Clim Dyn
DOI 10.1007/s00382-006-0160-6

Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations

B. Decharme · H. Douville

Received: 21 November 2005 / Accepted: 18 May 2006

Abstract The Global Soil Wetness Project (GSWP) is an international initiative aimed at producing global data sets of soil wetness and energy and water fluxes by driving land surface models with state-of-the-art 1° by 1° atmospheric forcing and land surface parameters. It also provides a unique opportunity to develop and test land surface parameterizations at the global scale, using multi-year off-line simulations that are not affected by the systematic errors found in atmospheric models. Nevertheless, the accuracy and reliability of the 10–year GSWP-2 atmospheric forcing remain questionable. A first comparison using the high-resolution Rhône-AGGregation (Rhône-AGG) database reveals that the baseline GSWP-2 precipitation forcing is drastically overestimated over the Rhône river basin. Hydrological simulations driven with each dataset and using the ISBA land surface model and the MODCOU river routing model are also compared. The simulated river discharges are validated against a dense network of river gauges and are generally less realistic when using the GSWP-2 instead of the Rhône-AGG precipitation forcing. Secondly, the GSWP-2 precipitation forcing is compared with three alternative data sets (GPCP-2, CRU-2, CMAP) at the global scale. Moreover, the results of a global sensitivity study to the precipitation forcing conducted with six land surface models are shown. The TRIP river routing model is used to convert daily runoff from all models into river discharges, which are compared at 80 gauging stations distributed over the globe. In agreement with the regional evaluation, the results reveal that the baseline GSWP-2 precipitation forcing is generally overestimated over the mid and high latitudes, which implies systematic errors in the simulated discharges. This study reveals that the empirical wind corrections applied to the GSWP-2 precipitation forcing are exaggerated, whereas the GPCP satellite adjustments seem to be useful for simulating realistic annual mean river discharges over the East Siberian river basins.

B. Decharme (✉)
CETP-IPSL-CNRS, 10 avenue de l'Europe,
78140 Vélizy-Villacoublay, France
e-mail: bertrand.decharme@cetp.ipsl.fr

H. Douville (✉)
Météo-France, CNRM/GMGEC/UDC,
Toulouse Cedex1 31057, France
e-mail: herve.douville@meteo.fr

1 Introduction

An important issue in climate modeling with atmospheric general circulation models (AGCMs) is the possible influence of soil wetness on climate variability and predictability (Dirmeyer 2000, 2001; Douville and Chauvin 2000; Douville et al. 2001; Douville 2002, 2003; Koster et al. 2000a, 2002). Soil wetness plays an important role in the global energy and water budget, but is still unknown over most of the globe because in situ measurements are very sparse, and remote sensing techniques are only partially effective. Soil moisture is a key land surface variable because it controls the partitioning of downward radiation and precipitation, affects the status of vegetation, and modulates the soil thermal and hydraulic properties. The space-time variability of soil moisture is usually represented in AGCMs through the use of land surface models

(LSMs). The complexity of these models ranges from the simple bucket model (Manabe 1969) to more sophisticated soil–vegetation–atmosphere transfer (SVAT) schemes with multiple parameterizations representing the physical processes linked to vegetation, soil, and snow.

The Global Soil Wetness Project (GSWP; http://www.iges.org/gswp/) was launched by the global energy and water cycle experiment (GEWEX) to provide high-resolution global soil moisture climatologies (Dirmeyer et al. 1999, 2002) by integrating LSMs in off-line mode using prescribed meteorological forcing, standardized soil and vegetation maps, and a common land-sea mask. In GSWP, the International Satellite Land Surface Climatology Project (ISLSCP) data set is used to supply boundary conditions, model parameters and meteorological forcing for more than twelve LSMs. The aim of the project is: (1) to produce global data sets of land surface fluxes, state variables, and related hydrological quantities; (2) to perform large-scale model intercomparison, validation and calibration; (3) to provide a large-scale validation and quality check of the ISLSCP data sets; (4) to conduct sensitivity studies of specific parameterizations and forcing variables that should aid future model and data developments. The aim of the present study is to address points 3 and 4 based on off-line land surface simulations over the French Rhône river basin.

During the pilot phase of GSWP (GSWP-1) (Dirmeyer et al. 1999), LSM intercomparisons have shown significant differences in the simulated soil moisture fields and in the partitioning of precipitation between total runoff and evapotranspiration (Entin et al. 1999). GSWP-1 has provided an opportunity to test and improve various parameterizations within LSMs at the global scale (Douville 1998; Boone and Wetzel 1999; Dirmeyer and Zeng 1999; Sud and Mocko 1999). A general conclusion is that sub-grid variability in infiltration, whether due to heterogeneity in soil properties and topography, or to the distribution of rainfall within a grid box, has a significant impact on the simulation of total runoff (Dirmeyer et al. 1999).

The Rhône-Aggregation (Rhône-AGG; http://www.cnrm.meteo.fr/mc2/projects/rhoneagg/) LSM intercomparison project was conduced at Météo-France (Boone et al. 2001, 2004). Rhône-AGG was an intermediate step leading up to the continuation of the GSWP project, GSWP-2. It includes a broader investigation of the impact of aggregation on the water budget simulations. The Rhône is the largest European river flowing into the Mediterranean Sea. The basin covers over 95,000 km^2 mostly in southeastern France. This domain was chosen because it contains a large variety of climate types (Mediterranean in the south, temperate in the north, mountainous in the east and the dryer climate in the west). Observed high-resolution soil and vegetation characteristics, subsurface parameters, and atmospheric forcing were mapped onto this domain as part of the GEWEX-Rhône project, which was conceived in recent years by the French research community in order to study the continental water cycle at the regional scale. Results from a series of scaling experiments show that LSMs taking into account land surface and/or atmospheric forcing spatial heterogeneities are able to reduce the scaling influence on the simulated water budget, in general agreement with former studies, for example Lohmann et al. (1998), Wood et al. (1998), Vérant et al. (2004), Decharme and Douville (2006), Decharme et al. (2006).

GSWP-2 encompasses the same core 10-year period as the ISLSCP-II data set (1986–1995). Like in GSWP-1, the spatial resolution is 1° by 1°, but the atmospheric forcing is now provided on a 3-hourly basis. The project also includes various LSM sensitivity studies to uncertainties in forcing data and surface parameters. The relevance of such uncertainties was indeed emphasized by several studies. Using the TRIP river routing model, Oki et al. (1999) revealed that the quality of the ISLSCP-I meteorological forcing data influenced directly the quality of the GSWP-1 runoff. More recently, Fekete et al. (2003) pointed out with a simple water balance model that the uncertainty in precipitation generally translates to at least the same and typically much greater uncertainty in total runoff. Following GSWP-1, a comparison over the Amazon river basin of the ISLSCP-I data set with an alternative precipitation and land surface data set was carried out by Chapelon et al. (2002) using the Interaction soil–biosphere–atmosphere (ISBA) LSM (Noilhan and Planton 1989). This study showed that possible deficiencies in the ISLSCP-I precipitation forcing and soil depth led to an underestimation of the simulated annual runoff, suggesting that uncertainties in the ISLSCP data set could still represent a major obstacle for the validation and improvement of LSM parameterizations over large river basins.

The objective of the present study is twofold. Firstly, the baseline GSWP-2 atmospheric forcing is evaluated over the French Rhône river basin, using the high-resolution Rhône-AGG dataset based on a dense observational network. A series of sensitivity experiments is performed to explore the impact of the GSWP-2 precipitation uncertainties on off-line hydrological simulations based on the ISBA land surface model. Such simulations represent an interesting tool to perform an indirect regional evaluation of the quality of the GSWP-2 data sets over a well-instrumented river catchment. Secondly, the baseline global precipitation forcing provided by GSWP-2 is compared to three other global datasets and evaluated using off-line hydrological simulations from six land surface models. This evaluation is based on the simulated discharge scores over the most important rivers of the globe using the total runoff integrating pathways (TRIP) river routing model (Oki and Sud 1998; http://www.hydro.iis.u-tokyo.ac.jp/~taikan/TRIPDATA/TRIPDATA.html). Both GSWP-2 and Rhône-AGG data sets are presented and compared in Sect. 2. Sections 3 and 4 describe the experimental design and the results of the discharge simulations performed at the regional and global scales respectively. In order to better understand the impact of precipitation uncertainties on the hydrological simulations, an idealized sensitivity test is also conducted over the Rhône basin and analyzed in Sect. 5. Finally, Sect. 6 provides a final discussion and gives the main conclusions of the study.

2 Data presentation and comparison

2.1 Brief overview of the GSWP-2 data set

The baseline meteorological forcing provided by GSWP-2 is based on the ISLSCP-II regridded National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) reanalysis. Corrections to the systematic biases in the 3-hourly reanalysis fields are made by hybridization with global observed monthly climatologies. The baseline total precipitation, *B0*, is a hybrid product of reanalysis, observations, and empirical wind corrections. GSWP-2 uses NCEP/DOE hybrid with Global Precipitation Climatology Centre (GPCC) gauge data for the validation period (1986–1995) and Climate Research Unit (CRU) data for the spin-up period (1982–1985) when GPCC data is not available:

$$P^{3h}_{hybrid} = P^{3h}_{NCEP} P^{m}_{obs} / P^{m}_{NCEP}, \quad (1)$$

where P^{3h}_{NCEP} and P^{m}_{NCEP} are the 3-hourly and monthly NCEP/DOE reanalysis respectively and P^{3h}_{hybrid} the hybridized 3-hourly version of monthly GPCC or CRU data (P^{m}_{obs}). Furthermore, adjustments are made using a ‘‘gauge catch ratio versus wind speed’’ algorithm based on the 3-hourly NCEP/DOE reanalysis of the 10-m wind speed (Zhao and Dirmeyer 2003; Dirmeyer et al. 2006). No detailed information about the precipitation versus wind correction algorithm has been provided by GSWP-2. In the regions of low GPCC rain gauge density (Fig. 1), this product, P^{3h}_{wind}, is combined with the hybridized 3-hourly version of the Global Precipitation Climatology Project (GPCP) data, $P^{3h}_{hybrid\_GPCP}$, that includes both in situ and satellite observations:

$$P^{3h}_{B0} = aP^{3h}_{wind} + (1-a)P^{3h}_{hybrid\_GPCP}$$
$$\left| \begin{array}{ll} a = 1.0 & \forall \text{gauge density} \geq 2, \\ a = 0.5 & \forall \text{gauge density} = 1, \\ a = 0.0 & \forall \text{gauge density} = 0, \end{array} \right. \quad (2)$$

where P^{3h}_{B0} is the 3-hourly baseline precipitation product used in GSWP-2.

Finally, total precipitation is partitioned into snowfall (S_r) and rainfall (R_r) according to a 0°C temperature threshold. Consequently, the wind corrections in the baseline GSWP-2 precipitation product do not depend on precipitation type. An alternative

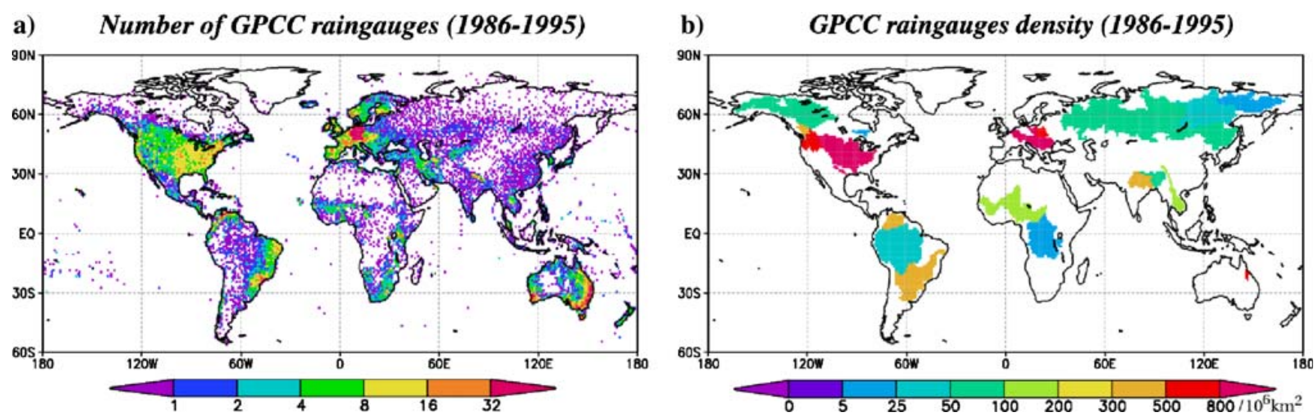

Fig. 1 **a** Mean number of GPCC rain gauges used to prepare the GSWP-2 precipitation forcing data sets over the period 1986–1995. **b** The rain gauges density (/10^6 km^2) for each basin used for the discharge validation is also shown

precipitation product, *P3*, provided by GSWP-2 where the NCEP/DOE reanalysis precipitation is only hybridized with the GPCC or CRU data, is also tested in the present study. According to Eq. 1, this alternative product does not take into account wind corrections and GPCP adjustments. Further details and/or additional informations on the other forcing variables [surface pressure (P_{surf}), 2-m air temperature (T_{air}), specific air humidity (Q_{air}), 10-m wind speed (U), downward shortwave radiation (SW) and downward longwave radiation (LW)] can be found in Zhao and Dirmeyer (2003) and Dirmeyer et al. (2006).

The land surface parameters are also specified according to the ISLSCP-II data set. The vegetation classification and the soil data come from the International Geosphere-Biosphere Programme (IGBP). Global 1° by 1° maps of sand, clay and silt fractions are also provided to derive soil hydraulic and thermal properties from soil texture, while specific vegetation parameters, such as leaf area index or minimum stomatal resistance, are specified for each vegetation category.

2.2 Brief overview of the GEWEX-Rhône project data set

The baseline Rhône-AGG data set is mapped onto the Rhône domain on a 8 km by 8 km grid. In the present study, all atmospheric forcing and surface parameters have been interpolated to the same 1° by 1° resolution as the GSWP2 data set. Note that the spatial interpolation from the high-resolution to the 1° by 1° grid has a direct influence on the quality of the atmospheric forcing. It reduces the grid cell precipitation rates, increases the spatial coverage of a rainfall event and leads to warmer and dryer conditions at lower resolution (Boone et al. 2004).

The original high-resolution atmospheric forcing is calculated using the analysis system for providing atmospheric information relevant to snow (SAFRAN) analysis system (Durand et al. 1993). This data set consists of standard screen-level observations at approximately 60 Météo-France weather network sites within the domain. The data covers over 249 homogeneous climatic zones. The total daily precipitation is analyzed using observational data from over 1,500 gauges, together with a vertical gradient of precipitation (with altitude) derived from climatology. It is then interpolated every 3 h following the daily cycle of the observed air humidity and temperature, the solar radiations, and the weather type (again from climatology). All the forcing variables are available at a 3-hourly time step. Four years of forcing are used in the current study, starting the 1 August 1985, and ending the 31 July 1989. SAFRAN computes the vertical profile of each atmospheric variable every 6 h within each climatic zone. The parameters are then interpolated to 1-h intervals and to the 8 km by 8 km grid. Further details and/or additional information on the other forcing variables can be found in Habets et al. (1999b) and Etchevers et al. (2001). Note that the quality of the SAFRAN analysis system was checked for more than 15 years and that SAFRAN is now used at Météo-France in an operational system for streamflow forecasting over all major French river basins.

Soil and vegetation data are available at the same resolution as the atmospheric forcing. Soil parameters are defined using the soil textural properties from the French National Institute for Agronomical Research (INRA) soil database (King et al. 1995). Vegetation parameters are defined using a vegetation map from the Corine Land Cover Archive and a 2-year satellite archive of the Advanced Very High Resolution Radiometer/Normalized Difference Vegetation Index (AVHRR/NDVI) (Champeaux et al. 2000). More detail can be found in Etchevers et al. (2001).

2.3 Comparison of GSWP-2 with Rhône-AGG

Given the large amount of observations and the high resolution used by the SAFRAN analysis the aggregated 1° by 1° Rhône-AGG data sets can be considered as a reference to evaluate the atmospheric forcing provided by GSWP-2. During Rhône-AGG, the SAFRAN atmospheric forcing was used by 15 land surface models (LSMs) to simulate river discharge over the Rhône river basin. Looking at the downstream Viviers gauging station (observations that were not provided to the modeling groups and can therefore be considered as a verification of the model ability to simulate runoff), the mean annual bias in simulated discharges from all LSMs was only +0.385% with a standard deviation of 6.1% (Boone et al. 2004). This result confirms the good quality of the atmospheric forcing provided by the SAFRAN analysis.

Table 1 comparing the mean values of the forcing in terms of low-level parameters, downward radiation and surface precipitation reveals that five variables are not in good agreement between the GSWP-2 and the Rhône-AGG data sets: the 10-m wind, the snowfall rate, the rainfall rate, the 2-m specific humidity, and the downward longwave radiation. Generally, the baseline GSWP-2 forcing overestimates the snowfall and rainfall rates compared to the Rhône-AGG data. The temporal correlations between both products are relatively poor at the daily and 3-hourly timescale. The overestimation

Table 1 Comparison between the GSWP-2 and Rhône-AGG (R-AGG) atmospheric forcing data sets: 4-year basin average values (*Mean*), $\text{Ratio} = (\text{Mean}_{\text{gswp2}}/\text{Mean}_{\text{r\_agg}} - 1) \times 100(\%)$, correlations (*r*) between both basin average on forcing (*frc*), daily (*day*) and monthly (month) time scale. The precipitation product without wind and satellite corrections is also showed (P3)

Variable description	Symbol	Units	Mean		Ratio (%)	*r*		
			R-AGG	GSWP-2		frc	Day	Month
Rainfall rate								
B0	R_r	mm/day	2.58	3.19	+23.6	0.36	0.66	0.95
P3				2.42	–6.6	0.36	0.67	0.98
Snowfall rate								
B0	S_r	mm/day	0.39	0.59	+50.3	0.37	0.58	0.87
P3				0.33	–18.6	0.37	0.58	0.90
Longwave radiation	LW	W m^2	296.6	312.8	+5.4	0.81	0.89	0.94
Shortwave radiation	SW	W m^2	143.0	142.9	–0.07	0.97	0.96	0.99
Specific humidity at 2 m	Q_{air}	10^{-3} kg/kg	6.09	7.07	+16.0	0.94	0.96	0.99
Wind Speed at 10 m	U	m/s	2.18	5.09	+133.7	0.73	0.78	0.80
Air temperature at 2 m	T_{air}	K	282.2	282.7	+0.15	0.97	0.98	0.99
Surface pressure	P_{surf}	10^3 Pa	93.15	93.58	+0.5	–	–	–

of precipitation is more pronounced during the cold season (Fig. 2b). This is entirely due to a stronger overestimation of wind speed than in the other seasons (Fig. 2e), since the selected algorithm for gauge undercatch does not depend on precipitation type. Furthermore, GPCC gauge network being dense, no GPCP corrections are applied according to Eq. 2. This is confirmed by the use of *P3* precipitation showing a better agreement with the Rhône-AGG data set. Over the four years, the frequency of rainy days (with precipitation intensities above or equal to 1 mm/day) is 47% in *B0*, 43% in *P3*, and only 33% in Rhône-AGG. Looking only at these rainy days (above or equal to 1 mm/day), Fig. 3 reveals that GSWP-2 favors the weak intensities to the detriment of the heavy precipitation events. This feature, which is more pronounced when the wind corrections are not applied, explains the poor daily correlation between GSWP-2 and the Rhône-AGG precipitation data sets shown in Table 1.

2.4 Global data intercomparison

Firstly, both *B0* and *P3* products are compared to the following precipitation data sets over the validation period (1986–1995):

- GPCP-2 (http://www.precip.gsfc.nasa.gov/): This global product combines estimations based on satellite microwaves (SSM/I), infrared (IR) data and GPCC gauge measurements (Adler et al. 2003). The current data set extends from 1979–2002 at 2.5° × 2.5° resolution and at a daily or monthly time scale.
- CMAP (http://www.cdc.noaa.gov/cdc/data.cmap.html): The CPC Merged Analysis of Precipitation produces pentad and monthly analyses of global precipitation in which observations from rain gauges are merged with precipitation estimates from several satellite-based algorithms (Xie and Arkin 1996). This global data set consists of monthly averaged precipitation rate values for the time period Jan 1979 to the Sep 2004 at 2.5° × 2.5° resolution. Note that, this data set used here does not include blended NCEP/NCAR reanalysis.
- CRU-2 (http://www.cru.uea.ac.uk/cru/data/hrg.htm): This data set is based only on raingauge measurements, for the period 1901–2002, and covering the global land surface at 0.5° resolution (Mitchell and Jones 2005).

The zonal mean comparison of the annual mean precipitation shown in Fig. 4 reveals that all products are in good agreement over the equatorial zone (30°N–30°S). A significant difference between *B0* and other data sets appears over northern mid and high latitudes (80°N–30°N). Indeed, B0 seems to be largely overestimated compared to other precipitation products. *P3* shows a better agreement with both CMAP and CRU-2 data set while GPCP precipitation is larger than these three products. Over the southern hemisphere, a drastic difference appears between *B0*, *P3* and other products. Nevertheless, over this domain, continental areas are very sparse and implications on hydrological simulations remain very small.

Figure 5 shows the comparison between the spatial distribution over mid and high latitudes of the *B0* annual mean precipitation and other products including *P3*. *B0* precipitation is higher than the other products everywhere, except the GPCP data set over Siberia. The most significant differences appear over

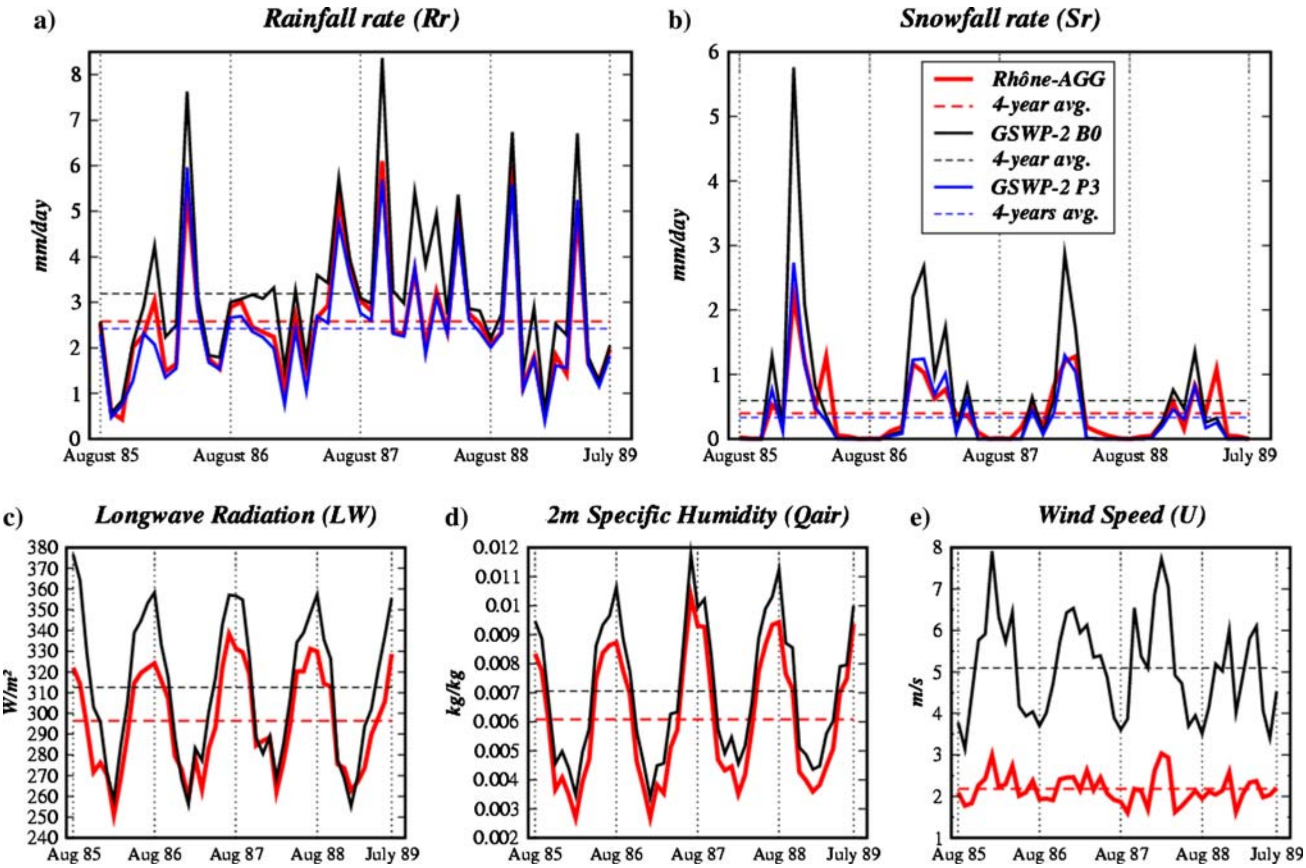

Fig. 2 Comparison between the baseline GSWP-2 *B0* (*black line*) and Rhône-AGG (*red line*) atmospheric data sets. Only monthly timeseries of the basin-average values for variables that differ significantly (see Table 2) are represented: **a** Rainfall rate (R_r), **b** Snowfall rate (S_r), **c** LongWave radiation (*LW*), **d** Specific humidity at 2 m (Q_{air}), and **e** Wind Speed at 10 m (*U*). The 4-year annual mean basin-average values are also shown for each data set. Note that an alternative GSWP-2 *P3* precipitation product without wind and satellite corrections is also shown (*blue line*)

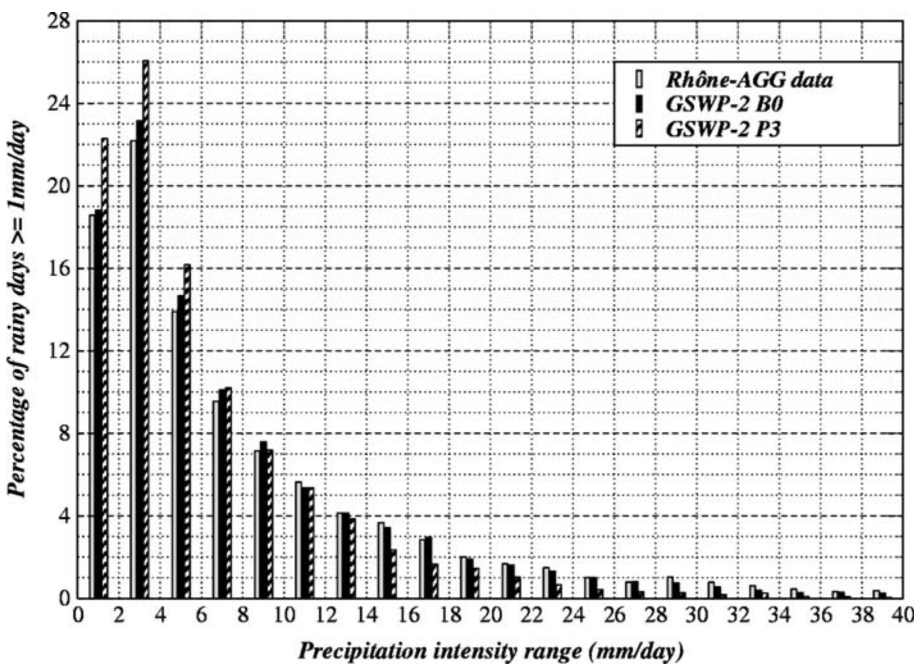

Fig. 3 Comparison between distributions of daily precipitation intensities for the baseline GSWP-2 *B0* (*black bar*), the alternative GSWP-2 *P3* (*dashed black bar*), and Rhône-AGG (*gray bar*) data sets. Note that the frequencies are expressed as a percentage of rainy days (i.e. days with precipitation intensities above or equal to 1 mm/day)

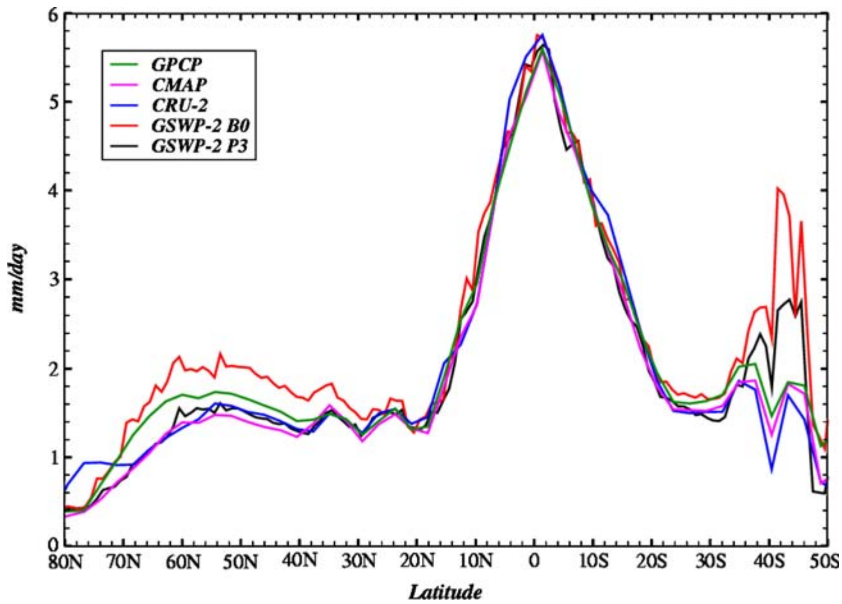

Fig. 4 Evaluation of GSWP-2 precipitation forcing data sets over the period 1986–1995. Annual zonal mean average (mm/day) of GSWP-2 baseline B0 (*red line*) and GSWP-2 alternative *P3* (*black line*) are compared to GPCP (*green line*), CMAP (*magenta line*) and CRU-2 (*blue line*) precipitation products

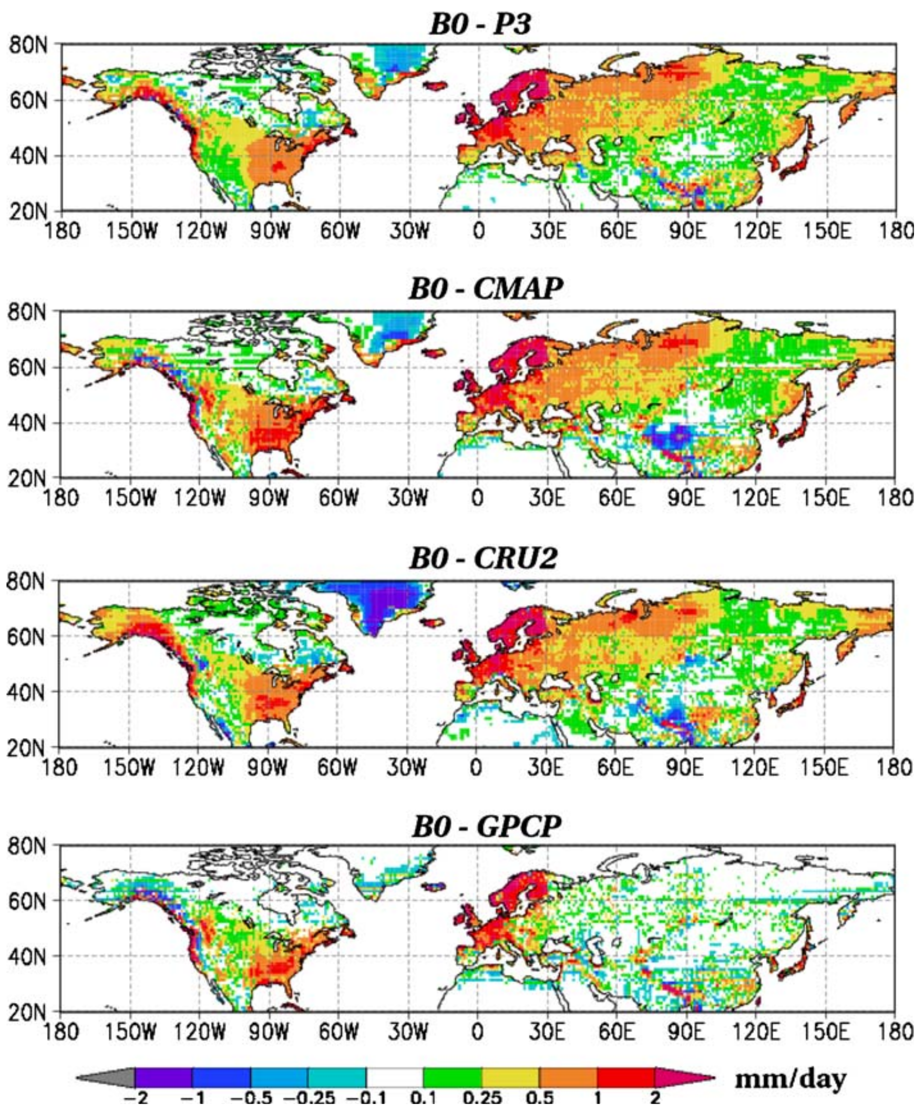

Fig. 5 Comparison of the baseline GSWP-2 *B0* annual mean precipitation rate (mm/day) over the period 1986–1995 with other products (*P3*, GPCC, CMAP, CRU-2)

North-East America and Europe. Over these regions, the density of the raingauge network (Fig. 1) reveals that the wind corrections applied to *B0* are responsible for these differences. Another important difference, especially due to GPCP corrections, appears over Siberia. Over this region, GPCC gauge density are very low (Fig. 1) and then, according to Eq. 2, this difference is mainly due to GPCP adjustments rather than to wind corrections.

3 Regional scale study: GSWP-2 versus Rhône-AGG

3.1 Experiment design

For the comparison of the off-line simulations driven with either the GSWP-2 or the Rhône-AGG atmospheric forcing, the ISBA land surface model is integrated over the Rhône river basin with a 5 min time step for four consecutive annual cycles where the first year is treated as a spin-up. Results are validated over the last 3 years (August 1986 to July 1989).

ISBA contains the basic physics of the land surface and needs only a limited number of parameters, which depend on the type of soil and vegetation (Noilhan and Planton 1989). It uses the force-restore method (Deardorff 1977, 1978) to calculate the temporal evolution of the surface and mean soil temperature, the water interception storage, the snow pack evolution (Douville et al. 1995), and the soil moisture budget (Noilhan and Mahfouf 1996) including freezing and thawing in the two uppermost layers (Boone et al. 2000). The ISBA LSM has a three-layer soil hydrology (Boone et al. 1999). The surface runoff is computed using the variable infiltration capacity (VIC) scheme (Zhao 1977; Dümenil and Todini 1992; Wood et al. 1992) introduced by Habets et al. (1999a) and the deep drainage is computed according to Mahfouf and Noilhan (1996). Note that, in this study ISBA is used as a global land surface model that has contributed to GSWP-2 and has not been tuned to give particularly good results over the Rhône basin.

Table 2 Summary of experiments over the Rhône river basin

Experiments	Precipitation forcing	Other atmospheric variables
RB0	*B0*	GSWP-2
RP3	*P3*	GSWP-2
RMR	*B0* hybridized with monthly Rhône-AGG data	GSWP-2
RDR	Rhône-AGG	GSWP-2
Control	Rhône-AGG	Rhône-AGG

A set of five simulations is performed and is summarized in Table 2:

- *Control*: this experiment consists of running ISBA with the whole Rhône-AGG atmospheric forcing aggregated onto the 1° by 1° GSWP-2 grid.
- *RB0*: this simulation consists of running ISBA with the whole GSWP-2 baseline atmospheric forcing.
- *RP3*: same as *RB0* but with an alternative GSWP-2 precipitation forcing without wind and GPCP corrections.
- *RMR*: same as *RB0* but with a new hybrid precipitation product based on the monthly Rhône-AGG data, P^{m}_{R-AGG}, and on the 3-hourly (Eq. 2) and monthly (P^{m}_{B0}) baseline products used in GSWP-2. This hybridization is simply made as follow:

$$P^{3h}_{RMR} = P^{3h}_{B0} P^{m}_{R-AGG} / P^{m}_{B0}. \tag{3}$$

- *RDR*: same as *RB0* but with the Rhône-AGG 3-hourly precipitation forcing and the GSWP-2 other atmospheric variables.

The purpose of these two last experiments (RMR and RDR) is to investigate the influence of the forcing variables other than precipitation and to assess the impact of the prescribed precipitation frequency. Note that all experiments are driven using the aggregated Rhône-AGG surface parameters.

For each simulation, MODCOU is used to convert the surface runoff and the drainage produced by ISBA into river discharge and water table variations (Habets et al. 1999b). The surface runoff is transferred to the river, and the routing from each grid cell is based on isochronous zones using a time step of 1 day. The drainage acts as a source for the water table, which is modeled using diffusivity equations. Note that, because the MODCOU model is integrated on the Rhône-AGG high-resolution grid, the simulated surface runoff and drainage are linearly disaggregated to the 8 km by 8 km grid before they are transferred to the hydrological model. Note that all components of the Rhône modeling system (SAFRAN-ISBA-MODCOU) have been developed and calibrated independently. Finally, simulated discharges are compared to observed data. One set of observations is used for evaluating the ISBA simulations over most of the Rhône basin. This set consists of daily streamflow data from 88 river gauges. Only sub-basins with drainage areas of at least 250 km^2, and where damming does not impact too much the flow, are used for the validation.

3.2 Results

As in the Rhône-AGG experiment, the simulated river discharges are compared to the observations at the Viviers gauging station. The statistics are shown in Table 3 where Eff represents efficiency or Nash–Sutcliffe criteria (Nash and Sutcliffe 1970) and Q_{sim}/Q_{osb} represents the ratio, or the mean bias, between simulated and observed annual discharge. Eff measures the skill of the model at capturing the observed variability of the discharges. It is defined, as follows:

$$\mathrm{Eff} = 1.0 - \frac{\sum_t \left(Q_{\mathrm{sim}}(t) - Q_{\mathrm{obs}}(t)\right)^2}{\sum_t \left(\overline{Q_{\mathrm{obs}}} - Q_{\mathrm{obs}}(t)\right)^2}, \qquad (4)$$

where $\overline{Q_{obs}}$ represented the observed temporal mean. Eff can be negative if the simulated discharge is very poor, is above 0.5 for a reasonable simulation, above 0.7 for a good one and 1 for a perfect model (Boone et al. 2004).

The time series of daily river discharge (from August 1987 to July 1988) and monthly mean river discharge (from August 1986 to July 1989) show a large overestimation of *RB0* during the rainy season (Fig. 6). *RP3*, as well as *RMR* and *RDR*, shows a better agreement with the observed discharges. This result is consistent with those shown in Fig. 2. The statistics given in Table 3 emphasize that the quality of the precipitation forcing is a crucial condition for producing a realistic simulation of the river discharge at Viviers.

A global discharge validation based on daily and monthly efficiencies is also performed with the help of the 88 gauging stations distributed over the entire basin. Figure 7 shows that efficiency scores are significantly better for *RP3* and the simulations with the Rhône-AGG precipitation forcing. Once again, the statistics of these distributions (Table 4) suggest that the quality of the GSWP-2 precipitation forcing, especially *B0*, is not sufficient to produce realistic hydrological simulations over the Rhône basin. These results indicate that the major sensitivity of our hydrological simulations is due to uncertainties in the monthly precipitation amounts. Nevertheless, the high-frequency precipitation intensity (comparison between *RMR* and *RDR*) plays a significant role at both daily and monthly time scales (*RDR* vs. *Control*) and must be reasonable to ensure a realistic simulation of the discharge dynamics, though its effect on annual discharge is negligible.

Table 3 Statistics of simulated monthly and daily river discharges at Viviers

Simulations	Q_{sim}/Q_{obs}	Efficiencies	
		Daily	Monthly
RB0	1.22	0.12	–0.13
RP3	0.88	0.71	0.75
RMR	0.91	0.75	0.80
RDR	0.92	0.83	0.85
Control	0.97	0.89	0.95

The monthly and daily efficiency (Eff), and the ratio of the simulated discharge to observation (Q_{sim}/Q_{obs}) are shown. The definition of Eff (Nash criteria) is given in Sect. 3

4 Global scale study

The previous regional study (Sect. 3) shows that *B0* precipitation leads to large discharge overestimations over the Rhône basin despite the high density of the GPCC observations (Fig. 1). This result is due to wind corrections because GPCP adjustments are not effective here according to Eq. 2. Nevertheless, the scope of our conclusions is limited since the Rhône basin is a small region of the globe. Moreover, only one land surface model was used to validate the GSWP-2 atmospheric forcing. For these reasons, this section presents a comparison between several global hydrological simulations performed with six different land surface models.

4.1 Experiment design

The GSWP-2 baseline precipitation data set is evaluated using comparisons between simulated and observed discharges at 80 gauging stations distributed among 33 basins over the globe (Table 5). Indeed, the TRIP river routing model is used to convert the daily total runoff from six LSMs (Table 6) into river discharge. TRIP is a linear model based on a single prognostic equation for the water mass within each grid cell of the hydrological network. The stream flow velocity is assumed constant and uniform at 0.5 m/s (Oki and Sud 1998; Oki et al. 1999). The global river channel network is provided at 1° resolution and adapted to GSWP-2 land-sea mask. Fifteen models representing the diversity of current land surface schemes have participated in GSWP-2. Nevertheless, only six LSMs have been driven with both *B0* and *P3* precipitation products. All of these six models are at least considered as "second-generation" models because they take into account the effect of vegetation and snow on surface hydrology and include realistic treatments of the surface energy balance. Furthermore, all these models take into

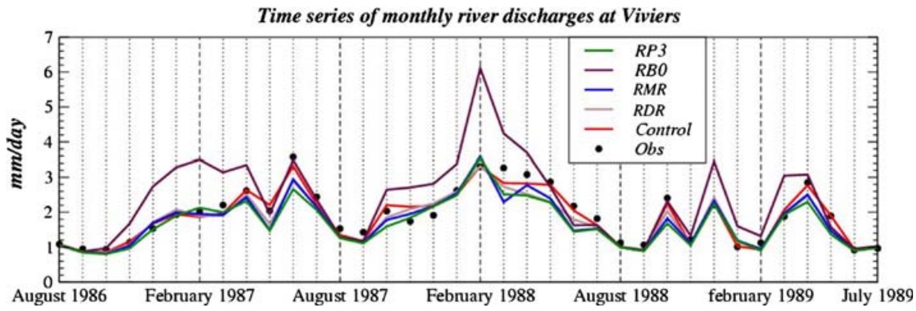

Fig. 6 Timeseries (August 1986–July 1989) of the observed (*black points*) and simulated monthly mean river discharges at Viviers. For all *panels, green, purple, blue, brown* and *red lines* represent, respectively, experiments *RP3*, *RB0*, *RMR*, *RDR* and *Control* defined in Table 2. The associated statistics are given in Table 3

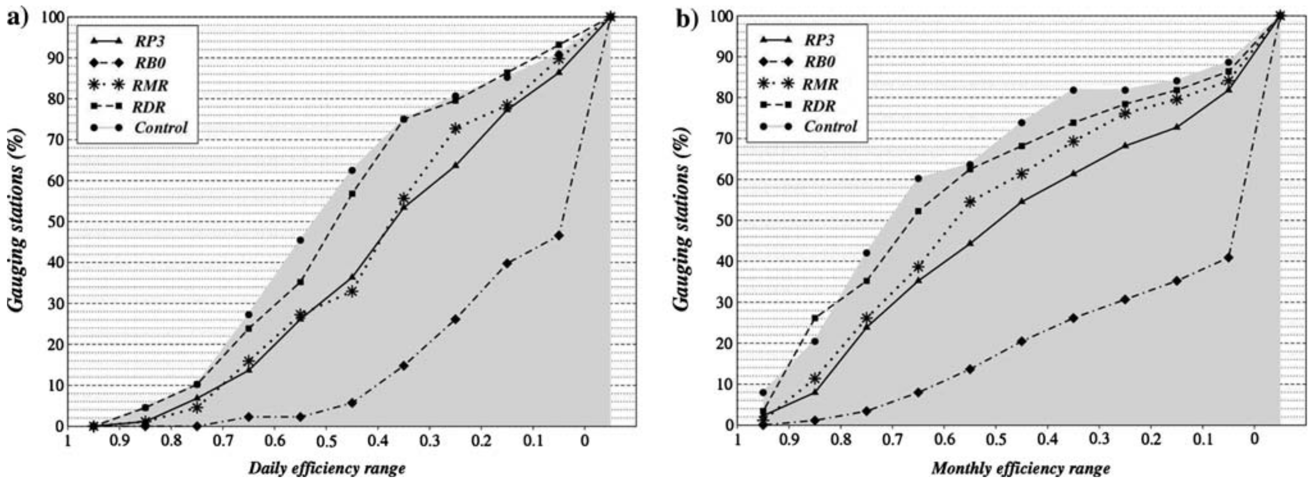

Fig. 7 Cumulative efficiency distribution of daily (**a**) and monthly (**b**) river discharges simulated in each experiment. The *plain, dotted-dashed, dotted, dashed*, and *gray plain lines* represent the *RP3*, *RB0*, *RMR*, *RDR* and *Control* experiments, respectively. The efficiencies are computed against daily discharges observed at 88 gauging stations. This number of stations is given in percentage on all panels. The corresponding statistics of the distributions are given in Table 4

consideration surface runoff and drainage mechanisms in order to produce a ‘‘realistic’’ total runoff. One of them, MOSES, is considered as a ‘‘third-generation’’ model because is the carbon cycle is described, and includes a more sophisticated representation of the functioning of vegetation than ‘‘second-generation’’ models. GSWP recommended that all LSMs start their runs from initial conditions of soil temperature based on the mean June air temperature, of soil moisture at 75% of saturation, and without snow cover over ice-free points. A spin-up period is considered using forcing data beginning 1 July 1982 and ending 31 December 1985 and the period 1986 to 1995 is used as the validation stage.

Table 4 Statistics of efficiency distributions of the simulated daily and monthly river discharges over the Rhône river basin at the Viviers gauging station. The minimum (Min), mean (Mean), maximum (Max), and standard deviation (Std) of each distribution are given

Efficiency statistics	RP3	RB0	RMR	RDR	Control
Min					
Daily	–1.58	–4.30	–1.75	–1.77	–1.58
Monthly	–1.23	–8.74	–1.30	–1.36	–1.16
Mean					
Daily	0.28	–0.53	0.30	0.38	0.39
Monthly	0.32	–1.20	0.39	0.46	0.51
Max					
Daily	0.86	0.66	0.86	0.85	0.90
Monthly	0.93	0.82	0.96	0.94	0.97
Std					
Daily	0.32	0.99	0.35	0.36	0.36
Monthly	0.47	2.00	0.44	0.44	0.43

Over the validation period (1986–1995), more than a hundred gauging station measurements are available via the Global Runoff Data Center (GRDC; http://www.grdc.sr.unh.edu/index.html), the R-ArticNet database (New Hampshire University; http://www.r-arcticnet.sr.unh.edu/v3.0/index.html) for high latitude basins and the HyBAm dataset (http://www.mpl.ird.fr/hybam/) for the Rio-Amazonas basin. Nevertheless, the set of monthly streamflow data used for the validation contains no more than 80 river gauges because only sub-basins with drainage areas of at least 10^5 km^2 and with a minimum observed period of 4 years have been selected (Table 5).

4.2 Results

As previously mentioned, outputs from six LSMs (Table 6) driven by both *B0* and *P3* precipitation

Table 5 Basins used for the comparison between simulated and observer discharges

Basins	NS	Downstream station	Area (km^2)	Lon	Lat	RD	Period
Rio Amazonas	12	Obidos	4,758,000	–55.5	–2.5	46	86–95
Congo	1	Brazzaville	3,649,000	15.5	–4.5	16	86–95
Mississippi	8	Vicksburg	3,011,000	–91.5	32.5	949	86–95
Ob	5	Salekhard	2,902,000	66.5	66.5	81	86–95
Parana	3	Timbues	2,596,000	–60.5	–32.5	317	86–94
Yenisei	2	Ygarka	2,502,000	86.5	67.5	62	86–95
Lena	2	Kusur	2,310,000	127.5	70.5	36	86–95
Mackenzie	8	Mackenzie	1,736,000	–133.5	67.5	64	86–95
Amur	1	Komsomolsk	1,772,000	137.5	50.5	93	86–90
Volga	1	Volvograd	1,326,000	44.5	48.5	71	86–90
Ganges	1	Harding Bridge	970,000	88.5	24.5	304	86–92
Yukon	2	Pilot station	826,000	–162.5	61.5	56	86–95
Orinoco	1	Puentes Angostura	820,000	–63.5	8.5	369	86–90
Niger	4	Niamey	799,000	2.5	13.5	185	86–91
Danube	5	Ceatal Izmail	797,000	28.5	45.5	1422	86–90
Columbia	1	The Dalles	634,000	–121.5	45.5	613	86–95
Chari	1	Ndjamena	558,000	15.5	11.5	122	86–91
Kolyma	1	Kolymskoye	536,000	158.5	68.5	22	86–95
Brahmaputra	1	Bahadurabad	519,000	89.5	25.5	50	86–92
Soa Francisco	1	Juazeiro	488,000	–40.5	–9.5	453	86–94
Mékong	5	Mukdahan	405,000	104.5	16.5	139	86–93
Severnaya Dvina	1	Ust Pinega	364,000	41.5	64.5	73	86–95
Pechora	2	Oksino	298,000	52.5	67.5	71	86–95
Indigirka	1	Vorontsovo	277,000	147.5	69.5	16	86–94
Sénégal	1	Kayes	239,000	–12.5	14.5	178	86–90
Yana	1	Ubileynaya	228,000	136.5	70.5	19	86–94
Fraser river	2	Hope	211,000	–121.5	49.5	485	86–95
Wisla	1	Tczew	194,000	18.5	53.5	514	86–94
Rhin	1	Rees	146,000	6.5	51.5	7636	86–95
Albany river	1	Near Hat Island	140,000	–83.5	51.5	14	86–95
Burdekin	1	Clare	126,000	147.5	–20.5	684	86–94
Colorado	1	Wharton	105,000	–96.5	29.5	897	86–95
Odra	1	Gozdowice	100,000	14.5	52.5	865	86–94

The number of gauging station for each basin (NS) is given. The name, the drainage area (area in km^2), the localizations (longitude, Lon, and latitude, Lat) into the TRIP river network, the rain gauge density (RD, /106 km^2) and the discharge observation period at each basin downstream station are also given

Table 6 Land surface model used for the global evaluation of the GSWP–2 forcing. The model structure gives the number of soil layer for hydrology (W) and temperature (T), and the maximum number of snow layers (S)

LSM	Institut	Nation	Model structure	Most recent references
ISBA	Météo-France	France	3W 2T 1S	Decharme and Douville (2006)
MOSES	Met office	UK	4W 4T 1S	Gedney and Cox (2003), Essery et al. (2003)
NOAH	NOAA/NCEP/EMC	USA	4W 4T 1S	Ek et al. (2003)
NSIPP	NASA/GSFC/GMAO	USA	3W 6T 3S	Koster et al. (2000b), Ducharne et al. (2000)
SsiB	IGES/COLA	USA	6W 6T 1S	Dirmeyer and Zeng (1999)
SWAP	RAS/IWP	Russia	2W 1T 1S	Gusev and Nasonova (2003)

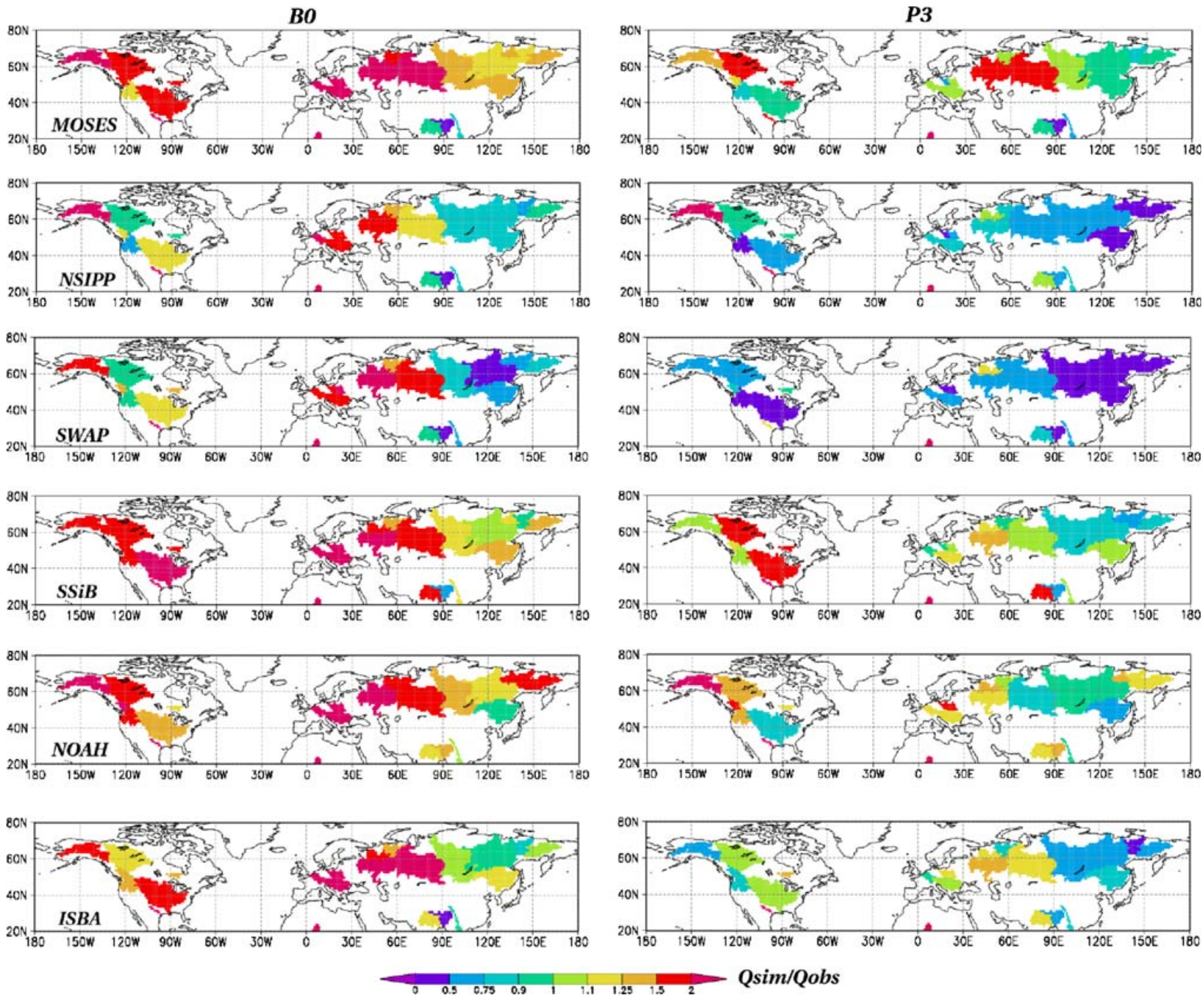

Fig. 8 Comparison between mean annual simulated to observed discharges ratios (Q_{sim}/Q_{obs}) for *B0* (*left panel*) and *P3* (*right panel*) simulations over 33 river basins. This ratio is shown for each downstream gauging station of each northern hemisphere basins given in Table 1

products are analyzed. For the *B0* runs, Fig. 8 shows that four models (MOSES, SSiB, NOAH, ISBA) drastically overestimate annual mean discharges over all North American basins while this trend is less pronounced for NSIPP and SWAP. This drastic overestimation is systematic over European and West Siberian basins, while the simulated discharges of East Siberian basins are in better agreement with observations. Note that, according to Eq. 2, the GPCP adjustment is particularly significant over East Siberia given the very low GPCC rain gauge density in this area (Fig. 1). For the *P3* runs, simulated runoff are closer to gauging measurements over North America except for NSIPP and SWAP that tend to underestimate observations. Over Europe and West Siberia, the simulated discharge quantities of all models, except SWAP, are generally in better agreement. Over East Siberia, only MOSES and NOAH show good results while other models tend to underestimated gauging measurements.

Figure 9 compares the simulated and observed annual cycles and synthesizes the monthly discharge dynamic with the efficiency, Eff, criteria. Over the Mississippi basin, all *B0* runs show overestimated discharges and poor dynamics scores (except NSIPP). *P3* runs show more acceptable scores except for NSIPP and SWAP that tend to underestimate gauging measurements and SSiB that misrepresents the annual dynamics of the discharge. Generally, the same results are observed over the Ob basin, while differences between *B0* and *P3* runs over the Rio-Amazonas basin are less pronounced.

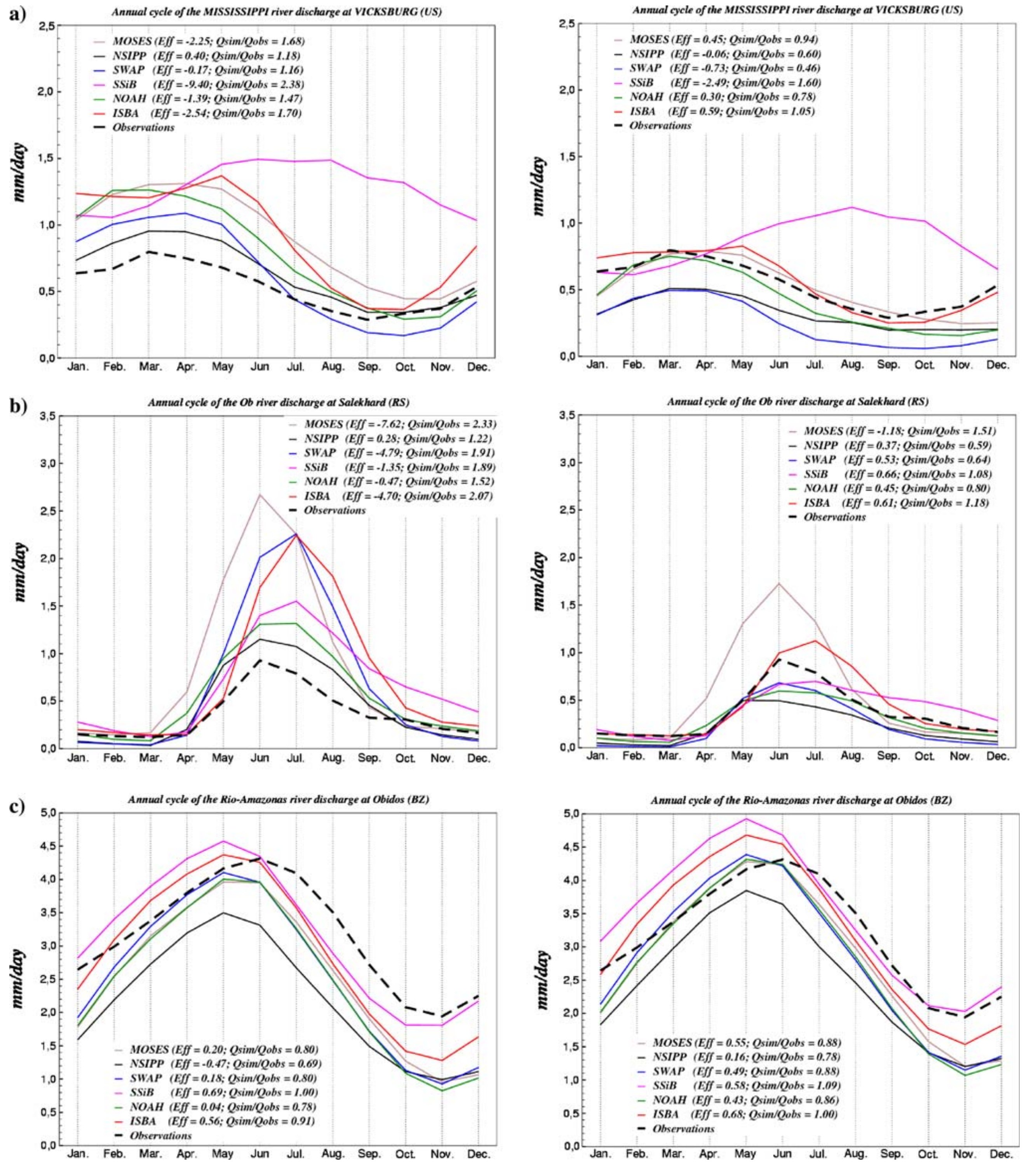

Fig. 9 Mean annual cycles comparison between simulated and observed discharges over the Mississippi (**a**), the Ob (**b**) and the Rio-Amazonas (**c**) river basins. Simulations forced with the *B0* precipitation product are shown on the *left* while simulations driven by the *P3* dataset are shown on the *right*. Note that monthly efficiencies (*Eff*) shown on each picture is calculated over the whole discharge observation periods (1986–1995) given in Table 5 and not over the annual cycle. The mean annual ratio between simulated and observed discharges (Q_{sim}/Q_{obs}) is also shown

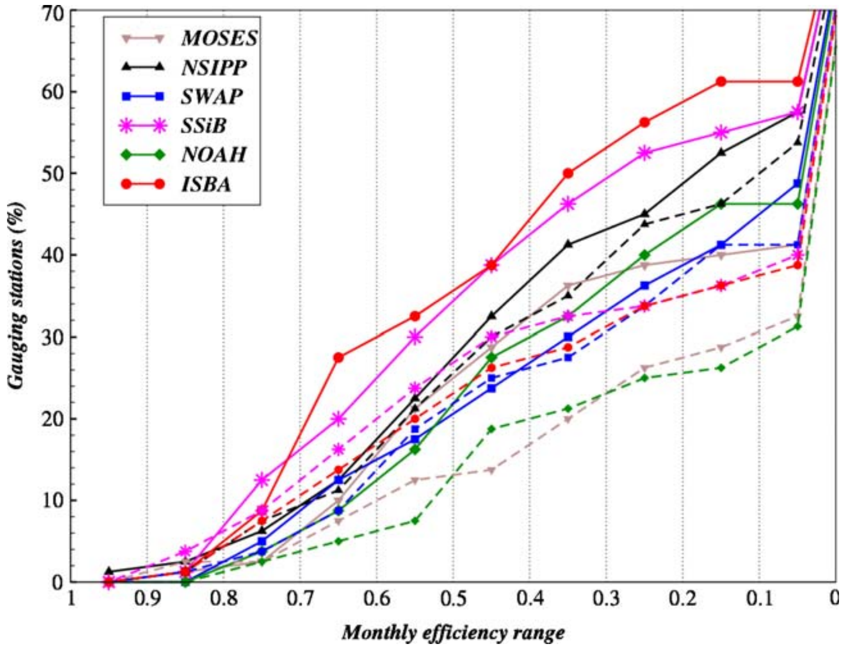

Fig. 10 Global comparison between *B0* (*dashed line*) and *P3* (*plain line*) simulated discharge scores. The cumulative efficiency distribution of monthly river discharges simulated by each land surface model is calculated with the helps of 80 gauging stations distributed over the globe (see Table 5)

Finally, a global discharge validation based on monthly efficiencies is also performed (Fig. 10) with the help of the 80 gauging stations distributed over the globe. This criterion confirm that generally *P3* runs are significantly better than *B0* simulations even if NSIPP and SWAP seem to be globally less sensitive to precipitation input than the other LSMs.

5 Idealized sensitivity test

The results in Sect. 3 and 4 suggest that the dominant source of uncertainty in our hydrological simulations is related to the precipitation forcing. Over the Rhône basin, the baseline GSWP-2 total precipitation, *B0*, is 27% greater than the SAFRAN analysis while the alternative product, *P3*, (without empirical corrections) is 9% lesser. At Viviers, the mean bias $\overline{Q}_{\mathrm{sim}}/\overline{Q}_{\mathrm{obs}}$ shows that the simulations driven with *B0* overestimate the observed discharge by 22% (*RB0*). With the alternative and more realistic *P3* precipitation forcing (*RP3*), the ratio is underestimated by 12% while using the SAFRAN precipitation forcing leads to an underestimation of 8% for *RDR* and 3% for *Control*. Furthermore, the mean annual discharge bias averaged over all the 88 gauging stations also reveals that the observed discharge is overestimated in *RB0* by about 43%, and underestimated by 26% in *RP3* against an underestimation of 17% for *RDR* and only 7% for *Control*. The bias shown by *Control* can be attributed to several sources: errors in the SAFRAN analysis, errors in the prescribed surface parameters, deficiencies in the ISBA LSM. The discharge errors found in *RDR* are also due to the GSWP-2 atmospheric variables other than precipitation. The main conclusion is that the suggested overestimation of the B0 precipitation leads to a strong overestimation of the annual mean discharge produced by ISBA, while the use of the *P3* precipitation forcing leads to more acceptable results. While this result alone is not sufficient to question the quality of the default GSWP-2 atmospheric forcing (*B0*), it is an additional item in the list of arguments that lead to this conclusion.

In order to better understand how precipitation uncertainties are translated into discharge uncertainties, one additional set of experiments has been performed over the Rhône basin. Several runs using ISBA driven with the Rhône-AGG atmospheric forcing have been compared in which the SAFRAN precipitation has been uniformly (in both space and time) increased or decreased by 5–50% with a step of 5%. The run where the precipitation error is considered equal to 0% corresponds to the *Control* run shown in Sect. 3. The response of the simulated discharges is analysed in term of mean annual relative errors and mean monthly efficiencies (Fig. 11). Three classes of results are presented considering the stations where the *Control* annual discharges are previously in good agreement with the observation (18% of the gauging stations), underestimated (50%) or overestimated (32%). The relationship between the relative errors in annual precipitation and annual discharges is quasi-linear over the Rhône basin, at least in this interval of precipitation biases (Fig. 11a). For example, at the stations where the *Control* annual discharges are correct, the mean slope of this relationship is equal to 1.7. Moreover, Fig. 11b shows that the efficiencies decrease more significantly when precipitation is overestimated than when it is underestimated. In other words, the efficiency criterion is more discriminating when the discharges are overestimated than when they are underestimated. These characteristics can be also observed for the *RMS* criterion (not shown). This feature partly explains that NSIPP and SWAP are globally less sensitive to precipitation input than the other LSMs when the efficiency criterion is used (Fig. 10). With the *B0* precipitation product, both models simulate a lesser overestimation of the annual mean discharges than the other LSMs over North America and Siberia (Fig. 8). Therefore, their efficiency scores are less affected. With the *P3* product, even if the discharges simulated by NSIPP and SWAP are generally in better agreement with observations compared to the *B0* run, some basins show a stronger underestimation than with the other LSMs and the improvement of the efficiency is less clear.

Some results of this last experiment are certainly model dependent. For another model than ISBA, the slope of the linear relationship between the relative errors in precipitation and discharge would probably be different. The linearity of this relationship also depends on the Rhône basin characteristics. For a basin like the Niger, having strong seasonal floodings, this relation is certainly not linear. Nevertheless, these results are consistent with the conclusion of Fekete et al. (2003): the uncertainty in precipitation is translated at least to the same or generally to a greater relative error in total runoff. The monthly efficiencies also point out that the quality of the hydrological simulations decreases drastically with increasing precipitation uncertainties. This result highlights once again the need for high-quality precipitation datasets for model validations.

6 Discussion and conclusions

The present study focuses on the validation of the atmospheric forcing proposed by GSWP-2 to drive land surface models and, thereby, produce global climatologies of soil moisture, land surface fluxes, and related hydrological quantities. First, this validation is conducted over the French Rhône river basin, encompassing a large variety of climate and vegetation types. The comparison of the GSWP-2 versus Rhône-AGG atmospheric forcing has revealed some systematic biases in precipitation, longwave radiation, air humidity, and wind speed, as well as a relatively poor temporal correlation between the GSWP-2 and Rhône-AGG timeseries at the 3-hourly and daily time scales. The precipitation shows the strongest discrepancies, with a general overestimation of both snowfall and rainfall over the Rhône domain. This deficiency is mainly due to the wind corrections that have been applied to the baseline GSWP-2 precipitation forcing, *B0*. An alternative precipitation product, *P3*, without wind corrections, reveals indeed a better agreement with the Rhône-AGG data set.

In addition to this direct comparison, a series of offline hydrological simulations has been conducted over

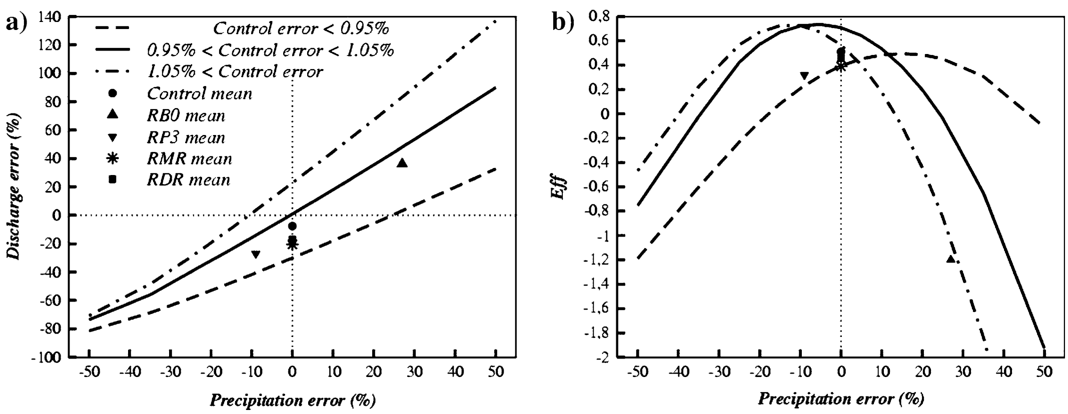

Fig. 11 Impact of precipitation errors on the simulated discharges over the Rhône river basin with the ISBA LSM. These results are obtained using the Rhône-AGG atmospheric forcing in which the SAFRAN precipitation has been uniformly (in both space and time) increased or decreased by 50% with a step of 5%. A precipitation error of 0% corresponds to the *Control* run shown in Sect.3. The responses of the ISBA discharge simulations are shown in term of mean annual errors (**a**) and mean efficiencies (**b**). For *each panel*, three classes of results are presented considering the stations where the *Control* annual discharges are in good agreement (*plain line*), underestimated (*dashed line*) or overestimated (*dashed-dotted line*). The mean discharge scores for each experiment presented in Sect. 3 are also shown

the Rhône basin with the ISBA land surface model driven with either the GSWP-2 or the Rhône-AGG precipitation data. These sensitivity tests represent an interesting tool for providing an indirect regional quality-check of the GSWP-2 database. The simulated runoff has been converted into river discharge with the MODCOU model (Habets et al. 1999b) and the simulated discharges have been validated against a dense network of gauging measurements. The results confirm that the *B0* precipitation product is the main problem of the GSWP-2 database over the Rhône domain. Simulations driven with the Rhône-AGG precipitation, and to a lesser extent with the *P3* product, are consistently better than the sensitivity experiments using the GSWP-2 baseline data. Moreover, other experiments have been performed in which the Rhône-AGG precipitation has replaced the GSWP-2 precipitation or has been used to hybridize this product. Results show reasonable discharge scores and also point out the relevance of daily precipitation intensities for accurate hydrological simulation. Prescribing more realistic precipitation intensities has a positive impact on the simulated discharges. Indeed, besides the monthly precipitation amounts, the intensity distribution of the daily precipitation also shows some systematic errors when the GSWP-2 precipitation products are compared to Rhône-AGG, suggesting that the NCEP/DOE reanalysis is not very reliable for prescribing the daily precipitation frequency. They favor low intensities at the expense of heavy precipitating events, which result in strong modification to the interception loss and to the simulated total runoff. Therefore, the quality of the simulated discharges, both at daily and monthly time scales, is better when the Rhône-AGG daily precipitation intensities are used (difference between the *RMR* and *RDR* runs). In other words, the impact of the daily precipitation intensities is significant on the quality of the hydrological simulations. A preliminary assessment of another precipitation product where the hybridization is made with the ECMWF 40-year reanalysis shows a relative improvement in this respect. These results summarize the main deficiencies of the GSWP-2 atmospheric forcing over the Rhône basin: the baseline precipitation forcing is drastically overestimated due to empirical wind corrections.

Secondly, a global comparison between the GSWP-2 baseline precipitation forcing and other available monthly data sets is also performed. This comparison reveals that *B0* precipitation is significantly stronger than the other products over the mid-and-high latitudes, except than the GPCP dataset over Siberia. The most significant discrepancies appear over North-East America and Europe where the raingauge network is yet relatively dense (Fig. 1). Off-line global hydrological simulations from six land surface models, for which the total runoff has been routed with the TRIP river routing model (Oki and Sud 1998) in order to simulate river discharges, are then analyzed. The comparison with gauging observations reveals that all *B0* runs overestimate significantly the discharges of mid and some high latitude basins. Both the Mississippi and Rhône basin results incriminate the wind corrections since GPCC rain gauge network is dense over these regions and then the GPCP corrections are not effective according to Eq. 2. Conversely, the corrections applied to the *B0* product seem to be useful over eastern Siberia basins for a reasonable discharge simulation. Over this domain, the GPCP corrections have the most significant effect considering the very low GPCC gauge density over these regions.

These results point out that the reliability and accuracy of the baseline precipitation forcing (*B0*) remains limited, especially in well-instrumented regions like Northeast America and Europe. These conclusions also suggest that the regional or global validation of a land surface model based on off-line simulations driven by the GSWP-2 baseline products remains a difficult task. Given the uncertainties in this precipitation forcing and their impacts on hydrological simulations, the validation of modifications to the model physics is not easy. This remark should however be tempered. Like it was already shown in GSWP-1, the GSWP-2 database remains very useful and valuable to detect major deficiencies in the land surface models (Dirmeyer et al. 1999; Oki et al. 1999; Boone and Wetzel 1999; Koster et al. 1999). This validation mode is much easier for interpretations than in coupled land–atmosphere experiments where systematic errors in the atmospheric forcing are generally much larger (Mahfouf et al. 1995). For this purpose, the alternative GSWP-2 precipitation product (*P3*) without wind and satellite corrections is strongly recommended. This recommendation is confirmed by the study of Ngo-Duc et al. (2005), which shows that a simple precipitation product (NCEP/DOE reanalysis hybridized with CRU-2) improves global hydrological simulations compared to baseline GSWP-2 data.

Note that this remark does not mean that the *P3* precipitation product is perfect. This product has also weaknesses, especially over some high latitude regions where the simulated annual mean discharge is systematically underestimated (Fig. 8). Indeed, it seems that the GPCP satellite adjustment included in *B0* is useful for a realistic discharge simulation over the eastern Siberian basins. Nevertheless, satellite precipitation retrieval is a difficult task over land since the signature is very indirect through the temperature of cold clouds in the infra-red or the scattering effect by cloud ice in the microwave. The satellite observations also have difficulty in detecting shallow, orographic precipitation (Adler et al. 2003). In addition, it has been recognized that uncertainties exist in the estimated precipitation climatologies due to biases of gauge measurements, such as wind-induced gauge undercatch, wetting and evaporation losses, and underestimation of trace precipitation amounts that lead to underestimated precipitation (Yang et al. 2001, 2005; Adam and Lettenmaier 2003; Ye et al. 2004) and consequently the annual mean discharges simulated by the LSMs. Some algorithms have been proposed in order to improve precipitation climatologies from gauge measurements using gauge catch ratio versus observed wind speed (Yang et al. 2001, 2005; Adam and Lettenmaier 2003; Ye et al. 2004). Wind corrections for gauge undercatch is however a difficult problem since they depend upon the type of instrument, the climatology of the studied areas, and the type of precipitation. The lack of a comprehensive description of the proposed methodology with scientific justifications in the open literature puts additional concern regarding the pertinence of the wind correction used in the *B0* product. The 10-m wind speed in the NCEP reanalysis over the Rhône-AGG domain show a clear overestimation compared to the SAFRAN analysis. This is probably due to the coarse resolution of the numerical weather prediction model. The orography of the Alps is probably too smooth and the effect of low-level blocking by the subgrid-scale orography as proposed by Lott and Miller (1997) does not seem to be included in the NCEP global model. Therefore, it is certainly not adequate to use low-level winds from a coarse resolution version of the NCEP model in order to correct in situ raingauge measurements for the undercatch problem. Moreover, it is well known that the gauge catch ratio is different for liquid versus solid precipitation (Yang et al. 2001; Adam and Lettenmaier 2003; Ye et al. 2004). This effect is not accounted for in the algorithm applied on the baseline GSWP-2 precipitation. These remarks emphasize the need of improving the global meteorological analyses that are currently available to drive land surface models. Corrections in monthly precipitation amounts seem to be a reasonable goal to reach in a near future checking and applying existing gauge undercatch algorithm adjustments (Yang et al. 2001, 2005; Adam and Lettenmaier 2003; Ye et al. 2004) or using new satellite observations (as EGPM, the European contribution to Global Precipitation Mission of the European Space Agency), while improving daily or 3-hourly precipitation intensities will be certainly a more difficult task.

Besides the atmospheric forcing, prescribed soil and vegetation properties are also uncertain at the global scale. Comparisons between ISBA simulations using the ISLSCP-2 or the ECOCLIMAP (Masson et al. 2003) database both at regional and global scale have shown that the simulated discharges can be very sensitive to the land surface parameters, at least over specific basins. Although soil and vegetation parameters have generally a weaker impact on the simulated discharges than the atmospheric forcing, they strongly affect the simulated soil moisture climatology. In other words, realistic simulations of river discharges do not guarantee realistic simulations of soil moisture. The improvement of land surface parameter databases is therefore also important. Nevertheless, the validation of the global hydrological simulations remains difficult for at least three reasons. The first, as already said, is due to the quality of land surface model inputs. The second is relative to the intrinsic uncertainties in land surface parameterizations like, for example, the representation of cold processes (snow pack, soil freezing) that can reduce the quality of the hydrological simulations over the high latitude basins. Finally, global land surface simulations are mainly validated using comparisons between simulated and observed discharges, which depend on river routing models that are also uncertain, and/or using in situ soil moisture observations, that are not necessarily representative of the model grid scale (Robock et al. 2000). Regarding this last point, recent and future satellite missions based on gravimetric or microwave measurement will enable to bring an additional constraint to the validation of the global hydrological simulations.

Acknowledgments The authors would like to thank all their colleagues at the many French laboratories that have participated in the development of the Rhône modeling system, as well as Paul Dirmeyer and his team at the Center for Ocean Land Atmosphere studies (COLA) that have participated in the development of the Global Soil Wetness Project. Finally, we also wish to thank A. Boone and F. Habets for their useful comments on the Rhône-AGG project. Thanks are also due to the anonymous reviewers for their very constructive comments. This work was supported by Météo-France/CNRM and by the ACI "Observation de la Terre" of the French Research Ministry.

References

Adam JC, Lettenmaier DP (2003) Adjustment of global gridded precipitation for systematic bias. J Geophys Res 108(D9), 4257. DOI:10.1029/2002JD002499, 2003

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P (2003) The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). J Hydrometeorol 4:1147–1167
Boone A, Wetzel PJ (1999) A simple scheme for modeling sub-grid soil texture variability for use in an atmospheric climate model. J Meteor Soc Jpn 77:317–333
Boone A, Calvet JC, Noilhan J (1999) Inclusion of a third soil layer in a land surface scheme using the force-restore method. J Appl Meteor 38:1611–1630
Boone A, Masson V, Meyers T, Noilhan J (2000) The influence of the inclusion of soil freezing on simulation by a soil–atmosphere-transfer scheme. J Appl Meteor 9:1544–1569
Boone A, Habets F, Noilhan J (2001) The Rhône-aggregation experiment. GEWEX News 11(3):3–5
Boone A, Coauthors (2004) The Rhône-aggregation land surface scheme intercomparison project: an overview. J Clim 17:187–208
Champeaux J, Acros LD, Bazile E, Giard D, Gourtorbe JP, Habets F, Noilhan J, Roujean JL (2000) AVHRR-derived vegetation mapping over western Europe for use in numerical weather prediction models. Int J Remote Sens 21:1183–1199
Chapelon N, Douville H, Kosuth P, Oki T (2002) Off-line simulation of the Amazon water balance : a sensitivity study with implications for GSWP. Clim Dyn 19:141–154
Deardorff JW (1977) A parametrization of ground-surface moisture content for use in atmospheric prediction model. J Appl Meteor 16:1182–1185
Deardorff JW (1978) Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. J Geophys Res 20:1889–1903
Decharme B, Douville H (2006) Introduction of a sub-grid hydrology in the ISBA land surface model. Clim Dyn 26:65–78
Decharme B, Douville H, Boone A, Habets F, Noilhan J (2006) Impact of an exponential profile of saturated hydraulic conductivity within the ISBA LSM: simulations over the Rhône basin. J Hydrometeorol 7:61–80
Dirmeyer PA (2000) Using a global soil wetness dataset to improve seasonal climate simulation. J Clim 13:2900–2922
Dirmeyer PA (2001) An evaluation of the strength of land–atmosphere coupling. J Hydrometeorol 2:329–344
Dirmeyer PA, Zeng FJ (1999), Precipitation infiltration in the simplified SiB land surface scheme. J Meteor Soc Jpn 77:291–303
Dirmeyer PA, Dolman AJ, Sato N (1999) The Global Soil Wetness Project : a pilot project for global land surface modeling and validation. Bull Am Meteor Soc 80:851–878
Dirmeyer PA, Gao X, Oki T (2002) The Second Global Soil Wetness Project GSWP2: science and implementation plan. IGPO publication Series 37, 65 pp, International GEWEX Project Office
Dirmeyer PA, Gao X, Zhao M, Guo Z, Oki T, Hanasaki N (2006) The Second Global Soil Wetness Project (GSWP-2): multi-model analysis and implications for our perception of the land surface. Bull Am Meteor Soc (accepted)
Douville H (1998) Validation and sensitivity of the global hydrologic budget in stand-alone simulations with the ISBA land-surface scheme. Clim Dyn 14:131–171
Douville H (2002) Influence of soil moisture on the Asian and African monsoons. Part II: Interannual variability. J Clim 15:701–720
Douville H (2003), Assessing the influence of soil moisture on seasonal climate variability with AGCMs. J Hydrometeorol 4:1044–1066
Douville H, Chauvin F (2000) Relevance of soil moisture for seasonal climate predictions: a preliminary study. Clim Dyn 16:719–736
Douville H, Royer J-F, Mahfouf J-F (1995) A new snow parameterization for the Météo-France climate model. Part I: validation in stand-alone experiments. Clim Dyn 12:21–35
Douville H, Chauvin F, Broqua H (2001) Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation. J Climte 14:2381–2403
Ducharne A, Koster DR, Suarez MJ, Stieglitz M, Kumar P (2000) A catchement-based approach to modeling land surface process in a general circulation model, 2, Parameter estimation and model demonstration. J Geophys Res 105:24823–24838
Dümenil L, Todini E (1992) A rainfall-runoff scheme for use in the Hamburg climate model. Adv Theor Hydrol 9:129–157
Durand Y, Brun E, Méridol L, Guyomarc'h G, Lesaffre B, Martin E (1993) A meteorological estimation of relevant parameters for snow schemes used with atmospheric models. Ann Glaciol 18:65–71
Ek MB, Mitchell KE, Lin Y, Grunmann P, Rogers E, Gayno G, Koren V, Tarpley JD (2003) Implementation of the upgraded Noah land-surface model in the NCEP operational mesoscale Eta model. J Geophys Res 108, 8851. DOI:10.1029/2002JD003296
Entin J. K., A. Robock, Vinnikov KYa, Zabelin V, Liu S, Namkhai A (1999) Evaluation of Global Soil Wetness Project soil moisture simulations. J Meteor Soc Jpn 77:183–198
Essery RL, The best MJ, Betts A, Cox PM, Taylor CM (2003), Explicit representation of subgrid heterogeneity in a GCM land surface scheme. J Hydrometeorol 4:530–543
Etchevers P, Colaz C, Habets F (2001) Simulation of the water budget and the rivers flows of the Rhône basin from 1981 to 1994. J Hydrol 244:60–85
Fekete BM, Vörösmarty CJ, Road JO, Willmott CJ (2003), Uncertainties in Precipitation and their impacts on runoff estimates. J Clim 17:294–304
Gedney N, Cox PM (2003), The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity. J Hydrometeorol 4:1265–1275
Gusev, Ye M, Nasonova ON (2003) The simulation of heat and water exchange in the boreal spruce forest by the land-surface model SWAP. J Hydrol 280:162–191
Habets F, Noilhan J, Golaz C, Goutorbe JP, Lacarrère P, Leblois E, Ledoux E, Martin E, Ottlé C, Vidal-madjar D (1999a) The ISBA surface scheme in a macroscale hydrological model applied to the HAPEX-MOBILHY area. Part I: model and database. J Hydrol 217:75–96
Habets F, Etchevers P, Golaz C, Leblois E, Ledoux E, Martin E, Noilhan J, Ottlé C (1999b) Simulation of the water budget and the river flows of the Rhône basin. J Geophys Res 104:31145–31172
King D, Lebas C, Jamagne M, Hardy R, Draoussin J (1995) Base de données géographiques des sols de France à l'échelle 1/1000000 (Geographical Soil Database for France at a scale of 1/1000000). Technical report, 100 pp, Institut National de Recherches Agronomiques (INRA), orleans, France
Koster DR, Oki T, Suarez MJ (1999), The off-line validation of land surface models: assessing success at the annual timescale. J Meteor Soc Jpn 77:257–263

Koster DR, Suarez MJ, Heiser M (2000a), Variability and predictability of precipitation at seasonal to interannual timescales. Hydrometeorol 1:26–46

Koster DR, Suarez MJ, Ducharne A, Stieglitz M, Kumar P (2000b) A catchement-based approach to modeling land surface process in a general circulation model, 1, Model structure. J Geophys Res 105:24809–24822

Koster DR, Dirmeyer PA, Hahmann AN, Ijpelaar R, Tyahla L, Cox P, Suarez MJ (2002) Comparing the degree of land-atmosphere interaction in four atmospheric general circulation models. J Hydrometeorol 3:363–375

Lohmann D, Coauthors (1998) The Project for intercomparison of land-surface parameterization schemes (PILPS) Phase-2c Red-Arkansas River Bassin experiment: III. Spatial and temporal analysis of water fluxes. Global Planet Change 19:161–180

Lott F, Miller MJ (1997) A new subgrid-scale orographic drag parametrization: Its formulation and testing. Quart J R Meteor Soc 123:101–127

Mahfouf J-F, Manzi AO, Noilhan J, Giordani H, Déqué M (1995) The land surface scheme ISBA within the Météo-France climate model ARPEGE. Part I: Implementation and preliminary result. J Clim 8:2039–2057

Mahfouf J-F, Noilhan J (1996) Inclusion of gravitational drainage in a land surface scheme based on the force-restore method. J Appl Meteor 35:987–992

Manabe S (1969) Climate and ocean circulation 1. The atmospheric circulation and the hydrology of the earth's surface. Mon Weather Rev 97:739–805

Masson V, Champeaux JL, Chauvin F, Mériguet C, Lacaze R (2003) A global database of land surface parameters at 1 km resolution for use in meteorological and climate models. J Clim 16:1261–1282

Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int J Climatol 25:693–712

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, 1, A discussion of principales. J Hydrol 10:282–290

Ngo-Duc T, Polcher J, Laval K (2005) A 53-year forcing data set for land surface models. J Geophys Res 110, D06116. DOI:10.1029/2004JD005434

Noilhan J, Planton S (1989) A simple parameterization of land surface processes for meteorological models. Mon Weather Rev 117:536–549

Noilhan J, Mahfouf J-F (1996), The ISBA land surface parametrization scheme. Global Planet Change 13:145–159

Oki T, et Sud YC (1998), Design of total runoff integrating pathways (TRIP). A global river chanel network. Earth Interact 2:1–36. (http://www.EarthInteractions.org/)

Oki T, Nishimura T, Dirmeyer P (1999) Assessment of annual runoff from land surface models using total runoff integrating pathways (TRIP). J Meteor Soc Jpn 77:235–255

Robock A, Vinnikov KY, Srinivasan G, Entin JK, Hollinger SE, Speranskaya NA, Liu S, Namkhai A (2000) The Global Soil Moisture Data Bank. Bull. Amer. Meteorol. Soc 81:1281–1299

Sud YC, Mocko DM (1999) New snow-physics to complement SsiB Part I: Design and evaluation with ISLSCP Initiative I datasets. J Meteor Soc Jpn 77:335–348

Vérant S, Laval K, Polcher J, De Castro M (2004) Sensitivity of the continental hydrological cycle to the spatial resolution over the Iberian peninsulia. J Hydrometeorol 5:267–285

Wood EF, Coauthors (1998) The Project for intercomparaison of land–surface parameterization schemes (PILPS) Phase-2c Red-Arkansas River Basin experiment: I. Experiment description and summary intercomparisons. Global Planet Change 19:115–135

Wood EF, Lettenmaier DP, Zartarian VG (1992) A land-surface hydrology parameterization with subgrid variability for general circulation models. J Geophys Res 97:2717–2728

Xie, Arkin (1996), Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. J. Clim 9:840–858

Yang D, Ohata T (2001) A bias-corrected siberian regional precipitation climatology. J Hydrometeorol 2:122–139

Yang D, Kane D, Zhang Z, Legates D, Goodison B (2005) Bias corrections of long-term (1973–2004) daily precipitation data over the northern regions. Geophys Res Lett 32, L19501. DOI:10.1029/2005GL024057

Ye B, Yang D, Ding Y, Han T, Koike T (2004) A bias-corrected precipitation climatology for China. J Hydrometeorol 5:1147–1160

Zhao RJ (1977) The Xinanjiang model applied in China. J Hydrol 134:317–381

Zhao M, Dirmeyer PA (2003) Production and analysis of GSWP-2 Near-Surface Meteorology Data Sets. COLA Technical report no 159 (Availabe online at http://www.grads.iges.org/gswp/), 36 pp, Center for Ocean-Land-Atmosphere studies, Calverton, US