Internal variability and model uncertainty components in future hydrometeorological projections: the Alpine Durance basin

M. Lafaysse,[1,2] B. Hingray,[1,3] A. Mezghani,[1,3,4] J. Gailhard,[5] L. Terray,[6]

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

[2]Météo-France/CNRS, CNRM-GAME, CEN, 1441 rue de la Piscine, 38400 Saint Martin d'Hères, France

[3]CNRS, LTHE UMR 5564, Grenoble, F-38000, France

[4]now at Meteorologisk Institutt, PO BOX 43 Blindern, Norway

[5]EDF-DTG, 21 avenue de l'Europe, BP 41, 38040 Grenoble CEDEX 9, France

[6]CERFACS/Sciences de l'Univers, CERFACS-CNRS URA 1875, 42 avenue Gaspard Coriolis, 31057 Toulouse CEDEX 1, France

Abstract.

A multireplicate multimodel ensemble of hydrological simulations covering the 1860-2099 period has been produced for the Upper Durance River basin (French Alps). An original quasi-ergodic analysis of variance was applied to quantify uncertainties related to General Circulation Models (GCMs), Statistical Downscaling Models (SDMs) and the internal variability of each GCM/SDM simulation chain. For temperature, GCM uncertainty prevails and SDM uncertainty is non-negligible. Significant warming and in turn significant changes are predicted for evaporation, snow cover and seasonality of discharges. For precipitation, GCM and SDM uncertainty components are of the same order. A high contribution of the large and small scale components of internal variability is also obtained, inherited respectively from the GCMs and the different replicates of a given SDM. The same applies for annual discharge. The uncertainty in values that could be experienced for any given future period is therefore very high. For both discharge and precipitation, even the sign of future realizations is uncertain at a 90% confidence level. These findings have important implications. Similarly to GCM uncertainty, SDM uncertainty cannot be neglected. The same applies for both components of internal variability. Climate change impact studies based on a single SDM realization are likely to be no more relevant than those based on a single GCM run. They may lead to poor decisions for climate change adaptation.

1. Introduction

Long-term water resources management at local and regional scales must account for the high potential impact of global change on the hydrological cycle. Hydrological scenarios required for climate change impact studies are commonly obtained by simulation with Hydrological Models (HMs) from future meteorological scenarios. To allow for a relevant impact assessment, meteorological scenarios have to fulfill some constraints imposed by the strong non-linearity and the high spatial and temporal variability of hydrological processes (e.g. strong dependence of temperature, radiative fluxes, precipitation, etc. on elevation and aspect in mountainous environments). Therefore, the meteorological scenarios have to be unbiased (e.g. with respect to space and seasonality) and to have high spatial and temporal resolution. Because such requirements are not fulfilled by General Circulation Model (GCMs) outputs, meteorological scenarios are classically obtained with regional downscaling models. Regional Climate Models (RCMs) can be used to physically increase the resolution. However, they still often require a post-processing step to remove remaining biases and reach the required resolution [*Teutschbein and Seibert*, 2012]. Statistical Downscaling Models (SDMs) are also frequently used for this. They generate meteorological scenarios using statistical relationships between local-scale meteorological variables and some large scale atmospheric predictors [e.g. *Maraun et al.*, 2010]. Because they are not computationally expensive, SDMs are commonly used to downscale a large number of climate model outputs. Future hydrological scenarios are classically obtained with a sequence of separate models (e.g. GCM/SDM/HM) used to cover from larger scales to smaller scales and hydrological impacts, for any greenhouse gas and aerosols emission

scenario. As summarized by *Dobler et al.* [2012], very different sources of uncertainties are involved when such a simulation chain is applied. They include scenario uncertainty, model uncertainty and uncertainty due to the internal variability of the simulation chains.

Scenario uncertainty is related to the poorly known future of greenhouse gas and aerosols emissions due to the highly uncertain trajectory of the future socio-economic development of human societies. Model uncertainty is due to the limitations of the model structure and parameterization used to represent geophysical processes. Different models usually simulate different responses to the same forcing data. This clearly concerns GCMs, SDMs and HMs: GCMs exhibit different sensitivities to perturbations of atmospheric composition, SDMs present different local meteorological responses to large scale atmospheric fields and HMs produce different hydrological responses to local meteorological conditions. Additional uncertainty may also arise from interactions between the different models of the simulation chain [e.g. *Yip et al.*, 2011]. These different model uncertainty sources will be referred to hereafter as GCM uncertainty, SDM uncertainty, HM uncertainty and Model interaction uncertainty. Model uncertainty and scenario uncertainty are usually explored by multimodel experiments [e.g. *Dequé*, 2007; *Chen et al.*, 2011] and by the comparison of projections resulting from several scenarios [*IPCC*, 2007].

The internal variability of simulation chains is expected to represent the natural variability of regional climate at decadal or multi-decadal time scales. This variability has long been observed even in a stationary climate. 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]. For a given simulation chain, a part of internal variability can be attributed to the chaotic variability of the climate at large scales.

Mainly produced by the GCM itself, it can be estimated from the variability obtained from GCM experiments over a long time period for a stationary climate [*Raisanen*, 2001] or non-stationary climate [*Hawkins and Sutton*, 2009; *Hingray and Saïd*, 2014]. It can also be estimated using several runs of the GCM corresponding to different initial conditions of the simulation [*Sansom et al.*, 2013; *Yip et al.*, 2011]. This uncertainty source will be referred to here as the large scale internal variability. Very similar large scale atmospheric circulation patterns can then lead to very different meteorological observations at the local scale. This can be seen as a natural variability of the influence of large scale processes on local meteorology. This local scale component of internal variability is in principle mainly accounted for by the downscaling model. Due to nonlinearities in model physics and dynamics, RCMs are for instance sensitive to small perturbations of initial and/or lateral boundary conditions [e.g. *Lucas-Picher et al.*, 2008; *Sanchez-Gomez et al.*, 2009]. When a statistical downscaling model is used, it classically includes a stochastic process that produces several meteorological replicates for one given large scale pattern [*Buishand and Brandsma*, 2001; *Mezghani and Hingray*, 2009]. The local scale internal variability corresponds in this case to the dispersion between these replicates.

Although simulation chains (GCM+SDM+HM) have been applied in numerous studies worldwide for projections over a large range of spatial scales, the different sources of uncertainties have not been equally investigated in recent literature. Table 1 summarizes some recent works on basins influenced by snowmelt. The uncertainties associated with emission scenarios and with GCMs (and RCMs when used) have been accounted for in a number of studies by multimodel analyses. However, GCM uncertainty and large scale internal variability were generally not separated (except by *Chen et al.* [2011] or

by *Harding et al.* [2012]). In contrast with GCMs, the uncertainties related to SDMs are usually ignored. In table 1, only 6 studies out of 56 performed projections based on several SDMs although some recent studies suggest a large dispersion of projections due to the choice of downscaling method [*Coulibaly*, 2009; *Teutschbein et al.*, 2011; *Chen et al.*, 2011]. The uncertainty produced by errors of hydrological model is also rarely investigated.

For recent years, long time series are available for the large majority of GCM experiments. GCM experiments from the ENSEMBLES-Stream2 European research project cover for instance 240 years from the pre-industrial period (1860-2000) to the end of the XXIst century [*Johns et al.*, 2011]. Therefore, long time series of regional hydrometeorological variables can be obtained for multiple model experiments with appropriate GCM/SDM/MH chains. These ensembles of experiments provide a major opportunity to significantly improve the estimation of internal variability and model uncertainty components.

The purpose of this paper is to estimate and compare the role of four sources of uncertainty for hydrometeorological projections in a meso-scale alpine catchment: GCM uncertainty, SDM uncertainty and large and small scale components of internal variability. The analysis is based on a multireplicate multimodel ensemble of climate experiments produced as part of RIWER2030 research project [*Lafaysse*, 2011; *Hingray et al.*, 2013].

The study area and dataset are described in section 2. Models and climate experiments are presented in section 3. Evaluations are carried out for each model of the simulation chain and then for the whole simulation chain. Then, the dispersion of future projections obtained from this ensemble of experiments is presented for various meteorological and hydrological variables in section 4. Uncertainties related to emission scenarios and to

hydrological model errors will not be investigated here. Total uncertainty and its components for the RIWER2030 ensemble of projections are quantified and discussed as a function of projection lead time in section 5. Section 6 presents our conclusions.

2. Study area and dataset

Projections were developed for the Upper Durance river basin, a 3580 km^2 basin located in the southern French Alps. Its outlet is at Serre-Ponçon lake, which is regulated by a large dam operated for hydropower production by Electricité de France (EDF). The river discharges are mostly natural upstream of the lake. Rainfall and snowfall spatial variability is usually high inside the catchment, due to the complex topography, the sensitivity to oceanic disturbances in the western part and the influence of the Mediterranean sea in the eastern and southern parts. With elevations ranging from 700 to 4100 meters, the catchment produces highly seasonal flows: minimum and maximum discharges are observed in winter and spring, respectively, mainly due to snow accumulation and melt. Nevertheless, major floods can be observed in the fall due to thunderstorms producing large amounts of liquid precipitation.

Projections of the multireplicate multimodel ensemble are obtained from $N_s = 6$ multivariate SDMs forced by the outputs of $N_e = 11$ GCMs experiments [*Lafaysse*, 2011; *Hingray et al.*, 2013]. GCM experiments come from $N_g = 5$ GCMs out of which in 3 cases an ensemble of 3 runs is available. The GCMs, SDMs and corresponding references are listed in table 2 and 3. For each GCM/SDM couple, an ensemble of $N_k = 100$ stochastic generations is available, resulting from the stochastic generation process associated with each SDM (see section 3.2). Each meteorological scenario covers the whole period 1860-2099 and was used as input to the physically based ISBA-Durance hydrological model

developed for the Upper Durance River catchment [*Lafaysse et al.*, 2011]. This gives a total of $N_s * N_e * N_k = 6600$ times series of 240-year multivariate meteorological and hydrological scenarios.

3. Models and evaluations

The different models are presented and evaluated below. The evaluation was performed to test their ability to reproduce, over a reference period, a number of statistics for key climatic or hydrometeorological variables. The aim was to identify which models were suitable for the simulation of relevant future projections and if weights could be eventually assigned to each with respect to their relative skills.

3.1. General Circulation Models

3.1.1. GCM climate experiments

GCM experiments are outputs from the STREAM2 experiment of the ENSEMBLES European project [*Johns et al.*, 2011]. GCMs were run for the 1860-2000 historical period with observed anthropogenic forcings only (greenhouse gases, aerosols and ozone concentrations and land-use fraction) and the 2000-2099 future period with the SRES-A1B emission scenario [*Nakicenovic et al.*, 2001]. Solar and volcanic forcings were held constant. The variations in simulated time series therefore have only two possible causes: the anthropogenic forcings and the large scale internal variability of GCMs. Internal variability is obtained in climate experiments with constant greenhouse gases, aerosols and ozone concentrations, constant solar and/or volcanic forcing and when the GCM run is not nudged by observed sea surface temperatures. This configuration is roughly that of STREAM2 experiments for the 1860-1979 historical period, for which the evolution of

anthropogenic forcings is sufficiently small to be considered not influential on the variations of a number of climatic variables. This assumption seems to be reasonable as the simulated atmospheric circulations (described for instance by the occurrency of a number of weather types) do not present statistically significant trends over these 120 years [*Lafaysse*, 2011] and thus appear to be independent of the anthropogenic forcings over this period.

3.1.2. GCMs evaluation

GCMs are expected to simulate the main statistical properties of large-scale atmospheric conditions correctly over a long historical period. The evaluation can focus on the climatological mean state of some variables, their daily distribution, seasonal cycle, interannual variability, or long-term trends [*Gleckler et al.*, 2008; *Santer et al.*, 2008; *Errasti et al.*, 2011; *Brands et al.*, 2011]. In the context of a regional scale impact study, evaluation criteria are usually chosen to test the ability of GCMs to reproduce large scale variables suitable for downscaling. As the synoptic scale atmospheric circulations are known to be the primary factor of local meteorological variability at middle-latitudes, we tested GCM ability to reproduce the observed occurrence frequencies of a selection of weather types for a given reference period [e.g. *Demuzere et al.*, 2009; *Anagnostopoulou et al.*, 2009].

We use 4 weather types for each season (December-January-February (DJF), March-April- May (MAM), June-July-August (JJA), September-October-November (SON)). They were identified with the k-means clustering algorithm [*Michelangeli et al.*, 1995] applied to the 10 first Empirical Orthogonal Functions (EOF, 90% of the total variance) of the daily sea level pressure anomaly fields extracted from the ERA-40 reanalysis [*Uppala et al.*, 2005], for the spatial domain defined by *Boé* [2007]. The winter weather type

composites are given in figure 1. For other seasons, the same 4 weather patterns can be identified, with less intense gradients. The GCM daily sea level pressure anomaly fields are interpolated on the same grid and projected on the ERA-40 EOF, allowing us to assign a weather type to each simulation day (minimizing the Euclidian distance to each weather-type centroid in the EOF space).

figure 1

For several GCMs, simulated frequencies obtained over the 1960-1979 reference period are significantly different from those observed for the same period. However, even if the GCM were perfect, observations and simulations are not expected to be equal due to internal variability. In the present case, the occurrence frequencies of each weather type were therefore also computed for all other 20-year sub-periods of the 1860 to 1979 historical runs. Figure 2 illustrates that 1960-1979 observed frequencies are included in the range of simulated frequencies. This occurs here for nearly all seasons and weather types. The only significant error is an imbalance between the occurrences of summer weather types in the IPCM4 model. Differences identified between simulated frequencies and observations can therefore not be assigned to GCM errors. This also precludes the identification of better or poorer performing GCMs within our selection.

Figure 2

3.2. Statistical Downscaling Models

3.2.1. SDMs description

As summarized in table 3, different versions of three multivariate SDMs are used in this work. They all are variants of the K-nearest neighbors (K-nn) resampling approach widely used for the generation of daily weather scenarios [e.g. *Buishand and Brandsma*, 2001; *Gangopadhyay et al.*, 2005]. Analogs of a state vector for the current generation day are searched for on the basis of similarity criteria in the historical database. The state vector

used for this identification is based on daily atmospheric and/or surface weather variables, referred to as predictors. The needed surface variables observed for one of the K-nn are then used as a weather scenario for the generation of the simulation time step. In the following, predictors are estimated from the NCEP atmospheric reanalysis [*Kalnay et al.*, 1996] and predictants are surface variables from the SAFRAN meteorological reanalysis [*Durand et al.*, 2009], which provides temperature, precipitation amount and phase, relative humidity, wind speed and shortwave and longwave radiation for 23 mountain regions in the French Alps for elevation bands with a resolution of 300 meters and 7 aspect classes, spanning the period from 1959 to 2006. In the K-nn approach, the only critical limitation for the generation of weather scenarios is the availability of observation data in the archive. The type and the time and space resolutions of surface meteorological variables available in the archive actually determine the type and the time and space resolutions of the generated weather scenarios. Weather scenarios generated in the present work therefore share the same characteristics as those of SAFRAN. The 3 SDMs were calibrated for the 1981-2005 period. In all cases, analog days are identified in the calibration period, with a search restricted to the days belonging to the same season as that of the target day (moving full window of 2 months). For the 3 SDMs, the identification of a weather scenario for a given day results from a stochastic process (e.g. random selection of an analog among the K-nn days). Different series of daily weather scenarios can therefore be generated. In the following, 100 series are generated for each SDM to represent the small scale internal variability. These 100 replicates cover the 1959-2006 period, when the model is forced with NCEP, and the 1860-2099 period, when it is forced with GCM outputs.

In the ANALOG model [*Hingray et al.*, 2013], the K-nn days are identified on the basis of large scale fields of 4 meteorological variables: the geopotential heights for 1000 and 700 hPa for two dates (days D and D+1). The choice of spatial domain, variables, levels and dates is the result of a detailed performance analysis [e.g. *Bontron*, 2004]. The *Teweles and Wobus* [1954] distance is used as a similarity criterion to measure the distance between days in terms of the spatial shapes of the fields. The ANALOG model is similar to the version used by EDF since the 1990s for operational probabilistic precipitation forecasts [*Obled et al.*, 2002]. For the present application, the day used in the scenario is randomly selected from the 10 nearest neighbors.

In the three versions of DSCLIM, the analog day is also drawn randomly from the 10 nearest neighbors. The nearest neighbors are identified within the days belonging to the weather type of the target day [*Boé et al.*, 2006]. In the historical version, ten weather types defined from daily sea level pressure fields and daily precipitation patterns over France are used for each season (DJF, MAM, JJA, SON). In DSCLIM10, the similarity criterion used to identify the 10 nearest neighbors is the Euclidian distance and the predictors are 4 regional precipitation indices obtained for each day with regression laws from the distances between the day and the centroids of the ten weather type clusters. In DSCLIM11, the state vector also includes the mean large scale temperature of the day. In DSCLIM21, weather types and precipitation indices are defined from the geopotential at 850 hPa and the spatial covariance of 500 hPa geopotential fields is included in the predictors vector. These three model versions were selected from complementary analyses by *Lafaysse* [2011].

In both versions of D2GEN, the K-nn are identified from daily indices of regional precipitation and temperature [*Mezghani and Hingray*, 2009]. These indices are generated by stochastic independent processes from daily atmospheric synthetic predictors. For regional precipitation, two Generalized Linear Models are used: one to model occurrence probability of daily precipitation (a binomial distribution with the logistic link function), and the other to model the distribution function of precipitation amount in case of wet days (a gamma distribution with the log link function). For D2GEN10, the predictors are mean sea level pressure and zonal and meridian geostrophic wind speeds at 700 hPa. For D2GEN32, relative and specific humidity predictors at 700 hPa are included respectively in the occurrence and amount models. For both versions of the model, regional temperature is modeled from the daily air temperature at 700 hPa with a first order auto-regressive model. These predictors were chosen on the basis of earlier downscaling studies in the Swiss Alps [*Brandsma and Buishand*, 1997; *Mezghani*, 2009]. One analog day is resampled from the 10 nearest neighbors selected in the previous step. Regional precipitation and temperature indices are used to rescale the SAFRAN high resolution precipitation and temperature data of this analog day.

In both ANALOG and DSCLIM, in order to follow the temperature increase induced by anthropogenic forcings independently of atmospheric circulation changes, the local temperatures resampled from the analog day are corrected according to the difference between the large scale temperature of this day and that of the target day. For the 3 SDMs, daily adjustments of some SAFRAN variables are also carried out to account for the differences in temperature between the target day and the analog day. The phase of precipitation is re-estimated: precipitation is assumed to be liquid if the simulated

temperature is greater than 1°C and solid otherwise. Incoming longwave radiation is also re-estimated according to temperature differences as proposed by *Etchevers* [2000].

3.2.2. SDMs evaluation

An extensive evaluation of the 6 SDMs was done by *Lafaysse* [2011]. The performance of each SDM was first evaluated on its ability to reproduce, for all meteorological variables available in SAFRAN analyses, their main statistical properties (e.g. seasonality, mean spatial structure, statistical distribution) as well as auxiliary statistical features determinant for hydrology (e.g. precipitation intermittency, wet and dry period persistence). The ability of each SDM to reproduce the observed time variations of the variables ("chronological" evaluation) was also investigated. The analysis was done for variables aggregated over different time scales (e.g. daily, monthly, seasonal, annual values). A good "climatological" performance was obtained for all SDMs for all considered statistical properties. More contrasted characteristics were obtained for the "chronological" evaluation which is more severe. Here, model performance is illustrated for the reproduction of winter and summer precipitation (respectively 23% and 22% of the total annual precipitation).

For each simulated time series of winter (resp. summer) precipitation, we calculated over the 1959-2006 the correlation with the observed series (R) and the ratio between the corresponding simulated and observed standard deviations (Q_{sdv}). These two criteria evaluate respectively the ability of the model to reproduce the observed time variations and the inter-annual variability of the seasonal variable. Taylor diagrams of figure 3 present for each SDM the 100 sets (R, Q_{sdv}) corresponding to the 100 simulated time series. First, the results highlight the high variability of both scores from one replicate to the other. This inter-replicate variability is induced by the day-to-day small scale internal variability of the

figure 3

SDMs mentioned in section 1, for the same large scale forcing. As highlighted by figure 3, a large variety of meteorological time series is also obtained for aggregated variables (e.g. seasonal precipitation here). Note that a direct consequence is that a robust evaluation of the SDM performance cannot be based on a single generation. Despite slight differences (e.g. lower correlations for DSCLIM10 and DSCLIM11 in summer), the results are very similar for all 6 SDMs. Their ability to reproduce climatological statistics of winter and summer precipitation is satisfactory (Q_{sdv} often close to 1, whatever the season). Note that unbiased results are a priori easy to obtain with statistical downscaling methods: this is the case for any simple bootstrap method using the data archive. The chronological evaluation is more severe. For the present case, correlations are quite high in winter (from 0.7 to 0.9) but rather low in summer (from 0.2 to 0.6), resulting from a significantly lower explanatory power of SDMs in this season. These results are similar to those of other studies in the region. Large scale atmospheric circulations used by the SDMs are known to be fairly good predictors of the synoptic scale disturbances that prevail in winter. Conversely, summer precipitation is primarily due to local convective events, which are difficult to relate to large scale fields.

In the present study, SDMs are applied for the generation of future scenarios. They have therefore also to present a good temporal transferability. The chronological evaluation carried out here gives a rough idea of it as different climate conditions are observed throughout the evaluation period (1959-2005). A more stringent evaluation should concern climatological temporal scales, where time variations are those of the decadal or multi-decadal means of the surface variable as a result of significantly different climate conditions. This evaluation is however not possible due to the limited length of available observation data.

Nevertheless, we tried to evaluate the temporal transferability by applying the models for a period (1959-1981) different from the learning period (1981-2005). Quite substantial biases appeared for some models and specific months (DSCLIM10, D2GEN10, D2GEN32) suggesting lower transferability for those models [*Lafaysse*, 2011]. Biases were however found to potentially result from time heterogeneities in predictors and predictants. These heterogeneities are well-known in SAFRAN analyses [*Vidal et al.*, 2010] and in NCEP (or ERA40) reanalyses [*Sturaro*, 2003; *Sterl*, 2004]. For instance, values of humidity variables in NCEP reanalyses are significantly higher for the 1959-1981 period than for the last few decades without physical explanation. This can partially explain a large overestimation of precipitation by D2GEN32 for 1959-1981. Therefore, a clear discrimination between data heterogeneities and poor model transferability was not possible here. At this stage, we cannot exclude from our analysis the models suspected not to be transferable. However, our results suggest that the scientific validity of transferring such algorithms over time is not guaranteed. A better evaluation methodology should be investigated to address the challenge of data heterogeneities, with a selection of learning days picked through the entire period but based on the variability of predictors [e.g. *Raje and Mujumdar*, 2010]. Another possibility could be to use the simulations of a Regional Climate Model as pseudo-observations to calibrate and evaluate SDMs [e.g. *Frias et al.*, 2006]. This method does not suffer from data heterogeneities and would allow us exploring a wider variety of climate contexts than those contained in the past few decades.

3.3. Hydrological Model

3.3.1. Model description

The hydrological model for the Upper Durance river basin was developed and evaluated by *Lafaysse et al.* [2011] and is referred to as ISBA-Durance. SAFRAN meteorological analyses are used to force the ISBA physically based Soil-Snow-Vegetation-Atmosphere interaction scheme [*Boone et al.*, 1999; *Boone and Etchevers*, 2001]. Energy and mass balances are computed for 3 snow-layers and 3 soil-layers in order to simulate snow accumulation, settling and melt, rainfall interception by vegetation, evaporation from bare soil, evapotranspiration, infiltration, liquid water freezing and melt, surface runoff and soil drainage. This scheme is used operationally on an 8-km grid covering the entire French territory for hydrological forecasting and many other applications [*Habets et al.*, 2008]. In our case, the scheme is applied to an irregular discretization based on 639 relatively homogeneous hydrological units defined by sub-basin contours, elevation bands, aspect classes and the ice-covered or ice-free character. The scheme allows simulation of ice melt over glacier ablation areas. A groundwater module with two storage reservoirs was also implemented to simulate water retention by aquifers. Full details are given in *Lafaysse et al.* [2011].

3.3.2. Model evaluation

As shown in *Lafaysse et al.* [2011], this model gives satisfactory results for the simulation of the seasonal and interannual variability of daily discharges, the timing and intensity of the spring snowmelt floods, the rising limb and recessions of autumn floods and winter low flows. The model has been shown to be robust over a 46-year simulation period. The scores are stable when the model is transposed outside the period for which groundwater module calibrations were performed and it reasonably reproduces the observed negative trends of late spring flows when driven by SAFRAN analyses over the period 1959-2006. These

results, the physical basis of the model and the successful transfer of the ISBA scheme to various climatic contexts in the world [e.g. *Lejeune et al.*, 2007; *Biancamaria et al.*, 2009; *Boone et al.*, 2010] suggest that the model should provide reasonable hydrological scenarios when used in a modified climate context.

3.4. Complete simulation sequence

The different models of the whole simulation chain were evaluated independently in the previous section. In this section, we first evaluate the ability of the different SDM+HM chains to generate relevant hydrological scenarios. The hydrological evaluation is next presented for the whole GCM+SDM+HM chains.

Due to the non-linearities already mentioned in section 1, forcing a hydrological model by SDM outputs requires consistency between all meteorological forcing variables (e.g. relevance of the simulated inter-variable correlations). The resampling step in each SDM automatically ensures this consistency for at least each generation day, i.e. sub-daily space-time patterns of all local meteorological variables are necessarily physically consistent as they correspond to an actual specific day in the observation period. Unrealistic time sequences over several consecutive days and/or weeks would be the only possible limitation of the generation process. One way to evaluate the space-time and inter-variable consistency of meteorological series is to proceed with a hydrological evaluation [e.g. *Bourqui et al.*, 2011]. Hydrological simulations obtained from the stochastic meteorological replicates are evaluated against a reference hydrological time series obtained by simulation from the meteorological observations. The observed discharge series is not used for the reference so as not to bias the evaluation through intrinsic errors introduced by the hydrological model. Results of this evaluation are presented in figure 4 for two particular

figure 4

SDMs and six recent years. In spite of a high inter-replicate variability, their median fits the daily variations of the reference discharge time series quite well. Their seasonal and interannual variations are particularly well reproduced. For instance, the simulation chain captures the significant differences between the 2001 and 2005 spring floods, which were a result of very different amounts of precipitation falling as snow. The *Nash and Sutcliffe* [1970] efficiency between generated and reference discharges ranges from 0.67 to 0.83 depending on the SDM and the evaluation period (1959-1981 or 1981-2005). For comparison, the Nash-Sutcliffe efficiency between reference and observed discharges is 0.85. Hydrological errors associated with SDMs are thus similar to those associated with the hydrological model, indicating good inter-variable consistency of generated weather scenarios. Keeping in mind that the only variables used for these hydrological simulations are large scale geopotential fields, these results are very satisfactory. A significant part of this success is of course due to the high seasonality of flows and to the important low pass filter role of the catchment, however the non-linearities of hydrological processes can also reduce or amplify some errors in the meteorological scenarios. For instance, the precipitation bias of some SDMs for the 1959-1981 period is enhanced for discharges and the interannual variance is overestimated by some SDMs. One of the processes involved is the positive feedback loop induced by a precipitation increase, which enhances snow accumulation, reduces annual evaporation and therefore leads to a larger increase of discharges [*Lafaysse*, 2011].

The final step consists of forcing the SDMs by GCM outputs and generating hydrological scenarios for a reference period. The evaluation possibilities have the same limitations as in section 3.1.2. They should concern only climatological averages and account for the

internal variability of the GCMs. With the whole simulation chain, no significant biases were found in the seasonal cycle of discharges (see figure 7 for the 1980-1999 period).

4. Future projections

As the performance of different GCMs was estimated to be fairly similar in previous evaluations, it is difficult to exclude or give higher confidence to future projections obtained with any specific GCM. This also applies to results obtained with SDM models. Different GCM/SDM combinations can however provide significantly different future projections. We first present changes in the 20-yr interannual mean of different hydrometeorological variables obtained between 1980-1999 and 2080-2099 for all replicates (100) generated for all GCM/SDM simulation chains. For selected simulation chains, we also present the time evolution of these changes over the whole 1860-2099 simulation period. The total uncertainty of changes and of the different components of uncertainty is estimated quantitatively as a function of lead time in section 5.

4.1. Distribution of changes

The cumulative distribution functions (CDFs) of changes from 1980-1999 to 2080-2099 are presented in figure 5 for mean annual temperature, mean annual precipitation and mean seasonal precipitation for winter and summer.

figure 5

In terms of temperature, major warming is obtained at the annual scale and simulated temperature changes appear to be rather consistent from one simulation chain to another. The inter-GCM dispersion is relatively high and the inter-SDM dispersion is non-negligible. The small scale internal variability corresponds to dispersion between the 100 replicates obtained for each chain. As shown by the spread of each CDF, it appears

to be negligible. The same results are obtained for changes in seasonal temperature, even if larger increases are obtained for summer and fall and lower increases for winter (not shown). In general and regardless the season, the simulated warming for the 2080-2099 projection lead time is very high for all simulation chains.

Results are rather different for precipitation. Depending on the simulation chain considered, annual precipitation is projected to increase or decrease. The same applies for changes in seasonal precipitation. In winter, the sign of the change is also very uncertain. Depending on the simulation chain, the CDFs indicate that a precipitation increase is very likely (median change of CDF up to -25%), very unlikely (median change up to +50%) or as likely as a precipitation decrease. For the other seasons, the sign of the change corresponds more often to a precipitation decrease, but a large dispersion of CDFs is still obtained and some combinations again suggest a significant precipitation increase. The inter-GCM dispersion is high. This can be explained mainly by the large differences simulated in the evolution of atmospheric circulation. The highest precipitation decreases, for IPCM4-3 in winter and for CNCM33-1 and DMIEH5C-3 in autumn, are for example linked to a clear decrease in the frequency of the south- westerly flows over France, a regime which is usually wet for this region [*Lafaysse*, 2011]. While GCMs are generally considered to be the most significant source of uncertainty, inter-SDM dispersion seems to be roughly as large. Simulated changes can be very different not only for different SDMs but even for different versions of a given SDM. This is the case for DSCLIM-10 and DSCLIM-11 (shift of about 20% of the CDF for all GCMs), based on different large scale atmospheric predictors (dynamical predictors are used in DSCLIM-10 and dynamical + thermo-dynamical predictors in DSCLIM-11). The same applies to D2GEN-10

and D2GEN-32. The strong predicted increase of winter precipitation in D2GEN32 is linked to the predicted increase of specific humidity in all GCMs, a predictor ignored in other SDMs. Small scale internal variability appears lower than inter-GCM or inter-SDM dispersion but is not negligible.

Future hydrological projections were simulated by the ISBA-Durance hydrological model from previous meteorological scenarios. As a result of the predicted temperature increase, the snow-rain transition elevation increases, leading in turn to a decrease in the portion of snowfall in the total basin precipitation (not shown). Important changes are thus expected for snowpack-related hydrological variables. Figure 6 (right, middle right) presents the CDFs of simulated changes between 2080-2099 and 1980-1999 in snow cover duration (SCD) for medium and high elevations. The simulated decrease ranges from -30% to -80% at 1650 meters. The amount of change is largely dependent on the rate of warming from the GCM/SDM chain. At 2550 meters, the simulated decreases ranges from -10% to -50%. At this elevation, due to the higher sensitivity of changes in snowpack to changes in precipitation, the influence of SDM is higher than at mid-elevations and large differences are obtained between DSCLIM10 and DSCLIM11 or between D2GEN10 and D2GEN32, consistent with the conclusions for precipitation. Globally, despite the high dispersion of projected precipitation, the predicted snow cover decrease is consistent between simulation chains as a consequence of the small dispersion between chains for temperature increase. Small scale internal variability is very low, in line with results obtained for temperature. Inter-GCM and inter-SDM dispersion seems to be of the same order, consistent with results obtained for precipitation and temperature. Note also that the predicted snow cover decrease is relatively homogeneous for different elevations and

Figure 6

aspects as illustrated in figure 9 for one specific simulation chain. Snowpack characteristics simulated at the end of the XXIst century correspond to those simulated at the end of the XXth century for elevations 400 to 800 meters lower, depending on the simulation chain.

figure 9

The predicted increase in annual evaporation is important for all simulation chains (figure 6, middle left). Dispersions obtained between chains mainly reflect those for temperature warming, as a result especially of the strong dependence of the annual evaporation on the snow cover extent (evaporation is much more efficient on snow-free surfaces). Small scale internal variability is again very low.

As annual discharges result from both annual precipitation and evaporation, most chains simulate a large decrease for this variable (figure 6, left). The dispersion between projections is consistently larger than that predicted for precipitation or evaporation only (changes in discharges range from -45% to +20% with a higher increase for the particular IPCM4-3/D2GEN32 combination). Inter-SDM and inter-GCM dispersions are again of the same order and small scale internal variability is again rather large. A more consistent signal between models is obtained for other hydrological variables. For instance, the seasonality of river flows would be strongly reduced by the decrease in snow cover extent and duration. The spring snowmelt flood is expected to be less intense and to occur earlier in the season regardless the simulation chain, even if the different changes in precipitation can modulate the magnitude of these changes (figure 7). The evolution of autumn discharges, mostly determined by direct rainfall-runoff processes, is rather different from one chain to the other.

figure 7

The dispersion among meteorological and hydrological projected changes provide insight on the uncertainty sources and on their relative contribution to total uncertainty. From previous graphs, we could for instance discuss the inter-GCM and the inter-SDM dispersions as well as the small scale internal variability obtained for each GCM/SDM chain. However, these graphs may be misleading due to large scale internal variability of each simulation chain. Let us temporarily neglect the small scale internal variability component of the projections. Let us thus only consider, for each simulation chain, the multireplicate median projection of the chain (the 50th percentile of each CDF in the graphs). For a given chain, this median projection is actually the combined result of: 1) the mean response of the chain to the climate change scenario under consideration and 2) the large scale internal variability of the chain results. The effect of large scale internal variability can be estimated by the dispersion between the different projections that would be obtained from different runs of the GCM of this chain. In the present case, the effect of large scale internal variability seems to be large for changes in precipitation and discharge as highlighted by the large differences obtained for the 3 GCM runs of DMIEH5C or MPIEH5C or IPCM4.

4.2. Time evolution of estimated changes

Another possibility to evaluate the importance of large scale internal variability is to plot for the studied variable the time evolution of its 20-yr interannual mean over the 1860-2099 simulation period. These time evolution plots are presented in figure 8 for annual temperature, annual, winter and summer precipitation, for all replicates of a given GCM/SDM chain (figure 8, bottom) and the multireplicate median of different chains with the same GCM (figure 8, middle) or with the same SDM (figure 8, top). The graphs

figure 8

first reveal the large dependence of decadal variations in all variables on the internal variability of atmospheric circulation regimes (see for instance the temporal co-variations of precipitation in figure 8 (middle), where the different SDMs are driven by one particular GCM) and the chaotic noise in figure 8 (top), where one particular SDM is driven by all GCMs). They also show that large scale internal variability highly depends on the variable. For temperature, it appears to be very low, at least when compared to the predicted warming for the future century. Similar results are obtained for snow cover variables and annual evaporation, highly influenced by temperatures (not shown). For precipitation, large scale internal variability is conversely very large, i.e. at least as large as estimated changes. The same applies for annual discharge. Large scale internal variability can also depend on the season, being for instance much larger for winter than for summer precipitation.

5. Uncertainty contributions analysis

The different uncertainty sources in the projections reported here were highlighted in the previous section. Their relative contributions to total uncertainty in projections are explored in the present section. The quantification of total uncertainty also allows computing confidence intervals of estimated changes. These analyses are done as a function of projection lead time, from 2000 to 2099. The quantification of total uncertainty and of the different uncertainty components is carried out using the Quasi-Ergodic ANOVA framework presented by *Hingray and Saïd* [2014] and summarized in Appendix A. It was applied in the present work to assess the different components of model uncertainty and internal variability for the different change variables discussed in section 4 (change in 20-year interannual mean temperature, precipitation, evaporation, discharge, snow cover

duration). The total variance $T(t)$ of the change variable for a given future projection lead time t is given by the sum of the following variance components:

$$T(t) = G(t) + S(t) + E(t) + LSIV(t) + SSIV(t)$$

where $G(t)$, $S(t)$, $E(t)$, $LSIV(t)$ and $SSIV(t)$ correspond respectively to the GCM uncertainty, SDM uncertainty, residual uncertainty and large and small scale internal variability. The residual term is attributable to GCM/SDM interaction effects. The magnitude and total uncertainty of one given change variable are characterized by the multichain mean of its climate change response (see Appendix A) at t and its total variance.

5.1. Model uncertainty and internal variability components

The left panels of figure 10 and figure 11 present for different variables the mean climate change response $\mu(t)$ and the limits of the interval $\mu(t) \pm 1.645\sqrt{T(t)}$. The size of each colored zone is a function of the fraction of total uncertainty explained by each uncertainty component.

figure 10

figure 11

For all considered variables, total uncertainty increases with lead time. The main contribution for this increase is that of model uncertainty, because both components of internal variability are constant or roughly constant over the whole period. The right panels of figures 10 and 11 present the relative contributions of the different model uncertainty and internal variability components to the total uncertainty variance. These relative contributions vary significantly for different projection lead times and variables. For the first 3 decades, the combined contribution of small scale and large scale internal variability is clearly predominant for all variables. It then decreases with lead time as model uncertainty increases. For temperature and snow cover duration, it drops sharply to less

than 10% at the end of the century, becoming negligible compared to model uncertainty. For evaporation, it also decreases rapidly down to roughly 20% in 2090. For annual precipitation and annual discharge, it is still roughly 50% of total uncertainty in 2090. For summer precipitation, this contribution is smaller (40%) but for winter, it is even greater than 65% (these seasonal differences are consistent with the differences in the variability of 1860-2099 trajectories already presented for both variables in figure 8).

For all variables, large scale internal variability is greater than small scale internal variability (excepted for summer precipitation). For temperature, the contribution of small scale internal variability to total internal variability (and in turn to total uncertainty) is even negligible regardless the time horizon. For all other variables (excepted for snow cover duration at 1550m), it is significant: for annual precipitation, discharge and evaporation, small scale internal variability represents about 30% of total internal variability and more than 10-20% of total uncertainty at the end of the century. An interesting result is obtained for seasonal precipitation, for which the contribution of small scale internal variability to total internal variability is roughly 20% for winter while exceeding 50% for summer. This results from the much lower predictive power of large scale predictors obtained for summer as discussed in section 3.2.2 and shown in figure 3.

As already mentioned, the contribution of model uncertainty to total uncertainty increases over the whole simulation period for all variables. The contribution of the GCM/SDM interaction term is significant but of second order compared to the contributions of GCM and SDM uncertainty. For temperature, snow cover duration and evaporation, GCM uncertainty is the main contribution after 2050. The contribution of the SDM is however significant. It is up to 20% in 2090 and often equivalent to that of

GCM for the middle of the century. For annual and seasonal precipitation, especially in summer, the contribution of SDM is higher than that of GCM. For annual discharge, the contribution of SDM uncertainty is even higher, resulting from the opposite effects of the respective SDMs on temperature and precipitation leading to an amplification of their effects on discharge (e.g. D2GEN32 leads to one of the smallest warming results, but to the wettest future scenarios regardless the driving GCM).

5.2. Confidence intervals

In the left panels of figures 10 and 11, the total area corresponds to the 90% confidence interval of the change variables, assuming the latter have a normal distribution. The lower (resp. upper) limit of this interval therefore corresponds to the value which has a 95% chance of not being exceeded. A significant non-zero change at a 90% confidence level is obtained when the zero value is outside this confidence interval. This is quickly the case for temperature, evaporation and snow cover duration after 2020-2030 because total uncertainty is low compared to the mean predicted change. For snow cover duration at the low elevation range, both the mean decrease and the total uncertainty are roughly twice than for the high elevation range. Conversely, the zero value of change is within the confidence interval for annual and seasonal precipitation and annual discharges. Except from winter precipitation, a different conclusion would be obtained if only the model uncertainty (i.e. blue and green areas) were considered: Although the sign of the mean expected effect of climate change is consistent between models, the internal variability adds a sufficiently large noise to compensate this effect. Therefore, and mainly as a consequence of internal variability, no assertion can be done on the sign of the future changes of mean interannual precipitation and discharges. This doesn't exclude that the

603 XXIst century climate will possibly exhibit significant changes relatively to the observed
604 recent precipitation and discharge variability, but these changes are not predictable. In
605 other words, the large internal variability associated to precipitation and discharge will
606 be able to temporarily either aggravate, moderate or even reverse the long-term impact
607 of climate change on these variables.

6. Discussion and conclusions

6.1. Methodology and interpretation of results

608 The multireplicate multimodel ensemble of hydrometeorological projections obtained for
609 the 1860-2099 period allows for the estimation of several sources of uncertainties: those
610 related to model errors (GCM, SDM, GCM/SDM interaction) and those related to the
611 internal variability of climate (large and small scale).

612 For a given projection lead time, estimating the total uncertainty of projections and the
613 contribution of the different uncertainty components from the sample of projections may
614 be misleading when the number of projection replicates is low and internal variability high.
615 A robust quantification has been carried out in the present work using the Quasi-Ergodic
616 ANOVA framework presented by *Hingray and Saïd* [2014], based on the quasi-ergodic
617 assumption for transient climate simulations. In our results, the contribution of the
618 different uncertainty components to total uncertainty is highly dependent on the variable
619 and on projection lead time. For temperature, GCM uncertainty prevails and, as opposed
620 to internal variability, SDM uncertainty is non-negligible. Significant warming and in turn
621 significant non-zero changes are expected for evaporation, snow cover and seasonality of
622 discharges. For precipitation, GCM and SDM uncertainty components are of the same
623 order. A high contribution of the large and small scale components of internal variability

is also obtained, inherited respectively from the GCMs and the different replicates of a given SDM. The same applies for annual discharge. The uncertainty in values that could be experienced for any given future period is therefore very high. For both discharge and precipitation, even the sign of future realizations is uncertain at a 90% level, under the A1B emission scenario. These finding have important implications.

1. Consistent with other recent studies [e.g. *Chen et al.*, 2011], SDMs have been found to produce a significant contribution to model uncertainty. Large dispersion may come from different SDM structures but also from different sets of predictors of the same SDM. Impact studies based on a single SDM, or on a single set of predictors for a given SDM, are likely to be of little value.

2. Depending on the studied variable, the contribution of internal variability to total uncertainty can be also very large. In the present case, the large internal variability associated to precipitation and discharge will be able to temporarily but significantly either aggravate, moderate or even reverse the long-term impact of global warming. Neglecting internal variability in impact studies is therefore expected to lead to erroneous estimates of possible climate changes.

3. The small scale component of internal variability, as simulated for instance by SDMs, has hardly ever been accounted for in impact studies. However, impact studies based on a single SDM realization (or small ensembles) are likely to be no more relevant than those based on a single run of available GCMs (or small ensembles). When they are intended to provide information for climate change adaptation, they may lead to poor decisions. In the present case, significant non-zero changes were obtained for variables mainly driven by temperature warming (snow cover decrease and discharge seasonality).

Such modifications are obviously of interest for adaptation issues. However, results also suggest that a relevant adaptation strategy should consider the possibility to adapt to internal variability of precipitation.

While our conclusions apply to the Upper Durance river basin, the dispersion associated with the different uncertainty sources studied is however likely to be encountered in other contexts and areas. As a significant large scale warming is expected for most mid-latitude regions, the robust decrease of snow cover and the associated modifications of evaporation and discharge seasonality are also expected in other high elevation basins of the Alps [e.g. *Jasper et al.*, 2004; *Bavay et al.*, 2009] or in other parts of the world [e.g. *Maurer*, 2007, Sierra Nevada, California]. In low elevation regions, where snow events are rare, the river discharge modifications will be driven more directly by precipitation and soil moisture changes. For snow-free mesoscale catchments, the high internal variability of precipitation could prevail for all hydrological variables. However, for larger basins, we would expect a lower variability for precipitation changes and therefore a possibly more significant response in future projections. Large uncertainties may however still remain at these scales.

6.2. Accounting for other uncertainty sources

The influence of additional uncertainty sources on the total uncertainty should obviously be investigated, especially those associated with greenhouse gas emission scenarios and with the choice of hydrological model. From recent works in Canadian basins with areas from 160 km^2 [*Grillakis et al.*, 2011] to 24600 km^2 [*Chen et al.*, 2011], hydrological uncertainties seem, however, to be much lower than the other uncertainties. Another po-

tentially important uncertainty source is related to the post-processing procedure applied to remove the biases of GCM large scale fields before applying SDMs [*Ehret et al.*, 2012].

The possibility of partitioning the different uncertainty sources remains of course a challenging issue when multiple sources of uncertainty are considered. The QE-ANOVA framework applied here could also be easily extended to hydrological projections from multiple hydrological models. It could also be used to estimate emission scenario uncertainty [*Hingray and Saïd*, 2014].

6.3. The potential to narrow uncertainty

Another critical question is the potential to narrow uncertainty. It will of course never be possible to remove uncertainties related to large and small scale natural variability because they are intrinsic to the Earth system [*Hawkins and Sutton*, 2009, 2011]. Conversely, model uncertainties can be reduced by a better understanding of geophysical processes and resulting improvements of numerical models. In the present study, for instance, the high contribution of SDM uncertainty for temperature change actually reflects the influence of the different large scale temperature predictors retained for the SDMs. The importance of this issue has been underestimated in the development of these SDMs. As all SDMs perform much better for temperature than for the precipitation, the main model development efforts have been focused on precipitation. More stringent tests need to be carried out for temperature in future SDM development work. They will likely narrow SDM uncertainty for this variable and for hydrological variables in snow dominated catchments.

The total uncertainty is also expected to be reduced if only the "best" models are taken into account. However, appropriate evaluation methodologies will be required for a

relevant selection of the "best" models. The selection could for instance be carried out on the basis of their robustness in temporal transposition. A number of other criteria could be critical, making this selection rather uncertain.

Acknowledgments.

This work is part of a PhD research project funded by Météo-France. It is linked to the RIWER2030 "Regional ClImate, Water, Energy Resources and uncertainties from 1960 to 2030" research project (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). Future hydrological simulations were run on the CIMENT computing grid (CIGRI). The authors also wish to thank Christian Pagé (CERFACS) for technical and scientific support on the DSCLIM model. The author also thanks Mériem Saïd (LTHE, Grenoble) for helpful discussions concerning this work. We finally thank the three anonymous reviewers for their careful revision, which helped improve this manuscript.

References

Anagnostopoulou, C., K. Tolika, and P. Maheras (2009), Classification of circulation types: a new flexible automated approach applicable to NCEP and GCM datasets, *Theor. Appl. Climatol.*, *96*(1-2), 3–15, doi:10.1007/s00704-008-0032-6.

Arora, M., P. Singh, N. K. Goel, and R. D. Singh (2008), Climate variability influences on hydrological responses of a large Himalayan basin, *Water Resour. Res.*, *22*(10), 1461–1475, doi:10.1007/s11269-007-9237-1.

Baltas, E. A. (2007), Impact of climate change on the hydrological regime and water resources in the basin of Siatista, *Int. J. Water Resour. Dev.*, *23*(3), 501–518, doi: 10.1080/07900620701485980.

Bavay, M., M. Lehning, T. Jonas, and H. Loewe (2009), Simulations of future snow cover and discharge in Alpine headwater catchments, *Hydrol. Process.*, *23*(1, Sp. Iss. SI), 95–108, doi:10.1002/hyp.7195.

Bethers, U., and J. Sennikovs (2009), Ensemble modelling of impact of climate change on runoff regime of Latvian rivers, in *18th World Imacs Congress and MODSIM09 International Congress on Modelling and Simulation - Interfacing modelling and simulation with mathematical and computational sciences*, edited by Anderssen, RS and Braddock, RD and Newham, LTH, pp. 3900–3906, Univ. Western Australia.

Biancamaria, S., P. D. Bates, A. Boone, and N. M. Mognard (2009), Large-scale coupled hydrologic and hydraulic modelling of the Ob river in Siberia, *J. Hydrol.*, *379*(1-2), 136–150, doi:10.1016/j.jhydrol.2009.09.054.

Boé, J. (2007), Changement global et cycle hydrologique : Une étude de régionalisation sur la France, Ph.D. thesis, Université Paul Sabatier, Toulouse.

Boé, J., L. Terray, F. Habets, and E. Martin (2006), A simple statistical downscaling scheme based on weather types and conditional resampling, *J. Geophys. Res.*, *111*.

Boé, J., L. Terray, E. Martin, and F. Habets (2009), Projected changes in components of the hydrological cycle in French river basins during the 21st century, *Water Resour. Res.*, *45*, doi:10.1029/2008WR007437.

Bontron, G. (2004), Prévision quantitative des précipitations. Adaptation probabiliste par recherche d'analogues., Ph.D. thesis, Institut National Polytechnique de Grenoble.

Boone, A., and P. Etchevers (2001), An intercomparaison of three snow schemes of varying complexity coupled to the same land surface and macroscale hydrologic models, *J. Hydrometeorol.*, *2*, 374–394.

Boone, A., J. Calvet, and J. Noilhan (1999), Inclusion of a third soil layer in a land surface scheme using the force-restore method, *J. Appl. Meteorol.*, *38*, 1611–1630.

Boone, A. A., I. Poccard-Leclercq, Y. Xue, J. Feng, and P. de Rosnay (2010), Evaluation of the WAMME model surface fluxes using results from the AMMA land-surface model intercomparison project, *Clim. Dynam.*, *35*(1, Sp. Iss. SI), 127–142, doi:10.1007/s00382-009-0653-1.

Bourqui, M., T. Mathevet, J. Gailhard, and F. Hendrickx (2011), Hydrological validation of statistical downscaling methods applied to climate model projections, in *HYDRO-CLIMATOLOGY: VARIABILITY AND CHANGE*, *IAHS Publication*, vol. 344, edited by Franks, SW and Boegh, E and Blyth, E and Hannah, DM and Yilmaz, KK, pp. 32–38, 25th General Assembly of the International Union of Geodesy and Geophysics, Melbourne, AUSTRALIA, JUN 28-JUL 07, 2011.

Boyer, C., D. Chaumont, I. Chartier, and A. G. Roy (2010), Impact of climate change on the hydrology of St. Lawrence tributaries, *J. Hydrol.*, *384*(1-2), 65–83, doi: 10.1016/j.jhydrol.2010.01.011.

Brands, S., S. Herrera, D. San-Martin, and J. M. Gutierrez (2011), Validation of the ENSEMBLES global climate models over southwestern Europe using probability density functions, from a downscaling perspective, *Clim. Res.*, *48*(2-3), 145–161, doi: 10.3354/cr00995.

Brandsma, T., and T. Buishand (1997), Statistical linkage of daily precipitation in Switzerland to atmospheric circulation and temperature, *J. Hydrol.*, *198*(1-4), 98–123, doi:10.1016/S0022-1694(96)03326-4.

Brasseur, G., and E. Roeckner (2005), Impact of improved air quality on the future evolution of climate, *Geophys. Res. Lett.*, *32*(23), doi:10.1029/2005GL023902.

Buishand, T., and T. Brandsma (2001), Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling, *Water Resour. Res.*, *37*(11), 2761–2776.

Caballero, Y., P. Chevallier, A. Boone, J. Noilhan, and F. Habets (2007), Calibration of the interaction soil biosphere atmosphere land-surface scheme on a small tropical high-mountain basin (Cordillera Real, Bolivia), *Water Resour. Res.*, *43*(7), doi:10.1029/2005WR004490.

Chang, H., and I.-W. Jung (2010), Spatial and temporal changes in runoff caused by climate change in a complex large river basin in Oregon, *J. Hydrol.*, *388*(3-4), 186–207, doi:10.1016/j.jhydrol.2010.04.040.

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 Resour. Res.*, *47*, doi:10.1029/2011WR010602.

Cherkauer, K. A., and T. Sinha (2010), Hydrologic impacts of projected future climate change in the Lake Michigan region, *J. Great Lakes Res.*, *36*(Sp. Iss. SI Suppl. 2), 33–50, doi:10.1016/j.jglr.2009.11.012.

Coulibaly, P. (2009), Multi-model approach to hydrologic impact of climate change, in *From headwaters to the ocean : hydrological changes and watershed management*, edited

by Taniguchi, M and Burnett, WC and Fukushima, Y and Haigh, M and Umezawa, Y, pp. 249–255, CRC PRESS-TAYLOR & FRANCIS GROUP.

Dankers, R., and O. Christensen (2005), Climate change impact on snow coverage, evaporation and river discharge in the sub-arctic Tana Basin, Northern Fennoscandia, *Clim. Change*, *69*(2-3), 367–392, doi:10.1007/s10584-005-2533-y.

Dankers, R., and L. Feyen (2008), Climate change impact on flood hazard in Europe: An assessment based on high-resolution climate simulations, *J. Geophys. Res.-Atmos.*, *113*(D19), doi:10.1029/2007JD009719.

Das, T., M. D. Dettinger, D. R. Cayan, and H. G. Hidalgo (2011), Potential increase in floods in California's Sierra Nevada under future climate projections, *Clim. Change*, *109*(1, SI), 71–94, doi:10.1007/s10584-011-0298-z.

Demuzere, M., M. Werner, N. P. M. van Lipzig, and E. Roeckner (2009), An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns, *Int. J. Climatol.*, *29*(12), 1796–1810, doi:10.1002/joc.1821.

Dequé, M. (2007), Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, *Glob. Planet. Change*, *57*(1-2), 16–26, doi:10.1016/j.gloplacha.2006.11.030.

Dibike, Y., and P. Coulibaly (2005), Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models, *J. Hydrol.*, *307*(1-4), 145–163, doi:10.1016/j.hydrol.2004.10.012.

Dobler, C., S. Hagemann, R. L. Wilby, and J. Stoetter (2012), Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed, *Hydrol. Earth Syst. Sci.*, *16*(11), 4343–4360, doi:10.5194/hess-16-4343-2012.

Driessen, T. L. A., R. T. W. L. Hurkmans, W. Terink, P. Hazenberg, P. J. J. F. Torfs, and R. Uijlenhoet (2010), The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model, *Hydrol. Earth Syst. Sci.*, *14*(4), 651–665.

Dufresne, J., J. Quaas, O. Boucher, S. Denvil, and L. Fairhead (2005), Contrasts in the effects on climate of anthropogenic sulfate aerosols between the 20th and the 21st century, *Geophys. Res. Lett.*, *32*(21), doi:10.1029/2005GL023619.

Durand, Y., M. Laternser, G. Giraud, P. Etchevers, B. Lesaffre, and L. Merindol (2009), Reanalysis of 44 Yr of Climate in the French Alps (1958-2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation, *J. Appl. Meteorol. Climatol.*, *48*(3), 429–449, doi:10.1175/2008JAMC1808.1.

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert (2012), "Should we apply bias correction to global and regional climate model data?", *Hydrol. Earth Syst. Sci.*, *16*(9), 3391–3404, doi:10.5194/hess-16-3391-2012.

Errasti, I., A. Ezcurra, J. Saenz, and G. Ibarra-Berastegi (2011), Validation of IPCC AR4 models over the Iberian Peninsula, *Theor. Appl. Climatol.*, *103*(1-2), 61–79, doi:10.1007/s00704-010-0282-y.

Etchevers, P. (2000), Modélisation de la phase du cycle continental de l'eau á l'échelle régionale : impact de la modélisation de l'enneigement sur l'hydrologie du bassin versant du Rhône., Ph.D. thesis, Université Paul Sabatier, Toulouse.

Frias, M. D., E. Zorita, J. Fernandez, and C. Rodriguez-Puebla (2006), Testing statistical downscaling methods in simulated climates, *Geophys. Res. Lett.*, *33*(19), doi:10.1029/2006GL027453.

Gangopadhyay, S., M. Clark, and B. Rajagopalan (2005), Statistical downscaling using K-nearest neighbors, *Water Resour. Res.*, *41*(2), doi:10.1029/2004WR003444.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *J. Geophys. Res.-Atmos.*, *113*(D6), doi:10.1029/2007JD008972.

Graves, D., and H. Chang (2007), Hydrologic impacts of climate change in the Upper Clackamas River Basin, Oregon, USA, *Clim. Res.*, *33*(2), 143–157.

Grillakis, M. G., A. G. Koutroulis, and I. K. Tsanis (2011), Climate change impact on the hydrology of Spencer Creek watershed in Southern Ontario, Canada, *J. Hydrol.*, *409*(1-2), 1–19, doi:10.1016/j.jhydrol.2011.06.018.

Habets, F., et al. (2008), The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *J. Geophys. Res.*, *113*.

Harding, B. L., A. W. Wood, and J. R. Prairie (2012), The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin, *Hydrol. Earth Syst. Sci.*, *16*(11), 3989–4007, doi:10.5194/hess-16-3989-2012.

Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *Bull. Amer. Meteorol. Soc.*, *90*.

Hawkins, E., and R. Sutton (2011), The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dynam.*, *37*(1-2), 407–418, doi:10.1007/s00382-010-0810-6.

Hingray, B., and M. Saïd (2014), Partitioning internal variability and model uncertainty components in a multimodel multireplicate ensemble of hydro-meteorological climate projections., *J. Climate, in revision*.

Hingray, B., N. Mouhous, A. Mezghani, K. Bogner, B. Schaefli, and A. Musy (2007), Accounting for global-mean warming and scaling uncertainties in climate change impact studies: application to a regulated lake system, *Hydrol. Earth Syst. Sci.*, *11*(3), 1207–1226.

Hingray, B., et al. (2013), RIWER2030: Climats Régionaux et Incertitudes, Ressource en Eau et Gestion associée de 1860 á 2100.Projet ANR VMCS 2009-2012., *Rapport Final*, CNRS/LTHE, EDF/LNHE, EDF/DTG. Grenoble.

Horton, P., B. Schaefli, A. Mezghani, B. Hingray, and A. Musy (2006), Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty, *Hydrol. Process.*, *20*(10), 2091–2109, doi:10.1002/hyp.6197.

Hreiche, A., W. Najem, and C. Bocquillon (2007), Hydrological impact simulations of climate change on Lebanese coastal rivers, *Hydrol. Sci. Journal*, *52*(6), 1119–1133, International Conference on Future of Drylands, Tunis, TUNISIA, JUN, 2006.

Huebener, H., U. Cubasch, U. Langematz, T. Spangehl, F. Niehrster, I. Fast, and M. Kunze (2007), Ensemble climate simulations using a fully coupled oceantroposphere-estratosphere general circulation model, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2089–2101, doi:10.1098/rsta.2007.2078.

Huss, M., D. Farinotti, A. Bauder, and M. Funk (2008), Modelling runoff from highly glacierized alpine drainage basins in a changing climate, *Hydrol. Process.*, *22*(19, Sp. Iss. SI), 3888–3902, doi:10.1002/hyp.7055.

Immerzeel, W. W., L. P. H. van Beek, M. Konz, A. B. Shrestha, and M. F. P. Bierkens (2012), Hydrological response to climate change in a glacierized catchment in the Hi-

malayas, *Clim. Change*, *110*(3-4), 721–736, doi:10.1007/s10584-011-0143-4.

IPCC (2007), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, *Tech. rep.*

Jasper, K., C. Pierluigi, D. Gyalistras, and F. J. (2004), Differential impacts of climate change on the hydrology of two alpine river basins, *Clim. Res.*, *26*, 113–129.

Johns, T., et al. (2006), The new Hadley Centre Climate Model (HadGEM1): Evaluation of coupled simulations, *J. Climate*, *19*(7), 1327–1353, doi:10.1175/JCLI3712.1.

Johns, T. C., et al. (2011), Climate change under aggressive mitigation: the ENSEMBLES multi-model experiment, *Clim. Dynam.*, *37*(9-10), 1975–2003, doi:10.1007/s00382-011-1005-5.

Juen, I., G. Kaser, and C. Georges (2007), Modelling observed and future runoff from a glacierized tropical catchment (Cordillera Blanca, Peru), *Glob. Planet. Change*, *59*(1-4), 37–48, doi:10.1016/j.gloplacha.2006.11.038.

Jung, I.-W., and H. Chang (2011), Assessment of future runoff trends under multiple climate change scenarios in the Willamette River Basin, Oregon, USA, *Hydrol. Process.*, *25*(2), 258–277, doi:10.1002/hyp.7842.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteorol. Soc.*, *77*(3), 437–471.

Keller, F., S. Goyette, and M. Beniston (2005), Sensitivity analysis of snow cover to climate change scenarios and their impact on plant habitats in alpine terrain, *Clim. Change*, *72*(3), 299–319, doi:10.1007/s10584-005-5360-2.

Kerkhoven, E., and T. Y. Gan (2011), Differences and sensitivities in potential hydrologic impact of climate change to regional-scale Athabasca and Fraser River basins of the leeward and windward sides of the Canadian Rocky Mountains respectively, *Clim. Change*, *106*(4), 583–607, doi:10.1007/s10584-010-9958-7.

Kingston, D. G., J. R. Thompson, and G. Kite (2011), Uncertainty in climate change projections of discharge for the Mekong River Basin, *Hydrol. Earth Syst. Sci.*, *15*(5), 1459–1471, doi:10.5194/hess-15-1459-2011.

Kunstmann, H., K. Schneider, R. Forkel, and R. Knoche (2004), Impact analysis of climate change for an Alpine catchment using high resolution dynamic downscaling of ECHAM4 time slices, *Hydrol. Earth Syst. Sci.*, *8*(6), 1030–1044.

Lafaysse, M. (2011), Changement climatique et régime hydrologique d'un bassin alpin. Génération de scénarios sur la Haute-Durance, méthodologie d'évaluation et incertitudes associées., Ph.D. thesis, Université Paul Sabatier, Toulouse.

Lafaysse, M., B. Hingray, P. Etchevers, E. Martin, and C. Obled (2011), Influence of spatial discretization, underground water storage and glacier melt on a physically-based hydrological model of the Upper Durance River basin, *J. Hydrol.*, *403*(1-2), 116–129, doi:10.1016/j.jhydrol.2011.03.046.

Larson, R. P., J. M. Byrne, D. L. Johnson, S. W. Kienzle, and M. G. Letts (2011), Modelling Climate Change Impacts on Spring Runoff for the Rocky Mountains of Montana and Alberta II: Runoff Change Projections using Future Scenarios, *Can. Water Res. J.*, *36*(1), 35–51.

Lejeune, Y., P. Wagnon, L. Bouilloud, P. Chevallier, P. Etchevers, E. Martin, J. Sicart, and F. Habets (2007), Melting of Snow Cover in a Tropical Mountain Environment in

Bolivia : Processes and Modeling, *J. Hydrometeorol.*, *8*, 922–937.

Li, L., Z.-C. Ha, J.-H. Wang, Z.-H. Wang, and Z.-B. Yu (2008), Impact of future climate change on runoff in the head region of the Yellow River, *J. Hydrol. Eng.*, *13*(5), 347–354, doi:10.1061/(ASCE)1084-0699(2008)13:5(347).

Liu, Z., Z. Xu, J. Huang, S. P. Charles, and G. Fu (2010), Impacts of climate change on hydrological processes in the headwater catchment of the Tarim River basin, China, *Hydrol. Process.*, *24*(2, Sp. Iss. SI), 196–208, doi:10.1002/hyp.7493.

Lopez-Moreno, J., S. Goyette, M. Beniston, and B. Alvera (2008), Sensitivity of the snow energy balance to climatic changes : prediction of snowpack in the Pyrenees in the 21st century, *Clim. Res.*, *36*, 203–217.

Lucas-Picher, P., D. Caya, R. de Elia, and R. Laprise (2008), Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain, *Clim. Dynam.*, *31*(7-8), 927–940, doi:10.1007/s00382-008-0384-8.

MacDonald, R. J., J. M. Byrne, S. W. Kienzle, and R. P. Larson (2011), Assessing the Potential Impacts of Climate Change on Mountain Snowpack in the St. Mary River Watershed, Montana, *J. Hydrometeorol.*, *12*(2), 262–273, doi:10.1175/2010JHM1294.1.

Magnusson, J., T. Jonas, I. Lopez-Moreno, and M. Lehning (2010), Snow cover response to climate change in a high alpine and half-glacierized basin in Switzerland, *Hydrol. Res.*, *41*(3-4), 230–240, doi:10.2166/nh.2010.115.

Maraun, D., et al. (2010), Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, *48*, doi:10.1029/2009RG000314.

935 Maurer, E. P. (2007), Uncertainty in hydrologic impacts of climate change in the Sierra
936 Nevada, California, under two emissions scenarios, *Clim. Change*, *82*(3-4), 309–325,
937 doi:10.1007/s10584-006-9180-9.
938 May, W. (2008), Climatic changes associated with a global "2 degrees C-stabilization"
939 scenario simulated by the ECHAM5/MPI-OM coupled climate model, *Clim. Dynam.*,
940 *31*(2-3), 283–313, doi:10.1007/s00382-007-0352-8.
941 Merritt, W., Y. Alila, M. Barton, B. Taylor, S. Cohen, and D. Neilsen (2006), Hydrologic
942 response to scenarios of climate change in sub watersheds of the Okanagan basin, British
943 Columbia, *J. Hydrol.*, *326*(1-4), 79–108, doi:10.1016/j.jhydrol.2005.10.025.
944 Mezghani, A. (2009), Génération multi-sites de variables météorologiques horaires en zone
945 alpine. Application á la simulation de scénarios de crues du Rhône supérieur suisse.,
946 Ph.D. thesis, Ecole Polytechnique Fdrale de Lausanne.
947 Mezghani, A., and B. Hingray (2009), A combined downscaling-disaggregation weather
948 generator for stochastic generation of multisite hourly weather variables over complex
949 terrain: Development and multi-scale validation for the Upper Rhone River basin, *J.*
950 *Hydrol.*, *377*(3-4), 245–260, doi:10.1016/j.jhydrol.2009.08.033.
951 Michelangeli, P., R. Vautard, and B. Legras (1995), Weather regimes - recurrence and
952 quasi stationnarity, *J. Atmos. Sci.*, *52*(8), 1237–1256.
953 Minville, M., F. Brissette, and R. Leconte (2008), Uncertainty of the impact of cli-
954 mate change on the hydrology of a nordic watershed, *J. Hydrol.*, *358*(1-2), 70–83,
955 doi:10.1016/j.jhydrol.2008.05.033.
956 Moore, K., D. Pierson, K. Pettersson, E. Schneiderman, and P. Samuelsson (2008), Effects
957 of warmer world scenarios on hydrologic inputs to Lake Malaren, Sweden and implica-

tions for nutrient loads, *Hydrobiologia*, *599*, 191–199, doi:10.1007/s10750-007-9197-8.

Moulin-Ollagnier, J. (1985), Ergodic Theory and Statistical Mechanics, in *Lecture Notes in Mathematics, 1115*, Springer-Verlag, Berlin, Heidelberg, New-York, Tokyo.

Nakicenovic, N., et al. (2001), A special report of IPCC Working Group III : Emissions scenarios, *Tech. rep.*, IPCC.

Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models, 1, a discussion of principles, *J. Hydrol.*, *10*(3), 282–290.

Obled, C., G. Bontron, and R. Garcon (2002), Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, *Atmos. Res.*, *63*(3-4), 303–324.

Pohl, S., P. Marsh, and B. R. Bonsal (2007), Modeling the impact of climate change on runoff and annual water balance of an Arctic headwater basin, *Arctic*, *60*(2), 173–186.

Raisanen, J. (2001), CO2-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability, *J. Climate*, *14*(9), 2088–2104, doi: 10.1175/1520-0442(2001)014<2088:CICCIC>2.0.CO;2.

Raisanen, J. (2008), Warmer climate: less or more snow?, *Clim. Dynam.*, *30*(2-3), 307–319, doi:10.1007/s00382-007-0289-y.

Raje, D., and P. P. Mujumdar (2010), Constraining uncertainty in regional hydrologic impacts of climate change: Nonstationarity in downscaling, *Water Resour. Res.*, *46*, doi:10.1029/2009WR008425.

Rasmus, S., J. Raisanen, and M. Lehning (2004), Estimating snow conditions in Finland in the late 21st century using the SNOWPACK model with regional climate scenario data as input, *Ann. Glaciol.*, *38*, 238–244.

Salas, D., F. Chauvin, M. Dqu, H. Douville, J. Gurmy, P. Marquet, S. Planton, J. Royer, and S. Tyteca (2005), Description and validation of CNRM-CM3 Global Coupled Model., *Working Note 103*, CNRM, Available at http://www.cnrm.meteo.fr/scenario2004/paper_cm3.pdf.

Sanchez-Gomez, E., S. Somot, and M. Deque (2009), Ability of an ensemble of regional climate models to reproduce weather regimes over Europe-Atlantic during the period 1961-2000, *Clim. Dynam.*, *33*(5), 723–736, doi:10.1007/s00382-008-0502-7.

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, *J. Climate*, *26*(12), 4017–4037, doi:10.1175/JCLI-D-12-00462.1.

Santer, B. D., et al. (2008), Consistency of modelled and observed temperature trends in the tropical troposphere, *Int. J. Climatol.*, *28*(13), 1703–1722, doi:10.1002/joc.1756.

Schäfli, B., B. Hingray, and A. Musy (2007), Climate change and hydropower production in the Swiss Alps : quantification of potential impacts and related modelling uncertainties, *Hydrology and Earth System Sciences*, *11*, 1191–1205.

Senatore, A., G. Mendicino, G. Smiatek, and H. Kunstmann (2011), Regional climate change projections and hydrological impact analysis for a Mediterranean basin in Southern Italy, *J. Hydrol.*, *399*(1-2), 70–92, doi:10.1016/j.jhydrol.2010.12.035.

Singh, P., and L. Bengtsson (2004), Hydrological sensitivity of a large Himalayan basin to climate change, *Hydrol. Process.*, *18*(13), 2363–2385, doi:10.1002/hyp.1468.

Singh, P., and L. Bengtsson (2005), Impact of warmer climate on melt and evaporation for the rainfed, snowfed and glacierfed basins in the Himalayan region, *J. Hydrol.*, *300*(1-4),

140–154, doi:10.1016/j.jhydrol.2004.06.005.

Somura, H., J. Arnold, D. Hoffman, I. Takeda, Y. Mori, and M. Di Luzio (2009), Impact of climate change on the Hii River basin and salinity in Lake Shinji: a case study using the SWAT model and a regression curve, *Hydrol. Process.*, *23*(13), 1887–1900, doi:10.1002/hyp.7321.

Stahl, K., R. D. Moore, J. M. Shea, D. Hutchinson, and A. J. Cannon (2008), Coupled modelling of glacier and streamflow response to future climate scenarios, *Water Resour. Res.*, *44*(2), doi:10.1029/2007WR005956.

Sterl, A. (2004), On the (in)homogeneity of reanalysis products, *J. Climate*, *17*(19), 3866–3873, doi:10.1175/1520-0442(2004)017<3866:OTIORP>2.0.CO;2.

Sturaro, G. (2003), A closer look at the climatological discontinuities present in the NCEP/NCAR reanalysis temperature due to the introduction of satellite data, *Clim. Dynam.*, *21*(3-4), 309–316, doi:10.1007/s00382-003-0334-4.

Sultana, Z., and P. Coulibaly (2011), Distributed modelling of future changes in hydrological processes of Spencer Creek watershed, *Hydrol. Process.*, *25*(8), 1254–1270, doi:10.1002/hyp.7891.

Teutschbein, C., and J. Seibert (2012), Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *J. Hydrol.*, *456*, 12–29, doi:10.1016/j.jhydrol.2012.05.052.

Teutschbein, C., F. Wetterhall, and J. Seibert (2011), Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale, *Clim. Dynam.*, *37*(9-10), 2087–2105, doi:10.1007/s00382-010-0979-8.

Teweles, S., and H. Wobus (1954), Verification of prognostic charts, *Bull. Amer. Meteorol. Soc.*, *35*(D7), 455–463.

Tisseuil, C., M. Vrac, S. Lek, and A. J. Wade (2010), Statistical downscaling of river flows, *J. Hydrol.*, *385*(1-4), 279–291, doi:10.1016/j.jhydrol.2010.02.030.

Uppala, S., et al. (2005), The ERA-40 re-analysis, *Quart. J. Roy. Meteorol. Soc.*, *131*(612, Part B), 2961–3012, doi:10.1256/qj.04.176.

Veijalainen, N., T. Dubrovin, M. Marttunen, and B. Vehvilainen (2010), Climate Change Impacts on Water Resources and Lake Regulation in the Vuoksi Watershed in Finland, *Water Resour. Res.*, *24*(13), 3437–3459, doi:10.1007/s11269-010-9614-z.

Vicuna, S., R. D. Garreaud, and J. McPhee (2011), Climate change impacts on the hydrology of a snowmelt driven basin in semiarid Chile, *Clim. Change*, *105*(3-4), 469–488, doi:10.1007/s10584-010-9888-4.

Vidal, J.-P., E. Martin, L. Franchisteguy, M. Baillon, and J.-M. Soubeyroux (2010), A 50-year high-resolution atmospheric reanalysis over France with the Safran system, *Int. J. Climatol.*, *30*(11), 1627–1644, doi:10.1002/joc.2003.

Weber, M., L. Braun, W. Mauser, and M. Prasch (2010), Contribution of rain, snow- and icemelt in the Upper Danube discharge today and in the future, *Geografia fisica e dinamica quaternaria*, *33*(2), 221–230.

Wiley, M. W., and R. N. Palmer (2008), Estimating the impacts and uncertainty of climate change on a municipal water supply system, *J. Water Res. Pl.-ASCE*, *134*(3), 239–246, doi:10.1061/(ASCE)0733-9496(2008)134:3(239).

Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins (2011), A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions, *J. Climate*, *24*(17),

4634–4643, doi:10.1175/2011JCLI4085.1.

Young, C. A., et al. (2009), Modeling the Hydrology of Climate Change in California's Sierra Nevada for Subwatershed Scale Adaptation1, *J. Am. Water Resour. Assoc.*, *45*(6), 1409–1423, doi:10.1111/j.1752-1688.2009.00375.x.

Zierl, B., and H. Bugmann (2005), Global change impacts on hydrological processes in Alpine catchments, *Water Resour. Res.*, *41*(2), doi:10.1029/2004WR003447.

Appendix A

The Quasi-Ergodic ANOVA framework is based on the quasi-ergodic assumptions for climate simulation in transient climate. Ergodicity applies to a dynamic system for which any sequence of events will pass through all the values accessible to the system, provided the duration of this sequence is sufficiently long [*Moulin-Ollagnier*, 1985]. Under the ergodic assumption, the statistics are expected to be invariant whether calculated over the ensemble of events for a given time or over time for a given sequence of events.

For any projection lead time, the climate response of a given simulation chain is classically defined as the ensemble mean of all possible replicates of the chain. The Quasi-Ergodic ANOVA (QE-ANOVA) framework considers that if the climate response of a particular simulation chain varies over the period, this variation should be gradual and smooth, the higher frequency variations of the time series being due to internal variability alone. It assumes also that the internal variability remains constant over the period for temperature or that it varies as a linear function of the mean effect of the simulation chain for the other variables, for which relative instead of absolute changes are studied. These assumptions correspond to the quasi-ergodic assumption for climate simulation in a transient climate.

In this context, partitioning model uncertainty and internal variability components for any projection lead time then simply consists in extracting the noise-free signal from the time series of each simulation chain. The noise-free signal defines the climate change response of the chain and its possible evolution with time. The climate change responses of all chains from the multimodel ensemble of projections can then be used to estimate the components of model uncertainty associated with the different models of the chains, i.e. uncertainty associated with GCMs and SDMs. It is additionally possible to estimate the uncertainty associated to the GCM-dependent deviations of each SDM. In the present case, these deviations combine the effects of systematic interaction between GCMs and SDMs plus the effects of random residuals [*Hingray and Saïd*, 2014]. This component is here therefore refered to as the residual / model interaction component of model uncertainty.

The noise around the climate change response of each chain defines the internal variability of the chain. In the QE-ANOVA framework, all runs and replicates available from the ensemble of projections can be accounted for, even when the number of runs and replicates is not the same for the different chains. In the present case, a large number of members is available for each chain ($100 * k$ where k is the number of GCM runs available for the chain). Simulations additionally cover a period as long as 240 years. This large dataset allows partitioning of both large and small scale components of internal variability. The estimation of the climate change response and the internal variability of the chain is in this case expected to be relatively robust even if internal variability is significant compared to the inter-chain dispersion of the climate change responses.

Table 1. Future hydrological projections in regions influenced by snow. Number of Greenhouse Gas Emissions Scenarios (GGES), General Circulation Models (GCM), Regional Circulation Models (RCM), Statistical Downscaling Models (SDM) and Hydrological Models (HM) used to assess the uncertainties.

Authors	GGES	GCM	RCM	SDM	HM	Basin/Region	Area (km^2)
Harding et al. [2012]	**3**	**16**		1	1	Colorado (USA)	600000
Immerzeel et al. [2012]	1	**5**		1	1	Langtang (Nepal)	360
Das et al. [2011]	1	**3**		1	1	Sierra Nevada (USA)	46080
Chen et al. [2011]	**2**	**6**		**4**	**3**	Manicouagan (Quebec)	24610
Teutschbein et al. [2011]	**2**	**2**		**3**	1	Vattolmaan (Sweden)	293
Grillakis et al. [2011]	1	**3**	**4**	1	**3**	Spencer Creek (Ontario, Canada)	160
Sultana and Coulibaly [2011]	1	1		**2**	1	Spencer Creek (Ontario, Canada)	291
Senatore et al. [2011]	**2**	**2**	**3**	1	1	Crati (Calabre, Italy)	1332
Kerkhoven and Gan [2011]	**4**	**7**		1	1	Athabasca et Fraser (Rocky M., Canada)	133000-230000
Kingston et al. [2011]	**8**	**7**		1	1	Mekong (China)	795000
Larson et al. [2011]	**2**	**2**		1	1	Ste Marie (Montana, Canada)	544
MacDonald et al. [2011]	**3**	**3**		1	1	Ste Marie (Montana, Canada)	544
Vicuna et al. [2011]	**2**	1	1	1	1	Limarí (Chile)	238-656
Jung and Chang [2011]	**2**	**8**		1	1	Willamette (Oregon, USA)	62-29700
Chang and Jung [2010]	**2**	**8**		1	1	Willamette (Oregon, USA)	62-29700
Veijalainen et al. [2010]	**3**	**3**	**2**	**2**	1	Vuoksi (Finland)	2430-61000
Tisseuil et al. [2010]	**2**	1		**4**		51 basins, France	10-51500
Driessen et al. [2010]	**3**	1	1	1	1	Ourthe (Belgium)	16
Boyer et al. [2010]	**2**	**3**		1	1	5 basins, Quebec	269-22000
Liu et al. [2010]	**2**	1		1	1	Tarim (China)	362000
Weber et al. [2010]	1	1	1	1	1	Upper Danube and tributaries (Switzerland, Germany)	165 - 77000
Cherkauer and Sinha [2010]	**3**	**2**		1	1	Michigan Lake (USA)	500-2000
Magnusson et al. [2010]	**2**	**2**	**6**	1	1	Dischma et Inn (Valais, Switzerland)	43-1945
Bavay et al. [2009]	**2**	**2**	**6**	1	1	Dischma et Inn (Valais, Switzerland)	43-1945
Somura et al. [2009]	1	1		1	1	Hii (Japan)	920
Bethers and Sennikovs [2009]	1	**7**	1	1	2	4 basins, Latvia	904-3220
Young et al. [2009]				1	1	Sierra Nevada, California	80000
Boé et al. [2009]	1	**14**	1	2	1	all France	500-100000
Coulibaly [2009]	1	1		**3**	**3**	Serpent (Quebec)	1760

Table 1. (continued)

Authors	GGES	GCM	RCM	SDM	HM	Basin/Region	Area (km^2)
Huss et al. [2008]	**2**	**3**	**9**	1	1	Zinal (Valais, Switzerland)	7-18
Dankers and Feyen [2008]	1	1	1		1	more than 200 basins in Europe	
Arora et al. [2008]				1	1	Chenab (India)	22400
Minville et al. [2008]	**2**	**5**		1	1	Chute du Diable (Quebec)	9700
Li et al. [2008]	**2**	**7**		1	1	Fleuve Jaune (China)	121000
Moore et al. [2008]	**2**	**2**	**2**	1	1	Mlaren Lake (Sweden)	8500
Wiley and Palmer [2008]	1	**7**		1	1	Cedar and Tolt (near Seattle, USA)	
Stahl et al. [2008]	**2**	1		1	1	Bridge (Brittanic Columbia, Canada)	152
Raisanen [2008]	1	**20**				global scale snow cover	
Lopez-Moreno et al. [2008]	1	1	**9**	1	1	Izas (Pyrenees, Spain)	loc.
Maurer [2007]	**2**	**11**		1	1	4 basins in California (USA)	4000-9350
Hreiche et al. [2007]				1	1	Nahr Ibrahim (Lebanon)	341
Caballero et al. [2007]	1	**7**		1	1	Adour-Garonne (France)	116000
Baltas [2007]	**2**	1		1	1	Sialista (Greece)	2700
Pohl et al. [2007]	**2**	**7**		1	1	Trail Valley Creek (Arctic)	
Hingray et al. [2007]	**2**	**3**	**9**	1	1	Aare, Broye, Orbe (Switzerland)	8300
Schäfli et al. [2007]	**2**	**3**	**9**	1	1	Mauvoisin (Switzerland)	169
Graves and Chang [2007]	1	**2**		1	1	Clackamas (Oregon, USA)	1260
Juen et al. [2007]	**4**	mean		1	1	Llanganuco (Peru)	86
Horton et al. [2006]	**2**	**3**	**9**	1	1	11 basins, Switzerland	39-185
Merritt et al. [2006]	**2**	**3**		1	1	Okanagan (Brittanic Columbia, Canada)	8000
Dankers and Christensen [2005]	1	1	1	1	1	Tana (Finland)	16000
Zierl and Bugmann [2005]	**4**	**4**		1	1	5 basins, Switzerland	44-77
Singh and Bengtsson [2004, 2005]				1	1	Satluj (India)	22275
Dibike and Coulibaly [2005]	1	1		**2**	**2**	Chute du Diable (Quebec)	9700
Keller et al. [2005]				1	1	Sntis (Switzerland)	0.25
Kunstmann et al. [2004]	1	1	1	1	1	Ammer (Germany)	710
Jasper et al. [2004]	**4**	**7**		**2**	1	Thur and Ticino (Switzerland)	1500
Rasmus et al. [2004]	1	1	1	1	1	6 local sites in Finland	loc.

Table 2. GCM experiments for 20C3M+SRES-A1B (1860-2099) from the ENSEMBLES STREAM2 EU research project [*Johns et al.*, 2011]. Amospheric (AR) and Oceanic (OR) Resolutions.

Institute	Model	Run	Acronym	AR	OR	Reference
CNRM (France)	CNRM-CM3.3	1	CNCM33-1	T63L31	2° L31	*Salas et al.* [2005]
DMI (Danemark)	ECHAM5-C	1 2 3	DMIEH5C-1 DMIEH5C-2 DMIEH5C-3	T31L19	3° L40	*May* [2008]
FUB (Allemagne)	EGMAM2	3	EGMAM2-3	T30L39	0.5/2.8° L20	*Huebener et al.* [2007]
IPSL (France)	IPSL-CM4_v2	1 2 3	IPCM4-1 IPCM4-2 IPCM4-3	N72L19	2° L31	*Dufresne et al.* [2005]
METO-HC (Roy. Uni)	HadGEM2-AO	1	HADGEM2-1	N96L38	1° L40	*Johns et al.* [2006]
MPI (Allemagne)	ECHAM5-C	1 2 3	MPEH5C-1 MPEH5C-2 MPEH5C-3	T31L19	3° L40	*Brasseur and Roeckner* [2005]

Table 3. Statistical Downscaling Models used in this study. Predictors used by the models are P_{SL} (sea level pressure), Z (geopotential height for the pressure level), T_a (surface temperature), u and v (zonal and meridian geostrophic wind), HU (relative humidity) and F_q (moisture flux). Subscripts indicate the pressure levels. One hundred generations were carried out for each GCM/SDM chain.

Model	Institute	Method	Predictors	Reference
ANALOG	EDF/LTHE	k nearest neighbors	Z_{700}, Z_{1000}	*Obled et al.* [2002]
DSCLIM-10	CERFACS	Weather types + transfer function	P_{SL}	*Boé et al.* [2006], *Lafaysse* [2011]
DSCLIM-11			P_{SL}, T_a	
DSCLIM-21			Z_{850}, Z_{500}	
D2GEN-10	LTHE	Transfer function + k nearest neighbors	P_{SL}, u_{700}, v_{700}	*Mezghani and Hingray* [2009]
D2GEN-32			P_{SL}, u_{700}, v_{700}, HU_{700}, F_{q700}	

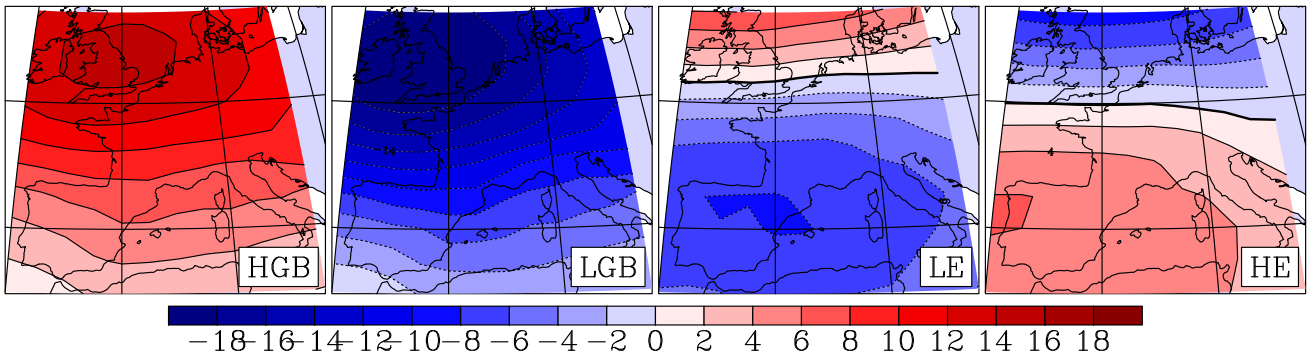

Figure 1. Sea level pressure composites for the 4 identified weather types in the ERA40 reanalysis in winter (DJF). LGB corresponds to Low pressure over Great Britain, HGB to High pressure over Great Britain, LE to Low pressure over Iberian peninsula, HE to High pressure over Iberian peninsula.

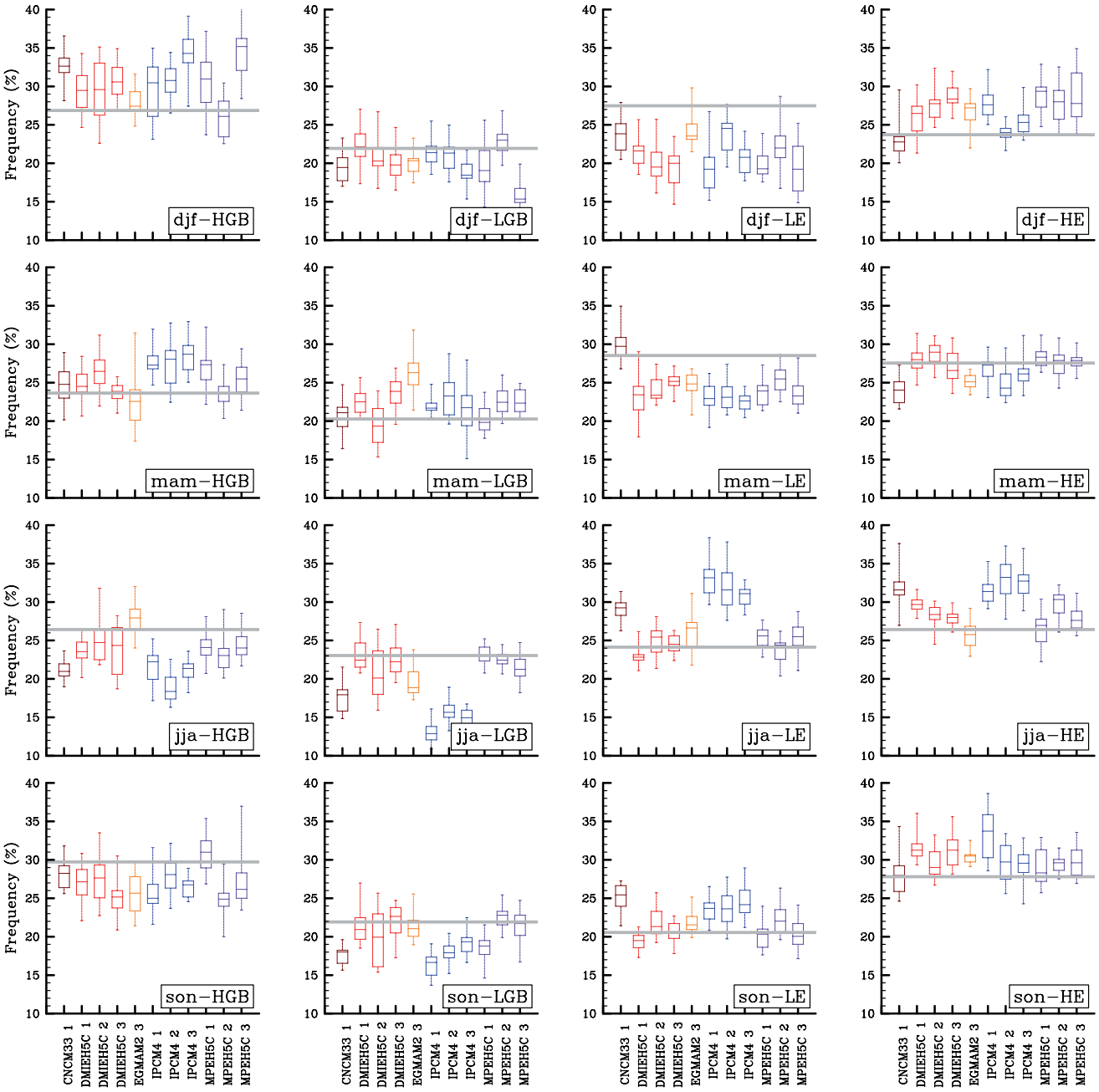

Figure 2. Observed occurrency frequency of weather types for all seasons (gray lines), 1960-1979. Boxplots for the simulated occurrency frequencies by the 12 climatic experiments in several 20-years periods between 1860 and 1979 (minimum, 25th percentile, median, 75th percentile, and maximum among all sub-periods).

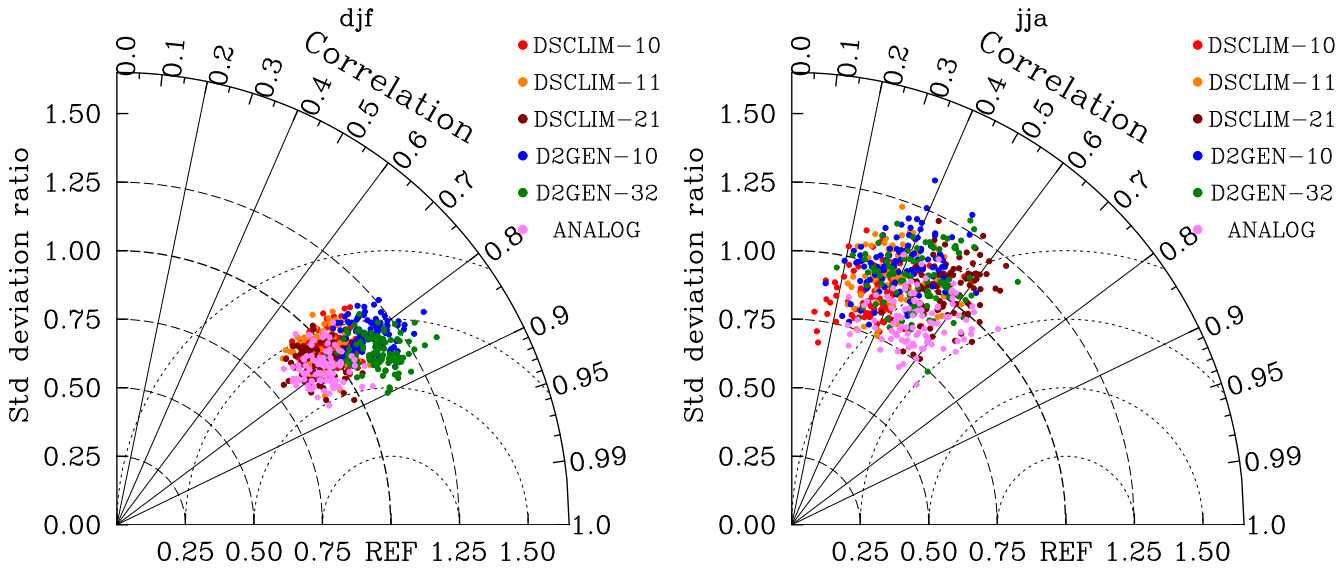

Figure 3. Taylor diagrams for the annual time series of seasonal basin precipitation (1959-2006). Correlation coefficient between simulated and observed time series and ratio between simulated and observed standard deviation. The colors correspond to the different SDMs, with 100 points per SDM (100 replicates per SDM). The 2 diagrams correspond to winter and summer (DJF, JJA).

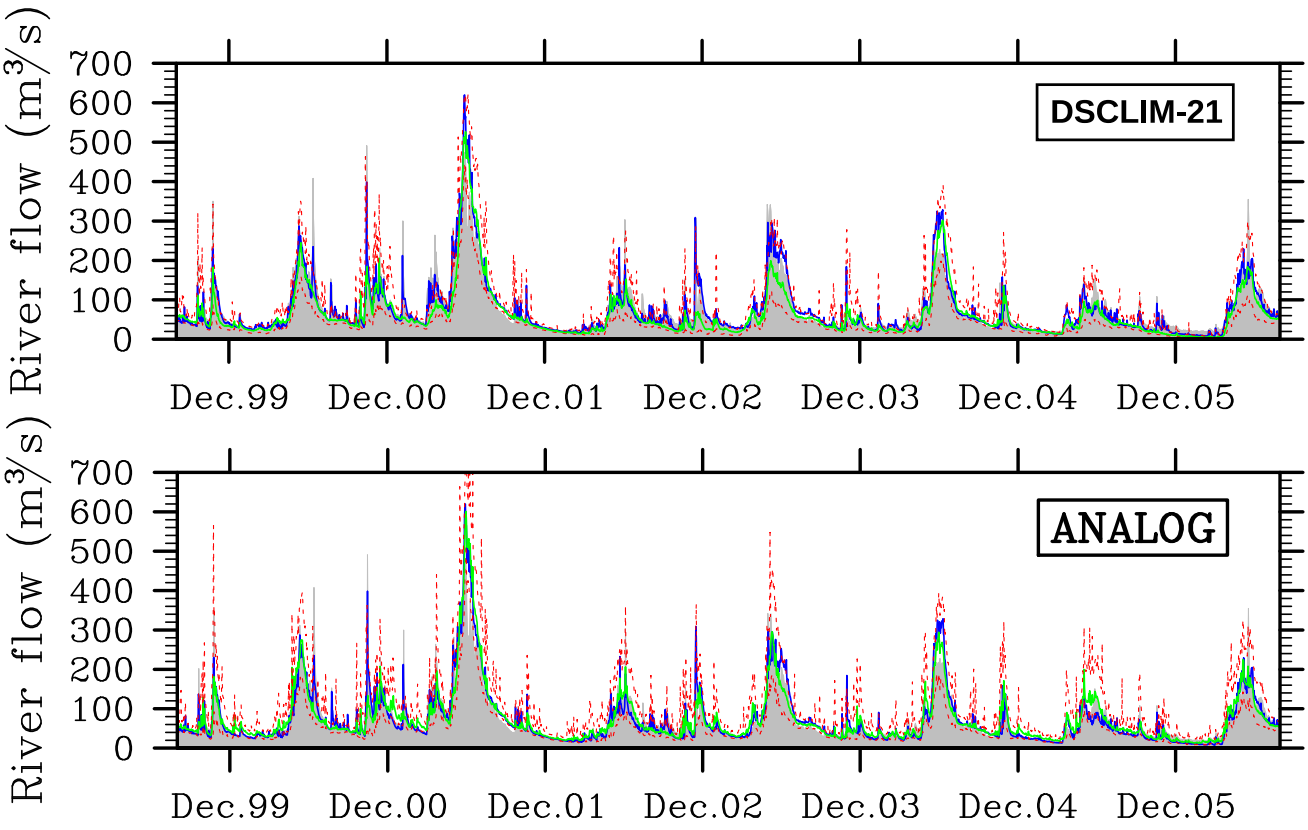

Figure 4. Observed (gray) discharges and simulated daily river discharges by forcing the hydrological model with meteorological observations (blue) or with 100 meteorological replicates (5th percentile, median and 95th percentile: red, green, red) from statistical downscaling models (DSCLIM-21 for top and ANALOG for bottom graph).

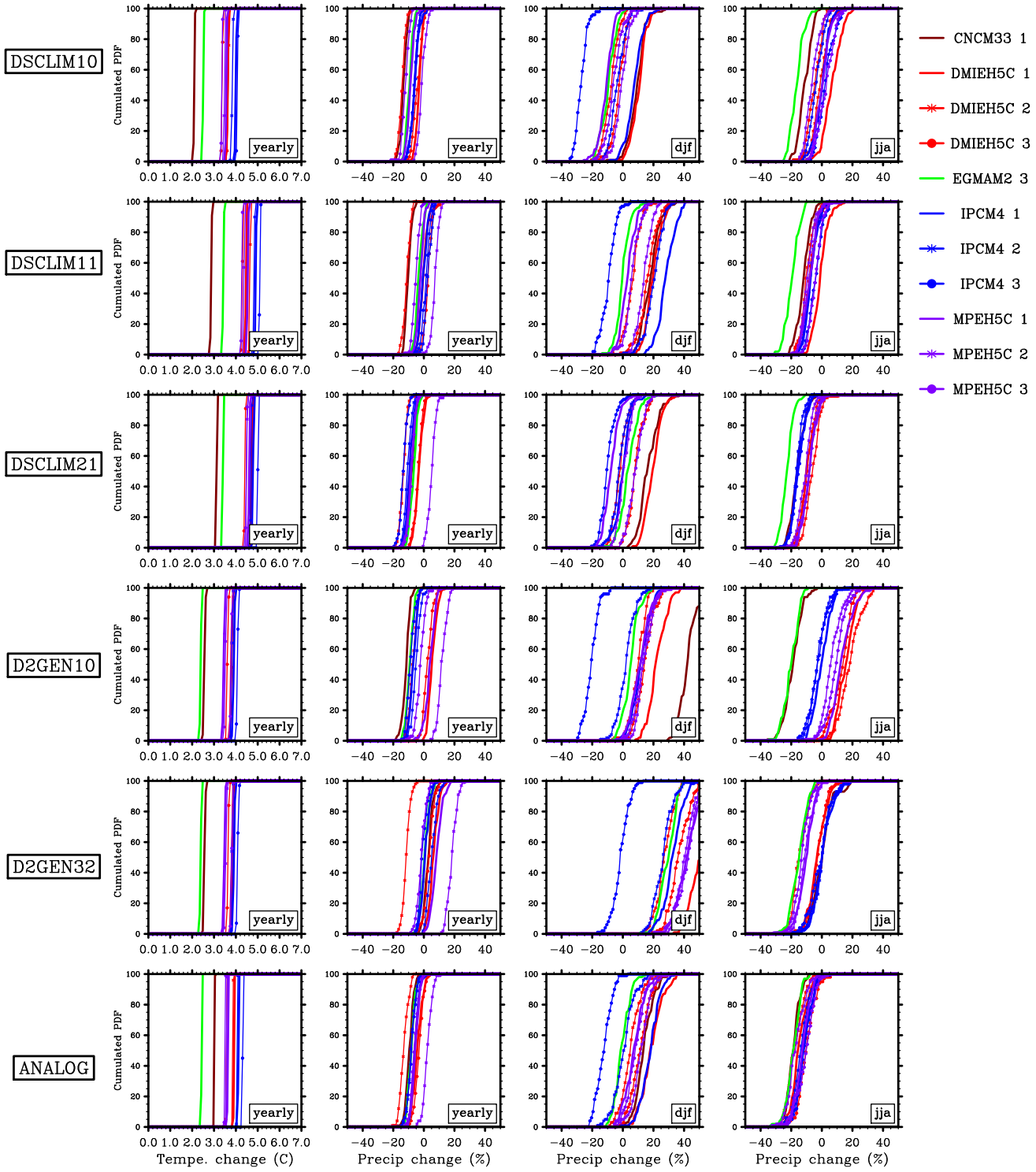

Figure 5. Cumulative distribution functions (CDFs) of mean changes in annual temperature (°C) and mean changes in annual, winter (djf) and summer (jja) precipitation (%) between 2080-2099 and 1980-1999, for all GCMs (different colors) and SDMs (different plots). CDFs correspond to the 100 replicates generated with each GCM/SDM modeling chain.

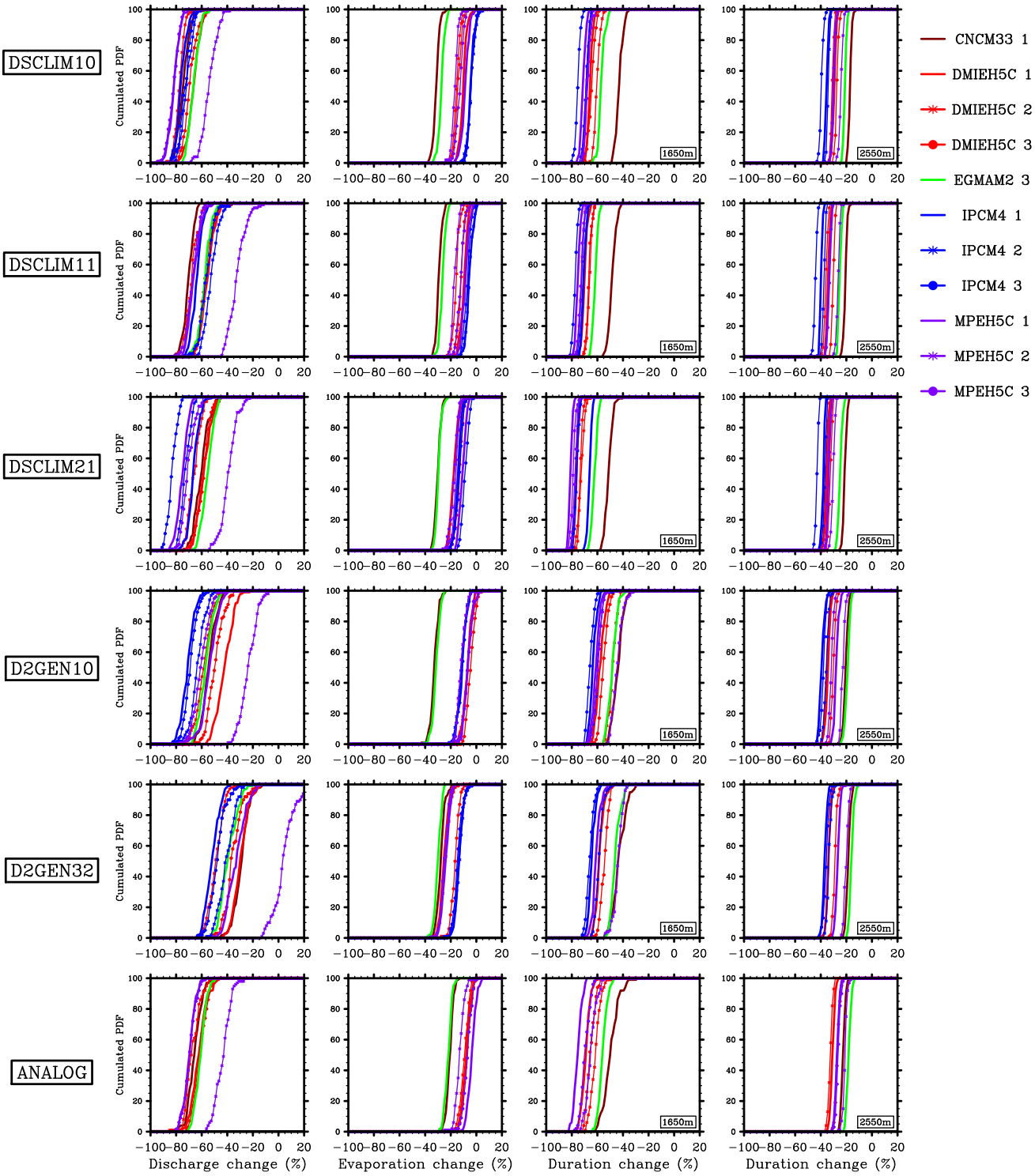

Figure 6. Same as figure 5 for mean changes (%) in annual discharge, annual evaporation and annual snow cover duration (1650 m; 2550 m). See figure 5 for caption details.

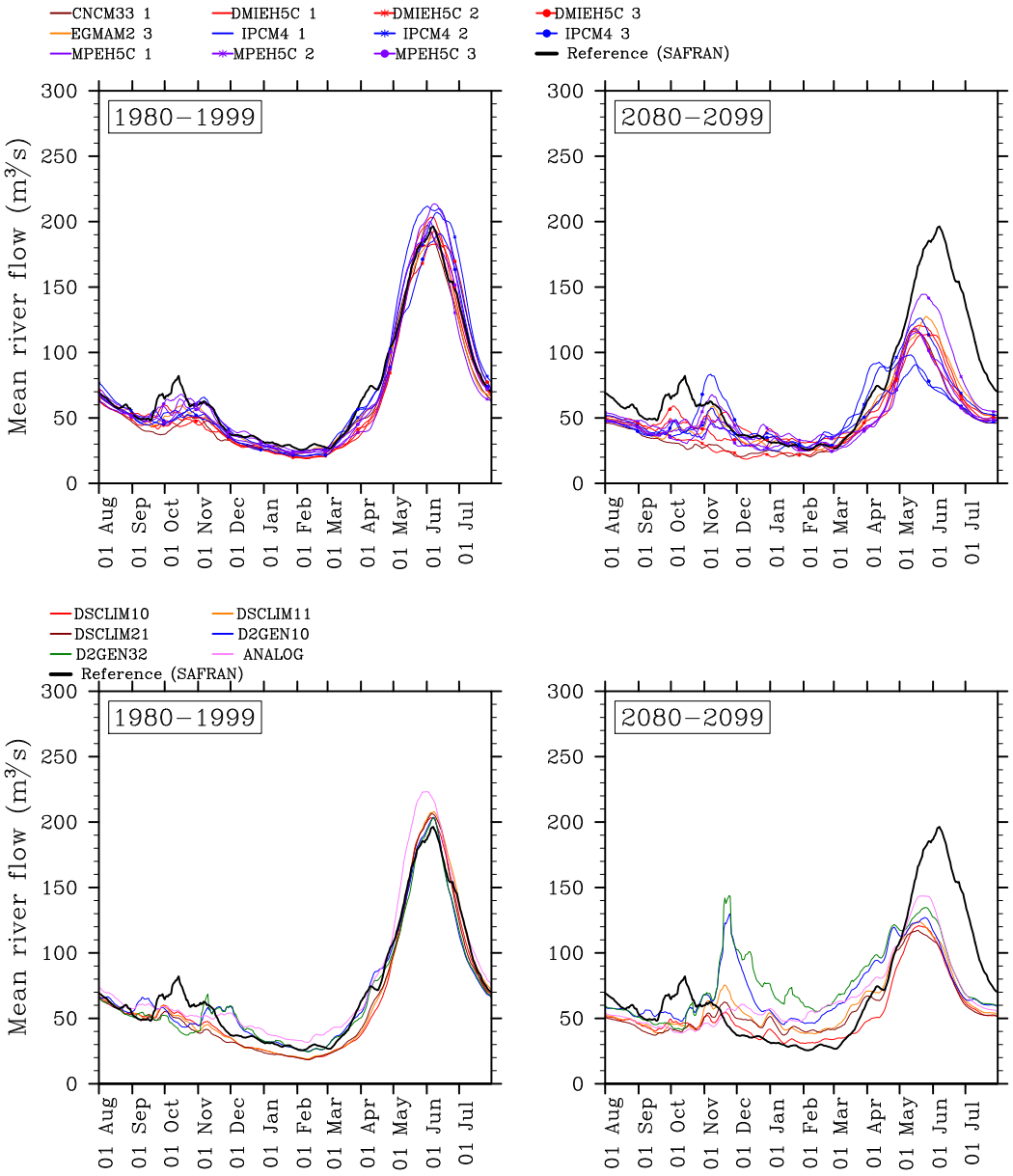

Figure 7. Mean seasonal cycle of simulated river flows for 2 periods (1980-1999, 2080-2099) ; Each curve is the multireplicate median cycle obtained from the 100 replicates generated for a particular GCM/SDM chain. Top: Cycles obtained with DSCLIM-10 forced respectively by the 11 GCMs. Bottom: Cycles obtained with the 6 SDMs forced by run 1 of DMIEH5C.

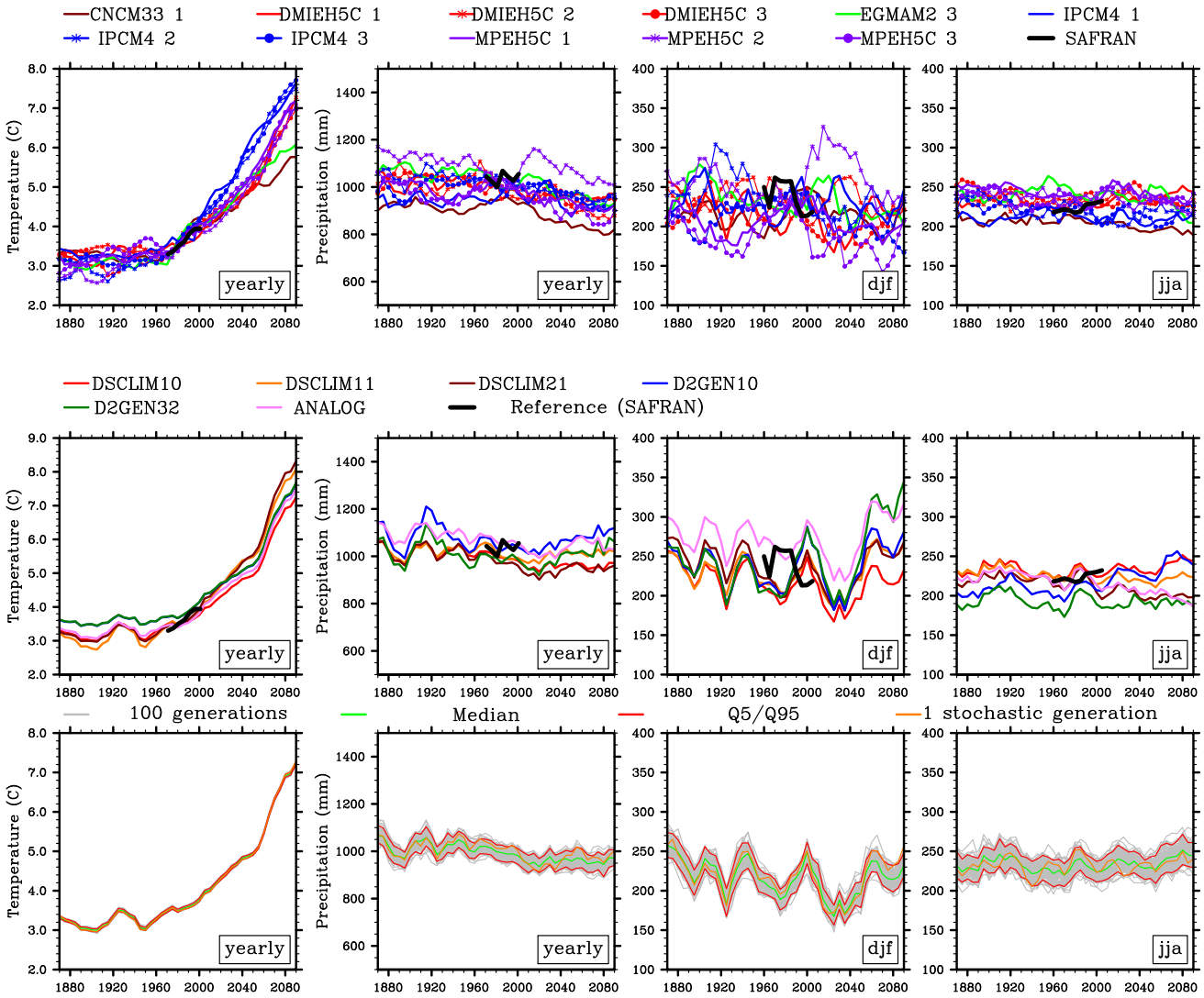

Figure 8. Evolution of the 20-year running mean of simulated annual temperature (°C), annual, winter (djf) and summer (jja) precipitation (mm) over the period 1860-2099 for a selection of scenarios. In the top and middle panels, each time trajectory is that of the multireplicate median value obtained for each 20-yr period from the 100 replicates simulated for a particular GCM/SDM chain. Top: Median trajectory of DSCLIM-10 forced by all GCMs. Middle: Median scenario of all SDMs forced by run 1 of DMIEH5C. Bottom: Median, 5th and 95th percentiles of the 100 replicates of DMIEH5C-1/DSCLIM-10 modeling chain. The years indicated on the time axis are the midpoints of the corresponding averaging future period.

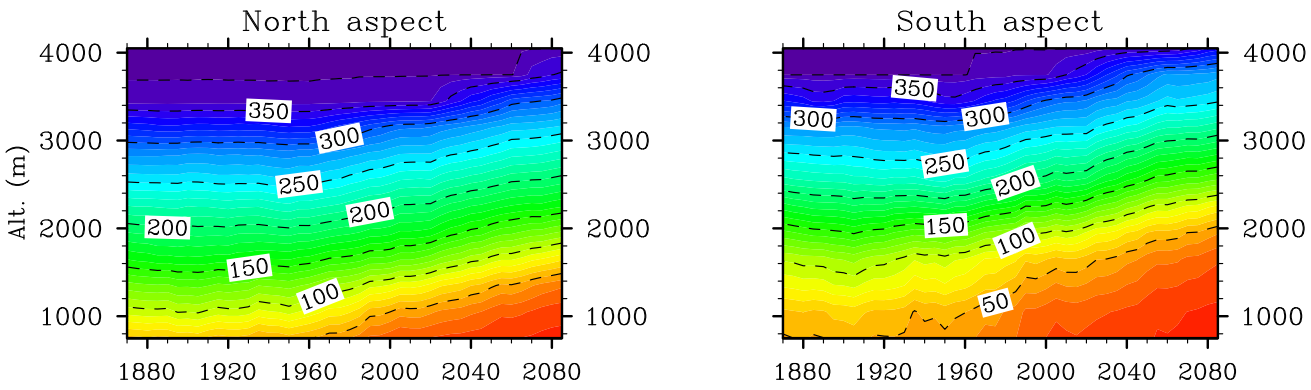

Figure 9. Evolution of the 20-year running mean of simulated snow cover duration over the period 1860-2099, for the median hydrological scenario forced by EGMAM2-3 and DSCLIM-11, as a function of elevation for North and South aspects.

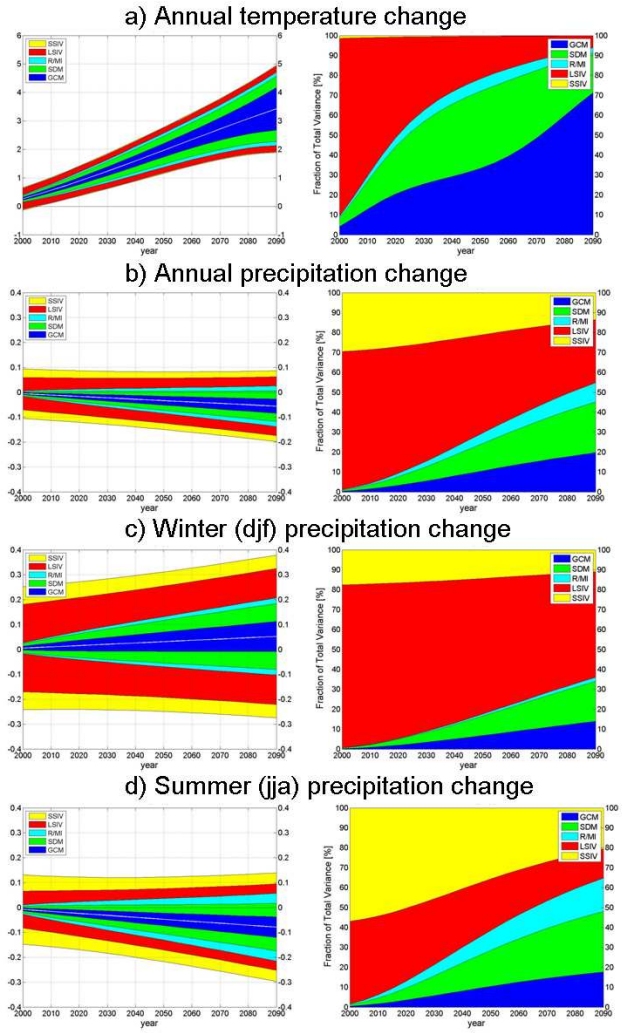

Figure 10. Left: multichain mean climate change signal $\mu(t)$ and total uncertainty in RIWER2030 projections for 20-yr mean annual temperature, annual, winter and summer precipitation as a function of prediction lead time (reference period: 1980-1999). The total colored area covered by all uncertainty components corresponds to $\mu(t) \pm 1.645\sqrt{T(t)}$ where $T(t)$ is the total uncertainty variance. 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). Right: fraction of total uncertainty variance explained by each uncertainty component. Dark blue: GCM uncertainty, green: SDM uncertainty, cyan: residual / GCM-SDM interaction (R/MI), red: large scale internal variability (LSIV), yellow: small scale internal variability (SSIV).

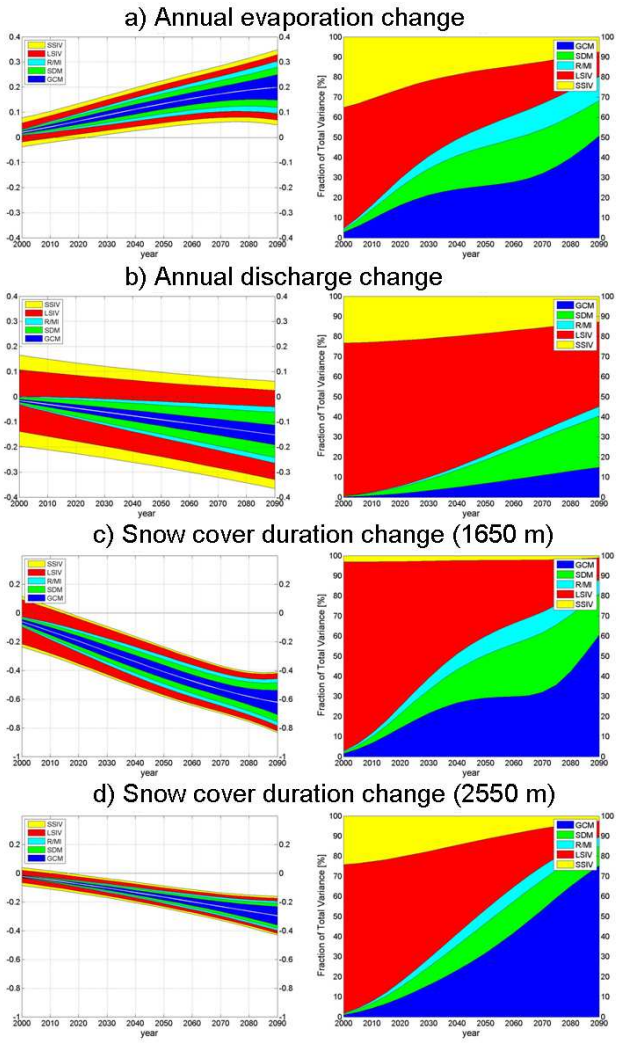

Figure 11. The same as figure 10 for projections of changes in evaporation, mean annual discharge, and Snow Cover Duration for two elevation ranges (1500-1800 m and 2400-2700 m). See Figure 10 for caption details