Interactive comment on "The Latest Improvements in SURFEX v8.0 of the Safran-Isba-Modcou Hydrometeorological Model over France" *by* Patrick Le Moigne et al.

Anonymous Referee #2

Received and published: 10 April 2020

1 - The colour black is used for the reviewer questions/remarks.
2 - The colour green is used to answer questions raised by the reviewer.
3 - The colour blue is used for the text added to the manuscript.

Summary:

This paper describes the new developments in the Safran-Isba-Modcou (SIM) hydrometeorological model as part of SURFEX v8.0. These new developments cover different aspects of the SIM system: atmospheric forcing, climate fields, land-surface model parameterizations and water budget parameterizations. The authors evaluate the impact of the new developments in an incremental way using different types of observations. They show that the new SIM system, considering all model changes, improves the simulation of daily river discharges and snow depth over a set of catchments and sites covering the French region, compared to the previous SIM system.

The discussion of the results is clear in most of the parts and conclusions are overall justified by the results shown. However, in my opinion the authors do not fully discuss the interactions among the different model changes, therefore explaining the physical mechanisms of some of the results. Also, I found the section on the soil temperature evaluation weak compared to the others. Finally, I have few comments on the introduction and model description parts: these sections are a bit difficult to read and can be improved. All these points are discussed in the main comments below. In summary, I would recommend the acceptance of the paper after major revisions, to make the paper stronger and more attractive to readers.

Main comments

1 - Introduction and model description: Sect. 1 and Sect. 2 are quite difficult to read. Sentences are not clear in some places, making it difficult to understand the message that the authors want to deliver. I would suggest the authors to improve the readability of these sections. Few examples are reported in the minor comments below.

2 - The manuscript has been revised to improve the English (request of the first referee), therefore the readability of all the sections of the manuscript has been improved. More comments can be found in the answers to the minor revisions.

1 - Discussion of the results: The authors performed a set of sensitivity experiments to extract the effect of each model change. However, I have the impression that they do not fully discuss the interactions among the different model changes, and how these relate to the results of the evaluation. For instance, why the SIM_PHY simulation deteriorate the scores in many of the presented metrics (Fig.8, Fig.9, Fig.11)? Is it because of errors in other components, like the atmospheric forcing, which then penalise a more physically complex model? If that is the case, are the other changes reducing such errors, therefore allowing to fully exploit the benefit from the new soil/snow parameterizations? Or are these unrelated? Another example is in the evaluation of the snow depth: the authors propose several hypothesis to explain the improvement of the simulation of the snow depth in SIM_NEW. However, the comparison of all experiments would clearly quantify which processes/changes are responsible for the improvements (see also one of the minor comments below).

2 - Thank you for the interesting question. In fact, what we see in Fig. 8 for instance is that SIM_REF is simulating the correct modelled to observed river flow (centred around 1) whereas in SIM_PHY this ratio is overestimated. However, in SIM_PHY, as explained in the description of the model, more complexity was added to the model based on a better representation of physics. Moreover, errors in the forcing data show that errors are compensating in SIM_REF. In SIM_PHY, the calculations performed on each of the vegetation type are using the A-gs parameterization of photosynthesis which tends to produce less evaporation over vegetation, leading to more water available in the rivers. On the other hand, it has already been mentioned that the radiation forcing is biased low. The combination of more water available in the soil and less radiative energy to evaporate it leads to an overestimation of the river flows. The more physically based SIM_PHY model is penalised by errors in the forcing. When correcting the IR radiation, SIM_FRC simulation exhibits a large improvement in the scores of river flows, with a modelled to observed ratio peak closer to 1, and a daily efficiency range improved in almost all cases, except perhaps for NSEs lower than 0.4, but the difference to SIM_REF is very small. The implementation of the subgrid topography with the use of elevation bands and of the subgrid hydrology with including a reservoir for snow are essentially impacting hydrology in mountains, and therefore snow and river flows that may be affected by snow melt.

The following paragraph has been added (Section 5.2 of the manuscript):

3 - The results show that SIM_REF simulates the correct ratio between modelled and observed river flow (centred around 1) whereas in SIM_PHY, this ratio is overestimated. However, in SIM_PHY, as explained in the model description, more complexity has been added to the model based on a better representation of physics. In addition, errors in the forcing data show that errors compensate for each other in SIM_REF, since despite a radiative deficit, river flow is rather well simulated. In SIM_PHY, the calculations performed on each of the vegetation types use the A-gs photosynthesis parameterization, which tends to produce less evaporation on the vegetation, leading to more water available in the rivers. On the other hand, it has already been mentioned that radiative forcing is underestimated. The combination of more water available in the soil and less radiative energy to evaporate it leads to an overestimation of river flows. This is the case of the more physical SIM_PHY model, which is more penalized by errors in forcing. By

correcting for IR radiation, the SIM_FRC simulation shows a clear improvement in river flow scores, with a peak of the modelled to observed ratio closer to 1, and an improved daily efficiency range in almost all cases, except perhaps for NSEs below 0.4, but in this case the difference with SIM_REF is very small. The implementation of the subgrid topography with the use of elevation bands (SIM_TOP) and the subgrid hydrology with the inclusion of a snow reservoir (SIM_NEW) essentially impacts the hydrology in the mountains, and thus the snow and river flows that are affected by snowmelt.

2 - For the evaluation of the snow depth, the comparison can only be made on the 9892 cells, which corresponds to the SIM_REF grid. The snow depth in SIM_FRC, SIM_TOP and SIM_NEW is the same because these experiments use the same IR correction and sub-grid processes related to topography or hydrology are not considered in the evaluation. In terms of snow depth, only SIM_PHY would be different from the other three simulations where the IR correction is applied below 1340m, which limits the interest of such a comparison.

The following paragraph has been added to section 5.3 of the manuscript:

3 - For the evaluation of the snow depth, the comparison can only be made on the 9892 cells, which corresponds to the SIM_REF grid. The snow depth in SIM_FRC, SIM_TOP and SIM_NEW is the same because these experiments use the same IR correction and sub-grid processes related to topography or hydrology which are not considered in the evaluation. In terms of snow depth, only SIM_PHY would be different from the other three simulations if the IR correction were to be applied below 1340 m, which limits the interest of such a comparison.

1 - Soil temperature evaluation: I found Sect. 4.6 on the analysis of soil temperature profiles rather weak and with not enough details. The scores of the new system, without a reference, cannot be put into the context of the paper and so are not adding valuable information to the results. A comparison between the different simulations would clarify at least the impact of each change on the bias. The authors state that such biases can be associated to incoming shortwave radiation or lack of geothermal heating, but what about the soil parameterization or soil/surface properties?

2 - This remark is very relevant, and that is true that the section on the soil temperature results looks too weak. As for the snow depth comparison, the ground temperature in SIM_FRC, SIM_TOP, and SIM_NEW is the same. Only SIM_PHY is different, and exhibits larger biases and rmses at all depths. Such biases can be of course also be associated to soil parameterization and soil and surface properties. It was decided to remove the section 4.6 and to add a discussion on the ground temperature results. Also, the table 3 and the figure showing the ground temperature observation network, as well as the description of the soil observations were removed.

The following paragraph has been added to section 5.1 of the manuscript (lines 471-473):

3 - We also compared the simulated soil temperatures to the observations made over France. The IR correction on soil temperature has a positive impact and significantly reduces biases and RMSEs (not shown). The results are consistent and of the same order of magnitude as those obtained by Decharme et al (2013).

Minor comments

1 - Abstract: The main scope of the paper is placed at the very end of the abstract. This could be placed earlier in the text to make clearer the main message of the paper.

2 - The abstract has been amended and the last sentence has been placed at the beginning:

3 - This paper describes the impact of the various changes made on the Safran-Isba-Modcou hydrometeorological system (SIM), and demonstrates that the new version of the model performs better than the previous one by making comparisons with observations of daily river flows and snow depths.

1 - Ln-35–40: This paragraph should be rephrased and clarified. Also references to previously published work on the evaluation of land surface and hydrological models should be introduced in the text.

2 - The introduction has been revised (following the first referee's remark) and the remark raised here has been dealt with.

1 - Ln.42: What do you mean by "independent" variable? As the authors stated few lines before, surface energy and water budgets form a coupled system. Please clarify/reformulate.

2 - The sentence was reformulated (line 72 of the revised manuscript) into:

3 - In addition, climate models have been evaluated at both global and regional scales through hydrology. Indeed, the coupling between their land surface model and hydrology allows a quantitative assessment to be made, through comparisons to variables such as river flow, groundwater levels and snow depth.

1 - Ln. 49-50: "... , where modelling contribution of SVAT . . . accounted for in models." this sentence is not clear, please reformulate.

2 - This paragraph was reformulated (lines 80-83) as:

3 - Recent initiatives to study the impact of anthropization on water availability, such as those supported by the Global Energy and Water Exchanges (GEWEX) project (Harding et al., 2015), where the contribution of LSMs to modelling appears to be important, show that irrigation needs to be considered in the models (Boone et al., 2019).

1 - Ln.65-73: The scope of the work can be improved, to make it more precise and easier to read. For instance the authors talk about “new parameterizations” at line 66, but changes to the atmospheric forcing (Ln. 68) are not a parameterization. Also at Ln.69 they talk about “these results”, but it is unclear which results are referring to at this point.

2 - This paragraph was reformulated (lines 97-103) as:

3 - The objective of this paper is to show how the development of new parameterizations and better atmospheric forcing prescription have improved the performance of the system. The current study, based on numerical simulations covering the period 1958-2018, shows how improvements in atmospheric forcing, land surface model physics and subgrid orography and hydrology improve the modelled river flow and snow depth of the SIM system. It also aims to describe how the model results are affected by each change separately and finally to demonstrate that the new model configuration performs better than the previous one in terms of river flow extremes, and when simulated snow depth or average river flow is compared to observed data.

1 - Ln. 83: could you be more specific on the horizontal resolution of the SAFRAN analysis?

2 - The SAFRAN analysis is performed on irregular areas of a few hundred square kilometres. In fact, the size of the area where the analysis is done varies from 400km2 to 1000km2, and then as indicated in the manuscript the horizontal interpolation is done to an 8 km regular grid. In the manuscript a reference to Le Moigne (2002) has been added (line 113) because a description of these areas is given.

3 - The analysis is carried out over geographical areas covering a few hundred square kilometres (Le Moigne, 2002), and the analysed fields are interpolated to hourly time steps.

1 - Ln. 83: is the 24h precipitation analysed every 6-hours?

2 - The analysis of 24h precipitation is performed daily. Manuscript was updated accordingly (line 111-113).

3 - SAFRAN (Durand et al., 1993; Quintana Segui et al., 2008) performs a 6-hourly analysis of near-surface meteorological variables such as temperature and relative humidity at 2 metres, wind speed, cloud cover and a daily analysis of 24-hour accumulated precipitation.

1 - Ln. 154: please clarify in which sense you mean “dry”, as this can be related to various processes.

2 - The term “dry” here refers to the soil moisture. The sentence in the manuscript was changed to (line 174):

3 - However, the Force-Restore scheme is known to be too dry in terms of soil moisture

1 - Ln. 163-167: this paragraph should be simplified/clarified, to make it clearer that the soil map is not changed between the previous SIM system and the new system (as far as I understood). Ln. 166-167: please clarify that this map was the one used in the “old” version of SIM.

2 - The beginning of section 2.3 was modified to account for these 2 remarks (lines 182-189):

3 - In addition to the changes in model physics described above, the land cover and topography databases have been updated to improve the realism of the external parameters of the ISBA model. The hydrogeological database representing the aquifer and the routing network was unchanged. In addition, the soil texture database for France is unchanged. In the former SIM system, the soil texture was based on a soil map provided by the Institut National de Recherches Agronomiques (INRA - King et al., 1995) at a resolution of 1 km. In the new SIM system, texture is defined by the Harmonized World Soil Database (HWSD - Nachtergaele et al., 2012) which is a soil map at 1 km resolution that combines several data sets available worldwide. In particular for France, the INRA soil map mentioned above has been integrated into the HWSD dataset (used in other applications using SURFEX outside France), so this change does not affect the SIM simulations.

1 - Sect. 2.3: as far as I understood, the main change in the climate fields is the update to ECOCLIMAP2. Hence, I would suggest to expand the discussion of this change, for instance explaining the impacts on the ISBA model? See also next comment.

2 - The last paragraph of section 2.3 was modified as follows (lines 201-211) to expand the discussion as required and also indicate the use of a new albedo parameterization.

3 - The impacts of modifying the vegetation fraction input to the ISBA model are multiple and will not be described here in detail (for a detailed comparison, see Faroux et al., 2013). ECOCLIMAP2 has definite advantages, the effects of which are directly reflected in the ISBA model. For example, ECOCLIMAP2 covers a larger time period than the previous version and therefore allows a better representation of the variability of surface parameters. Also, it distinguishes different types of crops that can be modelled separately, and therefore more accurately, with ISBA. The sensors on board satellites have better accuracy and the uncertainty of the measurement is reduced. The vegetation fraction in particular is improved and with it the roughness length of the vegetation which impacts the surface wind by the obstacle effect on near-surface flows. The leaf area index is also improved and its increase leads to a better description of the evaporative fraction, which is key for the energy partitioning in the model. The more realistic surface albedo developed by Carrer et al. (2014) was also used, as Decharme et al. (2013) showed that it improved results at the global scale.

1 - Ln. 187-193: this sentence is very long. Coud you break it in multiple sentences, better explaining the impact of this change?

2 - The answer is contained in the respond to the previous remark.

1 - Sect. 2.4, title: could you specify in the Section title that this is downward infrared radiation?

2 – Done (line 212).

3 - 2.4 Evolution of downward infrared radiation

1 - Ln. 195-198: is the bias related to a problem in the analysis (for instance cloud cover) or a RT model issue?

2 - The bias is likely due to a problem in the analysis and in the RT model. The cloud cover analysis is computed using T and q profiles from a large-scale atmospheric model that contains biases. The model used to solve the RT is an old model, with a rather low vertical resolution and therefore probably sub-optimal, but which was state-of-the-art in the 1990s.

A sentence has been added to the manuscript (line 218):

3 - The bias is likely due to a problem in the analysis and in the radiative transfer (RT) model. The cloud cover analysis is computed using temperature and humidity profiles from a large-scale atmospheric model that contains biases. Moreover, the model used to solve the RT is an old model, with a rather low vertical resolution and therefore probably sub-optimal, but which was state-of-the-art in the 1990s.

1 - Ln. 214: annual mean over which time period?

2 - Annual mean over the 60-years periods. This was added to the manuscript (end section 2.4, lines 230-231).

3 - Figure 2 shows the annual average over the 60-year period initial infrared radiation (left panel) and the amount of energy supplied when the correction is applied (right panel).

1 - Ln. 215-217: could you clarify this sentence? is the analysis done every 300m in the vertical direction?

2 - This is clarified in the manuscript (lines 233-234):

3 - In SAFRAN, the analysis is performed on homogeneous zones of several hundred square kilometres and the vertical component is explicitly considered with to a 300-metre slicing along the vertical.

1 - Ln. 233: why between 3 to 5 layers are necessary, and not more? What it is the vertical discretization between each band? Please clarify.

2 - The original attempt was to have 10 layers of 300 m from ground to 3000 m. However, this solution appeared to be too expensive and a solution based on the distribution of altitudes in

each grid box in 5 classes using deciles q20, q40, q60 and q80 was adopted. The vertical discretization varies from one grid point to another, but is at least equal to 300 m.

The sentence was changed in the manuscript to explain it better (lines 247-248):

3 - Using a vertical discretization of 300 m at each grid point to represent topographic variability was ideal but too costly. A solution based on the distribution of elevations in each grid cell into five bands represented by the quintiles q20, q40, q60 and q80 was adopted. For each of the 1044 grid points, the vertical discretization varies and is at least equal to 300m. In the end this gives a total of 3878 grid points involved in the calculations of the mountain simulation. Figure 1 (right panel) shows the elevation of the 1044 grid points where the elevation band method is applied.

1 - Ln. 264: How does the relatively low horizontal resolution of the ERA-40/ERA-I data impact the simulations? The horizontal resolution of ERA-I is ~80km, that is one order of magnitude less than the one used by the SIM grid. I am thinking for instance at regions with a low coverage of surface stations used in the analysis.

2 - In fact, the SAFRAN analysis does not suffer from a coarse guess as input. SAFRAN tries to be as close as possible to the observations, for temperature, humidity, …, precipitation. The analysis of precipitation even doesn't use large scale information as input. In France, the density of the observation network is very high, because a network dedicated to climatology complements the synoptic network which is less dense. So, there are almost no region with low coverage especially for precipitation which is key for hydrologic purposes.

The manuscript was modified as follows (lines 276-279)

3 - In France, the density of the observation network is very high, because a network dedicated to climatology completes the less-dense synoptic network. There are therefore practically no regions with poor coverage, especially for precipitation, which is essential for hydrology, and the coarse resolution of the analysis first guess is not an issue.

1 - Ln. 299-300: were the data cleaned in some way? For instance removing stations with a few number of observations? Or all data have been used to compute the statistics? The latter could introduce some artifact in the statistics. This should be better explained in the text.

2 - The length of the series is a source of variability in the scores (in particular the number of seasons the stations are open can vary from 1 to 32) but since very few series are complete, it was felt that it was nevertheless more robust to assess the performance of the model to consider as many stations as possible rather than trying to homogenize the length of the series.

The manuscript was modified (lines 308-311) as follows:

3 - The length of the measurement series and the number of seasons that stations are open are sources of variability in the scores. However, since very few series are complete, the choice was made to evaluate the performance of the model by considering as many stations as possible rather than trying to homogenize the length of the series.

1 - Ln.318-320: could you specify clearly in the text when the transition from ERA-40 to ERA-I occur in SAFRAN?

2 - The question raised here pushed the authors to verify in more details which ERA data was used. And it turned out that ERA-40 is used until 2002, then replaced by the operational ECMWF analysis. This means that ERA-I is not used, and the manuscript was corrected (line 275-276) as follows:

3 - In SAFRAN, the guess of the analysis used is ERA-40 until 2002 and the ECMWF operational analysis thereafter.

2 - And (line 336-342) as follows:

3 - In addition to this physical reason, a more technical reason is the change in the large-scale analysis used as boundary conditions to the ERA-40 reanalysis (Uppala et al., 2005), with a priori small changes in the analysed fields. During the production of the ERA-40 reanalysis, the ECMWF operational data assimilation system has evolved considerably and switched to a 4D-var variational method compared to the 3D-var method previously used. This new system has proven to be more accurate and the assimilation of a much larger number of satellite observations has led to a significant improvement in analysis and forecasting, in particular, for the vertical profiles of temperature and relative humidity.

2 - And (line 457-458) as follows:

3 - As explained in section 4.1, the calculations of these profiles have varied over time as a result of improvements in the global data assimilation systems used in the ERA-40 reanalysis production.

2 - And the reference to ERA-Interim was suppressed whereas the reference to ERA-40 was added:

3 - Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. Q. J. R. Meteorol. Soc. 131: 2961– 3012, 2005.

1 - Ln. 335-345: What is the reason for the deterioration in the lower part of the CCDF of NSE in SIM_PHY?

2 - The answer to that question is not straightforward. Most of the stations concerned by a deterioration in the lower part of the CCDF of NSE have a NSE lower than 0.55 and represent approximately 57% of the total number of stations. One part of the explanation comes from the calibration of the subgrid drainage in SIM_REF and not in SIM_PHY as explained in the manuscript. Then, a NSE lower than 0.5 can be considered as a bad simulation. So, both SIM_REF and SIM_PHY have problems in simulating the river flow at those stations. Several reasons can be proposed, and the first one is that some basins are urbanized and this is not well represented in the model. Then as we have seen, there are compensating errors in SIM_REF (correct Qsim to Qobs ratio and too low IR downward radiation).

The manuscript was modified to add this comment (lines 356-359):

3 - Most of the stations affected by deterioration in the lower part of the NSE CCDF have an NSE below 0.55 and represent about 57% of the total number of stations. Part of the explanation comes from the calibration of the subgrid drainage in SIM_REF which is not done in SIM_PHY.

1 - Sect.4.3: I would suggest to rename this subsection as it is quite vague at the moment: most of the paper regards the comparison to old SIM.

2 - The section was renamed (line 387):

3 - 4.3 Seasonal river flows

1 - Ln.405: why not adding a third box for SIM_PHY to evaluate the effect of the new snow/soil schemes on the snow depth?

2 - As already mentioned in the response to the main comments, the comparison to observations and involving SIM_REF can only be made over the 9892 grid boxes and is limited to those below 1340 m (elevation below which IR correction is applied). Therefore, only snow at mid-altitude would be considered and adding a third box with SIM_PHY would not help highlighting the effect of snow and soil schemes on snow depth.

1 - Ln. 413: “baresoil”–> bare soil

2 - Section 4.6 has been removed

1 - Ln.414: At which depth the soil temperature observations are taken? Is any interpolation applied to the data?

2 - Section 4.6 has been removed (As indicated in the original manuscript, temperatures are measured at 10 cm, 20 cm, 50 cm, and 100 cm, and no interpolation was applied)

1 - Ln.440-449: It would be nice to explicitly link this discussion on the changes of Evap/Precip with the changes in the discharge mean bias.

2 - A sentence has been added to make this link (lines 447-449):

3 - In SIM_NEW, the ratio of simulated to observed flow is in excess whereas it is better simulated in SIM_REF with a peak centred around 1. This result is consistent with an evaporation deficit in SIM_NEW compared to SIM_REF.

2 - And the last sentence of the paragraph was changed to (lines 450-452):

3 - This result shows that the sensible heat flux in SIM_NEW is much higher than in SIM_PHY, mainly due to the increased incoming infrared radiation, which partially compensates for the evaporation deficit.

1 - Ln. 489: I would rephrase this sentence for readers not familiar with detailed snow models.

2 - Sentence rephrased to (lines 510-512):

3 - At the same time, as described in section 2, the snow model has been revised to improve vertical layering, snow compaction and solar energy transmission within the snowpack through the use of spectral albedo, as is done in more advanced models.

Comments on the figures

1 - Generally, the figure captions should be improved to make them more self-explanatory.

2 - The figure's captions were improved

3 -

Figure 1: Height of the topography of the 9892 cells of the SIM grid (left) and the 3878 cells of the mountain SIM grid (right). The cells of the mountain grid correspond to the 1044 points having an altitude greater than 500 m and described vertically by several layers. Zones in yellow correspond to the Seine and Rhone aquifers. The dotted line delimits the Alps mountain.

Figure 2: Annual average of uncorrected (left) and corrected (right) downward longwave infrared radiation from SAFRAN analysis.

Figure 3: Maps of annual average of the SAFRAN analysis for the period 1958-2018 of (a) air temperature at 2 meters, (b) specific air humidity at 2 meters, (c) wind speed at 10 meters, (d) total annual precipitation, (e) direct solar radiation, and (f) diffuse solar radiation.

Figure 4: Annual average of the SAFRAN analysis of (a) air temperature at 2 meters, (b) specific air humidity at 2 meters, (c) wind speed at 10 meters, (d) direct solar radiation, (e) diffuse solar radiation, (f) infrared radiation, and (e) total precipitation rate.

Figure 5: Comparison of the NSE CCDF (left panel) and the simulated to observed flow ratio (right panel) for SIM_REF (dashed blue line), SIM_PHY (solid red line), SIM_FRC (solid cyan line), SIM_TOP (solid green line), and SIM_NEW (solid orange line).

Figure 6: Maps of the difference in mean NSE for NSE>0 between simulations: (a) SIM_PHY and SIM_REF, (b) SIM_FRC and SIM-PHY, (c) SIM_TOP and SIM_FRC, (d) SIM_NEW and SIM_TOP.

Figure 7: Map of the difference in mean NSE for NSE>0 between SIM_NEW and SIM_REF (left panel), and SIM_NEW NSE map (right panel).

Figure 8: Taylor diagrams of seasonal river flows for the different experiments over the period 1958-2018.

Figure 9: Taylor diagram of Q10 and Q90 deciles of river flows over the period 1958-2018.

Figure 10: Comparison of monthly river flows with SIM_NEW for the Garonne at Lamagistère over the period 1958-2018.

Figure 12: Maps of mean annual Bowen ratio (a) and evaporation to precipitation ratio (b) for SIM_NEW on average over period 1958-2018.

Figure 13: Mean annual evaporation to precipitation ratio (a), and Bowen ratio (b), for experiments SIM_REF, SIM_PHY, and SIM_NEW.

1 - In the caption of Figure 1, the definition of a “mountain grid cell” should be added.

2 - Done

1 - Some of the figures could be merged together, for instance Fig. 7 with Fig. 8 and Fig. 9 with Fig. 10, for conciseness.

2 - As compared to the initial manuscript, some figures have been removed (2, 6, 7) and 16 and 17 have been merged together.