Objective Calibration of Regional Climate Models: Application over Europe and North America

OMAR BELLPRAT

Institut Catala de Ciènces del Clima (IC3), Barcelona, Spain, and Institute for Atmospheric and Climate Science (IAC), ETH Zurich, Zurich, Switzerland

SVEN KOTLARSKI AND DANIEL LÜTHI

Institute for Atmospheric and Climate Science (IAC), ETH Zurich, Zurich, Switzerland

RAMÓN DE ELÍA AND ANNE FRIGON

Consortium sur la Climatologie Régionale et L'Adaptation aux Changement Climatiques (OURANOS), Montréal, Québec, Canada

RENÉ LAPRISE

Centre Étude et Simulation du Climat à l'Échelle Régionale, Montréal, Québec, Canada

CHRISTOPH SCHÄR

Institute for Atmospheric and Climate Science (IAC), ETH Zurich, Zurich, Switzerland

(Manuscript received 27 April 2015, in final form 8 November 2015)

ABSTRACT

An important source of model uncertainty in climate models arises from unconfined model parameters in physical parameterizations. These parameters are commonly estimated on the basis of manual adjustments (expert tuning), which carries the risk of overtuning the parameters for a specific climate region or time period. This issue is particularly germane in the case of regional climate models (RCMs), which are often developed and used in one or a few geographical regions only. This study addresses the role of objective parameter calibration in this context. Using a previously developed objective calibration methodology, an RCM is calibrated over two regions (Europe and North America) and is used to investigate the transferability of the results. A total of eight different model parameters are calibrated, using a metamodel to account for parameter interactions. The study demonstrates that the calibration is effective in reducing model biases in both domains. For Europe, this concerns in particular a pronounced reduction of the summer warm bias and the associated overestimation of interannual temperature variability that have persisted through previous expert tuning efforts and are common in many global and regional climate models. The key process responsible for this improvement is an increased hydraulic conductivity. Higher hydraulic conductivity increases the water availability at the land surface and leads to increased evaporative cooling, stronger low cloud formation, and associated reduced incoming shortwave radiation. The calibrated parameter values are found to be almost identical for both domains; that is, the parameter calibration is transferable between the two regions. This is a promising result and indicates that models may be more universal than previously considered.

1. Introduction

Information about the climate at regional scales has become a major need for society and policy makers (Christensen and Christensen 2007). A sophisticated

Corresponding author address: Omar Bellprat, Institut Catala de Ciènces del Clima (IC3), Doctor Trueta 203 3a, 08005 Barcelona, Spain.
E-mail: omar.bellprat@ic3.cat

DOI: 10.1175/JCLI-D-15-0302.1

way to determine this information at local scales is to use regional climate models (RCMs) that allow downscaling of climate information available at the global scale at high spatial resolution for a specific region (Laprise 2008). Comprehensive ensembles of RCM simulations are available for several continents, in particular over Europe (PRUDENCE and ENSEMBLES; Christensen and Christensen 2007; van der Linden and Mitchell 2009), South America [Climate Change Assessment and Impact Studies (CLARIS; Menéndez et al. 2010)], the United States [Project to Intercompare Regional Climate Simulations (PIRCS; Takle et al. 1999)], North America (NARCCAP; Mearns et al. 2012), the Arctic [Arctic Regional Climate Model Intercomparison (ARCMIP; Curry and Lynch 2002)], Asia [Regional Climate Model Intercomparison Project (RMIP) for Asia (Fu et al. 2005)], while for some continents RCM information is very limited. Within the ongoing Coordinated Regional Climate Downscaling Experiment (CORDEX; Giorgi et al. 2009), this gap is now filled by simulating all continents on the globe with ensembles of different RCMs.

The new direction of globally coordinated regional downscaling requires different regional models to simulate the climate over regions for which they have not previously been applied. This opens up the question of whether RCMs are transferable to other climate regions, since for most models the development and tuning have been constrained on a specific region of origin. The transferability is particularly challenging, as the specification of model parameters is a major source of model uncertainty (Allen 1999; Knutti et al. 2002; Murphy et al. 2004; Stainforth et al. 2005; Yokohata et al. 2010; Sanderson 2011; Klocke et al. 2011; Bellprat et al. 2012a).

The transferability of RCMs has been studied previously in the Inter-CSE (continental-scale experiment) Transferability Study (ICTS; Takle et al. 2007) as part of the Global Energy and Water Cycle Experiment (GEWEX). The experiments demonstrated that RCMs are able to simulate different climates outside of the domain for which the models have historically been developed (native domains), yet model biases are typically larger over domains not previously considered (nonnative domains) (Rockel et al. 2008; Rockel and Geyer 2008; Jacob et al. 2012; Mearns et al. 2012).

As a result Giorgi et al. (2012) argue that customization exercises (tuning of model parameters) need to be conducted in order to reduce model biases over nonnative domains. This implies that model parameters should be allowed to change depending on the climate considered, which is (depending on the model parameter considered) not always evident from physical principles. On the other hand, retuning of parameters for different climates might lead to overtuning, as raised by Jacob et al. (2012). Such an overtuning of parameters to compensate for structural deficiencies poses considerable risks, as climate change projections might be altered by parameter configurations (Murphy et al. 2007). Overtuning of regional climate models has thus been a concern of the Working Group on Numerical Experimentation (WGNE) of the World Climate Research Programme (WRCP) (Laprise et al. 2008). The reasonable transferability of RCMs found by ICTS has been considered as an argument against the criticism of overtuning of RCMs, yet the lack of objective approaches to determine model parameters has so far hindered the discussion on how to deal with model parameters related to both RCMs and GCMs.

Tuning of model parameters is presently still widely performed manually based on expert knowledge, and often without following a predefined strategy. Documentation on how climate models are being tuned is scarce (but see, e.g., Mauritsen et al. 2012). An objective and transparent methodology to calibrate model parameters in RCMs has been presented in Bellprat et al. (2012b, hereafter B12). The approach relies on a calibration procedure developed for global climate models (Neelin et al. 2010; Bracco et al. 2013). The main idea of the approach is to approximate the model response resulting from parameter perturbations using a computationally efficient statistical regression model (metamodel) that is estimated on the basis of a minimum set of model simulations. The estimated metamodel is consequently evaluated in terms of prediction accuracy using independent simulations of the climate model and is used to sample the parameter space for optimal model configurations. In this way large parameter spaces become tractable, since the metamodel is computationally efficient and hence millions of parameter experiments can be conducted. Using an automated approach reduces the risk of compensating errors, as it considers simultaneously a large number model parameters, observational datasets, and sources of uncertainty.

Here we use the same methodology and apply it over two different model domains. The aim of this study is to independently and objectively calibrate an RCM over Europe and North America. This will address the question of whether the model formulation can be considered universal, or whether different parameter configurations are objectively justifiable. These questions are also relevant when addressing the potential danger of model overtuning. By extending the number of parameters in comparison to B12 (from five to eight model parameters), we further aim to reduce the remaining model biases.

The structure of the paper is as follows: Section 2 provides a description of the RCM setup for both model

domains and a summary of the calibration framework. Section 3 shows the calibration results and discusses some of the important underlying processes. The results are discussed and conclusions are made in section 4.

2. Methods

a. *Regional model approach: Europe and North America*

We simulate the climate over the European and North American continents using the regional climate model Consortium for Small-Scale Modeling (COSMO) Model in Climate Mode [CCLM (Rockel et al. 2008)] following the procedures established in the CORDEX framework (Giorgi et al. 2009). CCLM is a climate modeling system relying on the numerical weather prediction model COSMO (Steppeler et al. 2003), and it has been used extensively to study the European climate at a wide range of resolutions and for a number of purposes (e.g., Suklitsch et al. 2008; Zubler et al. 2011; Kotlarski et al. 2012; Ban et al. 2014). Therefore, we entitle the European domain (EU) as the native domain of the model and the North American domain (NA) as the nonnative domain.

As in previous transferability studies (comparison of RCM simulations over a native and nonnative domains; Takle et al. 2007), the same model physics are used for both domains as schematically illustrated in Fig. 1 [reference configuration (REF) for both domains]. The comparison of these two simulations allows the assessment of the commonalities and differences of model biases between the two continents. Subsequently, a model parameter calibration is performed independently for each domain to obtain an optimized configuration (OPT) for the same set of parameters. The comparison of the determined optimal parameter distributions allows the assessment of how the calibration of the RCM depends upon the model domain.

The basic setup of the model follows B12. We use a horizontal resolution of 0.44° (approximately 50 km) and 32 atmospheric layers in the vertical. The model is used with its standard suite of parameterizations (see B12 for details). The simulations cover the period 1994–98 for the calibration and 1990–2008 for the validation, including the years of the calibration. Sea surface temperatures and prognostic atmospheric variables at the lateral boundaries are prescribed by reanalysis data from ERA-Interim (Dee et al. 2011), using a relaxation zone with a width of 12 grid points over each lateral boundary.

The two computational domains with topography and analysis domains are shown in Fig. 2. The two domains over EU and NA consist of 109 × 121 and 175 × 150 grid

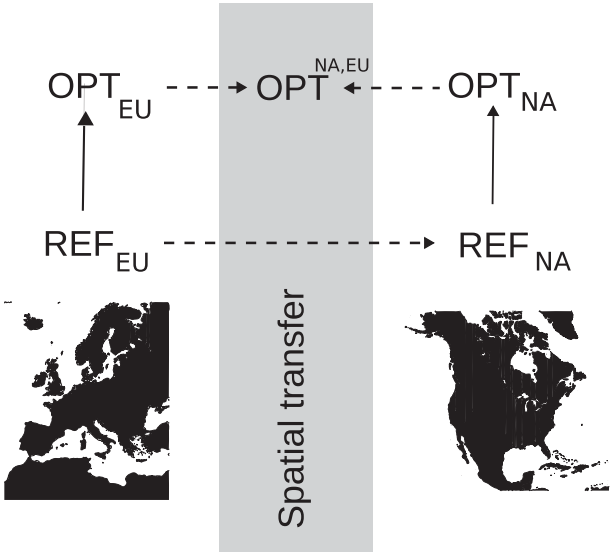

FIG. 1. Calibration approach studying the spatial transferability of an RCM for the native domain (Europe) and nonnative domain (North America). Both domains are first simulated using default parameter settings (REF EU and REF NA). Based on the reference simulation, a parameter calibration is performed in both domains independently (OPT EU and OPT NA). The analysis includes an assessment of the models' improvement due to calibration and a comparison of the optimal parameter settings found in the two different domains.

points in the horizontal, respectively. The black frames or outlines in Fig. 2 show analysis regions, for which spatial averages are considered in the calibration. The NA domain is approximately 2.5 times larger than EU and therefore contains more regions, which results in approximately equally sized regions over both domains. It is important to note that the consideration of a larger domain for NA implies a larger error growth, since the influence of the boundary conditions weakens with increasing domain size. As a result, the magnitude of internal variability is expected to be larger over NA, purely due to the model configuration (Alexandru et al. 2007). Further, the North American domain spans a wider band of latitudes reaching tropical zones that entail weather regimes not occurring over Europe, such as the formation of tropical cyclones and the North American monsoon (Adams and Comrie 1997; Bukovsky et al. 2015).

b. *Validation and performance measure*

The simulations for both domains are validated against observations of 2-m temperature (T2M), total precipitation (PR), and total cloud cover (CLCT) using the datasets listed in Table 1. For each variable we use one reference dataset and additional datasets to estimate the observational uncertainty (three for Europe

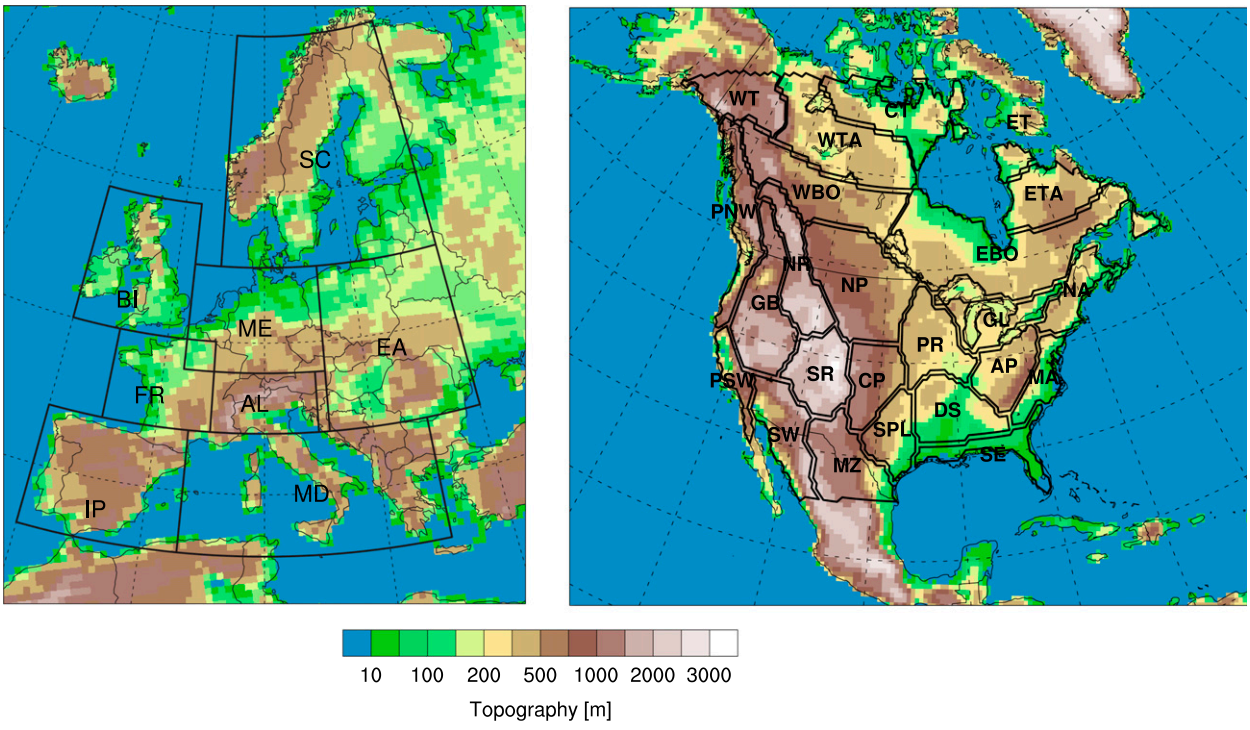

FIG. 2. Model domains and topographies for Europe and North America at a resolution of 0.44°. The area (number of grid points) of the North American domain is approximately 2.5 times larger than the European domain. (left) The analysis domains (black lines) for Europe are from the PRUDENCE project (Christensen and Christensen 2007): Iberian Peninsula (IP), France (FR), British Isles (BI), mid-Europe (ME), Alps (AL), Mediterranean region (MD), eastern Europe (EA), and Scandinavia (SC). (right) For the North America analysis, the domains presented in Bukovsky (2011) are used: Appalachia (AP), central plains (CP), central tundra (CT), Deep South (DS), eastern boreal (EBO), eastern taiga (ETA), eastern tundra (ET), Great Basin (GB), Great Lakes (GL), Mezquital Valley region (MZ), mid-Atlantic region (MA), northern plains (NP), northern Rockies (NR), North Atlantic region (NA), Pacific Northwest (PNW), Pacific southwest region (PSW), prairie (PR), southern Rockies (SR), U.S Southeast (SE), southwest region (SW), western boreal (WBO), western taiga (WTA), and western tundra (WT).

and two for North America). An integrated measure that summarizes the agreement between the model and the observations is defined. This measure is maximized by the calibration. For this purpose we employ an integrated performance score (PS; Bellprat et al. 2012a) that is an extension of the climate prediction index (CPI) defined in Murphy et al. (2004). PS computes the average-squared errors between a model m and the observations o for each validated variable ($V = 3$; T2M, PR, and CLCT), each month of the year ($T = 12$) for all simulated years ($Y = 5$), and for each regional average (regions shown in Fig. 2, where $R = 8$ regions for Europe and $R = 24$ regions for North America). These errors are scaled by the observed interannual variability σ_o,

TABLE 1. Observational datasets used for validation purposes over Europe and North America. The datasets denoted with a region specification (EU or NA) define the reference dataset. The remaining observations are used to estimate the observational uncertainty. (Expansions of acronyms are available at http://www.ametsoc.org/PubsAcronymList.)

Dataset	Variables	Resolution	Period	Reference
E-OBS (v.7)	Temperature (EU) Precipitation (EU)	0.44°	1950–2013	Haylock et al. (2008)
CRU Time Series (v3.1)	Temperature (NA) Precipitation (NA) Cloud cover (EU and NA)	0.5°	1901–2009	Harris et al. (2013)
University of Delaware (v3.01)	Temperature Precipitation	0.5°	1900–2006	Willmott and Matsuura (2009)
ISCCP flux data	Cloud cover	2.5°	1983–2008	Zhang et al. (2004)
HIRS	Cloud cover	1°	1989–2010	Wylie et al. (2005)

TABLE 2. Selected model parameters and short description of the involved processes. Some parameters have a unit denoted in parentheses. The default values are shown in boldface, accompanied with a chosen uncertainty range.

Acronym	Parameter or property	Value and range
rlam_heat	Scalar resistance for the latent and sensible heat fluxes in the laminar surface layer	**1**, [0.1, 5]
entr_sc	Entrainment rate for shallow convection (m^{-3})	$\mathbf{3 \times 10^{-4}}$, $[3 \times 10^{-5}, 3 \times 10^{-3}]$
qi0	Threshold for conversion of cloud ice to snow	**0**, $[0, 10^{-4}]$
uc1	Parameter controlling the vertical variation of critical relative humidity for subgrid cloud formation	**0.8**, [0, 1.6]
root_dp	Uniform factor for the root depth field	**1**, [0.5, 1.5]
tkhmin	Minimum vertical turbulent diffusion rate ($m^2\,s^{-1}$)	**1**, [0.1, 2]
radfac	Fraction of cloud water and ice considered by the radiation scheme	**0.5**, [0.3, 0.9]
soilhyd	Factor for the hydraulic conductivity and diffusivity	**1**, [1, 6]

allowing the computation of an average across all errors. The scaling of the errors is further extended in comparison to Murphy et al. (2004) by adding a measure of the observational uncertainty σ_ε and the internal variability σ_{iv}. In cases where the observational uncertainty or the internal noise of the model is large, the errors are downweighted. The final score is formed using an exponential transformation of the average-squared error, which scales the PS between 0 and 1 (1 being the highest score):

$$\mathrm{PS} = \exp\left[-0.5 \cdot \frac{1}{VRTY} \sum_{v}^{V} \sum_{r}^{R} \sum_{t}^{T} \sum_{y}^{Y} \frac{(m_{v,r,t,y} - o_{v,r,t,y})^2}{(\sigma_{o_{v,r,t}} + \sigma_{\mathrm{iv}_{v,r,t,y}} + \sigma_{\varepsilon_{v,r,t,y}})^2}\right]. \quad (1)$$

The total score is formed by a total of $VRTY = 1440$ and 4320 squared errors for Europe and North America, respectively, and thus considers much more spatial and temporal dimensions than a standard validation using seasonal mean biases. The total scaling of the errors in (1) is formed by adding the standard deviations instead of the variances to be consistent with B12; however, the effect of adding the squared terms has been found to be negligible.

c. Parameter calibration

Using the objective calibration methodology described in B12, an independent calibration of uncertain model parameters over both domains is conducted. A short description of the methodology is provided here, yet for more details and full equations the reader is referred to the original publication. The current application is an extension of the study presented in B12, where we considered five independent parameters. The current application considers three additional parameters for the calibration over Europe and conducts the same calibration over North America. The final list of parameters calibrated in this study with the respective predetermined uncertainty ranges are shown in Table 2 (tkhmin, soilhyd, and radfac are the additional parameters to those found in B12).

The selected parameters affect a wide range of important processes in the model: shallow convection (entr_sc), subgrid-scale cloud formation (uc1), interaction of clouds with radiation (radfac), autoconversion of cloud ice (qi0), turbulent transport of heat and moisture (tkhmin), exchange of heat and moisture between the atmosphere and the land surface (rlam_heat), strength of transpiration of the vegetation related to depth of rooting zone (root_dp), and hydraulic cycling of soil moisture (soilhyd). The selection of these parameters is based on sensitivity studies as presented in Bellprat et al. (2012a).

The calibration approach relies on a statistical approximation of the climate model (i.e., metamodel) that predicts the response of the model to parameter configurations. This statistical approximation is computationally much more efficient than the physical model, which allows the sampling of large numbers of parameter configurations and thereby determining optimal parameter values. The estimation of the metamodel requires RCM simulations that sample the edges and the center of the eight-dimensional parameter space. These simulations are restricted to a short period of five years (1994–98 for EU and 1992–96 for NA), which has proven to be sufficiently long to reach equilibrium in the model performance score (Bellprat et al. 2012a). Initialization of these simulations takes place on 1 January, from the long-term simulation (REF, 1990–2008).

The chosen metamodel is a multivariate quadratic regression model proposed by Neelin et al. (2010) that

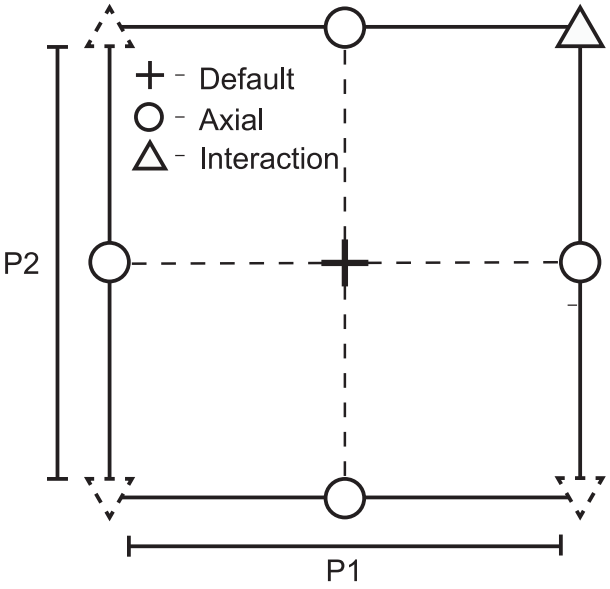

FIG. 3. Illustration of design points required in order to estimate the metamodel. The center is given by the REF simulation using default parameter values. For each parameter a min and a max value needs to be simulated to sample the borders (axial points) of each pairwise parameter plane. The interaction of the two parameters is sampled by one of the four corner points.

has been applied to coarse-resolution global model simulations (Bracco et al. 2013) and high-resolution regional model simulations over Europe (B12). For both types of applications, the metamodel proved to reproduce with high accuracy the response of the climate model when model parameters were altered. The number of model simulations required to estimate the metamodel is small, which makes this emulator suitable for computationally demanding climate models such as RCMs. More specifically, for each parameter two simulations using a minimum and maximum value have to be carried out to sample the boarders of the multidimensional parameter space (see Fig. 3 for illustration of simulation design). Parameter interactions are accounted for by changing two parameter values to either the minimum value or the maximum value at the same time for all possible parameter pairs. This gives a total number of $2N + N(N - 1)/2 = 44$ simulations, each 5-yr long, which are required to estimate the metamodel with eight model parameters ($N = 8$, the number of parameters considered).

The interactions of parameters can hence be sampled with four different experiments capturing all four corners in the pairwise plane. To increase the accuracy of the parameter interaction terms, additional simulations accounting for all different combinations have been carried out as described in B12. This leads to an additional 84 simulations for the eight parameters considered in this study. The effect of the parameter interactions is, however, small (B12; Bracco et al. 2013) and thus additional simulations have only been carried out for the calibration of EU.

Finally, one million parameter configurations are evaluated with the metamodel to determine the optimal parameter configuration. The parameter configurations are sampled using a Latin hypercube design (McKay et al. 2000; Gregoire et al. 2011). The verification of the calibration is based on a long RCM simulation using the OPT settings. This simulation spans the same period as REF (1990–2008). It includes the 5-yr calibration period but also 14 additional and independent years.

3. Results

a. *Calibration results*

1) EUROPE

We describe in this section the calibration results by comparing REF, which has not previously been calibrated using an objective approach (Rockel et al. 2008), with the calibrated simulation (OPT) over both continents. The simulation OPT is based on the calibration framework over Europe and North America (section 1) using the two optimal parameter configurations that have been determined. The two settings will be compared in section 3b. The mean seasonal biases of REF and OPT are shown in Figs. 4 and 5 for Europe and Figs. 6 and 7 for North America, and the corresponding seasonal biases in interannual temperature variability are shown in Figs. 8 and 9, respectively. The biases are related to the magnitude of biases simulated by other RCMs and, when reported in the literature they are accompanied with suggested reasons leading to these biases.

The REF simulations over Europe show a large warm bias in summer over the Mediterranean region, eastern Europe, and the Iberian Peninsula (for regions refer to definitions in Fig. 2). Suggested reasons are diverse, although they have mainly been discussed in the context of biases in land surface coupling (e.g., Rowell and Jones 2006; Vidale et al. 2007; Bellprat et al. 2013; Seneviratne et al. 2013). The overestimation of temperature is accompanied by an underestimation of total precipitation and cloud cover as shown in the middle and bottom panels of Fig. 4 as well as by a moisture deficit in the soils as discussed in Fischer et al. (2007). The correlation of the pattern of these biases illustrates the complex interactions of processes involved and disentangling these has been the focus of several recent studies (Fischer et al. 2007; Jaeger et al. 2008; Sutton et al. 2007; Davin et al. 2011; Cattiaux et al. 2013; Boé and Terray 2014).

A large fraction of the summer temperature and precipitation biases is reduced over the Mediterranean

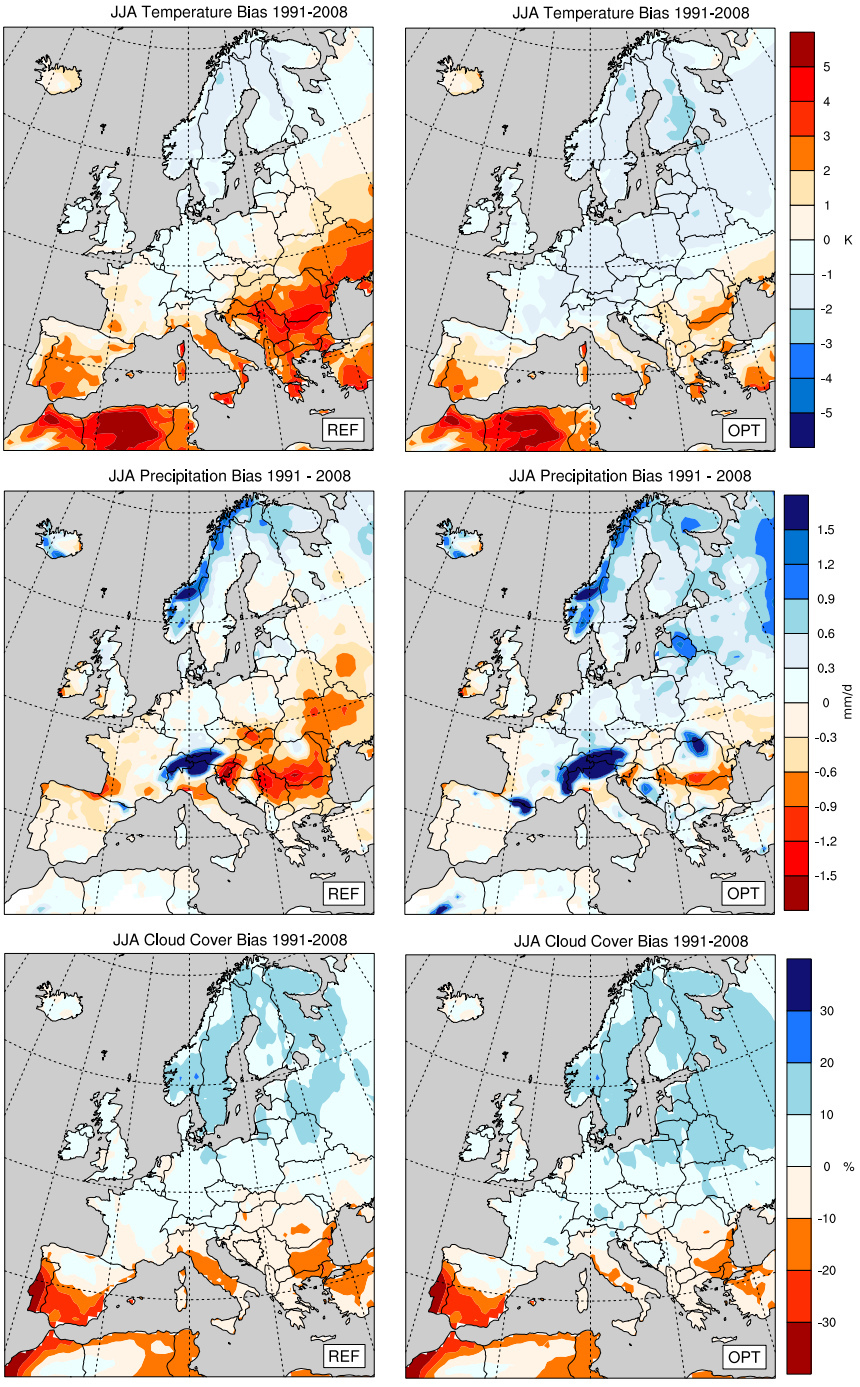

FIG. 4. Mean summer [June–August (JJA)] biases for the simulations (left) REF and (right) OPT for the period 1991–2008 over Europe. The biases are shown for (top) temperature, (middle) precipitation, and (bottom) total cloud cover.

region and eastern Europe in the calibrated simulation (OPT). This improved representation of the summer climate to a colder and moister state is a notable result, as it has persisted previous expert tuning efforts and remains prominent in the majority of global and regional climate models over semiarid continental regions (Vidale et al. 2007; Christensen et al. 2008; Mearns et al. 2012; Cattiaux et al. 2013; Bellprat et al. 2013; Kotlarski et al. 2014; Mueller and Seneviratne 2014). The achieved improvement is hence of wider

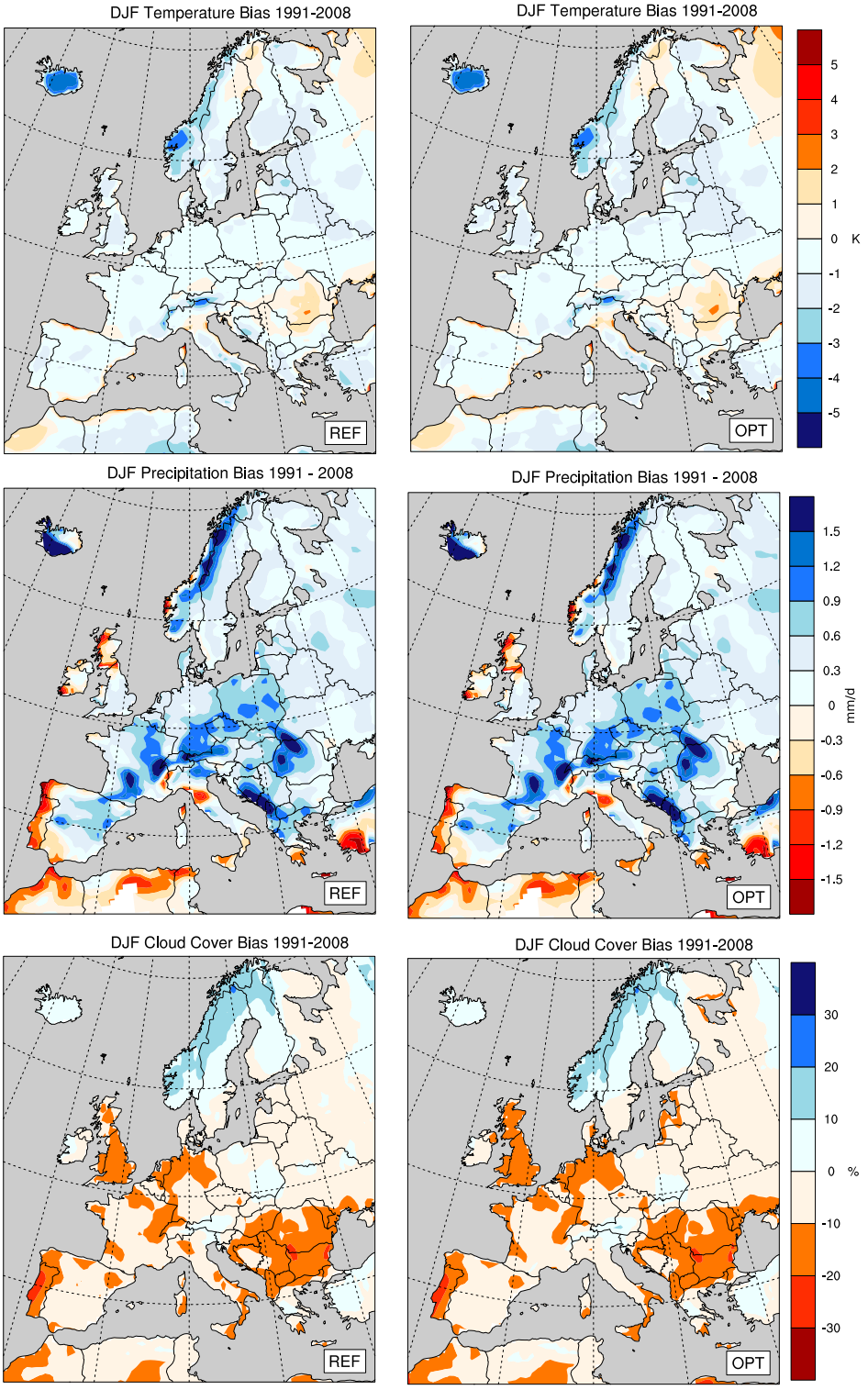

FIG. 5. As in Fig. 4, but for mean winter [December–February (DJF)].

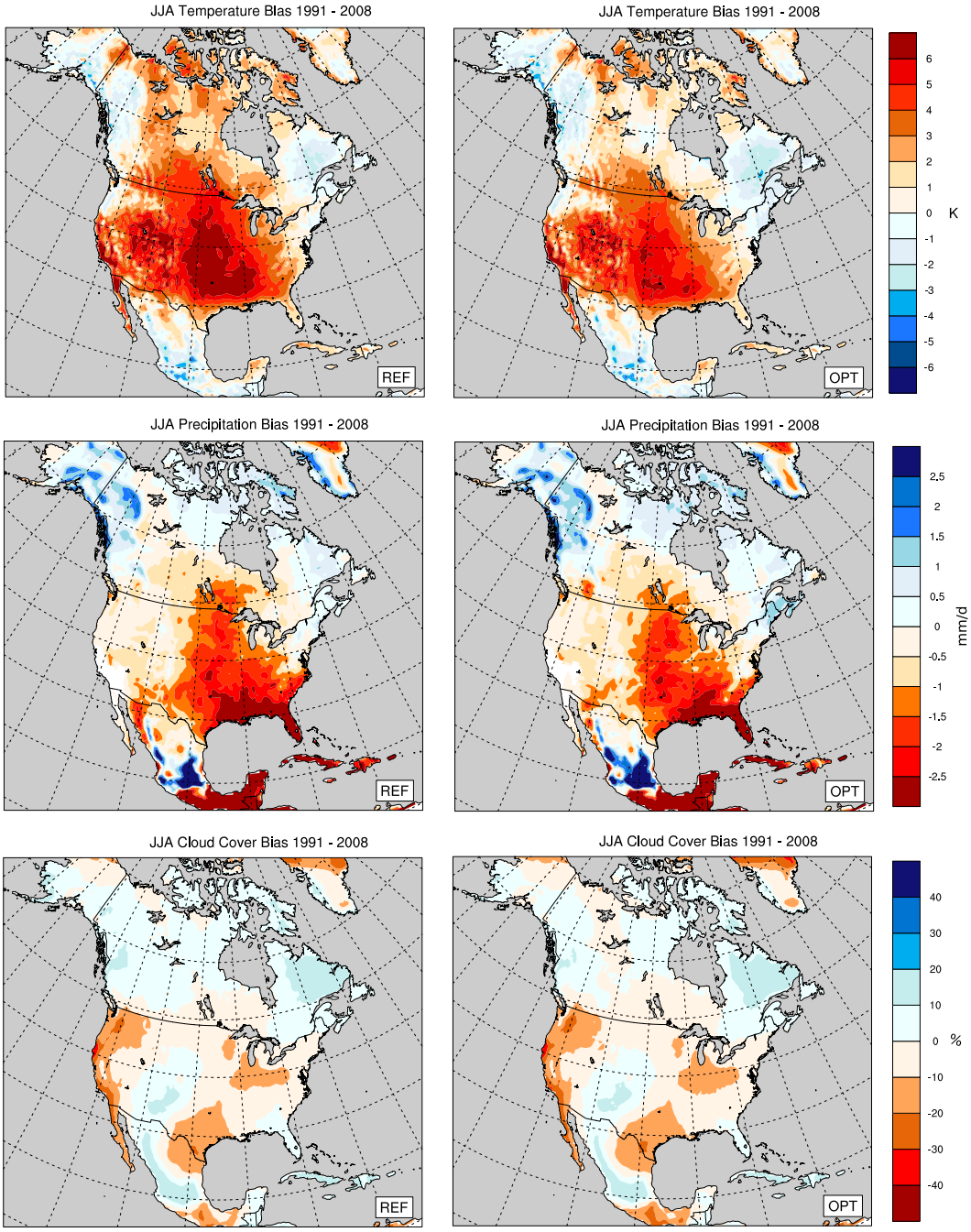

FIG. 6. As in Fig. 4, but for North America and a different observational reference dataset (see Table 1).

interest and is discussed in sections 3c and 3d in more detail.

The too warm mean summer conditions in REF are accompanied by an overestimation of the interannual summer variability (IASV) of temperature shown over large parts of the domain in Fig. 8. The highest bias occurs over eastern Europe, where the variability is overestimated by approximately 100% (1 K). This overestimation of IASV, again common to many models over semiarid regions (Vidale et al. 2007; Fischer et al. 2012), improves strongly in OPT—particularly over eastern Europe, where the variability is heavily overestimated in REF but also over the Mediterranean region, where the variability is reproduced much more realistically.

The overestimation of simulated mean seasonal temperature and its interannual variability is of significant

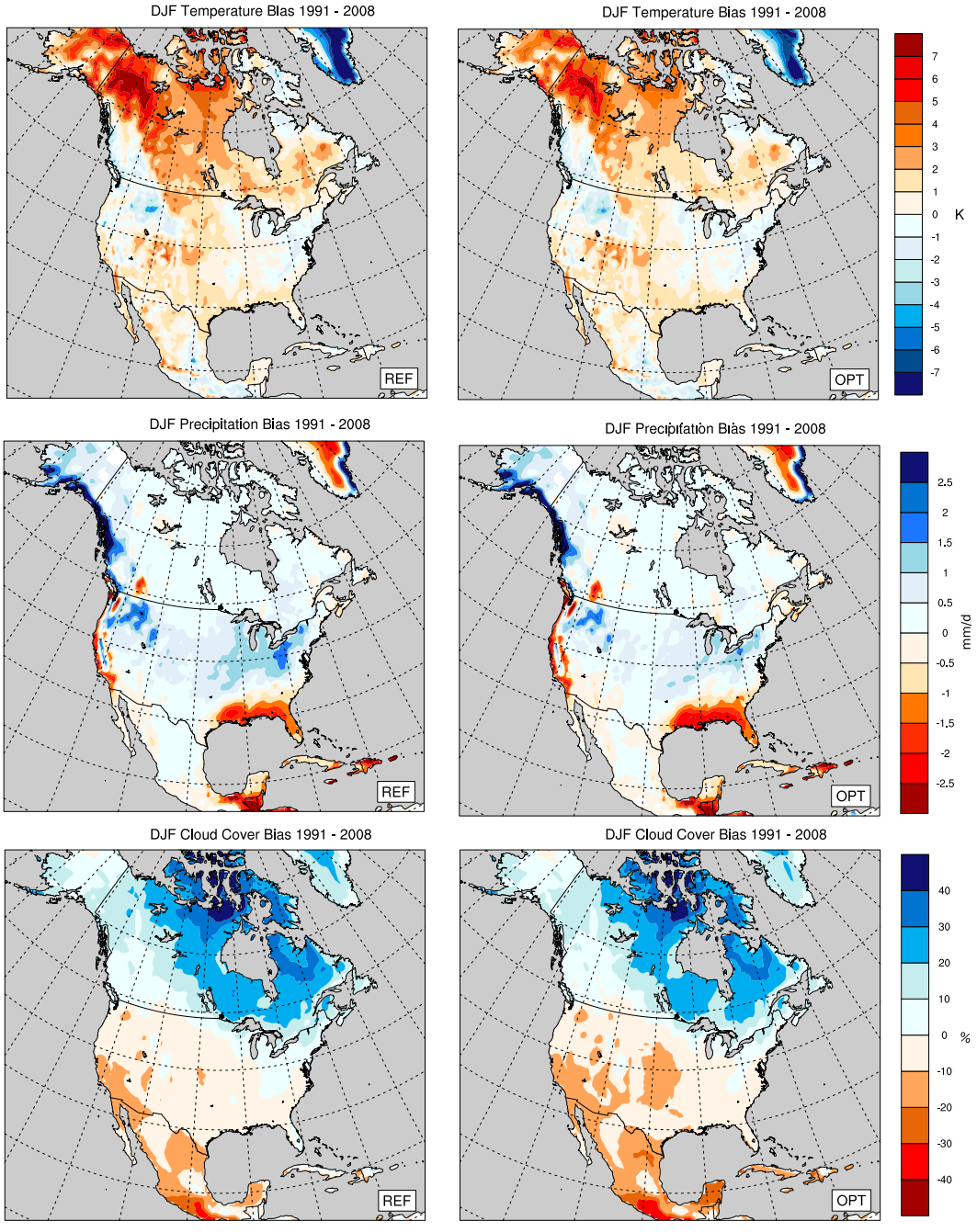

FIG. 7. As in Fig. 4, but for mean winter (DJF) over North America and a different observational reference dataset (see Table 1).

concern in relation to climate change projections (Schär et al. 2004). It has been shown that summer temperature biases are accentuated in warmer, drier climates and that climate models may therefore overestimate the projected warming (Boberg and Christensen 2012). The biases reach an upper limit due to constraints of soil moisture depletion (Bellprat et al. 2013) and a linear bias assumption hence leads to unphysical conditions. The achieved reduction of summer temperature biases in the control period is therefore of high importance to reduce uncertainties of temperature projections associated with nonstationarities of the biases.

In winter the temperature biases are much smaller than in summer and only enhanced biases in regions of

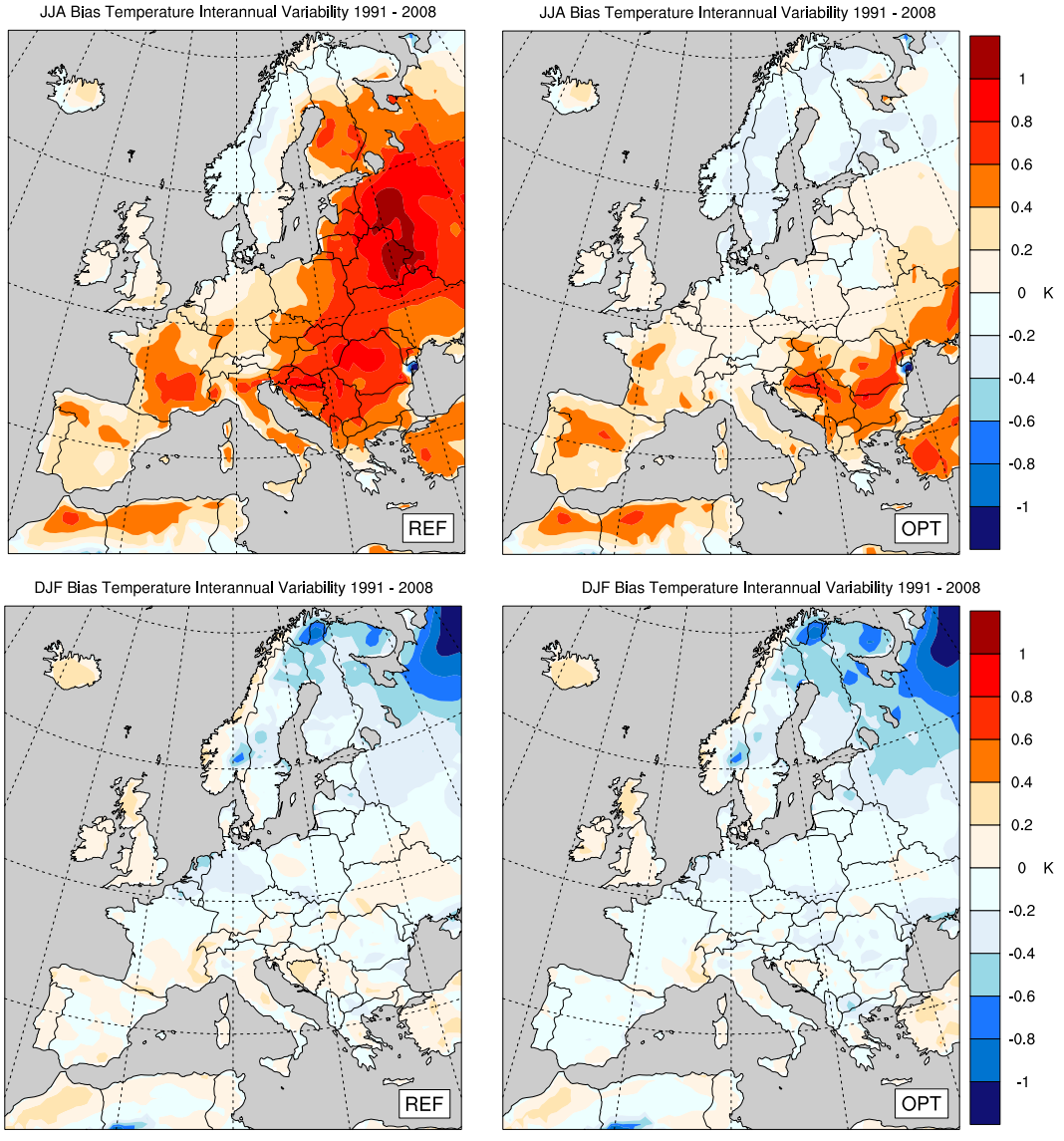

FIG. 8. Bias in (top) summer (JJA) and (bottom) winter (DJF) temperature interannual variability for the simulations (left) REF and (right) OPT for the period 1991–2008 over Europe.

high topography are present. The model further overestimates precipitation over a large part of the domain in winter, which is also apparent in the majority of RCMs over Europe (Kotlarski et al. 2014). The biases in the winter season remain almost unaffected by the calibration over Europe for both the mean and the interannual variability. The temperature bias is already small in winter and the remaining biases might result from missing a height correction to the model data. The interannual variability in winter is much larger than in summer and is captured accurately in REF and OPT, as it is largely determined by the lateral boundary forcing for the European domain (Lüthi et al. 1996). Similar to summer, the total cloud cover is underestimated over southern Europe and overestimated over northern Europe in both REF and OPT, which causes biases in the incoming shortwave radiation (Jaeger et al. 2008).

Biases of winter precipitation remain large in both simulations as in the majority of RCMs. The enhanced bias has so far not been discussed in the context of individual drivers. Some of the bias results from the systematic undercatch of true precipitation by rain gauges, particularly in mountain areas (Kotlarski et al. 2014). A further possible explanation for the overestimation of precipitation in the models might result from a seasonal balance of compensating biases. Unduly large precipitation amounts in winter increase the infiltration of water into the soil and therefore reduce the drying of the soils in summer, which is a main

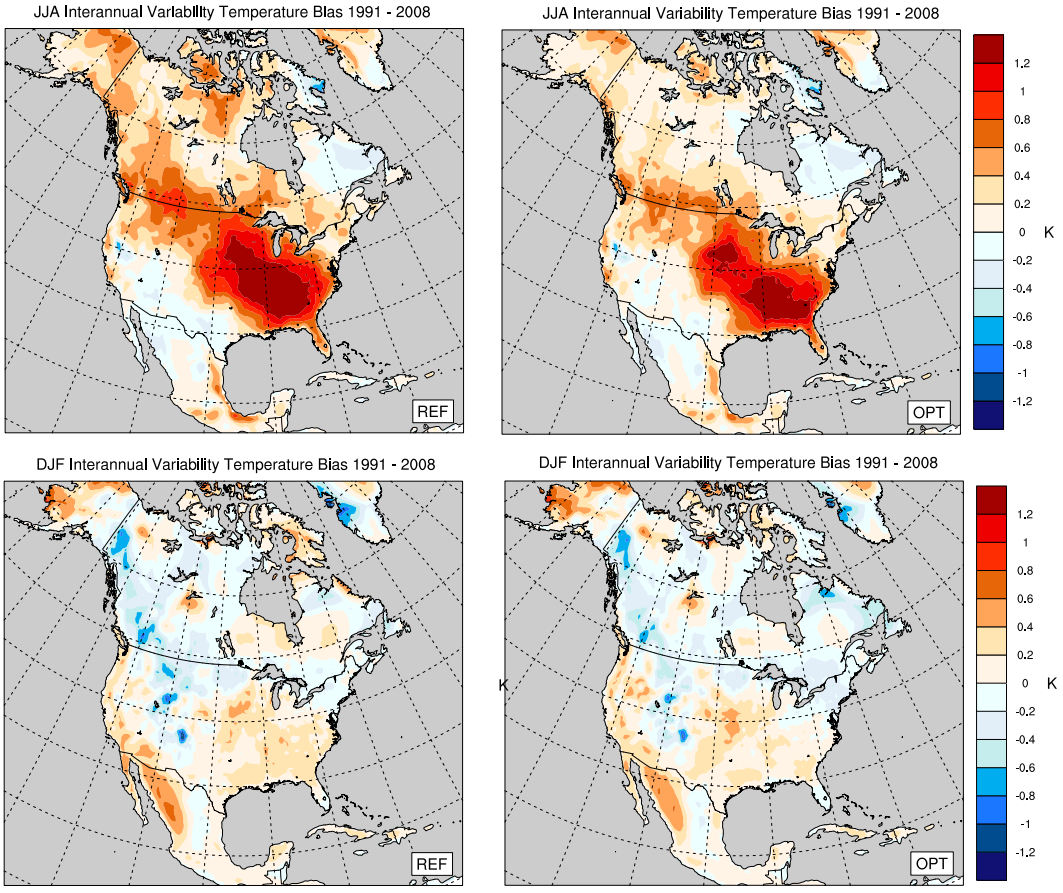

FIG. 9. As in Fig. 8, but for North America and a different observational reference dataset (see Table 1).

driver for hot temperature biases in CCLM. Considering only the winter season in the calibration yields a different optimal configuration that leads to a fundamental reduction of the winter precipitation bias according to the metamodel (not shown). However, this finding does not demonstrate a seasonal dependence on the model parameters. The required setup to prove a seasonal variation would require two parameter values for each parameter, one in winter and in summer, that are calibrated simultaneously (given the seasonal interplay as, e.g., in land surface coupling). Calibrating these two parameters might result in similar optimal values, as identified for the different model domains, even though the optimal values differ when calibrated for each respective season independently. This aspect of temporal "transferability" is not further explored here and should be placed in a physical context that supports a temporal dependence.

2) NORTH AMERICA

Summer temperatures over North America (Fig. 6) are strongly overestimated over semiarid regions, such as the Great Plains (CP and NP), the prairies, and the Deep South (geographic names refer to the regions defined in Fig. 2), but also over high latitudes, such as over the western boreal and western tundra regions. Central North America is a region of strong land surface coupling (Koster et al. 2004) and temperature biases in these regions have been related to erroneous simulation of land surface fluxes (Klein et al. 2006; Mueller and Seneviratne 2014). Other possible origins of the dry bias have been reported in relation to the formation of the low-level jet, which is an important source of moisture in central North America (Helfand and Schubert 1995) and a trigger of convection (Arritt et al. 1997; Seth and Giorgi 1998). An incorrect simulation of this moisture transport could result in a lack of moisture transport into central North American regions, leading to a too dry climate. Consistent with this picture, precipitation and cloud amounts are underestimated in central North America. The structure of this hot and dry bias is common among RCMs and GCMs over North America (Mearns et al. 2012; Mueller and Seneviratne 2014).

The identified summer biases over North America decrease for temperature and precipitation in OPT.

However, a large part of the bias remains. This indicates that, for North America, either neglected processes in the calibration or structural deficiencies need to be addressed in order to achieve validation results comparable to those over Europe. This is also evident in the bias of the interannual variability in summer (Fig. 9), which decreases in OPT but remains large over the Great Plains and Florida.

Winter temperature biases over North America (Fig. 7) are smaller compared to summer temperature biases, yet they are overestimated over large parts of the domain, particularly in high-latitudes. The model generally overestimates precipitation in winter, yet the dry bias for precipitation over Florida identified for summer remains. The bias in cloud cover shows again a latitudinal gradient which is yet more pronounced over the Boreal and Arctic regions. The winter biases over North America reduce in winter as shown over high latitudes in Fig. 7. The overestimation of the Boreal temperature is reduced by 1–2 K, yet remains particularly over Alaska and the Arctic. There is some indication for a reduction of precipitation biases south of the Great Lakes and a small increase in cloud cover bias over the Great Plains. The biases remain generally similar as in REF.

The origin of the winter biases has not been investigated in the model and hence the mechanisms remain unknown. It is worth noting that the spatial density of observational stations is low in high latitudes over North America (New et al. 2000) and thus biases should be considered with caution. The interannual temperature variability (Fig. 9) is strongly overestimated over the Great Plains and also over coastal areas near Florida. In winter the model captures the interannual variability very accurately and outperforms many other RCMs simulating the North American climate (Mearns et al. 2012).

The overall magnitude of the biases is larger over North America compared to Europe, which could be a result of either modeling a larger domain or deficiencies of the model to simulate a different climate. Generally, the model biases lie within the range of biases identified in other RCMs for both Europe and North America (Kotlarski et al. 2014; Mearns et al. 2012), although a particular deficiency of CCLM is the overly hot and dry climate in summer.

b. *Optimal model parameters*

In this section we discuss the changes in model parameter settings identified by the calibration. We show for this purpose in Fig. 10 the posterior parameter distributions (range of parameter values, blue histograms) for both domains in comparison to the default values (red vertical lines). The parameter ranges are constructed by sampling all parameter configurations that yield the best performance according to the metamodel and lie within an uncertainty estimate of the metamodel determined by B12.

The optimal parameter configurations (dashed black vertical lines in Fig. 10) for both domains agree very well, and so do the width and shape of the density of the distributions. As evident from Fig. 10, there is no indication that the calibration over tunes the selected parameters for a specific domain given the similarity of the regional parameter distributions. This risk has been raised in relation to CORDEX (Jacob et al. 2012), and it has also been argued that climate change projections could be affected by parameter uncertainty (Murphy et al. 2007). The lack of evident overtuning in our calibration exercise is a remarkable result, considering the large number of parameters that are calibrated. The result thus supports the idea that biases inherent to the definitions of the model physics are corrected by model calibration, in ways that are at least partly universal and independent of the region considered.

It is of interest to conduct a comparison against B12, where only the first five parameters have been considered in an otherwise identical calibration over Europe. Note that the optimal values for these five parameters remain virtually unchanged. The consideration of three additional parameters did not alter the settings found in B12, even though two of the additional parameters change notably in OPT (compared to their default values). This is a convenient but most likely not a general result, which might partly arise from a weak interaction between the parameters as they act in different physical parameterizations. However, the contribution of the parameter interactions is not negligible. Excluding the interactions in the metamodel increases the error to predict independent model simulations (10 simulations with random parameter values) by approximately 20%.

For both domains, the calibration finds higher values for the amount of cloud water seen by the radiation (radfac from default of 0.5 to 0.7) and much higher values for the hydraulic conductivity (soilhyd from default of 1 to 6). The increase in the interaction between clouds and radiation implies a 40% increase in the cloud water and ice amount that interacts with the radiation. Considering the experiment where only radfac is increased, we find a net reduction of the surface energy balance due to reduced incoming shortwave radiation. This cooling is partly balanced by reduced outgoing longwave radiation.

c. *The role of the hydraulic conductivity*

The substantial increase of the soil hydraulic conductivity deserves a more detailed analysis, as the respective adjustments are most significant for the

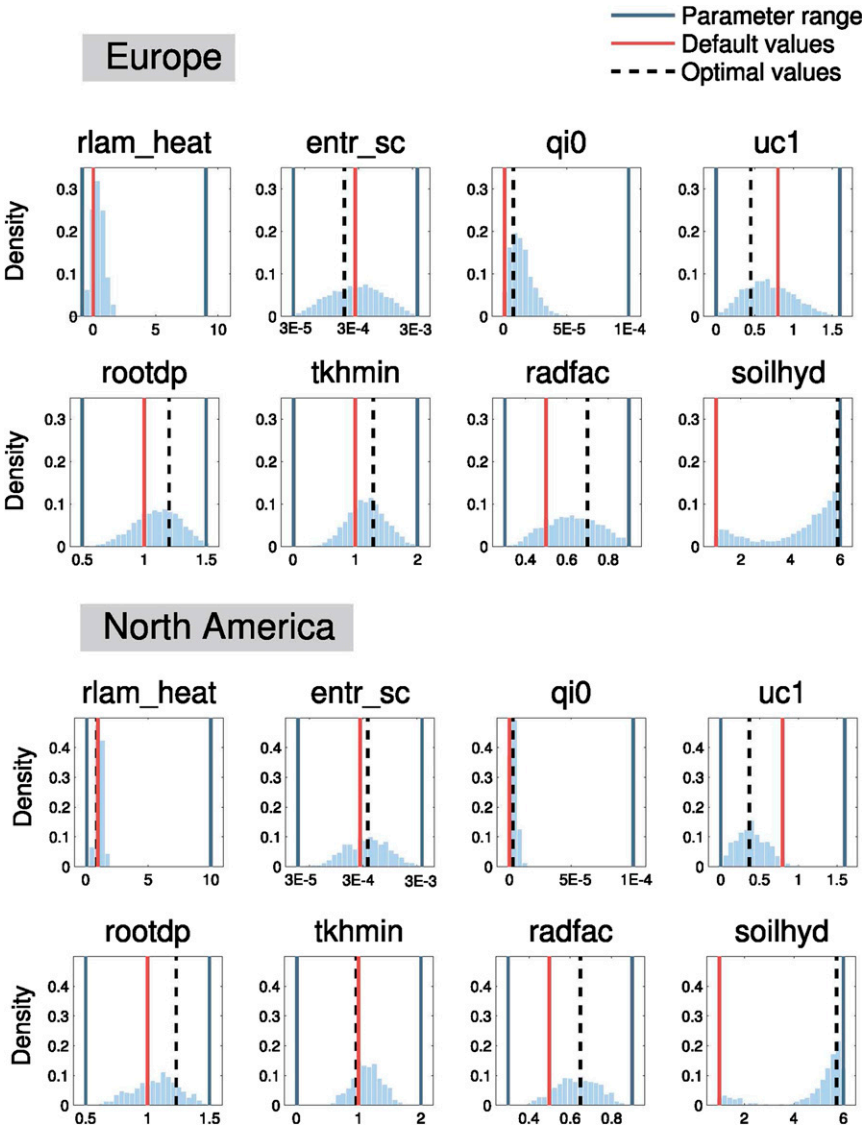

FIG. 10. Empirical densities (blue histograms) of the calibrated parameter values, which perform equally well, given the uncertainty of the metamodel in predicting the model performance. The calibration results for the eight calibrated parameters are shown for (top) Europe and (bottom) North America. In each panel, the dark blue lines show the parameter uncertainty range, the red line the default parameter value (REF), and the black dashed line the parameter combination of the best-performing simulation (OPT).

performance of the model—from all parameter changes considered. More specifically the calibration process yields a substantial increase (by about a factor of 6) in both model domains. This leads to higher availability of soil moisture in the deeper soil layers and thereby increases the evaporative fraction. Before discussing these aspects in the next subsection in more detail, we here address potential interpretations behind such an increase of hydraulic conductivity. In particular, can it be justified in terms of process understanding?

The hydraulic conductivity in CCLM is specified depending upon soil type and soil water content, and varies by several orders of magnitude (Doms et al. 2007). More specifically, the saturated hydraulic conductivity covers a wide range, from 4.7×10^{-5} to $1.7 \times 10^{-8}\,\mathrm{m\,s^{-1}}$, between sand and peat soils. For unsaturated soils the conductivity quickly drops by many orders of magnitude with decreasing relative soil water content. A comparison against the ECMWF Integrated Forecast System (IFS) shows that the hydraulic

conductivities are slightly larger in CCLM but generally of a similar magnitude (Cloke et al. 2011). In comparison to the represented variability, an increase of the hydraulic conductivity by about half an order of magnitude, as proposed by our calibration, is thus a relatively small perturbation.

The conductivity values used in our and other models originate from a collection of measurements from the literature (e.g., Rijtema 1969; van Genuchten 1980). Are these values, which are largely based on laboratory measurements using soil columns, representative of the grid boxes of climate models? As in most weather and climate models, the grid-scale conductivities are represented according to the most common soil type, and no consideration is given to subgrid-scale variability.

A possible interpretation for increasing the hydraulic conductivities comes from the fact that the hydraulic conductivity is a scale-dependent quantity that increases with horizontal scale. Information on this scaling comes from a multitude of field measurements that have effective horizontal scales of typically 1–10 or 100–500 m. Using data from a number of field studies, Rovey and Cherkauer (1995) found that the effective hydraulic conductivities applicable to horizontal scales of $O(100)$ m can be two orders of magnitude larger than the laboratory values obtained from soil column experiments.

The interpretation of this result relies on the heterogeneity of soils, which may include spatial variations in soil type, micro- and macroporosity, subsurface aquifers, etc. Heterogeneity may lead to a much higher hydraulic transport than expected from laboratory measurements of the hydraulic conductivity of the dominating soil type (e.g., Gelhar 1986). In essence, the effective hydraulic conductivity valid at the grid scale of the model (i.e., 10^4–10^5 m) cannot be inferred from laboratory measurement of the most common soil type. Rather a comparatively small soil fraction with high conductivity may imply a higher effective conductivity at the grid scale where the hydrological processes are explicitly represented. Recently, objective methodologies to upscale small-scale heterogeneity effects have been explored (Samaniego et al. 2010), but exploitation of these results for the current application are beyond the scope of the current study.

It is further important to note that, in contrast to all other parameters, the posterior distribution of the hydraulic conductivities is bimodal (Fig. 10). The calibration yields two optimal states of the parameter, one close to the default and another at the upper bound with higher density. This is an indication that the specified calibration range is too narrow. To check on this potential difficulty, we have repeated both calibrations using an extended calibration range ($1 \leq \text{soilhyd} \leq 10$) by extrapolating the simulated limit with the metamodel and finding only very minor changes to the optimized parameter.

d. Detailed analysis of European summer climate

The contribution to each parameter on this improvement can be studied using the metamodel as the model result can be decomposed into each parameter and each parameter interaction term. Using the metamodel we find that approximately 60% of the cooling of summer temperatures and 80% of the increased precipitation are caused by an increase in hydraulic conductivity. The remaining cooling can be attributed to an enhanced interaction of clouds and radiation, as well as an increase in depth of the rooting zone.

For a better interpretation of these results, we analyze the changes in the soil water balance for the Mediterranean region, where the strongest changes occur in OPT compared to REF by illustrating the change in the soil water distribution of the first seven hydraulic soil layers in Fig. 11. The figure shows the difference between OPT and REF of the relative soil moisture (soil moisture divided by the field capacity) in the course of the year. In OPT the surface layers are less saturated in winter and moister in summer by about 10% of the field capacity. The lower layers moisten throughout the entire year, which leads to an increase in the relative soil moisture of about 20%.

To explain these changes, the effect of the hydraulic conductivity on the vertical transport of unsaturated soils needs be considered. This vertical transport is described by Richards equation, which quantifies the changes in soil water content Θ due to a vertical gradient of the water potential $\delta\Psi/\delta z$ scaled by the hydraulic conductivity K_w,

$$\frac{\delta\Theta}{\delta t} = \frac{1}{\delta z}\left[K_w\left(\frac{\delta\Psi}{\delta z} + 1\right)\right], \tag{2}$$

with the vertical axis z pointing upward, and where $-K_w(\delta\Psi/\delta z + 1)$ represents the vertical flux of water. In the case where the vertical gradient of the water potential is small, the near-surface soil water content decreases because of gravity-induced downward flux into lower levels. When the gradient is strongly negative (i.e., if the water content strongly increases with depth), the near-surface soil water content increases as a result of capillary forces that induce an upward (diffusive) flux of soil water. The hydraulic conductivity affects both these processes.

Next, we investigate how the increase in hydraulic conductivity resulting from calibration affects the mean seasonal cycle of soil moisture, by considering the

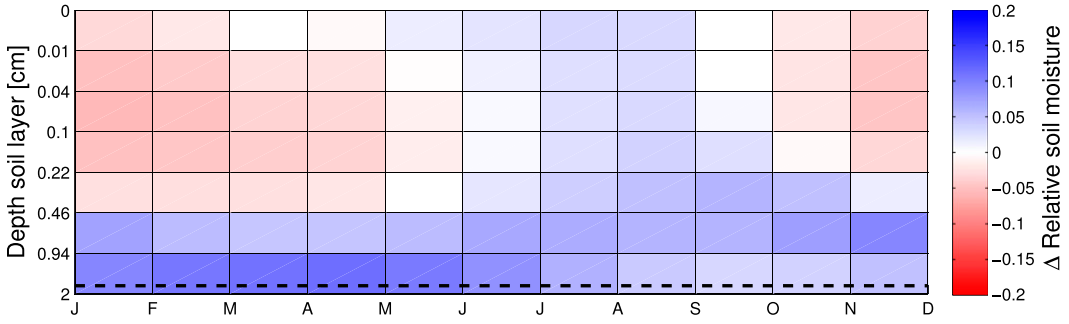

FIG. 11. Difference in monthly (period 1991–2008) relative soil moisture content (relative with respect to field capacity; right-hand axis) for each layer of the soil as simulated by the calibrated simulation minus the reference simulation (OPT − REF) over MD. The rows from top to bottom show the seven first hydraulically active soil layers with increasing depth of the layer (left-hand axis). The dashed line shows the depth of the rooting zone.

difference OPT − REF (Fig. 11). Overall, the gravitational flux is increased with higher K_w, yielding higher relative soil moisture content in the deep soil. In winter this increase is at the expense of soil moisture in the upper soil layers. In summer, when the upper soil layers dry out, increased hydraulic diffusivity leads to an enhanced upward flux of soil moisture through capillary forces, which increases the soil water content in the upper layers relative to reference conditions.

The overall soil water content is increased as a consequence of increased infiltration of precipitated shown in Fig. 12. The infiltration rate increases linearly with the hydraulic conductivity according to Darcy's law in unsaturated soils and thus enhanced infiltration is expected in OPT. The solid line shows the annual cycle of the water content integrated over 2 m of the soil for OPT compared to REF (dashed line) and the interannual variability of OPT is shown by the gray band. The figure shows that the total water content is increased in OPT by about 20% for the entire annual cycle. Although this change might seem large, it is small in comparison to the range of model-based observations driven by the Water and Global Change (WATCH) Forcing Data ERA-Interim (WFDEI) meteorological forcing (Weedon et al. 2014). The hydraulic conductivity is hence an important factor for controlling the vertical distribution of soil moisture and also the total amount of soil water.

The modification of the soil moisture balance has important consequences for summer climate due to surface–atmosphere coupling (Vidale et al. 2007; Fischer et al. 2007; Seneviratne et al. 2010). The changes in the surface water balance and the boundary layer are illustrated in Fig. 13, showing the change in low cloud cover, in the incoming shortwave radiation, and in the evaporative fraction. The boundary layer in OPT is substantially moister and more insulated with a strong increase in evaporative fraction of 30% and increased low cloud cover up to 15%. Higher evaporative fraction results in stronger evaporative cooling, while the increase in low cloud cover reduces the incoming shortwave radiation. This decrease in the incoming shortwave radiation is shown in the center panel of Fig. 13, where downward radiation is reduced strongly by to 20 W m^{-2}, lowering the net surface radiation budget at the surface. Stronger evaporative cooling and reduced net downward radiation leads to a cooling of the summer climate, which agrees with the geographical pattern of reduced bias of mean summer temperature (Fig. 4) and interannual summer temperature variability (Fig. 8). Qualitatively similar impacts are also evident over North America (Fig. 6), but the reduction of the summer temperature bias amounts to only about 1–2 K and is substantially smaller than over Europe (Fig. 4).

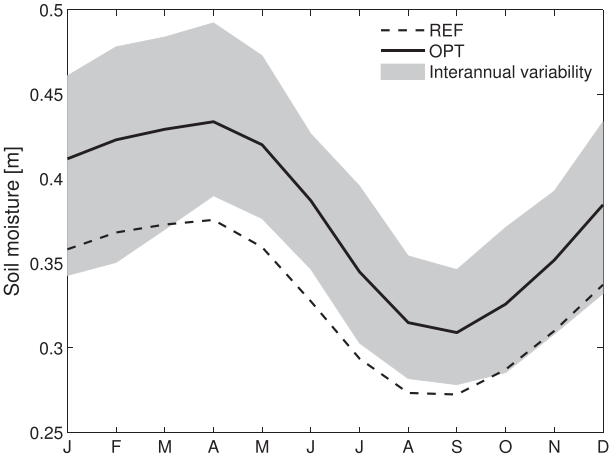

FIG. 12. Annual cycle of the absolute soil moisture content in the top 2 m of the soil for the simulations REF (dashed) and OPT (solid line), and its interannual variability (standard deviation, gray shading).

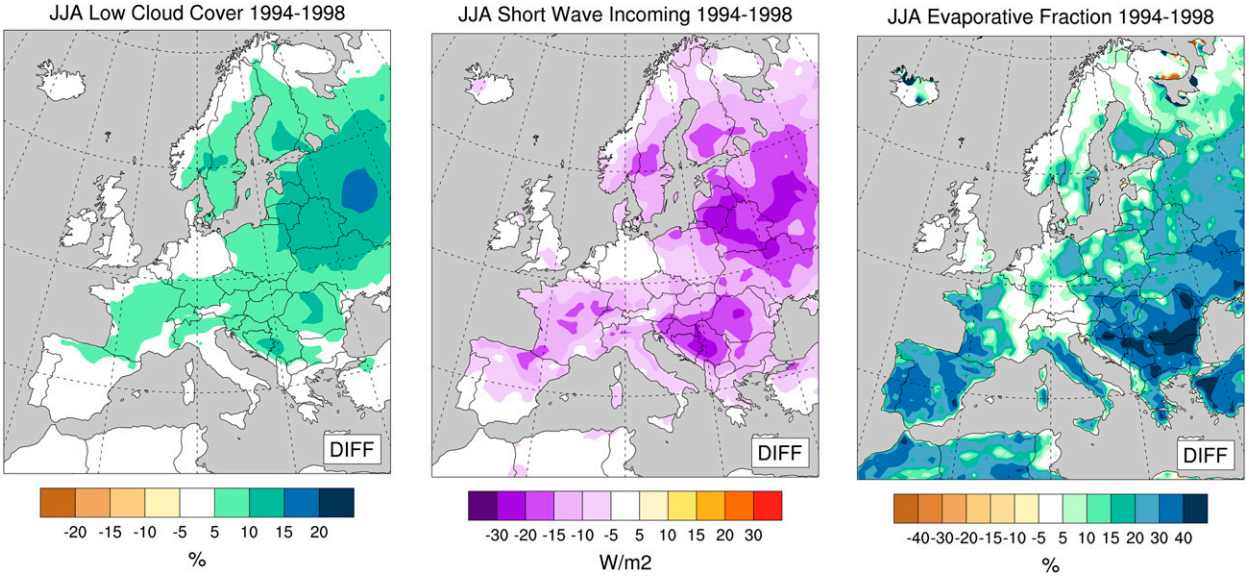

FIG. 13. Mean summer (JJA) differences between OPT and REF for (left) low cloud cover, (center) incoming shortwave radiation, and (right) evaporative fraction (latent heat flux divided by the sum of sensible heat and latent heat flux) for the period 1994–98.

4. Discussion and conclusions

We have analyzed the objective parameter calibration for a regional climate model over a native model domain (which was involved in previous expert tuning) and over a nonnative domain (with no previous tuning). The calibration simultaneously addressed eight different model parameters and was independently applied in both domains.

The calibration yielded important improvements in the representation of the summer climate for Europe and North America that has resisted previous expert tuning efforts. Most climate models suffer from an overly dry summer climate with strong warm temperature biases and dry precipitation biases widely discussed in recent studies (Boberg and Christensen 2012; Fischer et al. 2012; Kotlarski et al. 2014). We demonstrate here that parameter optimization strongly reduces these biases with respect to both the mean and the interannual variability. A key role is played by increased hydraulic conductivity, which increases soil moisture availability at the surface and moistens the boundary layer in summer. The determined increased hydraulic conductivity can be qualitatively explained by the scale dependency of the hydraulic conductivity.

Our results also show that the calibration yields almost identical optimal values over the two domains. This supports the robustness of the calibration methodology and indicates that it addresses uncertainties in the model physics that are common among different regions. It also indicates that no overtuning with respect to the model domain occurs, which has been a concern in the context of regional climate modeling (Laprise et al. 2008). The RCM used in this study is hence transferable between the domains considered. We argue that as a consequence, the transfer of an RCM to a nonnative domain should not necessarily include a change in model parameter values, unless there is some justifiable evidence from validation and/or physical considerations.

A multitude of biases remain, despite the calibration of parameters, particularly over North America, where the summer climate remains overly warm and dry over the Great Plains and Florida. Other prominent biases of CCLM include an overestimation of precipitation in winter over central Europe, a warm bias in winter over boreal and tundra regions over North America and an imbalance of total cloud cover, with too much cloud formation in the north and too little in the south for both model domains and both seasons. These biases should be addressed in future model development.

An objective calibration of a climate model as the one presented is certainly not a finite exercise. Recalibration of the model with new model releases comprising new physical parameterizations is fundamental to demonstrate its added value. Likewise, changes in the model configuration, such as an increase in horizontal resolution, would benefit from a revisited calibration as a variety of model parameters are known to be resolution dependent.

Acknowledgments. We would like to acknowledge the useful discussions with Marc Bierkens regarding the role of the soil hydraulic conductivity in large-scale hydrology.

Also we thank Xing Yuan and two anonymous reviewers for their constructive comments on the manuscript. We are further indebted to the COSMO consortium and the CLM community for providing access to and support of the CCLM. We would also like to acknowledge the technical support provided by the staff of MeteoSwiss and the Center for Climate Systems Modeling (C2SM). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under Project ID s432.

REFERENCES

Adams, D. K., and A. C. Comrie, 1997: The North American monsoon. *Bull. Amer. Meteor. Soc.*, **78**, 2197–2213, doi:10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2.

Alexandru, A., R. de Elia, and R. Laprise, 2007: Internal variability in regional climate downscaling at the seasonal scale. *Mon. Wea. Rev.*, **135**, 3221–3238, doi:10.1175/MWR3456.1.

Allen, M., 1999: Do-it-yourself climate prediction. *Nature*, **401**, 642, doi:10.1038/44266.

Arritt, R. W., T. D. Rink, M. Segal, D. P. Todey, C. A. Clark, M. J. Mitchell, and K. M. Labas, 1997: The Great Plains low-level jet during the warm season of 1993. *Mon. Wea. Rev.*, **125**, 2176–2192, doi:10.1175/1520-0493(1997)125<2176:TGPLLJ>2.0.CO;2.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res. Atmos.*, **119**, 7889–7907, doi:10.1002/2014JD021478.

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär, 2012a: Exploring perturbed physics ensembles in a regional climate model. *J. Climate*, **25**, 4582–4599, doi:10.1175/JCLI-D-11-00275.1.

——, ——, ——, and ——, 2012b: Objective calibration of regional climate models. *J. Geophys. Res.*, **117**, D23115, doi:10.1029/2012JD018262.

——, ——, ——, and ——, 2013: Physical constraints for temperature biases in climate models. *Geophys. Res. Lett.*, **40**, 4042–4047, doi:10.1002/grl.50737.

Boberg, F., and J. H. Christensen, 2012: Overestimation of Mediterranean summer temperature projections due to model deficiencies. *Nat. Climate Change*, **2**, 433–436, doi:10.1038/nclimate1454.

Boé, J., and L. Terray, 2014: Land–sea contrast, soil-atmosphere and cloud-temperature interactions: Interplays and roles in future summer European climate change. *Climate Dyn.*, **42**, 683–699, doi:10.1007/s00382-013-1868-8.

Bracco, A., J. D. Neelin, H. Luo, J. C. McWilliams, and J. E. Meyerson, 2013: High dimensional decision dilemmas in climate models. *Geosci. Model Dev. Discuss.*, **6**, 2731–2767, doi:10.5194/gmdd-6-2731-2013.

Bukovsky, M. S., 2011: Masks for the Bukovsky regionalization of North America. Regional Integrated Sciences Collective, Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, accessed 1 January 2013. [Available online at http://www.narccap.ucar.edu/contrib/bukovsky/.]

——, C. M. Carrillo, D. J. Gochis, D. M. Hammerling, R. R. McCrary, and L. O. Mearns, 2015: Toward assessing NARCCAP regional climate model credibility for the North American monsoon: Future climate simulations. *J. Climate*, **28**, 6707–6728, doi:10.1175/JCLI-D-14-00695.1.

Cattiaux, J., H. Douville, and Y. Peings, 2013: European temperatures in CMIP5: Origins of present-day biases and future uncertainties. *Climate Dyn.*, **41**, 2889–2907, doi:10.1007/s00382-013-1731-y.

Christensen, J. H., and O. B. Christensen, 2007: A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change*, **81**, 7–30, doi:10.1007/s10584-006-9210-7.

——, F. Boberg, O. B. Christensen, and P. Lucas-Picher, 2008: On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.*, **35**, L20709, doi:10.1029/2008GL035694.

Cloke, W. A., A. Weisheimer, and F. Pappenberger, 2011: Representing uncertainty in land surface hydrology: Fully coupled simulations with the ECMWF land surface scheme. *Proc. ECMWF/WMO/WCRP Workshop on Representing Model Uncertainty and Error in Numerical Weather and Climate Prediction Models*, Reading, United Kingdom, ECMWF, 109–120. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2011/8740-representing-uncertainty-land-surface-hydrology-fully-coupled-simulations.pdf.]

Curry, J. A., and A. H. Lynch, 2002: Comparing arctic regional climate model. *Eos, Trans. Amer. Geophys. Union*, **83**, 87, doi:10.1029/2002EO000051.

Davin, E., R. Stöckli, E. Jaeger, S. Levis, and S. Seneviratne, 2011: COSMO-CLM^2: A new version of the COSMO-CLM model coupled to the Community Land Model. *Climate Dyn.*, **37**, 1889–1907, doi:10.1007/s00382-011-1019-z.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Doms, G., and Coauthors, 2007: A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization. Consortium for Small-Scale Modelling Tech. Rep. LM-F90, 4.20, 154 pp. [Available online at http://www.cosmo-model.org/content/model/documentation/core/default.htm.]

Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, 2007: Soil moisture–atmosphere interactions during the 2003 European summer heat wave. *J. Climate*, **20**, 5081–5099, doi:10.1175/JCLI4288.1.

——, J. Rajczak, and C. Schär, 2012: Changes in European summer temperature variability revisited. *Geophys. Res. Lett.*, **39**, L19702, doi:10.1029/2012GL052730.

Fu, C., and Coauthors, 2005: Regional climate model intercomparison project for Asia. *Bull. Amer. Meteor. Soc.*, **86**, 257–266, doi:10.1175/BAMS-86-2-257.

Gelhar, L. W., 1986: Stochastic subsurface hydrology: From theory to applications. *Water Resour. Res.*, **22**, 135S–145S, doi:10.1029/WR022i09Sp0135S.

Giorgi, F., C. Jones, and G. Asrar, 2009: Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bull.*, **58**, 175–183.

——, and Coauthors, 2012: RegCM4: Model description and preliminary tests over multiple CORDEX domains. *Climate Res.*, **52**, 7–29, doi:10.3354/cr01018.

Gregoire, L. J., P. J. Valdes, A. J. Payne, and R. Kahana, 2011: Optimal tuning of a GCM using modern and glacial constraints. *Climate Dyn.*, **37**, 705–719, doi:10.1007/s00382-010-0934-8.

Harris, I., P. Jones, T. Osborn, and D. Lister, 2013: Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *Int. J. Climatol.*, **34**, 623–642, doi:10.1002/joc.3711.

Haylock, M. R., N. Hofstra, A. M. G. Klein 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. *J. Geophys. Res.*, **113**, D20119, doi:10.1029/2008JD010201.

Helfand, H. M., and S. D. Schubert, 1995: Climatology of the simulated Great Plains low–level jet and its contribution to the continental moisture budget of the United States. *J. Climate*, **8**, 784–806, doi:10.1175/1520-0442(1995)008<0784:COTSGP>2.0.CO;2.

Jacob, D., and Coauthors, 2012: Assessing the transferability of the regional climate model REMO to different Coordinated Regional Climate Downscaling Experiment (CORDEX) regions. *Atmosphere*, **3**, 181–199, doi:10.3390/atmos3010181.

Jaeger, E. B., I. Anders, D. Luethi, B. Rockel, C. Schaer, and S. I. Seneviratne, 2008: Analysis of ERA40-driven CLM simulations for Europe. *Meteor. Z.*, **17**, 349–367, doi:10.1127/0941-2948/2008/0301.

Klein, S. A., X. Jiang, J. Boyle, S. Malyshev, and S. Xie, 2006: Diagnosis of the summertime warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a weather forecasting approach. *Geophys. Res. Lett.*, **33**, L18805, doi:10.1029/2006GL027567.

Klocke, D., R. Pincus, and J. Quaas, 2011: On constraining estimates of climate sensitivity with present-day observations through model weighting. *J. Climate*, **24**, 6092–6099, doi:10.1175/2011JCLI4193.1.

Knutti, R., T. F. Stocker, F. Joos, and G. K. Plattner, 2002: Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, **416**, 719–723, doi:10.1038/416719a.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, doi:10.1126/science.1100217.

Kotlarski, S., T. Bosshard, D. Lüthi, P. Pall, and C. Schär, 2012: Elevation gradients of European climate change in the regional climate model COSMO-CLM. *Climatic Change*, **112**, 189–215, doi:10.1007/s10584-011-0195-5.

——, and Coauthors, 2014: Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model Dev. Discuss.*, **7**, 217–293, doi:10.5194/gmdd-7-217-2014.

Laprise, R., 2008: Regional climate modelling. *J. Comput. Phys.*, **227**, 3641–3666, doi:10.1016/j.jcp.2006.10.024.

——, and Coauthors, 2008: Challenging some tenets of regional climate modelling. *Meteor. Atmos. Phys.*, **100**, 3–22, doi:10.1007/s00703-008-0292-9.

Lüthi, D., A. Cress, H. C. Davies, C. Frei, and C. Schär, 1996: Interannual variability and regional climate simulations. *Theor. Appl. Climatol.*, **53**, 185–209, doi:10.1007/BF00871736.

Mauritsen, T., and Coauthors, 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, M00A01, doi:10.1029/2012MS000154.

McKay, M. D., R. J. Beckman, and W. J. Conover, 2000: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **42**, 55–61, doi:10.1080/00401706.2000.10485979.

Mearns, L. O., and Coauthors, 2012: The North American Regional Climate Change Assessment Program: Overview of phase I results. *Bull. Amer. Meteor. Soc.*, **93**, 1337–1362, doi:10.1175/BAMS-D-11-00223.1.

Menéndez, C. G. C., A. Sorensson, and J. P. Boulanger, 2010: CLARIS Project: Towards climate downscaling in South America. *Meteor. Z.*, **19**, 357–362, doi:10.1127/0941-2948/2010/0459.

Mueller, B., and S. I. Seneviratne, 2014: Systematic land climate and evapotranspiration biases in CMIP5 simulations. *Geophys. Res. Lett.*, **41**, 128–134, doi:10.1002/2013GL058055.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, and M. Collins, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772, doi:10.1038/nature02771.

——, B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J. Webb, 2007: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos. Trans. Roy. Soc. London*, **365A**, 1993–2028, doi:10.1098/rsta.2007.2077.

Neelin, J. D., A. Bracco, H. Luo, J. C. McWilliams, and J. E. Meyerson, 2010: Considerations for parameter optimization and sensitivity in climate models. *Proc. Natl. Acad. Sci. USA*, **107**, 21 349–21 354, doi:10.1073/pnas.1015473107.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.

Rijtema, P., 1969: Soil moisture forecasting. Instituut voor Cultuurtechniek en Waterhuishouding Tech. Rep. 513, 18 pp.

Rockel, B., and B. Geyer, 2008: The performance of the regional climate model CLM in different climate regions, based on the example of precipitation. *Meteor. Z.*, **17**, 487–498, doi:10.1127/0941-2948/2008/0297.

——, C. L. Castro, R. A. Pielke Sr., H. von Storch, and G. Leoncini, 2008: Dynamical downscaling: Assessment of model system dependent retained and added variability for two different regional climate models. *J. Geophys. Res.*, **113**, D21107, doi:10.1029/2007JD009461.

Rovey, C. W., and D. S. Cherkauer, 1995: Scale dependency of hydraulic conductivity measurements. *Groundwater*, **33**, 769–780, doi:10.1111/j.1745-6584.1995.tb00023.x.

Rowell, D., and R. Jones, 2006: Causes and uncertainty of future summer drying over Europe. *Climate Dyn.*, **27**, 281–299, 10.1007/s00382-006-0125-9.

Samaniego, L., R. Kumar, and S. Attinger, 2010: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.*, **46**, W05523, doi:10.1029/2008WR007327.

Sanderson, B. M., 2011: A multimodel study of parametric uncertainty in predictions of climate response to rising greenhouse gas concentrations. *J. Climate*, **24**, 1362–1377, doi:10.1175/2010JCLI3498.1.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Haberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, 332–336, doi:10.1038/nature02300.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.*, **99**, 125–161, doi:10.1016/j.earscirev.2010.02.004.

——, and Coauthors, 2013: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophys. Res. Lett.*, **40**, 5212–5217, doi:10.1002/grl.50956.

Seth, A., and F. Giorgi, 1998: The effects of domain choice on summer precipitation simulation and sensitivity in a regional

climate model. *J. Climate*, **11**, 2698–2712, doi:10.1175/1520-0442(1998)011<2698:TEODCO>2.0.CO;2.

Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406, doi:10.1038/nature03301.

Steppeler, J., G. Doms, U. Schättler, H. Bitzer, A. Gassmann, U. Damrath, and G. Gregoric, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteor. Atmos. Phys.*, **82**, 75–96, doi:10.1007/s00703-001-0592-9.

Suklitsch, M., A. Gobiet, A. Leuprecht, and C. Frei, 2008: High resolution sensitivity studies with the regional climate model CCLM in the Alpine region. *Meteor. Z.*, **17**, 467–476, doi:10.1127/0941-2948/2008/0308.

Sutton, R. T., B. Dong, and J. M. Gregory, 2007: Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophys. Res. Lett.*, **34**, L02701, doi:10.1029/2006GL028164.

Takle, E. S., and Coauthors, 1999: Project to Intercompare Regional Climate Simulations (PIRCS): Description and initial results. *J. Geophys. Res.*, **104**, 19 443–19 461, doi:10.1029/1999JD900352.

——, W. J. Gutowski, R. W. Arritt, J. Roads, I. Meinke, B. Rockel, C. G. Jones, and A. Zadra, 2007: Transferability intercomparison: An opportunity for new insight on the global water cycle and energy budget. *Bull. Amer. Meteor. Soc.*, **88**, 375–384, doi:10.1175/BAMS-88-3-375.

van der Linden, P., and J. F. B. Mitchell, 2009: ENSEMBLES: Climate change and its impacts at seasonal, decadal and centennial timescales; Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre Tech. Rep., 160 pp.

van Genuchten, M. Th., 1980: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Amer. J.*, **44**, 892–898, doi:10.2136/sssaj1980.03615995004400050002x.

Vidale, P. L., D. Lüthi, R. Wegmann, and C. Schär, 2007: European summer climate variability in a heterogeneous multi-model ensemble. *Climatic Change*, **81**, 209–232, doi:10.1007/s10584-006-9218-z.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.*, **50**, 7505–7514, doi:10.1002/2014WR015638.

Willmott, C. J., and K. Matsuura, 2009: Terrestrial air temperature and precipitation: Monthly and annual time series (V2.01). Subset used: 1990–2008, Center for Climatic Research, Department of Geography, University of Delaware, accessed 6 January 2012. [Available online at http://climate.geog.udel.edu/climate/.]

Wylie, D., D. L. Jackson, W. P. Menzel, and J. J. Bates, 2005: Trends in global cloud cover in two decades of HIRS observations. *J. Climate*, **18**, 3021–3031, doi:10.1175/JCLI3461.1.

Yokohata, T., M. J. Webb, M. Collins, K. D. Williams, M. Yoshimori, J. C. Hargreaves, and J. D. Annan, 2010: Structural similarities and differences in climate responses to CO_2 increase between two perturbed physics ensembles. *J. Climate*, **23**, 1392–1410, doi:10.1175/2009JCLI2917.1.

Zhang, Y. C., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, 2004: Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res.*, **109**, D19105, doi:10.1029/2003JD004457.

Zubler, E. M., D. Folini, U. Lohmann, D. Lüthi, A. Muhlbauer, S. Pousse-Nottelmann, C. Schär, and M. Wild, 2011: Implementation and evaluation of aerosol and cloud microphysics in a regional climate model. *J. Geophys. Res.*, **116**, D02211, doi:10.1029/2010JD014572.