

Geosci. Model Dev. Discuss., author comment AC2
<https://doi.org/10.5194/gmd-2021-53-AC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Tobias Stacke and Stefan Hagemann

Author comment on "HydroPy (v1.0): a new global hydrology model written in Python" by
Tobias Stacke and Stefan Hagemann, Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2021-53-AC2>, 2021

We like to thank the reviewer for reading and commenting on our manuscript. Here, we will repeat the reviewer's remarks (*italic font*) and add our response directly below (*standard font*).

In this manuscript, the authors provide a detailed description and validation of the global hydrological model HydroPy. The model is based on an existing model (MPI-HM) but has been completely rewritten in Python and made publicly available (along with the appropriate input data) under a GNU GPL license. The paper is well written and clearly explains all relevant processes in the model. Strengths and shortcomings are discussed, and potential targets for future model improvement are identified.

Thank you, we very much appreciate this positive evaluation of our work!

Lines 111-113: "It assumes ... storage overflow". These two sentences appear contradictory. Also, I have difficulties to relate them to the following technical description. Why is $S_{so,sg}$ scaled because $S_{so,max}$ is assumed to vary? This section would benefit from a more extensive description of the infiltration/runoff scheme and the concept behind it. In the end, this is one of the main components of the model.

We will add some more information to this paragraph to make it less confusing. The new version states:

It assumes that $S_{so,max}$ is not homogeneously distributed within a grid cell, but varies on subgrid scale. Thus, parts of the grid cell where the local storage capacity is low, can already generate surface runoff even though the cell average soil moisture state is still below its average maximum moisture holding capacity. Therefore, a fraction of R_{tr} is converted into R_{srf} as soon as the minimum soil moisture content is exceeded. This is realized by mapping S_{so} onto the sub-grid soil moisture capacity distribution parameters denoted by the index $_{sg}$.

Lines 182-186: It is very prudent to include such a feature. However, I would consider letting the simulation fail or at least throw a warning if a balance error occurs. Not all potential users of the model may be aware that the balance needs regular checking.

Indeed, we implemented the water balance check in such way that at the end of every simulation the global water balance components (both averaged over all grid cells and as

global sum) are displayed and color coded in green and red to indicate a closed or violated water balance. Additionally, the data file including the water balance fields is only written in case the water balance is not closed and therefore it can be seen at once if any problems occurred and in which year this happened. We will extend the manuscript to include this information.

Lines 206-207: "...wetlands were restricted to...". I am aware that GLWD identifies huge areas as wetlands in North America, which apparently doesn't correspond well to your model assumptions. But it feels a bit arbitrary to exclude the largest part of global wetlands just because you are unhappy with the results. What are the implications in other parts of the world? Since this is just a test setup, I see no urgent need to change this decision within the current study. But you should clearly flag it as a mismatch between the model and available data and come back to it in the discussion on wetlands in the section 4.4.

That is a very valid remark and we agree that some more information on our reasoning should be provided. The basic idea is that (at this stage) we wanted HydroPy to reproduce MPI-HM simulations as best as possible to have a robust reference for all further development. At the same time, we found it important to update some of the rather old datasets used for the land surface characteristics. Our GLWD wetland class selection was not so much guided by optimizing the selection to get the best possible discharge, but rather to best resemble the wetland distribution in the Matthews and Fung dataset (used for MPI-HM) to allow for a fair comparison between both. We will modify the sentence to emphasize this reasoning:

The lake fractions are used without any modifications, but wetlands were restricted to the classes floodplains and peatlands for this study. Thus, we can best resemble the general wetland distribution used for the predecessor of HydroPy and facilitate a clean comparison between both (see Sect. 4.4).

Lines 210-211: Note that such interpolation alters the effective monthly averages, with largest effects in months with minima and maxima. However, it is common practice and not easy to correct for.

We fully agree with the reviewer. Still, we are confident that the error caused by this approach should not affect the simulations too much for the rather large time scales the model is intended for. Of course, any results for short term application would need to take this effect into account.

Line 244: "...distribution parameter b_{sg} ...". Can you give a value for this parameter?

This parameter is a spatially distributed field provided in the land surface dataset, not a single value. For the majority of grid cells, the value is rather low (<5) but in more extreme (and only very few) cases it can reach values up to 100.

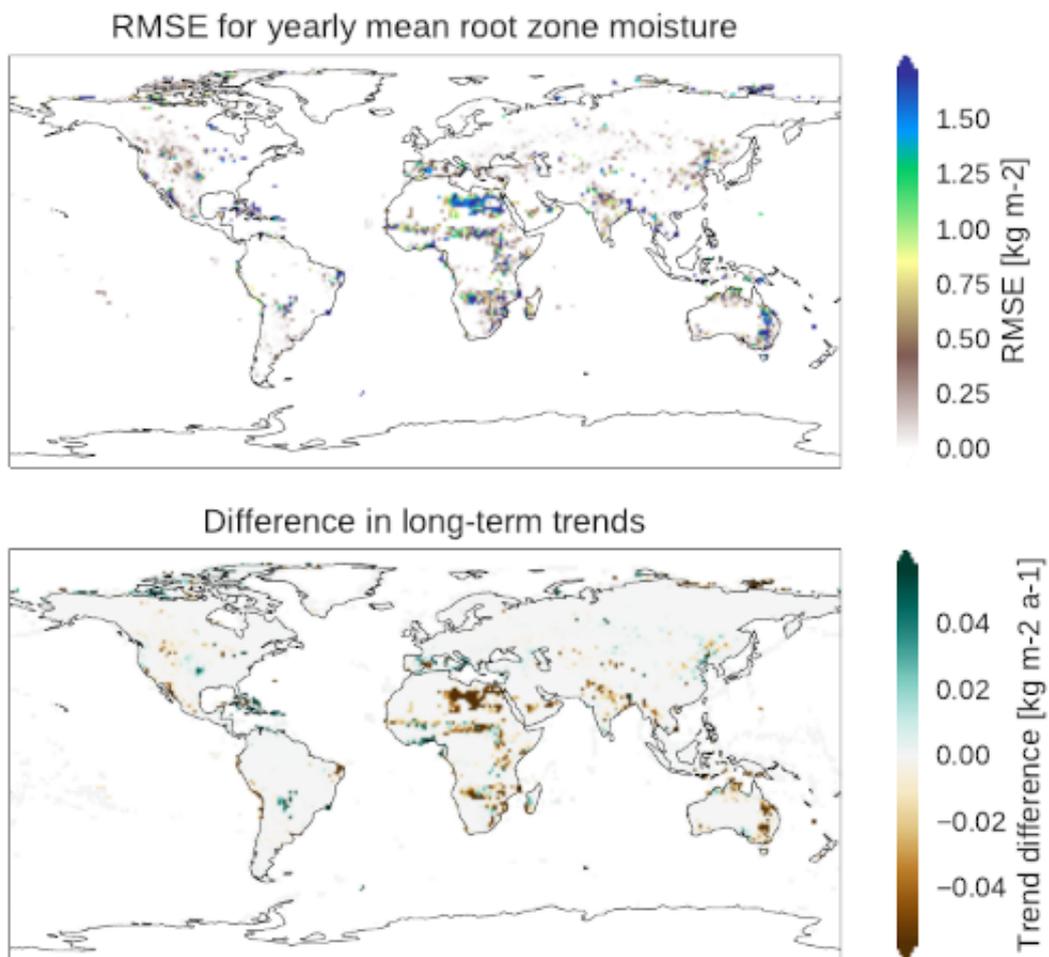
Lines 249-281: Please assure consistent use of subscript x in text and equations.

We are not sure whether we recognized the issue you are pointing at. Is it, because we are also using a subscript y in this paragraph? Here, it is necessary, because Eq 39 and 40 are 2-dimensional using the subscript x for flows in certain water body types (surface water and river) and y for the land cover types (lakes and wetlands) the water flows through. Anyway, we will check all subscripts for consistency and consider replacing generic subscripts wherever appropriate.

Lines 286-287: "The simulation ... until 2014". If I understand correctly, you are trying to bring your model to an equilibrium using conditions in a single year (1979). Wouldn't it be

better to use a range of years or at least a climatology?

Yes, our intention is to provide the model with a storage state where changes are dominated by climate input and not by any residual trends. There are several (small) advantages and disadvantages to the methods: using a climatology might reduce the day to day variability and thus limit the model's exposure to any daily extremes. Picking just one year (as we did) could lead to a bias in storage state in case the particular year happens to be an extreme year in a given region but allows for a very clean setup with a constant storage target, which the model approaches. Using a time series for spinup generates the exact opposite, e.g. extreme years and therefore an anomalous storage state are less likely. However, the target state is less well defined due to inter-annual variability. We decided to use the approach with the best defined target state because it allows for a more straight forward evaluation of the spinup behavior. Anyway, we claim that our model is not very sensitive to either method as long as the number of spin-up years and the general climate are similar. In order to test this hypothesis, we performed another spin-up simulation using a time series from 1930 until 1979 (instead of a single year) as well as another production simulation from 1979 until 2014 based on this spinup. We then compared our original simulation and the new simulation (Fig 1). This comparison demonstrates that indeed both methods cause differences in the production simulation. However, these differences are very low resulting in an RMSE which rarely exceeds 5 kg m⁻². Furthermore, the long term trends found in both simulations are very similar with only a small numbers of grid cells that show different trends due to the different spinup methods. From this, we conclude that the choice of spinup method has only a minor effect on our results.



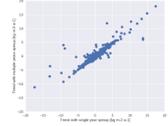


Figure 1: Impact of different spinup methods on root zone soil moisture during the simulation period. The left panels show the spatial distribution of root zone moisture RMSE and differences in long term trends between simulations initialized with multi-year or single-year spinup. The right panel displays the relation between trends in the different simulations at grid cell scale. Note, that all panels displays root zone moisture and not TWS as the latter is affected by single cells of steadily increasing snow cover (see Sec 4.1 in the manuscript).

Lines 306-320: I very much doubt that checking for the pixel with the largest absolute residual trend is sufficient proof that the model is in equilibrium. In many regions, the absolute storage trends are small because the involved water flows (precipitation, runoff, recharge, discharge) are small. Thus, storages (and fluxes) will take much longer to reach equilibrium in regions where initial storage trends are small (a fairly large part of the world, according to the map in Figure 3). This is of minor concern when considering globally aggregated fluxes, which are dominated by regions with large fluxes. But I suspect that many of the other results presented in the remainder of the paper are affected by storages not being in equilibrium at the beginning of the simulation phase. I strongly suggest using a different indicator for quantifying the proximity to equilibrium and adjusting the spin-up protocol accordingly.

We used this method exactly for the reason you mentioned: any residual trend is most likely only found in regions where the annual water turnover is much smaller than the storage capacity. However, we think such regions are very few and do not interfere with any of our analysis, either on global nor on catchment scale. To some part this is already demonstrated in Fig 1 (right panel) which shows the similarity of simulated long-term trends in spite of using different spin-up methods. Additionally, you can see in Fig. 2 the trends in total water storage for the last 10 years of spinup up simulation. Apart from the already discussed grid cells in glaciated regions, there are very few grid cells which show any residual trend (still being rather low with $< 5 \text{ kg m}^{-2} \text{ a}^{-1}$), but these are located in rather wet regions like the Ganges Delta and the Pantanal wetlands. Based on this analysis, we are confident that none of our results are significantly affected by spinup. We will extent Fig 3 of our manuscript to include this plot as well.

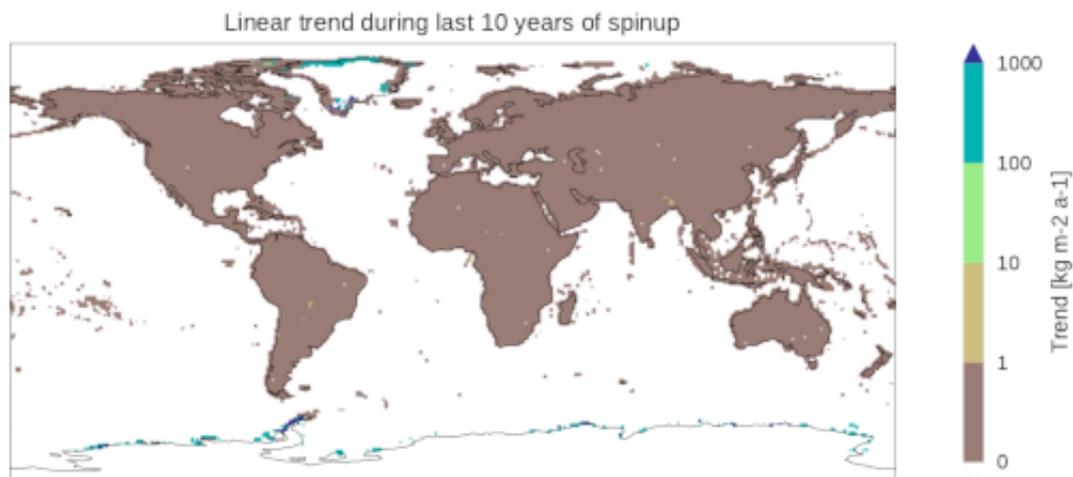


Figure 2: Spatial distribution of yearly trends during the last 10 cycles of the spinup simulation.

Line 370: "...temporal correlation...". What time step is used here?

The analysis time step is monthly (because we have only monthly data available for observations). We will add this information.

Line 374: What do you mean by "percentile bias..."?

The percentile bias is the ratio of the sum of differences between simulation and observation to the sum of observed values given as percentile. We will add the equation to the analysis section.

Line 416: What do you mean by "mitigated flow curve"?

We mean that the flow curve has a delayed and lower peak than its reference. As we say this in the next part of the sentence anyway, we will remove 'mitigated' from the sentence, to avoid any confusion. Sorry for this poor phrasing.

Lines 439-441: This would also solve the problem that led to the exclusion of wetlands, correct?

Yes, we think this would very likely be the case.