



Bayesian multi-model projections of climate: generalization and application to ENSEMBLES results

C. M. Buser¹, H. R. Künsch^{1,*}, C. Schär²

¹Seminar for Statistics and ²Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland

ABSTRACT: In a previous study, we developed a Bayesian methodology for combining multi-model climate change simulations into a single probabilistic projection which addresses changes in inter-annual variability beyond changes in mean temperature, and which explicitly considers time-dependent model biases. We tested 2 different but equally plausible bias assumptions referred to as 'constant bias' and 'constant relationship'. The former assumes that the biases in control and scenario periods are approximately constant, following the implicit assumption in most climate change studies. The latter approach follows seasonal forecasting procedures by assuming an approximate linear relationship between observed and simulated seasonal temperatures. In the present study we generalized this approach by combining the 2 bias assumptions into a single probabilistic projection. In cases where the 2 assumptions yield conflicting results, our methodology implicates a broader probability density function, thereby reflecting the increased level of uncertainty. We applied the new method to area-mean seasonal temperature distributions from global/regional climate model simulations of the ENSEMBLES project. Results are presented for changes in mean and variability between control (1961–1990) and scenario (2021–2050) periods. In comparison to the multi-model mean, the generalized Bayes method projected a considerably weaker warming during summer and autumn in much of continental Europe, a stronger winter warming in Scandinavia, France, eastern and central Europe, and a weaker warming in both summer and winter in the Mediterranean. These differences can be traced back to the models' difficulties in representing the natural interannual variability in these regions.

KEY WORDS: Climate projections · Regional climate model · Europe · Uncertainty · Bayesian methods · Interannual variability · Model bias · Bias change

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Climate projections and associated applications in impact studies have become an important topic of scientific and public interest during the last decades. Increasingly, they are based on large multi-model ensembles. For instance, the fourth Intergovernmental Panel on Climate Change (IPCC) assessment report summarizes data from 21 different coupled atmosphere–ocean global climate models (GCMs; Meehl et al. 2007). Similarly, regional projections are increasingly based on ensembles of high-resolution regional climate model (RCM) simulations. Over Europe, this approach has

been pioneered in the PRUDENCE and ENSEMBLES projects (Hewitt & Griggs 2004, Christensen & Christensen 2007, see also <http://ensembles-eu.metoffice.com>). Typically, ensemble mean and ensemble spread are used to visualize and present the consensus and the uncertainty of the models.

The ensemble approach was originally derived in probabilistic medium-range weather forecasting, and is now broadly used in numerical weather prediction, seasonal forecasting, and climate research on a wide range of time scales. Recent developments include the exploitation of multi-model ensembles and the investigation of single-model perturbed-physics ensembles,

*Corresponding author. Email: kuensch@stat.math.ethz.ch

as well as the use of stochastic forcings in promoting ensemble spread. Applications geared towards climate projections are usually based on a heterogeneous ensemble (with sometimes only a handful of ensemble members, stemming from different models in an only partly coordinated framework).

An important issue in ensemble approaches for climate research is the inability to rigorously quantify climate model biases. While biases of climate models are monitored for the control period, the lack of long-term comprehensive observations (on the centennial time scales considered) implies that it is difficult to decide how the model biases will change with the climate state. As a result of these difficulties, it is more challenging to provide an estimate in the form of a probability density function (PDF) from climate ensembles. Ideally, one would like to employ an objective methodology that takes into consideration the individual model results as well as their performance under current climatic conditions.

The Bayesian framework is particularly attractive for combining several models. Tebaldi et al. (2005) were among the first to use the Bayesian framework to analyze multi-model climate predictions. They derived PDFs for the mean temperature changes in 22 regions covering global land mass and 4 seasons by combining 30 yr regional climate averages from observations and from output of several GCMs. Their approach can be viewed as a weighted average of the individual GCM results, with weights similar to those used by the reliability ensemble average (REA) of Giorgi & Mearns (2002). The framework of Tebaldi et al. (2005) has been generalized in many directions. Furrer et al. (2007) and Smith et al. (2009) developed methods to study several regions simultaneously. Jun et al. (2008) analyzed the spatial variability of the additive biases in detail for the control climate. Tebaldi & Sanso (2008) and Buser et al. (2010) introduced multivariate generalizations for analyzing temperature and precipitation simultaneously. Tebaldi & Knutti (2007) reviewed multi-model climate projections and different types of uncertainty. They also discussed the problems of model dependence, tuning, and evaluation.

This paper continues the recent study of Buser et al. (2009), which extended the previous methodologies in 2 important respects. First, we considered not only changes in mean temperature, but also changes in interannual variability. Second, we explicitly considered changes of climate model biases, i.e. the biases were assumed to depend upon the climate state and were thus allowed to change between control and scenario periods. More specifically, we considered both additive and multiplicative biases, which relate to the climate model's ability to represent the mean climate and its interannual variations, respectively.

Two plausible assumptions about the extrapolation of additive biases into the future are introduced, referred to as the 'constant bias' and 'constant relation' assumptions. The former is used implicitly in most studies of climate change. The latter asserts that over-/underestimation of the interannual variability in the control period also leads to over-/underestimation of climate change, and this assumption is closely related to the statistical post-processing of seasonal climate predictions. Buser et al. (2009) explicitly allowed the additive and multiplicative model biases to change between control and scenario periods and used the Bayesian methodology to estimate the bias changes. The resulting lack of identifiability was resolved by using informative priors.

The results of Buser et al. (2009) showed that bias assumptions may critically affect the results. For instance, the application of the methodology to Alpine summer temperatures using PRUDENCE simulations for the periods 1961–1990 and 2071–2100, yielded different mean warmings, i.e. 5.4 K and 3.4 K for the 'constant bias' and 'constant relation' assumption, respectively. These significant differences demonstrate that bias changes should explicitly be considered as an additional source of uncertainty in climate projections.

The main contribution of the present study is an extension of Buser et al.'s (2009) approach using a method that includes the 2 bias assumptions as special cases. The generalization yields a single probabilistic projection, instead of 2 different projections that depend upon the underlying bias assumptions. Nevertheless, uncertainties due to bias assumptions are implicitly taken into account. For instance, if there is a disagreement between the different bias assumptions, the generalized approach increases the uncertainty of the resulting projections.

We applied the new Bayesian methodology to the recent set of GCM/RCM simulations from the ENSEMBLES project, <http://ensembles-eu.metoffice.com>, which used transient simulations from 1950–2050 with observed forcings for the past and an SRES A1B emission scenario for the future. In our study, we restricted attention to the 1961–1990 control and 2021–2050 scenario periods.

The paper is structured as follows. In Section 2, the data and the aggregation procedure are described. In Section 3, the methods and the Bayesian model setup are explained. In Section 4, results for the 8 regions and 4 seasons are shown, and in Section 5, we draw our conclusions.

2. DATA

In this paper, both observational data and output of RCMs are summarized by the term 'data.' We distinguish between 'known data', namely the observations

for the control period and the model output for both the control and the scenario period, and ‘unknown data’, namely the observations for the scenario period. The variable of interest is the 2 m temperature (above ground). The control period is 1961–1990 (CTL) and the scenario period is 2021–2051 (SCEN). The aim was to predict changes in mean temperature and inter-annual temperature variability between the 2 periods.

2.1. RCM data

For the statistical analysis, we used the output of 7 RCMs which are all part of the ENSEMBLES project (Hewitt & Griggs 2004, see also <http://ensembles-eu.metoffice.com>). All models considered were driven by transient GCM runs for the 1951–2051 period, as forced by observed greenhouse gas, aerosol, solar and volcanic forcings for the past (1951–2000) and an A1B emission scenario for the future (Nakicenovic et al. 2000). As our methodology required independent simulations, we used a subset of ENSEMBLES simulations whose members were driven by different atmospheric GCM simulations. The 7 simulations considered are listed in Table 1. Further information on the individual GCM and RCM simulations is available in Niehörster et al. (2008), Déqué (2009), and from the ENSEMBLES website. The 3 HadCM3 simulations used were based on the same model but with different parameter settings, in order to yield a small perturbed-physics ensemble with different climate sensitivities (see Murphy et al. 2007). For the purpose of the current paper, these simulations can be considered as independent.

The integration area of the RCM models varied, but in all cases covered the majority of Europe. The focus

of our analysis was on 8 standard regions (PRUDENCE regions) within Europe (Christensen & Christensen 2007). Table 2 provides the corresponding coordinates of the regions. The spatial resolution of the models was around 0.22° (~ 25 km). Model output was interpolated on a regular latitude–longitude grid (see Section 2.2) so that it could easily be compared to observations from the control period.

2.2. Observational data

We use the observed surface temperature data (E-OBS) from the ENSEMBLES project. The data were located on a regular $0.22^\circ \times 0.22^\circ$ latitude–longitude grid and covered the period 1950–2006. Haylock et al. (2008) provided a detailed description of the data set. The data can be accessed via <http://eca.knmi.nl/download/ensembles/download.php>. In the analysis, we assumed that the observations represent the true climate.

Table 2. Coordinates (degrees) of analysis regions, based on Christensen & Christensen (2007)

| Area | Latitude | Longitude |
|--------------------------|-------------|-------------|
| 1 British Isles (BI) | 50° N–59° N | 10° W–2° E |
| 2 Iberian Peninsula (IP) | 36° N–44° N | 10° W–3° E |
| 3 France (Fr) | 44° N–50° N | 5° W–5° E |
| 4 Middle-Europe (ME) | 48° N–55° N | 2° E–16° E |
| 5 Scandinavia (Sc) | 55° N–70° N | 5° E–30° E |
| 6 Alps (Al) | 44° N–48° N | 5° E–15° E |
| 7 Mediterranean (Md) | 36° N–44° N | 3° E–25° E |
| 8 Eastern Europe (EE) | 44° N–55° N | 16° E–30° E |

Table 1. ENSEMBLES data overview. We used a subset of models driven by different atmospheric global climate model (GCM) runs. The 3 HadCM3 simulations are members of a perturbed-physics ensemble. RCM: regional climate model

| Institute and location | Model (source) | |
|--|-------------------------------|---|
| | RCM | GCM |
| C41, Dublin, Ireland | RCA3 (Kjellström et al. 2005) | HadCM3 high sensitivity (Collins et al. 2006) |
| Centre National de Recherches Météorologiques (CNRM), Toulouse, France | ALADIN (Radu et al. 2008) | Arpege (Gibelin & Déqué 2003) |
| Eidgenössische Technische Hochschule (ETH), Zurich, Switzerland | CLM (Jaeger et al. 2008) | HadCM3 standard sensitivity (Collins et al. 2006) |
| Hadley Centre, Exeter, UK | HadRM | HadCM3 low sensitivity (Collins et al. 2006) |
| Koninklijk Nederlands Meteorologisch Instituut (KNMI), Amsterdam, Netherlands | RACMO (Lenderink et al. 2003) | ECHAM5 (Roeckner et al. 2003) |
| Ouranos, Montreal, Canada | CRCM (Plummer et al. 2006) | CGCM3 (Scinocca et al. 2008) |
| Sveriges Meteorologiska och Hydrologiska Institut (Swedish Meteorological and Hydrological Institute, SMHI), Norrköping, Sweden | RCA (Kjellström et al. 2005) | BCM (Furevik et al. 2003) |

2.3. Aggregation

We averaged the data both temporally over the 3 months of each season (spring: MAM, summer: JJA, autumn: SON, winter: DJF) and spatially over all land grid points within the region. For the spatial average, a grid point was considered a land point if at least 50 % of the corresponding area was landmass. Water grid points were excluded from all models and the E-OBS data set to avoid a mixing of the sea and land temperatures.

The statistical analysis was done separately for each season and each region. We thus ignored correlations between neighboring seasons and regions. We did not average over the years, but rather retained the information about interannual variability. A potential difficulty of this is that trends during the 2 periods become confounded with the interannual variability. In order to avoid this, we included linear trends in our statistical model.

3. METHODS

3.1. Bayesian formalism

Our main goal was the estimation of the *a priori* unknown parameters that represent changes in mean and variance of the temperature, given the known (simulated and observed) data as described in the previous section. In the Bayesian formalism, uncertainty about these and other parameters is represented in the form of probability distributions. It starts with prior probability distributions, i.e. probability distributions representing the *a priori* knowledge, and converts these by Bayes formula into posterior distributions which also contain the information from the data.

Whenever possible, we used uninformative priors, but for parameters representing bias changes, we had to use informative priors in order to resolve a lack of identifiability. This is discussed in more detail in Section 3.3 when we specify our prior distributions.

According to Bayes formula, the posterior density of the parameters given the known data is obtained as

$$p(\text{parameters} | \text{known data}) \propto p(\text{known data} | \text{parameters}) \times p(\text{parameters})$$

Here $p(\text{parameters})$ is the prior density and $p(\text{data} | \text{parameters})$ is the so-called likelihood. Its specification is discussed in Section 3.2.

Often, one is interested not only in parameters, but also in the distribution of unknown data, for instance the distribution of the temperature in the scenario period. Since this distribution depends on the unknown parameters, we computed the so-called posterior predictive density by integrating

$$p(\text{unknown data} | \text{parameters}) \times p(\text{parameters} | \text{known data})$$

over the parameters. It represents the combined uncertainty due to random variability of the unknown data and due to unknown parameters.

Although these formulas look simple, the resulting densities are typically not of a standard form, and explicit results require high dimensional integration. Common practice in modern statistics is to rely on Markov chain Monte Carlo methods instead. Monte Carlo methods replace analytical calculations with empirical estimates, computed with an artificially generated sample from the posterior distribution. For complicated, high-dimensional distributions, it is not feasible to generate an independent sample, but it is possible to generate a dependent sample with a suitable Markov chain. This means that each member of the sample is constructed recursively from its predecessor (see e.g. Gilks et al. 1996).

3.2. Distribution of data

All variables are mean temperatures for a specific region and season. $X_{0,t}$ denotes the observations in year $1960+t$, $X_{i,t}$ the control output of model i in year $1960+t$, and $Y_{i,t}$ the scenario output of model i in year $2020+t$ where $t = 1, 2, \dots, 30$ and $i = 1, 2, \dots, 7$. These are the known data. The unknown data are the observations in year $2020+t$, which we denote by $Y_{0,t}$ ($t = 1, \dots, 30$). No indices are needed for season and region since the analysis is done individually for each region and season.

We assumed that all data were normally distributed and independent. The assumption of normal distributions is reasonable due to the aggregation over a season and a region. In addition, quantile plots of observations and model data against the theoretical normal distribution did not show strong discrepancies. 'Independent' means that we ignored, on the one hand possible correlations between years, and on the other hand possible correlations between models or between models and observations. The autocorrelation plots of the series confirmed that the first part is reasonable. The absence of correlations between models and observations was justified because the models were supposed to reproduce only the climate and not the interannual variations of a specific year. We justify the independence between models by the fact that there was no explicit dependence between the models used in our analysis. Still, all of our models were based on the same scientific knowledge and thus may have contained the same systematic errors. This could have led to an underestimate of the uncertainty, which we are unable to assess and quantify.

Because we assumed independent normal distributions, we only had to specify the means (E) and vari-

ances (Var). We begin with the variables $X_{0,t}$, $X_{i,t}$, and $Y_{0,t}$ where the specification is rather straightforward:

$$E(X_{0,t}) = \mu + \gamma(t - T_0), \text{Var}(X_{0,t}) = \sigma^2 \quad (1)$$

$$E(X_{i,t}) = \mu + \beta_i + (\gamma + \gamma_i)(t - T_0), \text{Var}(X_{i,t}) = b_i^2 \sigma^2 \quad (2)$$

$$E(Y_{0,t}) = \mu + \Delta\mu + (\gamma + \Delta\gamma)(t - T_0), \text{Var}(Y_{0,t}) = q^2 \sigma^2 \quad (3)$$

Here $T_0 = 15.5$, so that μ is the expected temperature in the middle of the control period. The trend parameter γ is the yearly increase of the mean temperature in the control period, and σ is the interannual variation in the control period. For model i , the parameters β_i and γ_i are additive and b_i multiplicative biases, respectively. $\Delta\mu$ is the change in mean temperature, $\Delta\gamma$ is the change in trend, and q is the change in interannual variability between the control and the scenario period.

Finally, the means and variances of the model outputs in the scenario period are

$$\begin{aligned} E(Y_{i,t}) &= \mu + \Delta\mu + \beta_i + \Delta\beta_i + \kappa(b_i - 1)\Delta\mu + \\ &[\gamma + \Delta\gamma + \gamma_i + \Delta\gamma_i + \kappa(b_i - 1)\Delta\gamma] \times (t - T_0), \\ \text{Var}(Y_{i,t}) &= \sigma^2 q^2 b_i^2 q_{b_i}^2 \end{aligned} \quad (4)$$

With the parameters $\Delta\beta_i$, $\Delta\gamma_i$, and q_{b_i} we allowed changes in the additive and multiplicative biases between the control and scenario periods. The additional bias change terms $\kappa(b_i - 1)\Delta\mu$ and $\kappa(b_i - 1)\Delta\gamma$ are not standard. The parameter κ is supposed to be between 0 and 1. If $\kappa = 0$, we have the ‘constant bias’ assumption of Buser et al. (2009): the biases β_i and γ_i cancel if we take the difference $Y_{i,t} - X_{i,t}$. Hence, up to additive bias changes, all models were assumed to predict the climate shift between control and scenario period correctly for $\kappa = 0$. This assumption is used implicitly in most climate change studies. The value $\kappa = 1$ corresponds to the ‘constant relation’ assumption of Buser et al. (2009), which says that if model i over- or underestimates the difference between a warm and a cold year in the control period by the factor b_i , then the mean climate change will be also over- or underestimated by the same factor. In other words, the expected value of the difference $Y_{i,t} - X_{i,t}$ is $b_i\Delta\mu + \Delta\beta_i + (b_i\Delta\gamma + \Delta\gamma_i)(t - T_0)$. A value of κ between 0 and 1 corresponds to a compromise between the ‘constant bias’ and the ‘constant relation’ assumptions: the over- or underestimation is attenuated by the factor κ .

We refer to Buser et al. (2009) for a more detailed discussion of the 2 assumptions ‘constant bias’ and ‘constant relation’ and their influence on the outcome of the prediction. Note that the ‘constant relation’ assumption implies that the models with larger interannual variability should also predict larger climate changes whereas under the ‘constant bias’ assumption, no relation between these 2 quantities exists. We can

therefore look at the correlation coefficient between projected climate change and estimated interannual variability. Although an estimated correlation coefficient based on 7 data pairs is highly variable, we can combine the evidence from 32 such correlations (1 for each region and season). The Wilcoxon test gives strong evidence against a symmetric distribution around 0 (p value = 0.0063 for a one-sided alternative shifted towards positive values). Because the independence of the 32 correlations is doubtful, we should not interpret this p value as proof that the constant bias assumption is violated; rather, it is an indication that the constant relation assumption cannot be discarded. Since this is an exploratory analysis, we also computed the correlation coefficients with all 16 available GCM/RCM chains from the ENSEMBLES database, leading to a 1-sided p value of 0.0018 in the Wilcoxon test. Since a decision between the 2 assumptions is difficult, we find the approach in Eq. (4) with a continuum of intermediate assumptions represented by the parameter κ attractive.

The presence of systematic bias changes in climate simulations was also detected by Christensen et al. (2008). They investigated the relationship between monthly biases of RCMs in reanalysis-driven simulations and found that temperature biases in southern European regions increase with temperature, thus rejecting the hypothesis of climate-state-independent biases. Their result is qualitatively consistent with our analysis, although it merely considers the RCM component, rather than the whole GCM/RCM model chain.

The introduction of bias changes leads to identifiability problems since the observations $Y_{0,t}$ are unknown. The climate shift and bias changes are confounded. Only the sums

$$\Delta\mu + \Delta\beta_i + \kappa(b_i - 1)\Delta\mu, \Delta\gamma + \Delta\gamma_i + \kappa(b_i - 1)\Delta\gamma \quad (5)$$

and the products qq_{b_i} are determined. A large value of $\Delta\mu$ can be compensated by opposite model bias changes $\Delta\beta_i$ for each model, and the same is true for q and q_{b_i} . To solve this problem, we restricted the values of $\Delta\beta_i$ to be small and the values q_{b_i} to be close to 1. In the Bayesian approach, this can be done by choosing informative prior distributions for these parameters. With the introduction of the parameter κ , an additional identifiability problem arises if all b_i are similar, because then larger values for $\Delta\mu$ can be compensated with small values for κ . In general, if the multiplicative biases b_i are different, this problem disappears.

3.3. Prior distributions

We assumed that all parameters are *a priori* independent so that only the marginal prior distributions are

needed. There are 2 classes of parameters: the first class consists of μ , $\Delta\mu$, β_i , $\Delta\beta_i$, γ , $\Delta\gamma$, γ_i and $\Delta\gamma_i$, which are related to the mean values of the assumed normal distributions of the data. It is common to take normal priors for these parameters since this simplifies the computations. The other class of parameters consists of σ^2 , q^2 , b_i^2 and $q_{b_i}^2$, which are variances or multiplicative changes of the variances. It is a common procedure (Gelman et al. 2003) to work with the precision, which is defined as the inverse of the variance, and to choose a gamma distribution for the prior of the precision. The same procedure is used for the multiplicative change factors.

In order to fully specify the priors, we had to choose the so-called hyperparameters, i.e. the mean and variance in case of a normal prior and the shape and scale in case of a gamma prior. In Table 3, the prior distributions and their hyperparameters are shown for all parameters.

For the parameters μ , $\Delta\mu$, β_i , γ , $\Delta\gamma$, γ_i , σ^{-2} , q^{-2} and b_i^{-2} , the hyperparameters are such that the priors are flat and thus carry little information. For the parameters $\Delta\beta_i$, $\Delta\gamma_i$ and $q_{b_i}^{-2}$, we took informative priors with small variances that are concentrated around 0 and 1, respectively, to solve the identifiability problem. The sensitivity of the results to the choice of the prior distributions and the hyperparameters was studied by Buser et al. (2009). The choice of the informative priors influences the results; for instance, the less informative the prior for $\Delta\beta_i$ is, the wider the posterior distribution of $\Delta\mu$ becomes. We defend our choice with the following

arguments: if we use the constraint $\sum_i \Delta\beta_i = 0$, we can estimate the parameter $\Delta\beta_i$ without any prior by a method of moments (see Buser et al. 2010). Since under this constraint $\sum_i \Delta\beta_i^2$ is minimal, a prior variance that is much smaller than the empirical variance of these estimates would be in conflict with the data. On the other hand, we believe it is reasonable to assume *a priori* that the bias change is smaller than both the bias and the climate change.

A special role is played by the additional parameter κ , that in essence selects a weighted average between the ‘constant bias’ assumption ($\kappa = 0$) and the ‘constant relation’ assumption ($\kappa = 1$). In a fully Bayesian framework, we assign a prior distribution to it, in our case the uniform distribution on $[0, 1]$, and compute the posterior distribution of κ and the other parameters. If the posterior for κ is concentrated on values close to 0, it means that the data favor the ‘constant bias’ assumption. Similarly, a posterior concentrated near 1 indicates that the ‘constant relation’ is more plausible given the information in the data. Finally, a posterior distribution that is close to uniform indicates that the data alone do not allow a decision between the 2 assumptions.

If κ is considered unknown and treated in a fully Bayesian approach, it also affects the posterior distributions of the other parameters, e.g. the climate change $\Delta\mu$: The posterior of $\Delta\mu$ for any fixed value of κ is averaged with respect to the posterior for κ :

$$p(\Delta\mu | \text{data}) = \int_0^1 p(\Delta\mu | \text{data}, \kappa) p(\kappa | \text{data}) d\kappa \quad (6)$$

Table 3. Hyperparameters for the prior distributions. For normal distributions, the hyperparameters are the expectation (μ_0) and the variance (σ_0^2). For gamma distributions, the first hyperparameter is the shape parameter and the second hyperparameter is the rate. The mean is shape \times rate $^{-1}$, the variance is shape \times rate $^{-2}$. The 95 % confidence intervals show the range of values that were *a priori* decided to be physically plausible. The intervals are quite large for the non-informative priors. For the 3 parameters $\Delta\beta_i$, $\Delta\gamma_i$ and $q_{b_i}^{-2}$, we chose informative priors to solve the identifiability problem. Their confidence intervals are therefore smaller

| Parameter | Unit | Distribution | Hyper-parameter 1 (μ_0 , shape) | Hyper-parameter 2 (σ_0^2 , rate) | 95 % confidence interval |
|------------------|----------------------------|--------------|---|---|--|
| μ | $^{\circ}\text{C}$ | Normal | 0 (Winter) 15 (Summer) 10 (Spring) 10 (Autumn) | 25 | -9.8, 9.8 5.2, 24.8 0.2, 19.8 0.2, 19.8 |
| $\Delta\mu$ | $^{\circ}\text{C}$ | Normal | 0 | 16 | -7.8, 7.8 |
| β_i | $^{\circ}\text{C}$ | Normal | 0 | 16 | -7.8, 7.8 |
| $\Delta\beta_i$ | $^{\circ}\text{C}$ | Normal | 0 | 0.5 | -1.4, 1.4 |
| γ | $^{\circ}\text{C yr}^{-1}$ | Normal | 0 | 0.1 | -0.6, 0.6 |
| $\Delta\gamma$ | $^{\circ}\text{C yr}^{-1}$ | Normal | 0 | 0.1 | -0.6, 0.6 |
| γ_i | $^{\circ}\text{C yr}^{-1}$ | Normal | 0 | 0.1 | -0.6, 0.6 |
| $\Delta\gamma_i$ | $^{\circ}\text{C yr}^{-1}$ | Normal | 0 | 0.0005 | -0.045, 0.045 |
| σ^{-2} | $^{\circ}\text{C}^{-2}$ | Gamma | 0.1 | 0.1 | 0, 9.8 |
| q^{-2} | | Gamma | 0.1 | 0.1 | 0, 9.8 |
| b_i^{-2} | | Gamma | 0.1 | 0.1 | 0, 9.8 |
| $q_{b_i}^{-2}$ | | Gamma | 3 | 3 | 0.2, 2.4 |

(and the same formula holds for any other parameter). Hence the effect of the extended approach with an unknown κ is a widening of the posterior distribution, when projections using the 2 bias assumptions yield differing results and the data cannot decide between the 2 assumptions.

However, assuming prior independence of the value of κ for different seasons and regions can be criticized, since values of κ should be similar in consecutive seasons and neighboring regions. On the other hand, assuming that the value of κ is the same for all seasons and regions would presumably not be adequate either. Moreover, when testing different approaches we found that the posterior of κ is concentrated at values near 1 if a single model projects a substantially higher climate change, see Section 3.4 of Buser (2009). This is due to the fact that the ‘constant relation’ assumption can explain such an outlier more easily by increasing the value of b_i . For these reasons, we also show results for a more conservative approach where the posterior for other parameters is averaged with respect to the uniform distribution for κ , i.e. the prior instead of the posterior. This leads to an even bigger uncertainty when projections using the 2 bias assumptions yield differing results.

4. RESULTS

4.1. Weighting of bias assumptions

We first compared the 2 different ways of handling the parameter κ , as explained in the previous section. We show the posterior distributions of the climate shift $\Delta\mu$ for the Alps in Fig. 1 and for Scandinavia in Fig. 2 since in these 2 regions the differences between the 2 assumptions are most pronounced. In the first 2 rows of these figures, the posterior distributions of $\Delta\mu$ are given for the ‘constant bias’ (gray dashed line) and the ‘constant relation’ (light gray dotted line) assumption. In the first row, the solid black line is the uniform average of the posterior densities of $\Delta\mu$ for fixed κ (we took a discrete average over $\kappa = 0, 0.1, 0.2, \dots, 0.9, 1$), whereas in the second row the solid black line is the posterior distribution of $\Delta\mu$ in the fully Bayesian approach. In the second row we therefore have a weighted average with weights given by the posterior of κ , which is shown in the last row of the 2 figures, along with the uniform distribution which is also used as a prior. Because this posterior is rather close to the prior, the black curves in the 2 rows are very similar. The posterior is in particular not sufficiently concentrated near 0 or near 1 to

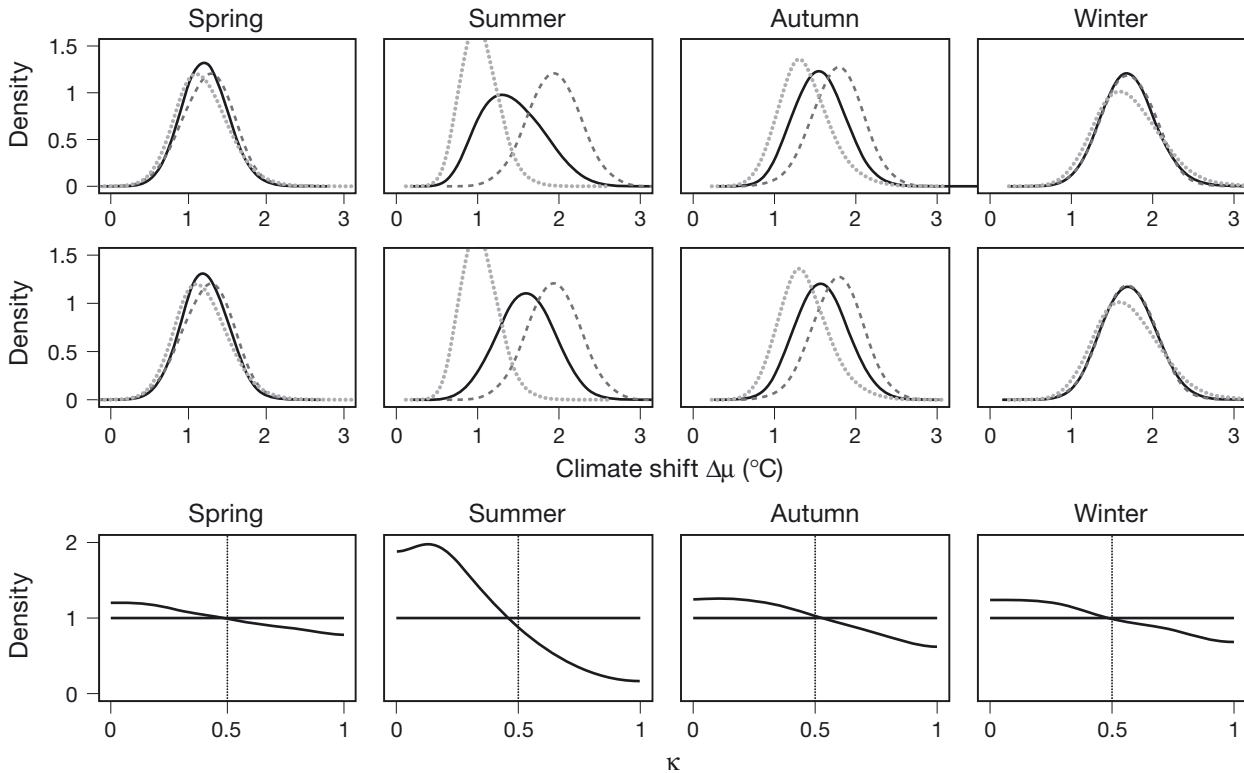


Fig. 1. Posterior distributions of the climate shift $\Delta\mu$ between 1975 and 2035 for the Alps (Al) for κ uniformly distributed (first row, black solid line), and for κ treated as a parameter in a fully Bayesian approach (second row, black solid line). For comparison, the ‘constant bias’ ($\kappa = 0$, gray dashed line) and the ‘constant relation’ ($\kappa = 1$, light gray dotted line) assumption are given. The last row shows the posterior distribution of κ compared to the uniform distribution

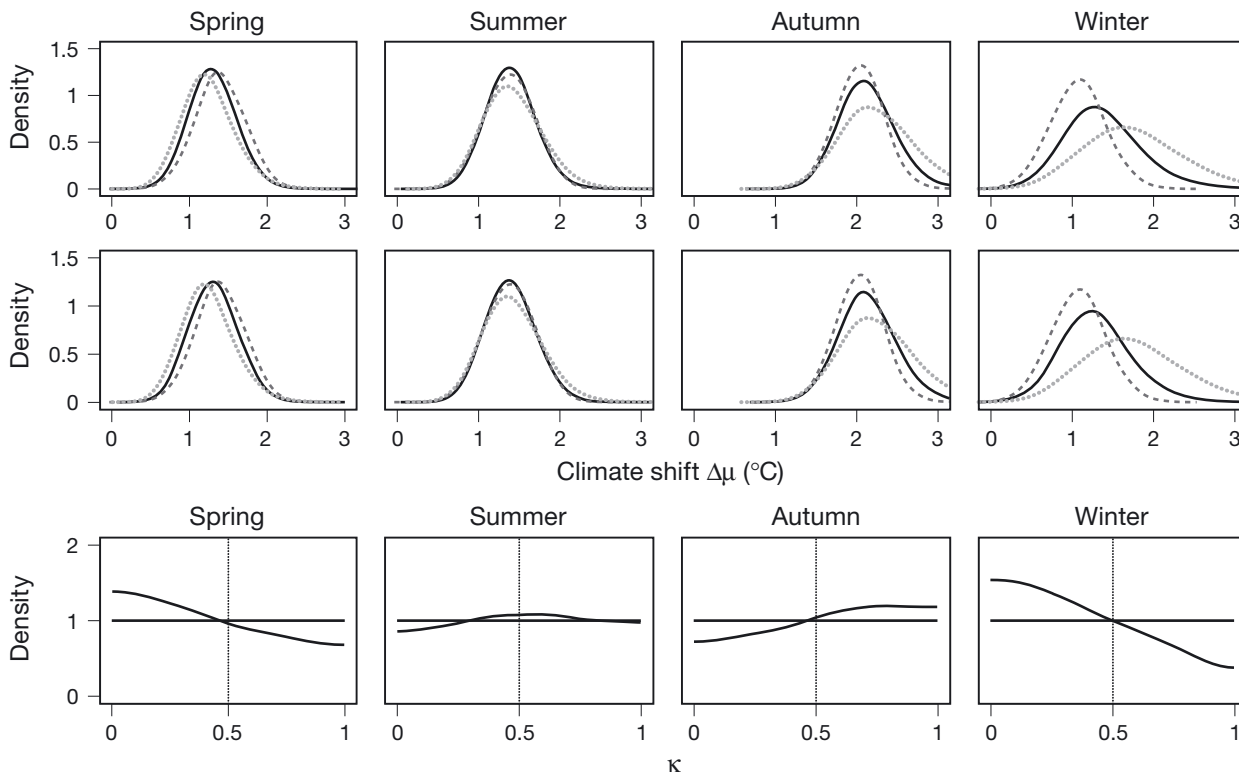


Fig. 2. Posterior distributions of the climate shift between 1975 and 20335 for Scandinavia (Sc). Other details in Fig. 1

clearly favor the 'constant bias' or the 'constant relation' assumption. Since the same is also true for the other regions, in the next subsection we only show the results of the fully Bayesian approach.

4.2. Results for all PRUDENCE domains

Fig. 3 shows the posterior distribution of the climate shift $\Delta\mu$ in a fully Bayesian analysis (black solid line) for all seasons and all regions. For comparison, the posterior distributions for the 2 extreme cases 'constant bias' ($\kappa = 0$) and 'constant relation' ($\kappa = 1$) are given as gray dashed and light gray dotted lines, respectively. Results are calculated for each season and each region separately.

For seasons and regions with a good agreement between the 'constant bias' and 'constant relation' assumptions (e.g. spring in France), also the approach with unknown κ resulted in a similar prediction. In such a situation, the interpretation is easy.

On the other hand, there were several regions and seasons where the posteriors under the 'constant bias' and 'constant relation' assumptions disagreed. Hence the finding of Buser et al. (2009) is not limited to the summer in the Alpine region, although the difference was largest there. The differences between the light

gray and the gray densities were most often a combination of different centers (as measured by the mean, median, or mode) and of different spreads (e.g. as measured by the standard deviation). Note that a larger spread in the posterior means a larger uncertainty about the climate shift. The black curve is in general a good compromise between the light gray and the gray curves, sometimes with a slight preference for the gray curve. When the difference between the 2 curves is mainly a shift, the black curve is wider than both the gray and the light gray curves, representing an increased uncertainty, whereas in other cases it also represents a compromise in the uncertainty.

The most pronounced differences between the 2 bias assumptions occurred as follows. In summer, the 'constant relation' assumption (light gray dotted lines) predicted a significantly weaker warming in the Alps (Al), eastern Europe (EE), and the Mediterranean (Md), and to a lesser extent also in France (Fr) and middle Europe (ME). The same pattern, but with weaker differences, was also evident for autumn. In contrast, in winter the 'constant relation' assumption predicted a significantly stronger warming in Scandinavia (Sc) and France, and to a lesser extent also in eastern and middle Europe. In spring, the differences were smallest in all regions. The largest difference in the posterior median warming of the 'constant bias' and 'constant relation' as-

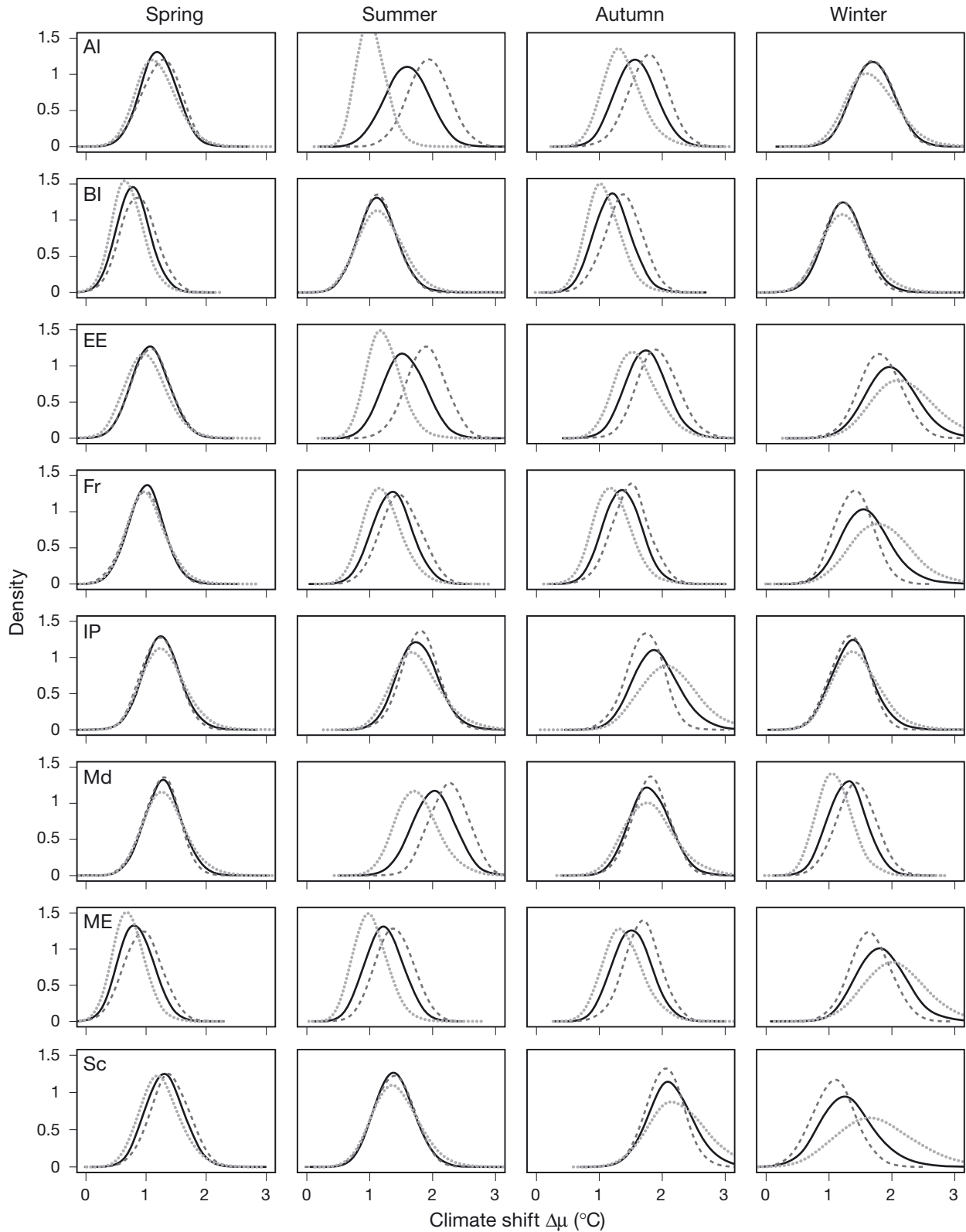


Fig. 3. Posterior distributions of mean warming $\Delta\mu$ between 1975 and 2035. Results are given for each region and season separately. Results for a fully Bayesian analysis with unknown κ are shown as a solid black line; the 'constant bias' ($\kappa = 0$) and 'constant relation' ($\kappa = 1$) assumptions are given as gray dashed and light gray dotted lines, respectively. Note the different projections resulting from the different bias assumptions, and how the full Bayesian analysis yields a suitable compromise between the 2 assumptions. See Table 2 for abbreviations

sumptions occurred in the Alpine summer (+1.9°C versus +1.0°C), while the biggest difference in the posterior 90 % quantile occurred for the Scandinavian winter (+2.5°C versus +3.5°C). These large differences demonstrate that bias assumptions matter.

The differences between the 2 assumptions can be traced back to a considerable over- or underestimation of the observed interannual variability during the control period. For the Scandinavian winter, all GCM/RCM chains underestimated the interannual variability during the control period. The ‘constant relation’ assumption as well as the generalization introduced here thus yielded a significantly stronger warming than the multi-model mean. In summer, the opposite was true for most regions. Because the majority of models overestimated the interannual variability, our method projected a weaker warming than the multi-model mean. This shows the necessity of a good representation of the true interannual variability by the models. For a detailed discussion of the influence of an inaccurate estimate of the interannual variability on the projected mean warming, we refer to Buser et al. (2009).

In Fig. 4, we show the analogue figure to Fig. 3 for q , the expected change of the interannual variability. There is a reduction of the interannual variability for most regions and seasons. This is in contrast to other publications that predicted an increase of the interannual variability (e.g. Schär et al. 2004, Vidale et al. 2007). However, these other studies addressed a later period (2071–2100), used a stronger greenhouse gas scenario (SRES A2 rather than A1B), and were based on a different simulation strategy than that used in the PRUDENCE project. We note that some of the RCMs in our analysis predicted an increase of the interannual variability, but this was not supported by the multi-model majority. For the later scenario period 2071–2100, a larger fraction of available ENSEMBLES models predicts some increase of the summer interannual variability in continental Europe, but the signal is not as pronounced as in the previous PRUDENCE scenarios. Also, most ENSEMBLES models show a significant increase in total daily temperature variability (Fischer & Schär 2010).

Further work is needed to reconcile the PRUDENCE and ENSEMBLES results. One potentially important factor relates to the different simulation strategies for the 2 sets of scenarios. The previous PRUDENCE scenarios were based on large-scale sea surface temperature (SST) distributions that prescribed the SST variability to be identical in control and scenario periods (see Christensen & Christensen 2007). In contrast, the ENSEMBLES scenarios are based on transient coupled GCM simulations, driven by observed forcings for the past and greenhouse gas and aerosol scenarios for the future. There is some intrinsic difficulty with the latter

approach. If the past forcing includes decadal and interannual variations in volcanic, solar, and aerosol forcing, this simplifies the comparison against observed climate variations. On the other hand, if the future forcing uses smooth volcanic, solar, and anthropogenic aerosol forcings, this leads to an artificial reduction of variability between the control and the scenario periods. As we used the period 1961–1990 as the control, this artificial reduction might be particularly important, as this period coincides with unusually strong volcanic activity.

In Fig. 5, the posterior distributions for the scenario trend $\gamma + \Delta\gamma$ are given. These distributions show that a substantial trend over the scenario period is likely. For the trend, the differences between the ‘constant bias’ and ‘constant relation’ assumption were so small that we could neglect them.

An important measure for a user is the posterior predictive density. It estimates the distribution of the scenario seasonal temperatures given all available data. It accounts for interannual variability and for all uncertainties considered, including those due to the bias assumptions. In Table 4, the 10, 50, and 90 % quantiles of the posterior predictive distribution (with Bayesian estimates of κ) are calculated for each region and season. For each region, the top row presents the probability distribution of mean seasonal temperatures, in terms of the 10, 50, and 90 % quantiles. The second row presents the distribution of the seasonal temperatures (including interannual variability) in the same format, with the 10 and 90 % quantiles representing the 1 in 10 yr seasonal mean temperature (hot and cold) for the scenario climate. Since the trends are not included, these distributions represent snapshots for 1975 and 2035, respectively. When comparing the control and scenario period, we note that for some regions (Alps, Mediterranean, and Spain), a cold summer (10 % quantile) of the scenario period 2021–2050 corresponds to a hot summer (90 % quantile) of the control climate (1961–1990).

5. CONCLUSIONS AND OUTLOOK

We generalized the Bayesian methodology of Buser et al. (2009) by combining the 2 different but equally plausible bias assumptions. The main conclusions on the methodological level are:

- The generalized Bayes methodology offers a single probabilistic projection for seasonal temperature distributions that includes the uncertainty of the different GCM/RCM chains, but also the uncertainty about the extrapolation of the biases into the scenario period.
- The posterior for κ tells us how well we can decide between the ‘constant bias’ and the ‘constant relation’

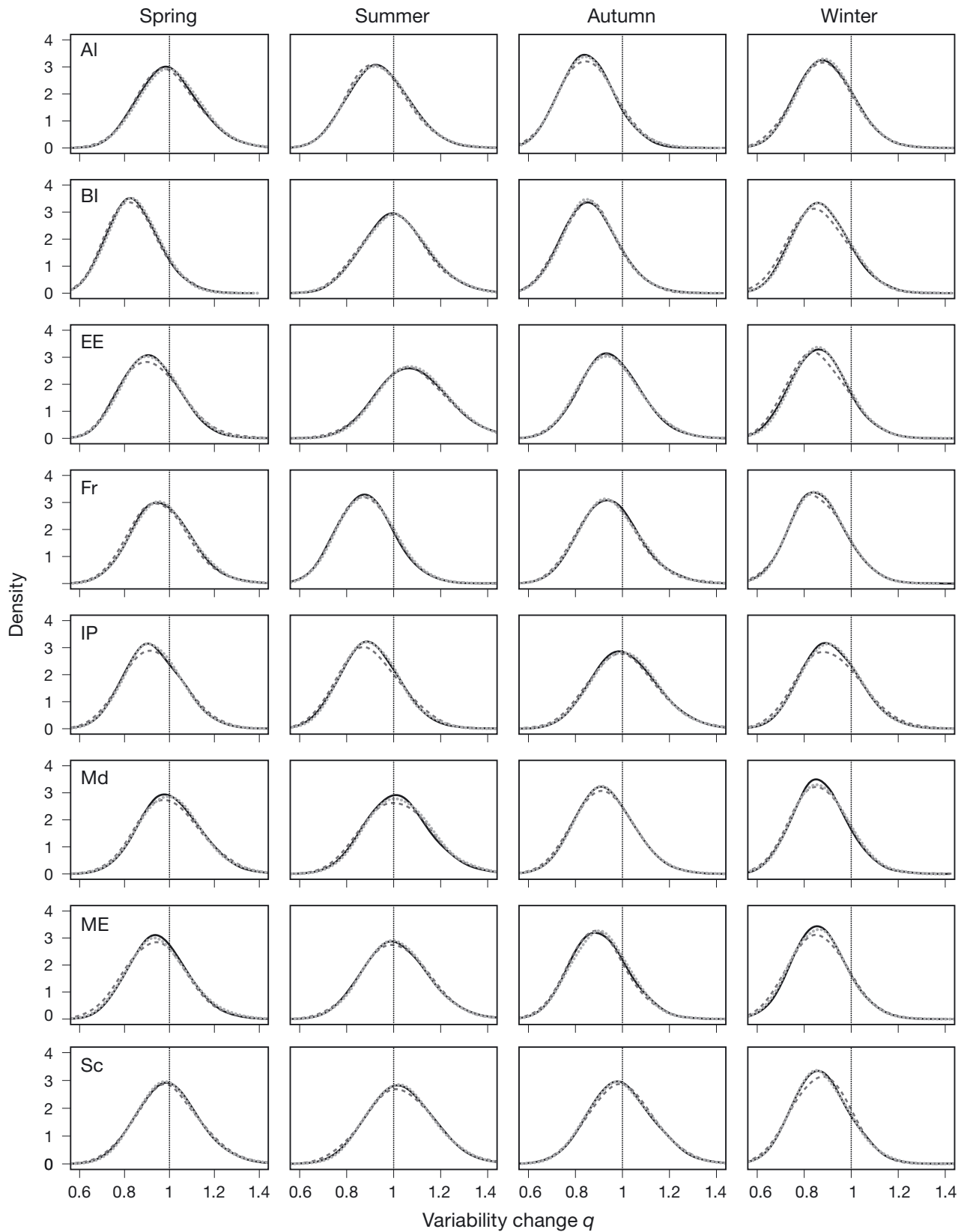


Fig. 4. Posterior distributions of the change of the interannual variability q between 1975 and 2035; other details as in Fig. 3. The different bias assumptions yield virtually identical posterior distributions. For the majority of seasons and regions, a reduction of the interannual variability is projected

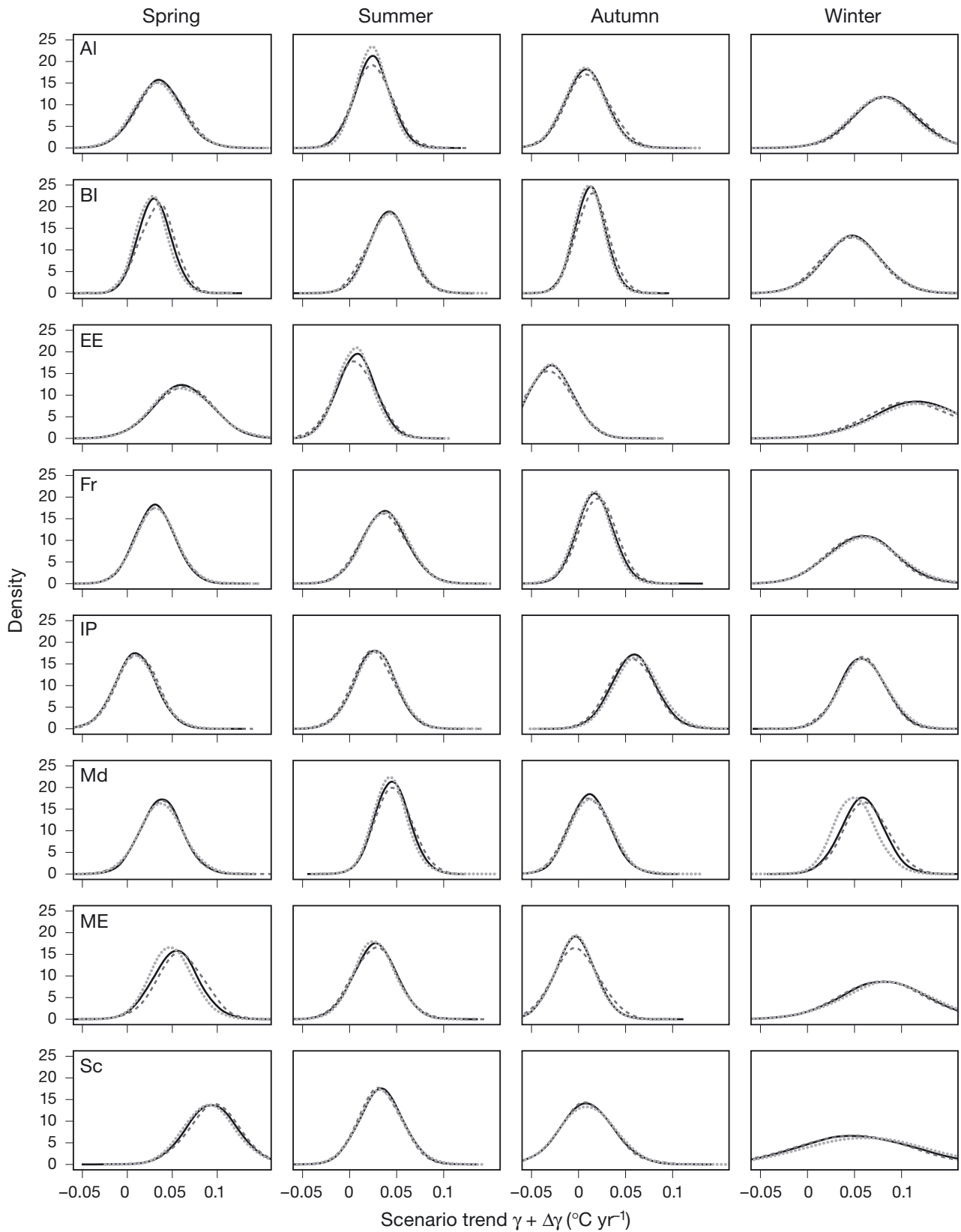


Fig. 5. Posterior distributions of the scenario trend $\gamma + \Delta\gamma$ between 1975 and 2035; other details as in Fig. 3. The differences between the 'constant bias' and 'constant relation' assumption are negligible

Table 4. Estimates of the 10, 50 and 90 % quantiles for the 2035 scenario climate. Scenario mean: posterior quantiles of the scenario climate mean $\mu + \Delta\mu$; scenario temperature: quantiles of posterior predictive distributions of seasonal temperatures, which also includes the interannual variability of the scenario climate; control: quantiles of the posterior predictive distribution of the control climate in 1975. See Table 2 for abbreviations

| Region | Spring | | | Summer | | | Autumn | | | Winter | | |
|----------------------|--------|------|------|--------|------|------|--------|------|------|--------|------|------|
| | 10% | 50% | 90% | 10% | 50% | 90% | 10 % | 50% | 90% | 10% | 50% | 90% |
| AI | | | | | | | | | | | | |
| Scenario mean | 7.5 | 8 | 8.4 | 17 | 17.5 | 17.9 | 9.8 | 10.2 | 10.7 | 0.8 | 1.2 | 1.8 |
| Scenario temperature | 6.7 | 8 | 9.3 | 16.6 | 17.5 | 18.4 | 9.3 | 10.2 | 11.3 | -0.4 | 1.2 | 2.9 |
| Control | 5.5 | 6.8 | 8 | 15 | 15.9 | 16.7 | 7.6 | 8.7 | 9.7 | -2.2 | -0.4 | 1.3 |
| BI | | | | | | | | | | | | |
| Scenario mean | 7.8 | 8.1 | 8.5 | 14.7 | 15.1 | 15.5 | 10.1 | 10.5 | 10.9 | 4.4 | 4.9 | 5.4 |
| Scenario temperature | 7.3 | 8.1 | 8.9 | 14 | 15.1 | 16.1 | 9.8 | 10.5 | 11.3 | 3.5 | 4.9 | 6.4 |
| Control | 6.5 | 7.3 | 8.2 | 13 | 13.9 | 14.9 | 8.6 | 9.3 | 10 | 2.1 | 3.7 | 5.2 |
| EE | | | | | | | | | | | | |
| Scenario mean | 8.8 | 9.3 | 9.8 | 19 | 19.4 | 19.9 | 10.2 | 10.6 | 11.1 | -0.9 | -0.3 | 0.4 |
| Scenario temperature | 7.8 | 9.3 | 10.9 | 18.4 | 19.4 | 20.5 | 9.4 | 10.6 | 11.8 | -2.6 | -0.3 | 2.1 |
| Control | 6.6 | 8.2 | 9.9 | 17 | 17.9 | 18.8 | 7.7 | 8.9 | 10 | -4.9 | -2.3 | 0.4 |
| Fr | | | | | | | | | | | | |
| Scenario mean | 10.2 | 10.6 | 11 | 18.5 | 18.9 | 19.4 | 12.6 | 13 | 13.4 | 5.2 | 5.8 | 6.4 |
| Scenario temperature | 9.5 | 10.6 | 11.7 | 17.8 | 18.9 | 20 | 12 | 12.9 | 13.9 | 4.1 | 5.8 | 7.5 |
| Control | 8.5 | 9.6 | 10.7 | 16.4 | 17.6 | 18.7 | 10.7 | 11.6 | 12.5 | 2.3 | 4.2 | 6.1 |
| IP | | | | | | | | | | | | |
| Scenario mean | 12.1 | 12.6 | 13 | 21.9 | 22.4 | 22.8 | 15.6 | 16.1 | 16.6 | 7.4 | 7.8 | 8.3 |
| Scenario temperature | 11.5 | 12.6 | 13.7 | 21.3 | 22.3 | 23.4 | 14.9 | 16.1 | 17.3 | 6.7 | 7.8 | 9 |
| Control | 10.2 | 11.3 | 12.4 | 19.5 | 20.6 | 21.7 | 13 | 14.2 | 15.3 | 5.3 | 6.4 | 7.6 |
| Md | | | | | | | | | | | | |
| Scenario mean | 11.9 | 12.3 | 12.8 | 22.2 | 22.7 | 23.1 | 15 | 15.4 | 15.8 | 5.3 | 5.7 | 6.1 |
| Scenario temperature | 11.1 | 12.3 | 13.5 | 21.7 | 22.6 | 23.6 | 14.3 | 15.4 | 16.5 | 4.7 | 5.7 | 6.8 |
| Control | 9.9 | 11 | 12.2 | 19.8 | 20.6 | 21.5 | 12.5 | 13.6 | 14.6 | 3.3 | 4.4 | 5.5 |
| ME | | | | | | | | | | | | |
| Scenario mean | 8.5 | 8.9 | 9.3 | 17.4 | 17.8 | 18.2 | 10.4 | 10.8 | 11.3 | 2 | 2.6 | 3.3 |
| Scenario temperature | 7.7 | 8.9 | 10.1 | 16.7 | 17.8 | 19 | 9.9 | 10.8 | 11.9 | 0.4 | 2.6 | 4.9 |
| Control | 6.9 | 8.1 | 9.3 | 15.5 | 16.6 | 17.7 | 8.3 | 9.3 | 10.4 | -1.8 | 0.8 | 3.3 |
| Sc | | | | | | | | | | | | |
| Scenario mean | 1.8 | 2.3 | 2.8 | 13.8 | 14.2 | 14.7 | 4.3 | 4.8 | 5.4 | -7 | -6.3 | -0.4 |
| Scenario temperature | 0.8 | 2.3 | 3.8 | 13.1 | 14.2 | 15.5 | 3.4 | 4.8 | 6.3 | -9.4 | -6.3 | -3 |
| Control | -0.4 | 1 | 2.4 | 11.8 | 12.9 | 14 | 1.3 | 2.7 | 4.1 | -12.1 | -8.6 | -5 |

bias assumptions, based on the fit of the GCMs/RCMs to the observational data and to each other. However, because artefacts in the data may favor the 'constant relation assumption', some care is necessary. Because of this, we also investigated, in addition to a fully Bayesian approach, a more conservative approach which considers all values of κ as equally likely, so that the 2 bias assumptions obtain the same weights. In our application, the posteriors for κ were flat and not very different from the prior. This means that we could not decide between the 2 assumptions, and the 2 approaches gave very similar results.

- The methodology does not make any *a priori* assumptions about climate change. The respective prior distribution is uninformative. Prior assumptions about the bias changes are necessary to solve an identifiability problem which occurs in any kind of climate projection. This approach may be criticized, but accounting for bias changes using informative priors is certainly

more conservative than assuming that biases do not change between scenario and control period, as is implicitly done in most climate change studies.

- Estimating bias changes to correct scenario runs of climate models may be seen as an alternative to the use of model weights in the evaluation of multi-model ensembles. Our approach does not use any explicit weights, but it may implicitly assign large bias changes to certain models, implying that the respective projections are not taken at face value. More specifically, this occurs if models are in strong disagreement with the majority of the other models, or if they exhibit large biases (in mean and/or variance) during the control period. In essence, such models are implicitly down-weighted in the determination of the multi-model Bayesian probabilistic projection.

Application of the Bayesian methodology to temperature scenarios of 7 ENSEMBLES GCM/RCM chains for 8 European regions demonstrated that the afore-

mentioned issues are not academic, but relevant to climate change projections. In particular, we found several regions and seasons where assumptions about bias changes were crucial. In comparison to the multi-model mean, the generalized Bayes method projected a considerably weaker warming during summer and autumn in much of continental Europe, a stronger winter warming in Scandinavia, France, eastern and central Europe, and a weaker warming in both summer and winter in the Mediterranean.

The generalized Bayesian methodology is able to deal with such discrepancies. If the 2 bias assumptions disagree, the methodology responds with a widening of the probability distribution, thereby forwarding the evident uncertainty to the user of the projections.

The current methodology could be generalized and made more versatile in a number of respects. This includes the use of multivariate approaches to yield joint precipitation and temperature scenarios (see Buser et al. 2010), the consideration of a hierarchical model (distinguishing biases at the GCM and RCM levels), and the application of the methodology at higher spatial resolution and to other multi-model data sets.

Acknowledgements. We acknowledge the ENSEMBLES project, funded by the European Commission's 6th Framework Programme through contract GOCE-CT-2003-505539. We are indebted to J. H. Christensen and O. B. Christensen for maintaining the ENSEMBLES RCM data base (<http://ensemblesrt3.dmi.dk/>). We also acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project (<http://eca.knmi.nl>). The ETH climate simulations were conducted on the computing facilities of the Swiss Center for Scientific Computing (CSCS), with support from the Swiss Ministry for Education and Research and from the Swiss National Science Foundation (NCCR Climate). We also acknowledge very useful comments from A. Fischer and A. Weigel and thank 3 anonymous reviewers for careful reading and helpful suggestions.

LITERATURE CITED

- Buser CM (2009) Bayesian statistical methods for the analysis of multi-model climate predictions. PhD dissertation, Eidgenössische Technische Hochschule, Zürich
- Buser C, Künsch HR, Lüthi D, Wild M, Schär C (2009) Bayesian multi-model projection of climate: bias assumptions and interannual variability. *Clim Dyn* 33:849–868
- Buser C, Künsch HR, Weber A (2010) Biases and uncertainty in climate projections. *Scand J Stat* 37:179–199
- Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Clim Change* 81:7–30
- Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (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
- Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb MJ (2006) Towards quantifying uncertainty in transient climate change. *Clim Dyn* 27:127–147
- Déqué M (2009) Temperature and precipitation probability density functions in ENSEMBLES regional scenarios. ENSEMBLES Tech Rep 5. Available at <http://ensembles-eu.metoffice.com>
- Fischer EM, Schär C (2010) High-impact European heatwaves in a changing climate. *Nat Geosci* 3:398–403
- Furevik T, Bentsen M, Drange H, Kindem IKT, Kvamsto NG, Sorteberg A (2003) Description and evaluation of the Bergen climate model: ARPEGE coupled with MICOM. *Clim Dyn* 21:27–51
- Furrer R, Sain SR, Nychka D, Meehl GA (2007) Multivariate Bayesian analysis of atmosphere–ocean general circulation models. *Environ Ecol Stat* 14:249–266
- Gelman A, Carlin JB, Stern HS, Rubin DB (2003) Bayesian data analysis, 2nd edn. Chapman & Hall/CRC, Boca Raton, FL
- Gibelin AL, Déqué M (2003) Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Clim Dyn* 20:327–339
- Gilks WR, Richardson S, Spiegelhalter DJ (1996) Markov Chain Monte Carlo in practice. Chapman and Hall, London
- Giorgi F, Mearns LO (2002) Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the 'reliability ensemble averaging' (REA) method. *J Clim* 15:1141–1158
- Haylock MR, Hofstra N, Tank AMGK, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation. *J Geophys Res* 113:D20119 doi:10.1029/2008JD010201
- Hewitt CD, Griggs DJ (2004) Ensembles-based predictions of climate changes and their impacts (ENSEMBLES). *Eos Trans Am Geophys Union* 85:556
- Jaeger EB, Anders I, Lüthi D, Rockel B, Schär C, Seneviratne SI (2008) Analysis of ERA40-driven CLM simulations for Europe. *Meteorol Z* 17:349–368
- Jun M, Knutti R, Nychka D (2008) Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *J Am Stat Assoc* 103:934–947
- Kjellström E, Barring L, Gollvik S, Hansson U and others (2005) A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3). SMHI Rep Meteorol Climatol 108. SMHI, Norrköping
- Lenderink G, van den Hurk B, van Meijgaard E, van Ulden AP, Cuijpers JH (2003) Simulation of present-day climate in RACMO2: first results and model developments. KNMI Tech Rep 252, KNMI, Amsterdam
- Meehl GA, Stocker TF, Collins WD, Friedlingstein P and others (2007) Global climate projections. In: *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge University Press, Cambridge, p 747–845
- Murphy JM, Booth BBB, Collins M, Harris GR, Sexton DMH, Webb MJ (2007) A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos Trans R Soc Lond A Math Phys Eng Sci* 365:1993–2028
- Nakicenovic N, Alcamo J, de Vries B, Fenhann J, Gaffin S, Gregory K (2000) Emission scenarios. A special report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Niehörster F, Fast I, Huebener H, Cubasch U (2008) The stream one ENSEMBLES projections of future climate change. ENSEMBLES Tech Rep 3. Available at <http://ensembles-eu.metoffice.com>
- Plummer D, Caya D, Cote H, Frigon A and others (2006) Climate and climate change over North America as simu-

- lated by the Canadian regional climate model. *J Clim* 19: 3112–3132
- Radu R, Déqué M, Somot S (2008) Spectral nudging in a spectral regional climate model. *Tellus Ser A Dyn Meteorol Oceanogr* 60:898–910
- Roeckner E, Bäuml G, Bonaventura L, Brokopf R and others (2003) The atmospheric general circulation model ECHAM 5. Part I: model description. Tech Rep 349, MPI, Hamburg
- Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger M, Appenzeller C (2004) The role of increasing temperature variability for European summer heat waves. *Nature* 427: 332–336
- Scinocca JF, McFarlane NA, Lazare M, Li J, Plummer D (2008) The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos Chem Phys* 8:7055–7074
- Smith RL, Tebaldi C, Nychka D, Mearns LO (2009) Bayesian modeling of uncertainty in ensembles of climate models. *J Am Stat Assoc* 104:97–116
- Tebaldi C, Knutti R (2007) The use of multi-model ensemble in probabilistic climate projections. *Philos Trans R Soc Lond A Math Phys Eng Sci* 365:2053–2075
- Tebaldi C, Sanso B (2008) Joint projections of temperature and precipitation change from multiple climate models: a hierarchical approach. *J R Stat Soc A* 172:83–106
- Tebaldi C, Smith RL, Nychka D, Mearns LO (2005) Quantifying uncertainty in projection of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. *J Clim* 18:1524–1540
- Vidale PL, Lüthi D, Wegmann R, Schär C (2007) European summer climate variability in a heterogeneous multi-model ensemble. *Clim Change* 81(Suppl 1):209–232

Submitted: November 2, 2009; Accepted: July 22, 2010

Proofs received from author(s): November 8, 2010