



HAL
open science

Influence of Bias Correcting Predictors on Statistical Downscaling Models

Mathieu Vrac, Pradeebane Vaithinada Ayar

► **To cite this version:**

Mathieu Vrac, Pradeebane Vaithinada Ayar. Influence of Bias Correcting Predictors on Statistical Downscaling Models. *Journal of Applied Meteorology and Climatology*, 2017, 56 (1), pp.5 - 26. 10.1175/JAMC-D-16-0079.1 . hal-01584269

HAL Id: hal-01584269

<https://hal.science/hal-01584269>

Submitted on 9 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Influence of Bias Correcting Predictors on Statistical Downscaling Models[✉]

MATHIEU VRAC AND PRADEEBANE VAITTINADA AYAR

Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL, CNRS), Centre d'Etudes de Saclay, Orme des Merisiers, Gif-sur-Yvette, France

(Manuscript received 12 February 2016, in final form 21 September 2016)

ABSTRACT

Statistical downscaling models (SDMs) and bias correction (BC) methods are commonly used to provide regional or debiased climate projections. However, most SDMs are utilized in a “perfect prognosis” context, meaning that they are calibrated on reanalysis predictors before being applied to GCM simulations. If the latter are biased, SDMs might suffer from discrepancies with observations and therefore provide unrealistic projections. It is then necessary to study the influence of applying bias correcting to large-scale predictors for SDMs, since it can have impacts on the local-scale simulations: such an investigation for daily temperature and precipitation is the goal of this study. Hence, four temperature and three precipitation SDMs are calibrated over a historical period. First, the SDMs are forced by historical predictors from two GCMs, corrected or not corrected. The two types of simulations are compared with reanalysis-driven SDM outputs to characterize the quality of the simulations. Second, changes in basic statistical properties of the raw GCM projections and those of the SDM simulations—driven by bias-corrected or raw predictors from GCM future projections—are compared. Third, the stationarity of the SDM changes brought by the BC of the predictors is investigated. Changes are computed over a historical (1976–2005) and future (2071–2100) time period and compared to assess the nonstationarity. Overall, BC can have impacts on the SDM simulations, although its influence varies from one SDM to another and from one GCM to another, with different spatial structures, and depends on the considered statistical properties. Nevertheless, corrected predictors generally improve the historical projections and can impact future evolutions with potentially strong nonstationary behaviors.

1. Introduction

The need for downscaling climate simulations is present in many studies, and in particular in many studies related to impacts of climate change. Indeed, if global climate models (GCMs) are the main tools for projecting future climate evolutions, the spatial resolution of their outputs is generally too coarse (from about 100 to 300 km) to drive impact models and to lead associated impact studies (e.g., IPCC 2013). Hence, downscaling models have been and are still extensively developed and applied to generate climate values (temperature, precipitation, etc.) at regional, local, or station scales. Now well established in the literature, those models are said to be either dynamical or statistical.

While dynamical downscaling explicitly solves the physical equations of the atmosphere at a regional scale—through so-called regional climate models (RCMs; e.g., Laprise 2008; Rummukainen 2010)—statistical downscaling is based on statistical relationships and models to link large-scale (atmospheric, environmental, etc.) information called predictors to local-scale variables of interest called predictands. Both approaches have been widely tested, evaluated, and intercompared on many climate variables (e.g., Schmidli et al. 2007; Vaittinada Ayar et al. 2015).

A statistical downscaling model (SDM), however, has some particularities with respect to RCMs. The main one is certainly that it is not directly based on physical equations. It therefore requires large- and local-scale data for its calibration [see, e.g., Maraun et al. (2010) and Vaittinada Ayar et al. (2015) for further discussions about RCMs and SDMs]. As such, when a statistical downscaling model is applied in a climate change context (i.e., driven by GCM future projections as predictors), some underlying hypotheses are made (Hewitson and Crane 2006): 1) the statistical model, calibrated

[✉] Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JAMC-D-16-0079.s1>.

Corresponding author e-mail: Mathieu Vrac, mathieu.vrac@lscce.ipsl.fr

under present or recent past conditions, remains valid under modified climate conditions; 2) the predictors used as input into the SDM are relevant and completely represent the climate change signal; and 3) the predictors used as input into the SDM are correctly represented and simulated by the GCM. Although the evaluation of those hypotheses is rarely performed before applying any SDM, some methods exist to assess the capability of SDMs to reproduce the statistical properties of observations when calibrated and driven by reanalysis data (Huth 1999; Robertson et al. 2004; Vrac et al. 2007b) or driven by GCM or RCM simulations (Charles et al. 1999; Wilby and Wigley 2000; Charles et al. 2004; Chen et al. 2014). Other approaches have been tested to evaluate the robustness of the SDMs in time, comparing SDMs' future projections with those from GCMs (e.g., Frías et al. 2006) or from RCMs (e.g., Wood et al. 2004; Haylock et al. 2006). Vrac et al. (2007c) have also developed a method for evaluating SDMs under control (CTRL) climate (reanalyses versus GCMs) and under future climate for which RCM future projections are considered as "pseudo observations" that can be compared with SDM future simulations [see also Gaitan et al. (2014) for applications]. Many of those studies showed that validating an SDM in present-day conditions does not imply legitimacy for the SDM projections in changed climate conditions (e.g., Charles et al. 1999). Moreover, the third hypothesis is of high importance. Indeed, most of the SDMs are part of the so-called perfect prognosis (PP) realm, as opposed to the model output statistics (MOS) approach [see Maraun et al. (2010) or Vaittinada Ayar et al. (2015) for details and references]. This PP context means that, for example, for the downscaling of future GCM projections, those SDMs need first to be calibrated based on reanalysis predictors to then be applied to the GCM simulations. If the latter are biased with respect to reanalyses, SDM outputs might themselves suffer from large discrepancies with respect to local-scale observations and therefore provide mistaken and unrealistic future projections (e.g., Charles et al. 2004; Frost et al. 2011; Bürger et al. 2012; Grouillet et al. 2016). It is then necessary to study the influence of bias correcting the large-scale predictors of SDMs, since it can have major impacts on the local-scale statistical simulations. Indeed, if climate model simulations have seen their quality improved over the last years and decades, they still have some biases in the sense that their statistical distribution differs from that of observations (Meehl et al. 2007; Christensen et al. 2008; White and Toumi 2013; Vrac and Friederichs 2015). Hence, in parallel to climate model developments, statistical bias correction (BC) methods have also been designed to adjust climate simulations by

transforming the simulated data into new data with fewer or no statistical biases with respect to reference (e.g., Haddad and Rosenfeld 1997; Gudmundsson et al. 2012; Vrac et al. 2012). It is therefore logical to wonder if and how BC methodologies applied to predictors derived from GCM simulations can affect the realism and the quality of the SDM projections under present and future conditions. In other words, even if the third hypothesis is not completely verified—that is, the predictors used as input into the SDM are not correctly simulated by the GCM—what are the impacts (in terms of SDM outputs) of bias correcting those GCM-derived predictors before performing an SDM? Some studies, such as Colette et al. (2012) or White and Toumi (2013), have investigated such a question for dynamical downscaling. Although they found that it can produce some undesirable features in the RCM simulations, their main conclusion was that such a prior correction of the large-scale inputs for RCMs with a quantile-association-based method clearly improves the quality of the RCM simulations. This question has never been addressed for statistical downscaling and is therefore the main goal of the present article, for temperature and precipitation.

To tackle this question, as SDM simulations may differ from one SDM to another, four (three) different SDMs are employed for temperature (precipitation). Those are applied at a daily time scale over inland Europe comprising approximately the region 20°–70°N, 40°W–80°E. Moreover, preliminary analyses indicated that specific bias correction methods do not strongly change the main conclusions of this study (not shown). Hence, only one BC method is used in the following. Both BC and SDMs are presented in section 2. The different datasets used are presented in section 3, including the reference datasets (one for the BC, one for the historical SDM evaluations) as well as the two GCMs and the reanalyses, both providing large-scale predictors. This section also describes the evaluations in historical and future contexts. Results are provided in section 4, not only for historical biases, but also for future comparisons in terms of changes in statistical properties for temperature and precipitation from present to future periods, as well as in terms of stationarity of the BC influence between present and future periods. General conclusions are given in section 5, along with discussions and perspectives for this study.

2. Statistical downscaling models and BC method

Four SDMs are tested for temperature and three for precipitation. Those models were selected based on the

evaluation study performed in Vaithinada Ayar et al. (2015). For temperature, the four SDMs are the following: the cumulative distribution function–transform (CDFt) approach (e.g., Michelangeli et al. 2009; Vrac et al. 2012), a conditional stochastic weather generator (SWG; e.g., Chandler and Wheeler 2002), a model based on analogs (ANALOG; e.g., Zorita and von Storch 1999), and a generalized additive model (GAM; Hastie and Tibshirani 1990). For precipitation downscaling, CDFt, SWG, and ANALOG are used, but not GAM, which performed poorly (Vaithinada Ayar et al. 2015). Those models are briefly described here; further details are given in Vaithinada Ayar et al. (2015, and references therein).

a. Cumulative distribution function–transform

The CDFt approach links the local-scale CDF of the variable of interest to the associated large-scale CDF through a quantile–quantile approach performed between the future large- and local-scale CDFs—not between present CDFs as in the classical quantile–quantile method. To do so, the future local-scale CDF is first estimated based on an assumption of a mathematical transformation to link the evolution of the large-scale CDF to the evolution of the local-scale one. Hence, CDFt and quantile–quantile methods are philosophically close to each other, but CDFt accounts for the CDF changes from the calibration to the projection (or future) time periods. All theoretical and technical details are given in Vrac et al. (2012). Note that, contrary to SWG, ANALOG, and GAM, which need to be calibrated with large-scale reanalysis data as predictors before being applied to GCM data, the CDFt model directly establishes a link between the GCM and local-scale CDFs. No calibration based on real-time predictors is needed. This is a specificity of some MOS approaches by opposition to PP models (for details, see, e.g., Maraun et al. 2010; Vaithinada Ayar et al. 2015).

In the present study, CDFt is used both as a statistical downscaling model and as a bias correction method. Indeed, CDFt will also be used to bias correct GCM outputs with respect to reanalysis data to provide bias corrected inputs to the different SDMs (i.e., CDFt included). This bias correction method has been chosen here 1) because it is derived from the quantile-mapping (QM) approach that is certainly one of the most widely used techniques and 2) because it allows us to account for the changes in the climate model CDFs.

b. Stochastic weather generator

In the SWG used in this study, a two-step approach is implemented to model precipitation in a vector generalized

linear model (VGLM) framework as in Chandler and Wheeler (2002). First, the rainfall occurrence is modeled through a logistic regression (e.g., Buishand et al. 2004; Fealy and Sweeney 2007), allowing to us characterize the probability p_i of rainfall occurrence for a given day i conditionally on atmospheric predictors (or covariates) $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,N})$:

$$p_i = \frac{\exp\left(P^0 + \sum_{j=1}^N P^j X_{i,j}\right)}{1 + \exp\left(P^0 + \sum_{j=1}^N P^j X_{i,j}\right)}, \tag{1}$$

where $(P^j)_{j=0,\dots,N}$ corresponds to coefficients to be estimated. Then, the probability density function (pdf) of the rain intensity (given that it rains) is assumed to be a Gamma distribution whose logarithms of the shape α_i and rate β_i parameters are linear functions of the large-scale predictors \mathbf{X}_i at day i :

$$\begin{cases} \log(\alpha_i) = \alpha^0 + \sum_{j=1}^N \alpha^j X_{i,j} \\ \log(\beta_i) = \beta^0 + \sum_{j=1}^N \beta^j X_{i,j} \end{cases}, \tag{2}$$

where $(\alpha_j)_{j=0,\dots,N}$ and $(\beta_j)_{j=0,\dots,N}$ correspond to coefficients to be estimated. For temperature, a single step is used, where temperature at day i is supposed to follow a Gaussian distribution with the mean m_i and the logarithm of the standard deviation σ_i linearly dependent on \mathbf{X}_i .

c. ANALOG

The ANALOG method is applied here in its deterministic form, as defined, for example, in Zorita and von Storch (1999), Yiou et al. (2007), and Vautard and Yiou (2009). For each given day i (with predictors \mathbf{X}_i) to be downscaled in the projection period, ANALOG consists of finding, in the calibration period Cal, the day d that has the closest atmospheric situation (i.e., the predictors \mathbf{X}_d the closest to \mathbf{X}_i) according to a chosen distance (or similarity) metric:

$$d = \operatorname{argmin}_{d \in \text{Cal}} \operatorname{dist}(\mathbf{X}_i, \mathbf{X}_d). \tag{3}$$

The temperature and precipitation values observed for this selected day d are then taken as the downscaled values for the day i . Many distances have been employed (e.g., Grenier et al. 2013), but the Euclidian distance is used here. Note that, as ANALOG samples days from the calibration dataset, it will not be able to create data out of this set. Hence, in a warming context, ANALOG

(at least in its present version) can shift the hottest historical days within the seasonal cycle, but it cannot generate new “record breaking” events. Therefore, an underestimation of the hottest days in a future climate context may be expected.

d. Generalized additive models

For temperature downscaling, a fourth model is also used, based on GAM (Hastie and Tibshirani 1990). GAM is a deterministic model characterizing the expectation of the local-scale predictand (here, the temperature) at day i , conditionally on the large-scale predictors \mathbf{X}_i as a sum of spline functions $f_j(X_j)$ (here cubic regression splines):

$$E(Y | X_1, \dots, X_N) = \sum_{j=1}^N f_j(X_j). \quad (4)$$

Such a downscaling method has been applied under present climate conditions, for instance, to downscale temperature (Vaithinada Ayar et al. 2015) or near-surface wind fields (Salameh et al. 2009), or under past climate like the Last Glacial Maximum time period (−21 kyr) to reconstruct monthly temperature and precipitation climatology over western Europe (Vrac et al. 2007a) or global permafrost (Levavasseur et al. 2011). Because of its poor performance for downscaling daily precipitation (Vaithinada Ayar et al. 2015), GAM is only used for temperature here. This approach is said to be data driven since it allows us to model both piecewise linearities and additive nonlinearities depending on the nature of the predictor–predictand dependence (in other words, the complexity of the model is not explicitly fixed by the user, who usually only specifies a smoothing penalty parameter).

Additional technical details about these four SDMs are given in Vaithinada Ayar et al. (2015, and references therein).

The bias correction of the GCMs through CDFt is applied to daily data but on a monthly basis (i.e., month by month), leading to 12 CDFt models for each predictor and GCM. CDFt as SDM is also applied to daily data and on a monthly basis (i.e., month by month). However, the other SDMs—SWG, ANALOG, and GAM—are calibrated and applied to daily data but on a seasonal basis (i.e., for each season separately) with the traditional Northern Hemisphere 3-month seasons: winter (DJF), spring (MAM), summer (JJA), and autumn (SON). Preliminary analyses with other configurations (e.g., monthly based calibrations for all models) did not show any significant differences in the conclusions presented later in this article (not shown).

3. Data and evaluation procedure

a. Reference and model data

For the calibration of most of the statistical downscaling models, two types of data are needed: local-scale observations of the variable of interest (predictand) and large-scale variables (predictors). In the present study, the daily local-scale predictands (i.e., temperature and precipitation) come from the E-OBS gridded dataset (Haylock et al. 2008) obtained through an interpolation procedure from the European Climate Assessment and Dataset (ECA&D) time series at meteorological stations (Klok and Klein Tank 2009). The E-OBS version used in this study is version 11 (release date April 2015) and covers inland Europe (comprising approximately the region 20°–70°N, 40°W–80°E) on a rotated grid with a spatial resolution of 0.44° from 1950 to 2014 at a daily time step.

The large-scale predictors come from ERA-Interim (ERA-I) reanalyses (Dee et al. 2011) taken at 1.125° × 1.125° spatial resolution from 1979 to 2005.

Except for CDFt, which only uses the large-scale variable of interest as predictors, the other SDMs tested in this study are using several predictors. For temperature, the predictors are the temperature at 2 m (T2); the sea level pressure (SLP); and the eastward wind, the northward wind, and the geopotential height at the 850-hPa pressure level (U850, V850, and Z850). For precipitation, the same predictors are employed with the addition of the dewpoint at 2 m (D2) and the large-scale precipitation (PR). All of the predictors are extracted over the region 20.25°–72.00°N, 52.875°W–76.50°E. More precisely, based on those data, for ANALOG the predictors are the fields of anomalies of those variables with respect to the seasonal cycle. For SWG and GAM, they are the first two principal components (PCs) of a principal component analysis (PCA; Barnston and Livezey 1987) performed on each predictor separately. For discussion on the choice of those predictors, see Vaithinada Ayar et al. (2015).

The daily predictors needed for the SDMs have also been extracted from two GCMs to drive the different SDMs over historical (1950–2005) and future (2006–2100) time periods under the RCP8.5 scenario (IPCC 2013): the IPSL-CM5A-MR global climate model with a 1.125° × 2.5° resolution, developed at L’Institut Pierre-Simon Laplace (Marti et al. 2010; Dufresne et al. 2013), and the CNRM-CM5 global climate model (Voldoire et al. 2013) with a 1.4° × 1.4° resolution, developed at Météo-France. All GCM variables used as predictors into the SDMs have been regridded to the 1.125° × 1.125° ERA-I spatial resolution. Hence, the SDMs (calibrated with reanalyses) are then forced either by the

raw (regridded) predictors or by the bias corrected (regridded) predictors, both over the historical and future (2006–2100) time periods. As mentioned in [section 2](#), the bias correction of those GCM-based predictors is performed through CDFt, and the BC references are the ERA-I reanalyses.

b. Evaluations over the historical period and comparison over a future period

The impact of bias correcting the GCM-derived predictors before incorporating them into SDMs is first studied over a historical period where observed data are available. In this context, the SDMs are first calibrated over the 1979–2005 time period to link the ERA-I predictors to the E-OBS data used as reference. Note that the predictors depend on the SDM of interest (see [section 3a](#)). In a second step, the predictors are taken from GCMs to drive the SDMs and provide downscaled simulations. Those GCM predictors can be used either directly as input into the SDMs, or in a corrected version, that is, with a BC method applied to them to make their distribution as close as possible to that of the ERA-I initial predictors. Those two types of predictors—hereafter referred to as raw and bias corrected predictors, respectively—are extracted or calculated over the 1979–2005 time period in order to drive the SDMs over the same time period. The evaluations are performed by comparison of those GCM-driven downscaled time series with the ERA-I–driven downscaled time series. Note that the real values of the predictands (i.e., E-OBS) are not used for evaluation since the goal here is to evaluate 1) the biases of SDM simulations when using GCM outputs as predictors instead of reanalyses and 2) the potential gain in SDM results when correcting those predictors. Indeed, the evaluation of those various simulations against real observations would make it difficult to distinguish the biases brought by the SDM itself from the biases brought by the predictors. Working with SDMs driven by predictors from ERA-I as evaluation reference allows us to ease the interpretation and to focus directly on the predictors themselves, without considering the quality (with respect to E-OBS) of the SDMs. Note that the quality and biases of the SDM with respect to E-OBS have been studied in [Vaithinada Ayar et al. \(2015\)](#).

In a future projection context, there is obviously no observed or reference data available for evaluating the simulations. However, the goal here is to evaluate the differences between the projections obtained with raw-GCM-predictor-driven SDMs and those obtained from the BC-GCM-predictor-driven SDMs, as well as between the different SDMs. To perform such a task, the SDMs are first calibrated as before over the 1979–2005

time period between the predictors from the ERA-I reanalyses and the reference E-OBS data. Then, GCM predictors—in their raw or bias corrected version—over the years 2006–2100 are then provided as input into the calibrated SDMs. In all the following, the analyses are performed according to one future time period, 2071–2100. Based on this approach, it is possible to compare, for a given SDM and the given future time period, the changes in the SDM simulations when driven by raw or corrected predictors. To do so in a coherent way, all temporal changes are computed with respect to the outputs of the associated method over the 1979–2005 time period: that is, the raw-GCM-driven SDM change is computed with respect to the raw-GCM-driven SDM simulations for 1979–2005, and equivalently for the corrected-GCM-driven SDM approaches.

Moreover, relying on the SDM historical and future projections (from raw and bias corrected predictors), it is also possible to study the temporal stationarity of the changes brought by the BC of the predictors. Those changes can indeed be computed over the historical (1976–2005) and future (2071–2100) time periods and compared to assess the potential nonstationarity.

4. Results

In all the following, results are only shown for the CNRM GCM for the winter (DJF) season. Indeed, conclusions are generally similar for the IPSL outputs. However, winter IPSL results (in Figs. SM3–SM14 of the online supplementary material) will be discussed in the following when significant differences with respect to CNRM are visible. For brevity, summer CNRM and IPSL results are provided for illustrations only as online supplementary material (in Figs. SM15–SM34). Note that hereinafter online supplementary figures are denoted with an SM prefix.

a. Mean and variance evaluations over the historical period

In this subsection, analyses will rely on “biases” of SDM outputs. Those biases are defined as differences between SDM projections driven either by corrected or noncorrected predictors and SDM outputs driven by reanalysis data as predictors. The underlying idea is that SDM outputs driven by reanalyses somehow provide the best results that can be obtained with SDMs. For an evaluation of the biases of reanalysis-driven SDMs, see [Vaithinada Ayar et al. \(2015\)](#). [Figures 1 and 2](#) show differences of mean winter biases, that is, biases of SDM(BC) minus biases of SDM(raw), from the SDMs

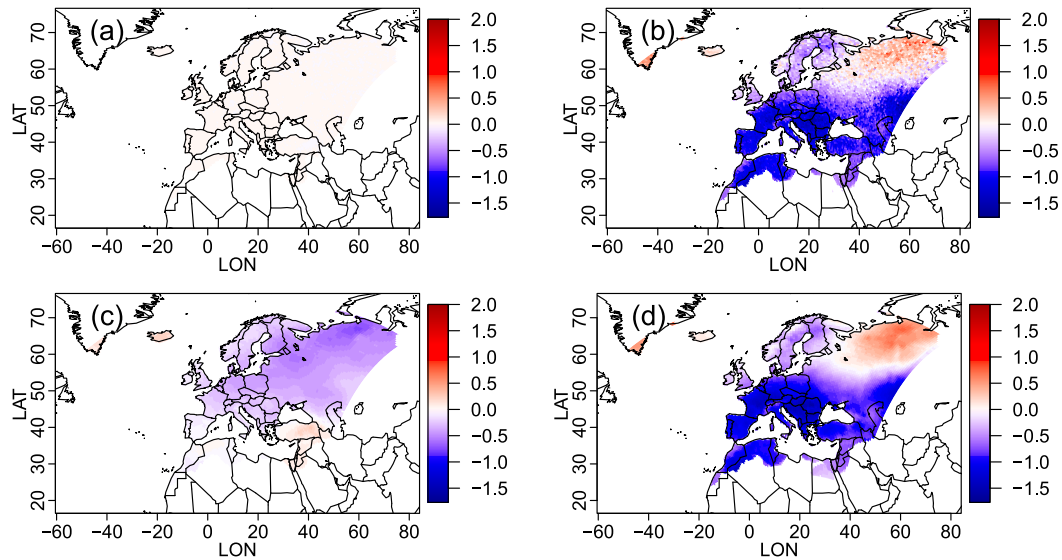


FIG. 1. Differences of temperature biases ($^{\circ}\text{C}$) between SDM simulations driven by corrected CNRM predictors [biases of SDM(BC)] or noncorrected CNRM predictors [biases of SDM(raw)], that is, biases of SDM(BC) – biases of SDM(raw): (a) CDFt, (b) SWG, (c) ANALOG, and (d) GAM.

driven by the CNRM outputs as predictors for temperature (in degrees Celsius) and precipitation (relative biases in percent), respectively. In general, a negative value does not mean that the bias itself (or its absolute value) is smaller for SDM(BC) than for SDM(raw) (e.g., -12 is lower than 3 but $|-12| > |3|$). However, in the present study, for temperature, the analysis of the differences of the absolute mean biases (given for temperature in Fig. SM1 of the online supplementary material) confirms that, for most locations, this does imply smaller biases for SDM(BC) than for SDM(raw). For precipitation (Fig. 2), patterns of differences of absolute

values of relative biases are somehow different for SWG (Fig. 2b). While in Fig. 2 positive and negative differences of biases are found, the computation of the differences of the absolute biases (Fig. SM2 of the online supplementary material) shows mostly negative values, indicating a bias (in absolute value) smaller for SDM(BC) than for SDM(raw) for SWG. Hence, for precipitation, positive values in Fig. 2b mostly indicate opposite signs of the biases. CDFt presents very close to zero differences of seasonal biases for both temperature and precipitation (Figs. 1a, 2a). This indicates that the correction of the large-scale temperature or

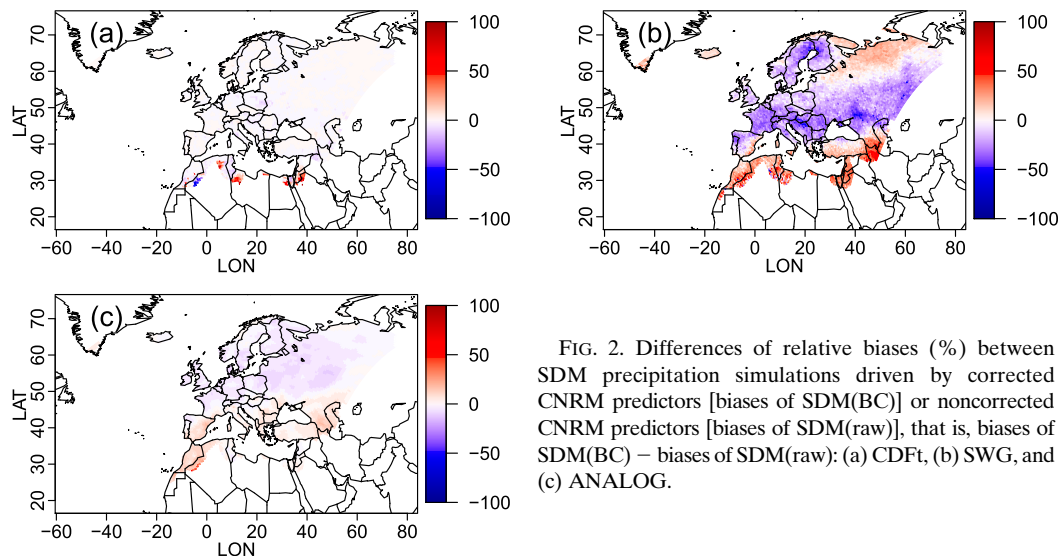


FIG. 2. Differences of relative biases (%) between SDM precipitation simulations driven by corrected CNRM predictors [biases of SDM(BC)] or noncorrected CNRM predictors [biases of SDM(raw)], that is, biases of SDM(BC) – biases of SDM(raw): (a) CDFt, (b) SWG, and (c) ANALOG.

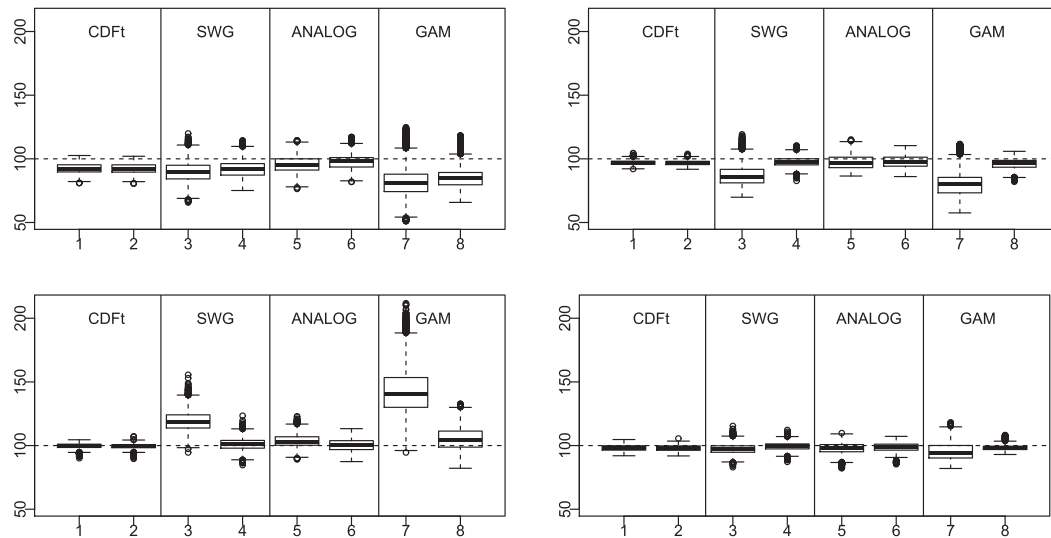


FIG. 3. Box plots of temperature seasonal variance ratios (%) for each SDM from corrected or noncorrected predictors for (top left) winter, (top right) spring, (bottom left) summer, and (bottom right) autumn. For each SDM, the left box plot is from noncorrected predictors while the right box plot is from corrected predictors. In each box plot, the box is constituted by the 25th, 50th, and 75th percentiles. The upper and lower fences are situated at 1.5 times the interquartile range from the box, and the dots are the values beyond those fences.

precipitation—or more generally the bias correction of the only predictor used as input in CDFt—does not impact the downscaling results provided by CDFt. This model simply adapts the CDF transformation to the inputs. Note also that the small differences between results from raw or corrected predictors can be explained by the already satisfying results of CDFt (at least in terms of biases) when driven by uncorrected predictors. For SWG (Figs. 1b, 2b), the differences of biases are not as close to zero as for CDFt. For the major part of the domain, the differences are significantly negative. Therefore, bias correcting the predictors used in SWG improves the quality of the downscaling results. This is not true everywhere. For example, the northeastern part of the domain for temperature and precipitation and the southern Mediterranean Sea basin for precipitation show positive differences of biases, implying that bias correcting the predictor degrades the downscaling for SWG over those regions in winter.

SWG presents differences of biases with a large spatial variability from one season to another, for both temperature and precipitation. Hence, if the bias correction globally brings additional quality to the final SWG downscaling, the gain varies in time and space. CDFt and SWG results for the other seasons and using the IPSL GCM (instead of CNRM) display equivalent features (e.g., see IPSL winter results in supplementary material Figs. SM3 and SM4) even if spatial specificities may change. For the ANALOG method, the temperature differences are all negative in winter (Fig. 1c), indicating a

clear gain in correcting the predictors. However, positive differences appear for the other seasons, over different regions depending on the season (not shown). Nevertheless, the positive or negative values of differences of biases are usually not as pronounced as those from SWG. The gains or degradations brought by bias correction of the predictors in ANALOG stay relatively small for all seasons (not shown), especially for precipitation. Differences of biases for GAM (Fig. 1d) are always similar to those of SWG. This was relatively expected since SWG can be seen here as a kind of stochastic version of the deterministic GAM. In other words, on average, the parameters of the distributions modeled by GAM and SWG are equivalent, but, by construction, SWG provides more variability than GAM through stochastic realizations, while GAM only provides the daily conditional expectations. This is probably the main reason that explains the pixelation in the differences of biases for SWG (Figs. 1b, 2b), while GAM does not have this problem.

In addition to mean seasonal biases, it is also informative to look at the variability of the temperature and precipitation projections. Figures 3 and 4 show the box plots of seasonal variance ratios (VR in percent) for temperature and precipitation, respectively. For each grid point, VR is defined as $VR = [\text{var}(S)/\text{var}(R)] \times 100$, where S and R correspond to simulated and reference daily time series for the season of interest, that is, to the downscaled values from the (raw or corrected) GCM and those from ERA-I, respectively. The main conclusions are the same for temperature and precipitation.

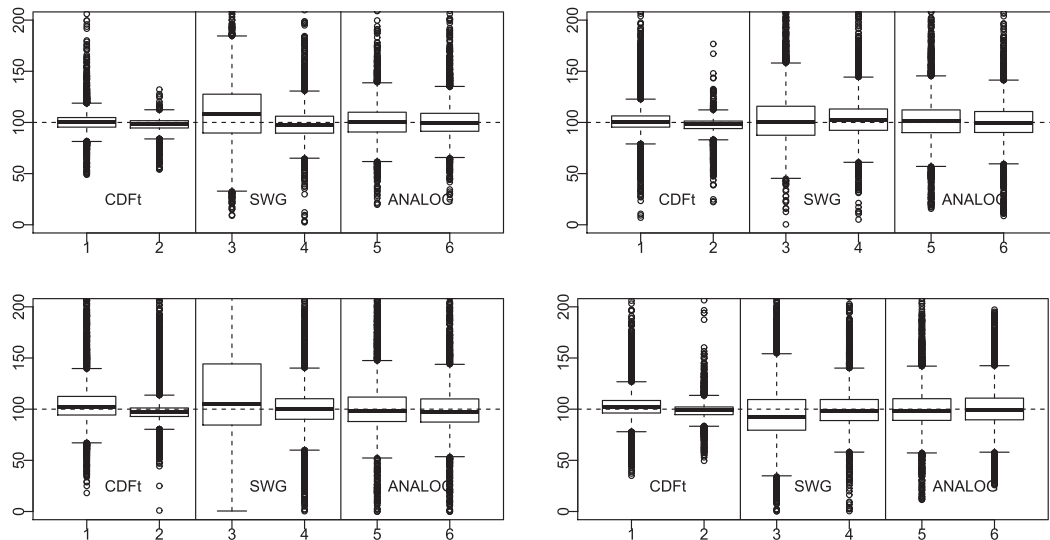


FIG. 4. As in Fig. 3, but for precipitation.

Whatever the season of interest, the mean VR from CDFt is not affected by the bias correction of the predictors, staying close to the perfect 100% VR score. Nevertheless, for precipitation, the variability of the VR values is clearly reduced around 100%, implying some gain induced by the bias correction. ANALOG does not show much difference between VR from corrected predictors or not for the precipitation variable and only slight improvements for temperature. However, the improvements are much stronger for SWG, especially in spring and summer for temperature (Fig. 3) and in winter and summer for precipitation (Fig. 4). Clear improvements (i.e., VR closer to 100%) are also visible for GAM (in temperature) for all seasons.

Generally, in terms of mean biases and variance ratios, the results are very similar when predictors come from IPSL or CNRM GCMs (see IPSL winter results in Figs. SM3–SM6 for direct comparisons with Figs. 1–4).

Moreover, an evaluation of the gain of bias correcting the SDM predictors has also been realized in terms of extreme values of temperature and precipitation. This has been conducted through the quantification of the percentage of simulated values higher than the reference 95th percentile (hereinafter referred to as $\%_{>Q_{95}}$) and through the calculation of the bias (with respect to the reference time series) of the 10-yr return levels (hereinafter referred to as BRL_{10y}). Both criteria provided equivalent patterns and features, similar to those given by the variance ratios. So, $\%_{>Q_{95}}$ and BRL_{10y} are not shown.

b. Comparisons in a future period context

Obviously, there are no reanalysis data to characterize future climate predictors. As a consequence, the

SDMs, driven by raw or corrected GCM data, cannot be compared with simulations from SDM driven by observed predictors or reanalyses. Therefore, this subsection consists of comparing, over a future time period, the raw and bias corrected GCM-driven SDM outputs. This comparison is performed according to two angles. First, an analysis about the influence of applying bias correcting to predictors on the changes in the mean and of the variability is performed for both temperature and precipitation. Second, the influence of such a correction on the temporal stationarity (i.e., from present to future context) is investigated, both in terms of mean climate and in terms of variability of temperature and precipitation. The future time period considered here is 2071–2100, under the RCP8.5 scenario (IPCC 2013).

1) INFLUENCES ON CHANGES IN MEAN AND VARIANCE

Figure 5 shows, in the left column, the mean winter temperature changes (future minus historical data) from the four SDMs driven by raw or bias corrected CNRM predictors, as well as the map of the mean temperature changes as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations. The right column presents the maps of the differences of mean winter temperature changes (from 1976–2005 to 2071–2100) between BC-CNRM-driven SDM simulations and raw-CNRM-driven SDM simulations [i.e., change of $SDM(BC \text{ predictors})$ minus change of $SDM(\text{raw predictors})$] for the four temperature SDMs. Hence, the right column allows us to visualize the influence of bias correcting the SDM predictors on

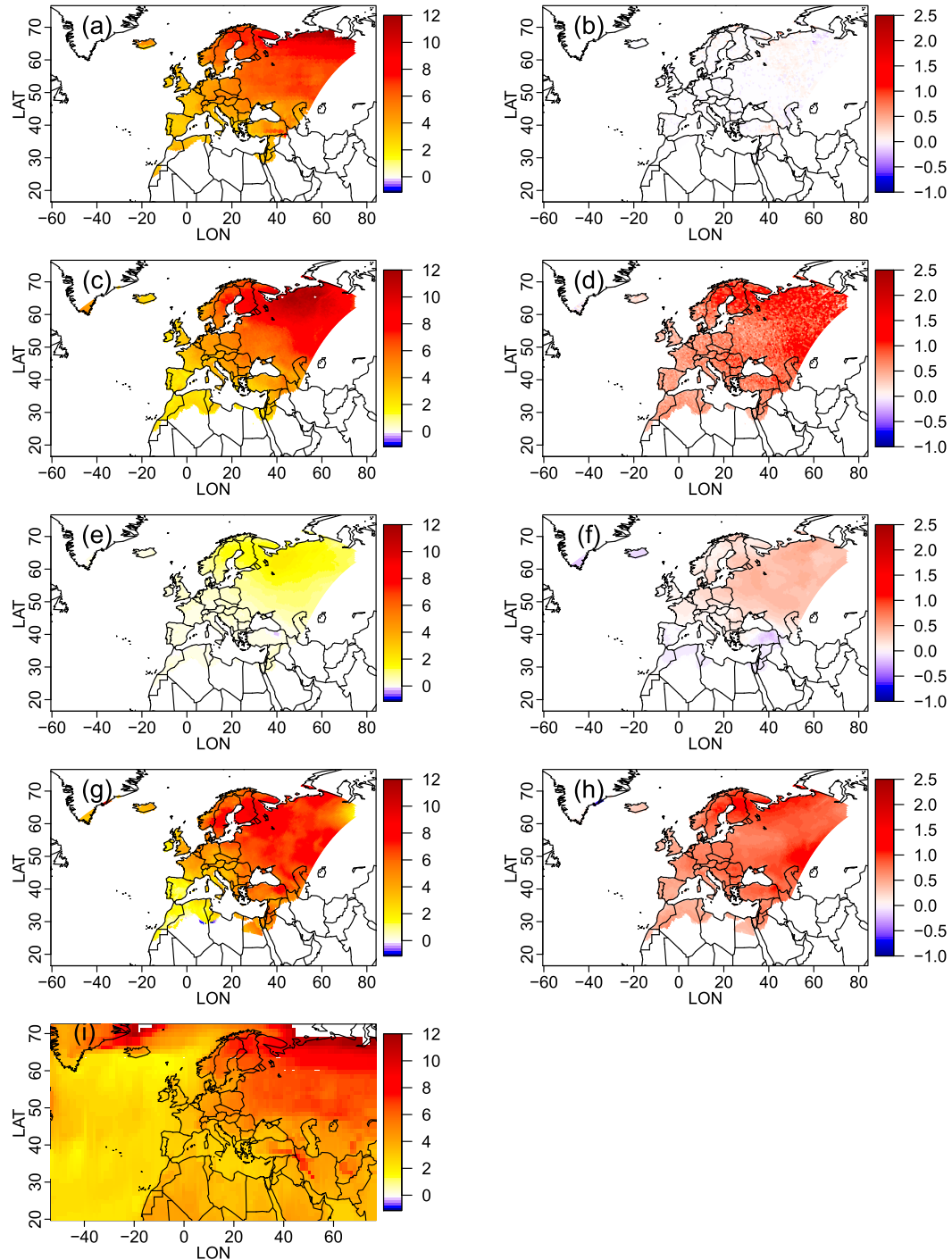


FIG. 5. (left) Maps of the mean winter temperature changes (future minus historical data) from the four SDMs driven by raw CNRM predictors [(a) CDFt, (c) SWG, (e) ANALOG, and (g) GAM] and (i) the map of the mean winter temperature changes as seen by the raw and not downscaled CNRM simulations. (right) Maps of the differences of mean winter temperature changes from 1976–2005 to 2071–2100 between BC-CNRM-driven and raw-CNRM-driven SDM simulations [i.e., change of SDM(BC predictors) – change of SDM(raw predictors)] for (b) CDFt, (d) SWG, (f) ANALOG, and (h) GAM. The units are degrees Celsius for all panels.

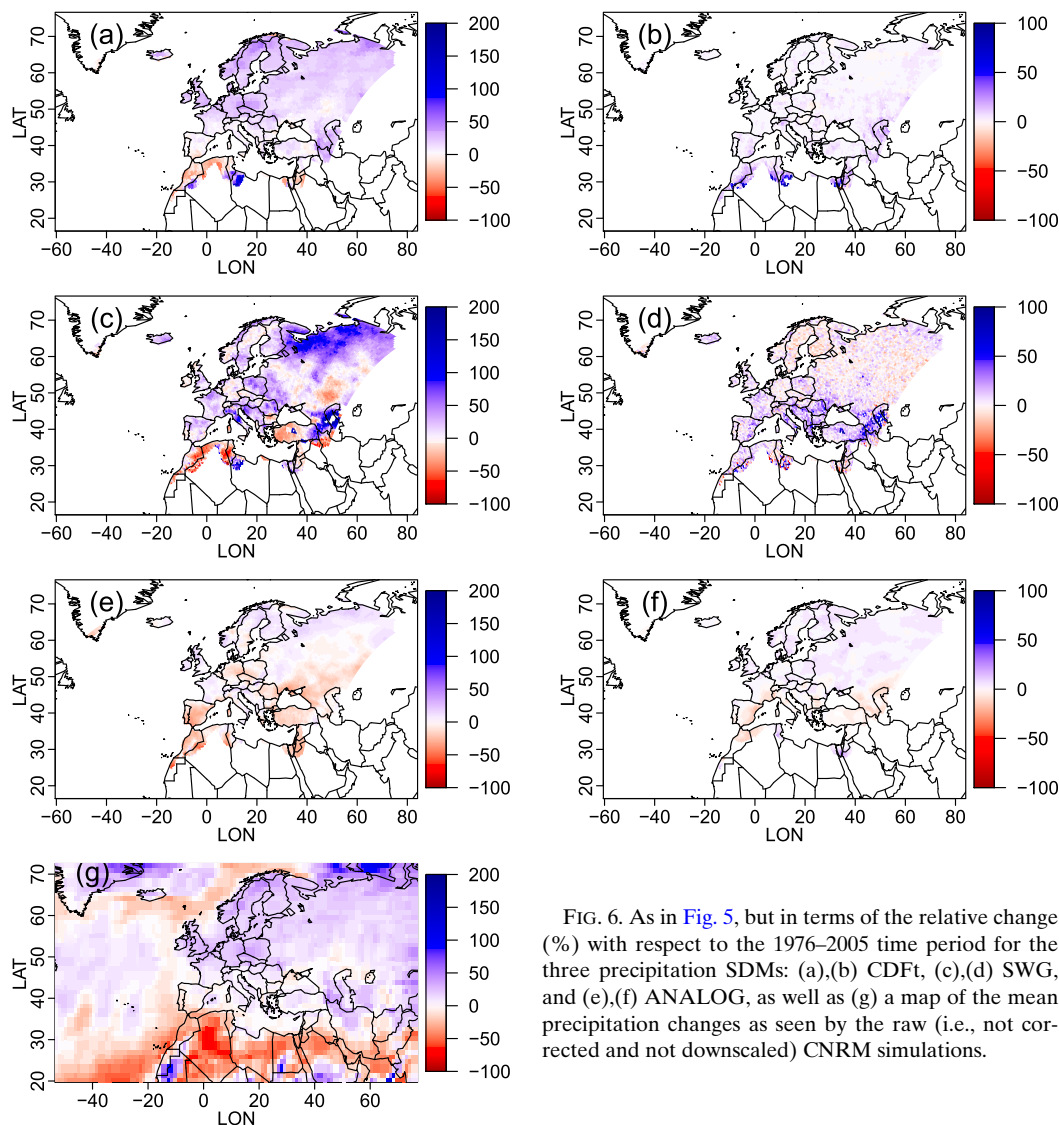


FIG. 6. As in Fig. 5, but in terms of the relative change (%) with respect to the 1976–2005 time period for the three precipitation SDMs: (a),(b) CDFt, (c),(d) SWG, and (e),(f) ANALOG, as well as (g) a map of the mean precipitation changes as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations.

the mean of the statistically downscaled temperature simulations.

All temperature changes seen by the SDMs driven by raw GCM predictors (left column) are consistent with the temperature change from the GCM, seeing a global increase of temperature. However, if the intensity of this warming is relatively coherent (with respect to the GCM temperature) for CDFt, SWG, and GAM (Figs. 5a,c,g), this is not the case for ANALOG (Fig. 5e), which strongly underestimates it. Moreover, when looking at the differences of temperature changes (right column) between BC-CNRM-driven SDM simulations and raw-CNRM-driven SDM simulations, it is clear that bias correction of the predictors has no impact on CDFt but slightly accentuates the warming projected by ANALOG and strongly emphasizes the warming simulated by SWG and GAM.

Figure 6 shows the equivalent of Fig. 5 but in terms of relative changes (in percent) for the three precipitation SDMs, as well as the map of the mean precipitation relative changes as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations (Fig. 6g). As in Fig. 5, the ANALOG method driven by raw predictors (Fig. 6e) underestimates the changes relative to those from the raw GCM (Fig. 6g). While SWG (Fig. 6c) projects too high mean precipitation changes, the CDFt model changes (Fig. 6a) are very close to that of the initial GCM. When bias correcting the predictors, the (relative) differences of changes (right column) are small: CDFt (Fig. 6b) is not much influenced by the BC except over some regions in North Africa, ANALOG is almost not influenced at all by bias correction, and SWG (Fig. 6d) shows the

biggest influence (with an increase of mean precipitation) around the Black Sea.

When comparing Fig. SM7 (from IPSL) with Fig. 5 (from CNRM), the main difference is observed with ANALOG for the differences (between BC-GCM-driven simulation and raw-GCM-driven simulation) of mean temperature changes in Fig. SM7f and Fig. 5f. Indeed, for IPSL-driven ANALOG simulations, the difference of changes is negative over the northeastern area of the domain, while it remains only positive for CNRM-driven ANALOG simulations. In terms of changes in (relative) mean precipitation, for IPSL-driven precipitation simulations (Fig. SM8) the mean changes (left column), as well as the differences of mean changes (right column), are more pronounced than for CNRM-driven simulations (Fig. 6). This seems to be due partially to more pronounced changes in mean IPSL projections. The major difference between Fig. 6 and Fig. SM8 is for SWG differences of mean changes (Fig. SM8d and Fig. 6d). While for CNRM (Fig. 6d), the differences are essentially positive—indicating an increase of the precipitation mean changes due to the bias correction step—they can be also negative over a large portion of the domain for IPSL (Fig. SM8d), showing opposite influences of the bias correction depending on the location.

Changes in variance can also be studied. The left column of Fig. 7 displays the changes (in percent) of the variance of the winter daily temperatures (i.e., $100 \times$ future variance/historical variance) from the four temperature SDMs driven by raw CNRM predictors as well as a map of the temperature variance changes as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations. The right column of Fig. 7 shows the ratios (in percent) of the winter temperature variance changes from the BC-CNRM-driven SDM simulations over those of the raw-CNRM-driven SDM simulations, for the four temperature SDMs. Hence, the right column allows us to visualize the influence of bias correcting the SDM predictors on the variability of the statistically downscaled temperature simulations.

When driven by raw predictors, all SDMs mostly show a decrease of temperature variability in winter [except GAM (Fig. 7g) around the Black Sea], in agreement with the raw GCM (Fig. 7i). Once more, ANALOG displays the weakest intensity of change. Looking at the right column of Fig. 7, the influence of the bias correction of the predictors is mainly to emphasize this decrease of temperature variability for SWG, ANALOG, and GAM (Figs. 7d,f,h), although GAM (Fig. 7h) can have a higher variance (due to bias correction) around the Mediterranean basin. CDFt (Fig. 7b) is the only SDM to see globally its BC-predictor-driven temperature projections

with a higher variability than its raw-predictor-driven projections.

Figure 8 is the same as Fig. 7 but for the three precipitation SDMs, and it also presents the map of the changes in the winter precipitation variances as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations. If CDFt (Fig. 8a) and SWG (Fig. 8c)—driven by the raw predictors—and the raw GCM (Fig. 8g) are in relative agreement with an increase of the precipitation variance, ANALOG (Fig. 8e) mainly displays a decrease of the daily precipitation variability. Note that, once more, SWG seems to emphasize the changes of variance seen by the GCM, while CDFt shows the changes most similar to those of the GCM. Moreover, bias correcting the predictors before applying the SDMs globally increases the variability even more for CDFt (Fig. 8b) and slightly more for SWG (Fig. 8d)—although some decreases with respect to raw-predictor-driven SWG are visible—and does not impact the variance changes from the ANALOG projections (Fig. 8f).

The results and conclusions are quite similar for both GCMs (Fig. 7 for CNRM and Fig. SM9 for IPSL), with the exception of the northeastern part of the domain for which the variances slightly increase in time with IPSL for CDFt, SWG and, to some extent, for ANALOG (Figs. SM9a,c,e). Interestingly, this feature is not present for GAM (Fig. SM9g) but is clearly visible on the variance changes from the raw IPSL simulations (Fig. SM9i). However, in terms of ratios of variance changes (right columns), the strongest dissimilarity appears for ANALOG (Figs. 7f and SM9f) over the northeastern area that shows ratios above 100% for IPSL (Fig. SM9f) while CNRM only has ratios below 100% (Fig. 7f).

Furthermore, like for changes in relative mean precipitation (Figs. 6 and SM8), when comparing the influences of the two GCMs in Fig. 8 (CNRM) and Fig. SM10 (IPSL), the major difference appears for the ratios of variance changes (right columns) with SWG in Figs. 8d and SM10d. With this method, while CNRM ratios are mostly above 100%, indicating that the bias correction of the CNRM predictors in SWG increases the final precipitation variance changes, IPSL ratios also have many values below 100%, meaning that the bias correction of the IPSL predictors decreases the final precipitation variance changes.

2) STATIONARITY OF INFLUENCES ON MEAN AND VARIANCE

A common question in statistical downscaling or in bias correction contexts is, “what is the strength of the stationarity assumption underlying those statistical

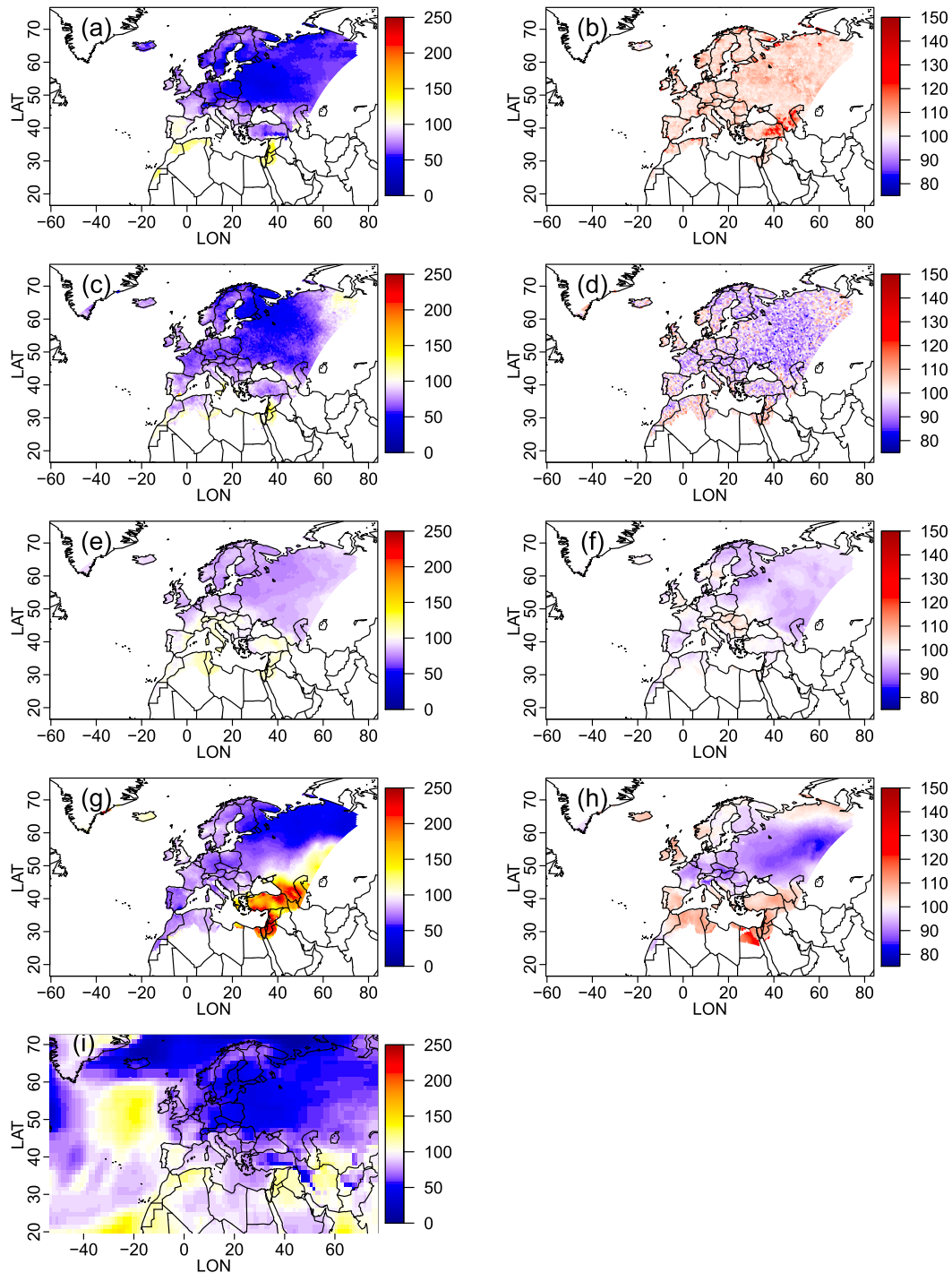


FIG. 7. (left) Maps of the changes (%) of the variance of the winter daily temperatures (i.e., $100 \times \text{future variance} / \text{historical variance}$) from the four temperature SDMs [(a) CDFt, (c) SWG, (e) ANALOG, and (g) GAM] driven by raw CNRM predictors, as well as (i) a map of the temperature variance changes as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations. (right) Ratios (%) of the winter temperature variance changes from the BC-CNRM driven SDM simulations over those of the raw-CNRM driven SDM simulations, for (b) CDFt, (d) SWG, (f) ANALOG, and (h) GAM.

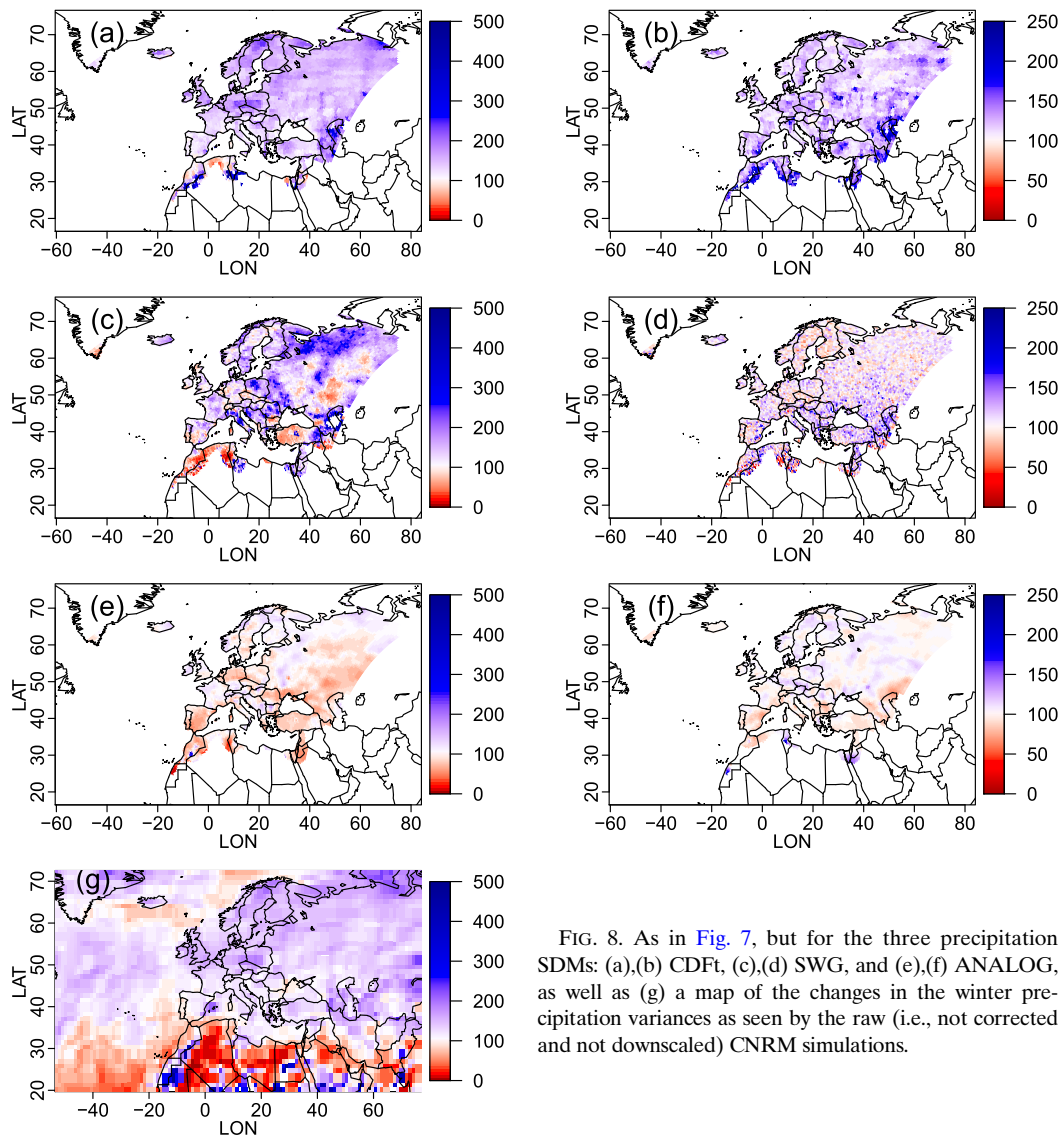


FIG. 8. As in Fig. 7, but for the three precipitation SDMs: (a),(b) CDFt, (c),(d) SWG, and (e),(f) ANALOG, as well as (g) a map of the changes in the winter precipitation variances as seen by the raw (i.e., not corrected and not downscaled) CNRM simulations.

models?” In this part, this question is studied in the following way: for each variable (temperature and precipitation) and for each statistical criterion (mean and variance), the difference between the SDM simulations driven by the corrected predictors and those driven by the raw predictors is calculated [i.e., $SDM(BC) - SDM(raw)$], not only over the historical time period (1976–2005) but also over the future time period (2071–2100). Those two differences (i.e., one for the historical period, one for the future period) can then be compared to evaluate the stationarity in time of the influence of applying bias correcting to predictors on the SDM results.

Figure 9 shows, for each temperature SDM, the differences between the mean winter SDM temperature

simulations driven by the corrected CNRM predictors and those driven by the raw CNRM predictors [i.e., $SDM(BC) - SDM(raw)$]. On the left column, the differences are calculated over the 1976–2005 period, while on the right column, it is over the 2071–2100 period. Those maps allow us to visualize the change in downscaled temperatures when the SDM is driven by bias corrected predictor, with respect to downscaled temperatures driven by raw GCM predictors.

Looking at the left column of Fig. 9 (i.e., historical differences), as expected from the evaluations on the historical time period and the previous results, differences are clearly visible from one temperature SDM to another, except for SWG and GAM (Figs. 9c,g), which are quite similar and have a strong spatial pattern.

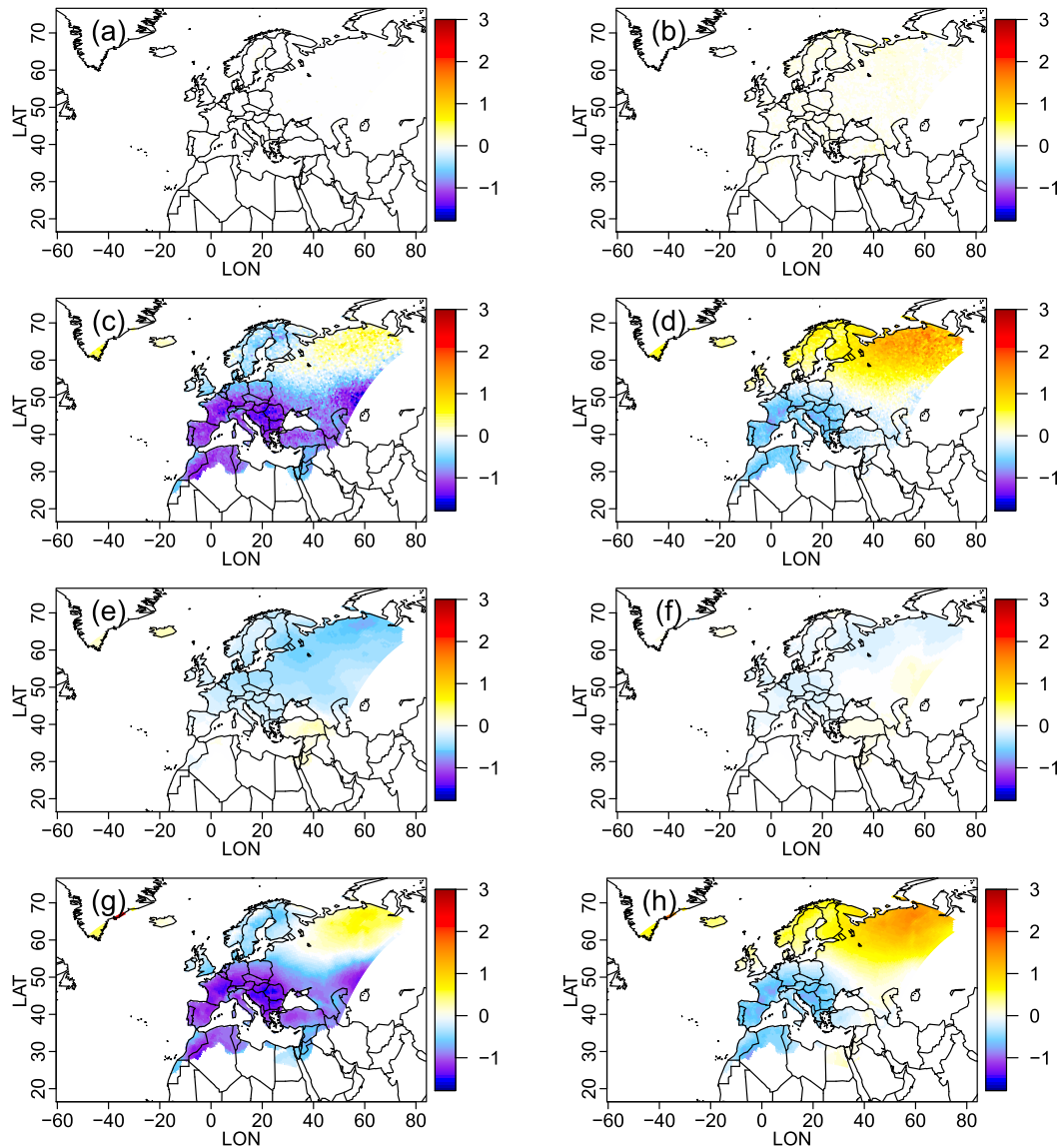


FIG. 9. Differences ($^{\circ}\text{C}$) between the mean winter SDM temperature simulations driven by the corrected CNRM predictors and those driven by the raw CNRM predictors [i.e., $\text{SDM}(\text{BC})$ minus $\text{SDM}(\text{raw})$]. The differences are calculated over (left) the 1976–2005 period and (right) the 2071–2100 period for (a),(b) CDFt, (c),(d) SWG, (e),(f) ANALOG, and (g),(h) GAM.

ANALOG (Fig. 9e) and CDFt (Fig. 9a) show only small negative differences and no differences, respectively. In the right column, we can see that the differences over the 2071–2100 period are not exactly the same as in the historical period. Globally, the patterns of differences are equivalent between the two periods but the intensities are not: SWG and GAM (Figs. 9d,h) still have equivalent patterns and intensities, but the latter ones have been shifted by about $+1^{\circ}\text{C}$. This is the opposite for ANALOG (Fig. 9f), which sees a smaller but negative difference (about -0.5°C) between BC-predictor-driven and raw-predictor-driven models, while CDFt

shows equivalent and very close to zero differences for both historical and future periods.

Figure 10 is the same as Fig. 9 but for the ratio (in percent) of mean precipitation from the three precipitation SDMs. This time, patterns are very similar from the historical and future time periods. CDFt (Figs. 10a,b) and ANALOG (Figs. 10e,f) show only small ratio values, while SWG (Figs. 10c,d) has stronger intensities (both lower and higher than 100% values), but the changes brought by the bias correction of the predictors in the three SDMs, for mean precipitation, are not much influenced by the considered time period.

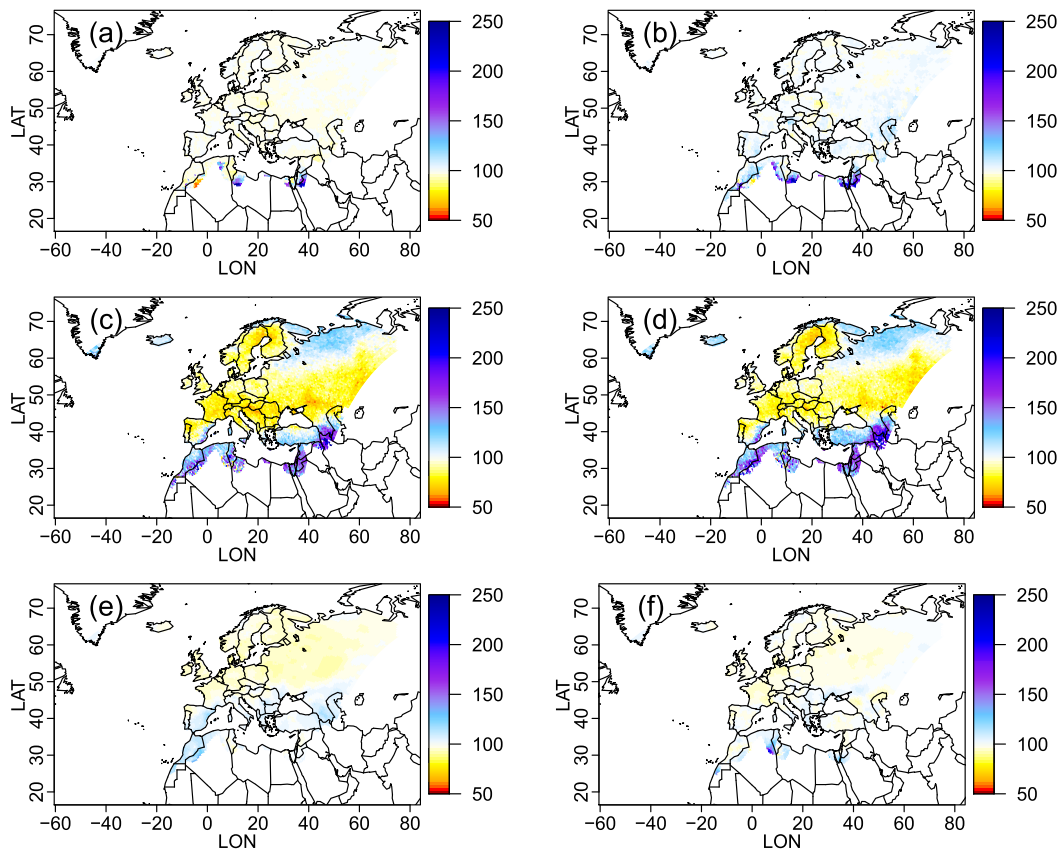


FIG. 10. As in Fig. 9, but for relative mean precipitation (%) from the three precipitation SDMs: (a),(b) CDFt, (c),(d) SWG, and (e),(f) ANALOG.

This is not necessarily the case when looking at the stationarity of the changes brought in bias correcting the predictors in terms of variance, for both temperature and precipitation. Figure 11 shows, for the four temperature SDMs, the ratios (in percent) of winter SDM temperature variance from SDMs applied to bias corrected or raw predictors {i.e., $100 \times \text{var}[\text{SDM}(\text{BC})]$ divided by $\text{var}[\text{SDM}(\text{raw})]$ }. In the left column, the ratios and variances are calculated over the 1976–2005 period, while in the right column, it is over the 2071–2100 period. As previously, in the left column, the variance ratios are different from one SDM to another, with CDFt having the ratio closest to 100%, indicating no changes (in terms of variance) over the historical period brought by the bias correction of the predictors. Comparing the left and right columns of Fig. 11, it is clear that the changes in winter daily temperature variances brought in correcting predictors are not stationary between 1976–2005 and 2071–2100. Bias correcting the predictors increases the variance over the future time period for CDFt (Fig. 11b) and mostly decreases the variance in future for SWG (Fig. 11d) and GAM (Fig. 11h), more in the future than in the historical period. Interestingly, if

BC of the predictors slightly increases the variance of the ANALOG historical projections (Fig. 11e), this has no impact (on the variance) over the future time period (Fig. 11f), implying a nonstationary behavior.

Figure 12 shows the variance ratio maps equivalent to Fig. 11 but for the precipitation ratios from three precipitation SDMs. Here, as for the relative mean precipitation in Fig. 10, over the historical period, CDFt (Fig. 12a) and ANALOG (Fig. 12e) are less influenced by the bias correction than SWG (Fig. 12c). However, comparing left (historical) and right (future) columns, if SWG (Figs. 12c,d) and ANALOGS (Figs. 12e,f) show temporal stationarity in the change of variances with very similar patterns and intensities of precipitation variance ratios, this is not the case for CDFt (Figs. 12a,b), which goes from almost no impact of the bias correction of the predictors over 1976–2005 (Fig. 12a close to 100%) to an increase of variance (ratio $\sim 150\%$ on average) over 2071–2100.

When comparing results from the two downscaled GCMs, first in terms of stationarity of influence on mean temperature, the conclusions are very similar for the two GCMs, except that the intensities of the patterns are

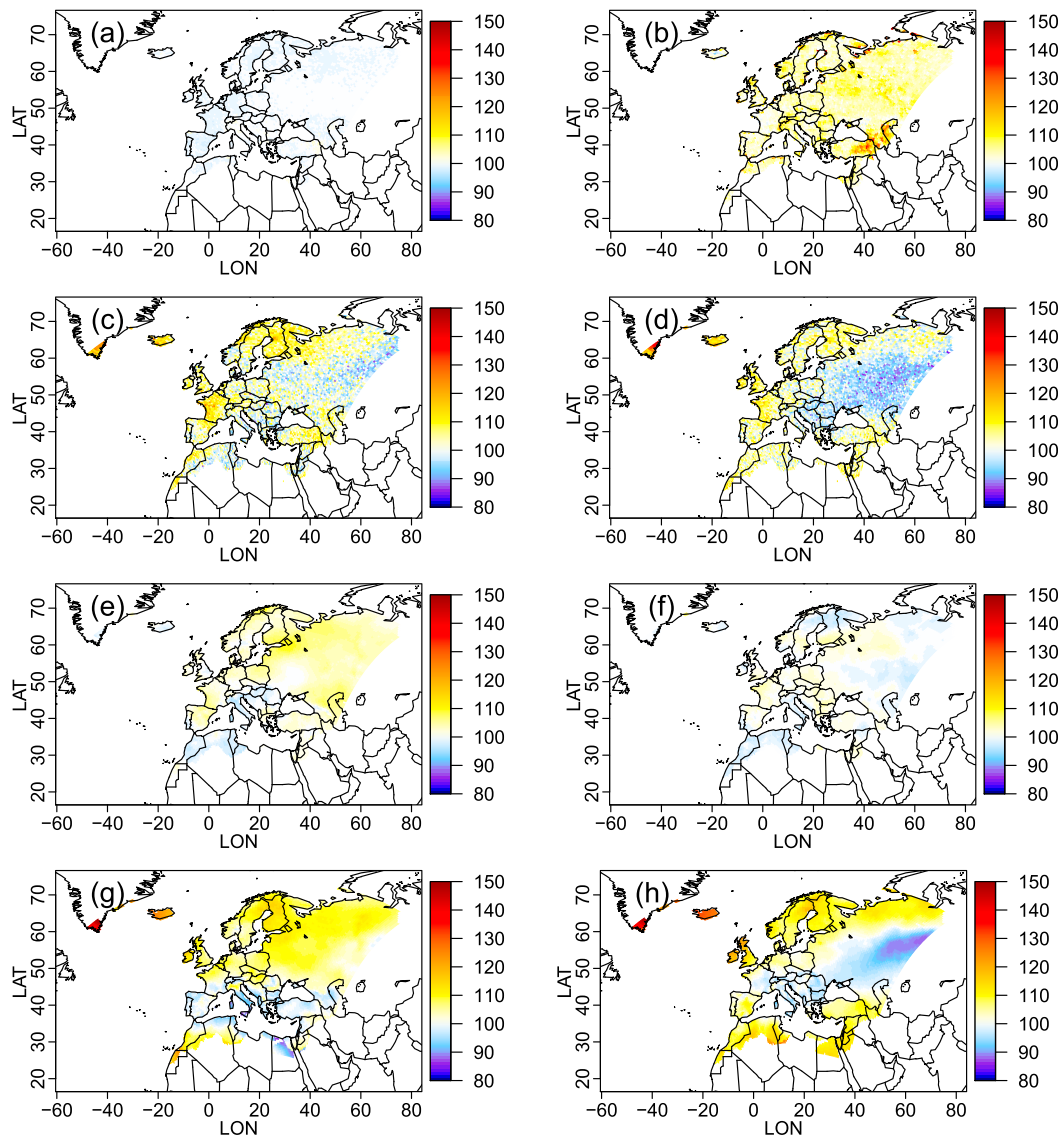


FIG. 11. Ratios (%) of winter SDM temperature variance from SDMs applied to bias corrected or raw predictors {i.e., $\text{var}[\text{SDM}(\text{BC})]$ divided by $\text{var}[\text{SDM}(\text{raw})]$ }. Ratios and variances are calculated over (left) the 1976–2005 period and (right) the 2071–2100 period for (a),(b) CDFt, (c),(d) SWG, (e),(f) ANALOG, and (g),(h) GAM.

stronger over the 2071–2100 (right columns) for SWG and GAM when driven by IPSL (Figs. SM11d,h) than when driven by CNRM (Figs. 9d,h). In other words, bias correcting the IPSL simulations tends to reinforce the temporal nonstationarity for SWG and GAM more than bias correcting the CNRM simulations. This is the same for mean precipitation (cf. Figs. 10 and SM12), where bias correcting the IPSL outputs reinforces the temporal nonstationarity for SWG more than bias correcting the CNRM simulations (Figs. 10d and SM12d).

In terms of stationarity of influence on temperature variance, although slight differences are visible, results

are relatively equivalent for CNRM-driven ratios (Fig. 11) and IPSL-driven ones (Fig. SM13), for the CDFt, SWG, and ANALOG models (Figs. 11a–f and SM13a–f). However, GAM presents stronger differences with IPSL (Fig. SM13) than with CNRM (Fig. 11), especially over 2071–2100 (Figs. 11h and SM13h), with temperature variance ratios close to 150% in the southeastern part of the Mediterranean basin, or close to 80% over Russia. For precipitation variance, the spatial patterns and conclusions are very similar for the two GCMs (cf. Figs. 12 and SM14), even if some changes are visible (e.g., more pronounced intensities for SWG in 2071–2100; Figs. 12d and SM14d).

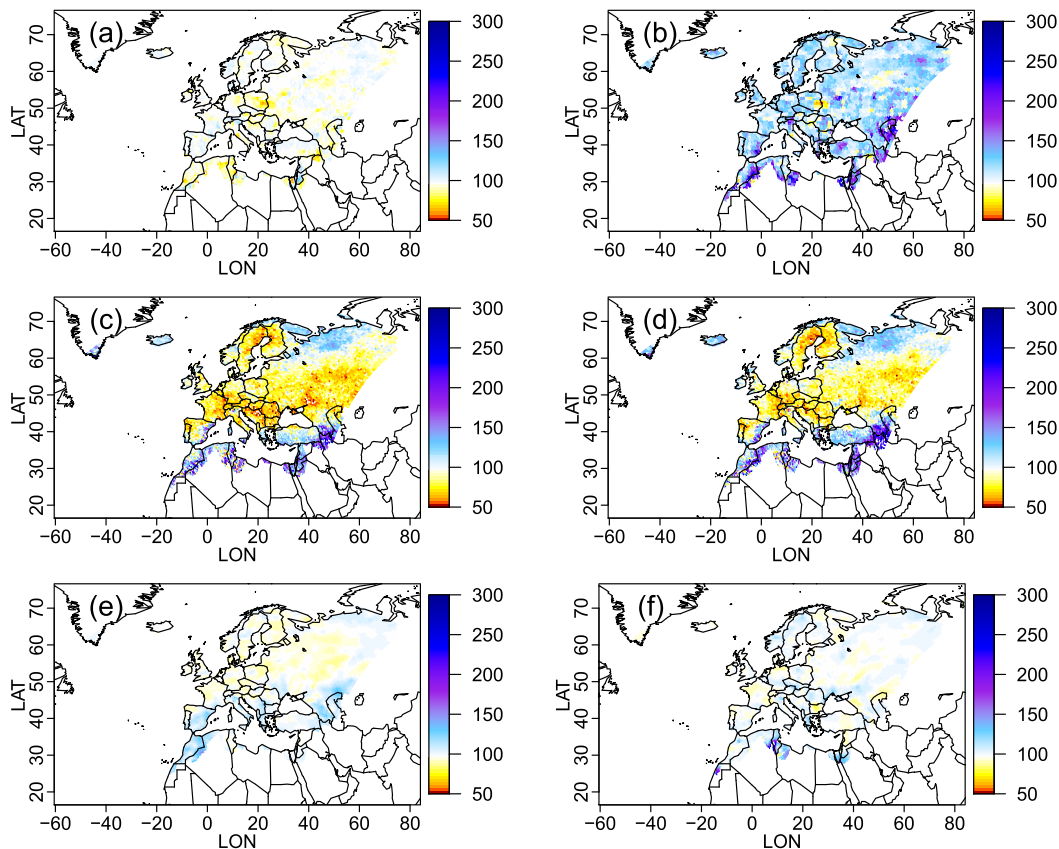


FIG. 12. As in Fig. 11, but for the precipitation ratios from the three precipitation SDMs: (a),(b) CDFt, (c),(d) SWG, and (e),(f) ANALOG.

As a global conclusion, the influences brought by the BC of the SDM predictors, both in terms of mean and variance, are not always the same from one SDM to another, from one variable to another, and from one GCM to another, but they also are not always the same over the historical period and over the end of the twenty-first century.

5. Conclusions and discussion

This study has investigated the influence of bias correcting (or not bias correcting) the large-scale predictors used as inputs into SDM simulations of daily temperatures and precipitation. To do so, four temperature SDMs (CDFt, SWG, ANALOG, GAM) and three precipitation SDMs (CDFt, SWG, ANALOG) have been calibrated based on reanalyses and observed data over a historical time period.

To evaluate the improvements or degradations of the quality of the statistical simulations depending on the correction or not of the predictors, the calibrated SDMs were then forced by historical predictors coming from two GCMs either corrected or not corrected, and

those two types of simulations were compared with SDM simulations driven by reanalyses. The main conclusions in terms of mean and variance biases (with respect to reanalysis-driven SDMs) were that, for both temperature and precipitation, the bias correction may have some important impacts on the SDM simulations. However, the influence of the BC of the predictors is not the same from one SDM to another—for example, CDFt is mostly not influenced while SWG is much improved—with different spatial structures and may also depend on the considered statistical properties (mean or variance). In order not to repeat the detailed results presented in section 4a, Table 1 summarizes the main conclusions of the “historical” evaluations.

The influence of bias correcting the SDM predictors on the future change in some basic statistical properties of the SDM simulations has also been studied. In this climate change context, the SDMs were driven by GCM future projections in their corrected or uncorrected versions. The temporal mean and variance changes of the BC-predictor-driven SDMs precipitation and temperature simulations were compared with those of the raw-predictor-driven SDMs, as well as with those of the

TABLE 1. Main conclusions about the influence of the bias correction of the GCM predictors on the mean and variance historical biases of the SDM simulations with respect to ERA-I-driven SDM simulations. See [section 4a](#) for details.

SDMs	Mean temperature	Variance temperature	Relative mean PR	Variance PR
CDFt	Not influenced	Not influenced	Not influenced	Not influenced (range reduced)
SWG	Influenced (improved)	Influenced (improved)	Influenced (improved)	Influenced (improved)
ANALOG	Moderately influenced (improved)	Not influenced	Moderately influenced (improved)	Not influenced
GAM	Influenced (improved)	Influenced (improved)	—	—

raw GCM. Globally, raw-predictor-driven CDFt, SWG, and GAM (for temperature only) changes are more or less in agreement with raw GCM changes—although some spatial-, variable-, or criterion-related specificities were visible—while raw-predictor-driven ANALOG had a much weaker intensity of changes. The influence of the bias correction on those SDM changes depends again on the SDM, the considered variable, and the statistical criterion. [Table 2](#) provides a synthetic overview of the main conclusions seen from [section 4b\(1\)](#). As somehow expected, in terms of mean and variance changes, the ANALOG method is only moderately influenced (at the maximum) by the bias correction of its predictors. Indeed, as ANALOG resamples data from the calibration dataset, it cannot create data out of this set, hence reducing the potential changes that the method can produce.

Finally, for the two variables, the temporal stationarity of the changes brought by the BC of the predictors has been investigated in terms of means and variances. Those changes (i.e., differences or ratios) have been computed over the historical (1976–2005) and future (2071–2100) time periods and compared to assess the (non)stationarity. Once more, the results were versatile, depending on the considered variable, criterion, and SDM, but some potentially strong nonstationarities were found. [Table 3](#) gives a short summary of the main findings of [section 4b\(2\)](#). Hence, this study brought some useful information to use to properly interpret the

simulations provided by statistical downscaling models, either with bias corrected or with uncorrected predictors from GCM as input.

Furthermore, one classical advantage of applying bias correcting to GCMs is to reduce the disagreements among the model outputs ([Ahmed et al. 2013](#)). Therefore, it is also interesting to see if removing biases from GCM predictors reduces the differences between the CNRM-driven and the IPSL-driven SDM performances.

First, in the “historical evaluation” context, the differences between “bias” values from the raw IPSL and the raw CNRM experiments are compared with the differences in the case of the BC IPSL and the BC CNRM experiments. This is presented via box plots in [Fig. 13](#) for winter, where the top row shows the results for temperature and the bottom row for precipitation, and the left column is for differences of mean biases while the right column is for differences of variance ratios. The equivalent box plots for summer are given in [Fig. SM35](#). For CDFt, the bias correction does not reduce the disagreements between the two GCMs (both in terms of differences of mean bias and variance ratio), for both seasons and variables. In winter ([Fig. 13](#)), for both temperature and precipitation, the disagreement (for mean and variance) does not change much by the bias correction for SWG and ANALOG, although the variability (i.e., spread of the box plots) is usually reduced. However, for GAM (i.e., temperature only), bias

TABLE 2. Main conclusions about the influence of the bias correction of the predictors on the mean and variance changes of the SDM simulations. See [section 4b\(1\)](#) for details.

SDMs	Mean temperature	Variance temperature	Relative mean PR	Variance PR
CDFt	No influence	Influence (increased variance)	No influence	Strong influence (increased variance)
SWG	Influence (increased warming)	Influence (decreased variance)	Moderate influence (increased PR)	Moderate influence (mostly increased variance)
ANALOG	Moderate influence (increased warming)	Moderate influence (decreased variance)	No influence	Weak influence (both signs)
GAM	Influence (increased warming)	Strong influence (both signs)	—	—

TABLE 3. Main conclusions concerning the temporal stationarity of changes—in means and variances—brought by the BC of the predictors for the SDM simulations. See section 4b(2) for details.

SDMs	Mean temperature	Variance temperature	Relative mean PR	Variance PR
CDFt	Stationary	Nonstationary (increased changes)	Stationary	Nonstationary (increased changes)
SWG	Nonstationary (increased changes)	Moderately nonstationary (some regions)	Stationary	Stationary
ANALOG	Moderately nonstationary (decreased changes)	Nonstationary (decreased changes)	Stationary	Stationary
GAM	Nonstationary (increased changes)	Moderately nonstationary (some regions)	—	—

correcting the predictors reduces this variability more strongly in terms of mean bias differences and, surprisingly, increases the disagreement in terms of variance ratios. In summer (Fig. SM35), except for CDFt where no clear difference is visible, the bias correction seems to reduce the disagreement between the two GCMs for most SDMs and both variables, including a generally smaller variability of the differences. This is clearly visible for the mean bias differences with SWG and GAM.

Second, the same analysis is also conducted in terms of differences of “changes” (from present to future) between the raw IPSL and the raw CNRM experiments, which are compared with the differences of “changes” in the case of the BC IPSL and the BC CNRM experiments. Results are shown in Fig. 14 for winter and in Fig. SM36 for summer. In both figures, for the two variables, the influence of bias correcting the predictors is very limited on reducing the disagreement between

IPSL and CNRM. This means that the differences of changes (from present to future) between the two downscaled GCMs, both in terms of mean and variance, are basically conserved with the bias correction of the predictors. Although the biases of the downscaling output can be reduced by the correction step, this indicates that the specificities of the GCM evolutions are preserved anyway.

Moreover, some choices have been made to design the present study. Those can of course potentially affect the results. For example, the choice was made to correct “raw” predictors, even for SWG and GAM that have PCs as predictors. Imbert and Benestad (2005) found that bias correcting the RCM PCs with respect to re-analyses (instead of adjusting the raw climate variables and then computing the PCs) improved the results of linear and analog-based statistical downscaling methods. However, we have not followed this approach

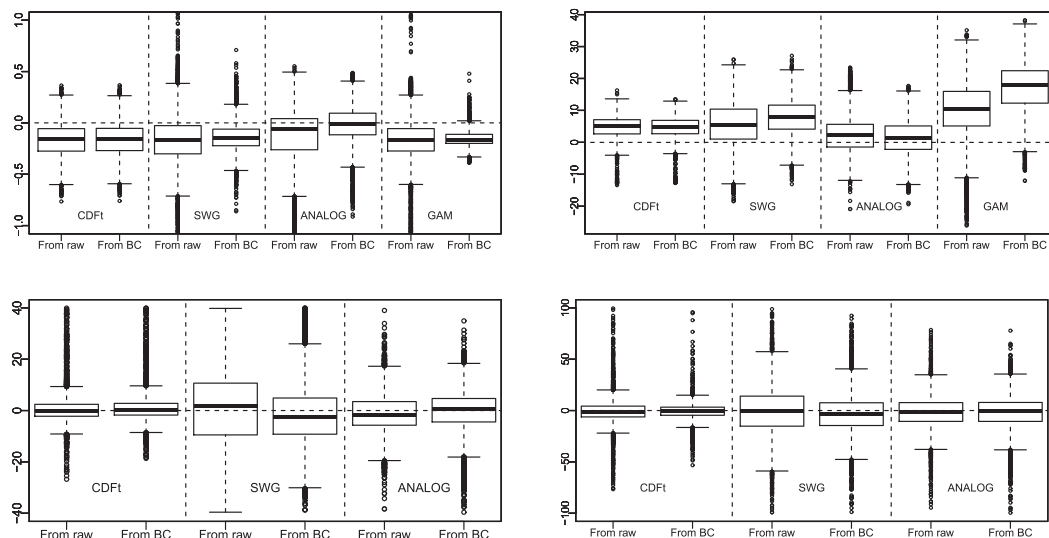


FIG. 13. In the “historical evaluation” context, box plots of differences between “bias” values from the raw IPSL and raw CNRM experiments, as well as box plots of differences in the case of the BC IPSL and the BC CNRM experiments: results for (top) temperature, and (bottom) precipitation for differences of (left) mean biases and (right) variance ratios. All panels are for winter (DJF).

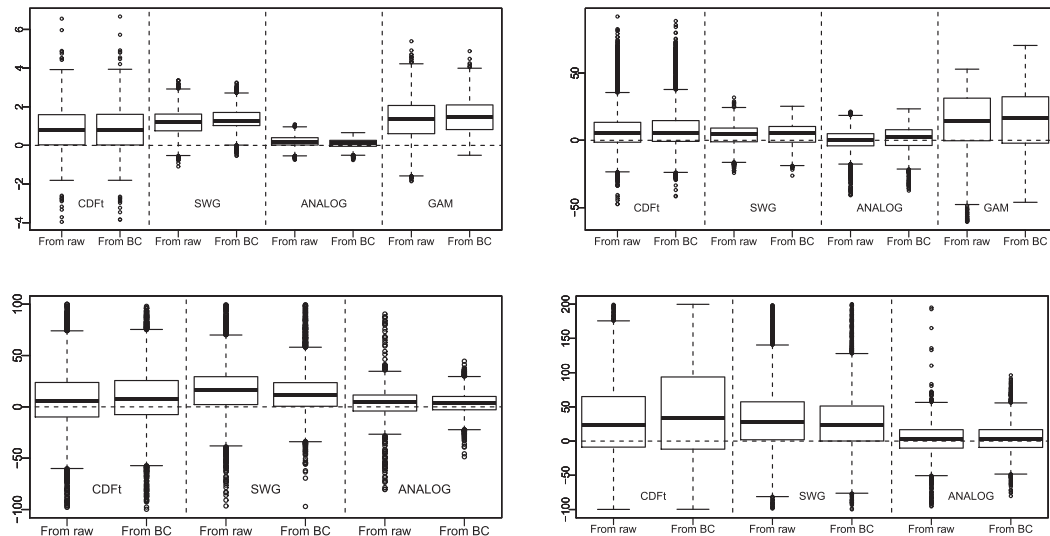


FIG. 14. As in Fig. 13, but in terms of differences of changes (from present to future) instead of biases.

in this study for two reasons. First, one can wonder if bias correcting the climate model first PC with respect to reanalyses first PC makes sense if they do not explain the same percentage of variance. This question is left for future work. Furthermore, we wanted the bias correction to be applied the same way for the different SDMs tested in this study. Since ANALOG and CDFt are applied respectively to the fields of anomalies and to raw values—that is, not to the PCs—the correction of the PCs for SWG and GAM would have made results less comparable between SDMs. However, this PC correction is an interesting perspective that would deserve further methodological research and analyses.

Another choice was to base the present study on a single bias correction method, CDFt. Indeed, the goal was not to demonstrate the added value of a specific BC technique over another one, but to highlight that the application of a BC method to predictors that will force/drive SDMs may have (positive or negative) consequences on the quality and properties of the SDM outputs. The CDFt method has been chosen here 1) because it is derived from the QM approach that is certainly one of the most widely used techniques; and 2) because it allows accounting for the changes in the climate model CDFs. However, Pierce et al. (2015) found that CDFt and QM “can alter the original GCM projected monthly mean change” for both temperature and precipitation. Hence, the use of other BC methods, such as trend or change preserving techniques (e.g., equidistant CDF matching or quantile delta mapping; Pierce et al. 2015; Cannon et al. 2015), can potentially lead to different results. Nevertheless, as CDFt and QM are quite

classical methods for bias correction, the analysis of the consequences of their use to bias correct predictors to enter in SDMs remains of high interest. Studying the impacts of multiple BC methods in such a context is out of the scope of the present article and is therefore left for future work.

Note that the statistically simulated data used in this study should eventually be available on the Earth System Grid Federation (<http://esgf.llnl.gov/>) node but are already available upon request to the authors.

The present work can obviously be extended further in many manners. For example, only one BC method has been employed here (CDFt). However, some preliminary analyses have also been performed with a classical quantile–quantile BC method, indicating that the main conclusions are robust (not shown). Nevertheless, increasing the number of BC models to realize the same analyses as done in this paper could be necessary to have a more global overview of the impact of bias correction in such a context.

Similarly, using more GCMs would help in generalizing the conclusions of such a study. Of course, the same analyses could also be carried out with RCMs instead of GCMs to understand the influence of the resolution (GCM versus RCM as predictors) or the influence of different coupling in RCM on the final statistically downscaled results (e.g., ocean–atmosphere coupled RCM versus stand-alone atmosphere RCM).

Moreover, this study and the associated analyses are not restricted to SDMs and can certainly be performed with (or including) RCMs instead of (or in addition to) SDMs. Some attempts have been made in this direction, for example, by Colette et al. (2012) or Bruyère et al.

(2014), evaluating RCM runs forced by bias corrected GCM data.

It is important to keep in mind that spatial or inter-variable dependencies were not accounted for, neither in the BC method nor in the SDMs tested in this study. The introduction of such specificities may have consequences on the results and conclusions brought by the present analyses. This therefore remains an important perspective for future work.

Acknowledgments. This work has been partially supported by the ANR-project StaRMIP (Grant ANR-12-JS06-0005-01); the CLIP-C project (FP7 Grant 607418); the ANR-project REMEMBER (Grant Agreement ANR-12-SENV-0001-01); the GICC REMedHE project (2012–15) funded by the French Ministry of Ecology, Sustainable Development and Energy; the EXTREMOSCOPE project (supported by the French Ministry of Ecology); and the LABEX-IPSL project (Grant ANR-10-LABX-18-01). It was supported by the IPSL group for regional climate and environmental studies. All computations have been made in R.

REFERENCES

- Ahmed, K. F., G. Wang, J. Silander, A. M. Wilson, J. M. Allen, R. Horton, and R. Anyah, 2013: Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. Northeast. *Global Planet. Change*, **100**, 320–332, doi:10.1016/j.gloplacha.2012.11.003.
- Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126, doi:10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2.
- Bruyère, C. L., J. M. Done, G. J. Hollanda, and S. Fredrick, 2014: Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dyn.*, **43**, 1847–1856, doi:10.1007/s00382-013-2011-6.
- Buishand, T. A., M. V. Shabalova, and T. Brandsma, 2004: On the choice of the temporal aggregation level for statistical downscaling of precipitation. *J. Climate*, **17**, 1816–1827, doi:10.1175/1520-0442(2004)017<1816:OTCOTT>2.0.CO;2.
- Bürger, G., T. Murdock, A. Werner, S. Sobie, and A. Cannon, 2012: Downscaling extremes—An intercomparison of multiple statistical methods for present climate. *J. Climate*, **25**, 4366–4388, doi:10.1175/JCLI-D-11-00408.1.
- Cannon, A., S. Sobie, and T. Murdock, 2015: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *J. Climate*, **28**, 6938–6959, doi:10.1175/JCLI-D-14-00754.1.
- Chandler, R. E., and H. S. Wheeler, 2002: Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. *Water Resour. Res.*, **38**, 1192, doi:10.1029/2001WR000906.
- Charles, S., B. Bates, P. Whetton, and J. Hughes, 1999: Validation of downscaling models for changed climate conditions: Case study of southern Australia. *Climate Res.*, **12**, 1–14, doi:10.3354/cr012001.
- , I. N. Smith, and J. P. Hughes, 2004: Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrol. Processes*, **18**, 1373–1394, doi:10.1002/hyp.1418.
- Chen, J., F. Brissette, and R. Leconte, 2014: Assessing regression-based statistical approaches for downscaling precipitation over North America. *Hydrol. Processes*, **28**, 3482–3504, doi:10.1002/hyp.9889.
- Christensen, J., F. Boberg, O. 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.
- Colette, A., R. Vautard, and M. Vrac, 2012: Regional climate downscaling with prior statistical correction of the global climate forcing. *Geophys. Res. Lett.*, **39**, L13707, doi:10.1029/2012GL052258.
- 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.
- Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, doi:10.1007/s00382-012-1636-1.
- Fealy, R., and J. Sweeney, 2007: Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modelling approach. *Int. J. Climatol.*, **27**, 2083–2094, doi:10.1002/joc.1506.
- Frías, M. D., E. Zorita, J. Fernández, and C. Rodríguez-Puebla, 2006: Testing statistical downscaling methods in simulated climates. *Geophys. Res. Lett.*, **33**, L19807, doi:10.1029/2006GL027453.
- Frost, A., and Coauthors, 2011: A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. *J. Hydrol.*, **408**, 1–18, doi:10.1016/j.jhydrol.2011.06.021.
- Gaitan, C. F., W. W. Hsieh, and A. J. Cannon, 2014: Comparison of statistically downscaled precipitation in terms of future climate indices and daily variability for southern Ontario and Quebec, Canada. *Climate Dyn.*, **43**, doi:10.1007/s00382-014-2098-4.
- Grenier, P., A.-C. Parent, D. Huard, F. Anctil, and D. Chaumont, 2013: An assessment of six dissimilarity metrics for climate analogs. *J. Appl. Meteor. Climatol.*, **52**, 733–752, doi:10.1175/JAMC-D-12-0170.1.
- Grouillet, B., D. Ruelland, P. Vaittinada Ayar, and M. Vrac, 2016: Sensitivity analysis of runoff modeling to statistical downscaling models in the western Mediterranean. *Hydrol. Earth Syst. Sci.*, **20**, 1031–1047, doi:10.5194/hess-20-1031-2016.
- Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, 2012: Technical note: Downscaling RCM precipitation to the station scale using statistical transformations—A comparison of methods. *Hydrol. Earth Syst. Sci.*, **16**, 3383–3390, doi:10.5194/hess-16-3383-2012.
- Haddad, Z., and D. Rosenfeld, 1997: Optimality of empirical Z-R relations. *Quart. J. Roy. Meteor. Soc.*, **123**, 1283–1293, doi:10.1002/qj.49712354107.
- Hastie, T., and R. Tibshirani, 1990: *Generalized Additive Models*. Monographs on Statistics and Applied Probability, Vol. 43, Chapman and Hall, 352 pp.
- Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess, 2006: Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.*, **26**, 1397–1415, doi:10.1002/joc.1318.
- , 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.
- Hewitson, B. C., and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. *Int. J. Climatol.*, **26**, 1315–1337, doi:10.1002/joc.1314.
- Huth, R., 1999: Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Climate Res.*, **13**, 91–101, doi:10.3354/cr013091.
- Imbert, A., and R. Benestad, 2005: An improvement of analog model strategy for more reliable local climate change scenarios. *Theor. Appl. Climatol.*, **82**, 245–255, doi:10.1007/s00704-005-0133-4.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., doi:10.1017/CBO9781107415324.
- Klok, E., and A. Klein Tank, 2009: Updated and extended European dataset of daily climate observations. *Int. J. Climatol.*, **29**, 1182–1191, doi:10.1002/joc.1779.
- Laprise, R., 2008: Regional climate modelling. *J. Comput. Phys.*, **227**, 3641–3666, doi:10.1016/j.jcp.2006.10.024.
- Levavasseur, G., M. Vrac, D. M. Roche, D. Paillard, A. Martin, and J. Vandenberghe, 2011: Present and LGM permafrost from climate simulations: Contribution of statistical downscaling. *Climate Past*, **7**, 1225–1246, doi:10.5194/cp-7-1225-2011.
- Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, doi:10.1029/2009RG000314.
- Marti, O., and Coauthors, 2010: Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dyn.*, **34**, 1–26, doi:10.1007/s00382-009-0640-6.
- Meehl, G., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 747–845.
- Michelangeli, P.-A., M. Vrac, and H. Loukos, 2009: Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophys. Res. Lett.*, **36**, L11708, doi:10.1029/2009GL038401.
- Pierce, D., D. Cayan, E. Maurer, J. Abatzoglou, and C. Hegewish, 2015: Improved bias correction techniques for hydrological simulations of climate change. *J. Hydrometeorol.*, **16**, 2421–2442, doi:10.1175/JHM-D-14-0236.1.
- Robertson, A., S. Kirchner, and P. Smyth, 2004: Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *J. Climate*, **17**, 4407–4424, doi:10.1175/JCLI-3216.1.
- Rummukainen, M., 2010: State-of-the-art with regional climate models. *Wiley Interdiscip. Rev.: Climate Change*, **1**, 82–96, doi:10.1002/wcc.8.
- Salameh, T., P. Drobinski, M. Vrac, and P. Naveau, 2009: Statistical downscaling of near-surface wind over complex terrain in southern France. *Meteor. Atmos. Phys.*, **103**, 253–265, doi:10.1007/s00703-008-0330-7.
- Schmidli, J., C. M. Goodess, C. Frei, M. R. Haylock, Y. Huncheda, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. *J. Geophys. Res.*, **112**, D04105, doi:10.1029/2005JD007026.
- Vaithinada Ayar, P., M. Vrac, S. Bastin, J. Carreau, M. Déqué, and C. Gallardo, 2015: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: Present climate evaluations. *Climate Dyn.*, **46**, 1301–1329, doi:10.1007/s00382-015-2647-5.
- Vautard, R., and P. Yiou, 2009: Control of recent European surface climate change by atmospheric flow. *Geophys. Res. Lett.*, **36**, L22702, doi:10.1029/2009GL040480.
- Volz, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.*, **40**, 2091–2121, doi:10.1007/s00382-011-1259-y.
- Vrac, M., and P. Friederichs, 2015: Multivariate—intervariable, spatial, and temporal—bias correction. *J. Climate*, **28**, 218–237, doi:10.1175/JCLI-D-14-00059.1.
- , P. Marbaix, D. Paillard, and P. Naveau, 2007a: Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe. *Climate Past*, **3**, 669–682, doi:10.5194/cp-3-669-2007.
- , M. Stein, and K. Hayhoe, 2007b: Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing. *Climate Res.*, **34**, 169–184, doi:10.3354/cr00696.
- , —, —, and X.-Z. Liang, 2007c: A general method for validating statistical downscaling methods under future climate change. *Geophys. Res. Lett.*, **34**, L18701, doi:10.1029/2007GL030295.
- , P. Drobinski, A. Merlo, M. Herrmann, C. Lavaysse, L. Li, and S. Somot, 2012: Dynamical and statistical downscaling of the French Mediterranean climate: Uncertainty assessment. *Nat. Hazards Earth Syst. Sci.*, **12**, 2769–2784, doi:10.5194/nhess-12-2769-2012.
- White, R., and R. Toumi, 2013: The limitations of bias correcting regional climate model inputs. *Geophys. Res. Lett.*, **40**, 2907–2912, doi:10.1002/grl.50612.
- Wilby, R., and T. M. L. Wigley, 2000: Precipitation predictors for downscaling: Observed and general circulation model relationships. *Int. J. Climatol.*, **20**, 641–661, doi:10.1002/(SICI)1097-0088(200005)20:6<641::AID-JOC501>3.0.CO;2-1.
- Wood, A., L. Leung, V. Sridhar, and D. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.
- Yiou, P., R. Vautard, P. Naveau, and C. Cassou, 2007: Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe. *Geophys. Res. Lett.*, **34**, L21808, doi:10.1029/2007GL031981.
- Zorita, E., and H. von Storch, 1999: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *J. Climate*, **12**, 2474–2489, doi:10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2.