AMERICAN METEOROLOGICAL SOCIETY

Journal of Climate

EARLY ONLINE RELEASE

The DOI for this manuscript is doi: 10.1175/JCLI-D-13-00629.1

Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections

Benoit Hingray *

CNRS, LTHE UMR 5564, Grenoble, F-38000, France

Univ. Grenoble Alpes, LTHE UMR 5564, Grenoble, F-38000, France

Mériem Saïd

CNRS, LTHE UMR 5564, Grenoble, F-38000, France

Univ. Grenoble Alpes, LTHE UMR 5564, Grenoble, F-38000, France

Now at Agence technique de l'information sur l'hospitalisation (ATIH), 117, bd Marius Vivier Merle 69329 Lyon cedex 03

* *Corresponding author address:* Benoit Hingray, CNRS, LTHE UMR 5564, Grenoble, F-38000, France.

E-mail: benoit.hingray@ujf-grenoble.fr

ABSTRACT

A simple and robust framework is proposed for the partitioning of the different components of internal variability and model uncertainty in an unbalanced multimember multimodel ensemble (MM2E) of climate projections obtained for a suite of statistical downscaling models (SDMs) and global climate models (GCMs). It is based on the quasi-ergodic assumption for transient climate simulations. Model uncertainty components are estimated from the noise-free-signals of the different modeling chains using a two-way ANOVA framework. The residuals from the noise-free-signals are used to estimate the large and small scale internal variability components associated with each considered GCM/SDM configuration. This framework makes it possible to take into account all members available from any climate ensemble of opportunity.

Uncertainty is quantified as a function of lead time for projections of changes in temperature and precipitation produced for a mesoscale alpine catchment. Internal variability accounts for more than 80% of total uncertainty in the first decades. This proportion decreases to less than 10% at the end of the century for temperature but remains greater than 50% for precipitation. Small scale internal variability is negligible for temperature however it is similar to the large scale component for precipitation, whatever the projection lead time. SDM uncertainty is always greater than GCM uncertainty for precipitation. It is also greater for temperature in the middle of the century. The response-to-uncertainty ratio is very high for temperature. For precipitation, it is always less than one, indicating that even the sign of change is uncertain.

1. Introduction

a. *Uncertainty sources*

A critical issue in climate change impact studies is the estimation of uncertainties associated with future projections along with the estimation of the contribution of the different uncertainty sources. As pointed out by Dobler et al. (2012), very different sources of uncertainty are involved including scenario uncertainty, model uncertainty and model's internal variability.

Scenario uncertainty is related to the poorly known future of greenhouse gas emissions. Model uncertainty, also termed response uncertainty, corresponds to the dispersion between the different climate responses classically obtained with different models for the same forcing configuration. It is due to the limitations of the model structure and parameterization used to represent geophysical processes. Model uncertainty clearly concerns Global Climate Models (GCMs) used to simulate at the global scale the response of the climate system to atmospheric composition perturbations. It also concerns Regional Downscaling Models (RDMs) used to derive corresponding climate scenarios at the regional scale. Model uncertainty and scenario uncertainty are classically explored via multimodel and multiscenarios experiments (IPCC 2007).

The internal variability of a given GCM/RDM chain supposedly represents the natural variability of regional climate at daily to multi-decadal time scales (e.g. Karoly and Wu 2005). Due to the chaotic and nonlinear nature of atmospheric processes, this variability has long been observed even in a stationary climate (Madden 1976). In a non-stationary climate, this variability can remain high above the trend related to a given forcing (e.g. greenhouse gases and aerosols) (e.g. Hawkins and Sutton 2011; Deser et al. 2012; Lafaysse et al. 2014). The internal variability of a given GCM/RDM chain can actually be partitioned into its large scale and local scale components (Braun et al. 2012; Lafaysse et al. 2014). The large scale component can be attributed to the chaotic variability of the climate at the global

scale. Mainly produced by the GCM itself, it corresponds to the variability obtained from a given GCM experiment over a long time period for a stationary climate (e.g. Räisänen 2001; Deser et al. 2012). Very similar large scale atmospheric circulation patterns can then lead to very different meteorological observations at the local scale. This local scale component of natural variability is in principle mainly accounted for by the RDM. For instance, when a statistical downscaling model (SDM) is used as a RDM, it classically includes a stochastic process that produces several possible meteorological scenarios for a given large scale pattern or a given sequence of large scale patterns (Buishand and Brandsma 2001; Mezghani and Hingray 2009). The local scale internal variability, sometimes referred to as the RDMs internal variability, corresponds to the dispersion between these scenarios.

b. Partitioning and estimating uncertainties

Estimating the total uncertainty of projections from a multimember multimodel ensemble (MM2E) of climate experiments and estimating the contribution of each particular uncertainty source requires an appropriate statistical framework. Various methods have been proposed for this in the recent years. Most are based on empirical statistical analysis or more formal Analysis of Variance of projections obtained for a particular MM2E opportunity (Hawkins and Sutton 2009; Hingray et al. 2007; Räisänen 2001; Yip et al. 2011). In all cases, a major difficulty lies in the limited number of members that have until now been classically available for most modeling chains and in particular for GCMs. When a large number of members is available for a given chain (e.g. GCM), the climate response of the chain for the considered projection lead time can be estimated from to the multimember mean of the projections. Similarly, the internal variability of the chain can be estimated from the inter-member variance of the projections (e.g. Deser et al. 2012). When only a few members are available, the climate response of the chain and the effect of its internal variability are actually difficult if not impossible to separate, especially when the internal variability is non-negligible compared to the chain's climate response. In such a case, signif-

icantly biased estimates are likely to be obtained for the different uncertainty components, total uncertainty and significance of projected changes (e.g. Deser et al. 2012; Sansom et al. 2013; Hingray and Saïd 2014).

Another consequence of a small number of members is that these estimates may also strongly vary from one projection lead time to another, depending on the projection values simulated for each lead time. Non-negligible temporal variations in GCM internal variability were for instance obtained over the next century by Yip et al. (2011) from regional warming projections based on 21 future climate experiments with 2 runs each. Such temporal variations are expected to be much greater for other variables presenting a non-negligible internal variability such as precipitation (Hingray and Saïd 2014). There is however obviously no reason to expect abrupt changes in internal variability in a transient climate.

In recent years, long time series have become available for the large majority of GCM experiments and in turn for a large number of GCM/RDM chains. GCM experiments from the ENSEMBLE-Stream2 European research project cover for instance 140 years of the pre-industrial period (1860–2000) and the 100 years of the next century (Johns et al. 2011). This provides a major opportunity to significantly improve the estimation of internal variability and model uncertainty components, total uncertainty and in turn the significance of projected changes.

The aim of this paper is to present a theoretical framework for the partitioning of the total uncertainty of climate projections into the four following components: GCM uncertainty, SDM uncertainty and large scale and local scale internal variability of the GCM/SDM modeling chains. The method in particular makes use of all data and simulation experiments available for the target region, whatever the number of runs available for instance for each GCM. For illustration, the method is applied to hydrometeorological projections based on an MM2E of experiments developed for a meso-scale alpine catchment as part of the RIWER2030 research project (Hingray et al. 2013).

Section 2 describes the regional climate experiments and explains some of the practical

choices made for processing the data. Section 3 introduces the two-part statistical framework proposed for the analysis (basically an analysis of variance). Some technical developments are relegated to the Appendix. Results are presented in Section 4. Section 5 discusses the results and presents our conclusions. The data and our code for the calculations in this application is available in the supplementary online materials, written in the Matlab computing environment (`http://www.lthe.fr/RIWER2030/`).

2. Data

We illustrate the method by applying it to projections of the change in 20-year means of precipitation and temperature obtained from the MM2E of opportunity derived within the RIWER2030 research project (Hingray et al. 2013; Lafaysse et al. 2014). The MM2E is extracted from an original ensemble of daily time series simulated for the years 1860–2099.

Projections were obtained for the Upper Durance River catchment, a 3850 km^2 wide drainage basin located in the southern French Alps. They come from $N_s = 6$ multivariate SDMs forced by the outputs of $N_e = 11$ climate simulations. The simulations correspond to $N_g = 5$ GCMs, an ensemble of 3 runs being available for 3 GCMs. They were obtained from the STREAM2 experiment conducted within the ENSEMBLE European project under 20CM3 historical forcing (with constant solar and volcanic forcing) for the 1860–2000 historical period and SRES-A1B emission scenario for the 2000–2100 future period (Johns et al. 2011). The GCMs, SDMs and corresponding references are listed in Table 1 and Table 2. The particular GCMs used for the analysis were chosen on the basis of data availability.

For each GCM/SDM chain, an ensemble of $N_k = 100$ scenarios is available, resulting from the stochastic generation process associated with each SDM (Lafaysse et al. 2014). This gives a total of $N_s \times N_e \times N_k = 6600$ times series of multivariate meteorological scenarios for 1860–2099. Figure 1 shows time series of 20-year mean temperatures and precipitation for selected GCM/SDM configurations and members.

[TABLE 1 HERE]

[TABLE 2 HERE]

[FIGURE 1 HERE]

As proposed by Hingray et al. (2007), the uncertainty analysis is carried out on future to reference control period changes X, expressed in terms of absolute changes for temperatures and in terms of relative changes for precipitation:

$$X = MT_t - MT_c, \quad (1)$$

$$X = MP_t/MP_c - 1, \quad (2)$$

where MT and MP are the 20-yr interannual mean values Y of the raw meteorological projections, where t refers to a given future or past 20-yr period centred on year t and where c refers to the 20-yr reference control period centred on year c. In the present work, the reference control period is the 1980–1999 period. The choice to work on change variables rather than on raw projections follows the choice made for most impact studies. More confidence is actually usually given to changes variables assumed to allow for partially coping with biases in simulation models (e.g. GCMs, SDMs).

3. Theoretical framework

a. The quasi-ergodic assumption

In recent years, long time series of control simulations plus transient projections have become available for the large majority of GCM experiments. Time variations due to a GCM over the whole simulation period are a combination of the climate response of the GCM and its internal variability. Similarly, time variations of a regional variable obtained over the period for a given GCM/SDM chain are the combination of the climate response of the chain and of its total internal variability.

Let us consider that the studied variable is the interannual mean over a given n-year sub-period of a given regional climate variable. Let us next consider one simulation member obtained with a given GCM/SDM chain for a stationary climate (e.g. a climate in the absence of any anthropogenic forcing or any external radiative forcing). Provided that the simulation period is long enough, the time average of this simulation for the studied variable corresponds to the climate response of the GCM/SDM chain for this variable in this climate context. It would also correspond to its ensemble average for any n-year sub-period of the total simulation period, where the ensemble is constituted by all members that could be achieved with this GCM/SDM chain for this sub-period. For any member, the statistical distribution of the climate variable over time would similarly define the internal variability of the chain for this variable. It would also correspond, for any sub-period of the total simulation period, to the distribution of the variable that would be achieved from an infinite number of simulation members.

Such dynamical system, for which the behavior averaged over time is the same as averaged over the space of all the system's states, is said to be ergodic. For an ergodic system, the time average of one sequence of events is the same as the ensemble average at a given time. The same applies also for higher order statistical moments such as variance.

If long time series were available for a stationary climate context for all GCM/SDM chains, the climate response and the internal variability of each chain for a given time would be therefore easy to estimate assuming that each chain behaves as an ergodic random process. Because solar and volcanic forcings were held constant in ENSEMBLES-STREAM2 experiments, this would here appear to be a reasonable assumption for the 1860–2000 extended control period (or at least for a relevant part of it, see section 3.b) for which GCM/SDM control experiments were produced. When 20-yr means are considered as the studied change variable and when stationarity can be assumed for the whole extended control period, the GCM/SDM climate response and its internal variability can be estimated from a sample of size 7 (140 years/20 years).

A sample with more data can be obtained assuming that the process is "quasi-ergodic" for any transient period over which solar and volcanic forcings are held constant. It seems reasonable to consider that if the climate response of each GCM/SDM chain varies over the transient period, this variation should be gradual and smooth, the higher frequency variations of the time series being due to internal variability alone (Figure 2). Another reasonable assumption is that the internal variability remains constant over the period or that it varies gradually as a linear function of the climate response of the GCM/SDM chain. The first assumption was used for instance by Hawkins and Sutton (2011) and Räisänen (2001) and in the present study for changes in surface temperature. The second assumption is used in the present study for changes in precipitation (see Appendix B).

Under this quasi-ergodic assumption, extracting of the noise-free-signal (NFS) from the time series makes it thus possible to estimate the climate response of the GCM/SDM chain and its possible change with time (the climate change response of the chain), as well as the noise around this response and in turn the GCM/SDM internal variability. In the present case, as simulations are provided for the 1860–2100 period, the GCM/SDM internal variability can be estimated from a sample of size 12 (240 years/20 years).

[FIGURE 2 HERE]

Finally, note also that when N different times series members are available for a given GCM/SDM chain, the quasi-ergodic assumption allows the estimation of both the GCM/SDM climate response and its internal variability from the "super-sample" constituted by all sub-period data of all times series members (Figure 2, modeling chain B). In the present case, and even if only one stochastic realization is considered, the size of this "super-sample" is for instance $12 \times 3 = 36$ for GCM/SDM chains with 3 available GCM runs. As shown by Hingray and Saïd (2014), the estimation of both the climate response and the internal variability of the chain is in this case expected to be much more robust than when estimated from the small number of members available for the considered lead time, even if internal variability is significant when compared to inter-chain dispersion.

These considerations are the base of the quasi-ergodic ANOVA framework used for partitioning the uncertainty sources for the present dataset. Let us note $X(g, s, r, k, t)$ the simulations outputs of the studied change variable for a given run r of each GCM g, each stochastic generation k of SDM s, each year t. These outputs can be written as:

$$X(g, s, r, k, t) = \mathrm{NFS}(g, s, t) + \eta(g, s, r, k, t), \tag{3}$$

where $\mathrm{NFS}(g, s, t)$ is the NFS of the change variable for GCM/SDM chain g–s and where $\eta(g, s, r, k, t)$ are the residuals of the k^{th} stochastic generation of SDM s for the r^{th} run of the GCM/SDM chain g–s.

The total uncertainty of the change variable, $\mathbb{V}\mathrm{ar}\{X(g, s, r, k, t)\}$, corresponds to the sum of the variances of both terms of the right hand side of Equation (3). They respectively correspond to the model uncertainty and to the internal variability of X for the modeling chains. The following framework, summarized in Figure 3, details how to estimate the different components of both terms. The main steps are are:

- The noise-free-signal $\mathrm{NFS}(g, s, t)$ in Equation (3) is first estimated for each GCM/SDM chain g–s (top of Figure 3). The estimation is done for absolute changes in temperatures or relative changes in precipitation (see Section 3.b).

- In a second step, we focus on model uncertainty, associated with the noise-free-signal (left side of Figure 3). The estimated noise-free-signal $\widehat{\mathrm{NFS}}(g, s, t)$ obtained from step 1 is modeled using a two-way ANOVA framework allowing to separate GCM uncertainty (G), SDM uncertainty (S) and residual uncertainty (R) (see Section 3.c).

- In a subsequent step, we focus on model internal variability, corresponding to the noise term $\eta(g, s, r, k, t)$ in Equation (3) (right side of Figure 3). The estimated counterpart of internal variability, say $\widehat{\eta}(g, s, r, k, t)$, comes from the estimation of $\mathrm{NFS}(g, s, t)$ and is modeled as the sum of two terms: $\delta(g, s, r, t)$ related to the large scale component of internal variability (LSIV) and $\varepsilon(g, s, r, k, t)$ related to the small scale component

of internal variability (SSIV). The estimation of the associated variances is detailed in Section 3.d.

- The total uncertainty is finally derived from these different model uncertainty and internal variability components (see Section 3.e).

[FIGURE 3 HERE]

b. Noise-free-signal

The NFS for the change variable X (e.g. temperature changes) is estimated for each modeling chain from a trend model fitted to the raw projections Y (e.g. temperature) of the chain as:

$$\widehat{\mathrm{NFS}}(g,s,t) = y(g,s,t) - y(g,s,c) \tag{4}$$

for absolute changes in temperature and:

$$\widehat{\mathrm{NFS}}(g,s,t) = y(g,s,t)/y(g,s,c) - 1 \tag{5}$$

for relative changes in precipitation, where $y(g,s,t)$ and $y(g,s,c)$ are respectively the trend estimates of the raw projections Y for the future time period t and the reference control period c.

The specific constraints and formulations retained for the estimation of the trends are presented in Appendix A. The trend models are composed from a constant value $y(g,s,C)$ over the extended control period [1860,year0] and from a trend function for the transient period [year0+1, 2099] where the pivot year0 is the year where the climate is estimated to become non-stationary (Figure 2).

A linear trend was chosen for precipitation. No satisfying adjustment was obtained with higher order monomials or polynomials. Higher order polynomials were observed to over-fit time fluctuations of raw data and often lead to no coherent NFSs between runs available for a same GCM/SDM configuration. As reported by Räisänen (2001), a linear trend is

expected to overestimate internal variability when the magnitude of this latter is small when compared to non-linearities in the climate change response. As highlighted in the "Results" section, this is obviously not the case for precipitation as internal variability is very large for this variable. The confidence in a linear trend is also supported by the fact that internal variability estimated for the extended control period is similar to that estimated for the detrended future period (not shown).

A linear trend is conversely not suited for changes in temperature which clearly highlight a progressive increase from the 1950's. A 3^{rd} order polynomial was found to be optimal, allowing additionally for adjusting for smaller increase rates at the end of the XXI^{th} century. The pivot year, estimated with a sensitivity analysis, is 1950 for temperature, 1980 for precipitation.

c. Partitioning model uncertainty

It is assumed that the noise-free change response $\widehat{\text{NFS}}(g, s, t)$ can be partitioned as:

$$\widehat{\text{NFS}}(g, s, t) = \mu(t) + \alpha(g, t) + \beta(s, t) + \gamma(g, s, t) \tag{6}$$

where $\mu(t)$ is the overall climate response representing the grand-ensemble mean of all experiments at projection lead time t, $\alpha(g, t)$ and $\beta(s, t)$ are respectively the mean deviations of GCM g and SDM s from the grand-ensemble mean $\mu(t)$, and where $\gamma(g, s, t)$ is the residual.

The parameters $\alpha(g, t)$ and $\beta(s, t)$ are called the main effects of the models, and $\mu(t)$ is the multichain mean. They are estimated using a classical two-way ANOVA framework without interactions (Berrington De González and Cox 2007; Searle 1971). Let us define the

following means where the symbol $\bullet$ indicates averaging over the particular index:

$$\widehat{\mathrm{NFS}}(g,\bullet,t) = \frac{1}{N_s}\sum_{s=1}^{N_s}\widehat{\mathrm{NFS}}(g,s,t), \tag{7}$$

$$\widehat{\mathrm{NFS}}(\bullet,s,t) = \frac{1}{N_g}\sum_{g=1}^{N_g}\widehat{\mathrm{NFS}}(g,s,t), \tag{8}$$

$$\widehat{\mathrm{NFS}}(\bullet,\bullet,t) = \frac{1}{N_g N_s}\sum_{g=1}^{N_g}\sum_{s=1}^{N_s}\widehat{\mathrm{NFS}}(g,s,t). \tag{9}$$

With a least squares estimation under constraints $\sum_{g=1}^{N_g}\widehat{\alpha}(g,t)=0$ and $\sum_{s=1}^{N_s}\widehat{\beta}(s,t)=0$, the parameter estimators in model (6) are given by:

$$\widehat{\alpha}(g,t) = \widehat{\mathrm{NFS}}(g,\bullet,t) - \widehat{\mathrm{NFS}}(\bullet,\bullet,t), \tag{10}$$

$$\widehat{\beta}(s,t) = \widehat{\mathrm{NFS}}(\bullet,s,t) - \widehat{\mathrm{NFS}}(\bullet,\bullet,t), \tag{11}$$

$$\widehat{\mu}(t) = \widehat{\mathrm{NFS}}(\bullet,\bullet,t), \tag{12}$$

residuals are given by:

$$\widehat{\gamma}(g,s,t) = \widehat{\mathrm{NFS}}(g,s,t) - \widehat{\mu}(t) - \widehat{\alpha}(g,t) - \widehat{\beta}(s,t). \tag{13}$$

and the total variability (or, in other words, total model uncertainty) of $\widehat{\mathrm{NFS}}(g,s,t)$ is given by:

$$\mathbb{V}\mathrm{ar}\left[\widehat{\mathrm{NFS}}(g,s,t)\right] = \mathbb{V}\mathrm{ar}\left[\widehat{\alpha}(g,t)\right] + \mathbb{V}\mathrm{ar}\left[\widehat{\beta}(s,t)\right] + \mathbb{V}\mathrm{ar}\left[\widehat{\gamma}(g,s,t)\right]. \tag{14}$$

The three model uncertainty components will hereafter be referred to as GCM uncertainty (G), SDM uncertainty (S) and residual uncertainty (R). The theoretical variances may be estimated by their empirical counterparts, namely

$$G(t) = \frac{1}{N_g-1}\sum_{g=1}^{N_g}\left\{\widehat{\alpha}(g,t)\right\}^2, \tag{15}$$

$$D(t) = \frac{1}{N_s-1}\sum_{s=1}^{N_s}\left\{\widehat{\beta}(s,t)\right\}^2, \tag{16}$$

$$R(t) = \frac{1}{(N_g-1)(N_s-1)}\sum_{g=1}^{N_g}\sum_{s=1}^{N_s}\left\{\widehat{\gamma}(g,s,t)\right\}^2. \tag{17}$$

The residuals $\gamma(g, s, t)$ describe any GCM-dependent deviations of SDMs. These deviations may be purely random. In this case, the deviations $\gamma(g, s, t)$ obtained for a given GCM/SDM chain depend on the members (run/generations) used to identify the NFS of the chains. GCM-dependent deviations of SDMs may however be purely due to GCM/SDM interactions (e.g. Yip et al. 2011). In this case, the deviation term of Equation (6) is independent of the members used to identify the NFS of the chains. For a given multimodel ensemble of NFSs, deviations come from the combined effect of randomness and interaction. Estimating the relative importance of each of these components can be obtained using a classical 2-way ANOVA with interaction. This would however require the availability of multiple members of NFSs for each GCM/SDM chain. In the present case, 3 runs are available for 3 GCMs. Three members of the NFS can therefore be obtained for each of the corresponding GCM/SDM chains. The interaction term estimated from this subset of chains was found to account for 20 to 30% of the variance of the deviation term. The contribution of interactions cannot be estimated for the full RIWER2030 MM2E given that only one run is available for 2 GCMs, however it is expected to be of the same order as that obtained from the subset. The uncertainty component associated with the deviation term in the ANOVA will therefore subsequently be referred to as the residual/model interaction uncertainty component (R/MI).

d. Partitioning internal variability

The noise term $\eta(g, s, r, k, t)$ of Equation (3) accounts for the internal variability (IV) of the change variable X associated to the selected run r of GCM/SDM chain *g–c*. It can be also partitioned in the following components:

$$\eta(g, s, r, k, t) = \delta(g, s, r, t) + \varepsilon(g, s, r, k, t), \tag{18}$$

where $\delta(g, s, r, t)$ is the deviation of the r^{th} run of GCM g from the climate response of the GCM/SDM chain *g–s* and $\varepsilon(g, s, r, k, t)$ is the deviation of the k^{th} stochastic generation of

SDM s from the r^{th} run of the chain.

The parameter $\delta(g, s, r, t)$ is related to the large scale component of internal variability (further noted as LSIV) resulting from the run r of GCM g used to force SDM s and $\varepsilon(g, s, r, k, t)$ is related to the small scale component of internal variability (further noted as SSIV) resulting from the stochastic generation k of SDM s when forced by run r of GCM g. It is assumed that the $\delta(g, s, r, t)$'s and $\varepsilon(g, s, r, k, t)$'s have zero mean.

Expression of LSIV and SSIV components are derived in the following for temperature. Similar expressions were derived for precipitation where relative changes are considered instead of absolute ones. They are presented in Appendix B.

(i) Small Scale internal variability.

For a given GCM/SDM chain g–s, the SSIV of the change variable X can be derived from the SSIV of the raw data Y. When only one run is available for the chain, this latter can be estimated for each projection lead time t as the inter-member variance of the 100 stochastic generations for Y. It was found to be roughly constant over the whole 1860–2099 simulation period validating the quasi-ergodic assumption for this component (see for instance the bottom graphs of Figure 1 where the inter-percentile distance is roughly independent on time). We could also check that for a given chain there is no correlation between the values of Y in the control and in a future climate. Then it follows that the SSIV of the change variable $X = Y_t - Y_c$, reads for any year t:

$$\mathbb{V}\text{ar}_k \{X(g, s, r, k, t)\} = \mathbb{V}\text{ar}_k \{Y(g, s, r, k, t)\} + \mathbb{V}\text{ar}_k \{Y(g, s, r, k, c)\} \tag{19}$$

where $\mathbb{V}\text{ar}_k \{Y(g, s, r, k, c)\}$ and $\mathbb{V}\text{ar}_k \{Y(g, s, r, k, t)\}$ denote the inter-generation variance of raw data Y simulated by GCM/SDM chain g–s for the reference control period c and any other period t. To estimate $\mathbb{V}\text{ar}_k \{Y(g, s, r, k, t)\}$, define the multi-generation mean at time

t obtained from the outputs of the r^{th} run of GCM/SDM combination g–s:

$$Y(g,s,r,\bullet,t) = \frac{1}{N_k}\sum_{k=1}^{N_k} Y(g,s,r,k,t). \tag{20}$$

Then $\mathbb{V}\text{ar}_k\left\{Y(g,s,r,k,t)\right\}$ is estimated by their empirical counterpart, that is

$$\widehat{\mathbb{V}\text{ar}}_k\left\{Y(g,s,r,k,t)\right\} = \frac{1}{N_k-1}\sum_{k=1}^{N_k}\left\{Y(g,s,r,k,t) - Y(g,s,r,\bullet,t)\right\}^2. \tag{21}$$

It also follows that the SSIV of the change variable X for the GCM/SDM chain g–s can be estimated as the multi-period mean of two times the variance for the raw data Y. The quasi-ergodic assumption also implies that when multiple GCM runs are available for the chain, the SSIV of the change variable X for this chain is the multi-run mean of the SSIV estimated for each run of the chain.

Finally, the multimodel mean of this estimate is taken to be the SSIV component for the change variable X for the ensemble of climate experiments. It finally reads:

$$SSIV = \frac{2}{N_g N_s}\sum_{g=1}^{N_g}\sum_{s=1}^{N_s}\left[\frac{1}{TN_{g,r}}\sum_{r=1}^{N_{g,r}}\sum_{t=1}^{N_t}\mathbb{V}\text{ar}_k\left\{Y(g,s,r,k,t)\right\}\right], \tag{22}$$

where $N_{g,r}$ is the number of runs available for GCM g (either 1 or 3 in the present case, see Table 1).

Note that for precipitation, for which relative changes are considered instead of absolute changes, the $SSIV$ expression is a function of projection lead time t (see Appendix B). For the sake of notation simplicity, this component will be referred to in the following as $SSIV(t)$ for both temperature and precipitation variables.

(ii) Large scale internal variability.

The LSIV is the variance component of the noise that remains after having removed the noise due to the stochastic downscaling generation process of the SDMs. It is here estimated for each modeling chain from the time series obtained for each run of the GCM from the multi-generation mean of the raw data. For the r^{th} run of the GCM/SDM chain

g–s, $Y(g,s,r,\bullet,t)$ is actually the sum of the noise-free change response for this chain and of the noise component of the run due to LSIV. Assuming that the system is quasi-ergodic over the whole simulation period, the LSIV of the GCM/SDM chain for the raw data Y can next be estimated as the variance over time of the residuals of $Y(g,s,r,\bullet,t)$ from the NFS for Y. Note that for temperature changes, this assumes that LSIV is constant over the entire simulation period. Results obtained for the control and the future period independently show this is a reasonable assumption. When multiple runs are available for a chain, the LSIV for Y is estimated from the variance over time of residuals from all runs. Recall that the NFS correspond in this case to the common trend of these multiple runs.

Following Hingray et al. (2007), we assume that for a given GCM/SDM chain g–s there is no correlation between the values of raw data $Y(g,s,r,\bullet,t)$ obtained in the control and in a future climate. The LSIV of the change variable X is thus also two times that of variable Y for this chain.

The multimodel mean of the large scale internal variability for the change variable X is finally taken to be the large scale internal variability component. It finally reads:

$$LSIV = \frac{2}{N_g N_s} \sum_{g=1}^{N_g} \sum_{s=1}^{N_s} \mathbb{V}\mathrm{ar}_{T,N_{g,r}} \left\{ Y(g,s,r,\bullet,t) - y(g,s,t) \right\}. \tag{23}$$

For precipitation, for which relative changes are considered instead of absolute changes, the $LSIV$ expression is also a function of projection lead time t (see Appendix B). For both variables, this component will be also referred to as $LSIV(t)$ in the following.

Note finally that by construction the $\delta(g,s,r,t)$'s and the $\varepsilon(g,s,r,k,t)$'s are expected to be uncorrelated and the total uncertainty variance due to internal variability is simply the sum of LSIV and SSIV variance components. The absence of correlation can be checked empirically, for each stochastic generation obtained with SDM s when forced by run r of GCM g, with the correlation coefficient obtained between time series $\delta(g,s,r,t)$ and $\varepsilon(g,s,r,k,t)$ available over the simulation period. For our dataset, the empirical correlation coefficient is in the $]-0.4, 0.4[$ interval for more than 90% of the cases.

e. Total uncertainty

Under the assumption that the different variables on the right hand side of equations (3), (6) and (18) are uncorrelated, the total variance for the change variable X for a given future period t is given by the sum of the different variance components introduced previously. It reads:

$$T(t) = G(t) + S(t) + R(t) + LSIV(t) + SSIV(t), \tag{24}$$

where $G(t)$, $S(t)$, $R(t)$, $LSIV(t)$ and $SSIV(t)$ are given by Equations (15), (16), (17), (23) and (22), respectively.

The grand-ensemble mean $\mu(t)$ of all experiments at t and the total variance characterize the magnitude and total uncertainty of the change variable X. Note that the total uncertainty corresponding to this total variance is the potential uncertainty that would be obtained from a large number of members for the considered set of GCM/SDM (i.e. large number of runs for each GCM and large number of generations for each GCM/SDM chain). It may be significantly different from the total variance that would be obtained from the MM2E sample of opportunity (Hingray and Saïd 2014).

We further define the fraction of total variance explained by uncertainty source $U(t)$ as

$$U(t)/T(t), \tag{25}$$

where $U(t)$ is either $G(t)$, $S(t)$, $R(t)$, $LSIV(t)$ or $SSIV(t)$.

4. Results

a. Total uncertainty and sources of contributions

Figure 4 presents, for 20-yr temperature and precipitation, the grand ensemble mean climate change response $\mu(t)$ and the limits of the interval $\mu(t) \pm 1.645\sqrt{T(t)}$ where $T(t)$ is the total uncertainty variance.

[FIGURE 4 HERE]

For both variables, total uncertainty increases with lead time, especially for projections of temperatures changes for which it is multiplied by a ratio of 3.5 from the beginning to the end of the century. It is conversely multiplied by a factor of only 1.4 for changes in precipitation. The main contribution for this increase is that of model uncertainty. Remember that the amplitudes of both IV components are assumed to remain constant over the whole period for temperature changes. For precipitations relative changes, they are conversely assumed to be a linear function of the grand-ensemble mean of the NFSs (see Appendix B). In the present case, as the grand-ensemble mean is expected to slightly decrease with time, both components of IV highlight therefore a slight decrease over the period.

The relative contributions of model uncertainty and internal variability components to the total uncertainty are given in Figure 5. They vary significantly with projection lead time and the variable considered. For the first 3 decades, the contribution of internal variability is highly predominant for both variables. For the first 2 decades, in particular, the combined contribution of LSIV and SSIV represents nearly all the total uncertainty (resp. more than 80% and 95% for temperature and precipitation). The contribution of internal variability then decreases with lead time as the total uncertainty increases. For temperature, it rapidly drops to less than 10% by the end of the century, then becoming negligible compared to model uncertainty. For precipitation, it is still around 60% of total uncertainty in 2090 and therefore remains the greatest contribution to total uncertainty over the whole simulation period.

For both variables, the LSIV component is greater than the SSIV component. For temperature, the contribution of SSIV is negligible whatever the lead time. It is less than 5% even at the beginning of the century. This clearly reflects the fact that the small scale variability in temperature is nearly fully explained by the large scale variability. For precipitation, the contribution of SSIV to cumulated internal variability is around 20% (as a consequence of the quasi-ergodic assumption, this contribution is roughly constant over the

simulation period). In 2090, LSIV is 50% of total uncertainty and thus remains the dominant uncertainty source. However, at this time, SSIV still represents more than 10% of total uncertainty (vs. 20% for the first decades). For precipitation, SSIV is therefore far from negligible, highlighting, as already discussed, that large scale variability cannot explain all of the variability observed for this variable (Lafaysse et al. 2014).

For both variables, the contribution of model uncertainty increases over the whole simulation period. The contribution of the error/interaction term is non-negligible, however it is far less than the contributions of GCM and SDM uncertainty. For temperature, GCM uncertainty is the main contribution for the end of the century. The SDM contribution is however non-negligible. It reaches up to 20% for the end of the century and even tends to be higher than the GCM contribution for the middle of the century.

For precipitation, the GCM contribution is slightly smaller than the SDM contribution for the whole period. The change in the main effect of each SDM with time was estimated for the next century. Results show that the inter-SDM dispersion from a given SDM with different predictors is as large as that from different SDMs (see Figure 8b in section 5b). Note also that the contributions of GCM and SDM uncertainty to total uncertainty are much smaller than the contribution of LSIV. They are also smaller than the contribution of SSIV except for the end of the century where the contribution of SDM uncertainty is slightly larger.

[FIGURE 5 HERE]

b. Significance of changes

Comparing for each projection lead time the grand-ensemble mean response to the total uncertainty provides a rough idea of the significance of the estimated changes with respect to any reference level of change. As mentioned previously, the total colored area in graphs of Figure 4 corresponds to the interval $\mu(t) \pm 1.645\sqrt{T(t)}$ where $T(t)$ is total uncertainty

variance. Assuming for convenience that all possible future climate projections are normally distributed, this interval corresponds to the confidence interval of possible future changes at the 90% confidence level. It allows estimating the significance at a 90% confidence level that all possible future realizations of the climate exceed a given reference level of change. In the following, we will discuss the significance for a non-zero climate change realization. A significant non-zero climate change realization is expected when the zero change value is outside the confidence interval or when the response-to-uncertainty ratio, expressed as $\mu(t)/\left\{1.645\sqrt{T(t)}\right\}$, is outside the $[-1, 1]$ interval. The response-to-uncertainty ratio (R/U) is presented for both variables in Figure 6.

[FIGURE 6 HERE]

For temperature, a significant non-zero change is predicted from the beginning of the century. The time of emergence of a significant warming, defined here as the first future lead time for which R/U is outside the $[-1, 1]$ interval, is found within the first decade. For precipitation, the non-zero change is not significant. The absolute value of R/U remains much smaller than 1 throughout the entire period, indicating that the sign of change is very uncertain at the 90% confidence level even at the end of the century. Precipitation that may be experienced for a given future period could be therefore higher, but also lower, than those observed for the control period. This is mainly due to the large internal variability value for this variable.

Note that Figure 4 also allows to discuss the significance of a non-zero climate change response. The total area covered by the GCM uncertainty (blue), SDM uncertainty (green) and residual / model interaction uncertainty (cyan) corresponds actually to the interval $\mu(t) \pm 1.645\sqrt{M(t)}$ where $M(t)$ is total model uncertainty expressed as the square root of total model uncertainty variance $G(t) + S(t) + R(t)$. Assuming that the predicted climate change response has a normal distribution, a significant non-zero climate response at a 90% confidence level is obtained when the zero climate change response value is outside this

interval, i.e. outside the model uncertainty area. For precipitation, it is here also interesting to note that, even if simulation chains roughly agree on the direction of the change response (a decrease), a non-zero climate change response is still not significant at a 90% confidence level.

c. The potential to reduce uncertainty

It will never be possible to remove uncertainties related to internal variability because they are intrinsic to the Earth system even if uncertainty in the large scale component can be reduced for coming decades by improving the initialization of climate projections with observations (Smith et al. 2007). On the other hand, GCM and SDM uncertainties might be reduced by a better understanding of climatic and hydrological processes and resulting improvements of numerical models.

Figure 6 also presents the response-to-uncertainty ratio assuming perfect models. This highlights the potential gain that would be achieved when improving both GCM and SDM. For temperature, the R/U increase is important especially for the second half of the century. This traduces again the very small impact of internal variability on the significance of changes. This traduces next the large potential to narrow uncertainty by improving simulation models. The high contribution of SDM uncertainty was for instance surprising as we though initially that a large fraction of uncertainty should be carried by GCMs. It even tends to be higher than that of GCM for the middle of the century. This reflects actually the influence of the different large scale temperature predictors retained for the SDMs (e.g. the extent of the large scale predictor domain in the different versions of dsclim which presents the most different effects). This critical issue was underestimated in the development of the SDMs. SDM development and selection were based on an extensive evaluation of the model ability to reproduce, in perfect prognosis conditions (where SDM are forced by NCEP/NCAR atmospheric reanalysis; Kalnay et al. 1996), a number of key statistical characteristics of the observed regional climate for the past 50 years. As all SDMs were

found to perform much higher for temperature than for the precipitation (Lafaysse 2011; Lafaysse et al. 2014), the main effort in model development was focused on precipitation. A deeper discussion is here obviously out of the scope of this paper but this clearly shows also the interest of the uncertainty analysis that allowed us to identify this likely drawback of some of the SDMs. This clearly shows also that an improved choice of the predictors would allow in this case for reducing the SDM contribution to total uncertainty in future projections. More constraining evaluation tests would have therefore to be carried out in future SDM developments for temperature. A special attention should be especially paid to the time transferability of the SDMs for this variable (e.g. comparison with temperature changes obtained from GCM outputs directly).

For precipitation, the potential to reduce uncertainty is also non-negligible however model uncertainty accounts for a much smaller part of total uncertainty than for temperature. The R/U increases obtained when assuming perfect models are therefore low to very low, especially for the first four decades. This again reflects the dominant role of internal variability. Even if all models would agree on the ensemble mean precipitation change, this mean change would be very small compared to the internal variability (see the potential R/U in Figure 6). This means that adapting management practices to the internal variability of precipitation should be a priority as it will likely contribute to successful adaptation to climate change.

5. Discussion and conclusion

a. Methodology and interpretation of results

We introduce a simple two-part approach for the partitioning of model uncertainty and internal variability components in multimember multimodel ensembles of climate experiments. It is based on the quasi-ergodic assumption for the outputs of climate simulations obtained over long control and transient climate period. Model uncertainty components are estimated from the noise-free responses of the different modeling chains using a classical

two-ways ANOVA framework. Internal variability components, produced by each modeling chain, are estimated from the time series of residuals for the multiple members available for the chain.

Our results for changes in precipitation and temperature are the following:

- SDM uncertainty is of same order to that GCM. For the present dataset, it was even found to be greater for precipitation during the whole simulation period and for temperature during some decades in the middle of the century.

- The contribution of the residual / model interaction term is much smaller than that of SDM and GCM uncertainty, but is not negligible.

- Internal variability is the main component of total uncertainty, especially for the first decades. It rapidly decreases to less than 10% at the end of the century for temperature. It roughly always remains the main uncertainty component for precipitation.

- The small scale component of internal variability is negligible for temperature. It is however of the same order than the large scale component for precipitation, whatever the projection lead time.

- The response-to-uncertainty ratio at the 90% confidence level is very high for temperature, whatever the lead time. Its absolute value is conversely always much smaller than one for precipitation, indicating that even the sign of precipitation change is uncertain. The time of emergence of change due to global warming is found to be as soon as the first decade for regional temperature but is not expected to be within this century for precipitation.

These findings have important implications. As already suggested by numerous works, neglecting model uncertainty associated to downscaling models is expected to lead to erroneous climate change estimates (e.g. Chen et al. 2011). Depending on the studied variable, the same is likely to apply when internal variability is ignored. For precipitation, the major

impact of the large scale component of internal variability is well established (e.g. Hawkins and Sutton 2011). It is however often disregarded. The same applies for the small scale component of internal variability, associated to downscaling models. Its impact on climate projections is however likely to be non-negligible as highlighted in the present study for precipitation projections. Similar conclusions could be expected with dynamical downscaling models. The presence of internal variability in ensembles of nested Regional Climate Model (RCM) simulations is actually widely acknowledged in the community working on dynamical downscaling (Nikiéma and Laprise 2013). Alexandru et al. (2007) found for instance that seasonal statistics of precipitation fields simulated with the Canadian Regional Climate Model are likely to be poorly estimated from a single model simulation and that a robust estimation, depending on season and region, could require a minimum number of 10 members. For climate projections, the internal variability associated to the RCM could be locally larger than the mean climate change response, as shown for instance by Braun et al. (2012). Part of this internal variability is expected to be reduced with a closer forcing of the RCM towards the driving circulation but part of it must also correspond to the freedom the system has of creating different local weather from different weather patterns. If the real contribution of this internal variability component seems to remain unclear at this time, its impact on climate projection is therefore worth more attention.

Multimodel (GCMs and Downscaling Models) experiments became a standard in Climate change impact studies. Another standard should be to rely on multimember experiments. Impact studies based on single members of SDMs or RCMs experiments (or small ensembles) are likely to be not more relevant than those based on single runs of available GCMs (or small ensembles). When they are intended to provide information for climate change adaptation, they may lead to poor decisions. A relevant strategy would be, in the present case, to adapt to internal variability of precipitation.

b. Partitioning uncertainty components with the quasi-ergodic ANOVA framework

A strong advantage of the quasi-ergodic ANOVA (QE-ANOVA) approach presented in the present work is that all periods and experiments available in the ensemble of climate experiments opportunity can be valorized. A critical and classical problem in similar studies is actually that the number of available runs differs from one GCM to the other. As an ANOVA in such an unbalanced case is no easily tractable (Sansom et al. 2013), a unique number of runs is next usually considered for all GCMs (e.g. Hawkins and Sutton 2009; Hingray et al. 2007; Yip et al. 2011). This simplifies the ANOVA framework and allows treating all GCM/SDM chains equally. This leads however to disregard a large number of available runs. Additionally, Deser et al. (2012) and Northrop (2013) pointed out the fact that results can depend strongly on the particular dataset chosen, so relying on a single dataset may be misleading. A major challenge is to have therefore the possibility to use all available experiments for the impact study under consideration (Sansom et al. 2013).

In the present framework, the model uncertainty analysis is carried out on the climate change responses estimated respectively for each GCM/SDM chain. The climate change response is unique for a given chain whatever the number of runs available to estimate it. In the classical ANOVA part of the method, all available runs can therefore be accounted for and all GCM can be treated equally, whatever the number of runs accounted for. The residuals from the climate response estimated for each modeling chain allows next for estimating the different components of internal variability.

In the present work, a large number of SDM generations was made available for each run of each GCM/SDM chain. This allowed for an easy separation of both large scale and small scale components of internal variability. Such a dataset is likely to be not always available. In such a case, the partition between both components would not be possible. The method would however allow estimating the total internal variability resulting from both components using an equation similar to Equation (23).

In all cases, all available data can again be used to estimate the large scale component of

internal variability. For a given modeling chain, this latter can be actually estimated from the residuals obtained for the entire control+transient simulation period from all members available for the chain.

c. Robustness of uncertainty components estimates

The more members we can rely on, the more robust the estimates of uncertainty components are obviously expected to be. The climate response for a given modeling chain can however be already identified from only one member. In this case, when a long control+transient period is available (e.g. 240 years in the present study), the quasi-ergodic assumption makes it already possible to estimate the large scale internal variability component from a quite large sample of residuals (7 data in our case).

In this work, results were presented using all experiments of the RIWER2030 MM2E, and especially the 3 runs available for 3 of the 5 considered GCMs (Table 1). The results are roughly the same when only one run is used for each GCM. This is for instance illustrated in Figure 7 by the roughly unchanged partitioning of total variance obtained when only the first, second or third run is used. The main difference is the slightly smaller contribution of the large scale internal variability component obtained for precipitation. The robustness of the method is also illustrated by the fairly unchanged patterns of main model effects shown in Figure 8 for such configurations. For both variables, the main effect of SDMs is similar whatever the run or multirun experiment considered. The main effect of the GCMs is somewhat more sensitive to the experiment but the general pattern of main effects is the same.

The robustness of uncertainty component estimates obviously depends on the possibility to achieve a robust estimate of each modeling chain climate response. A potentially critical problem is therefore the potential errors associated to the identification of the climate response of each modeling chain. A first difficulty concerns the choice of the type of trend used for the NFS estimation. For convenience, analytical trend functions were chosen in the

present work and a same analytical model (with chain-dependent parameters) was retained for all GCM/SDM chains. For some chains, better adjustments could have been probably obtained with higher order polynomials or with trend functions estimated with non-parametric methods. An overly adaptable trend function is however expected to underestimate the internal variability of the chain and to be too sensible to the sample of data used for its identification (i.e. to the member retained for the estimation).

Even when a simple trend analytical function is retained, the dependence of the trend to the data sample may be high and it could be therefore difficult to identify the true climate change response of the modeling chain. This problem is of course expected as soon as internal variability is high compared to the climate response of the chain. This is likely to explain differences in the results presented here, for precipitation especially (e.g. the rather large dispersion of the main GCM effects between one-run configurations or the smaller contribution of large scale internal variability component to total uncertainty for one-run configurations).

From a 40-member climate experiment, Deser et al. (2012) found that the minimum number of members required for an accurate estimation of the climate change response (linear least-squares trends fit to the period 2005–2060) can be quite large depending on the considered variable. Due to the relative amplitudes of the climate change response and natural variability, they highlighted that 1 member is needed to detect a significant (at the 95% confidence level) warming in the 2050s decade compared to the 2010s at nearly all locations, compared to approximately $3-6$ (>15) ensemble members for tropical and high latitude (middle latitude) precipitation.

For a number of impact studies, the relevance and significance of the climate responses identified from available MM2E of climate projections are therefore expected to be rather problematic. For this critical identification issue, robust methods and clear consensual guidelines would be highly welcomed from the climate research community. Further methodological developments are also required to complement the QE-ANOVA framework. They should

allow for a thorough analysis and quantification of the different errors sources liable to impact uncertainty estimates.

[FIGURE 7 HERE]

[FIGURE 8 HERE]

d. Extensions

The QE-ANOVA framework can be easily applied for a large number of different data sets. The main requirement is that long term simulations are available to allow for a robust estimation of the climate response of each modeling chain and of its internal variability. In all cases, the NFS would allow for the estimation of all model uncertainty components and the residuals from this NFS would allow for the estimation of the different internal variability components.

A straightforward application is for instance to quantify the uncertainty in hydrometeorological projections obtained by simulation with a given hydrological model from a dataset similar to the one used in the present work. An application is presented by Lafaysse et al. (2014) for hydrological projections obtained for the Upper River Durance basin from the present MM2E of climate projections. The same framework could be also easily extended to hydrological projections from multiple hydrological models (HM). The NFSs resulting from all GCM/SDM/HM chains would make it possible to partition the main model effects. This could be done using a classical three-way ANOVA. As the hydrological behavior of catchments is mainly deterministic, as reflected by the large majority of hydrological models used for climate change impact studies, the HM is not expected to introduce any additional component to the internal variability of the chain. Internal variability could be partitioned, in the same way as presented here, into its large scale and small scale components related mainly to GCMs and SDMs respectively.

As long as climate experiments would be available for a number of different emission

scenarios, another straightforward extension of this framework would be also to additionally estimate scenario uncertainty. Each emission scenario would actually influence the NFSs of the modeling chains. Scenario uncertainty would be thus considered as an additional "model uncertainty" factor to be estimated from the classical ANOVA part of the framework.

Note finally also that the QE-ANOVA framework presented here does not necessarily require all the model components of the simulation chain. For instance, it does not require that GCM experiments are downscaled with SDM or RCMs. It can be applied directly on GCM outputs if only experiments from different GCM are available. The partition would here only concern the uncertainty due to the GCM and that due to large scale internal variability. As already mentioned, the framework does also not require all the members used in the present work (e.g. different runs of the same GCM, a large number of stochastic generations from the SDM). It could be also easily applied on the outputs of GCM/SDM experiments with a single SDM generation for each GCM/SDM chain. In such a case, the different components of internal variability could however not be partitioned and the estimate of internal variability would contain both components.

Acknowledgments.

This work was initiated during the research project RIWER2030 "Regional ClImate, Water, Energy Resources and uncertainties from 1960 to 2030" (`http://www.lthe.fr/RIWER2030/`) funded by Electricité de France (EDF), the Centre National de la Recherche Scientifique (CNRS) and the French National Research Agency (ANR). It was finalized within the R2D2-2050 research project (`https://r2d2-2050.cemagref.fr/`) founded by the Ministère de l'Ecologie, du Développement Durable et de l'Energie. The author wishes to thanks Matthieu Lafaysse (Météo-France, Grenoble), Joël Gaihlard (EDF, Grenoble), Abdelkader Mezghani (MetNo, Oslo) for providing climate projections. We thank also Jerome Buzzi (Orsay, Paris), Douglas David Baptista De Souza (LTHE, Grenoble) and Jean-Philippe Vidal (Irstea, Lyon) for interesting discussions on this work as well as Michel Slivitzky for

his enthusiastic appreciation on a former manuscript version. We finally thank the three anonymous reviewers for constructive suggestions.

APPENDIX A

Trend models

Different trend models were used in previous works to fit the time evolution of future projections from climate experiments. To estimate internal variability in 17 CMIP2 experiments, Räisänen (2001) assumed that the climate change response for the 2000–2080 period was a linear function of time, like the greenhouse run radiative forcing that was proportional to the logarithm of atmospheric CO_2. Hawkins and Sutton (2009) assumed the noise-free-signal to be a fourth order polynomial of time over the 1950–2099 period.

In the present work, an analytical trend model is also used for convenience. It applies on the raw temperature or precipitation projections Y. For each GCM/SDM combination, the trend model is estimated for the whole 1860–2099 period with following constraints:

- It is composed of a constant value $y(g, s, C)$ over the extended control period [1860–year0] and of a linear or polynomial trend for the transient period [year0+1, 2099]. $y(g, s, C)$ is the mean value of the variable estimated for the extended control period and constitutes the starting value of the adjustment for the transient period.

- The pivot year and the polynomial degree are assumed to be the same for all GCM/SDM chains. They vary from one variable to the other. The coefficients of the linear or polynomial trend are the same for all members of a given GCM/SDM chain. They vary from one chain to the other.

- The fit is made using ordinary least squares. In case of a polynomial trend, the fit insure that the derivative of the polynomial is zero for the pivot year (see equations below).

A basic assumption there is thus that the climate response of a given GCM/SDM configuration varies progressively over the 1860–2100 period with a continuous derivative for the pivot year especially. Fitting a trend without these constraints potentially leads to a noise-free climate change that over-fits the raw series around the 2000 year especially, leading in turn to overestimate the model uncertainty components and to underestimate the internal variability components.

Note that $y(g, s, C)$ is not necessary equal to the raw projection Y obtained for the reference control period c from which future changes are estimated. On the one hand, the pivot year does actually not necessarily correspond to the middle year of the reference control period c. In turn the trend estimate $y(g, s, c)$ of the raw projection Y for the reference control period c is a priori different from the constant value $y(g, s, C)$ estimated over the extended control period C (see Figure 2). On the other hand, the raw value $Y(c)$ is also expected to be significantly different from the trend estimate $y(g, s, c)$ as soon as internal variability is non-negligible (e.g. for precipitation in the present case). The trend function identified using $y(g, s, C)$ as starting value for the transient climate period is therefore expected to be much more robust and thus relevant that the one which could have been obtained using the sample mean of the reference control period c instead.

The analytical expressions of the coefficients of trend function have been derived to meet the above mentioned constraints. Different trend models were tested for the raw variables considered in this work: a monomial trend of degree n ($1 \leq n \leq 3$) and a polynomial trend of degree 3. In the case of a monomial trend, the trend has the following expression:

$$x(t) = \begin{cases} x_0 & \text{for } t < t_0, \\ x_0 + a\tau^n & \text{for } t \geq t_0, \end{cases} \tag{A1}$$

where x_0 is the constant value estimated for the extended control, $\tau = t - t_0$ and where the least square estimate of coefficient a is .

$$a = \sum_{t_i=t_0}^{T} (\tau_i^n Y_i) \Big/ \sum_{t_i=t_0}^{T} \tau_i^{2n}, \tag{A2}$$

where Y_i is the raw projection of the studied variable for future period centred at year t_i and where T is the final future projection period (2080–2099 here). This expression holds for every strictly positive real number. For $n > 1$, it insures that the derivative is zero when $t = t_0$.

In case of a polynomial of degree 3, the trend expression reads:

$$x(t) = \begin{cases} x_0 & \text{for } t < t_0, \\ x_0 + a\tau^2 + b\tau^3 & \text{for } t \geq t_0, \end{cases} \tag{A3}$$

where the least-square estimation of both coefficients with constraints mentioned above are:

$$b = \sum_{t_i=t_0}^{T} \left(Y_i - \frac{m_y}{m_2}\,\tau_i^2 \right) \left(\tau_i^3 - \frac{m_3}{m_2}\,\tau_i^2 \right) \Big/ \sum_{t_i=t_0}^{T} \left(\tau_i^3 - \frac{m_3}{m_2}\,\tau_i^2 \right)^2, \tag{A4}$$

$$a = \frac{m_y - b\,m_3}{m_2} \tag{A5}$$

with

$$m_y = \sum_{t_i=t_0}^{T} Y_i, \tag{A6}$$

$$m_2 = \sum_{t_i=t_0}^{T} \tau_i^2, \tag{A7}$$

$$m_3 = \sum_{t_i=t_0}^{T} \tau_i^3. \tag{A8}$$

The choice of the trend model was made studying the homoscedasticity of the residuals and the relevance of the trend shape over the whole simulation period. The sensitivity analysis considered different types of models and different pivot years year0.

APPENDIX B

Internal variability components for relative changes

Uncertainty components in the case of relative changes are derived in this appendix following Appendix A of Hingray et al. (2007). Let consider the simulations outputs over the 1860–2099 period of modeling chain m defined by the r^{th} run of GCM/SDM combination g–c. Assuming that at any time the noise-free estimate of Y is a good approximation of the expected value of Y, approximations to the variance of the change variable X for this chain m can be obtained with the delta method (Stuart and Ord 1987, Sections 10.5 and 10.6):

$$\widehat{\mathbb{V}\mathrm{ar}}\left(X_{m,t}\right) \approx \left\{\frac{y(g,s,t)}{y(g,s,c)}\right\}^2 \left[\frac{\widehat{\mathbb{V}\mathrm{ar}}\left(Y_{m,t}\right)}{\left\{y(g,s,t)\right\}^2} + \frac{\widehat{\mathbb{V}\mathrm{ar}}\left(Y_{m,c}\right)}{\left\{y(g,s,c)\right\}^2}\right] \tag{B1}$$

where $\widehat{\mathbb{V}\mathrm{ar}}\left(X_{m,t}\right)$ denote the estimated total internal variability of the change variable X for modeling chain m and where $\widehat{\mathbb{V}\mathrm{ar}}\left(Y_{m,t}\right)$ and $\widehat{\mathbb{V}\mathrm{ar}}\left(Y_{m,c}\right)$ are the estimated variances of the raw data Y for the future period t and reference control period c, respectively.

Assuming that the large scale and small scale internal variability components are non correlated, both terms in the parentheses of the right hand side of this equation can be partitioned in their large and small scale components.

An estimate of the small scale variability component therefore reads:

$$\widehat{\mathbb{V}\mathrm{ar}}_k\left(X_{m,t}\right) \approx \left\{\frac{y(g,s,t)}{y(g,s,c)}\right\}^2 \left[\frac{\widehat{\mathbb{V}\mathrm{ar}}_k\left(Y_{m,t}\right)}{\left\{y(g,s,t)\right\}^2} + \frac{\widehat{\mathbb{V}\mathrm{ar}}_k\left(Y_{m,c}\right)}{\left\{y(g,s,c)\right\}^2}\right] \tag{B2}$$

Both two terms of Equation (B2) are equivalent to a coefficient of variation of Y with respect to the inter-generation variance. This coefficient of variation was assumed to be roughly constant for a given chain over the whole 1860–2099 simulation period. Its multi-period mean was used for estimating the SSIV of the relative change variable X for the modeling chain m when only one run is available. It reads:

$$\widehat{\mathbb{V}\mathrm{ar}}_k\left(X_{m,t}\right) \approx \left\{\frac{y(g,s,t)}{y(g,s,c)}\right\}^2 \frac{2}{T}\sum_{t=1}^{T} \widehat{\mathbb{V}\mathrm{ar}}_k\left(\frac{Y_{m,t}}{y(g,s,t)}\right). \tag{B3}$$

When multiple runs are available for a given GCM/SDM chain, the multirun mean of these estimates was retained as SSIV for this chain. In both cases, the SSIV for chain m is a function of time via the term $y(g,s,t)$ in the equation. The SSIV component of the relative change variable for the projection ensemble was next estimated for each future lead time t as the multimodel mean of these variances, from the following equation:

$$SSIV(t) \approx \frac{2}{N_g N_s} \sum_{g=1}^{N_g} \sum_{s=1}^{N_s} \left[\frac{1}{T N_{g,r}} \left\{ \frac{y(g,s,t)}{y(g,s,c)} \right\}^2 \sum_{r=1}^{N_{g,r}} \sum_{t=1}^{T} \widehat{\mathbb{V}\mathrm{ar}}_k \left(\frac{y(g,s,r,k,t)}{y(g,s,t)} \right) \right]. \quad \text{(B4)}$$

Note that Equation (B4) has a form similar to Equation (22). The raw variable is replaced by the raw variable normalized by its expected value and a multiplicative coefficient accounts for the relative change in expected values from the control to the future period. Note that for $t \leq c$, the SSIV is two times the multimodel mean of the coefficient variation of Y with respect to the inter-generation variance.

The large scale internal variability component has the same expression as that of SSIV in Equation (B2) but, due to the limited number of runs available, the inter-run variance cannot be estimated. Similarly to the assumption made for changes in temperature, we assume that the LSIV for Y with respect to the inter-run dispersion is, in term of coefficient of variation, constant over the whole simulation period. It follows that for any time t:

$$\widehat{\mathbb{V}\mathrm{ar}}_r \left(\frac{Y_{m,t}}{y(g,s,t)} \right) \approx \widehat{\mathbb{V}\mathrm{ar}}_T \left(\frac{Y_{m,t}}{y(g,s,t)} \right). \quad \text{(B5)}$$

When multiple runs are available for a chain, this variance component is estimated accounted for data from all runs. The LSIV component of the relative change variable X was next be estimated from the multimodel mean of the inter-period inter-run variance of $Y(g,s,r,\bullet,t)$ as:

$$LSIV(t) = \frac{2}{N_g N_s} \sum_{g=1}^{N_g} \sum_{s=1}^{N_s} \left[\left\{ \frac{y(g,s,t)}{y(g,s,c)} \right\}^2 \widehat{\mathbb{V}\mathrm{ar}}_{T,N_{g,r}} \left(\frac{Y(g,s,r,\bullet,t) - y(g,s,t)}{y(g,s,t)} \right) \right]. \quad \text{(B6)}$$

REFERENCES

Alexandru, A., R. De Elia, and R. Laprise, 2007: Internal variability in regional climate downscaling at the seasonal scale. *Monthly Weather Review*, **135**, 3221–3238.

Berrington De González, A. and D. R. Cox, 2007: Interpretation of interaction: A review. *The Annals of Applied Statistics*, **1 (2)**, 371–385.

Boé, J., L. Terray, F. Habets, and E. Martin, 2006: A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling. *Journal of Geophysical Research: Atmospheres*, **111**, D23106, doi:10.1029/2005JD006889.

Brasseur, G. P. and E. Roeckner, 2005: Impact of improved air quality on the future evolution of climate. *Geophysical Research Letters*, **32**, L23704, doi:10.1029/2005GL023902.

Braun, M., D. Caya, A. Frigon, and M. Slivitzky, 2012: Internal variability of the Canadian RCMs hydrological variables at the basin scale in Quebec and Labrador. *Journal of Hydrometeorology*, **13 (2)**, 443–462.

Buishand, T. A. and T. Brandsma, 2001: Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling. *Water Resources Research*, **37 (11)**, 2761–2776.

Chardon, J., B. Hingray, A. C. Favre, P. Autin, J. Gailhard, I. Zin, and C. Obled, 2014: Spatial similarity and transferability of analog dates for precipitation downscaling over France. *J.Climate (in revision)*.

Chen, J., F. P. Brissette, A. Poulin, and R. Leconte, 2011: Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research*, **47**, W12509, doi:10.1029/2011WR010602.

824 Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate
825 change projections: the role of internal variability. *Climate Dynamics*, **38 (3-4)**,
826 527–546, doi:10.1007/s00382-010-0977-x, URL `http://link.springer.com/10.1007/`
827 `s00382-010-0977-x`.

828 Dobler, C., S. Hagemann, R. L. Wilby, and J. Stötter, 2012: Quantifying different sources
829 of uncertainty in hydrological projections in an Alpine watershed. *Hydrology and Earth*
830 *System Sciences*, **16 (11)**, 4343–4360.

831 Dufresne, J.-L., J. Quaas, O. Boucher, S. Denvil, and L. Fairhead, 2005: Contrasts in the
832 effects on climate of anthropogenic sulfate aerosols between the 20th and the 21st century.
833 *Geophysical Research Letters*, **32**, L21703, doi:10.1029/2005GL023619.

834 Hawkins, E. and R. Sutton, 2009: The potential to narrow uncertainty in regional climate
835 predictions. *Bulletin of the American Meteorological Society*, **90 (8)**, 1095–1107.

836 Hawkins, E. and R. Sutton, 2011: The potential to narrow uncertainty in projections of
837 regional precipitation change. *Climate Dynamics*, **37 (1–2)**, 407–418.

838 Hingray, B., A. Mezghani, and T. A. Buishand, 2007: Development of probability distribu-
839 tions for regional climate change from uncertain global mean warming and an uncertain
840 scaling relationship. *Hydrology and Earth System Sciences*, **11 (3)**, 1097–1114.

841 Hingray, B. and M. Saïd, 2014: Partitioning internal variability and model uncertainty in a
842 multimodel ensemble of climate projections: influence of the number of model runs. *To be*
843 *submitted to Geophys. Res. Let.*

844 Hingray, B., et al., 2013: RIWER2030: Climats Régionaux et Incertitudes, Ressource en Eau
845 et Gestion associée de 1860 à 2100. Tech. rep., CNRS/LTHE, EDF/LNHE, EDF/DTG,
846 60p.+Annexes pp., Grenoble, FR. Projet ANR VMCS 2009-2012. Rapport Final.

847 Huebener, H., U. Cubasch, U. Langematz, T. Spangehl, F. Niehörster, I. Fast, and M. Kunze,
848 2007: Ensemble climate simulations using a fully coupled ocean-troposphere-stratosphere
849 general circulation model. *Philosophical Transactions of the Royal Society A: Mathemati-*
850 *cal, Physical and Engineering Sciences*, **365 (1857)**, 2089–2101.

851 IPCC, 2007: Climate Change 2007: The Physical Science Basis. *Contribution of Working*
852 *Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate*
853 *Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor,
854 and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New
855 York, NY, USA, pp. 976.

856 Johns, T., et al., 2011: Climate change under aggressive mitigation: the ensembles multi-
857 model experiment. *Climate Dynamics*, **37 (9–10)**, 1975–2003.

858 Kalnay, E., et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bulletin of the Amer-*
859 *ican Meteorological Society*, **77 (3)**, 437–471.

860 Karoly, D. J. and Q. Wu, 2005: Detection of regional surface temperature trends. *Journal*
861 *of Climate*, **18 (21)**, 4337–4343, doi:10.1175/JCLI3565.1, URL `http://libra.msra.cn/`
862 `Publication/5645608/detection-of-regional-surface-temperature-trends`.

863 Lafaysse, M., 2011: Changement climatique et régime hydrologique d'un bassin alpin.
864 Génération de scénarios sur la Haute-Durance, méthodologie d'évaluation et incertitudes
865 associées. Ph.D. thesis, Université Toulouse III – Paul Sabatier, FR.

866 Lafaysse, M., B. Hingray, L. Terray, J. Gailhard, and A. Mezghani, 2014: Internal vari-
867 ability and model uncertainty components in a multireplicate multimodel ensemble of
868 hydrometeorological projections : the alpine durance basin. *Water Resources Research*,
869 **50**, doi:10.1002/2013WR014897.

870 Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level
871 pressure. *Monthly Weather Review*, **104 (7)**, 942–952, doi:10.1175/1520-0493(1976)

872 104⟨0942:EOTNVO⟩2.0.CO;2, URL http://libra.msra.cn/Publication/2143198/
873 estimates-of-the-natural-variability-of-time-averaged-sea-level-pressure.

874 May, W., 2008: Climatic changes associated with a global "2C–stabilization" scenario sim-
875 ulated by the ECHAM5/MPI-OM coupled climate model. *Climate Dynamics*, **31 (2–3)**,
876 283–313.

877 Mezghani, A. and B. Hingray, 2009: A combined downscaling-disaggregation weather gener-
878 ator for stochastic generation of multisite hourly weather variables over complex terrain:
879 Development and multi-scale validation for the Upper Rhone River basin. *Journal of Hy-*
880 *drology*, **377 (3–4)**, 245 – 260.

881 Nikiéma, O. and R. Laprise, 2013: An approximate energy cycle for inter-member variability
882 in ensemble simulations of a regional climate model. *Climata Dynamics*, **41 (3–4)**, 831–
883 852.

884 Northrop, P. J., 2013: Comments on A simple, coherent framework for partition-
885 ing uncertainty in climate predictions. *Journal of Climate*, **26 (12)**, 4375–4376, doi:
886 10.1175/JCLI-D-12-00527.1, URL http://journals.ametsoc.org/doi/abs/10.1175/
887 JCLI-D-12-00527.1.

888 Obled, C., G. Bontron, and R. Garçon, 2002: Quantitative precipitation forecasts: a sta-
889 tistical adaptation of model outputs through an analogues sorting approach. *Atmospheric*
890 *Research*, **63 (3–4)**, 303 – 324.

891 Räisänen, J., 2001: Co2-induced climate change in cmip2 experiments: Quantification of
892 agreement and role of internal variability. *Journal of Climate*, **14 (9)**, 2088–2104.

893 Salas-Mélia, D., et al., 2005: Description and validation of CNRM-CM3 Global Coupled
894 Model. Tech. Rep. 103, CNRM, Météo-France, 36p. pp., Toulouse, FR. Available at http:
895 //www.cnrm.meteo.fr/scenario2004/paper_cm3.pdf.

Sansom, P. G., D. B. Stephenson, C. A. T. Ferro, G. Zappa, and L. Shaffrey, 2013: Simple uncertainty frameworks for selecting weighting schemes and interpreting multimodel ensemble climate change experiments. *Journal of Climate*, **26 (12)**, 4017–4037.

Searle, S. R., 1971: *Linear Models*. John Wiley & Sons, New York, NY, USA.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317 (5839)**, 796–799.

Stuart, A. and J. Ord, 1987: *Kendall's Advanced Theory of Statistics*, Vol. 1: Distribution Theory. 5th ed., Charles Griffin, London, UK.

Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins, 2011: A simple, coherent framework for partitioning uncertainty in climate predictions. *Journal of Climate*, **24 (17)**, 4634–4643.

List of Tables

TABLE 1. **GCM experiments for the 20C3M+SRESA1B (1860–2100) from the ENSEMBLES Stream 2 EU research project (Johns et al. 2011).**

Centre	**Model**	**Acronym**	**Run(s)**	**Reference(s)**
CNRM (France)	CNRM-CM3.3	CNCM33	1	Salas-Mélia et al. (2005)
DMI (Denmark)	ECHAM5-C	DMIEH5C	3	May (2008)
MPI (Germany)	ECHAM5-C	MPIEH5C	3	Brasseur and Roeckner (2005)
IPSL (France)	IPSL-CM4v2	IPCM4v2	3	Dufresne et al. (2005)
FUB (Germany)	EGMAM2	EGMAM2	1	Huebener et al. (2007)

TABLE 2. **Statistical Downscaling Models from the RIWER2030 ANR project (Lafaysse et al. 2014; Hingray et al. 2013).** Predictors used by the models are: SLP: sea level pressure, Z: geopotential height for the pressure level, Ta: surface temperature, u and v: zonal and meridian geostrophic wind, HU: relative humidity, Fq: moisture flux, subscript refers to pressure level. The number of generation for each GCM/SDM chain is 100.

Centre	**Method**	**Acronym**	**Predictors**	**Reference(s)**
EDF (France)	k-nearest neighbors	analog20	Z_{700}, Z_{1000}	Obled et al. (2002), Chardon et al. (2014)
Cerfacs (France)	Weather Type + Transfer function	dsclim10	SLP	Boé et al. (2006), Lafaysse (2011)
		dsclim11	SLP, Ta	
		dsclim21	Z_{850}, Z_{500}	
LTHE (France)	Transfer function + k-nearest neighbors	d2gen10	SLP, u_{700}, v_{700}, Ta_{700}	Mezghani and Hingray (2009)
		d2gen32	SLP, u_{700}, v_{700}, Ta_{700}, HU_{700}, Fq_{700}	

List of Figures

4 **Grand-ensemble mean climate change response $\mu(t)$ and total uncertainty in 20-yr mean temperature and precipitation projections for the Upper Durance River Catchment as a function of projection lead time (reference period 1980–1999).** The total colored area covered by all uncertainty components corresponds to $\mu(t) \pm 1.645\sqrt{T(t)}$ where $T(t)$ is the total uncertainty variance. The time axis gives the midpoint of the averaging future period. Uncertainty components are shown in dark blue for GCM uncertainty, green for SDM uncertainty, cyan for Residuals/Model Interaction (R/MI), red for large scale internal variability (LSIV) and yellow for small scale internal variability (SSIV) of the GCM/SDM chains. For each model uncertainty and internal variability component, the vertical extent of the corresponding area is proportional to the fraction of total uncertainty explained by the component. This fraction is obtained from the ratio (standard deviation of uncertainty component)/(standard deviation of total uncertainty).

5 **Fraction of total variance explained by each source of uncertainty for expected mean change in 20-yr mean temperature and precipitation projections as a function of projection lead time.** For details see caption of Figure 1.

6 **Response-to-uncertainty ratio for expected mean change in 20-yr mean temperature and precipitation projections as a function of time (reference period is 1980–1999).** The continuous black line with circles is the potential response-to-uncertainty ratio assuming zero model uncertainty. In both case, the climate change response is the grand-ensemble mean $\mu(t)$ of the ensemble of experiments obtained from Equation (9). The uncertainty corresponds to the 90% confidence level.

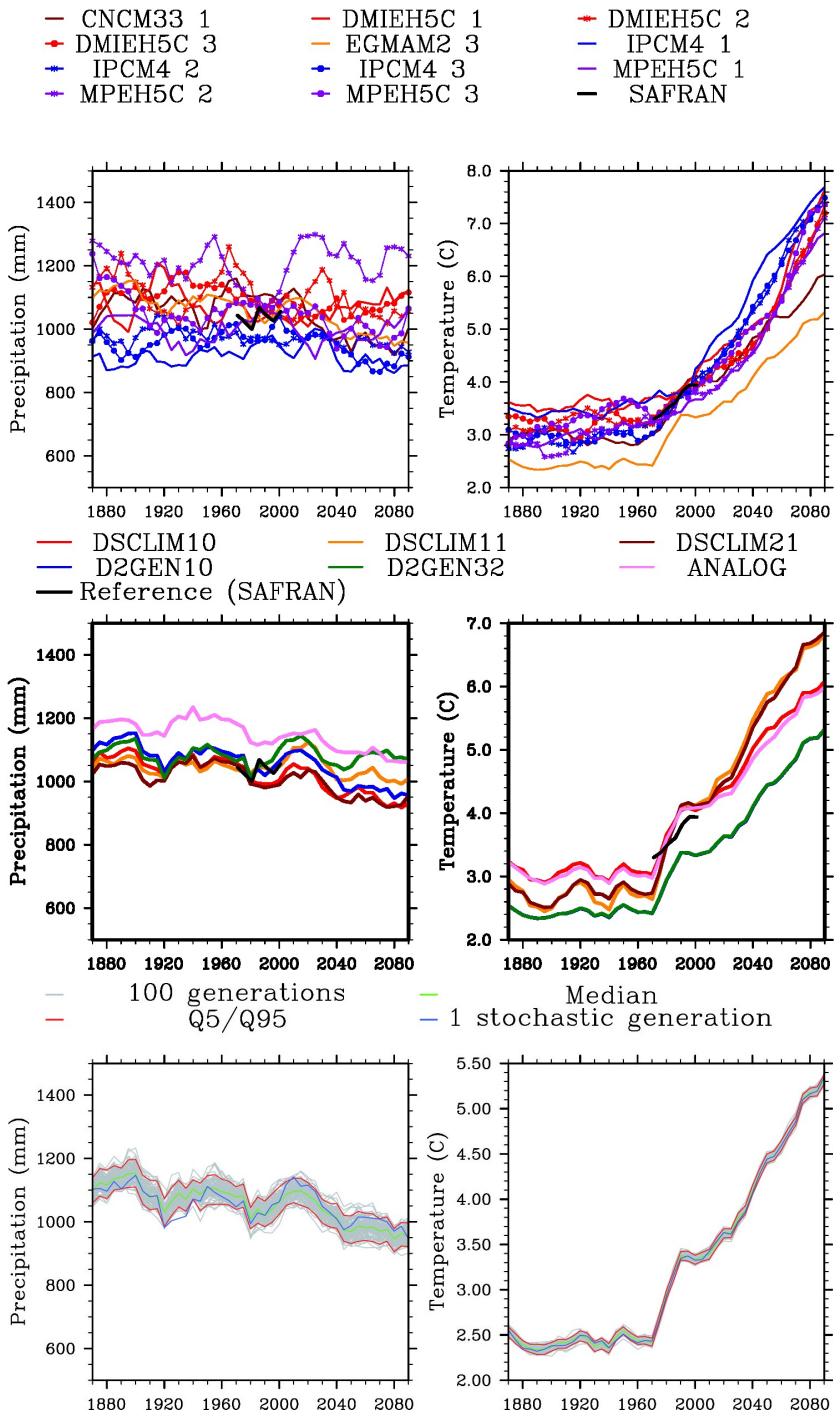

FIG. 1. **20-year running mean of annual temperature (right) and precipitation (left) versus time over the period 1860–2100.** Bottom: 100 scenarios (gray lines) obtained from the DMIEH5C-1/DSCLIM-10 modeling chain (the bottom red, green and top red curves correspond respectively to 5th percentile, median and 95th percentile; the blue curve corresponds to one scenario randomly selected from the set of 100). Middle and top: Multi-generation mean of the 100 scenarios obtained with DSCLIM-10 when forced by each of the 11 GCM experiments (top) or with all SDMs forced by the DMIEH5C-1 GCM experiment (middle).

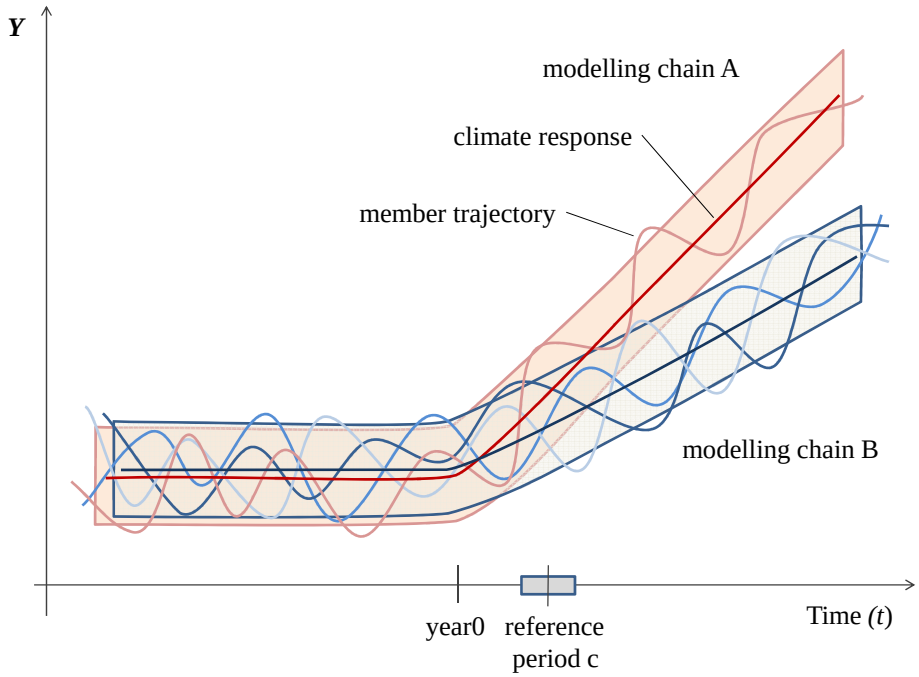

FIG. 2. **Time series members, climate responses and internal variability for two modeling chains over a control+transient simulation period.**

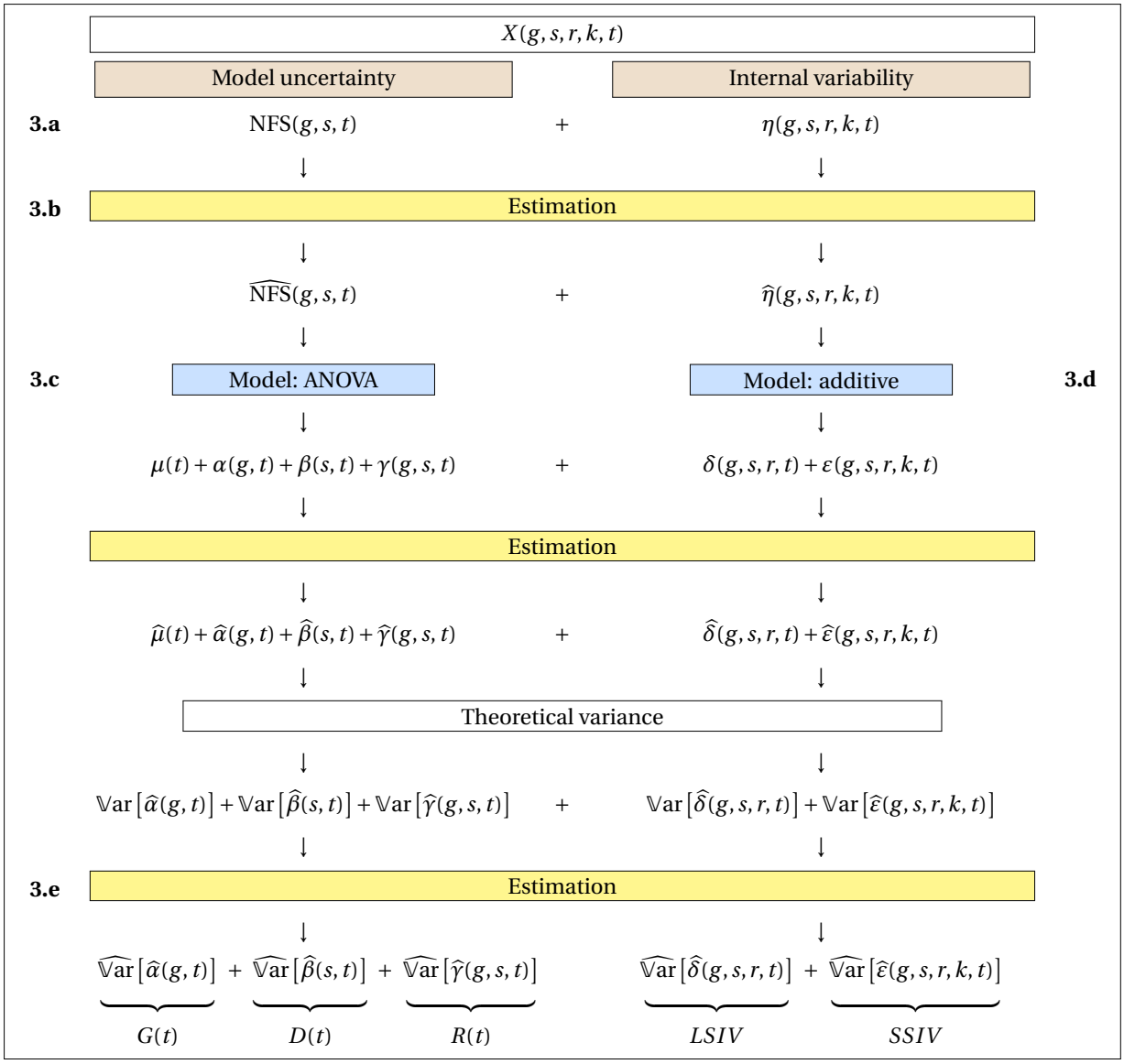

FIG. 3. **Schematic flow diagram of the Quasi-Ergodic ANOVA (QE-ANOVA) framework applied for partitioning model uncertainty components** (GCM uncertainty, SDM uncertainty, Residual-GCM/SDM interaction) **and internal variability components** (large scale, small scale).The number of the section where the methodology and/or estimation is presented is mentioned in the margins of the figure.

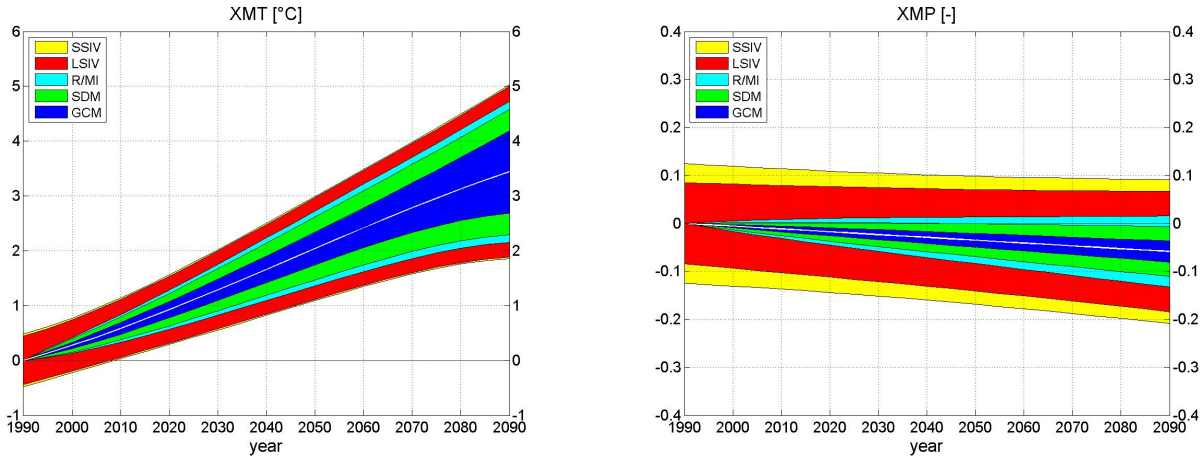

FIG. 4. **Grand-ensemble mean climate change response $\mu(t)$ and total uncertainty in 20-yr mean temperature and precipitation projections for the Upper Durance River Catchment as a function of projection lead time (reference period 1980–1999).** The total colored area covered by all uncertainty components corresponds to $\mu(t) \pm 1.645\sqrt{T(t)}$ where $T(t)$ is the total uncertainty variance. The time axis gives the midpoint of the averaging future period. Uncertainty components are shown in dark blue for GCM uncertainty, green for SDM uncertainty, cyan for Residuals/Model Interaction (R/MI), red for large scale internal variability (LSIV) and yellow for small scale internal variability (SSIV) of the GCM/SDM chains. For each model uncertainty and internal variability component, the vertical extent of the corresponding area is proportional to the fraction of total uncertainty explained by the component. This fraction is obtained from the ratio (standard deviation of uncertainty component)/(standard deviation of total uncertainty).

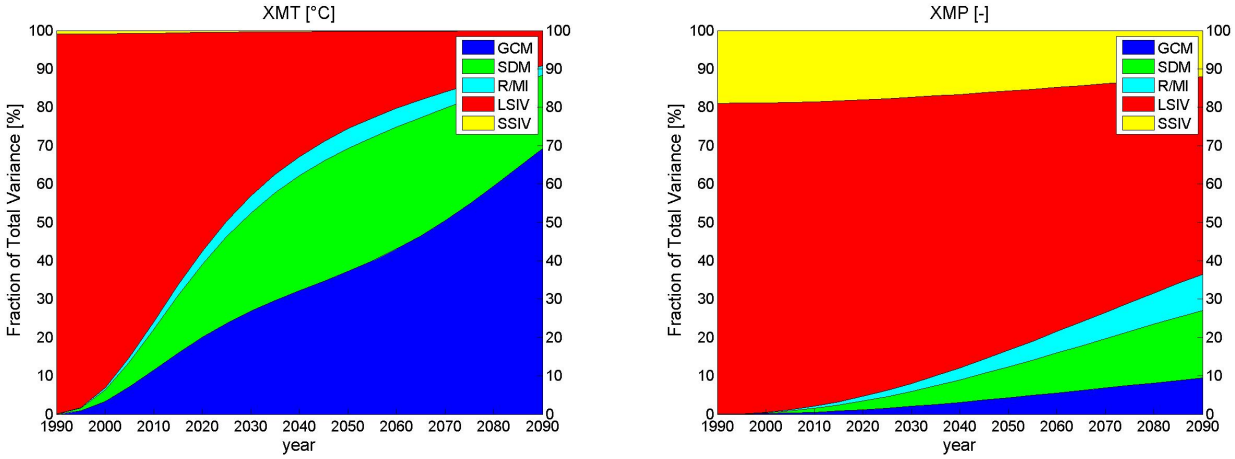

FIG. 5. **Fraction of total variance explained by each source of uncertainty for expected mean change in 20-yr mean temperature and precipitation projections as a function of projection lead time.** For details see caption of Figure 1.

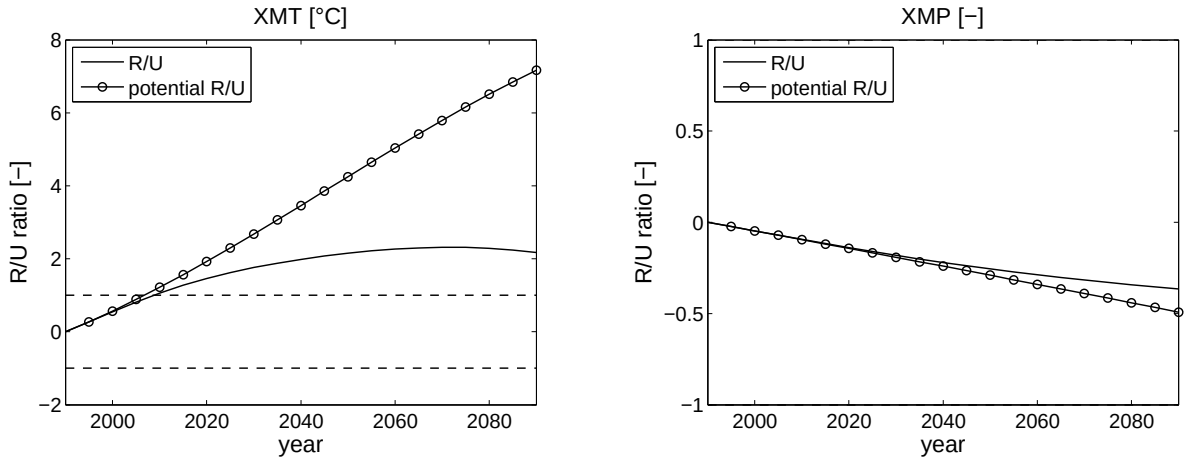

FIG. 6. **Response-to-uncertainty ratio for expected mean change in 20-yr mean temperature and precipitation projections as a function of time (reference period is 1980–1999).** The continuous black line with circles is the potential response-to-uncertainty ratio assuming zero model uncertainty. In both case, the climate change response is the grand-ensemble mean $\mu(t)$ of the ensemble of experiments obtained from Equation (9). The uncertainty corresponds to the 90% confidence level.

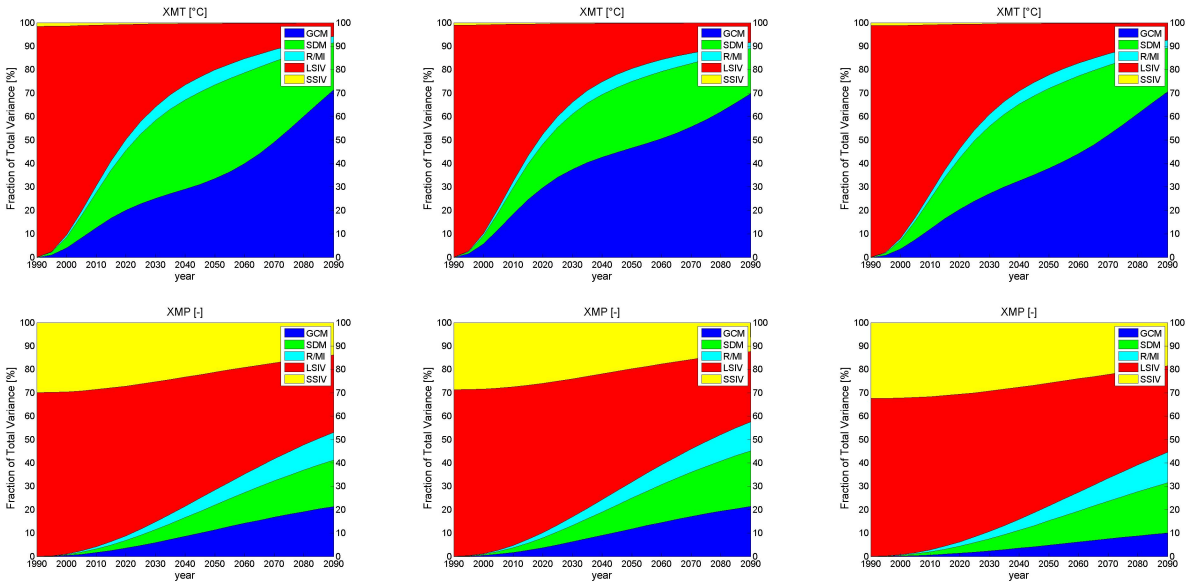

FIG. 7. **Influence of GCM run selection on fraction of total variance explained by each source of uncertainty for expected mean change in 20-yr mean temperature and precipitation projections as a function of time.** Analysis is based with either the first (left), the second (middle) or the third runs (right) of each GCM for which 3 runs are available. Graphs have to be compared to those of Figure 5. For details see caption of Figure 4.

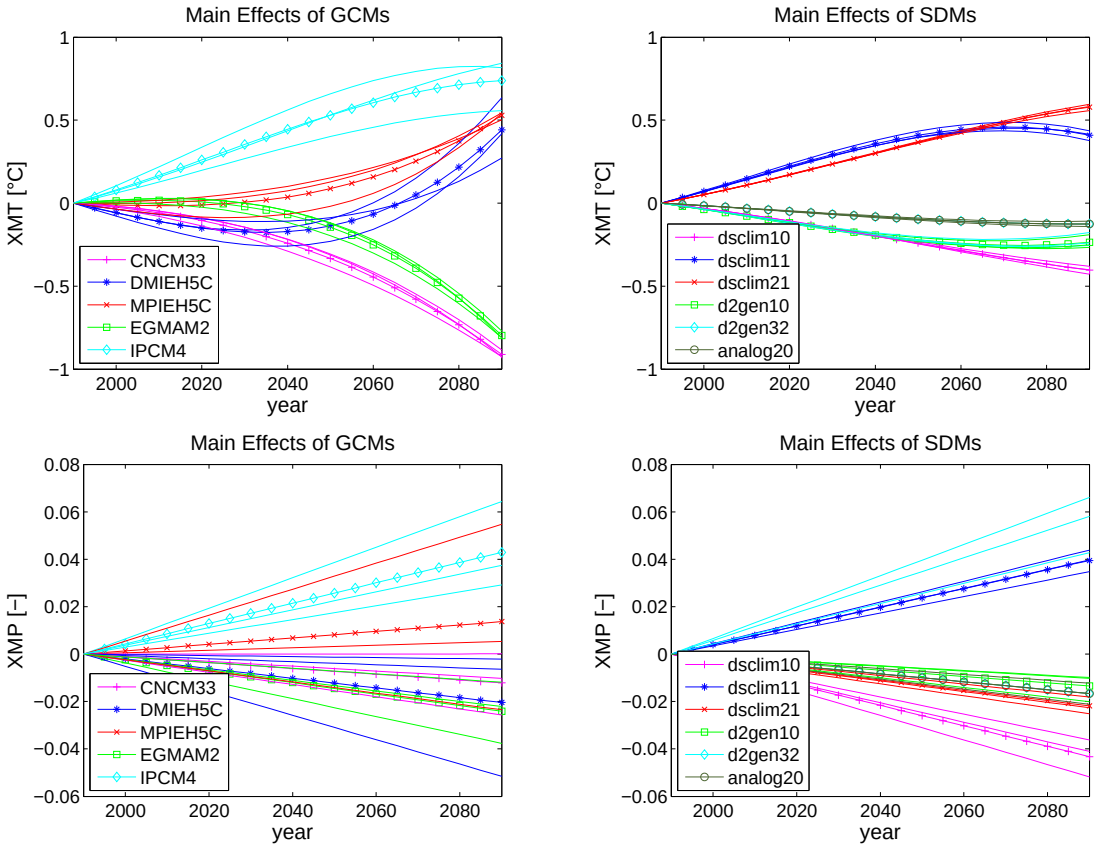

FIG. 8. **Influence of GCM run selection on the main effects of GCM (left) and SDM (right) for expected mean change in 20-yr mean temperature (top) and precipitation (bottom) projections as a function of time.** Model effects are either obtained from the multiple runs analysis (lines with markers) or from a single run analysis (lines without markers). Lines of a same color correspond to the effect of a same model. Note that for precipitation, the main effects of both GCM and SDM present a linear dependency to time as a result of the trend model used to estimate the NFS of each GCM/SDM chain.