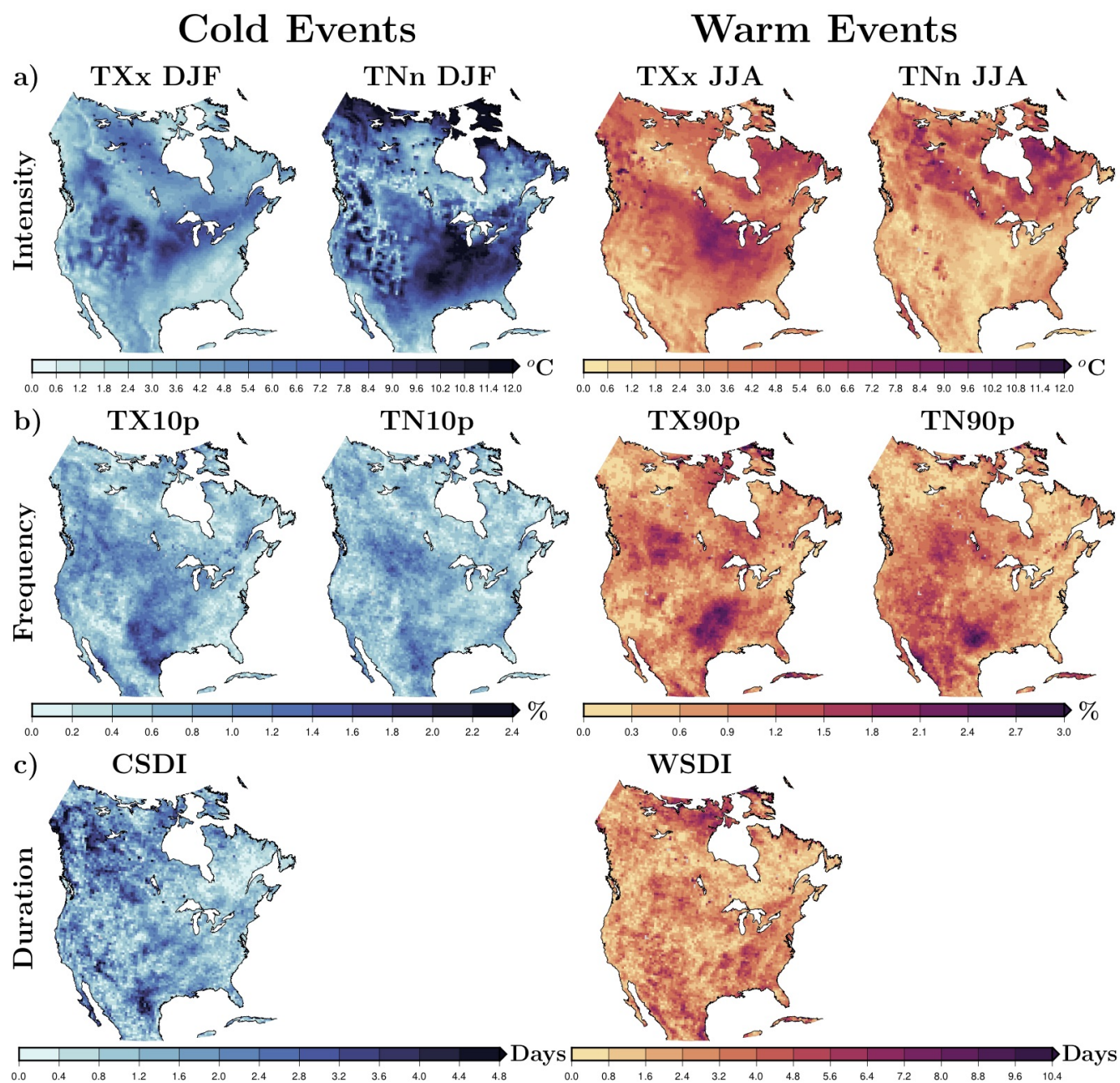


Figure 7. Inter-model range across three CORDEX simulations (i.e., difference between the highest value and the lowest value of the CORDEX ensemble at each grid cell) of extreme indices associated with intensity (a), frequency (b), and duration (c) of cold and warm extreme temperature events (Table 2). The range across simulations is computed using the mean of each index from 1980 to 2012 for each simulation.

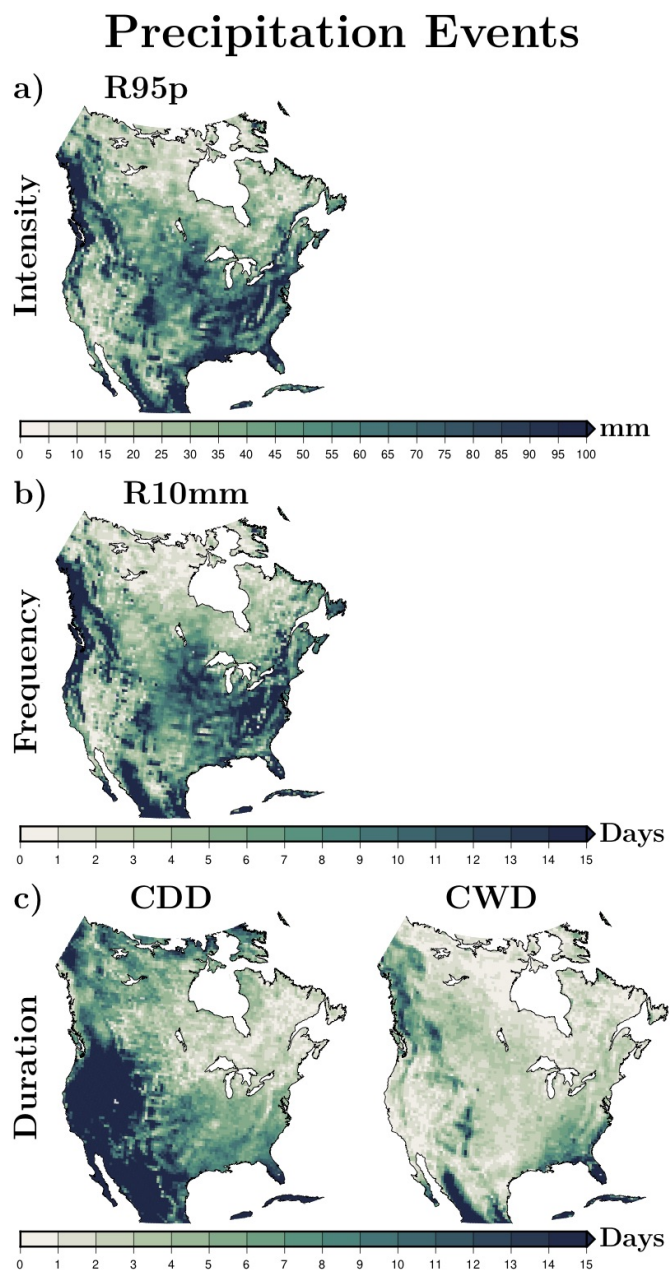


Figure 8. As in Figure 7 but for extreme precipitation events.



inter-model range across the CORDEX simulations over the central US, also shown in the WRF simulations for the TX90p index (Figures 5b and 7b). The uncertainty in the duration of warm spells among the CORDEX simulations does not show large spatial differences, although the ranges are slightly larger at low latitudes coinciding with regions marked by the WRF ensemble and at very high latitudes (Figures 5c and 7c). The simulation of precipitation extreme statistics is generally more uncertain across the CORDEX simulations than across the WRF simulations (Figures 6, 8, and S17). Interestingly, all regions with large uncertainties in the simulation of precipitation extremes among the WRF simulations are also identified as areas with large uncertainty across the CORDEX ensemble. There are, however, additional areas with large uncertainty in the CORDEX ensemble, particularly for the consecutive dry days index and the frequency index at middle and high latitudes (Figures 6 and 8). Thus, the comparison between the WRF and CORDEX ensembles suggests that results from this study may be applicable to other model ensembles over some areas, particularly for the simulation of warm temperature and precipitation extremes.

5 Discussion

5.1 Comparison of inter-model ranges across the WRF and CORDEX ensembles

In order to provide context for the applicability of these results to other model ensembles, we compared the inter-model range across the WRF simulations with the inter-model range across three CORDEX simulations in representing extreme events (Figures 5-8). Since CORDEX simulations were performed by three structurally different RCMs (the WRF, the RCA4, and the CRCM-UQAM models), we expected a broader inter-model range of the simulated extreme indices across CORDEX simulations. Differences in the representation of extreme events among the CORDEX simulations arise from several factors, such as different atmospheric and ocean parameterizations, land surface model components, the representation of land cover, treatment of boundary conditions and the application of nudging techniques. In addition to all these factors, internal variability may be another important component for the inter-model range of the simulated extreme events. However, previous analyses have showed that the spread of extreme events among ensemble members of an individual model is generally small compared to inter-model spreads (Kharin et al., 2007; Sillmann et al., 2013a).

Although CORDEX simulations were performed using boundary conditions from the ERA reanalysis product, the comparison with the WRF simulations is possible because we compute inter-model ranges across ensembles as a measure of the uncertainty in each model ensemble. Thus, we compare model's uncertainty in both ensembles finding common areas with large inter-model ranges for the simulation of cold and warm temperature extremes and precipitation extremes, despite they used different products as boundary conditions. The similar uncertainties of extreme events in the CORDEX ensemble relative to the WRF simulations suggest that the LSM component may be an important source of uncertainty in the CORDEX ensemble. That is, the LSM component employed in each CORDEX simulation (Table S1) may be simulating different land-atmosphere interactions and affecting the simulation of extreme events over those regions.

Despite there being more sources of uncertainty in the CORDEX simulations than across the WRF simulations, the comparison between these ensembles displays larger inter-model ranges across the WRF simulations than across the CORDEX ensemble over certain areas and for certain extreme indices (Figures S16 and S17). This suggests the possible existence of bias



375 compensation inside the CORDEX simulations. Moreover, each RCM may have a different sensitivity to the employed LSM
component as well as to other components and parameterizations. Additional sensitivity studies using the WRF model or an-
other climate model with different settings and parameterizations may help to discern other important sources of uncertainties
in the simulation of extreme events, such as horizontal resolution.

5.2 Climatology of extreme events as represented by the WRF simulations and by the CMIP5 simulations

380 Sillmann et al. (2013a) presented an evaluation of the CMIP5 models in simulating some of the extreme indices defined by
ETCCDI; this information was used in the Intergovernmental Panel on Climate Change (IPCC) chapter on models' evaluation
(Flato et al., 2013). The analysis period employed by Sillmann et al. (2013a), 1981-2000, differs from the one used in this
analysis, but a rough comparison can be done between our results and theirs for some extreme indices. For example, the
spatial patterns of DJF coldest night and JJA hottest day are similar for the WRF and CMIP5 ensemble means (Figure 3 and
385 Figure 2 in Sillmann et al. 2013a). Sillmann et al. (2013a) also provides regional averages over six NA regions, adapted from
Giorgi and Francisco (2000). These spatial averages allow identification of some regional differences between the WRF and
the CMIP5 ensembles, for example over the eastern US coast (ENA region) where the WRF simulations yield warmer JJA
maximum temperatures than the CMIP5 ensemble (Figure 4 and Figure 3 in Sillmann et al. 2013a). The spatial patterns of the
WRF and CMIP5 ensembles for CSDI and WSDI indices are also similar, although the WRF ensemble reaches longer cold
390 and warm events (Figure 3 and Figures S6-S7 in Sillmann et al. 2013a). The representation of the intensity index for heavy
precipitation events (R95p) also shows similar spatial patterns between both ensemble means, although the WRF ensemble is
generally more intense over most regions (Figures 3 and 4, and Figures 6 and 7 in Sillmann et al. 2013a). Similar results are
found for the simulation of consecutive dry days, showing similar spatial patterns with some regional differences especially at
low latitudes (CAM region, Figures 3 and 4, and Figures 6 and 7 in Sillmann et al. 2013a). The variability across the CMIP5
395 ensemble for the simulation of precipitation indices seems to be particularly large at low latitudes (CAM region) similar to
WRF uncertainty in the representation of precipitation extremes associated with the LSM component (Figure 6, and Figure 7
in Sillmann et al. 2013a). Although this is a rough comparison between results presented in this article and in Sillmann et al.
(2013a), this comparison suggests that our conclusions could be also applicable to the CMIP5 ensemble as it was the case for
the CORDEX ensemble.

400 5.3 Implications of these results

Increases in heat-related events have been directly and robustly associated with increases in mortality, for example in Europe
during the heatwave of 2003 (Fischer et al., 2007) or in India (Mazdiyasni et al., 2017). Heavy precipitation events often lead to
floods, which also are directly associated to economic loss and death toll (Hu et al., 2018). All climate change projections point
out to a future increase in temperature and precipitation extreme events (Sillmann et al., 2013b), thus developing mitigation
405 strategies will become necessary to preserve human health. Climate model simulations are our best source of information to
inform measure against climate change impacts. However, the results presented here indicate that the simulation of several
extreme indices varies largely depending on the employed LSM component. This means that a climate model may simulate the



climatology of heat extremes 5°C warmer and 6 days longer depending on the employed LSM component, and similarly for cold extremes and heavy precipitation events. The accuracy of climate models in simulating extreme events will likely affect
410 climate change policy, therefore having repercussions for society and environment.

The indices employed here to study the climatology of extreme temperature events were based on minimum and maximum temperature outputs. However, many studies have proven that the study of compound events using indices based on multiple variables, such as temperature and moisture outputs, are more representative of thermal stress in humans and ecosystems than standard indices (Zscheischler et al., 2018). The large LSM influence on the climatology of extreme temperature and
415 precipitation events, suggests that the uncertainty arising from the LSM component could be higher on extreme indices based on multiple variables. However, the analysis of the LSM influence on compound events is beyond the scope of this work, and constitutes an interesting line for future research.

6 Conclusions

WRF simulations coupled to different LSM components showed similar spatial patterns of land-atmosphere interactions, in-
420 dicative of atmospheric control over surface conditions at middle and high latitudes and land surface control over lower latitudes, particularly in JJA. However, the simulation of land-atmosphere interactions differs at regional scales depending on the LSM choice in two directions; by altering land control on surface processes (VAC_c and VAC_d categories) and by altering atmospheric forcing and its influence on surface conditions (VAC_a and VAC_b categories). Thus, the NOAH LSM is associated with the weakest representation of land control on surface conditions, while the CLM4 LSM simulates one of the strongest land
425 effect on surface conditions. The use of different LSM components leads to large ranges of represented extreme temperature and precipitation events, affecting their simulation in intensity, frequency and duration. The CLM4 LSM yields the weakest cold events, the warmest hot days, and the heaviest precipitation events, while the NOAH simulation yields the weakest land control on surface conditions, the weakest warm temperature events and the weakest heavy precipitation events. This relationship between the degree of land control on surface conditions and the intensity of extreme events is in agreement with two case
430 studies during the Russian 2010 heat wave and the Amazon 2010 drought (Zscheischler et al., 2015). Meanwhile, the NOAH-MP LSM produces the driest simulation, yielding slightly wetter conditions when using dynamic vegetation at middle and low latitudes. Despite small differences between simulations with prescribed and dynamic vegetation, differences are much more marked among the WRF simulations due to different LSM components.

Previous studies using GCM simulations suggested a dependence of the simulated land-atmosphere interactions on the em-
435 ployed LSM component with possible consequences for the simulation of extreme events (García-García et al., 2019). Results from four WRF simulations differing only in the LSM component support that hypothesis. Additionally, areas with large uncertainties in the simulation of temperature and precipitation extremes across the WRF simulations due to different LSM components appear in the NA-CORDEX model ensemble, which indicates the possible LSM influence on the simulation of extreme events within other model ensembles. This work reinforces the important role of the LSM component in climate sim-
440 ulations, supporting the urgency of on-going research focused on improving this model component and their implementation



in regional and global climate models as well as in reanalysis products. The strong LSM dependency of climate model simulation of extremes is also of special importance for international reports focused on land, such as the IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse gas fluxes in Terrestrial Ecosystems (Arnell, 2019). Future sensitivity analyses to the LSM component using different regional and
445 global climate models would be useful to understand models' differences in simulating temperature and precipitation extremes, helping to narrow the inter-model range across reanalyses and climate model projections in simulating extreme events.

Code and data availability. The source code of the Weather Research and Forecasting model (WRF v.3.9) is http://www2.mmm.ucar.edu/wrf/users/download/get_source.html (access date: August, 2017). Extreme indices were calculated using ETCCDI definitions (<https://www.climdex.org/learn/indices/>) and software from the R package `climdex.pcic` (<http://cran.r-project.org/web/packages/climdex.pcic/index.html>,
450 November 2019). The daily temperature and precipitation outputs of all simulations together with the code used to estimate the presented climate extreme indices are available at <https://doi.org/10.5281/zenodo.3745510>. The NARR product was obtained from <https://nomads.nccdc.noaa.gov/data/narr/> (access date: August, 2017). Evaluation simulations from the North American component of the CORDEX project were downloaded from <https://www.earthsystemgrid.org/search/cordexsearch.html> (access date: December, 2018).

Author contributions. AGG designed the modeling experiment, performed the simulations and analyzed model outputs. All authors contributed to the interpretation and discussion of results. AGG wrote the manuscript with continuous feedback from all authors.
455

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. We thank to the Mesoscale and Microscale Meteorology (MMM), the National Center for atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA) for making the WRF code and the NARR product available as well as to the World Climate Research Program's Working Group on Coupled modeling, which oversees CORDEX, and the individual groups and
460 institutions for making their data available. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada Discovery Grant (NSERC DG 140576948), the Nova Scotia Research and Innovation Trust (NSRIT), Compute Canada and The Canadian Foundation for Innovation (CFI) to H. Beltrami. H. Beltrami holds a Canada Research Chair (CRC 230687). A. García-García and F.J. Cuesta-Valero are partially financed by Dr. Beltrami's CRC and Memorial University of Newfoundland.



References

- 465 Arneth, A. e. a., ed.: Summary for Policymakers, pp. 1–43, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, www.ipcc.ch/report/srccl/, 2019.
- Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., and Pappenberger, F.: ERA-Interim/Land: a global land surface reanalysis data set, *Hydrology and Earth System Sciences*, 19, 389–407, 2015.
- 470 Barlage, M., Zeng, X., Wei, H., and Mitchell, K. E.: A global 0.05° maximum albedo dataset of snow-covered land based on MODIS observations, *Geophysical Research Letters*, 32, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL022881>, 2005.
- Cannon, A. J., Sobie, S. R., and Murdock, T. Q.: Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?, *Journal of Climate*, 28, 6938–6959, <https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-14-00754.1>, 2015.
- Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Kiehl, J. T., Briegleb, B., Bitz, C., and Lin, S.-J.: 475 Description of the NCAR community atmosphere model (CAM 3.0), NCAR Tech. Note NCAR/TN-464+ STR, 226, 2004.
- Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., et al.: The community climate system model version 3 (CCSM3), *Journal of Climate*, 19, 2122–2143, 2006.
- Davin, E. L., Maisonnave, E., and Seneviratne, S. I.: Is land surface processes representation a possible weak link in current Regional Climate Models?, *Environmental Research Letters*, 11, 074 027, <http://stacks.iop.org/1748-9326/11/i=7/a=074027>, 2016.
- 480 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, 485 <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828>, 2011.
- Diro, G. T., Sushama, L., Martynov, A., Jeong, D. I., Verseghy, D., and Winger, K.: Land-atmosphere coupling over North America in CRCM5, *Journal of Geophysical Research: Atmospheres*, 119, 11,955–11,972, <https://doi.org/10.1002/2014JD021677>, 2014.
- Diro, G. T., Sushama, L., and Huziy, O.: Snow-atmosphere coupling and its impact on temperature variability and extremes over North America, *Climate Dynamics*, 50, 2993–3007, <https://doi.org/10.1007/s00382-017-3788-5>, 2018.
- 490 Donat, M. G., King, A. D., Overpeck, J. T., Alexander, L. V., Durre, I., and Karoly, D. J.: Extraordinary heat during the 1930s US Dust Bowl and associated large-scale conditions, *Climate Dynamics*, 46, 413–426, <https://doi.org/10.1007/s00382-015-2590-5>, 2016.
- Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", *Hydrol. Earth Syst. Sci.*, 16, 3391–3404, <https://doi.org/10.5194/hess-16-3391-2012>, 2012.
- Fischer, E., Seneviratne, S., Lüthi, D., and Schär, C.: Contribution of land-atmosphere coupling to recent European summer heat waves, 495 *Geophysical Research Letters*, 34, 2007.
- Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, book section 9, pp. 741–866, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, <https://doi.org/10.1017/CBO9781107415324.020>, www.climatechange2013.org, 2013.



- 500 García-García, A., Cuesta-Valero, F. J., Beltrami, H., and Smerdon, J. E.: Characterization of Air and Ground Temperature Relationships within the CMIP5 Historical and Future Climate Simulations, *Journal of Geophysical Research: Atmospheres*, pp. 3903–3929, <https://doi.org/https://doi.org/10.1029/2018JD030117>, 2019.
- Gevaert, A. I., Miralles, D. G., Jeu, R. A. M., Schellekens, J., and Dolman, A. J.: Soil Moisture-Temperature Coupling in a Set of Land Surface Models, *Journal of Geophysical Research: Atmospheres*, 123, 1481–1498, <https://doi.org/10.1002/2017JD027346>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD027346>, 2018.
- 505 Giorgi, F. and Francisco, R.: Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM, *Climate Dynamics*, 16, 169–182, <https://doi.org/10.1007/PL00013733>, 2000.
- Giorgi, F. and Gutowski Jr., W. J.: Regional Dynamical Downscaling and the CORDEX Initiative, *Annual Review of Environment and Resources*, 40, 467–490, <https://doi.org/10.1146/annurev-environ-102014-021217>, <https://doi.org/10.1146/annurev-environ-102014-021217>, 2015.
- 510 Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, *Atmos. Chem. Phys*, 14, 5233–5250, <https://doi.org/https://doi.org/10.5194/acp-14-5233-2014>, 2014, 2014.
- Hauser, M., Orth, R., and Seneviratne, S. I.: Role of soil moisture versus recent climate change for the 2010 heat wave in western Russia, *Geophysical Research Letters*, 43, 2819–2826, <https://doi.org/10.1002/2016GL068036>, 2016.
- 515 Hicks Pries, C. E., Castanha, C., Porras, R. C., and Torn, M. S.: The whole-soil carbon flux in response to warming, *Science*, 355, 1420–1423, <https://science.sciencemag.org/content/sci/355/6332/1420.full.pdf>, 2017.
- Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, *Nature Geosci*, 4, 17–21, <http://dx.doi.org/10.1038/ngeo1032>, 2011.
- 520 Hong, S.-Y. and Lim, J.-O. J.: The WRF single-moment 6-class microphysics scheme (WSM6), *J. Korean Meteor. Soc*, 42, 129–151, 2006.
- Hong, S.-Y., Noh, Y., and Dudhia, J.: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes, *Monthly Weather Review*, 134, 2318–2341, <https://doi.org/10.1175/MWR3199.1>, 2006.
- Hu, P., Zhang, Q., Shi, P., Chen, B., and Fang, J.: Flood-induced mortality across the globe: Spatiotemporal pattern and influencing factors, *Science of The Total Environment*, 643, 171 – 182, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.06.197>, <http://www.sciencedirect.com/science/article/pii/S0048969718322745>, 2018.
- 525 IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, <https://doi.org/10.1017/CBO9781107415324>, www.climatechange2013.org, 2013.
- Jeong, D. I., Sushama, L., Diro, G. T., Khaliq, M. N., Beltrami, H., and Caya, D.: Projected changes to high temperature events for Canada based on a regional climate model ensemble, *Climate Dynamics*, 46, 3163–3180, <https://doi.org/10.1007/s00382-015-2759-y>, 2016.
- 530 Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., and García-Bustamante, E.: A Revised Scheme for the WRF Surface Layer Formulation, *Monthly Weather Review*, 140, 898–918, <https://journals.ametsoc.org/doi/abs/10.1175/MWR-D-11-00056.1>, 2012.
- Karl, T. R., Nicholls, N., and Ghazi, A.: CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary, pp. 3–7, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-015-9265-9_2, 1999.
- 535 Katragkou, E., García-Díez, M., Vautard, R., Sobolowski, S., Zanis, P., Alexandri, G., Cardoso, R. M., Colette, A., Fernandez, J., Gobiet, A., Goergen, K., Karacostas, T., Knist, S., Mayer, S., Soares, P. M. M., Pytharoulis, I., Tegoulis, I., Tsikerdekis, A., and Jacob, D.: Regional



- climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble, *Geoscientific Model Development*, 8, 603–618, <https://doi.org/10.5194/gmd-8-603-2015>, <https://www.geosci-model-dev.net/8/603/2015/>, 2015.
- 540 Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in Temperature and Precipitation Extremes in the IPCC Ensemble of Global Coupled Model Simulations, *Journal of Climate*, 20, 1419–1444, <https://journals.ametsoc.org/doi/abs/10.1175/JCLI4066.1>, 2007.
- Knist, S., Goergen, K., Buonomo, E., Christensen, O. B., Colette, A., Cardoso, R. M., Fealy, R., Fernández, J., García-Díez, M., Jacob, D., Kartsios, S., Katragkou, E., Keuler, K., Mayer, S., van Meijgaard, E., Nikulin, G., Soares, P. M. M., Sobolowski, S., Szepszo, G., Teichmann, C., Vautard, R., Warrach-Sagi, K., Wulfmeyer, V., and Simmer, C.: Land-atmosphere coupling in EURO-CORDEX evaluation
545 experiments, *Journal of Geophysical Research: Atmospheres*, <https://doi.org/10.1002/2016JD025476>, 2016.
- Laguë, M. M., Bonan, G. B., and Swann, A. L. S.: Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and un-coupled land-atmosphere system, *Journal of Climate*, 32, 5725–5744, <https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-18-0812.1>, 2019.
- Li, M., Ma, Z., Gu, H., Yang, Q., and Zheng, Z.: Production of a combined land surface data set and its use to assess land-atmosphere
550 coupling in China, *Journal of Geophysical Research: Atmospheres*, 122, 948–965, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025511>, 2017.
- Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., Chen, F., Chen, L., Clark, M., Dai, A., Dudhia, J., Eidhammer, T., Gochis, D., Gutmann, E., Kurkute, S., Li, Y., Thompson, G., and Yates, D.: Continental-scale convection-permitting modeling of the current and future climate of North America, *Climate Dynamics*, 49, 71–95, <https://doi.org/10.1007/s00382-016-3327-9>, 2017.
- 555 Lorenz, R., Argüeso, D., Donat, M. G., Pitman, A. J., van den Hurk, B., Berg, A., Lawrence, D. M., Chéruy, F., Ducharne, A., Hagemann, S., Meier, A., Milly, P. C. D., and Seneviratne, S. I.: Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble, *Journal of Geophysical Research: Atmospheres*, 121, 607–623, <https://doi.org/10.1002/2015JD024053>, 2016.
- Martynov, A., Laprise, R., Sushama, L., Winger, K., Šeparović, L., and Dugas, B.: Reanalysis-driven climate simulation over CORDEX
560 North America domain using the Canadian Regional Climate Model, version 5: model performance evaluation, *Climate Dynamics*, 41, 2973–3005, <https://doi.org/10.1007/s00382-013-1778-9>, 2013.
- Mazdiyasi, O., AghaKouchak, A., Davis, S. J., Madadgar, S., Mehran, A., Ragno, E., Sadegh, M., Sengupta, A., Ghosh, S., Dhanya, C. T., and Niknejad, M.: Increasing probability of mortality during Indian heat waves, *Science Advances*, 3, e1700066, <https://advances.sciencemag.org/content/advances/3/6/e1700066.full.pdf>, 2017.
- 565 Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E. H., et al.: North American regional reanalysis, *Bulletin of the American Meteorological Society*, 87, 343–360, <https://doi.org/https://doi.org/10.1175/BAMS-87-3-343>, 2006.
- Michalakes, J., Chen, S., Dudhia, J., Hart, L., Klemp, J., Middlecoff, J., and Skamarock, W.: Development of a next generation regional weather research and forecast model, vol. 1, World Scientific, 2001.
- 570 Miralles, D. G., den Berg, M. J., Teuling, A. J., and Jeu, R. A. M.: Soil moisture-temperature coupling: A multiscale observational analysis, *Geophysical Research Letters*, 39, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL053703>, 2012.
- Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *Journal of Geophysical Research: Atmospheres*, 116, 2156–2202, <https://doi.org/10.1029/2010JD015139>,
575 2011.



- Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z.-L., Zeng, X., and Zeng, X.: Technical description of version 4.0 of the Community Land Model (CLM), Tech. rep., NCAR, Boulder, 2010.
- 580 Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., and Takahashi, K.: The JRA-25 reanalysis, *Journal of the Meteorological Society of Japan. Ser. II*, 85, 369–432, 2007.
- Orlowsky, B. and Seneviratne, S. I.: Global changes in extreme events: regional and seasonal dimension, *Climatic Change*, 110, 669–696, <https://doi.org/10.1007/s10584-011-0122-9>, 2012.
- Philip, S. Y., Kew, S. F., Hauser, M., Guillod, B. P., Teuling, A. J., Whan, K., Uhe, P., and Oldenborgh, G. J. v.: Western US high June
585 2015 temperatures and their relation to global warming and soil moisture, *Climate Dynamics*, 50, 2587–2601, <https://doi.org/10.1007/s00382-017-3759-x>, 2018.
- Reichle, R. H., Koster, R. D., Lannoy, G. J. M. D., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A.: Assessment and Enhancement of MERRA Land Surface Hydrology Estimates, *Journal of Climate*, 24, 6322–6338, <https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-10-05033.1>, 2011.
- 590 Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, 85, 381–394, <https://journals.ametsoc.org/doi/abs/10.1175/BAMS-85-3-381>, 2004.
- Samuelsson, P., Jones, C. G., Willén, U., Ullerstig, A., Gollivik, S., Hansson, U., Jansson, C., Kjellström, E., Nikulin, G., and Wyser, K.: The Rossby Centre Regional Climate model RCA3: model description and performance, *Tellus A*, 63, 4–23, <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.2010.00478.x>, 2011.
- 595 Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., and Rahimi, M.: Changes in climate extremes and their impacts on the natural physical environment, in: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, pp. 109–203, Cambridge University Press, 2012.
- Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part
600 1. Model evaluation in the present climate, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50203>, 2013a.
- Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, *Journal of Geophysical Research: Atmospheres*, 118, 2473–2493, <https://doi.org/10.1002/jgrd.50188>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50188>, 2013b.
- Sippel, S., Zscheischler, J., Mahecha, M. D., Orth, R., Reichstein, M., Vogel, M., and Seneviratne, S. I.: Refining multi-model pro-
605 jections of temperature extremes by evaluation against land–atmosphere coupling diagnostics, *Earth System Dynamics*, 8, 387–403, <https://doi.org/10.5194/esd-8-387-2017>, <https://www.earth-syst-dynam.net/8/387/2017/>, 2017.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, NCAR Technical Note, Tech. rep., National Center for Atmospheric Research, Boulder, Colorado, USA, 2008.
- 610 Stieglitz, M. and Smerdon, J. E.: Characterizing land-atmosphere coupling and the implications for subsurface thermodynamics, *Journal of climate*, 20, 21–37, <https://doi.org/https://doi.org/10.1175/JCLI3982.1>, 2007.



- Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R.: Implementation and verification of the unified NOAH land surface model in the WRF model, 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, pp. 11-15, 2004.
- 615 Vertenstein, M., Craig, T., Middleton, A., Feddema, D., and Fischer, C.: CESM1. 0.4 user's guide, UCAR Doc, 2012.
- Vogel, M. M., Orth, R., Cheruy, F., Hagemann, S., Lorenz, R., Hurk, B. J. J. M., and Seneviratne, S. I.: Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks, *Geophysical Research Letters*, 44, 1511–1519, <https://doi.org/10.1002/2016GL071235>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL071235>, 2017.
- Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., and
620 Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, *Journal of Geophysical Research: Atmospheres*, 116, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015140>, 2011.
- Zscheischler, J., Orth, R., and Seneviratne, S. I.: A submonthly database for detecting changes in vegetation-atmosphere coupling, *Geophysical Research Letters*, 42, 9816–9824, <https://doi.org/10.1002/2015GL066563>, 2015GL066563, 2015.
- Zscheischler, J., Westra, S., van den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch,
625 D. N., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, *Nature Climate Change*, 8, 469–477, <https://doi.org/10.1038/s41558-018-0156-3>, 2018.