Stefano Ferraris, 23 Aug 2022

The paper address a urgent need, the modeling of spatial and temporal water balance at the continental scale. Continental droughts like the one is occurring now make this need even more urgent. I fully agree that only streamflow fitting is not meaningful, and we need also hydrologic states and fluxes with available observations such as SM, evapotranspiration (ET), water table depth (WTD), snow water equivalent (SWE) and total water storage,

The paper is very detailed and well written, but some part of the process modeling make it necessary to be better explained.

We thank Stefano Ferraris for his positive comments on our manuscript. We have revised the manuscript based on your constructive comments and suggestions. We replied to your comments in the blue text below.

One first problem is overland flow:

I wonder about the sense of overland flow modeling with kinematic wave at 3 km spatial scale. It is also mentioned a "two-way overland flow routing" what is it?
In ParFlow-CLM, overland flow, which is generated by saturation or infiltration excess, is implemented as a two-dimensional kinematic wave equation approximation of the shallow water equations. The overland flow direction is determined through the D-4 flow routing approach. We revised the text in the manuscript for clarity.

Are Manning's coefficient or hydraulic conductivity you mention possible to be defined at the 3km scale?
As stated in the manuscript, in the current modeling setup, distributed parameters describing the soil properties, saturated hydraulic conductivity, van Genuchten parameters, and porosity were assigned to each hydrofacies and soil classes and were estimated based on the pedotransfer functions. In the revised manuscript, we now included a Table in supplementary material with complete parameter values used in the current study.

Vegetation is almost absent in the text. It is modeled with a single layer, but no more is detailed.
We appreciate your comment. As stated in the manuscript, land cover classes were based on the MODIS dataset (Friedl et al., 2002) and each class has unique parameters such as leaf area index, roughness length and reflectance. We provided more details about the vegetation parameters in the revised manuscript.
I have seen that an area intensively irrigated in summer shows quite low ET fluxes. Only the rice part of it have high fluxes, therefore I wonder if irrigation is taken into account in ET fluxes.
Irrigation is not taken into account in this model setup; hence also the ET fluxes are unaffected.

Snow has a very detailed coding, with up to 5 layers, how can be given such a description at the continental scale?
Detailed description of snow model in ParFlow-CLM model is given in Ryken et al., 2020. In the revised manuscript we now briefly described the main processes as:
"ParFlow-CLM simulates snow water equivalent using thermal, vegetation, canopy and snow age processes which determine the amount of precipitation falling as snow. Changes in snow through time is simulated through albedo decay, snow compaction, sublimation, and melt

processes. Snow layer is initialized when snow is present on the ground and can be divided up to 5 snow layers based on prescribed thickness and the amount of snow present on the ground."

The paper speaks in more details of soil moisture, but the first 3 centimeters say nothing about subsurface water flow. Field data are "from 19 stations from four networks and In case that more than 1 station is located within one 3 km grid cell, the average of those stations was used for comparison". Does it mean that less than 19 pixel in all Europe has a SM ground validation?
Thanks for pointing this out. For the time period of 1997–2006, we only have data available for 41 stations (please see Table 3 of Naz et al., 2020), however, for some pixels if there is more than 1 station located within the gridcell then the average of those stations were used resulting in 19 grid cells over Europe. We modified the text for clarity in the revised manuscript.

You mention "consistently higher mean SM": I think that are much more important the dynamics of SM. I agree to perform a montly average anomalies comparison, but the dynamics is partly lost.
We agree with the reviewer's comment. However, because of the data limitation (e.g. sparse in-situ data and only surface information can be compared with remote sensing observations), makes it difficult to perform more detailed comparison of SM dynamics at the deeper soil layers.
Also, I know that having information abut soil structure is impossible at the continental scale, but it has to be remarked that only texture cannot give enough information.
In the revised manuscript, we provided more information about the soil data limitations over larger scales.

Less important, a figure has no number, but only ?? at line 417.
It has been corrected in the revised manuscript.

References:
Friedl, M. A., McIver, D. K., Hodges, J. C., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., and Cooper, A.: Global land cover mapping from MODIS: algorithms and early results, Remote Sensing of Environment, 83, 287–302, 2002.
Naz, B. S., Kollet, S., Franssen, H.-J. H., Montzka, C., and Kurtz, W.: A 3 km spatially and temporally consistent European daily soil moisture reanalysis from 2000 to 2015, Sci Data, 7, 111, https://doi.org/10.1038/s41597-020-0450-6, 2020.
Ryken, A., Bearup, L. A., Jefferson, J. L., Constantine, P., and Maxwell, R. M.: Sensitivity and model reduction of simulated snow processes: Contrasting observational and parameter uncertainty to improve prediction, Advances in Water Resources, 135, 103473, https://doi.org/10.1016/j.advwatres.2019.103473, 2020.