Responses to comments from anonymous Referee 1

On “Should altitudinal gradients of temperature and precipitation inputs be inferred from key parameters in snow-hydrological models?” by D. Ruelland (HESS-2019-556)

Referee’s comment

This paper presents a nicely conceived study where several alternative approaches for distributing temperature and precipitation are compared in a mountain region (French Alps). In addition to standard interpolation approaches based on inverse-distance weighting and Kriging, the author explores the possibility of optimizing lapse rates as part of a snow-hydrologic-model calibration procedure. Results of a split-sample test show that the latter approach provides improved results for the target variables considered during calibration, that is, fractional snow cover (FSC) from MODIS, streamflow, and the water balance. Also, optimizing the temperature and precipitation distribution algorithm with hydrologic data results in the temperature and precipitation fields being colder and wetter than those obtained by using only in-situ measurements of temperature and precipitation, respectively; this result agrees with expectations, especially since the considered ground-based network is not representative of high elevations in the study catchments.

I enjoyed reading the manuscript and I think it represents an interesting contribution for HESS, especially because the investigated topic is a clear open issue in mountain hydrology. I do have several general and specific comments, which I attach below. Overall, I think that the revision is feasible.

Authors’ response

I would like to sincerely thank the referee for the time and effort he/she spent in reading the initial manuscript and for making many clear, pertinent and constructive suggestions for improvement. This helped a lot to re-write the paper.

General comments

Referee’s comment

1. in my understanding, the author essentially compares two main strategies: the first is to regionalize temperature and precipitation based on in-situ data and various interpolation/extrapolation schemes based on IDW or Kriging (Section 3); the second is to ‘’embed” part of the distribution process into the snow-hydrologic models via lapse rates that will correct a first-guess distribution based on IDW (see Section 4, Table 2 and 5 and Equations 9 and 10). While the first approach is independent from hydrologic data like fractional snow cover and streamflow, the second does take advantage of these data to adjust some of the distribution parameters. In my understanding, the main point of the paper is that the second strategy is superior to the first, especially since adjusting snow parameters rather than precipitation-distribution parameters does not allow the model to significantly improve its performance (see Table 5 and Fig. 5). Unless I am missing something here, this improvement was however assessed based on the same hydrologic variables that were used to calibrate the snow-hydrologic models, rather than on independent measurements of the two variables of interest: temperature and precipitation. This left me wondering if this experiment shows that “calibrating the local gradients using an inverse snow-hydrological modelling framework” improves actual temperature and precipitation estimates, or if it shows that it improves hydrologic predictions. In principle, one would expect the obtained altitudinal gradients to be both more effective in terms of hydrologic predictions and in terms of temperature and precipitation, but the improvement obtained by ‘’embedding” part of the distribution process into the snow-hydrologic models is quantified in terms of modeling skills for fractional snow cover, streamflow, and the water balance (Figs. 5 to 8) rather than for independent estimates of temperature and precipitation. If independent data of temperature and precipitation at high elevations are not available, then I would recommend the

author to clarify the extent to which these results apply to temperature and precipitation in addition to hydrologic variables.

Authors' response and modifications to manuscript

I agree that an improved fit for hydrologic variables may not automatically mean that the model is also better representing weather patterns of temperature and precipitation. Since independent data of temperature and precipitation at high elevations were indeed not available, it was not possible to clarify the extent to which these results apply to temperature and precipitation in addition to hydrologic variables. As a result and following the relevant referee comment and argumentation, the following text has been added in the Section 6.2. Recommendations:

"….However the differences in the two compared approaches are worth discussing. The first is to regionalize temperature and precipitation based on in-situ data and various interpolation/extrapolation schemes based on IDW or Kriging; the second is to "embed" part of the distribution process into the snow-hydrologic models via calibrated lapse rates correcting a first-guess distribution based on IDW. While the first approach is independent from hydrological data like fractional snow cover and streamflow, the second does take advantage of these data to adjust some of the distribution parameters. The second strategy prove superior to the first, especially since calibrating distribution parameters rather than adjusting snow parameters allowed the models to significantly improve their performance. This improvement was however assessed based on the same hydrological variables that were used to calibrate the snow-hydrologic models, rather than on independent measurements of temperature and precipitation. This left wondering if improving hydrologic predictions by calibrating the local gradients using an inverse snow-hydrological modelling framework also improves actual temperature and precipitation estimates. In principle, one would expect the obtained altitudinal gradients to be both more effective in terms of hydrologic predictions and in terms of temperature and precipitation, but the improvement obtained by "embedding" part of the distribution process into the snow-hydrologic models was only quantified in terms of modelling skills. An improved fit for hydrologic variables may not automatically mean that the model is also better representing weather patterns of temperature and precipitation. A good example is that the optimized lapse rates (Fig. 9) can locally be different between the two hydrologic models considered. Since independent data of temperature and precipitation at high elevations are not available, it was not possible to clarify the extent to which these results apply to temperature and precipitation in addition to hydrologic variables."

Referee's comment

2. The point above is particularly important since hydrologic models may suffer from several sources of conceptual and parametric uncertainties, some of which are visible in the interesting Figure 8. It follows that an improved fit for hydrologic variables may not automatically mean that the model is also better representing weather patterns of temperature and precipitation. A good example here is that the obtained lapse rates (Fig. 9) can locally be quite different between the two hydrologic models considered. To me, this may challenge the idea that this approach could be used to "infer local altitudinal gradients from a sparse network of gauges based on key parameters in the snow-hydrological models" (L 592ff). It does suggest that the method improves hydrologic predictions, but implications for actual temperature and precipitation are more elusive to me and should be discussed more extensively.

Authors' response and modifications to manuscript

Agreed. The sentence (L 592 ff) has been removed and an entire paragraph (based on the argumentation from the referee in his/her general comment #1 and #2 has been added in the Section 6.2. Recommendations (see answer to the preceding comment).

Referee's comment

3. Related to this, both hydrologic models were used in lumped mode (L340), even if several other modeling approaches explicitly account for spatial variability in hydrologic processes (e.g., raster-based models). At least some discussion on this point would be interesting.

Authors' response and modifications to manuscript

Agreed. This is part of an on-going research. To address the referee comment, the following text has been added at the end of the conclusion section (6.3. Prospects):

"…Finally, it is worth mentioning that spatial variability was only considered along five elevation bands in each catchment since preliminary tests showed no improvement in the hydrologic predictions when applying the SAR in a full distribution mode. However, the SAR was not designed to explicitly account for topographic effects (slope, aspect and shading) on snow redistribution, accumulation and melt (see e.g. Frey and Holzmann, 2015). For instance, a grid-based temperature-index model could be implemented to include potential clear-sky direct solar radiation at the surface, thus considering both the seasonal variations of melt rates and the geometric effects on melt attributable to terrain (see e.g. Hock, 1999). It would thus be interesting to assess whether accounting for the influence of such effects can further improve the daily hydrologic predictions at the basin scale."

Referee's comment

4. Spatial variability was considered in the snow model, which was implemented along five elevation bands in each catchment. This model does include all fundamental snow processes, but in my understanding does not include a specific provision for wind drift. Relying on FSC from MODIS may sometimes lead to confounding effects in this regard, where wind-driven accumulation and erosion is mistakenly assumed as due to precipitation or melt. Was this somehow taken into account here, or could the author suggest how to include this in the framework?

Authors' response and modifications to manuscript

The snow accounting routine (SAR) is only based on (semi-)distributed temperature and precipitation inputs to simulate the main snow processes related to accumulation and melt. As a result, it does not include a specific provision for wind drift within each catchment, which would probably require an additional distributed input regarding wind speed and direction. Given the challenge to distribute temperature and precipitation from a sparse gauge network, distributing wind (from even more scarce measures than for temperature/precipitation) may be unrealistic at the spatio-temporal scales considered (daily analysis in the French Alps over the period 1998–2016). It would also probably require applying the SAR in a full distribution mode. But, even doing so, it is unlikely that accounting for wind-driven accumulation and erosion would have a significant impact on the streamflow and FSC simulations at the basin scale, because it can be assumed that these processes are somewhat averaged at the basin scale. Although these aspects are worth discussing, it seems difficult to integrate them into the text without adding weight and making the discussion too large and complex. However, in link with the preceding referee comment, note that the following text (dealing with spatial variability) has been added at the end of the conclusion section:

"…Finally, it is worth mentioning that spatial variability was only considered along five elevation bands in each catchment since preliminary tests showed no improvement in the hydrologic predictions when applying the SAR in a full distribution mode. However, the SAR was not designed to explicitly account for topographic effects (slope, aspect and shading) on snow accumulation, redistribution and melt (see e.g. Frey and Holzmann, 2015). For instance, a grid-based temperature-index model thus be implemented to include potential clear-sky direct solar radiation at the surface, thus considering both the seasonal variations of melt rates and the geometric effects on melt attributable to terrain (see e.g. Hock, 1999). It would thus be interesting to assess whether

accounting for the influence of such effects can further improve the daily hydrologic predictions at the basin scale."

Referee's comment

5. I am also interested in the different outcomes of this analysis for precipitation between the daily and the annual time scales (table 4). Maybe one key to interpret this result is that summer vs. winter precipitation patterns are different, and the in-situ network might be more representative of the former than of the latter (or vice versa). I am thinking to convective precipitation here, which sometimes show significantly different elevational gradient from stratiform or orographic precipitation. Some more discussion on precipitation regimes could be interesting in this paper.

Authors' response and modifications to manuscript

I could not see how the analysis of seasonal patterns of precipitation (see Figure 1 below) can explain why interpolation performance is improved by the external drift at the annual (and also monthly) time scale, while it is not at the daily time scale. As pointed out in the text, this shows that the correlation between precipitation and topography increases with the increasing time aggregation as already reported in other studies (e.g., Bárdossy and Pegram, 2013; Berndt and Haberlandt, 2018). The elevation-dependency of precipitation thus depends significantly on the accumulation time. At the daily time scale, the orographic enhancement is limited because on a given day there is no monotonic relationship between elevation and precipitation amount: it depends on where the precipitation event occurs in the first place.

The two preceding sentences were added in the Section 5.1 to try to better explain the different outcomes for precipitation between the daily and the annual time scales (Table 4). The following sentence was also added in the Section 6.3: "The problem is even more challenging for precipitation whose lapse rates could not be related to seasonal or other types of systematic variations as they are strongly dependent on the synoptic meteorological conditions and therefore highly variable."

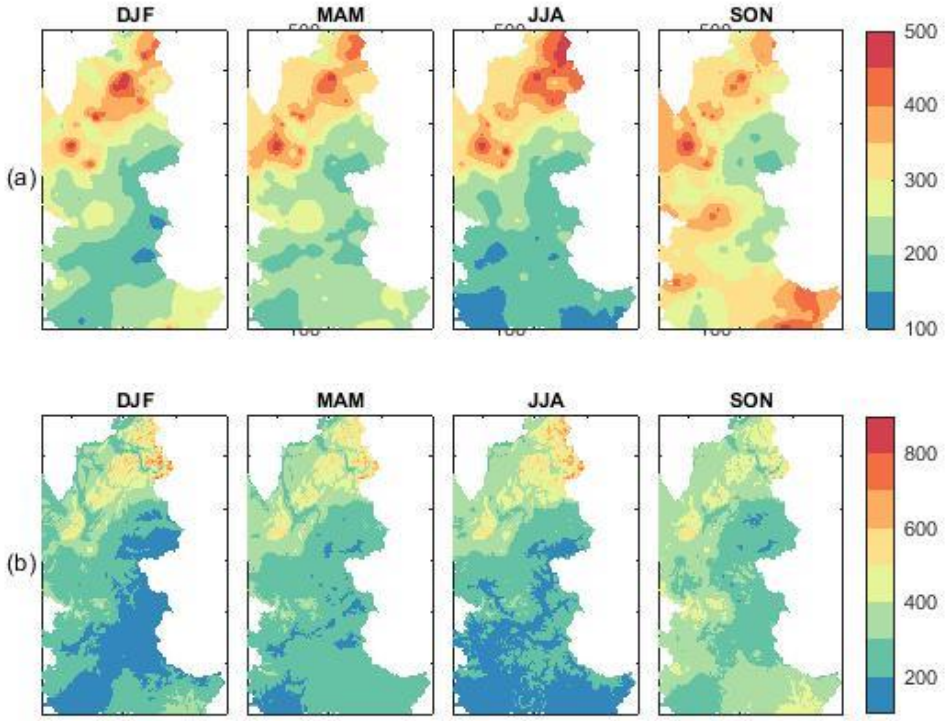


Fig. 1 Seasonal patterns of precipitation with (a) IDW applied to in-situ network (without elevation dependency) and (b) IED applied to in-situ network (with elevation dependency via external drift). Values are in mm per season (i.e. 3 months).

Specific comments

Referee's comments in grey – Authors' responses and modifications to manuscript in blue

Line 149: is this because no gap was originally present in the dataset, or because these gaps were filled? If the second, maybe briefly mention how.
These meteorological gauges were selected because no gap was originally present in their original time series from the 1st September 2000 to the 31st of August 2016. The text has been modified to make the purpose clearer.

Section 2.1: a histogram with the elevation distribution of in-situ stations may be helpful, along with more details about the climatology of the study period (annual mean temperature and precipitation, annual runoff etc). Doing so may help the author to set the context of the analysis, especially for non-local readers.
Done. Figure 1 was modified to incorporate elevation distributions of in-situ stations, DEM and basins (see below). Details about the "estimated" climatology of the study period are now provided in Table 1 (see below), which has been modified to include mean annual temperature (T), total precipitation (P), snowfall fraction (S) and streamflow (Q) for each basin. Note however that these values are very delicate to provide since they necessarily rely on approximations depending on the method used to distribute temperature and precipitation (in link with the paper issue). This is why they were not included in the initial submitted paper. As indicated in the modified caption of Table 1, catchment areal temperature, total precipitation and snowfall fraction were estimated after calibrating local altitudinal gradients over 2000–2016 using the snow-hydrological inverse approach proposed in the current paper (see Test #4 in Table 5).

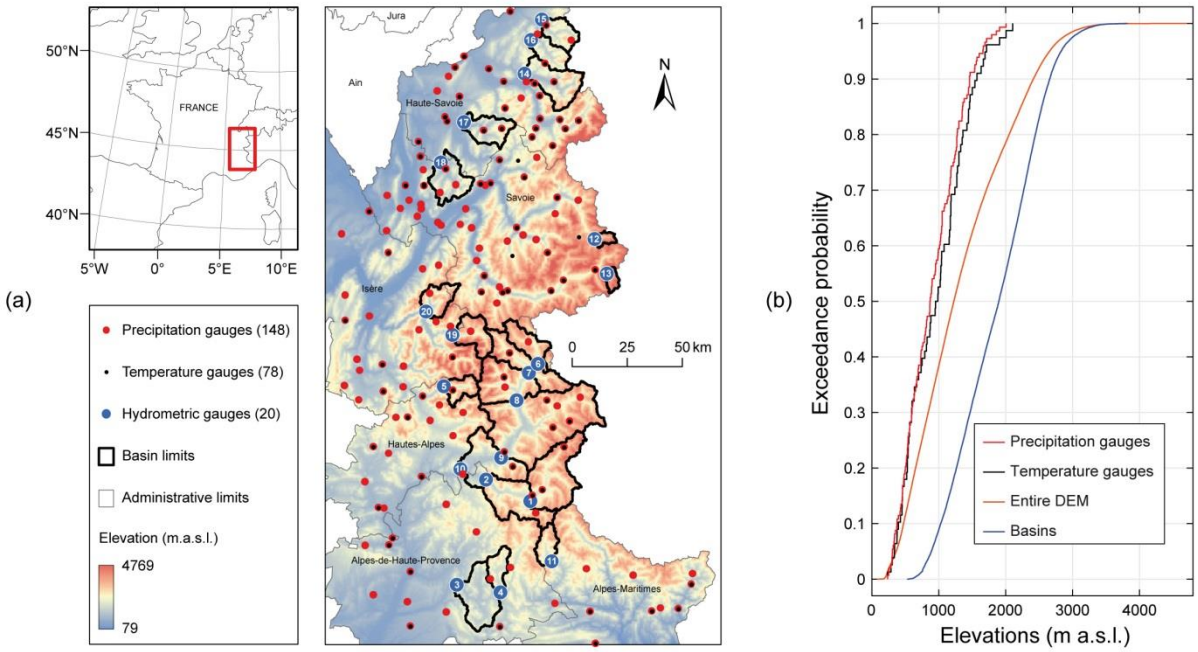


Fig. 1 Study area and data: (a) Location of the selected precipitation, temperature and streamflow stations, as well as elevations from a SRTM digital elevation model (DEM resampled to a grid with 0.5x 0.5 km cells) in the French Alps; (b) Elevation distributions of in-situ stations, DEM and basins. The station numbers refer to Table 1.

Table 1 Streamflow gauging stations and main catchment characteristics. Percentages of glacierized area were estimated from the World Glacier Inventory (NSIDC, 2012). Mean annual precipitation (*P*), snowfall fraction (*S*) and temperature (*T*) were estimated after calibrating local altitudinal gradients over 2000–2016 using the snow-hydrological inverse approach proposed in the current paper (see Test #4 in Table 5).

#	Station	River	Area	Glacierized area	Elevations (m.a.s.l.)		Mean annual precip. (*P*)	Snowfall fraction (*S*)	Mean annual temp. (*T*)	Mean annual streamflow (*Q*)
			(km²)	(%)	Min	Max	(mm/year)	(%)	(°C)	(mm/year)
1	Barcelonnette	Ubaye	549	0	1132	3308	856	49	1.9	521
2	Lauzet-Ubaye	Ubaye	946	0	790	3308	979	45	3.0	654
3	Beynes	Asse	375	0	605	2273	921	15	8.9	344
4	Saint-André	Issole	137	0	931	2392	1013	23	7.0	481
5	Villar-Lourbière	Séveraisse	133	4	1023	3623	1781	49	2.4	1317
6	Val-des-Prés	Durance	207	0	1360	3059	847	55	0.9	688
7	Briançon	Durance	548	1	1187	3572	811	52	1.6	714
8	Argentière-la-Bessée	Durance	984	3	950	4017	1064	52	2.3	765
9	Embrun	Durance	2170	2	787	4017	1075	48	3.1	693
10	Espinasses	Durance	3580	1	652	4017	1054	45	3.5	654
11	Villeneuve	Var	132	0	926	2862	1072	37	4.6	650
12	Val-d'Isère	Isère	46	9	1831	3538	1245	63	-1.5	1119
13	Bessans	Avérole	45	12	1950	3670	1635	66	-2.3	1311
14	Taninges	Giffre	325	0	615	3044	2315	35	5.0	1771
15	Vacheresse	Dranse d'Abondance	175	0	720	2405	1671	30	4.8	1088
16	La Baume	Dranse de Morzine	170	0	690	2434	1637	32	4.6	1285
17	Dingy-Saint-Clair	Fier	223	0	514	2545	1649	26	6.6	1243
18	Allèves	Chéran	249	0	575	2157	1486	23	6.9	819
19	Mizoën	Romanche	220	9	1057	3846	1369	56	0.9	978
20	Allemond	L'Eau Dolle	172	2	713	3430	1553	47	2.4	1164

Section 2.2: is any of these catchments glacierized? If so, how were glaciers considered in this framework? If not, may glaciers hamper the applicability of this method in other regions, especially with regard to the mass-balance-closure term in Eq. 12 and Fig. 6?

During the catchment selection process, I tried to minimize possible interactions with non-snow related processes that could also influence streamflow. Therefore, we tried avoiding glacierized basins, basins with known inter-catchment groundwater flows, and catchments with documented flow diversions. However, it must be acknowledged that some basins are partly glaciated. Table 1 now includes the percentage of glaciated areas estimated for each catchment from the World Glacier Inventory (NSIDC, 2012).

The following sentence was also added in the section 4.3.2: "As the majority of basins had negligible glacierized areas (see Table 1), no specific glacier model was activated. This led to ignore the late summer contribution of glacier melt to river discharge in the three basins having 9–12% glacierized areas".

This did not affect significantly the mass-balance-closure term at the annual scale. However, it is worth mentioning that more important contribution of ice melt would require a glacier component to not hamper the use of WB in the objective function. However, please note that the objective function was changed to address a comment from the referee 2 as regards to the mass-balance-closure term. The simulations were re-run with a more classical objective function (i.e. without using the WB constraint) based only on NSE_{FSC} and NSE_{sqrtQ}. The results were the same as regards to the modelling distribution performances between the various tests. Please also note that this did not change the main findings of the paper (see detailed answers to the referee2 comments #9 and #12).

Section 2.3: the approach by Gascoin et al. 2015 was, to my knowledge, developed in the Pyrenees, a mountain range with significantly lower elevations than the Alps. How was the method adapted for the French Alps? Is the performance similar to that originally published by Gascoin et al. 2015 in a different mountain range?

The referee is right. The approach by Gascoin et al. (2015), itself inspired by previous works from Parajka and Blöschl (2008) in Austria, and Gafurov and Bárdossy (2009) in Afghanistan, was adapted for the French Alps. To do so, the MODIS snow-products were gap-filled in various mountain ranges (including the Pyrenees and the French Alps). The missing data (due mainly to cloud obscuration) were less important in the Pyrenees (46%) than in the Alps (52%) for the same period (2000–2016). As a result, for the temporal deduction by sliding the time filter, we allowed the window size to be incremented up to 9 days in the Alps (versus up to 6 days in the Pyrenees) to account for the differences in cloud obscuration. Moreover, Gascoin et al. (2015) used an adjacent spatial deduction as a second step: each no-data pixel was reclassified as snow (no-snow) if at least five of the eight adjacent pixels were classified as snow (no-snow). We did not use this step considering it was a too rough approximation. Finally, we validated our adapted method on the two mountain ranges. Validation based on confusion matrices with 1 image/month (i.e. about 200 cloud-free images over the studied period) showed that the gap-filling technique applied to the MODIS snow-products led to the reconstruction of validation images with average accuracies of 98% in the Pyrenees and 94% in the Alps.

To address the referee comment, more details regarding the gap-filling method and validation have been given through additional text and a new figure (see below) in the Section 2.3.

"MOD10A1 (Terra) and MYD10A1 (Aqua) snow products version 5 were downloaded from the National Snow and Ice Data Center for the period 24 February 2000–1 January 2017. This corresponds to 6157 dates among which 98.8% are available for MOD10A1 and 85.8% for MYD10A1 since Aqua was launched in May 2002 and became operational in July 2002. These snow products are derived from a Normalised Difference Snow Index (NDSI) calculated from the near-infrared and green wavelengths, and for which a threshold was defined for the detection of snow (Hall et al., 2006, 2007). Cloud cover represents a significant limit for these products, which are generated from instruments operating in the visible-near-infrared wavelengths. As a result, the grid cells were gap filled to produce daily cloud-free snow cover maps of the study area. The different classes in the original products were first merged into three classes: no-snow (no snow or lake), snow (snow or lake ice), no-data (clouds, missing data, no decision, or saturated detector). The missing values were then filled according to a gap filling algorithm inspired by techniques proposed in the literature (Parajka and Blöschl, 2008; Gafurov and Bárdossy, 2009; Gascoin et al., 2015). The algorithm works in three sequential steps:

(i) Aqua/Terra combination: for every pixel, if no-data was found in MOD10A1 then the value from MYD10A1 was used instead. Priority was given to MOD10A1 because MYD10A1 was found to be less accurate (see Gafurov and Bárdossy, 2009).

(ii) Temporal deduction by sliding time filter: a no-data pixel was reclassified as snow (no-snow) if the same pixel was classified as snow (no-snow) in both the preceding and following grids. The preceding and following grids were searched within a sliding temporal window, whose size was incremented up to 9 days in order to reduce the remaining fraction of no-data pixels to below 10% at least (Fig. 2a). It should be noted that three periods of gaps in an upper time window (11, 13 and 18 days) were present in the data because of technical failures of the MODIS sensor. In these cases, a longer time deduction was used beforehand to specifically fill these periods.

(iii) Spatial deduction based on elevation and neighbourhood filter: for each date and each pixel, a 3x3 neighbourhood spatial filter was used to account for the elevation and the data in the neighbouring pixels to fill the remaining no-data pixels. Two configurations were considered: either the central pixel has no-data and the algorithm tries to attribute a neighbouring value,

or the central pixel has a value that can be assigned to some of its neighbours. The two configurations were repeated until there were no more gaps (Fig. 2a).

The resulting database consists of 5844 binary (snow/no-snow) maps at 500 m spatial resolution for the period 2000–2016 (16 hydrological years, from the 1st of September 2000 to the 31st of August 2016). As a synthesis of these maps, snow cover durations over the study area are presented in Fig. 2c.

In order to validate the gap-filling technique, a daily snow product with less than 10% of no-data pixels was selected for each month of the studied period. These images were "blackened" (i.e. with 100% no-data pixels), before applying the algorithm over the entire period to fill all gaps, including validation images. Filling accuracy was estimated for each image removed by computing confusion matrices which compared the pixels of the removed validation images and the filler reconstructions of these images. Validation based on confusion matrices with 1 image/month showed that the gap-filling technique applied to the MODIS snow-products led to the reconstruction of images with average accuracies of 94% (Fig. 2b). The mean monthly accuracies show greater ease in filling gaps in summer than in winter due to the differences in cloud obscuration. However it should be noted that the actual accuracy of the MODIS gap-filling technique is necessarily greater than that of the validation procedure, in which many quality images needed to fill the gaps were missing."

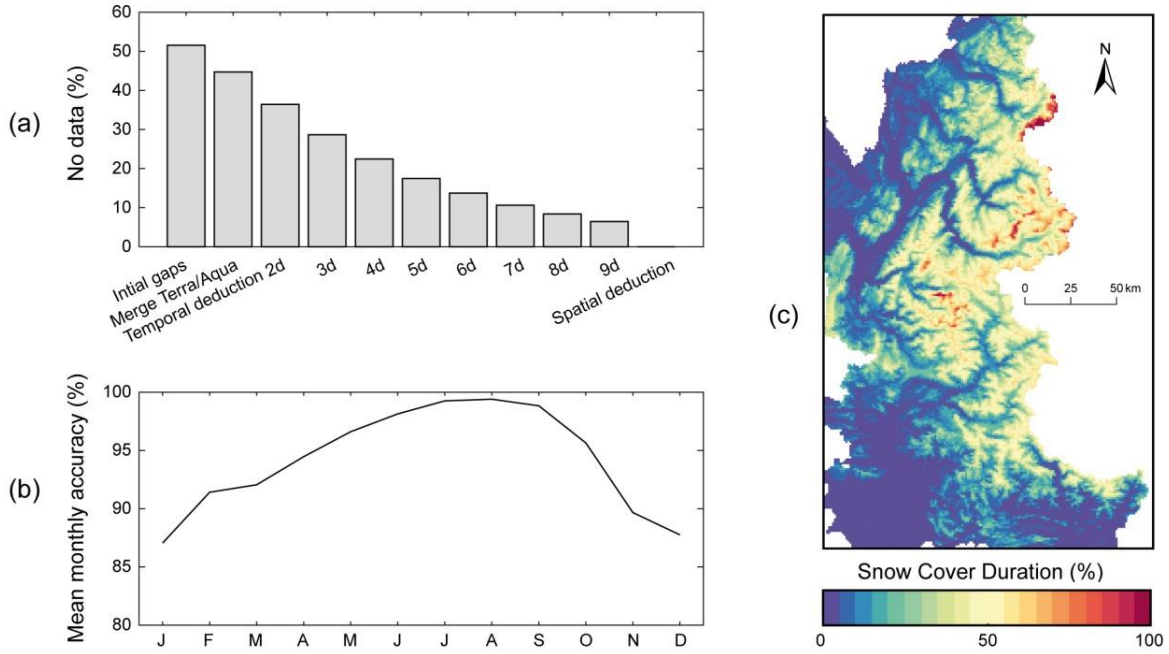


Fig. 2 Results of gap-filling applied to MODIS snow products: (a) Evolution of the number of pixels classified as no-data (e.g., clouds) during the gap-filling procedure; (b) Mean monthly accuracies according to validation based on confusion matrices with 1 image/month, i.e. ~200 cloud-free images over the 2000–2016 period; (c) Snow cover duration based on gap-filled MODIS snow products over the 2000–2016 period.

Line 263ff: was mean precipitation computed across the whole study region? Might doing so exclude more localized precipitation events in favor of more widespread stratiform events?

In the initial submission, parameters were only estimated for days with at least 1 mm mean precipitation, i.e. approximatively 41% of the daily sample (whereas parameters were calculated for all months and years since there were no locations with dry months or dry years in the dataset). The 1 mm mean precipitation was indeed computed across the whole study region. The initial motivation was to limit the effect of precipitation zero values during the optimization of the daily local regression between precipitation and elevation for external drift. Doing so might indeed bias the optimization of the *n(u)* surrounding observations during the leave-one-out cross-validation, thus potentially affecting the computation of the external drift with the KED and IED technique when

further used to interpolate daily precipitation in the study region (Fig. 5 – Fig. 4 in the initial submission).

Following the referee comment, the leave-one-out cross-validation was re-run with the whole set of daily samples using the four interpolation techniques (IDW, ORK, KED and IED). As a result, the RMSE, MAE and NSE values changed for the cross-validation of the interpolation methods against precipitation daily series (see new values in Table 4). However, interestingly, this did affect neither the ranking between methods nor the optimized interpolation parameters. For instance, 17 surrounding neighbors were still found during optimization to compute altitudinal gradients of precipitation based on the daily, local linear regression with KED and IED. This means that the initial results (optimized parameters and daily interpolated precipitation) were not affected by the 41% subset. In other words, the subset did not exclude more localized precipitation events in favor of more widespread events. However, as it revealed unnecessary, the sentence about it in the section 3.3 was deleted and the RMSE, MAE and NSE scores were changed in Table 4 accordingly to the computation of the cross-validation with the all daily precipitation samples.

Title of Section 4: ASSESSSMENT -> ASSESSMENT.
Corrected in the revised manuscript.

Section 4: a table with the list of all parameters considered by the snow and hydrologic models would be helpful, including an explicit statement of which parameters where calibrated. Some of these parameters are only mentioned at the very end of the manuscript (Section 5.3).
The referee probably missed Table 2 that lists all (fixed or calibrated) parameters regarding the snow accounting routine (SAR) and Table 3 that lists all calibrated parameters regarding the two hydrological models. The two models were run with the SAR on top. Since the SAR parameters were calibrated differently (from 2 to 5 free parameters) in each modelling experiment, it was necessary to introduce two separated tables for the parameters. Note however that these two tables are presented in two consecutive sections, making them theoretically understandable.

Line 321: this should in fact be evaporation to me, since there is no transpiration in the snow module (correct?)
There is no transpiration in the snow module. However, PE is computed based on the temperature-based formulation proposed by Oudin et al. (2005) which explicitly mentioned the "evapotranspiration" term ("…towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling"). The formula aims at estimating the maximum amount of evapotranspirable water taking into account the meteorological context and for a plant cover corresponding to grass. As a result, it is common to use the term potential evapotranspiration when one refers to the Oudin formula.

Section 4.3.1: where the two periods similar in terms of snow conditions and streamflow, as well as mean temperature and mean precipitation across the in-situ network?
Mean annual precipitation increased by around 17% between the two periods (1104 mm/yr vs. 1294 mmm/yr), while mean annual temperature were stable (9.2 °C vs. 9.3 °C) across the in-situ network presented in Section 2.1. Although the second period was generally wetter, this hides differences in between the catchments. At the basin scale, the differences between the two periods ranged from -10% to 15% for precipitation, -0.5 °C to +0.5 °C for temperature, and -11% to +50% for streamflow. These details are now indicated in Section 4.3.1.

As far as the snow conditions are concerned, differences in mean annual snowfall between the two periods ranged from -10% to +50% according to the best-performing simulations. As these values were obtained through simulations, I feel it is not correct to mention them in Section 4.3.1. I hope the above values regarding precipitation, temperature and streamflow between the two periods are sufficient to address the referee comment. The following text was added in the beginning of Section 4.3.1:

"Mean annual precipitation increased by around 17% between the two periods, while mean annual temperature were stable (9.2 °C vs. 9.3 °C) across the in-situ network presented in Section 2.1. Although the second period was generally wetter, differences can be observed locally. At the basin scale, the differences between the two periods ranged from -10% to +15% for precipitation, -0.5 °C to +0.5 °C for temperature, and -11% to +50% for streamflow."

Eq 12 and Section 4.3.2: does the third component of the OF assume that interannual variability in subsurface storage is negligible? This might not be an issue in the studied area, but it may be worth mentioning this in case interested readers would like to apply this approach somewhere else. In fact, results in Section 5.3 do suggest that interannual sub-surface dynamics are worth discussing.
The referee is right. The third component (WB) of the objective function assumes that interannual variability in subsurface storage is negligible.

As explained above, please note that the objective function was changed to address a comment from the referee 2 as regards to the mass-balance-closure term WB. The simulations were re-run with a more classical objective function (i.e. without using the WB constraint) based only on NSE_{FSC} and NSE_{sqrtQ}. The results were the same as regards to the modelling distribution performances between the various tests. Please also note that this did not change the main findings of the paper (see detailed answers to the referee2 comments #9 and #12). However, the following paragraph (and associated new table) was added in the section 5.3 (Identifiability of the parameters) to further discuss on the IGE (inter-catchment groundwater exchanges) issue and suggest the findings of the initial submission using the WB term in the OF:

"…Equifinality is also reduced in Tests #4–6 for the parameters controlling runoff generation and routing (*X1*, *X3* and *X4*). On the opposite, the parameter of the inter-catchment groundwater flows (*X2*) is poorly identifiable with variation coefficients of 24.8%, 20.3% and 143.1% with Test #4, Test #5 and Test #6, respectively. This suggests that inter-catchment groundwater exchanges (IGE) do not play a key role in the studied catchments. Indeed, fixing *X2* to a value of 0 (i.e. without potential IGE) with an alternative GR3J model provided similar mean validation efficiency on the set of catchments as compared to the GR4J associated with the 2-parameter SAR (Table 7). However, other objective functions may result in other findings as far as IGE are concerned. For instance, additional tests (not shown here for brevity sake) confirmed that it was possible to greatly reduce the X2 equifinality without decreasing the model efficiency by adding a water balance term in the objective function to constrain the proportion of years respecting the water and energy balance in the Turc-Budyko non-dimensional graph (see Andréassian and Perrin, 2012). These tests suggested that it may be relevant to explicitly represent inter-catchment groundwater transfers in association with correcting or scaling factors applied to the precipitation input data to render the distribution between evapotranspiration, streamflow and underground fluxes more realistic, as already reported by Le Moine et al. (2007)."

Table Mean validation efficiency on the set of 20 catchments with the GR4J model and the GR3J model in association with the 2-parameter SAR.

Model	Total number of free parameters	Mean NSE_{SNOW}	Mean NSE_{Q}	Mean NSE_{lnQ}	Mean VE_{C}
2-parameter SAR/GR4J	6 (2 + 4)	0.86	0.79	0.82	0.95
2-parameter SAR/GR3J	5 (2 + 3)	0.86	0.78	0.81	0.94

Section 5.1: do statistics reported in Fig. 4 and at lines 406ff consider areas outside the studied catchments too, including Italy and Switzerland? It might be better to report statistics for the French Alps only here since this is where data were available to this study.

Yes, they do. To address the referee comment, a geographical mask (representing only the six French administrative departments from which data were selected) was applied on the initial maps. Statistics were re-computed accordingly. Note that only means have changed since the minimum and maximum values were still in the masked maps. Note that the same mask was also applied to the DEM in Fig. 1 to make coherent the presentation.

Section 5.2: the first paragraph of this section and Table 5 should be moved to the Methods. It should also be clarified that each re-calibration mode included hydrologic parameters too (correct?)
I acknowledge that this first paragraph and Table 5 could somehow be moved to the Methods. However, as indicated in the text, "for sake of brevity, here we only present the results obtained with the datasets interpolated with the IDW and IED procedures, since cross-validation at the daily time scale showed that they slightly outperformed the ORK and KED methods, respectively". This means that the short "methodological" paragraph and associated Table 5 are strongly linked to the results presented in the previous section 5.1 and cannot be moved to the Methods since they depend on these results. Note also that all the modelling experiment (snow-accounting routine, hydrological models and calibration/validation methods) are fully described in the Methods. Here, we further combine this modelling experiment to part of the results of the interpolation methods to introduce the tests to account for elevation dependency in the T and P inputs via the modelling experiment. Since the tests (#1-6) are rather complex and since their results are immediately presented and discussed in the Section 5.2, I believe that reading is probably easier in that way.

I hope these arguments will convince the referee.

Regarding the second comment, the parameters of the SAR and the hydrological model were indeed optimized simultaneously, as already indicated in Section 4.3.2. Following the referee comment, it was however repeated and clarified that each re-calibration mode included hydrologic parameters too, by adding the following text in the Table 5 caption: "Note that each calibration tests included also the hydrological parameters of GR4J or HBV9 (the parameter ranges tested are listed in Table 2 for the SAR and in Table 3 for the hydrological models)."

Fig. 6: it seems like all data are within the boundaries given by the water and energy limit. I am not an expert of this approach and was wondering why one should aim to obtain "the least stretched and dispersed cluster". More details on this might be helpful for other readers too.
A closer look to the Figure shows that the dots are systematically within the boundaries given by the water and energy limits only with Test #4. It can also be observed that Test #4 led to the least stretched and dispersed cluster, which was not particularly intended. This is only a finding which adds to the more important fact that the dots are all located within the "physical" limits by the water and energy limits.

However, as explained above, all the simulations were re-run using a more classical objective function (i.e. without the WB term) to address comments from referee2. As a result, the water balance was not closed systematically and it was not interesting anymore to present the Budyko graphs for the different tests. The Budyko graphs and associated comments were therefore removed from the manuscript.

Line 490ff and other similar passages of the manuscript: in fact, this result suggests to me that correcting for precipitation and temperature distribution has a stronger impact on model predictions than adjusting for other snow-related processes like phase partitioning or melt, rather than that "adapting to local snow processes is not indispensable". To me, other processes are important too, but correctly estimating total accumulation is likely the most important one here.
Done. The sentence has been modified as follows (see also other modifications in link with other comments notably in Section 6.2 Recommendations):
'…This suggests that correcting for temperature and precipitation distribution has a stronger impact on model predictions than adjusting for snow-related processes like phase partitioning or melt, and

that correctly estimating total accumulation is likely to play a first-order role in the snow-hydrological responses of the studied catchments.'

Section 5.3: I would probably add more details about how parameter identifiability is quantified from Figure 8.

Agreed. Figure 8 was modified to indicate more clearly which parameters belong to the SAR or to the hydrological models (GR4J). Moreover, more details have been provided in the figure caption and in the text. Variation coefficients (in %) of the 20% best parameter solutions compared to the optimised values for each parameter have also been introduced in the figure and in the text to bring more details about how parameter identifiability can be quantified from the figure.

The following paragraph (and associated new table) was also added in the section 5.3 (Identifiability of the parameters) to further discuss on the IGE issue and suggest the findings of the initial submission using the WB term in the OF:

"…Equifinality is also reduced in Tests #4–6 for the parameters controlling runoff generation and routing (*X1*, *X3* and *X4*). On the opposite, the parameter of the inter-catchment groundwater flows (*X2*) is poorly identifiable with variation coefficients of 24.8%, 20.3% and 143.1% with Test #4, Test #5 and Test #6, respectively. This suggests that inter-catchment groundwater exchanges (IGE) do not play a key role in the studied catchments. Indeed, fixing *X2* to a value of 0 (i.e. without potential IGE) with an alternative GR3J model provided similar mean validation efficiency on the set of catchments as compared to the GR4J associated with the 2-parameter SAR (Table 7). However, other objective functions may result in other findings as far as IGE are concerned. For instance, additional tests (not shown here for brevity sake) confirmed that it was possible to greatly reduce the X2 equifinality without decreasing the model efficiency by adding a water balance term in the objective function to constrain the proportion of years respecting the water and energy balance in the Turc-Budyko non-dimensional graph (see Andréassian and Perrin, 2012). These tests suggested that it may be relevant to explicitly represent inter-catchment groundwater transfers in association with correcting or scaling factors applied to the precipitation input data to render the distribution between evapotranspiration, streamflow and underground fluxes more realistic, as already reported by Le Moine et al. (2007)."

Table Mean validation efficiency on the set of 20 catchments with the GR4J model and the GR3J model in association with the 2-parameter SAR.

Model	Total number of free parameters	Mean NSE_{SNOW}	Mean NSE_{Q}	Mean NSE_{lnQ}	Mean VE_{C}
2-parameter SAR/GR4J	6 (2 + 4)	0.86	0.79	0.82	0.95
2-parameter SAR/GR3J	5 (2 + 3)	0.86	0.78	0.81	0.94

Line 610 and, earlier, line 490: how were these "physical or general values" obtained?

To make the purpose clearer, the initial sentence has been completed in Section 4.1 as follows:

"….The aim of using this mode was to account for elevation dependency in the T and P inputs from constant, calibrated orographic gradients while fixing the parameters that control snow accumulation and melt to physical or general values: precipitation phase determined based on a linear separation between -1 °C and +3 °C (see USACE, 1956), temperature threshold for snowmelt fixed at 0 °C, degree-day melt factor set at 5 mm. $°C^{-1}.d^{-1}$ (mean general value taken from Hock, 2003)."

The additional references have been also inserted in the reference list.

New references

Frey, S. and Holzmann H.: A conceptual, distributed snow redistribution model, Hydrol. Earth Syst. Sci., 19, 4517–4530, https://doi.org/10.5194/hess-19-4517-2015, 2015.

Gafurov, A. and Bárdossy, A.: Cloud removal methodology from MODIS snow cover product, Hydrol. Earth Syst. Sci., 13, 1361–1373, https://doi.org/10.5194/hess-13-1361-2009, 2009.

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation. J. Glaciology 45, 101–111, https://doi.org/10.3189/S0022143000003087, 1999.

Hock, R.: Temperature index melt modelling in mountain areas, J. Hydrol., 282, 104–115, https://doi.org/10.1016/S0022-1694(03)00257-9, 2003.

NSIDC: World Glacier Inventory, Version 1. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5/NSIDC-WGI-2012-02, 2012.

USACE: Snow Hydrology: Summary Report of the Snow Investigation. Portland, Oregon, North Pacific Division, Corps of Engineers, U.S. Army, 1956.

Responses to comments from anonymous Referee 2

On “Should altitudinal gradients of temperature and precipitation inputs be inferred from key parameters in snow-hydrological models?” by D. Ruelland (HESS-2019-556)

Referee’s comment

The article analyzes the sensitivity of a snow accounting procedure and hydrological modeling results to the evaluation of temperature and precipitation in space and time in mountainous catchments. The study is based on a set of 20 catchments in the French Alps and two hydrological models. The author evaluates the interplay between the lapse rate, snow routine and hydrological model parameters.

I found this is a clear and interesting paper. I have a few suggestions for improvement detailed below, some of which are quite major and requiring new calculations. I suggest considering the paper for possible publication in HESS after major revision.

Authors’ response

I would like to thank the referee for the time spent in reviewing the initial paper and making interesting suggestions. Most of them were judged useful, I provided a point-by-point response to the reviewer’s comments and tried to bring modifications to the manuscript accordingly.

Detailed comments

Referee’s comment

1. I found that the literature review could have been more exhaustive, to better stress the originality of the work compared to existing studies on similar or close topics. Some recent works could be discussed, for example the work by Le Moine et al. (2015) on the link between snow and hydrological sub-models in model parameterization, some studies on using snow data to calibrate hydrological models (Besic et al. 2014, Henn et al. 2016, Riboust et al. 2019), some studies with physical approaches to estimate lapse rates (Rahman et al. 2014, Zhang et al. 2015, Naseer et al. 2019). The review could also be extended on how gauge undercatch factors are estimated. The author should further discuss to which extent the proposed approach is original compared to these past findings.

Authors’ response and modifications to manuscript

I thank the referee for these additional references, some of which I did not know. They were cited in the text and added to the reference list.

Following the referee comment, the following paragraph was added in the introduction section:

“…Several studies proposed approaches to estimate lapse rates based on physically-based or conceptual models on specific catchments. Zhang et al. (2013) showed that the runoff simulation results involving snowmelt and rainfall runoff were highly sensitive to the temperature and precipitation lapse rates in a Tibetan catchment. Rahman et al. (2014) calibrated the SWAT model in a snow-dominated basin in the Swiss Alps and found also that temperature lapse rate was significantly important for hydrological performance. Naseer et al. (2019) considered a dynamic lapse rate based on a vertical profile of temperature in a catchment in Japan and succeeded to improve the precipitation phase in a distributed hydrological modelling framework. Henn et al. (2016) investigated the value of snow data to constrain the inference of precipitation from streamflow, using lumped hydrologic models and an elevation-band snow model in a Californian basin. Their results suggested that multiple types of hydrologic observations, such as streamflow and SWE, may help to constrain the water balance of high-elevation basins. Le Moine et al. (2015) proposed a calibration strategy where the parameters of both an interpolation model and a daily snow-hydrological model are jointly inferred in a multi-variable approach applied in a catchment in the French Alps. Using a hydro-meteorological modelling chain involving 31 calibrated parameters, they

showed the potential of using different types of observations (rain gauges, snow water equivalent measurements and streamflow data) to help assess temperature and precipitation lapse rates according to different weather types. These examples encourage testing whether an inverse modelling approach based on calibrated constant lapse rates can perform well with parsimonious conceptual models applied in numerous basins."

The recent reference from Riboust et al. (2019) was mentioned later in the introduction:
"…Moreover, other authors (Franz and Karsten, 2013; He et al., 2014; Riboust et al., 2019) showed that adding snow data information to the calibration procedure enabled the identification of more robust snow parameter sets by making the snow models less dependent on the rainfall-runoff model with which they are coupled.
It was also acknowledged in section 4.1 about the snow accounting routine (se answer to the referee comment #7.

Some references were also acknowledged in the conclusion section:
"…Accurate estimate of these parameters greatly helps in determining the form of precipitation and spatial distribution of temperature and precipitation, and are critical for snow cover and runoff modelling in high mountain catchments, as already reported in other regions (Zhang et al., 2013; Naseer et al., 2019)."

Referee's comment

2. Section 2.1: It would be useful to add a figure showing the distributions of mean precipitation and temperature over the set of gauges, to give an idea of the variability across the study domain.

Authors' response and modifications to manuscript

Details about the "estimated" climatology of the study period are now provided in Table 1 (see below), which has been modified to include mean annual temperature (T), total precipitation (P), snowfall fraction (S) and streamflow (Q) for each basin. Note however that T, P and S values are very delicate to provide since they necessarily rely on approximations depending on the method used to distribute temperature and precipitation (in link with the paper issue). This is why they were not included in the initial submitted paper. As indicated in the modified caption of Table 1, catchment areal temperature, total precipitation and snowfall fraction were estimated after calibrating local altitudinal gradients over 2000–2016 using the snow-hydrological inverse approach proposed in the current paper (see Test #4 in Table 5).

Table 1 Streamflow gauging stations and main catchment characteristics. Percentages of glacierized area were estimated from the World Glacier Inventory (NSIDC, 2012). Mean annual precipitation (P), snowfall fraction (S) and temperature (T) were estimated after calibrating local altitudinal gradients over 2000–2016 using the snow-hydrological inverse approach proposed in the current paper (see Test #4 in Table 5).

Station	River	Area	Glacierized area	Elevations (m.a.s.l.)		Mean annual precip. (P)	Snowfall fraction (S)	Mean annual temp. (T)	Mean annual streamflow (Q)
		(km²)	(%)	Min	Max	(mm/yr)	(%)	(°C)	(mm/yr)
Barcelonnette	Ubaye	549	0	1132	3308	802	48	1.9	521
Lauzet-Ubaye	Ubaye	946	0	790	3308	947	44	3.0	654
Beynes	Asse	375	0	605	2273	920	16	8.7	344
Saint-André-Les-Alpes	Issole	137	0	931	2392	965	24	6.8	481
Villar-Lourbière	Séveraisse	133	4	1023	3623	1561	47	2.3	1317
Val-des-Prés	Durance	207	0	1360	3059	836	54	0.9	688
Briançon	Durance	548	1	1187	3572	844	51	1.7	714
Argentière-la-Bessée	Durance	984	3	950	4017	1014	52	2.1	765
Embrun	Durance	2170	2	787	4017	990	48	2.9	693
Espinasses	Durance	3580	1	652	4017	964	45	3.4	654

Villeneuve-d'Entraunes	Var	132	0	926	2862	989	37	4.8	650
Val-d'Isère	Isère	46	9	1831	3538	1245	63	-1.5	1119
Bessans	Avérole	45	12	1950	3670	1399	66	-2.4	1311
Taninges	Giffre	325	0	615	3044	2031	36	4.7	1771
Vacheresse	Dranse d'Abondance	175	0	720	2405	1669	29	4.9	1088
La Baume	Dranse de Morzine	170	0	690	2434	1636	32	4.7	1285
Dingy-Saint-Clair	Fier	223	0	514	2545	1649	26	6.5	1243
Allèves	Chéran	249	0	575	2157	1486	23	6.9	819
Mizoën	Romanche	220	9	1057	3846	1205	56	0.8	978
Allemond	L'Eau Dolle	172	2	713	3430	1460	46	2.7	1164

Referee's comment

3. Section 2.2: Reference could be given to the work by Leleu et al. (2014).

Authors' response and modifications to manuscript

I could not access this reference from the journal "*La Houille Blanche*", although I contacted the authors to obtain a hard copy. As a result, I could not judge if it was appropriate to reference this work here. However, Section 2.2 deals mainly with the streamflow series gathered from the French hydrological database (www.hydro.eaufrance.fr) for 20 catchments, as indicated in the text. In other publications, the French hydrological database is usually acknowledged this way by indicating the web site from which data (and metadata) can be freely accessed.

Referee's comment

4. Table 1: Please explain the meaning of abbreviations in the last column. Is this information useful here?

Authors' response and modifications to manuscript

The information was indeed not very useful. It has been removed from the revised Table 1 (see Table above).

Referee's comment

5. Section 3.3: The author calculates the efficiency criteria on precipitation values. However, the criteria may be strongly influenced by a few large rainfall events, which may not be representative of the average characteristics of precipitations. It may be useful to consider computing the efficiency criteria on transformed precipitation (e.g. root square transformation) to avoid putting too much weight on outlier values. Would this change something in results?

Authors' response and modifications to manuscript

The Jack-Knife cross-validation procedure on precipitation series is usually performed on RMSE (e.g. Kyriadis et al., 2001; Le Moine et al., 2013; Yang et al., 2018 to cite just a few). I could not find in the literature an example where RMSE was applied based on a root square transformation of precipitation in such a procedure. Of course, it can be argued than computing the objective function (here RMSE) on transformed precipitation may lead to different interpolation parameters. However, it can also be assumed that large rainfall events are critical for elevation/precipitation regressions when looking at the optimized surrounding gauges to consider in the KED and IED methods. This means that using an efficiency criterion on transformed precipitation may also put less weight on large rainfall events to compute the regressions.

Following the referee comment, the JK cross-validation was re-run with IED using $RMSE_{sqrt}$ on daily, monthly and yearly precipitation series. No significant differences could be found in the optimized interpolation parameters as shown in the following Table 1. Only the number of optimized number of surrounding neighbours changed slightly at the daily time scale from 17 to 15. This cannot

be not judged significant as this parameter presents a low sensitivity between 12 and 20 neighbours due to the intrinsic compromise looking on the whole study domain.

Table 1 Cross-validation of the IED method with RMSE or $RMSE_{sqrt}$ (root square transformation of precipitation) against yearly, monthly and daily series from precipitation gauges over the period 2000–2016. The values of $n(u)$ and ω represent the interpolation parameters, which were optimised using the leave-one-out procedure.

	Current paper (RMSE as EC)			Alternative test ($RMSE_{sqrt}$ as EC)		
	Efficiency criterion	IED parameters		Efficiency criterion	IED parameters	
	RMSE	$n(u)$	ω	$RMSE_{sqrt}$ (RMSE)	$n(u)$	ω
Yearly	150.31 mm/year	12	3	2.13 (150.31)	12	3
Monthly	22.20 mm/month	12	2	1.05 (22.2)	12	2
Daily	2.90 mm/day	17	2	0.51 (2.91)	15	2

For these different reasons, and also because the $RMSE_{sqrt}$ is difficult to interpret since units are not allowed, the usual RMSE criterion was kept in the article as efficiency criteria on precipitation series. Note also that the two other criterions (MAE and NSE which were not used for optimization) are not presented anymore in Table 4 as they seemed to cause confusion in the result interpretation (see answer to the Referee's comment #10 below).

Referee's comment

6. L261: The name "RMSE" given to the normalized RMSE is a bit confusing. The author may choose another name, e.g. NRMSE.

Authors' response and modifications to manuscript

There was an error in the text, which is now corrected. In fact the models were cross-validated against the usual RMSE (root mean square error) without any normalization, as follows:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N} (Vpre_i - Vobs_i)^2 / N} \quad (1)$$

where $Vpre_i$ and $Vobs_i$ are the predicted and observed variables respectively at time scale *i* and *N* the total number of time steps.

Referee's comment

7. Section 4.1: Some modifications in this snow module were recently proposed by Riboust et al. (2019), to account for snow-covered area. This should be shortly commented, to better explain how the proposed approach compares to this existing work.

Authors' response and modifications to manuscript

Agreed. This is now commented in the Section 4.1, as follows:

"In the original version of CEMANEIGE, fractional snow-covered area (FSC) is calculated as follows:

$$FSC_i(t) = \min\left(\frac{SWE_i(t)}{SWE_{th}}, 1\right) \quad (11)$$

where *SWE* is the quantity of snow accumulated on the catchment in snow water equivalent (a state variable of the model, in mm), and SWE_{th} is the model's melting threshold. SWE_{th} is calculated as being equal to 90% of mean annual solid precipitation on the catchment considered (Valéry et al., 2014). Alternative approaches have been proposed to account for the hysteresis that exists between *FSC* and *SWE* during the accumulation and melt phases (Riboust et al., 2019). However, introducing such a hysteresis adds two additional free parameters to the SAR. Instead, SWE_{th} was fixed to 40 mm

since preliminary sensitivity analyses showed that this value gave very satisfactory *FSC* values when compared to the MODIS observations in the studied catchments."

Referee's comment

8. Fig.3: Maybe add the meaning of the key variables (at least inputs/output) in the figure caption. If UZL is the threshold for the upper output, maybe the arrow should stop at the level of this output.

Authors' response and modifications to manuscript

Agreed. The meaning of the key variables has been added in the Figure caption, as follows: "*R*, *M*, *PE* and *Q* stand for rainfall, melt, potential evapotranspiration and streamflow, respectively". Following the referee comment, the arrow for *UZL* was also slightly modified in the figure.

Referee's comment

9. L376-378: This is a point I did not understand in the proposed methodology. By introducing this criterion WB in the objective function, the author forces the model to close the water balance in the sense of Budyko. This is quite successful when looking at results shown in Fig. 6, since no data lies outside the boundaries of balance closure in the plot. However, I do not understand the physical rationale behind putting this constraint. There are many catchments where the water balance cannot be closed in the Budyko sense for good reasons, mainly because of underground water exchanges. The author artificially constrains the models using WB. I think a more classical bias criterion would be better to consider instead.

Authors' response and modifications to manuscript

The rationale behind using WB in the objective function was to enhance the parameter identifiability without decreasing the model efficiency. I agree that there are many catchments where the water balance cannot be closed in the Budyko sense for good reasons due notably to inter-catchment groundwater exchanges (IGE). There are also several bad reasons for which water balance is sometimes not closed at the basin scale: errors in the precipitation volumes, wrong estimate of potential evapotranspiration, inaccurate knowledge of the catchment area, etc. Since the paper deals with the lapse rates of the temperature and precipitation inputs, it can be assumed that using a more classical objective function (i.e. without WB) may also lead to optimize the lapse rates while water balance is not closed for the above "bad" reasons (errors in the precipitation volumes, wrong estimate of potential evapotranspiration).

Nevertheless, I decided to follow the referee comment because sensitivity analyses to the objective function are far beyond the paper issue and because other readers may not be convinced by the proposed constraint. I re-run all the simulations with a more classical objective function (i.e. without using the WB constraint) based only on NSE_{FSC} and NSE_{sqrtQ}. The results were the same as regards to the modelling distribution performances between the various tests (see Figures 1 and 2 below). Changing the objective function thus did not change the main findings of the paper notably as regards to the interest of calibrating the temperature and precipitation lapse rates via a parsimonious 2-parameter SAR. Obviously, the water balance was not closed systematically and it was not interesting anymore to present the Budyko graphs for the different tests. This led also to deteriorate the general parameter identifiability (see Figure 3 below). This particularly affected the identifiability of the *X2* parameter of GR4J. However, the ranking of parameter identifiability in between tests did not change and the parsimonious 2-parameter SAR still led to the best parameter identifiability, while remaining among the best-performing models. Finally, the optimized lapse rates were slightly changed: the precipitation gradients were notably found more similar between the two hydrological models tested (see Figure 4 below). This is another reason that convinced me to renounce to the WB constraint in the objective function.

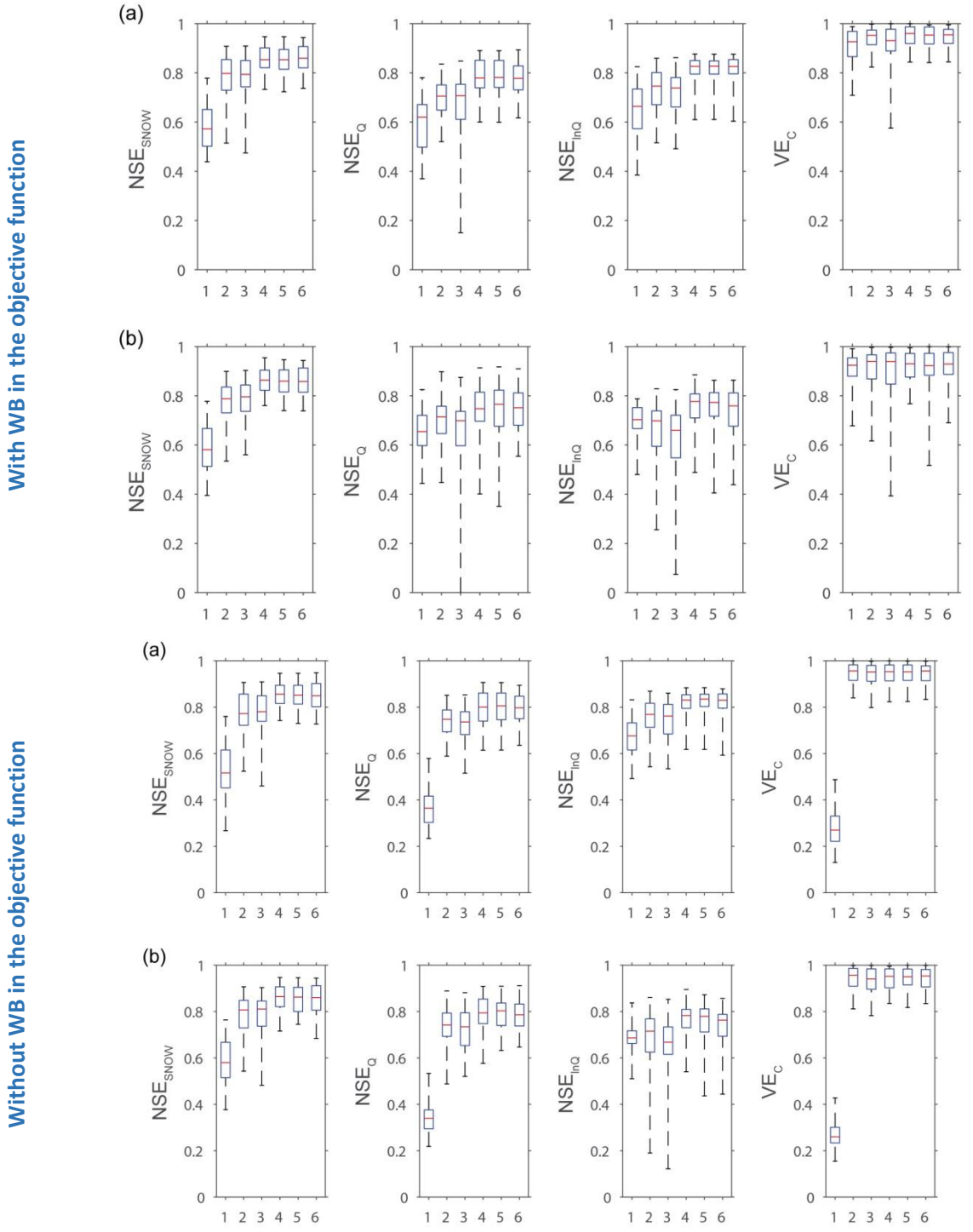
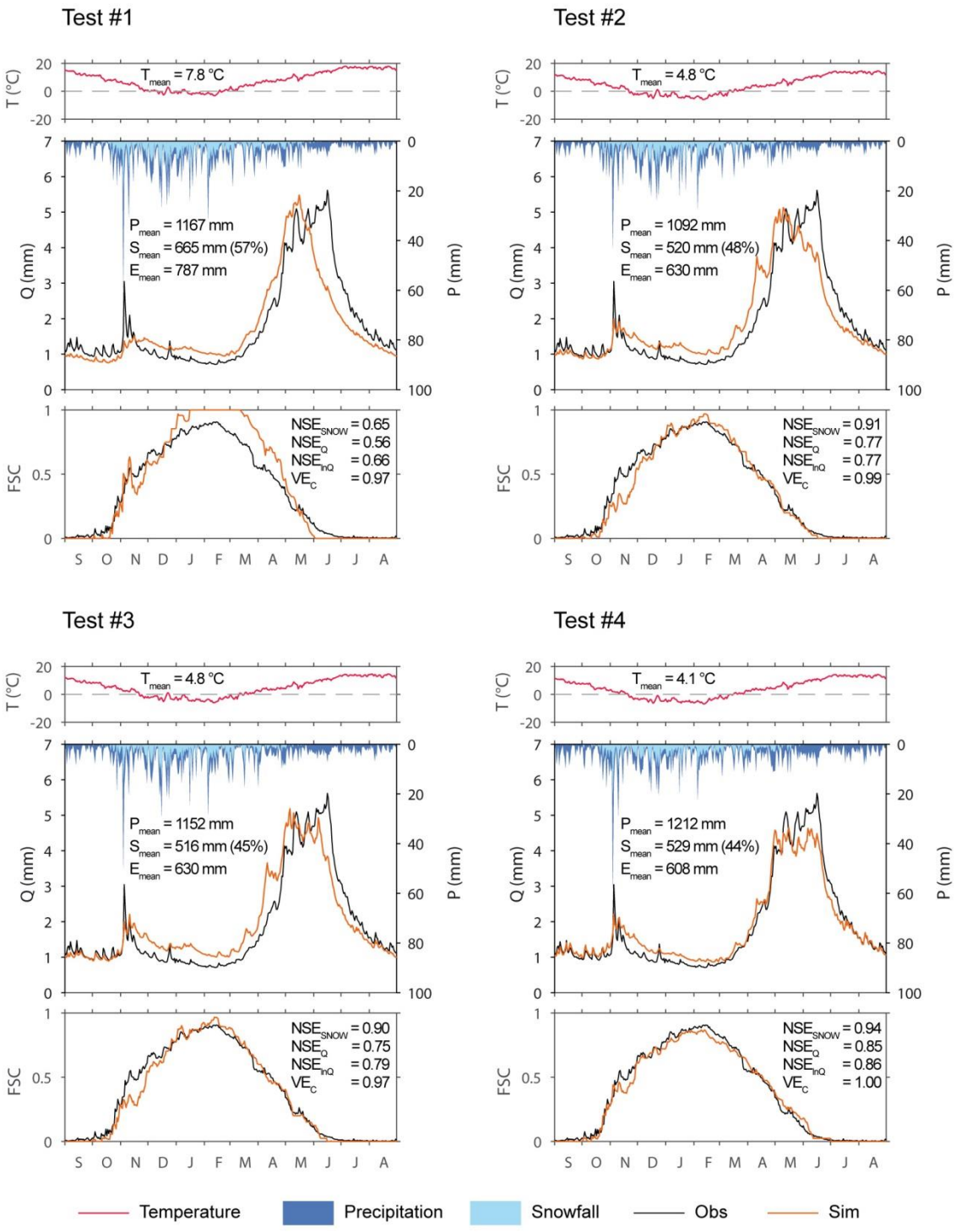


Fig. 1 Boxplots (showing 0.00, 0.25, 0.50, 0.75 and 1.00 percentiles) of the efficiency distributions obtained in validation by the (a) GR4J and (b) HBV9 models combined with the snow model according to six different tests (see Table 5) to account for elevation dependency in the *T* and *P* inputs on the 20 snow-affected Alpine catchments.

With WB in the objective function



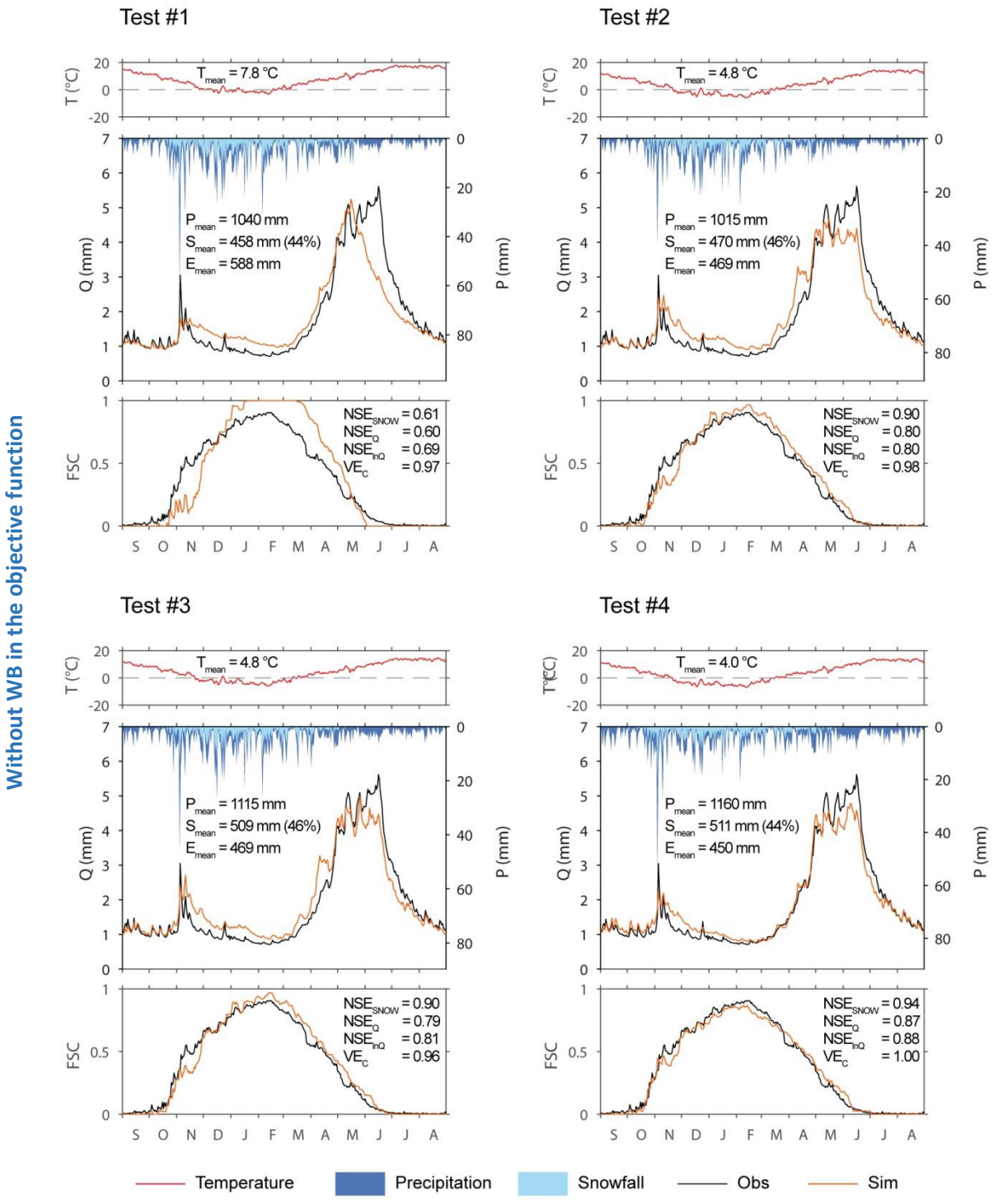
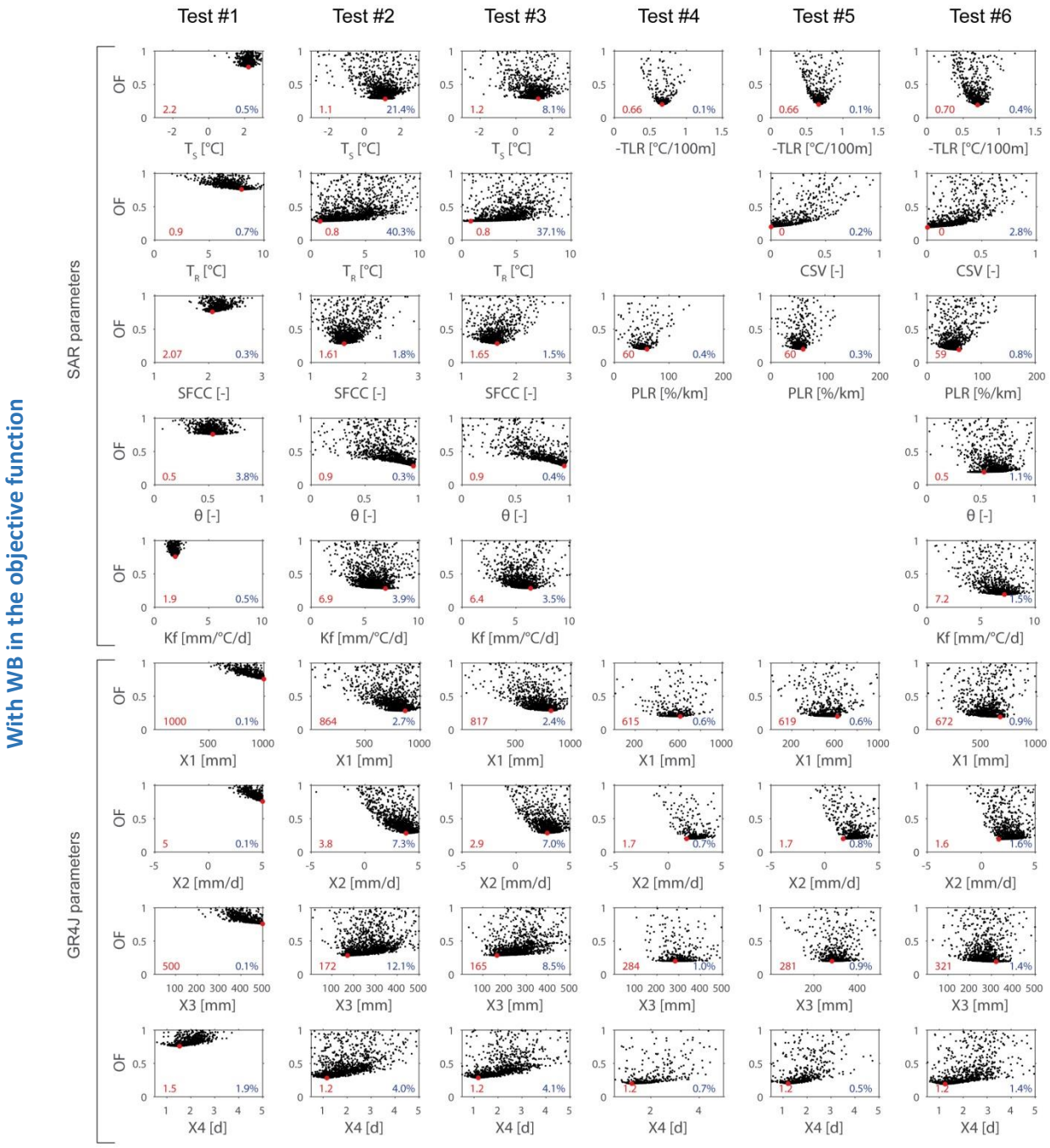


Fig. 2 Comparison of snow-hydrological simulations with elevation dependency according to Tests #1 to #4 (see Table 5) with GR4J for the Durance at Serre-Ponçon. The graphs show mean inter-annual time-series of temperature, precipitation, streamflow and fractional snow cover at the catchment scale in validation over the period 2008–2016. T_{mean}, P_{mean} and S_{mean} stand for mean annual temperature, precipitation, and snowfall, respectively. The efficiency criterions NSE_{SNOW}, NSE_{Q}, NSE_{lnQ} and VE_{C} are computed from continuous (not mean seasonal) series over 2008–2016.


With WB in the objective function
Test #1
Test #2
Test #3
Test #4
Test #5
Test #6
SAR parameters
GR4J parameters
OF
T_S [°C]
T_R [°C]
SFCC [-]
θ [-]
Kf [mm/°C/d]
-TLR [°C/100m]
CSV [-]
PLR [%/km]
X1 [mm]
X2 [mm/d]
X3 [mm]
X4 [d]

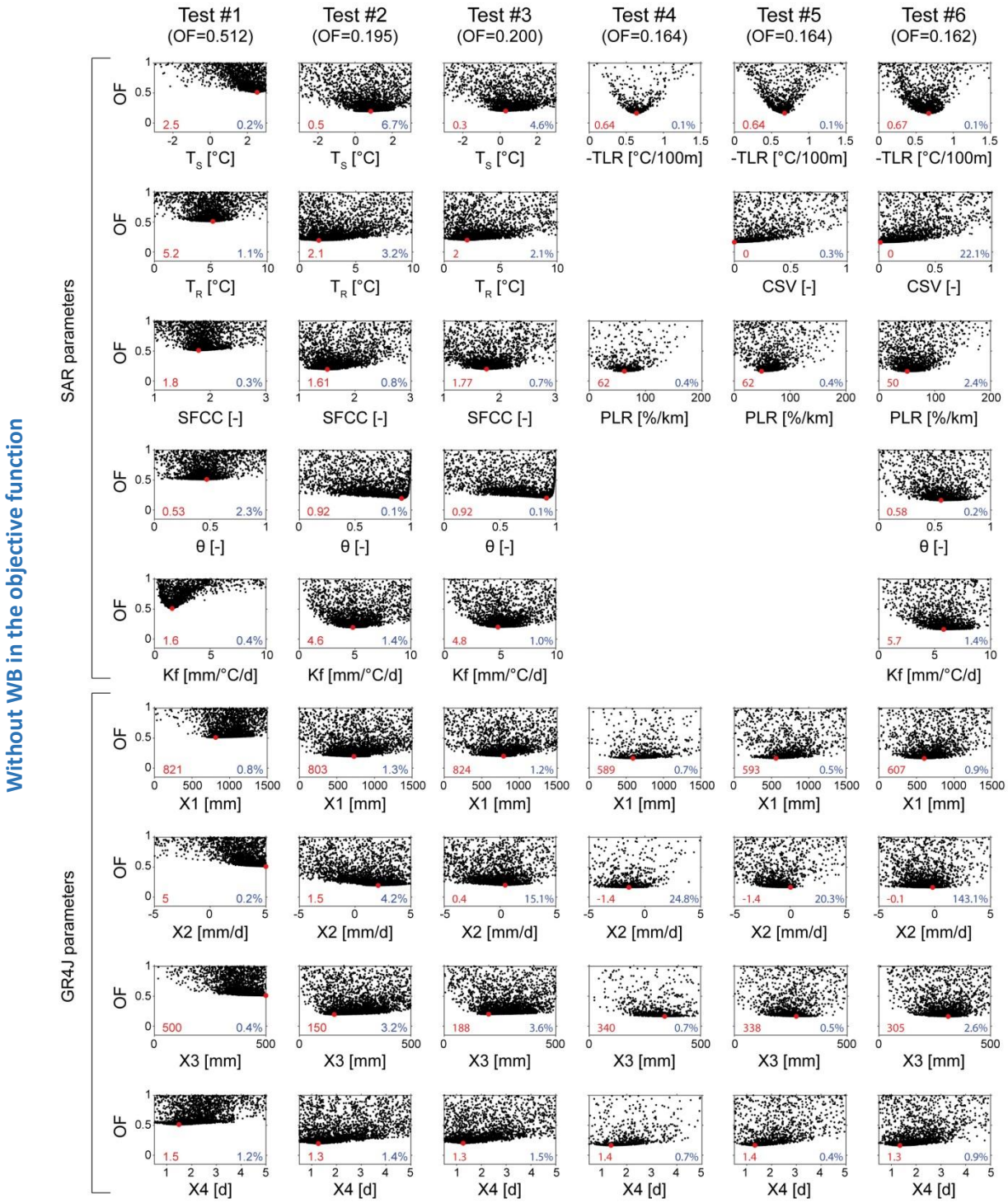


Fig. 3 Parameter sensitivity to the objective function (OF) according to Tests #1 to #6 (see Table 5) with GR4J combined with the snow accounting routine (SAR) on the Durance at Serre-Ponçon. The values and dots in red indicate the optimised calibrated parameters when minimising OF, the black dots represent trials of the SCE-UA optimisation algorithm, and the values in blue are the variation coefficients (in %) of the 20% best parameter solutions compared to the optimised values for each parameter (the lowest value, the easiest parameter identifiability). Note that depending on the tests, the calibrated parameters of the SAR vary from 2 to 5 (see Table 5 and Table 2 in the manuscript), while the GR4J hydrological models has 4 parameters (see Table 3 in the manuscript).

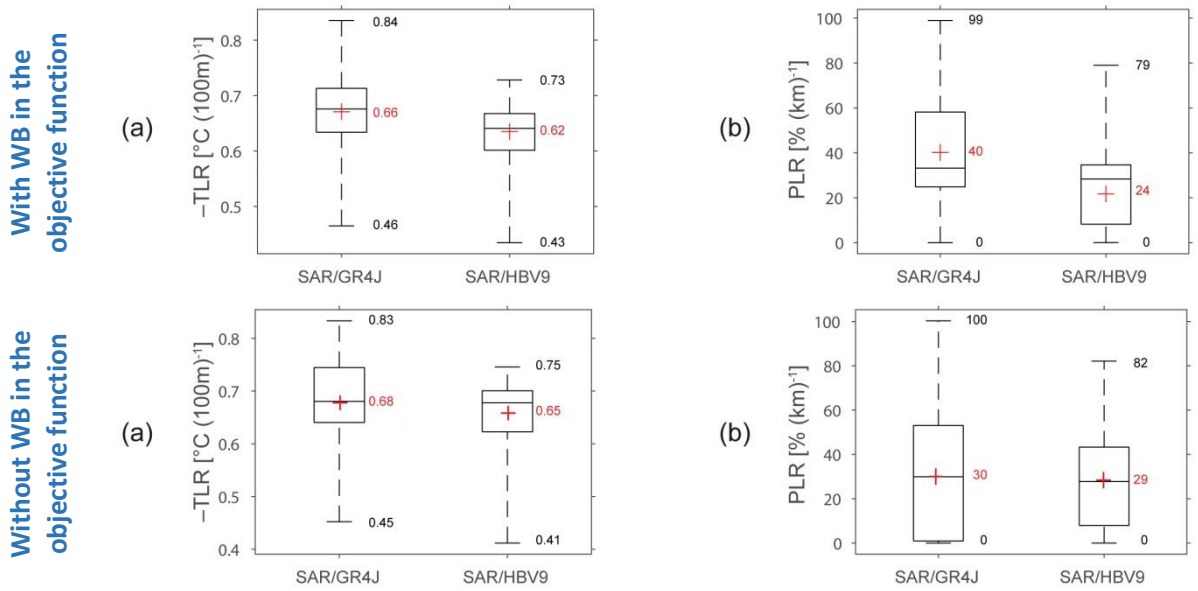


Fig. 4 Boxplots (showing 0.00, 0.25, 0.50, 0.75 and 1.00 percentiles) of the ranges of (a) temperature and (b) precipitation lapse rates calibrated with the 2-parameter SAR (Test #4) in association with the GR4J and HBV9 models on the 20 snow-affected Alpine catchments. The red crosses indicate mean values.

Referee's comment

10. Table 4: There is a strong drop in the NSE criterion for temperature when going from monthly to daily time steps for IDW and ORK. How this drop can be explained?

Authors' response and modifications to manuscript

The drop in the NSE criterion for temperature was in fact when going from monthly (or daily) to yearly time scale for IDW and ORK. As it can be seen from the following figure 5, the NSE criterion is very sensitive to the number of considered time steps, and further on the range of sampled temperatures (which is quite different at the yearly versus monthly time step). As a result, the NSE values between the different methods should be compared only for a given time scale, and not in between time scales.

On the opposite, the RMSE criterion (which was used as objective function in the JK cross-validation) is better representative for the comparison of temperature (whatever the time scales) since units are directly comparable. Following the referee comment, and as NSE values seemed to cause confusion in the result interpretation, only the RMSE values are now presented in Table 4.

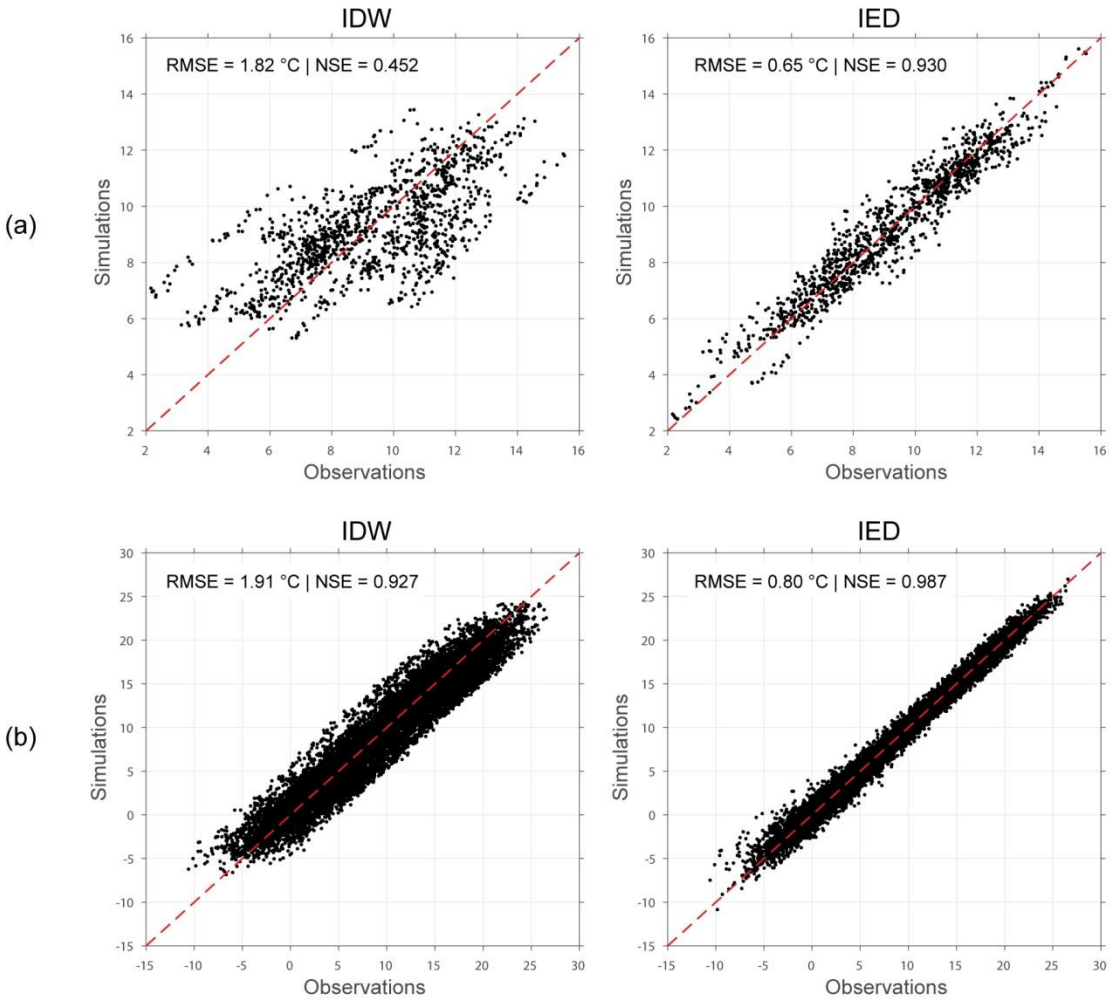


Fig. 5 Cross-validation of the IDW and IED methods against (a) yearly and (b) monthly series from temperature gauges over the period 2000–2016. The NSE criterion is very sensitive to the number of considered time step, and further on the range of sampled temperatures (which is quite different at the yearly versus monthly time step). As a result, there is drop in NSE values for temperature when going from monthly to yearly time step, particularly with the IDW method. Therefore, the NSE values between the different methods should be compared only for a given time step, and not in between time steps. On the opposite, the RMSE criterion (used as objective function for cross-validation) is better representative for the comparison of temperature whatever the time step.

Referee's comment

11. L472-476: I think this result is the consequence of using WB in the objective function. As mentioned above, this constraint is artificial and potentially counterproductive for the efficiency of the model.

Authors' response and modifications to manuscript

This result was indeed the consequence of using WB in the objective function. Please note however that this constraint was not counterproductive for the efficiency of the model as it can be seen clearly from Figures 1 and 2 of the revision notes: with or without WB in the objective function, the hydrological predictions are significantly improved as regards to the efficiency criterions when using a SAR targeting for the temperature and precipitation lapse rates (Tests #4, #5 and #6).

Following the referee comment, a more classical objective function was used (i.e. without the WB term in the OF). Obviously, the water balance was not closed systematically and it was not interesting anymore to present the Budyko graphs for the different tests. The Budyko graphs and associated comments were therefore removed from the manuscript.

Referee's comment

12. L510-516: I find this a bit contradictory with the WB constraint. If the author makes the hypothesis that underground water exchanges between catchments may play a key role, why does the author constrain water balance not to account for such exchanges in the optimization phase?

Authors' response and modifications to manuscript

As explained above, inter-catchment groundwater exchanges (IGE) are not the only reason why the water balance may not be closed in the Budyko sense. Other reasons (maybe more important) may play a key role such as errors in the precipitation volumes or wrong estimate of potential evapotranspiration. Since the paper deals with the optimization of temperature (impacting snow accumulation and melt, but also evapotranspiration estimates) and precipitation gradients, constraining the water balance in the objective function aimed mainly at enhancing the parameter identifiability (see Fig. 3 of the revision note) without deteriorating the modelling efficiency (see Figs. 1 & 2 of the revision note). While the HBV model considers the catchments as closed systems, GR4J allows potential IGE via its X2 parameter.

However, since sensitivity analyses to the objective function are far beyond the paper issue and because other readers may not be convinced by the proposed constraint, I renounced to the WB constraint in the objective function (see answers to the referee comment #9) and I re-run all the simulations with a more classical *OF*. Figures and comments were changed accordingly. Please note that it did not change the main findings of the paper.

The following paragraph (and associated new table) was also added in the section 5.3 (Identifiability of the parameters) to further discuss on the IGE issue and suggest the findings of the initial submission using the WB term in the OF:

"…Equifinality is also reduced in Tests #4–6 for the parameters controlling runoff generation and routing (*X1*, *X3* and *X4*). On the opposite, the parameter of the inter-catchment groundwater flows (*X2*) is poorly identifiable with variation coefficients of 24.8%, 20.3% and 143.1% with Test #4, Test #5 and Test #6, respectively. This suggests that inter-catchment groundwater exchanges (IGE) do not play a key role in the studied catchments. Indeed, fixing *X2* to a value of 0 (i.e. without potential IGE) with an alternative GR3J model provided similar mean validation efficiency on the set of catchments as compared to the GR4J associated with the 2-parameter SAR (Table 7). However, other objective functions may result in other findings as far as IGE are concerned. For instance, additional tests (not shown here for brevity sake) confirmed that it was possible to greatly reduce the X2 equifinality without decreasing the model efficiency by adding a water balance term in the objective function to constrain the proportion of years respecting the water and energy balance in the Turc-Budyko non-dimensional graph (see Andréassian and Perrin, 2012). These tests suggested that it may be relevant to explicitly represent inter-catchment groundwater transfers in association with correcting or scaling factors applied to the precipitation input data to render the distribution between evapotranspiration, streamflow and underground fluxes more realistic, as already reported by Le Moine et al. (2007)."

Table Mean validation efficiency on the set of 20 catchments with the GR4J model and the GR3J model in association with the 2-parameter SAR.

Model	Total number of free parameters	Mean NSE_{SNOW}	Mean NSE_{Q}	Mean NSE_{lnQ}	Mean VE_{C}
2-parameter SAR/GR4J	6 (2 + 4)	0.86	0.79	0.82	0.95
2-parameter SAR/GR3J	5 (2 + 3)	0.86	0.78	0.81	0.94

Referee's comment

13. Fig. 8 is interesting. However there are some cases which reveal that the optimum is probably outside the preset parameter range. This is typically the case for Test#1 for parameters X1 to X3. Therefore the ranges should be extended.

Authors' response and modifications to manuscript

For Test#1 (and only for Test#1), the parameters *X1* to *X3* indeed reached the maximum allowed range. Please note however that Test#1 serves as a benchmark. As explained in Table 5 and in the text, it differs from the other tests because no elevation dependency in the *T* and *P* inputs are considered. As a result, hydrologic predictions with Test#1 are significantly (and rather logically) outperformed by the other approaches accounting for elevation-dependency (see e.g. Figures 1 and 2 in the revision note). Extending the range of the parameters would be both poorly efficient in improving the simulations and incorrect from a numerical point of view. The referee has to be aware that the parameter ranges were preset to values recommended by the models' authors (Perrin et al., 2003 for GR4J and Beck et al., 2016 HBV9). They have been found after numerous simulations in very different contexts and can be judged as large enough. By the way, it can be seen in Figure 3 of the revision note, that no parameter limits are reached in the other tests, thus suggesting that the preset parameter range are adequate.

To address the referee comment, I only extended the range of the *X1* parameter (from 10-1000 mm to 0-1500 mm) of GR4J to ensure a better correspondence with the UZL parameter range of HBV9. All simulations were re-run with this new range (and also with an objective function without WB, see answer to the referee's comment #9), and Figures and comments were modified accordingly. Please note that this did not change the results (see Figure 1 of the revision note) and the parameters X2 and X3 still reached the maximum allowed range (see Figure 3 in the revision note) with Test#1 (and only with Test #1) for the reasons explained above. The following comment was also added in the beginning of section 5.3:

"…The maximum allowed parameter range is only reached for the parameters *X1* and *X2* with Test #1. This test differs from the others because no elevation dependency in the *T* and *P* inputs are considered. Consequently, hydrologic predictions of Test #1 are significantly outperformed by the other approaches. Extending the parameter ranges beyond the tested values would be both poorly efficient in improving the simulations and incorrect from a numerical point of view since they were set to values recommended by the models' authors. Moreover, no parameter limits were reached in the other tests, thus suggesting that the parameter ranges are adequate…"

Referee's comment

Cited references:

Besic, N., et al. (2014). "Calibration of a distributed SWE model using MODIS snow cover maps and in situ measurements." Remote Sensing Letters 5(3): 230-239.

Henn, B., et al. (2016). "Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation." Water Resour. Res. 52(11): 8700-8723.

Le Moine, N., et al. (2015). "Hydrologically Aided Interpolation of Daily Precipitation and Temperature Fields in a Mesoscale Alpine Catchment." J. Hydrometeorol. 16(6):2595-2618.

Leleu, I., et al. (2014). "Re-founding the national information system designed to manage and give access to hydrometric data." La Houille Blanche(1): 25-32.

Naseer, A., et al. (2019). "Distributed Hydrological Modeling Framework for Quantitative and Spatial Bias Correction for Rainfall, Snowfall, and Mixed-Phase Precipitation Using Vertical Profile of Temperature." J. Geophys. Res.-Atmos. 124(9): 4985-5009.

Rahman, K., et al. (2014). "Streamflow response to regional climate model output in the mountainous watershed: a case study from the Swiss Alps." Environmental Earth Sciences 72(11): 4357-4369.

Riboust, P., et al. (2019). "Revisiting a simple degree-day model for integrating satellite data: implementation of SWE-SCA hystereses." Journal of Hydrology and Hydromechanics 67(1): 70-81.

Zhang, F., et al. (2013). "Snow cover and runoff modelling in a high mountain catchment with scarce data: effects of temperature and precipitation parameters." Hydrol. Processes 29(1): 52-65.

Authors' response and modifications to manuscript

The proposed references were judged useful. Therefore, they were cited in the text and added to the reference list, except that of Leleu et al. (2014) which I could not find.

Should altitudinal gradients of temperature and precipitation inputs be inferred from key parameters in snow-hydrological models?

Denis Ruelland

CNRS, HydroSciences Montpellier, University of Montpellier, Place E. Bataillon, 34395 Montpellier Cedex 5, France

Correspondence to: *denis.ruelland@umontpellier.fr*

Abstract. This paper evaluates whether snow-covered area and streamflow measurements can help assess altitudinal gradients of temperature and precipitation in data-scarce mountainous areas more ~~realistically~~ efficiently than using the usual interpolation procedures. A dataset covering 20 Alpine catchments is used to investigate this issue. Elevation dependency in the meteorological fields is accounted for using two approaches: (i) by estimating the local and time-varying altitudinal gradients from the available gauge network based on deterministic and geostatistical interpolation methods with an external drift; and (ii) by calibrating the local gradients using an inverse snow-hydrological modelling framework. For the second approach, a simple 2-parameter model is proposed to target the temperature/precipitation-elevation relationship and to regionalise air temperature and precipitation from the sparse meteorological network. The coherence of the two approaches is evaluated by benchmarking several hydrological variables (snow-covered area, streamflow) computed with snow-hydrological models fed with the interpolated datasets and checked against available measurements. Results show that accounting for elevation dependency from scattered observations when interpolating air temperature and precipitation cannot provide sufficiently accurate inputs for models. The lack of high-elevation stations seriously limits correct estimation of lapse rates of temperature and precipitation, which, in turn, affects the performance of the snow-hydrological simulations due to imprecise estimates of temperature and precipitation volumes. Instead, retrieving the local altitudinal gradients using an inverse approach enables increased accuracy in the simulation of snow cover and discharge dynamics, while limiting problems of over-calibration and equifinality.

1. INTRODUCTION

1.1. Providing accurate meteorological inputs in mountainous regions

Regionalising air temperature and precipitation is a critical step in producing accurate areal inputs for hydrological models in high altitude catchments. The ability to correctly reproduce areal precipitation is vital to avoid the failure of hydrological models, which are sensitive to input volumes at the catchment scale (e.g. Oudin et al., 2006; Nicótina et al., 2008). Accurate temperature fields are also particularly important in mountainous regions because temperature is the main driver for snow/rain partition and snowmelt, and consequently influences seasonal discharge (e.g. Hublart et al., 2015; 2016).

However, in areas with complex topography, the characteristic spatial scales of temperature and precipitation estimates are typically poorly captured, notably when the network of measurements used is sparsed. Gridded datasets obtained by interpolating measurements taken at meteorological stations are thus affected by inaccuracies, which are spatially and temporally variable and difficult to quantify (Haylock et al., 2008; Isotta et al., 2015). Measurement errors depend on local conditions and increase with terrain elevation, as the operational conditions become more extreme (Frei and Schär, 1998). In the case of precipitation, a well-known problem arises from the systematic errors associated with precipitation under-catch during snowfall (Strasser et al., 2008), especially in windy conditions (Sevruk, 2005). In addition, temperature and precipitation are under-sampled at high elevations, because meteorological stations are mainly located at low elevations for logistical reasons (Hofstra et al., 2010). This makes it difficult to derive the local and seasonal relationship between

meteorological observations and topography, even though this is indispensable for accurate spatial temperature and precipitation estimates (Masson and Frei, 2014). Indeed, atmospheric uplift caused by relief tends to increase precipitation with elevation through the so-called orographic effect (Barry and Chorley, 1987). Nevertheless, precipitation accumulation trends can show considerable scatter with altitude depending on the region's exposure to wind and synoptic situations (Sevruk, 1997). The relationship between temperature and elevation is generally more obvious. The rate at which air cools with a change in elevation ranges from about -0.98 °C (~~hm~~100m)$^{-1}$ for dry air (i.e., the dry-air adiabatic lapse rate) to about -0.40 °C (~~hm~~100m)$^{-1}$ (i.e., the saturated adiabatic lapse rate; Dodson and Marks, 1997). Average temperature gradients of -0.60 °C (Dodson and Marks, 1997) or -0.65 °C (~~hm~~100m)$^{-1}$ (Barry and Chorley, 1987) are often used when high precision is not required. However, such average values are known to be rough approximations which are not suitable for more precise studies (see e.g. Douguédroit and De Saintignon, 1984). Notably they mask significant variations in different meteorological conditions and in different seasons. For instance, temperature lapse rates are generally lower in winter than in summer, as shown by Rolland (2003) for Alpine regions.

1.2. Schemes for spatial interpolation of air temperature or precipitation

The mapping of air temperature and precipitation using discrete observations based on gauge networks has been extensively studied. Readers can refer to, for instance, Ly et al. (2013) for a review on the different deterministic and geostatistical methods designed for operational hydrology and hydrological modelling at the catchment scale.

Schemes for spatial interpolation of meteorological variables vary in three ways (Stahl et al., 2006): (1) the model used to characterise the spatial variation of the variable of interest, (2) the method used to choose the surrounding points (number or distance, angular position relative to the prediction point) and (3) the approach used to adjust for elevation. The simplest approach is to choose the nearest station and adjust for elevation according to an assumed lapse rate. However, this method is fairly crude and ignores fine-scale spatial variations. Where more than one station is used in the prediction, a model is required to determine how to interpolate from them. Interpolation weights have been estimated using approaches including inverse distance weighting (IDW) (e.g. Dodson and Marks, 1997; Shen et al., 2001; Frei, 2014) and geostatistical methods based on kriging (e.g. Garen and Marks, 2005; Spadavecchia and Williams, 2009). Kriging relies on statistical models involving autocorrelation, which refers to the statistical relationships between measured points. Ordinary kriging (OKR) is well-known among kriging algorithms (see e.g. Goovaerts, 1997 for a detailed presentation of these algorithms). Different methods have been developed to deal with the statistical relationship between temperature/precipitation and elevation like regression analysis (Drogue et al., 2002), or more elaborate geostatistical techniques including simple kriging with local means, kriging with external drift (KED) and co-kriging (CKR): see Goovaerts (2000) for a comparison of these approaches. Among these techniques, KED has been widely used to generate temperature and precipitation maps. For instance, Masson and Frei (2014) showed that KED led to much smaller interpolation errors than linear regressions in the Alps. This was achieved with a single predictor (local topographic height), whereas the incorporation of more extended predictor sets (slope, circulation-type dependence of the relationship, inclusion of a wind-aligned predictor) enabled only marginal improvement. For daily precipitation, interpolation accuracy improved considerably with KED and the use of a simple digital elevation compared to OKR (i.e., with no predictor). These results confirm that accounting for topography is important for spatial interpolation of daily precipitation in high-mountain regions. Conversely, other authors showed that, even though taking topography into account was indispensable for temperature reconstruction whatever the temporal resolution, it was less clear for daily precipitation. For example, Ly et al. (2011) reported no improvement in precipitation estimated at a daily time scale if topographical information was taken into account with KED and CKR, compared to simpler methods such as ORK and IDW. In a recent and very complete comparative study, Berndt and Haberlandt (2018) analysed the influence of temporal resolution and network density on the spatial interpolation of climate variables. They showed that KED using elevation

performed significantly better than ORK for temperature data at all temporal resolutions and station densities. For precipitation, using elevation as additional information in KED improved the interpolation performance at the annual time scale, but not at the daily time scale.

Theoretically, KED can account for local differences in topographic influence in different seasons and synoptic situations. Indeed, the regression coefficients computed between the primary variable (temperature or precipitation) and the secondary variable (elevation) are implicitly estimated through the kriging system within each search neighbourhood (Goovaerts, 2000). The relation between variables is thus assessed locally, meaning changes in correlation across the study area can be taken into account. However, as suggested by Stahl et al. (2006) concerning temperature and by Ly et al. (2011) concerning precipitation, care should be taken in applying KED when interpolating daily variables with very few neighbouring sample points. Indeed, methods that compute local lapse rates from the surrounding control points can perform poorly in regions with insufficient high-elevation data, due to inaccurate estimation of local lapse rates.

1.3. Placing meteorological fields in a hydrological perspective

A subject that requires further investigation is which methods that produce daily temperature and precipitation fields can provide the best snow cover and streamflow simulations. The usual cross validation for the inter-comparison of interpolation methods is limited, especially in ungauged areas, like the highest parts of mountainous areas. As stressed by Gottardi et al. (2012), a method can perform well in interpolation (at the ground network altitudes) but poorly in extrapolation (higher). This is because the observed set is not representative of the entire feature space. As a result, estimations at high elevations are difficult to check due to the lack of meteorological data. To go further, the use of other data like streamflow measurements may be a good alternative way to validate temperature and precipitation estimations at high-elevation sites.

To date, few studies have compared the performance of different interpolation methods evaluated by hydrological modelling in mountainous areas. Among the few that have, Tobin et al. (2011) showed that kriging (and more specifically KED) can be used effectively to estimate temperature and precipitation fields in complex alpine topography during flood events. Their comparative analyses of the different interpolation techniques suggested that geostatistical methods performed better than IDW. In particular, with elevation as auxiliary information, KED gave the overall best validation statistics for the set of events under study. However, it can be hypothesised that, in many mountainous areas, gauge observations do not include sufficient information to accurately account for the elevation dependency of air temperature and precipitation using interpolation techniques, which are thus limited to providing accurate inputs for snow-hydrological models. On the other hand, numerous calibration parameters controlling snow accumulation (the temperature threshold between the solid and liquid phase, temperature range of phase separation, snowfall gauge under-catch factor) and melt (temperature threshold for snowmelt, degree-day melt factor, snowpack thermal state, etc.) have been introduced in most of the snow accounting routines (SAR) used in operational hydrology: see e.g. HBV (Bergström, 1975), MOHYSE (Fortin and Turcotte, 2007), CEMANEIGE (Valéry et al., 2014), MORDOR (Garavaglia et al., 2017). The aim of using these parameters is to adapt to local snow processes, but they could be used primarily to compensate for errors in the input data without satisfactorily achieving it.

1.4. Inverting the hydrological cycle

In contrast, inverting the hydrological cycle with snow-hydrological models may help identify the dependency of the areal inputs on elevation more realistically and enable more accurate snow-hydrological simulations, while simultaneously limiting the number of free parameters. The idea is not completely new and was notably introduced by Valéry et al. (2009) in an attempt to use streamflow measurements to improve knowledge of yearly precipitation in data-sparse mountainous regions. Their results suggested that it was possible to unambiguously identify the altitudinal precipitation gradients from streamflow at a yearly time scale. In another paper, Valéry et al. (2010) proposed regionalisation of daily air temperature and

precipitation to better estimate inputs over high-altitude catchments as regards to the water balance. In their conclusion, the authors claimed that their regionalisation approach also significantly improved the performance of a rainfall-runoff model at a daily time scale. However, the lapse rates in the temperature and precipitation inputs were estimated from gauge observations at the regional scale based on a leave-one-out procedure. This leaves room for potential improvement by locally inferring the lapse rates based on inverse modelling applied at the catchment scale.

A few studies proposed approaches to estimate lapse rates based on hydrological modelling in specific catchments. Zhang et al. (2013) showed that the runoff simulation results involving snowmelt and rainfall runoff were highly sensitive to the temperature and precipitation lapse rates in a Tibetan catchment. Rahman et al. (2014) calibrated the SWAT model in a snow-dominated basin in the Swiss Alps and found also that temperature lapse rate was significantly important for hydrological performance. Naseer et al. (2019) considered a dynamic lapse rate based on a vertical profile of temperature in a catchment in Japan and succeeded to improve the precipitation phase in a distributed hydrological modelling framework. Henn et al. (2016) investigated the value of snow data to constrain the inference of precipitation from streamflow, using lumped hydrologic models and an elevation-band snow model in a Californian basin. Their results suggested that multiple types of hydrologic observations, such as streamflow and SWE, may help to constrain the water balance of high-elevation basins. Le Moine et al. (2015) proposed a calibration strategy where the parameters of both an interpolation model and a daily snow-hydrological model are jointly inferred in a multi-variable approach applied in a catchment in the French Alps. Using a hydro-meteorological modelling chain involving 31 calibrated parameters, they showed the potential of using different types of observations (rain gauges, snow water equivalent measurements and streamflow data) to help assess temperature and precipitation lapse rates according to different weather types. These studies encourage testing whether an inverse modelling approach based on calibrated constant lapse rates can perform well in numerous basins with parsimonious conceptual models.

Mis en forme : Anglais (États Unis)

Improvement is also to be expected from the use of ~~auxiliary~~ observations such as remotely-sensed snow cover data to calibrate and validate models in addition to the runoff measurements, which can help better assess the reliability of the modelled snow processes (see e.g. Parajka and Blöschl, 2008; Thirel et al., 2013). Moreover, other authors (Franz and Karsten, 2013; He et al., 2014; Riboust et al., 2019) showed that adding snow data information to the calibration procedure enabled the identification of more robust snow parameter sets by making the snow models less dependent on the rainfall-runoff model with which they are coupled. Using both streamflow and snow-cover observations in an inverse modelling approach could thus provide further insights into the most relevant snow parameters, while improving our knowledge of the altitudinal temperature and precipitation gradients in data-sparse mountainous regions.

1.5. Objectives

Based on the above issues, this paper investigates whether altitudinal gradients should be inferred from available gauge information when interpolating air temperature and precipitation, or from key-parameters of snow-hydrological models in mountainous areas. To address this question, we use a large dataset of mountainous, snow-affected catchments in the French Alps and we propose a framework to assess the hydrological coherence of gridded datasets and for inferring orographic gradients based on snow-hydrological observations. The rest of the paper is organised as follows. Section 2 describes the study region, the data and their pre-processing. Section 3 provides a brief description of the interpolation procedures tested. Section 4 presents the model assessment methodology. The results are presented and discussed in Section 5, and the main findings, recommendations and future outlooks are summarised in Section 6.

2. STUDY AREA AND DATASET

2.1. Meteorological data

The study was carried out in the French Alps whose altitudes range from 793 to nearly 4800 m a.s.l. (Fig. 1). A dataset of 78 temperature gauges and 148 precipitation gauges was gathered from the RADOME (*Réseau Automatisé d'Observations Météorologiques Etendues*) database of Météo-France (https://publitheque.meteo.fr) for six administrative departments (*Alpes-de-Haute-Provence*, *Hautes-Alpes*, *Alpes-Maritimes*, *Isère*, *Savoie* and *Haute-Savoie*). The extracted series are the mean daily air temperature and the daily liquid equivalent water depth of total precipitation for each station over an 18-year period from the 1st of September 1998 to the 31st of August 2016. These gauges were selected because no gap was originally present in their time seriestheir series present no missing data from the 1st September 2000 to the 31st of August 2016, thus allowing a coherent and stable signal to be represented over the 16-year period of analysis. The corresponding gauge density is ~3 stations per 1000 km² for temperature and ~5 stations per 1000 km² for precipitation, which is close to the recommended minimum density for mountainous areas (~4 stations per 1000 km², WMO, 2008). Although the spatial distribution of the available meteorological stations is reasonably balanced, high altitudes (above 2000 m a.s.l.), which represent approximately 20% of the whole study area and 45% of the catchment surface area (Fig. 1b), remain under-representedare not monitored, as temperature and precipitation gauges are mainly located at low and mid elevations: between 235 and 2105 m for temperature, and between 235 and 2006 m for precipitation.

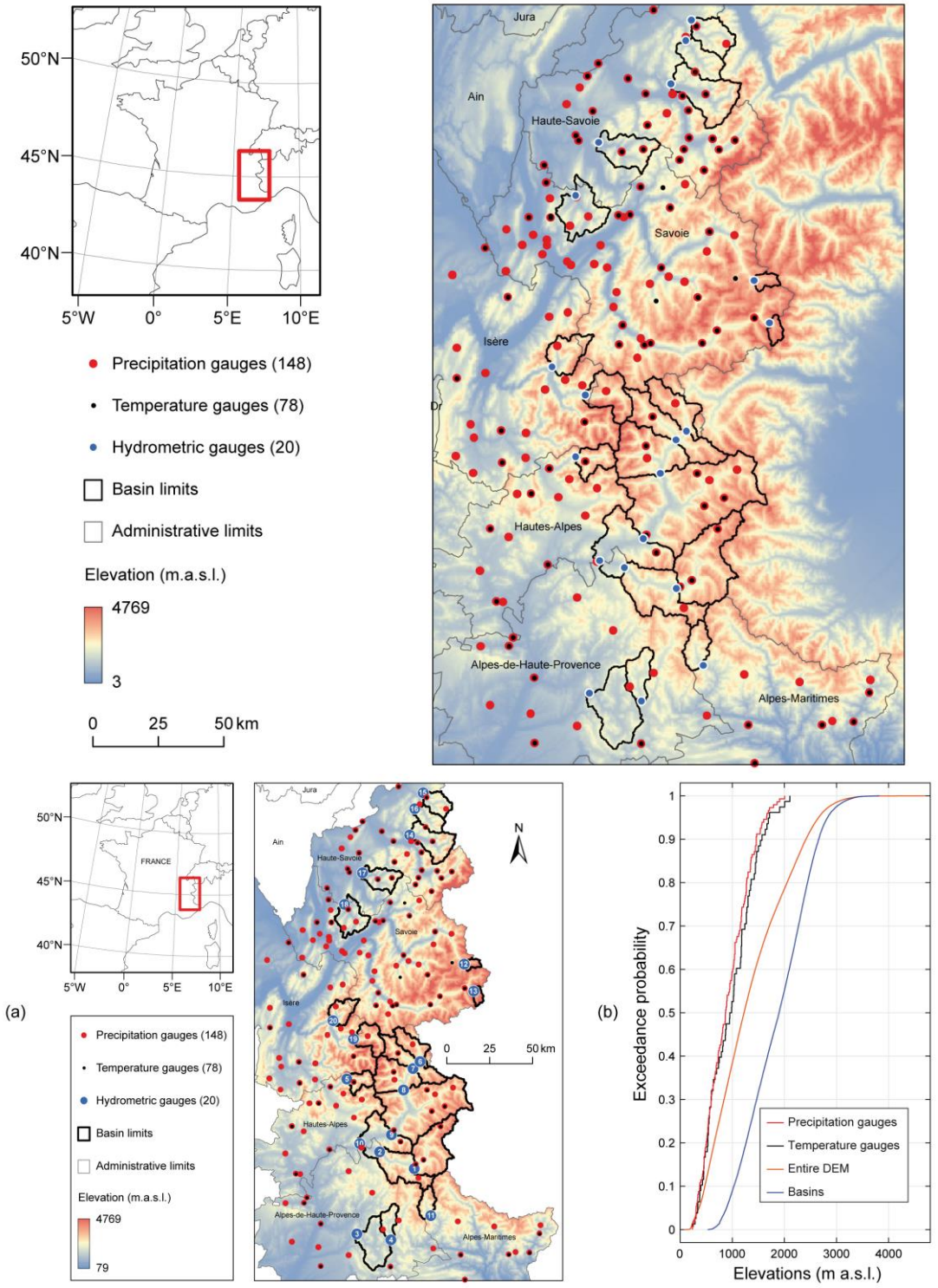


Fig. 1 Study area and data: (a) Location of the selected precipitation, temperature and streamflow stations, as well as elevations from a SRTM digital elevation model (DEM resampled to a grid with 0.5x 0.5 km cells) in the French Alps; (b) - Elevation distributions of in-situ

stations, DEM and basins~~Elevations are from a SRTM digital elevation model which was resampled to a grid with 0.5x 0.5 km cells~~. The station numbers refer to Table 1.

2.2. Streamflow data

In order to avoid case-specific results, a dataset of 20 catchments was gathered from the French hydrological database (www.hydro.eaufrance.fr) over the study area (Fig. 1 and Table 1). The catchments were selected based on the following criteria: (i) their streamflow regime is considered to be natural since they are located upstream from any major hydraulic installations, such as dams and water transfers; (ii) their streamflow regime is moderately to strongly affected by snow; and (iii) their streamflow series present good quality measurements according to the hydrological reports, with less than 10% daily missing values for the period 2000–2016. The catchments are located on high reliefs (the median range of altitude is around 2000 m a.s.l.) and are differently affected by snow (between 15% and 66% of their total precipitation falls in the solid form). During the catchment selection process, we tried avoiding glacierized catchments to minimize possible interactions with non-snow related processes that could also influence streamflow. Therefore, most catchments have zero or limited glacierized areas. Catchment areal precipitation and temperature were estimated after inferring altitudinal gradients, as detailed in the current paper. The dataset cover a large range of hydrological conditions with mean annual precipitation, temperature and streamflow ranging from 811 to 2315 mm, -2.3 °C to +8.9 °C, and 344 mm to 1771 mm, respectively.

Table 1 Streamflow gauging stations and main catchment characteristics. Percentages of glacierized area were estimated from the World Glacier Inventory (NSIDC, 2012). Mean annual precipitation (P), snowfall fraction (S) and temperature (T) were estimated after calibrating local altitudinal gradients over 2000–2016 using the snow-hydrological inverse approach proposed in the current paper (see Test #4 in Table 5).

#	Station	River	Area	Glacierized area	Elevations (m.a.s.l.)		Mean annual precip. (P)	Snowfall fraction (S)	Mean annual temp. (T)	Mean annual streamflow (Q)
			(km²)	(%)	Min	Max	(mm/year)	(%)	(°C)	(mm/year)
1	Barcelonnette	Ubaye	549	0	1132	3308	856	49	1.9	521
2	Lauzet-Ubaye	Ubaye	946	0	790	3308	979	45	3.0	654
3	Beynes	Asse	375	0	605	2273	921	15	8.9	344
4	Saint-André	Issole	137	0	931	2392	1013	23	7.0	481
5	Villar-Lourbière	Séveraisse	133	4	1023	3623	1781	49	2.4	1317
6	Val-des-Prés	Durance	207	0	1360	3059	847	55	0.9	688
7	Briançon	Durance	548	1	1187	3572	811	52	1.6	714
8	Argentière-la-Bessée	Durance	984	3	950	4017	1064	52	2.3	765
9	Embrun	Durance	2170	2	787	4017	1075	48	3.1	693
10	Espinasses	Durance	3580	1	652	4017	1054	45	3.5	654
11	Villeneuve	Var	132	0	926	2862	1072	37	4.6	650
12	Val-d'Isère	Isère	46	9	1831	3538	1245	63	-1.5	1119
13	Bessans	Avérole	45	12	1950	3670	1635	66	-2.3	1311
14	Taninges	Giffre	325	0	615	3044	2315	35	5.0	1771
15	Vacheresse	Dranse d'Abondance	175	0	720	2405	1671	30	4.8	1088
16	La Baume	Dranse de Morzine	170	0	690	2434	1637	32	4.6	1285
17	Dingy-Saint-Clair	Fier	223	0	514	2545	1649	26	6.6	1243
18	Allèves	Chéran	249	0	575	2157	1486	23	6.9	819
19	Mizoën	Romanche	220	9	1057	3846	1369	56	0.9	978
20	Allemond	L'Eau Dolle	172	2	713	3430	1553	47	2.4	1164

Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique

~~**2.2. Streamflow data**~~

~~In order to avoid case-specific results, an extensive dataset of 20 catchments was gathered from the French hydrological database (www.hydro.eaufrance.fr) over the study area (Fig. 1 and Table 1). The catchments were selected based on the~~

~~following criteria: (i) their streamflow regime is considered to be natural since they are located upstream from any major hydraulic installations, such as dams and water transfers; (ii) their streamflow regime is moderately to strongly affected by snow (i.e. between 10% and 70% of their total precipitation falls in solid form); and (iii) their streamflow series present good quality measurements according to the hydrological reports, with less than 10% daily missing values for the period 2000–2016.~~

2.3. Snow cover data

MOD10A1 (Terra) and MYD10A1 (Aqua) snow products version 5 were downloaded from the National Snow and Ice Data Center for the period 24 February 2000–1 January 2017. This corresponds to 6157 dates among which 98.8% are available for MOD10A1 and 85.8% for MYD10A1 since Aqua was launched in May 2002 and became operational in July 2002. These snow products are derived from a Normalised Difference Snow Index (NDSI) calculated from the near-infrared and green wavelengths, and for which a threshold was defined for the detection of snow (Hall et al., 2006, 2007). Cloud cover represents a significant limit for these products, which are generated from instruments operating in the visible-near-infrared wavelengths. As a result, the grid cells were gap filled to produce daily cloud-free snow cover maps of the study area. The different classes in the original products were first merged into three classes: no-snow (no snow or lake), snow (snow or lake ice), no-data (clouds, missing data, no decision, or saturated detector). The missing values were then filled according to a gap-filling algorithm inspired by techniques proposed in the literature (Parajka and Blöschl, 2008; Gafurov and Bárdossy, 2009; Gascoin et al., 2015). The algorithm works in three sequential steps:

(i) Aqua/Terra combination: for every pixel, if no-data was found in MOD10A1 then the value from MYD10A1 was used instead. Priority was given to MOD10A1 because MYD10A1 was found to be less accurate (see Gafurov and Bárdossy, 2009).

(ii) Temporal deduction by sliding time filter: a no-data pixel was reclassified as snow (no-snow) if the same pixel was classified as snow (no-snow) in both the preceding and following grids. The preceding and following grids were searched within a sliding temporal window, whose size was incremented up to 9 days in order to reduce the remaining fraction of no-data pixels to below 10% (Fig. 2a). It should be noted that three periods of gaps in an upper time window (11, 13 and 18 days) were present in the data because of technical failures of the MODIS sensor. In these cases, a longer time deduction was used beforehand to specifically fill these periods.

(iii) Spatial deduction based on elevation and neighbourhood filter: for each date and each pixel, a 3x3 neighbourhood spatial filter was used to account for the elevation and the data in the neighbouring pixels to fill the remaining no-data pixels. Two configurations were considered: either the central pixel has no-data and the algorithm tries to attribute a neighbouring value, or the central pixel has a value that can be assigned to some of its neighbours. The two configurations were repeated until there were no more gaps (Fig. 2a).~~described in Gascoin et al. (2015). The algorithm works in three sequential steps: (i) Aqua/Terra combination; (ii) temporal deduction by sliding the time filter up to 9 days; (iii) spatial deduction by elevation and neighbourhood filter to fill the remaining gaps. The resulting database consists of 5844 binary (snow/no-snow) maps at 500 m spatial resolution for the period 2000–2016 (16 hydrological years, from the 1st of September 2000 to the 31st of August 2016).~~

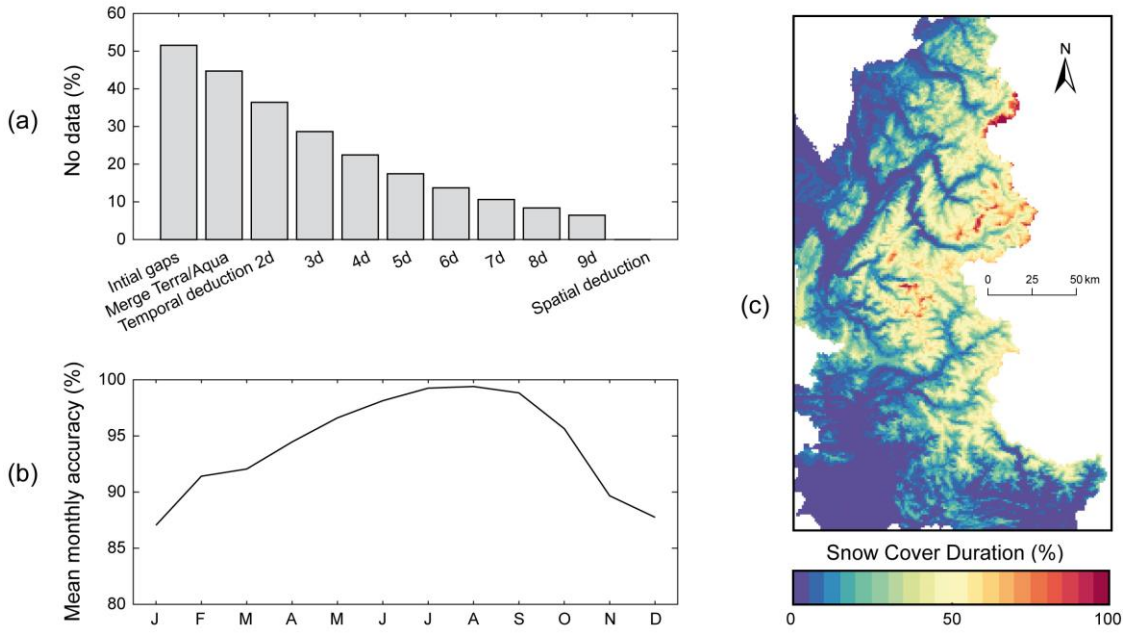


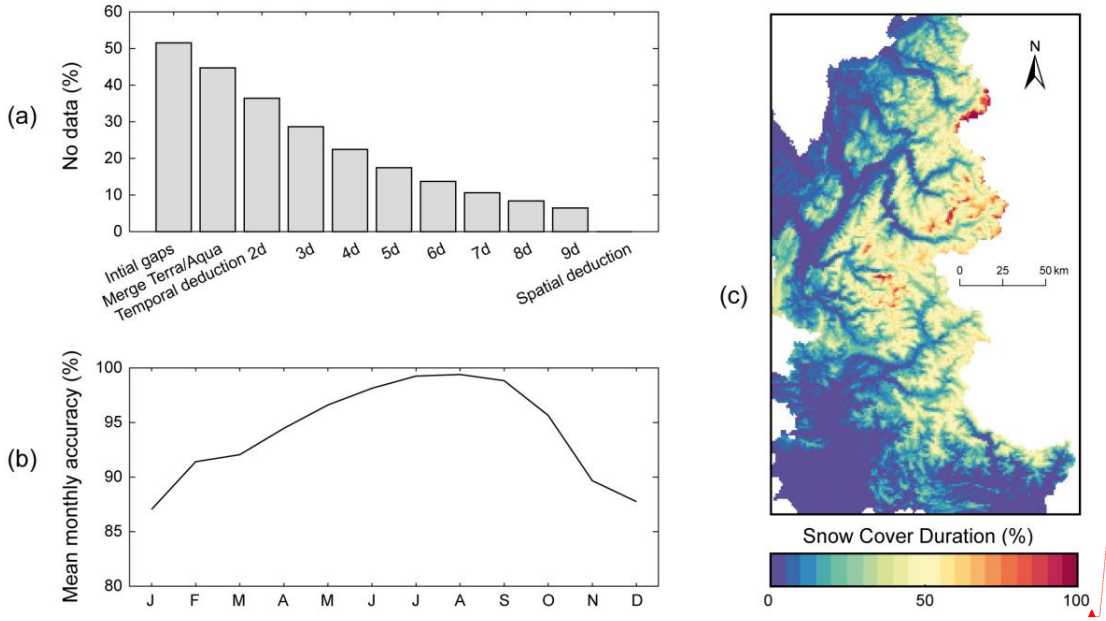
Fig. 2 Results of gap-filling applied to MODIS snow products: (a) Evolution of the number of pixels classified as no-data (e.g., clouds) during the gap-filling procedure; (b) Mean monthly accuracies according to validation based on confusion matrices with 1 image/month, i.e. ~200 cloud-free images over the 2000–2016 period; (c) Snow cover duration based on gap-filled MODIS snow products over the 2000–2016 period.

The resulting database consists of 5844 binary (snow/no-snow) maps at 500 m spatial resolution for the period 2000–2016 (16 hydrological years, from the 1st of September 2000 to the 31st of August 2016). As a synthesis of these maps, snow cover durations over the study area are presented in Fig. 2c.

In order to validate the gap-filling technique, a daily snow product with less than 10% of no-data pixels was selected for each month of the studied period. These images were "blackened" (i.e. with 100% no-data pixels), before applying the algorithm over the entire period to fill all gaps, including validation images. Filling accuracy was estimated for each image removed by computing confusion matrices which compared the pixels of the removed validation images and the filler reconstructions of these images. Validation based on confusion matrices with 1 image/month showed that the gap-filling technique applied to the MODIS snow-products led to the reconstruction of images with average accuracies of 94% (Fig. 2b). The mean monthly accuracies show greater ease in filling gaps in summer than in winter due to the differences in cloud obscuration. However it should be noted that the actual accuracy of the MODIS gap-filling technique is necessarily

Mis en forme : Normal, Espace Avant : 0 pt, Après : 0 pt

greater than that of the validation procedure, in which many quality images needed to fill the gaps were missing.



Mis en forme : Police :10 pt

3. INTERPOLATION PROCEDURES

This section briefly presents the different spatial estimators used in the present study. The interpolation methods analysed include inverse distance weighted (IDW), ordinary kriging (ORK), kriging with external drift (KED) and IDW with external drift (IED). Interest readers can refer to Goovaerts (2000) for a detailed presentation of the different kriging algorithms, and to Diggle and Ribeiro (2007) for their implementation in the public domain in the *GeoPackage* in *R*.

3.1. Spatial interpolation methods

3.1.1. Inverse distance weighting

Let us consider the problem of estimating the given variable *z* at an unsampled location *u* using only surrounding observation data. Let $\{z(u_\alpha), \alpha = 1, \ldots, n\}$ be the set of data measured at *n* surrounding locations u_α. The inverse distance weighting (IDW) method estimated *z* as a linear combination of *n(u)* surrounding observations with the weights being inversely proportional to the square ω distance between observations and *u*:

$$Z_{\mathrm{IDW}}(u) = \frac{1}{\sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)} \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u) z(u_\alpha) \quad \text{with} \quad \lambda_\alpha(u) = \frac{1}{|u - u_\alpha|^\omega} \tag{1}$$

The basic idea behind the weighting scheme is that observations that are close to each other on the ground tend to be more alike than those located further apart, hence observations closer to *u* should receive a larger weight.

3.1.2. Ordinary kriging

Instead of Euclidian distance, geostatistics uses the semivariogram as a measure of dissimilarity between observations. The experimental semivariogram is computed as half the average squared difference between the components of data pairs:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2 \quad (2)$$

where *N(h)* is the number of pairs of data locations a vector *h* apart. The hypotheses of spatial variability were here homogeneity and an isotropic spatial pattern due to the lack of sufficient sampled points, and hence identical variability in all directions.

Kriging is a generalized least-squares regression technique that makes it possible to account for the spatial dependence between observations, as revealed by the semivariogram, in spatial prediction. Like the inverse distance weighting method, ordinary kriging (ORK) estimates the unknown variable *z* at the unsampled location *u* as a linear combination of neighbouring observations:

$$Z_{\mathrm{ORK}}(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ORK}(u) z(u_\alpha) \quad \text{with} \quad \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{ORK}(u) = 1 \quad (3)$$

The ordinary kriging weights $\lambda_\alpha^{ORK}(u)$ are determined such as to minimise the estimation variance $Var\{Z_{\mathrm{ORK}}(u) - \mathrm{z(u)}\}$,while ensuring the unbiasedness of the estimator $E\{Z_{\mathrm{ORK}}(u) - \mathrm{z(u)}\} = 0$. These weights are obtained by solving a system of linear equations known as the ordinary kriging system:

$$\begin{cases} \sum_{\beta=1}^{n(u)} \lambda_\beta(u) \gamma(u_\alpha - u_\beta) - \mu(u) = \gamma(u_\alpha - u) & \alpha = 1, \dots, n(u) \\ \sum_{\beta=1}^{n(u)} \lambda_\beta(u) = 1 \end{cases} \quad (4)$$

where $\mu(u)$ are Lagrange parameters accounting for the constraints on the weights. The only information required by the kriging system (4) are semivariogram values for different lags, and these are readily derived once a semivariogram model has been fitted to experimental values. In this study, we dealt with the fitting of the semivariogram using two existing theoretical models, as presented below:

- Exponential model

$$\gamma(h;\theta) \begin{cases} 0, & \mathrm{h} = 0, \\ \theta_0 + \theta_1[1 - \exp(-3(\|h\|/\theta_2], & \mathrm{h} \neq 0, \end{cases} \quad (5)$$

for $\theta_0 \geq 0,\ \theta_1 \geq 0\ and\ \theta_2 \geq 0$.

- Spherical model

$$\gamma(h;\theta)\begin{cases}0, & \mathrm{h}=0,\\ \theta_0+\theta_1\left(\dfrac{3\|h\|}{2\theta_2}-\dfrac{1}{2}\left(\dfrac{\|h\|}{\theta_2}\right)^3\right), & 0<\|h\|\le\theta_2,\\ \theta_0+\theta_1, & \mathrm{h}>\theta_2,\end{cases} \tag{6}$$

for $\theta_0 \geq 0,\ \theta_1 \geq 0\ and\ \theta_2 \geq 0$.

The spherical model was tested because it is the most widely used semivariogram model and is characterised by linear behaviour (Goovaerts, 2000). The exponential model was selected in addition because it is recommended in the literature for spatial analysis of temperature (Tobin et al., 2011) and precipitation (Bárdossy and Pegram, 2013; Masson and Frei, 2014) in high-mountain regions. Each of these models was combined with a nugget effect, sill and range as parameters. An automatic procedure was necessary to fit the semivariogram model to experimental values over the study period (1998–2016). The models were fitted using regression such that the weighted sum of squares of differences between the experimental and model semivariogram is minimum (see Goovaerts, 2000).

3.2. Accounting for elevation dependency

3.2.1. Kriging with external drift

Kriging with an external drift (KED) predicts sparse variables which are poorly correlated in space by considering that there is a local trend within the neighbourhood; primary data is assumed to have a linear relation with auxiliary information exhaustively sampled over the study area (Ahmed and de Marsily, 1997). KED thus uses secondary information (such as elevation) to derive the local mean of the primary attribute *z* and then performs kriging on the corresponding residuals:

$$Z_{\mathrm{KED}}(u) - m_{\mathrm{KED}}(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{KED}(u)[z(u_\alpha) - m_{\mathrm{KED}}(u_\alpha)] \tag{7}$$

with $m_{\mathrm{KED}}(u) = a_0(u) + a_1(u)y(u)$

where $y(u)$ are elevation data available at all estimation points, a_0 and a_1 are two regression coefficients estimated from the set of collocated variable of interest and elevation data $\{z(u_\alpha), y(u_\alpha), \alpha = 1, \dots, n\}$ using a simple linear relation.

The KED procedure was applied at each time step independently and within each search neighbourhood when the time series were interpolated. The coefficients a_0 and a_1 thus varied in space and time, which makes possible to consider a variable space-time relationship between the primary variable (temperature or precipitation) and the secondary variable (elevation).

3.2.2. Inverse distance weighting with external drift (IED)

The external drift approach was also tested using the inverse distance weighting procedure to propose an original technique, which we called IDW with external drift (IED), as follows:

$$Z_{\mathrm{IED}}(u) - m_{\mathrm{IED}}(u) = \frac{1}{\sum_{\propto=1}^{n(u)} \lambda_\alpha(u)} \sum_{\propto=1}^{n(u)} \lambda_\alpha(u)[z(u_\alpha) - m_{\mathrm{IED}}(u_\alpha)] \tag{8}$$

with $m_{\mathrm{IED}}(u) = a_0(u) + a_1(u)y(u)$

3.3. Leave-one-out procedure

The interpolation parameters ($n(u)$ and ω for IDW and IED, and $n(u)$ and theoretical models for ORK and KED) were calibrated and the interpolation performance was assessed by "leave-one-out" cross validation, which consists of the following principle: a successive estimation of all sampled locations was performed by using all other stations while always excluding the sample value at the location concerned. The spatial models were validated against RMSE (root mean square error~~ normalized with the average of observations~~) for temperature and precipitation at daily, monthly and yearly time ~~steps~~scales. ~~For precipitation, parameters were only estimated for days with at least 1 mm mean precipitation, i.e. approximatively 41% of the daily sample, whereas parameters were calculated for all months and years since there were no locations with dry months or dry years in the dataset.~~ Since the external drift computation and kriging weights can sometimes lead to negative precipitation amounts (Deutsch, 2006), *a posteriori* correction was performed to replace all negative-estimated precipitation values with a zero value.~~ The following performance measures were used to compare the estimations and observations for n locations and t time steps: the RMSE, the absolute simple bias (expressed in °C or mm) and the Nash-Sutcliffe Efficiency criterion (NSE; Nash and Sutcliffe, 1970).~~

The elevations of the gauging stations were used when applying the KED and IED procedures for the "leave-one-out" cross validation. When interpolating temperature and precipitation exhaustively over the study area, the elevation predictors were based on the digital elevation model (DEM) of the shuffle radar topography mission (SRTM; Farr et al., 2007). SRTM originally has a resolution of about 90 m. In this study, we used the SRTM elevation model resampled to a grid with 0.5x0.5 km cells from the UTM32N coordinate reference system. This spatial resolution was judged as a good balance between computational constraints and elevation accuracy.

4. MODEL ASSESS~~S~~MENT METHODOLOGY

The way of accounting for orographic gradients in the temperature and precipitation datasets was also assessed with respect to its ability to contribute to simulations of snow covered area and streamflow at the catchment scale using the following modelling experiment.

4.1. Snow accounting routine (SAR)

The selected SAR (Fig. 3~~2~~a) is a modified version of CEMANEIGE proposed by Valéry et al. (2014). The original version was modified to account for: a snowfall under-catch correction factor as used in the HBV snow routine (see Beck et al., 2016), the computation of fractional snow-covered area (FSC) from a snow water equivalent (SWE) threshold, and possible integration of temperature and precipitation altitudinal gradients.

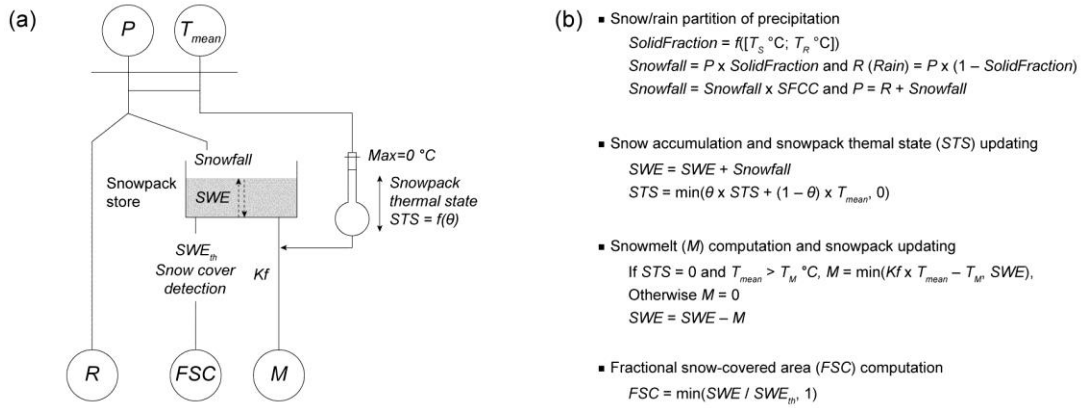


Fig. 3~~2~~ Snow accounting routine: (a) conceptual scheme and (b) associated equations (modified from Valéry et al., 2014). *P*, *R*, T_{mean}, *M* and *FSC* stand for total precipitation, rainfall, mean temperature, melt and factional snow cover, respectively.

Depending on the objectives, the model can be run in a full distributed mode or according to elevation bands. As distributed (or semi-distributed) inputs, it requires the daily liquid equivalent water depth of total precipitation (P) and mean daily air temperature (T_{mean}). In the case of a semi-distributed application at the catchment scale, the first step is to divide the catchment into elevation zones of equal area. Mean areal inputs (P and T_{mean}) are then extracted for each elevation zone from gridded temperature and precipitation datasets. In the present study, the number of elevation zones was set at five due to computational constraints and because preliminary tests showed no significant improvement in the snow-hydrological simulations when a higher spatial resolution (more elevation bands or full distribution) was used.

In each elevation band, the functions of the SAR described in Figure 3~~2~~b are applied with a unique set of parameters. Internal states (snowpack represented according to snow water equivalent (*SWE*) and its thermal state *STS*) vary independently in each elevation zone according to the differences in input values. When gridded temperature and precipitation datasets interpolated without elevation dependency are used, the SAR enables forcing data for each elevation zone to be modified based on two orographic gradients (*TLR* and *PLR*) used as key parameters:

$$T_i(t) = T_i^{IDW}(t) + \left[\frac{\left(TLR + \frac{1}{2}TLR \times Si \times CSV\right)}{100} \times (y_i - y_i^{IDW}(t))\right] \quad (9)$$

with:

$$Si \begin{cases} \sin\left(\frac{2\pi \times (d - 80.5)}{366}\right), & lat > 0 \\ -\sin\left(\frac{2\pi \times (d - 80.5)}{366}\right), & lat < 0 \end{cases}$$

$$P_i(t) = P_i^{IDW}(t) \times [1 + \frac{PLR}{1000} \times (y_i - y_i^{IDW}(t))] \quad (10)$$

where $T_i^{IDW}(t)$ and $P_i^{IDW}(t)$ are, respectively, the mean areal temperature and precipitation interpolated based on the IDW procedure in elevation zone i at time step t; $y_i^{IDW}(t)$ is the mean areal elevation interpolated based on the IDW procedure in elevation zone i from the available gauges at time step t; y_i is the mean areal elevation extracted from the DEM in elevation zone i; *TLR* and *PLR* are the constant temperature and precipitation lapse rates to be calibrated; *CSV* is a coefficient of

Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique, Indice
Mis en forme : Police :Italique
Mis en forme : Police :Italique

seasonal variation due to solar radiation to be applied to *TLR* (when set to 0, no seasonal variation is applied); *Si* is an index of seasonal change in solar radiation accounting for daytime length and ranges from -1 on the 21st of December (winter solstice) to 1 on the 21st of June (summer solstice) in non-leap years in the Northern Hemisphere ($lat > 0$), *d* is the number of days since the 1st of January of the current year.

In the original version of CEMANEIGE, fractional snow-covered area (FSC) is calculated as follows:

$$FSC_i(t) = \min\left(\frac{SWE_i(t)}{SWE_{th}}, 1\right) \quad (11)$$

where *SWE* is the quantity of snow accumulated in snow water equivalent (a state variable of the model, in mm), and SWE_{th} is the model's melting threshold. SWE_{th} is calculated as being equal to 90% of mean annual solid precipitation on the catchment considered (Valéry et al., 2014). Alternative approaches have been proposed to account for the hysteresis that exists between *FSC* and *SWE* during the accumulation and melt phases (Riboust et al., 2019). However, introducing such a hysteresis adds two additional free parameters to the SAR. Instead, SWE_{th} was fixed to 40 mm since preliminary sensitivity analyses showed that this value gave satisfactory *FSC* values when compared to the MODIS observations in the studied catchments.

To ensure insightful comparison with the modelling experiment, the SAR was calibrated according to different modes and degrees of freedom (Table 2). In mode *M1*, elevation dependency, which was accounted for (or not) in the *T* and *P* inputs based on the interpolation procedures, was tested by calibrating five parameters (T_S, T_R, *SFCC*, θ, *Kf*) which control snow accumulation and melt. This mode is usually used to allow snow processes to be adjusted to local conditions and/or the errors in the *T* and *P* inputs. In mode *M2*, all parameters of the SAR were fixed in order to introduce two parameters (*TLR* and *PLR*) as orographic gradients. The aim of using this mode was to account for elevation dependency in the *T* and *P* inputs from constant, calibrated orographic gradients while fixing the parameters that control snow accumulation and melt to physical or general values: precipitation phase determined based on a linear separation between -1 °C and +3 °C (USACE, 1956), temperature threshold for snowmelt fixed at 0 °C, degree-day melt factor set at 5 mm. $°C^{-1}.d^{-1}$ (mean general value taken from Hock, 2003). In mode *M3*, the same approach was chosen but an additional parameter (*CSV*) was associated with the *TLR* gradient in order to test the value of introducing a seasonal variation in the temperature lapse rates (see Eq. 9). In mode *M4*, elevation dependency in the *T* and *P* inputs was also accounted for based on three (*TLR*, *CSV* and *PLR*) parameters and two other parameters (θ, *Kf*) were calibrated in addition to allow for snowmelt adjustment.

In each altitudinal band, five outputs (rainfall, snowfall, snowmelt, potential evapotranspiration and fractional snow-covered area) are computed at each daily time step. Rainfall (*R*) and snowmelt (*M*) are summed to compute the total quantity of water available for production and transfer in the catchment. Potential evapotranspiration (*PE*) is computed for each altitudinal band using the temperature-based formulation proposed by Oudin et al. (2005):

$$PE_i(t) = \frac{R_e}{\lambda\rho}\frac{T_i(t)+5}{100} \quad \text{if } (T_i(t)+5) > 0; \quad else\ PE_i(t) = 0 \quad (12)$$

where R_e is the extra-terrestrial solar radiation ($MJ.m^{-2}.day^{-1}$) which depends on the latitude of the basin and the Julian day of the year, λ is the net latent heat flux (fixed at 2.45 $MJ.kg^{-1}$), ρ is the water density (set at 11.6 $kg.m^{-3}$) and $T_i(t)$ is the air temperature (°C) estimated in the elevation zone *i* at time step *t*.

Table 2 Parameters of the snow accounting routine and their associated fixed values or ranges tested in each modelling experiment.

			Fixed values or ranges tested			
Param.	Meaning	Unit	*M1*	*M2*	*M3*	*M4*

Mis en forme : Police :10 pt, Vérifier l'orthographe et la grammaire
Mis en forme : Police :(Par défaut) Times New Roman, 10 pt, Non Gras, Vérifier l'orthographe et la grammaire
Mis en forme : Police :Italique
Mis en forme : Police :Italique, Indice
Mis en forme : Indice
Mis en forme : Police :Italique
Mis en forme : Police :Italique, Indice
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Non Surlignage

T_S	Temperature between the solid and liquid phase	°C	[-3; 3]	-1	-1	-1
T_R	Thermal range for the phase separation above T_S	°C	[0; 10]	4	4	4
SFCC	Snowfall gauge under-catch correction factor	-	[1; 3]	1	1	1
θ	Weighting coefficient for snowpack thermal state	-	[0; 1]	0	0	[0; 1]
T_M	Temperature threshold for snowmelt	°C	T_S + 1	0	0	0
Kf	Degree-day melt factor	mm.°C^{-1}.d^{-1}	[0; 10]	5	5	[0; 10]
SWE_{th}	Snow water equivalent threshold to compute FSC	mm	40	40	40	40
TLR	Temperature lapse rate	°C (100m)$^{-1}$	-	[0;-1.5]	[0;-1.5]	[0;-1.5]
CSV	Coefficient of seasonal variation applied to *TLR*	-	-	0	[0; 1]	[0; 1]
PLR	Precipitation lapse rate	% (km)$^{-1}$	-	[0; 200]	[0; 200]	[0; 200]

The outputs of each band are averaged to estimate the total liquid output of the SAR and *PE* at the catchment scale in order to feed the combined hydrological models (Fig. 4~~3~~).

4.2. Hydrological models

To avoid model-specific results, two well-known hydrological models (see structures in Fig. 4~~3~~ and parameters in Table 3) were chosen in association with the SAR: the 4-parameter GR4J presented by Perrin et al. (2003) and a 9-parameter lumped version of the HBV model (Bergström, 1995; Beck et al., 2016), here referred to as HBV9 to avoid confusion with the original version.–

The two models were run at a daily time step and used in lumped mode with the SAR on top. The structure and the number of degrees of freedom differ between GR4J and HBV9, which should make ~~our~~ results more generalizable.

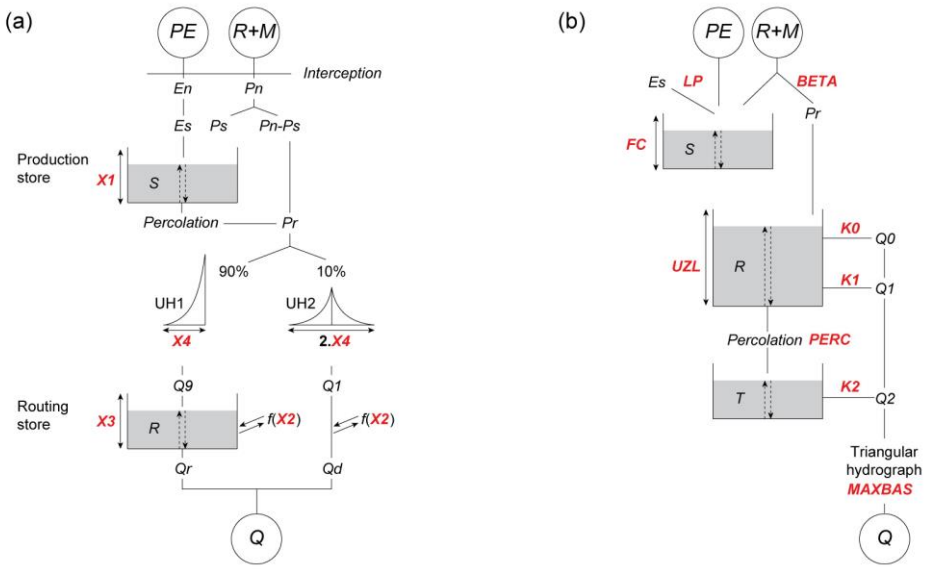


Fig. 4~~3~~ Diagrams of the two hydrological models used: (a) GR4J and (b) HBV9. Calibrated parameters are in red and are further described in Table 3. *R*, *M*, *PE* and *Q* stand for rainfall, snowmelt, potential evapotranspiration and streamflow, respectively.

Table 3 Parameters of the hydrological models and their associated ranges tested.

Model	Parameter	Meaning	Unit	Tested range
GR4J	*X1*	Maximum capacity of the production store *S*	mm	[~~1~~0; 1~~5~~000]
	X2	Inter-catchment exchange coefficient	mm.d^{-1}	[-5; 5]
	X3	Maximum capacity of the non-linear routing store *R*	mm	[0; 500]
	X4	Unit hydrograph (UH) time base	d	[0.5; 5]

Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique

HBV9	*BETA*	Shape coefficient of recharge function	-	[0.5; 5]
	FC	Maximum water storage in the unsaturated-zone store *S*	mm	[10; 1500]
	LP	Fraction of soil moisture above which actual evapotranspiration reaches PE	-	[0.3; 1]
	K0	Additional recession coefficient of the upper groundwater store *R*	-	[0.05; 1]
	K1	Recession coefficient of the upper groundwater store *R*	-	[0.1; 0.8]
	UZL	Threshold value for extra flow from the upper zone	mm	[0; 500]
	PERC	Maximum percolation to the lower zone	$mm.d^{-1}$	[0; 6]
	K2	Recession coefficient of the lower groundwater store *T*	-	[0.01; 0.15]
	MAXBAS	Length of the equilateral triangular weighting function	-	[1; 7]

4.3. Calibration and validation methods

4.3.1. General model assessment

The models (GR4J and HB9 with the SAR on top) were cross-validated using a split-sample test procedure. The simulation period (2000–2016) was split into two sub-periods alternatively used for calibration and validation. -Thus two calibration and two validation tests were performed to provide ~~results in validation~~evaluation on all available data. Mean annual precipitation increased by around 17% between the two periods, while mean annual temperature were stable (9.2 °C vs. 9.3 °C) across the in-situ network presented in Section 2.1. Although the second period was generally wetter, differences can be observed locally. At the basin scale, the differences between the two periods ranged from -10% to +15% for precipitation, -0.5 °C to +0.5 °C for temperature, and -11% to +50% for streamflow.

- The models were run in a continuous way for the whole reference period, while only hydrological years (from the 1st of September to the 31st of August) corresponding to the calibration and validation periods were considered to compute the efficiency criteria. The 3-year period before the simulation period was used for model warm-up to limit the effect of the storage initialisation, and was not included in performance computation.

4.3.2. Optimisation algorithm and objective function

The parameters of the SAR and the hydrological model were optimised simultaneously, using the Shuffled Complex Evolution (SCE) algorithm (Duan et al., 1992). The algorithmic parameters of SCE were set to the values recommended by Duan et al. (1994) and Kuczera (1997) to reduce the risk that SCE fails in local optimal solutions. The objective function (OF) used was a multi-criteria composite function focusing simultaneously on variations in snow-covered area~~,~~ and streamflow dynamics ~~and water balance~~ at the basin scale, as follows:

$$\mathrm{OF} = 1 - (0.5 \times \mathrm{NSE}_{\mathrm{SNOW}} + 0.5 \times \mathrm{NSE}_{\mathrm{sqrQ}})$$

with:

$$\mathrm{NSE}_{\mathrm{SNOW}} = \frac{1}{E}\sum_{i=1}^{E}\left(1 - \frac{\sum_{t=1}^{N}(FSC_{sim,t}^{i} - FSC_{obs,t}^{i})^2}{\sum_{t=1}^{N}(FSC_{obs,t}^{i} - \overline{FSC_{obs}^{i}})^2}\right) \quad (1~~3~~2)$$

$$\mathrm{NSE}_{\mathrm{sqrQ}} = 1 - \frac{\sum_{t=1}^{N}(\sqrt{Q_{sim,t}} - \sqrt{Q_{obs,t}})^2}{\sum_{t=1}^{N}(\sqrt{Q_{obs,t}} - \overline{\sqrt{Q_{obs}}})^2}$$

where $FSC_{obs,t}^{i}$ and $FSC_{sim,t}^{i}$ are the observed and simulated fractional snow-covered area (FSC) in elevation zone *i* at daily time step *t*, *N* is the total number of time steps, $\overline{FSC_{obs}^{i}}$ is the mean observed FSC in elevation zone *i* over the test period, *E* is

the total number of elevation zones (fixed to 5 for the study), $Q_{obs,t}$ and $Q_{sim,t}$ are the observed and simulated streamflows at daily time step t, $\overline{\sqrt{Q_{obs}}}$ is the mean observed square root transformed flows over the test period.

NSE$_{SNOW}$ relies on the Nash-Sutcliffe Efficiency criterion. Perfect agreement between the observed and simulated values gives a score of 1, whereas a negative score represents lower reproduction quality than if the simulated values had been replaced by the mean observed values. NSE$_{sqrQ}$ can be considered as a multi-purpose criterion focusing on the simulated hydrograph. It puts less weight on high flows than the standard NSE on non-transformed discharge (Oudin et al., 2006). As the majority of basins had negligible glacierized areas (see Table 1), no specific glacier model was activated. This led to ignore the late summer contribution of glacier melt to river discharge in the three basins having 9–12% glacierized areas.

4.3.3. Efficiency criteria in validation

Four criteria were used to evaluate model performance during validation. The first one was the NSE$_{SNOW}$ criterion. To put more emphasis on high and low flow conditions, we used the NSE on non-transformed streamflows (NSE$_{Q}$) that gives more weight to large errors generally associated with peak flows, and the NSE on log-transformed streamflows (NSE$_{lnQ}$). The absolute cumulated volume error (VE$_{C}$) was also computed to obtain information on the agreement between observed and simulated total discharge over the test periods:

$$\mathrm{VE_C} = 1 - \frac{\left|\sum_{t=1}^{N} Q_{sim,t} - \sum_{t=1}^{N} Q_{obs,t}\right|}{\sum_{t=1}^{N} Q_{obs,t}} \quad (14\text{~~3~~})$$

A value of 1 indicates perfect agreement while values less than 1 indicate over- or underestimation of the volume.

5. RESULTS AND DISCUSSION

5.1. Cross-validation of the interpolation methods

Table 4 lists the results of cross-validation of the interpolation methods against yearly, monthly and daily series from temperature and precipitation gauges. Kriging with ORK led to an improvement over IDW only for precipitation interpolation at the yearly and monthly time scales. Considering elevation dependency with external drift (KED and IED) ~~significantly~~ improved the performance of the kriging and inverse-distance methods, except for precipitation estimated at the daily time scale. This shows that the correlation between precipitation and topography increases with the increasing time aggregation as already reported in other studies (e.g., Bárdossy and Pegram, 2013; Berndt and Haberlandt, 2018). The elevation-dependency of precipitation thus depends significantly on the accumulation time. At the daily time scale, the orographic enhancement is limited because on a given day there is no monotonic relationship between elevation and precipitation amount: it depends on where the precipitation event occurs in the first place.

Of all the methods tested, IED provided the best performance in terms of ~~both~~ lower RMSE ~~and MAE, and higher NSE~~ for each variable (temperature and precipitation) and at all temporal resolutions (yearly, monthly and daily), except for precipitation at the daily time scale for which IDW performed best.

The exponential variogram model for the ORK and KED method performed systematically better than the spherical model whatever the variable and the time scale. In contrast, the exponent ω used with the IDW and IED methods varies from 1 to 3 depending on the considered variable and time scale. The optimised number of surrounding neighbours $n(u)$ also varies depending on the method and time scale. At the daily time scale, $n(u)$ ranged from 6 when interpolating temperature with ORK to 17 when interpolating precipitation with KED and

IED. Hence, 10 (17) surrounding neighbours were used to compute altitudinal gradients of temperature (precipitation) based on the daily linear regressions with KED and IED.

Table 4 Cross-validation of the interpolation methods against yearly, monthly and daily series from meteorological gauges over the period 2000–2016. The best efficiency criteria for each analytical time scale and each variable of interest (temperature and precipitation) are in bold. The values of $n(u)$ and ω (for IDW and IED) and of $n(u)$ and *model* (for ORK and KED) represent the interpolation parameters, which were optimised using the leave-one-out procedure, as described in section 3.3.

		Temperature (78 gauges)				Precipitation (148 gauges)			
		Without elevation dependency		With elevation as external drift		Without elevation dependency		With elevation as external drift	
		IDW	ORK	KED	IED	IDW	ORK	KED	IED
Yearly	RMSE	1.82 °C	1.82 °C	0.65 °C	**0.65 °C**	177.05 mm	174.27 mm	153.75	**150.31 mm**
	$n(u)$	6	6	8	10	4	15	12	12
	ω	1	-	-	2	3	-	-	3
	model	-	exponential	exponential	-	-	exponential	exponential	-
Monthly	RMSE	1.91 °C	1.92 °C	0.86 °C	**0.80 °C**	23.19 mm	22.73 mm	22.35 mm	**22.20 mm**
	$n(u)$	7	6	8	10	5	15	12	12
	ω	1	-	-	2	2	-	-	2
	model	-	exponential	exponential	-	-	exponential	exponential	-
Daily	RMSE	2.16 °C	2.18 °C	1.21 °C	**1.20 °C**	**2.83 mm**	2.86 mm	2.91 mm	2.90 mm
	$n(u)$	7	6	10	10	10	10	17	17
	ω	1	-	-	2	2	-	-	2
	model	-	exponential	exponential	-	-	exponential	exponential	-

Figure 5~~4~~ shows the annual temperature and precipitation maps obtained by interpolation daily data from the meteorological gauges with the period 2000–2016 using the IDW, ORK, KED and IED methods and their optimised parameters (Table 4). The maps of IED estimates of mean temperature and annual precipitation closely resemble the KED maps. Temperature estimates range from -9.2 °C to 16.0 °C with KED, and from -11.2 °C to 16.6 °C with IED. Precipitation estimates range from 630 mm to 3273 mm with KED, and from 642 mm to 3184 mm with IED. As expected, these ranges are wider than those obtained with the IDW and ORK procedures, which, unlike the KED and IED methods, do not consider either local or seasonal elevation dependency. As a result, the ranges obtained with KED and IED are probably more realistic with respect to temperature, but not necessarily with respect to precipitation, for which daily cross-validation shows that the simple IDW method provided better results. However, cross-validation was based on ~~precipitation~~ gauges sampled only below 2006 m a.s.l. (2105 m a.s.l.), meaning ~~validation~~ evaluation of the ~~temperature and~~ precipitation (temperature) gridded datasets at higher altitudes was not possible. Another approach is thus needed to further explore whether elevation dependency should be disregarded when estimating daily precipitation (as suggested by cross-validation), and, if not, whether this dependency should be accounted for in the interpolation process or by inverting the hydrological cycle. A sensitivity analysis of snow-hydrological simulations to the orographic gradients was thus conducted.

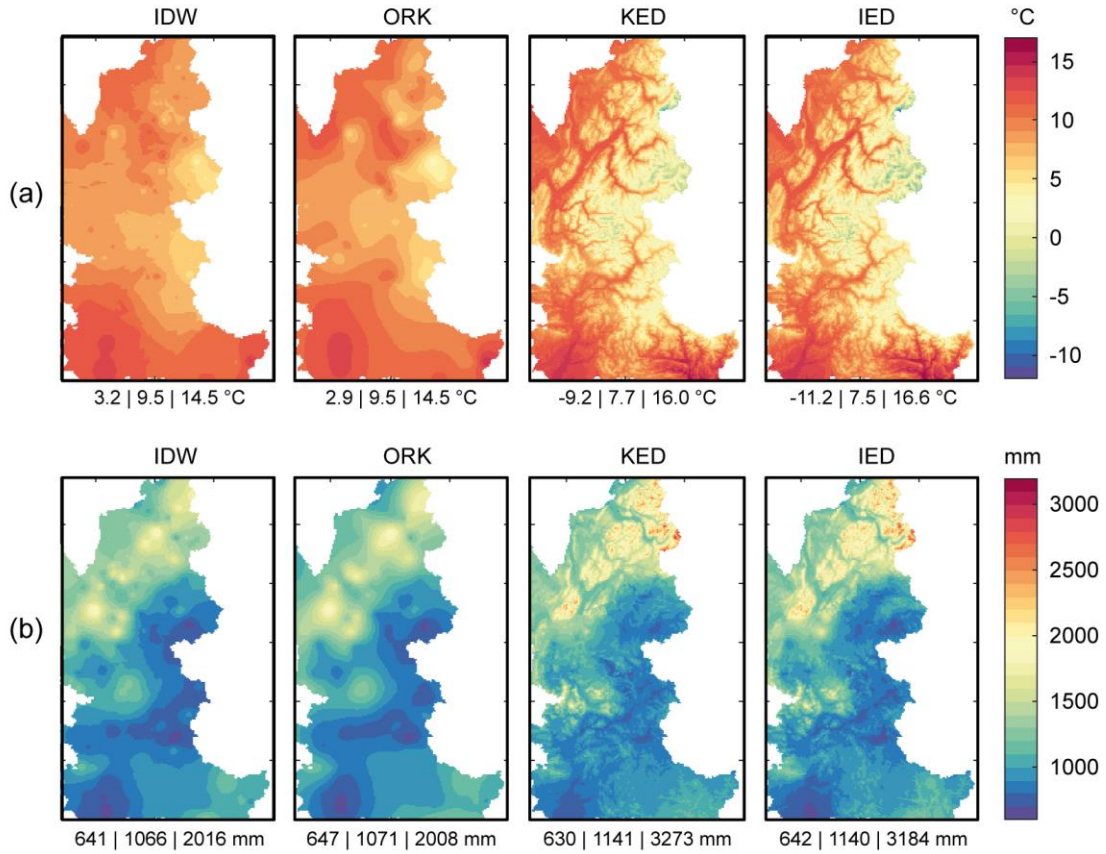


Fig. 54 Maps of mean annual temperature (°C) and total precipitation (mm per year) obtained by interpolation of (a) 78 gauges (temperature) and of (b) 148 gauges (precipitation) with daily data for the period 2000–2016 using the inverse distance weighted (IDW), ordinary kriging (ORK), kriging with external drift (KED) and IDW with external drift (IED). The numbers below each map stand respectively for minimum, mean and maximum values (expressed in °C for temperature and in mm/year for precipitation) in the maps.

5.2. Sensitivity of the snow-hydrological simulations to the orographic gradients

For the sake of brevity, here we only present the results ~~we~~ obtained with the datasets interpolated with the IDW and IED procedures, since cross-validation at the daily time scale showed that they slightly outperformed the ORK and KED methods, respectively. Table 5 summarises the six tests performed to account for elevation dependency in the *T* and *P* inputs via the modelling experiment described in section 4.

Table 5 Description of the tests to account for elevation dependency in the T and P inputs via the modelling experiment described in section 4. Note that each calibration tests included also the hydrological parameters of GR4J or HBV9 (the parameter ranges tested are listed in Table 2 for the SAR and in Table 3 for the hydrological models).

Mode	Test number	*T* input	*P* input	Calibrated parameters (excluding hydrological models)	Principle
M1	1	T-IDW	P-IDW	T_S, T_R, *SFCC*, θ, *Kf*	No elevation dependency in the *T* and *P* inputs, and five calibrated parameters for adjustment of snow accumulation and melt
	2	T-IED	P-IDW	T_S, T_R, *SFCC*, θ, *Kf*	Elevation dependency only in the *T* input based on the IED interpolation procedure, and five calibrated parameters for adjustment of snow accumulation and melt
	3	T-IED	P-IED	T_S, T_R, *SFCC*, θ, *Kf*	Elevation dependency in the *T* and *P* inputs based on the IED interpolation procedure, and five calibrated parameters for adjustment of snow accumulation and melt
M2	4	T-IDW	P-IDW	*TLR*, *PLR*	Elevation dependency in the *T* and *P* inputs considered based on two calibrated

					parameters in the SAR, and fixed parameters for snow accumulation and melt
M3	5	T-IDW	P-IDW	*TLR, CSV, PLR*	Elevation dependency in the *T* and *P* inputs considered based on three calibrated parameters in the SAR, and fixed parameters for snow accumulation and melt
M4	6	T-IDW	P-IDW	*TLR, CSV, PLR, θ, Kf*	Elevation dependency in the *T* and *P* inputs considered based on three calibrated parameters in SAR, and two calibrated parameters for adjustment of snow melt

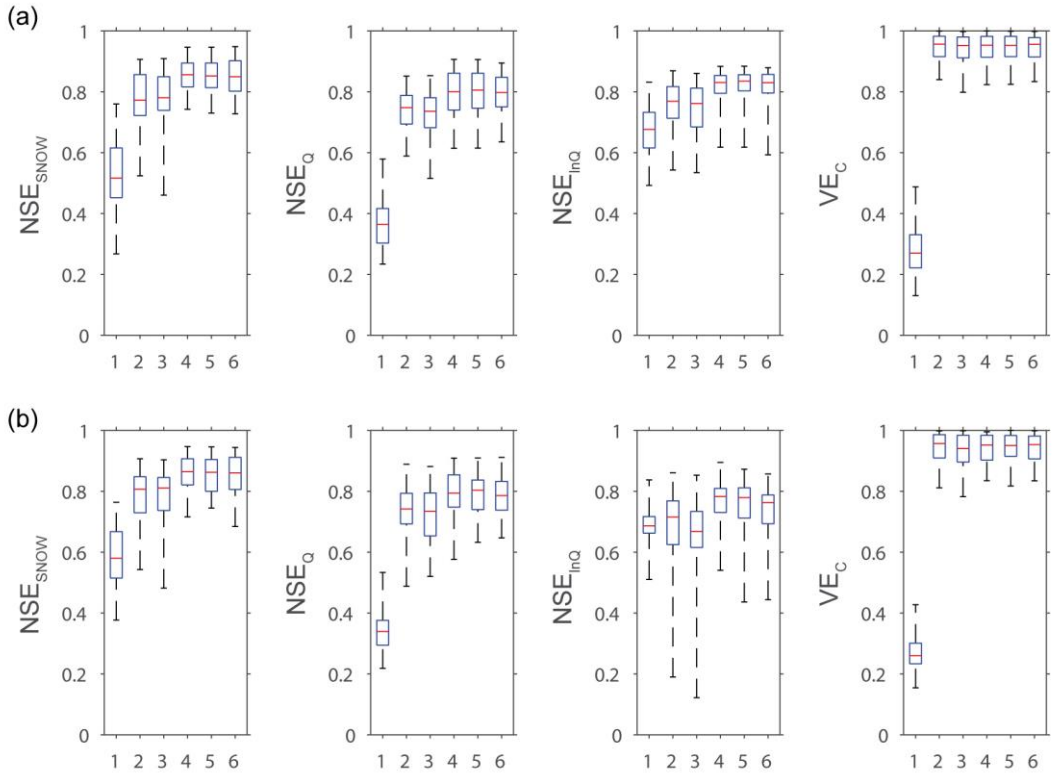


Fig. ~~6~~5 Boxplots (showing 0.00, 0.25, 0.50, 0.75 and 1.00 percentiles) of the efficiency distributions obtained in validation by the (a) GR4J and (b) HBV9 models combined with the snow model according to six different tests (see Table 5) to account for elevation dependency in the *T* and *P* inputs on the 20 snow-affected Alpine catchments.

Figure ~~6~~5 and Table 6 summarise the efficiency distributions obtained in validation with the GR4J and HBV9 models combined with the snow model in the different tests on the 20 snow-affected Alpine catchments. The results produced by the two hydrological models are in agreement and highlight the following main findings:

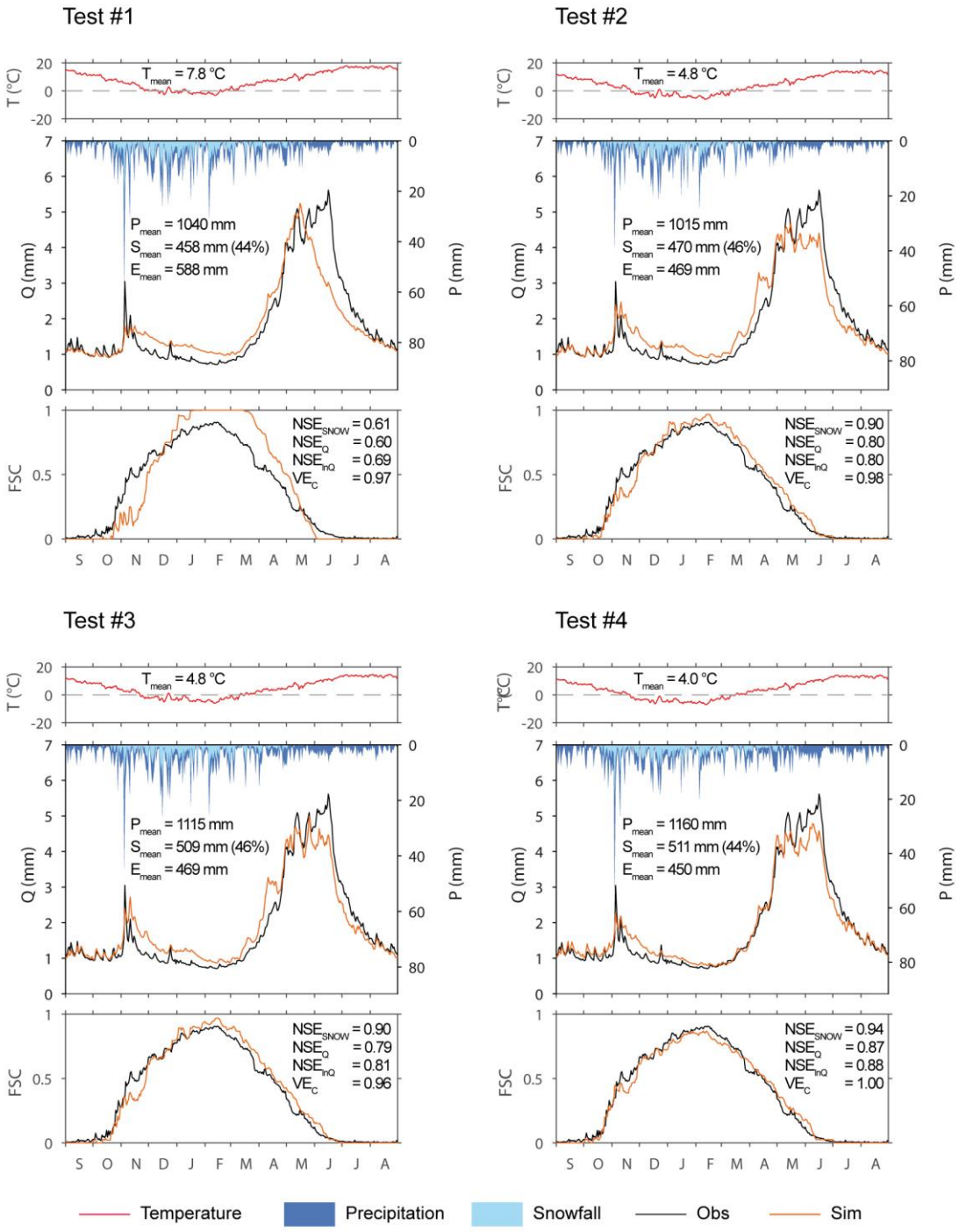
- Not considering elevation dependency in either the *T* or *P* inputs (Test #1) leads to notable failures of the snow-hydrological models, due to incorrect snow/rainfall partitioning and snowmelt in space and over time caused by too high temperatures and insufficient input volumes of precipitation, which cannot be offset by the free parameters of the SAR. Notably, calibrating temperature thresholds and ranges for snow/rain partition and snow melt, as well as snow under-catch (using the *SFCC* parameter) is clearly unsatisfactory.
- Considering elevation dependency only in the *T* inputs based on the IED procedure (Test #2) significantly improves the snow-hydrological simulations, but considering elevation dependency in the *P* inputs based on the same procedure (Test #3) is not as efficient, notably for streamflow simulations. This shows that the estimated precipitation with IED over the catchments is of limited accuracy.

- Improving the areal temperature and precipitation estimation clearly requires the calibration of altitudinal temperature and precipitation gradients. The snow-hydrological simulations are considerably improved when using the parsimonious 2-parameter SAR based only on the calibration of *TLR* and *PLR* (Test #4).~~,~~
- Compared to Test #4 based only on a 2-parameter SAR, only limited improvements in the performance distributions are obtained by introducing additional free parameters to account for the seasonal variability of the temperature gradients (Test #5) and for local adjustment of snowmelt (Test #6).

Table 6 Mean validation efficiency of the 6 modelling tests (see Table 5) on the set of 20 catchments with the GR4J model and the HBV9 model.

Model	Test number	Number of free parameters of the SAR	Mean NSE_{SNOW}	Mean NSE_{Q}	Mean NSE_{lnQ}	Mean VE_{C}
GR4J	1	5	0.539	0.3759	0.66	0.2991
	2	5	0.789	0.740	0.764	0.9494
	3	5	0.79	0.7266	0.741	0.940
	4	2	0.86	0.798	0.821	0.95
	5	3	0.856	0.8079	0.821	0.95
	6	5	0.856	0.8079	0.812	0.95
HBV9	1	5	0.597	0.3466	0.6970	0.2791
	2	5	0.798	0.740	0.687	0.941
	3	5	0.789	0.7263	0.651	0.9388
	4	2	0.867	0.795	0.764	0.942
	5	3	0.867	0.797	0.756	0.942
	6	5	0.857	0.797	0.756	0.943

As a representative example of the studied catchments, Figure 7~~7~~ illustrates the differences in the simulations obtained by Tests #1 to #4 for the Durance at Serre-Ponçon. This 3580 km² catchment with altitudes ranging between 652 and 4017 m a.s.l. ensures inflows to one of the biggest dams in Europe (maximum capacity of 1.3 km^3). Dynamics of fractional snow cover area and streamflow are better simulated when considering elevation dependency of the *T* and *P* inputs via two calibrated, altitudinal gradients (Test #4). Compared to the other tests, mean annual temperature (4.0~~1~~ °C) is lower and mean annual precipitation (~~1212~~ 1160 mm) is higher. Less precipitation is considered in solid form (44% of total precipitation on average) and accumulation is longer during winter, which fits streamflow observations better, both for low flows from December to April and for flood peaks between May and July. It is worth noting that these ~~more realistic~~improved simulations were obtained with a SAR calibrated on only two parameters targeting the local lapse rates whereas the other simulations were based on a SAR calibrated on five parameters. This shows that calibrating the usual snow parameters to compensate for errors in the input data and/or to adapt to local snow-related processes is less efficient in the simulations than inferring only temperature and precipitation lapse rates while ~~setting~~ fixing all the other parameters~~ to physical or general values~~. This suggests that correcting for temperature and precipitation distribution has a stronger impact on model predictions than adjusting for snow-related processes like phase partitioning or melt and that correctly estimating total accumulation is likely to play a first-order role in the snow-hydrological responses of the studied catchments.



585 **Fig. 7** Comparison of snow-hydrological simulations with elevation dependency according to Tests #1 to #4 (see Table 5) with GR4J for the Durance at Serre-Ponçon. The graphs show mean inter-annual time-series of temperature, precipitation, streamflow and fractional snow cover at the catchment scale in validation over the period 2008–2016. T_{mean}, P_{mean} and S_{mean} stand for mean annual temperature, precipitation, and snowfall, respectively. The efficiency criterions NSE_{SNOW}, NSE_{Q}, NSE_{lnQ} and VE_{C} are computed from continuous (not mean seasonal) series over 2008–2016. ~~This suggests that adapting to local snow processes is not indispensable and that~~
590 ~~compensating for the errors in the input data is better achieved by simply calibrating the local temperature and precipitation gradients.~~

Mis en forme : FigCaption

Mis en forme : Police :9 pt

5.3. Identifiability of the parameters

Figure 8~~8~~ shows a~~n~~ representative example of parameter sensitivity to the objective function (*OF*) according to the six tests (see Table 5) with the GR4J model on the Durance at Serre-Ponçon. The maximum allowed parameter range is only reached for the parameters *X1* and *X2* with Test #1. This test differs from the others because no elevation dependency in the *T* and *P* inputs are considered. Consequently, hydrologic predictions of Test #1 are significantly outperformed by the other approaches. Extending the parameter ranges beyond the tested values would be both poorly efficient in improving the simulations and incorrect from a numerical point of view since they were set to values recommended by the models' authors. Moreover, no maximum parameter limits were reached in the other tests, thus suggesting that the parameter ranges are adequate.

As already shown, considering elevation gradients (Tests #4, #5 and #6) minimises OF and significantly improves model performance. It also improves the parameter identifiability. The temperature altitudinal gradient (*TLR*) is easily identifiable with values ranging from -0.6~~6~~ °C/100~~h~~m (Test #4 and Test #5) to -0.~~70~~ 67 °C/100m (Test #6), and with variation coefficients of 0.1% for the 20% best-performing parameter solutions. It clearly ~~appears to be~~reveals as a key parameter for improving snow and streamflow simulations compared to parameters calibrated using elevation gradients inferred from usual interpolation method with external drift (Tests #2 and #3). The optimum value of the *CSV* parameter (Tests #5 and #6) is zero, clearly indicating no need to account for the seasonal variation in the temperature lapse rate in the catchment studied here (which is also the case in most catchments). The precipitation lapse rate (*PLR*) is also easy to identify with optimised values~~(around~~ of 6~~2~~0%/km (Tests #4–5) and variation coefficients 0.4%~~ in the catchment studied)~~. Introducing additional parameters controlling snowmelt (θ and *Kf* in Test #6) does not significantly improve the simulations and decreases the parameter identifiability (variation coefficients increase compared to Tests #4 and Test #5 based on a 2-parameter and 3-parameter SAR, respectively). This shows that model performance is mainly sensitive to the use of parameters for temperature and precipitation lapse rates and that a 2-parameter SAR based on *TLR* and *PLR* (Test #4) on top of the hydrological models tested is both essential and sufficient to ~~produce~~ achieve satisfactory simulations.

-Equifinality is also reduced in Tests #4–6 for the~~more marked for the~~ parameters controlling runoff generation and routing (*X1*, ~~*X2,*~~ *X3* and *X4*). On the opposite, the parameter of the inter-catchment groundwater flows (*X2*) is poorly identifiable with variation coefficients of 24.8%, 20.3% and 143.1% with Test #4, Test #5 and Test #6, respectively. This suggests that inter-catchment groundwater exchanges (IGE) do not play a key role in the studied catchments. Indeed, fixing *X2* to a value of 0 (i.e. without potential IGE) with an alternative GR3J model provided similar mean validation efficiency on the set of catchments as compared to the GR4J associated with the 2-parameter SAR (Table 7). However, other objective functions may result in other findings as far as IGE are concerned. For instance, additional tests (not shown here for brevity sake) confirmed that it was possible to greatly reduce the *X2* equifinality without decreasing the model efficiency by adding a water balance term in the objective function to constrain the proportion of years respecting the water and energy balance in the Turc-Budyko non-dimensional graph (see Andréassian and Perrin, 2012). These tests~~, suggesting that these parameters somehow interact. However, the parameter of the inter-catchment groundwater flows (*X2*) shows a clear optimum towards positive values, indicating the need for additional water, which cannot be totally offset by the calibrated precipitation lapse rate. This result~~ suggest~~s~~ed that it ~~remains~~ may be relevant~~important~~ to explicitly represent inter-catchment groundwater transfers in association with correcting or scaling factors applied to the precipitation input data to render the distribution between evapotranspiration, streamflow and underground fluxes more realistic, as already reported by Le Moine et al. (2007).

Table 7 Mean validation efficiency on the set of 20 catchments with the GR4J model and the GR3J model in association with the 2-parameter SAR.

Model	Total number of free	Mean	Mean	Mean	Mean

Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Non Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Mis en forme : Police :Italique
Tableau mis en forme

	parameters	NSE_{SNOW}	NSE_{Q}	NSE_{lnQ}	VE_{C}
2-parameter SAR/GR4J	6 (2 + 4)	0.86	0.79	0.82	0.95
2-parameter SAR/GR3J	5 (2 + 3)	0.86	0.78	0.81	0.94

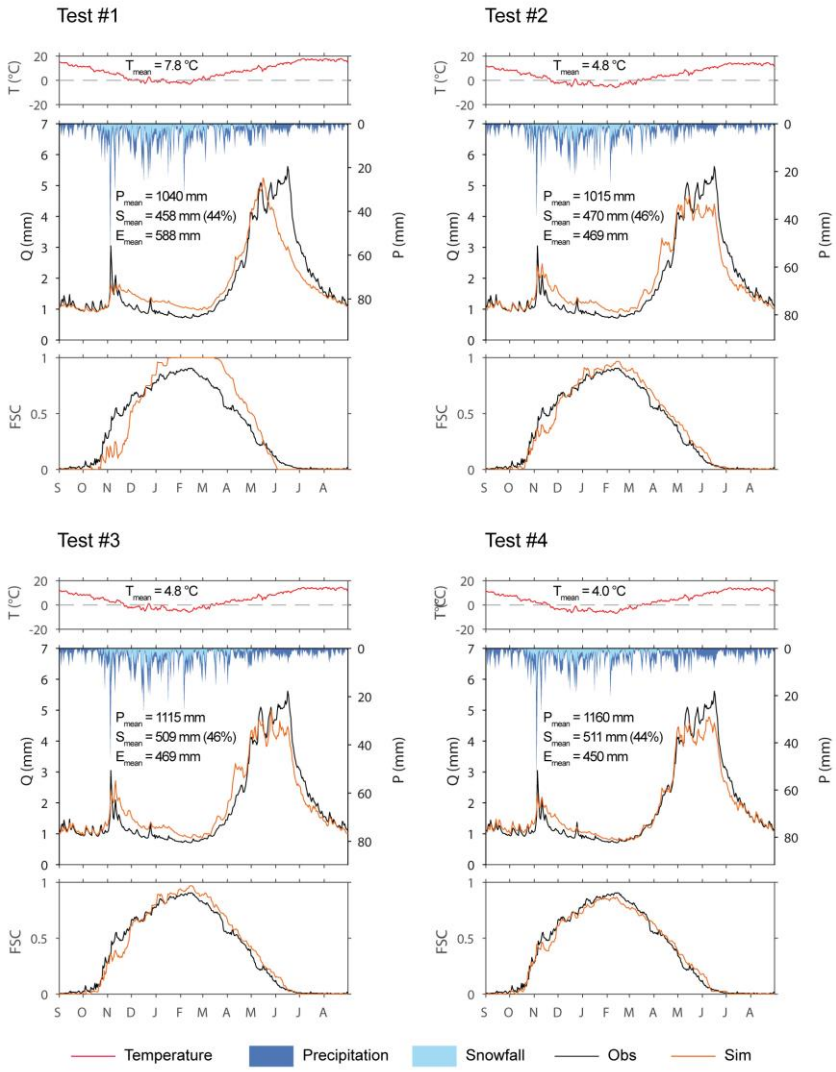


Fig. 7 Comparison of snow-hydrological simulations with elevation dependency according to Tests #1 to #4 (see Table 5) with GR4J for the Durance at Serre-Ponçon. The graphs show mean inter-annual time series of temperature, precipitation, streamflow and fractional snow cover at the catchment scale in validation over the period 2008–2016. T_{mean}, P_{mean} and S_{mean} stand for mean annual temperature, precipitation, and snowfall, respectively. The efficiency criterions NSE_{SNOW}, NSE_{Q}, NSE_{lnQ} and VE_{C} are computed from continuous (not mean seasonal) series over 2008–2016.

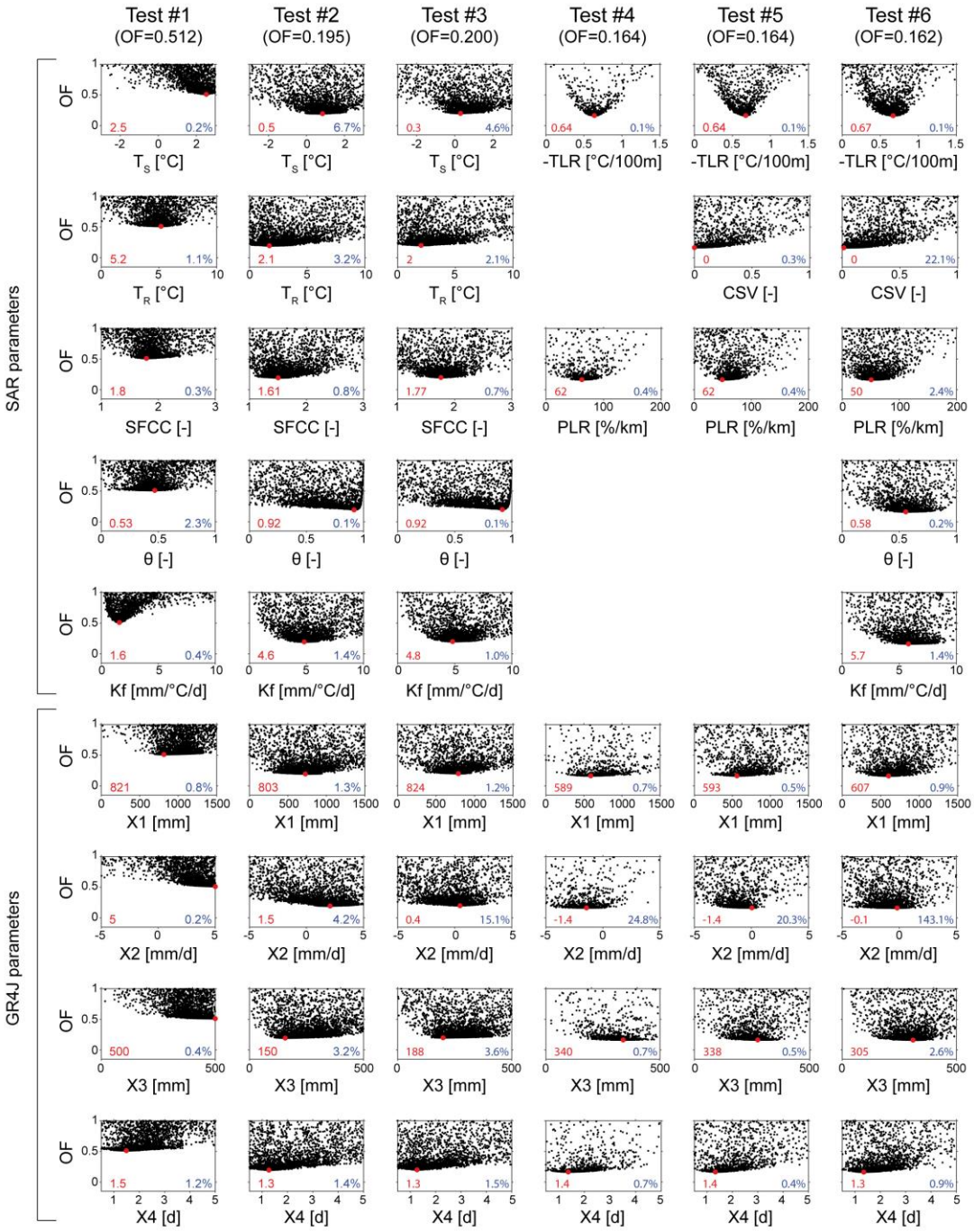


Fig. 8~~8~~ Parameter sensitivity to the objective function (OF) according to Tests #1 to #6 (see Table 5) with GR4J combined with the snow ~~model~~ accounting routine (SAR) on the Durance at Serre-Ponçon. The values and dots in red indicate the optimised calibrated parameters when minimising OF, ~~while~~ the black dots represent trials of the SCE-UA optimisation algorithm, and the values in blue are the variation coefficients (in %) of the 20% best-performing parameter solutions compared to the optimised values for each parameter (the lowest value, the easiest parameter identifiability). Note that depending on the tests, the calibrated parameters of the SAR vary from 2 to 5 (see Table 5 and Table 2), while the GR4J hydrological models has 4 free parameters (see Table 3).

5.4. Ranges of the calibrated altitudinal gradients

Figure 9 shows that the temperature and precipitation lapse rates vary considerably from one catchment to another.

The mean value of the calibrated temperature lapse rates is -0.686 °C (~~hm~~100m)$^{-1}$ and -0.652 °C (~~hm~~100m)$^{-1}$ with GR4J and HBV9, respectively. These values are higher than the yearly lapse rates identified by Rolland (2003) from gauge observations in Alpine regions, which ranged from -0.54 to -0.58 °C (100-m)$^{-1}$ in the Italian and Austrian Tyrol. Instead, the mean calibrated values are close to the average temperature gradients generally proposed as approximations in the literature (~~-0.60 °C in Dodson and Marks, 1997; -~~0.65 °C (~~hm~~100m)$^{-1}$ in Barry and Chorley, 1987). They can be used as suitable estimates for daily snow-hydrological purposes in the French Alps. However, to better account for local meteorological conditions, it may be advisable to calibrate them since the *TLR* parameter ranges from -0.413 to -0.834 °C (~~hm~~100m)$^{-1}$ depending on the catchments and on the models, and is easily identifiable (see Fig. 8 and section 5.3.).

The mean value of the calibrated precipitation lapse rates is 340% (km)$^{-1}$ and 294% (km)$^{-1}$ with GR4J and HBV9, respectively. The differences in ranges between the two models may be due to the GR4J ability to gain (or loose) water from inter-catchment groundwater flows through its *X2* parameter (see section 5.3.), unlike HBV9 which considers the catchment as a closed system. On the other hand, HBV9 relies on more parameters for production and transfer, thus enabling to compensate differently for the errors in the precipitation volumes. Whatever the model, the calibrated lapse rates indicate the need for increased precipitation volumes in most catchments, either to counterbalance for erroneous measurements such as the systematic errors associated with precipitation under-catch during snowfall, or to consider the orographic effect that cannot be sufficiently accounted for by the gauges used for interpolating the precipitation fields. However the ranges of the precipitation lapse rates, from 0 to 10099% (km)$^{-1}$ with GR4J and from 0 to 7982% (km)$^{-1}$ with HBV9, suggest that the required correction is catchment-specific and depends either on the local meteorological conditions or on data from the available surrounding stations to interpolate the daily precipitation.

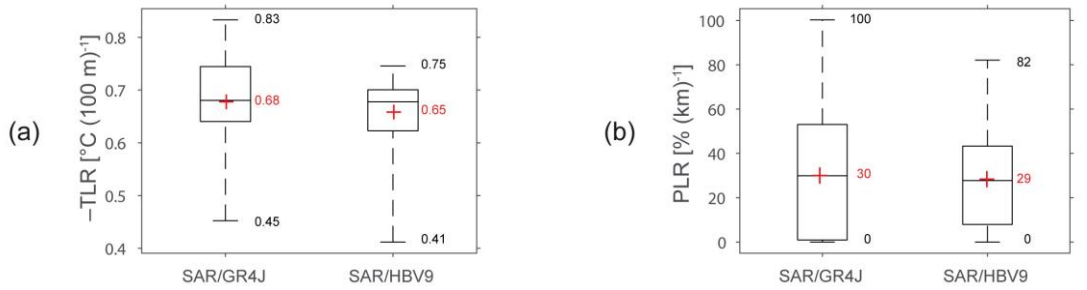


Fig. 99 Boxplots (showing 0.00, 0.25, 0.50, 0.75 and 1.00 percentiles) of the ranges of (a) temperature and (b) precipitation lapse rates calibrated with the 2-parameter SAR (Test #4) in association with the GR4J and HBV9 models on the 20 snow-affected Alpine catchments. The red crosses indicate mean values.

6. SUMMARY AND CONCLUSIONS

6.1. Synthesis~~ummary~~

~~Elevation is a key factor in spatial climate variability in mountainous areas, for both temperature and precipitation. However, it is still difficult to establish the elevation dependency locally at a daily time scale from scattered observations.~~ In this paper, several alternative approaches for distributing daily temperature and precipitation are compared in the French Alps. ~~The aim of this paper was thus to assess whether snow-covered area and streamflow measurements can help assess altitudinal gradients of temperature and precipitation more realistically than using standard interpolation procedures in data scarce mountainous areas. To investigate this issue, we used an extensive dataset based on 78 temperature gauges, 148 precipitation gauges, 500 m MODIS gap-filled snow products and 20 streamflow gauges covering the period 2000 to 2016 in the French~~

~~Alps~~. Elevation dependency in the temperature and precipitation fields was accounted for using two ~~approaches~~main strategies: (1) by estimating the local and time-varying altitudinal gradients from the available gauge network based on deterministic (inverse distance weighted) and geostatistical (kriging) ~~interpolation~~ methods with external drift; and (2) by calibrating the local gradients using an inverse snow-hydrological modelling ~~framework~~procedure. ~~In the second approach, we assumed a simple two-parameter correction model to regionalise air temperature and precipitation from the sparse meteorological network: the first parameter (*TLR*) targeted the temperature-elevation relationship, while the second parameter (*PLR*) targeted the precipitation-elevation relationship. The coherence of the two approaches was evaluated by benchmarking several hydrological variables of interest (snow-covered area, streamflow and water balance) computed with snow hydrological models fed by the interpolated datasets and applied in a modelling framework against available measurements. The advantage of this approach is that it integrates the complex benchmarking process of observational meteorological datasets in an easy-to-grasp metric (Laiti et al., 2018).~~

Cross-validation of the mapping methods showed that, whatever the time scale, temperature estimates can clearly benefit from taking altitude into account with interpolation methods based on external drift. For precipitation, incorporating elevation in the interpolation methods was helpful for yearly and monthly accumulation times but could not achieve an improvement for daily time resolution. Results also showed that accounting for elevation dependency from gauge networks when interpolating air temperature and precipitation was not sufficient to provide accurate inputs for the snow-hydrological models tested here. The lack of high-elevation stations seriously limited correct estimation of local, time-varying lapse rates of temperature and precipitation, which, in turn, affected the performance of the snow-hydrological simulations due to too imprecise estimates of temperatures and of precipitation volumes. Conversely, optimizing lapse rates as part of a snow-hydrological modelling procedure ~~retrieving the local altitudinal gradients *via* an inverse modelling approach~~ provided evidence for increased accuracy in the simulation of snow cover and discharge dynamics ~~(including discharge volumes)~~, while limiting the problems of over-calibration and equifinality through parsimonious parametrisation.

6.2. Recommendations

These results suggest that interpolation methods using elevation as external drift such as those tested (KED and IED) should be used with caution in the absence of sufficient high-elevation data. Although the gauge density in the French Alps is close to the minimum density recommended by WMO (2008) for mountainous areas, the number of weather stations is insufficient for a complete cover of the altitude ranges. This seriously limits estimates of local and seasonal relations with elevation, notably for daily precipitation, but also for temperature, which was initially not apparent when using the leave-one-out procedure against available gauges. Placing meteorological fields in a snow-hydrological perspective thus proved indispensable to confirm the limited suitability of standard interpolation methods for generating reliable spatially distributed modelling inputs in mountainous areas. It also made it possible to propose a modelling approach to ~~infer~~ correct meteorological inputs in complex, mountainous environments and showed that it is possible (and even advisable) to use remotely-sensed snow-cover and streamflow measurements to improve our knowledge of temperature and precipitation inputs in data-scarce mountainous regions. Using auxiliary observations of snow cover notably proved to be useful to give additional insights into the reliability of the modelled snow processes.

However the differences in the two compared approaches are worth discussing. The first is to regionalize temperature and precipitation based on in-situ data and various interpolation/extrapolation schemes based on IDW or Kriging; the second is to "embed" part of the distribution process into the snow-hydrologic models via calibrated lapse rates correcting a first-guess distribution based on IDW. While the first approach is independent from hydrological data like fractional snow cover and streamflow, the second does take advantage of these data to adjust some of the distribution parameters. The second strategy prove superior to the first, especially since calibrating distribution parameters rather than adjusting snow parameters

allowed the models to significantly improve their performance. ~~We thus suggest using the proposed modelling framework to infer local altitudinal gradients from a sparse network of gauges based on key parameters in the snow-hydrological models. More generally, following Tobin et al. (2011), we also recommend using such a framework for a preliminary assessment of the hydrological coherence of gridded datasets to be used in large-scale hydro-climatic studies.~~This improvement was however assessed based on the same hydrological variables that were used to calibrate the snow-hydrologic models, rather than on independent measurements of temperature and precipitation. This left wondering if improving hydrologic predictions by calibrating the local gradients using an inverse snow-hydrological modelling framework also improves actual temperature and precipitation estimates. In principle, one would expect the obtained altitudinal gradients to be both more effective in terms of hydrologic predictions and in terms of temperature and precipitation, but the improvement obtained by "embedding" part of the distribution process into the snow-hydrologic models was only quantified in terms of modelling skills. An improved fit for hydrologic variables may not automatically mean that the model is also better representing weather patterns of temperature and precipitation. A good example is that the optimized lapse rates (Fig. 9) can locally be different between the two hydrologic models considered. Since independent data of temperature and precipitation at high elevations are not available, it was not possible to clarify the extent to which these results apply to temperature and precipitation in addition to hydrologic variables. As pointed out by Dettinger (2014), we are largely blind to what is happening in these high-altitude regions.

Another ~~recommendation~~ key-issue concerns the level of complexity required to control snow accumulation and melt. Most degree-day snow models in the literature use free parameters to adjust snowpack processes and streamflow responses, including the whole water balance. Some parameters (temperature thresholds for snow/rain partition and snowmelt, solid precipitation correction factor) aim to compensate for the errors in the *T* and *P* inputs, while others (thermal state of the snowpack, degree-day melt factor) aim to fit snowmelt to local conditions. Our results showed that calibrating these parameters based on a 5-parameter SAR was much less efficient in improving the modelling performance than fixing them and calibrating only local temperature and precipitation altitudinal gradients based on a simple 2-parameter SAR. These results ~~suggest~~ show that altitudinal gradients of temperature and precipitation inputs should be inferred from key parameters in snow-hydrological models since they play a first-order role in snow-hydrological simulations. Accurate estimate of these parameters greatly helps in determining the form of precipitation and spatial distribution of temperature and precipitation, and are critical for snow cover and runoff modelling in high mountain catchments, as already reported in other regions (Zhang et al., 2013; Naseer et al., 2019). ~~Instead of compensating for the errors in the meteorological inputs, i~~Inferring the gradients reduces the input errors originating from the non-representative vertical distribution of stations while allowing the parameters of snow accumulation and melt to be set at general or physical values (see Table 2 in Section 4.1). ~~Indeed, i~~Introducing additional free parameters to account for the seasonal variability of the temperature gradients and for adjustment of snowmelt led to only limited improvements in the performance distributions compared to the simulations based on the parsimonious 2-parameter SAR. This finding suggests that correcting errors in the model inputs is more critical than adapting the SAR to local snow processes. It also suggests limiting the degree of freedom allowed in degree-day snow models in order to reduce the risk of over-parametrisation.

Mis en forme : Non Surlignage

6.3. Prospects

It would be instructive to further explore the sensitivity of snow-hydrological simulations to seasonal variations in the lapse rates, e.g. by using daily altitudinal gradients instead of a uniform constant gradient at the basin scale. However, as shown in the present paper, establishing the relationship between temperature/precipitation and elevation at the daily time scale from a sparse network of gauges is challenging in mountainous regions. For temperature, the methods tested for computing local and daily lapse rates for each prediction point (KED, IED) outperformed the methods that did not account for altitudinal

gradients (IDW, ORK) in the leave-one-out procedure. ~~However~~On the other hand, using only a constant lapse rate calibrated from the inverse modelling approach performed substantially better as regards to the snow-hydrological predictions than using the interpolated datasets of temperature with external drift. This shows either that the local temperature lapse rates (including their seasonal variation) were not correctly captured by the daily application of interpolation methods with external drift, or that, for our experiment, accurately estimating a constant, uniform gradient for temperature was more important than estimating its seasonal variations for hydrological simulations. A seasonal variation in temperature gradient was also tested with a sinusoidal approach, which required an additional free parameter to determine the variation interval. When we compared the modified 3-parameter SAR version (Test #5 in Table 5) with the 2-parameter SAR ~~on the whole dataset~~, we found that the snow-hydrological performance distributions of the two SARs were very similar (Fig. ~~6~~5). This means that, although the seasonal variation in the temperature altitudinal gradient can be put in evidence from gauge networks, as shown by Rolland (2003) for alpine regions, it did not appear indispensable for the daily snow-hydrological processes represented in our modelling experiment. Alternatively, improving the snow-hydrological simulations could consist in using minimum and maximum air temperature rather than daily mean temperature (see e.g. Turcotte et al., 2007) in order to better determine the snow/rain partition. For the regionalisation of these extreme temperatures, one challenge that remains will be characterising the high variability of daily lapse rates, which reflects temperature inversions as well as rapidly changing circulation patterns, as reported in Stahl et al. (2006). ~~The same problem applies to precipitation since the seasonal relationship between precipitation and elevation also depends on exposure to atmospheric flows.~~The problem is even more challenging for precipitation whose lapse rates could not be related to seasonal or other types of systematic variations as they are strongly dependent on the synoptic meteorological conditions and therefore highly variable. Further research could thus build on the works of Jarvis and Stuart (2001) for temperature and Gottardi et al. (2012) for precipitation and focus on methods for interpolation and extrapolation that are capable of accounting for differences in the influence of topography in different seasons and synoptic situations.

Finally, it is worth mentioning that spatial variability was only considered along five elevation bands in each catchment since preliminary tests showed no improvement in the hydrologic predictions when applying the SAR in a full distribution mode. However, the SAR was not designed to explicitly account for topographic effects (slope, aspect and shading) on snow accumulation, redistribution and melt (see e.g. Frey and Holzmann, 2015). For instance, a grid-based temperature-index model could be implemented to include potential clear-sky direct solar radiation at the surface, thus considering both the seasonal variations of melt rates and the geometric effects on melt attributable to terrain (see e.g. Hock, 1999). It would thus be interesting to assess whether accounting for the influence of such effects can further improve the daily hydrologic predictions at the basin scale.

Data availability. The hydro-meteorological data and MODIS snow products used in this study are available via the respective websites of the dataset producers: Météo-France (https://publitheque.meteo.fr), Banque Hydro (http://www.hydro.eaufrance.fr), and NASA's National Snow and Ice Data Center (NSIDC, https://nsidc.org) Distributed Active Archive Center (DAAC).

Author contribution. Denis Ruelland conceived the study, performed the analysis and wrote the paper.

Competing interest. The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements The author is very grateful to *Météo-France* (https://publitheque.meteo.fr) and *Banque Hydro* (http://www.hydro.eaufrance.fr) for providing the necessary public hydro-meteorological data for the study. NASA's National Snow and Ice Data Center (NSIDC) Distributed Active Archive Center (DAAC) is also acknowledged for providing MODIS snow products (https://nsidc.org), which were used for model calibration and validation. Finally, the

author is sincerely grateful to the two anonymous reviewers for they careful reading of the original manuscript and their insightful comments and constructive suggestions for improvements.

REFERENCES

Ahmed, S., and de Marsily, G.: Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, Water Res. Research, 23, 1717–1737, https://doi.org/10.1029/WR023i009p01717, 1987.

Andréassian, V., and Perrin, C.: On the ambiguous interpretation of the Turc-Budyko nondimensional graph, Water Res. Research, 48, W10601, https://doi.org/10.1029/2012WR012532, 2012.

Bárdossy, A., and Pegram, G.: Interpolation of precipitation under topographic influence at different time scales, Water Res. Research, 49, 4545–4565, https://doi.org/10.1002/wrcr.20307, 2013

Barry, R. G., and Chorley, R. J.: Atmosphere, Weather and Climate. 5th ed. Routledge, 448 pp., 1987.

Beck, H., van Dijk, A. I. J. M., de Roo, A., Miralles, D. G. McVicar, T. R., Schellekens, J., and Bruijnzeel, L. A.: Global-scale regionalization of hydrologic model parameters, Water Res. Research, https://doi.org/10.1002/2015WR018247, 2016.

Bergström, S.: Development of a snow routine for the HBV-2 model, Nordic Hydrol., 6, 73–92, https://doi.org/10.2166/nh.1975.0006, 1975.

Berndt, C., and Haberlandt, U.: Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density, J. Hydrol. Regional Studies, 15, 184–202, https://doi.org/10.1016/j.ejrh.2018.02.002, 2018.

Dettinger, M.: Impacts in the third dimension. Nat. Geosci., 7, 166–167, doi:10.1038/ngeo2096, 2014.

Deutsch, C. V.: Correcting for negative weights in ordinary kriging, Comput. Geosci., 22, 765–773, https://doi.org/10.1016/0098-3004(96)00005-2, 1996.

Diggle, P. J., and Ribeiro, P. J.: Model-Based Geostatistics, Springer Series in Statistics, Springer, https://doi.org/10.1007/978-0-387-48536-2, 2007.

Dodson, J., and Marks, D.: Daily air temperature interpolated at high spatial resolution over a large mountainous region, Climate Res. 8, 1–20, https://doi.org/10.3354/cr008001, 1997.

Douguédroit, A., and de Saintignon, M. F. : Les gradients de température et de précipitation en montagne, Rev. Geogr. Alp., 72, 225–240, doi :10.3406/rga.1984.2566, 1984.

Drogue, G., Humbert, J., Deraisme, J., Mahr, N., and Freslon, N. : A statistical topographic model using an omnidirectional parameterization of the relief for mapping orographic rainfall, Int. J. Climatol. 22, 599–613, https://doi.org/10.1002/joc.671, 2002.

Duan, Q. Y., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Res. Research, 28, 1015–1031, https://doi.org/10.1029/91WR02985, 1992.

Duan, Q., Sorooshian, S., and Gupta, V.: Optimal use of the SCE-UA global optimization method for calibrating watershed models, J. Hydrol., 158, 265–284. https://doi.org/10.1016/0022-1694(94)90057-4, 1994.

Kuczera, G.: Efficient subspace probabilistic parameter optimization for catchment models, Water Res. Research, 33, 177–185, https://doi.org/10.1029/96WR02671, 1997.

Farr, T. G., Rosen, P. A. , Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The shuttle radar topography mission. Rev. Geophys., 45, RG2004, https://doi.org/10.1029/2005RG000183, 2007.

Franz, K. J., and Karsten, L. R.: Calibration of a distributed snow model using MODIS snow covered area data, J. Hydrol., 494, 160–175, https://doi.org/10.1016/j.jhydrol.2013.04.026, 2013.

Frei, C., and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatology, 18, 873–900, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Frei, C.: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, Int. J. Climatolotology, 34, 1585–1605, https://doi.org/10.1002/joc.3786, 2014.

Mis en forme : Interligne : Multiple 0.85 li, Autoriser lignes veuves et orphelines

Frey, S. and Holzmann H.: A conceptual, distributed snow redistribution model, Hydrol. Earth Syst. Sci., 19, 4517–4530, https://doi.org/10.5194/hess-19-4517-2015, 2015.

Gafurov, A. and Bárdossy, A.: Cloud removal methodology from MODIS snow cover product, Hydrol. Earth Syst. Sci., 13, 1361–1373, https://doi.org/10.5194/hess-13-1361-2009, 2009.

Garavaglia, F., Le Lay, M., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E., and Mathevet, T.: Impact of model structure on flow simulation and hydrological realism from lumped to semi-distributed approach, Hydrol. Earth Syst. Sci., 21, 3937–3952, https://doi.org/10.5194/hess-21-3937-2017, 2017.

Gascoin, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J.-F., Szczypta, C., Marti, R., and Sánchez, R.: A snow cover climatology for the Pyrenees from MODIS snow products, Hydrol. Earth Syst. Sci., 19, 2337–2351, https://doi.org/10.5194/hess-19-2337-2015, 2015.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113–129, https://doi.org/10.1016/S0022-1694(00)00144-X., 2000.

Gottardi, F., Obled, C., Gailhard, J., and Paquet, E.: Statistical reanalysis of precipitation fields based on ground network data and weather patterns: Application over French mountains, J. Hydrol., 432–433, 154–167, https://doi.org/10.1016/j.jhydrol.2012.02.014., 2012.

Hall, D., Riggs, G., and Salomonson, V.: MODIS/Terra Snow Cover Daily L3 Global 500m Grid V005, National Snow and Ice Data Center, Boulder, Colorado, USA., 2006.

Hall, D., Riggs, G., and Salomonson, V.: MODIS/Aqua Snow Cover Daily L3 Global 500m Grid V005, National Snow and Ice Data Center, Boulder, Colorado, USA., 2007.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded dataset of surface temperature and precipitation, J. Geophysical Research, 113, D20119, https://doi.org/10.1029/2008JD10201, 2008.

He, Z. H., Parajka, J., Tian, F. Q., and Blöschl, G.: Estimating degree-day factors from MODIS for snowmelt runoff modeling, Hydrol. Earth Syst. Sci., 18, 4773–4789, https://doi.org/10.5194/hess-18-4773-2014, 2014.

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation. J. Glaciology 45, 101–111, https://doi.org/10.3189/S0022143000003087, 1999.

Hock, R.: Temperature index melt modelling in mountain areas, J. Hydrol., 282, 104–115, https://doi.org/10.1016/S0022-1694(03)00257-9, 2003.

Hofstra, N., New, M., and McSweeney, C.: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data, Climate Dynamics, 35, 841–858, https://doi.org/10.1007/s00382-009-0698-1, 2010.

Hublart, P., Ruelland, D., Dezetter, A., and Jourde, H.: Reducing structural uncertainty in conceptual hydrological modeling in the semi-arid Andes, Hydrol. Earth Syst. Sci., 19, 2295–2314, https://doi.org/10.5194/hess-19-2295-2015, 2015.

Hublart, P., Ruelland, D., Garcia de Cortázar-Atauri, I., Gascoin, S., Lhermitte, S., and Ibacache, A.: Reliability of lumped hydrological modelling in a semi-arid mountainous catchment facing water-use changes, Hydrol. Earth Syst. Sci., 20, 3691–3717, https://doi.org/10.5194/hess-20-3691-2016, 2016.

Isotta, F. A., Frei, C., Weilguni, V., Percec Tadic, M., Lassègues, P., Rudolf, B. et al.: The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatology, 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

~~Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., Nikulin, G., and Majone, B.: Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets, Water Res. Research, 54, https://doi.org/10.1002/2017WR021633, 2018.~~

Le Moine, N., Andréassian, V., Perrin, C., and Michel, C.: How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, Water Res. Research, 43, W06428, https://doi.org/10.1029/2006WR005608, 2007.

Le Moine, N., Hendrickx, F., Gailhard, J., Garçon, R., and Gottardi, F.: Hydrologically aided interpolation of daily precipitation and temperature fields in a mesoscale Alpine catchment, J. Hydrometeorol, 16, 2595–2618, https://doi.org/10.1175/JHM-D-14-0162.1, 2015.

Mis en forme : Interligne : Multiple 0.85 li, Autoriser lignes veuves et orphelines

Ly, S., Charles, C., and Degré, A.: Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium, Hydrol. Earth Syst. Sci., 15, 2259–2274, https://doi.org/10.5194/hess-15-2259-2011, 2011.

Ly, S., Charles, C., and Degré, A.: Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: A review, Biotechnol. Agron. Soc. Environ. 17, 392–406, http://hdl.handle.net/2268/136084, 2013.

Masson, D., and Frei, C.: Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types, Hydrol. Earth Syst. Sci., 18, 4543–4563, https://doi.org/10.5194/hess-18-4543-2014, 2014.

Naseer, A., Koike, T., Rasmy, M., Ushiyama, T., and Shrestha, M.: Distributed hydrological modeling framework for quantitative and spatial bias correction for rainfall, snowfall, and mixed-phase precipitation using vertical profile of temperature, J. Geophysical Research: Atmospheres, 124, 4985–5009. https://doi.org/10.1029/2018JD029811, 2019.

~~Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models – Part I: A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.~~

Nicótina, L., Alessi Celegon, E., Rinaldo, A., and Marani, M.: On the impact of rainfall patterns on the hydrologic response, Water Res. Research, 44, W12401, https://doi.org/10.1029/2007WR006654, 2008.

NSIDC: World Glacier Inventory, Version 1. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5/NSIDC-WGI-2012-02, 2012.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2: towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, J. Hydrol., 303, 290–306, https://doi.org/10.1016/j.jhydrol.2004.08.025, 2005.

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C., and Michel, C.: Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, Water Res. Research, 42(7), https://doi.org/10.1029/2005WR004636, 2006.

Parajka, J., and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, J. Hydrol., 358, 240–258, https://doi.org/10.1016/j.jhydrol.2008.06.006, 2008.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, J. Hydrol., 279, 275–289, https://doi.org/10.1016/S0022-1694(03)00225-7, 2003.

Riboust, P., Thirel, G., Le Moine, N., and Ribstein, P.: Revisiting a simple degree-day model for integrating satellite data: implementation of SWE-SCA hystereses, J. Hydrol. Hydromech., 67, 70–81, https://doi.org/10.2478/johh-2018-0004, 2019.

Rolland, C.: Spatial and seasonal variations of air temperature lapse rates in alpine regions, J. Climate, 16, 1032–1046, https://doi.org/10.1175/1520-0442, 2003.

Sevruk, B.: Regional dependency of precipitation–altitude relationship in the Swiss Alps, Clim. Change 36, 355–369, https://doi.org/10.1023/A:1005302626066, 1997.

Sevruk, B.: Hydrometeorology: rainfall measurement, gauges. In: Anderson, M. G. (Ed.), Encyclopedia of Hydrological Sciences, Vol. 1. Wiley&Sons Ltd., Chichester, UK, pp. 529–535. Ch. 40., 2005.

Shen, S. S. P., Dzikowski, P., Li, G. L., and Griffith, D.: Interpolation of 1961–1997 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada, J. Appl. Meteorol. 40, 2162–2177, https://doi.org/10.1175/1520-0450(2001)040<2162:IODTAP>2.0.CO;2, 2001.

Sinclair, M., Wratt, D., Henderson, R. and Gray, W.: Factors affecting the distribution and spillover of precipitation in the Southern Alps of New Zealand – a case study, J. Appl. Meteorol. 36, 428–442, https://doi.org/10.1175/1520-0450(1997)036<0428:FATDAS>2.0.CO;2, 1997.

Spadavecchia, L., and Williams, M.: Can spatio-temporal geostatistical methods improve high resolution regionalisation of meteorological variables? Agr. Forest Meteorol., 149, 1105–1117, https://doi.org/10.1016/j.agrformet.2009.01.008, 2009.

Stahl, K., Moore, R. D., Floyer, J. A., Asplin, M. G., and McKendry, I. G.: Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density, Agricultural and Forest Meteorology, 139, 224–236, https://doi.org/10.1016/j.agrformet.2006.07.004, 2006.

Thirel, G., Salamon, P., Burek, P., and Kalas, M.: Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter, Remote Sens., 5, 5825–5850, https://doi.org/10.3390/rs5115825, 2013.

Tobin, C, Nicótina, L., Parlange, M. B., Berne, A., and Rinaldo, A.: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region, J. Hydrol., 401, 77–89, https://doi.org/10.1016/j.jhydrol.2011.02.010, 2011.

USACE: Snow Hydrology: Summary Report of the Snow Investigation. Portland, Oregon, North Pacific Division, Corps of Engineers, U.S. Army, 1956.

Valéry, A., Andréassian, V., and Perrin, C.: Inverting the hydrological cycle: when streamflow measurements help assess altitudinal precipitation gradients in mountain areas. In: New approaches to hydrological prediction in data-sparse regions (Proc. of Symposium HS.2 at the Joint IAHS & IAH Convention, Hyderabad, India, September 2009), IAHS Publ., 333 281–286, 2009.

Valéry, A., Andréassian, V., and Perrin, C.: Regionalization of precipitation and air temperature over high-altitude catchments - learning from outliers, Hydrol. Sci. J., 55, 928–940, https://doi.org/10.1080/02626667.2010.504676, 2010.

Valéry, A., Andréassian, V., and Perrin, C.: As simple as possible but not simpler: What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments, J. Hydrol., 517, 1176–1187, https://doi.org10.1016/j.jhydrol.2014.04.058, 2014.

WMO: Guide to hydrological practices. Volume I: Hydrology – From measurement to hydrological information, 6th edition. World Meteorological Organization, Geneva, 2008.

Zhang, F., Zhang, H., Hagen, S. C., Ye, M., Wang, D., Gui, D., Zeng, C., Tian, L., and Liu, J.: Snow cover and runoff modelling in a high mountain catchment with scarce data: Effects of temperature and precipitation parameters, Hydrol. Proc., 29, 52–65, doi:10.1002/hyp.10125, 2015.