Reviewer's comment on manuscript gmd-2020-31

Summary:

The authors have improved the original manuscript, in particular in the introduction and model's description sections. They also addressed my main points, with the inclusion of a discussion regarding the role of compensating errors among the different simulations. However, I still have some comments on specific points of the revised manuscript, which I would like the authors to consider. For this reason, I recommend the acceptance of the manuscript after minor revisions.

Specific comments

Page 2 Ln 36: "coupled to hydrology" is not clear, I would suggest "... coupled to hydrological models"

Page 2 Ln 38-41: These two sentences could be improved, as it sounds a bit of a repetition.

Page 2 Ln 42: Atmospheric biases would be introduced in land surface model simulations even in offline simulations, for instance a precipitation bias in the atmospheric model leads to a spurious anomaly in soil moisture. Could you please clarify?

Page 5 Ln 145: "as realistically as possible" seems a too strong statement. Boone and Etchevers' (2001) 3-layer snow scheme is an intermediate complexity snow scheme, that is in between simple bulk models and very detailed snow schemes used for avalanche forecasting.

Page 8 Ln 249-250: I would suggest to put the range of values used in the vertical discretization in the text, in order to be clearer.

Page 10 Ln 309-311: Whilst I understand that all data were used, it is important to highlight if any stations contained large periods of missing data during the analysis period, and if so how this influences the conclusions drawn.

Page 11 Ln 336-341: As far as I understand, reanalysis use a fixed model and data assimilation system, for consistency, throughout the production. So I am not sure how these sentence relate to the behaviour in the forcing fields discussed in this subsection. Could the authors clarify this?

Page 14 Ln 420-421: The authors suggest that the improvement in altitude resolution, I think due to the new orography map implemented, can be responsible for the improvements. However, at Page 6, Ln 94-96 they state that "the impact of using SRTM90 is rather limited". Could the authors clarify the link between these two statements?

Page 14 Ln 421-423: Is this a hypothesis to explain the improvement in SIM_NEW? Or is it a diagnostic technique that could be used to further improve the results? Can the authors clarify this sentence?

Page 14 Ln 423-424: I would say that also the new vegetation maps can play a role in the improvements.

Page 15, Sect. 5.2: Many statements in this subsection could be improved for clarity.

Page 16, Ln 482: "...less radiative energy to evaporate ~~**it**~~ leads to an overestimation of river flows" (remove "it")

Page 16 Ln 492-494: This sentence needs more clarity, as it is currently difficult to follow.

Page 16 Ln 491-494: So the improvements in snow depth could be mainly attributed to the improvements in the snow/soil physics and the new climate fields, is it right? It would be interesting to know which one (new physics or new land cover fields) play the larger role.

Page 17 Ln 512: "as **it** is done in…" (add "it")

Page 18 Ln 555: the word "Formula" be more appropriate than "Principle".