

Hydrol. Earth Syst. Sci. Discuss., author comment AC3
<https://doi.org/10.5194/hess-2021-527-AC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Alemu Yenehun et al.

Author comment on "Spatial and temporal simulation of groundwater recharge and cross-validation with point measurements in volcanic aquifers with variable topography" by Alemu Yenehun et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-527-AC3>, 2022

Response to Referee 2

General

In this study Alemu Yenehun et al. estimate groundwater recharge and study the spatial and temporal recharge patterns for the Lake Tana region in Ethiopia. Three (well-) established recharge estimation methods are applied, a model approach using the physically-based WetSpa model, the groundwater table fluctuation method, and the chloride mass balance method. I applaud the Authors for conducting such an extensive study in a research area where data acquisition I assume can often be challenging. I think the main contributions from this paper are 1) an improved understanding of the groundwater recharge fluxes in the case study area, and 2) the comparison of different recharge estimation methods. These contributions are valuable, as groundwater is an important source of drinking water in the region and a better understanding of the resource in this region would serve many. I found the manuscript generally well written (minor textual improvements are required), but some restructuring and changes to the figures might be required to improve the readability of the manuscript. The main concern I have for this manuscript are related to the methodology and the description thereof, discussed in detail below and in separate line comments. Additionally, while I think the study is worth publishing, I am unsure whether the contribution fits the scope of HESS ("substantial new concepts, ideas, methods, or data") in its current form. In my opinion, it would fit much better in other journals where it can be submitted as a case study (e.g., Hydrogeology journal), and reviewed as such. If the Authors wish to publish in HESS, I think a more elaborate analysis and discussion of the uncertainties in recharge estimation could be done to better fit within the scope of the journal.

Reply: Thank you so much. We have tried to do an uncertainty analysis of the model parameters and have added it in the revised version.

Methods description

From the description of the methodology and data following in this manuscript, it is hard to reproduce the results and gain a full understanding of the modeling procedure. This could partly be solved by sharing the scripts, input, and output data used in this study. However, often a more detailed description of the modeling procedure is also required.

The WetSpass model requires substantial data input, which is not always clearly described in the manuscript. From the description of the data, it seems that different time periods were used to generate the input data for WetSpass. Perhaps this could be clarified using a table that summarizes the different data sources and time series characteristics (e.g., measurement interval, period, operator). Given the high temporal variability of the different hydrometeorological variables, I assume the same time periods for all variables are used, but I could not verify this from the current manuscript. A few times it is mentioned "expert judgement" is used, but it remains unclear what values were applied and why. The calibration process is only briefly described, and the calibrated parameters are not reported. In general, I think it would be good to rewrite the methods section with reproducibility in mind.

Reply: Yes, indeed, a table containing the different data sources and time period of the WetSpass model input variables will be added in the revised final manuscript. It is true that for some of the model input variables, the data for different time periods have been considered. As the model is using long-term averages of the different variables as an input (in our case: long-term average summer and winter), hence, observing long-term time-series temporal variations of the water balance components with the WetSpass model is impossible. The model is limited as it is simulating only long-term monthly averages as a fine time scale (time series results are not possible). In other words, the model is basically developed to simulate long-term average monthly/seasonal conditions, and most importantly the spatial variations. The model is basically developed for spatial simulation of groundwater recharge (which is later coupled with groundwater modeling-MODFLOW) whose short-time (day to day or hour to hour variations) is less important compared to the other water balance terms.

Except for the groundwater level map (which is based on recently collected data), meteorological variables (rainfall, average temp., wind speed, all meteorological parameters used to calculate PET using the Penman-Monteith method) were more or less in similar time periods. Yes, it is true that the meteorological variables have high temporal variations (most importantly the seasonal or monthly, or daily variations); year-to-year variation is relatively small. In other words, there is a small variation for a given season or month for different years. Given the objective of this paper i.e. the spatial variation of recharge and to a certain extent other water balance terms, these would not be a major problem.

The other non-meteorological (physical) parameter that changes with time is the land use. For the long-term average data (2005 to 2018), we used the land use map using the 2014 satellite image and ground truth data, for the modeling; land use of 1980 and 2000 for the evaluation of the effect of land use change on the water balance terms. The use of one-time data for the land use and long-term average for the other parameters might have slightly departed from reality and may have caused some uncertainties. However, the model is not robust enough to use different land uses for calibration. Furthermore, the change in land use is not so fast, and hence, the land use of 2014 might have represented more or less the possible land use status of the area during 2005-2018.

Explicit statements for what do we mean by "**expert judgment**" will be added in the revised version. We mean some adjustment for the parameters' values on the default land use parameter table has been made, for example, the root depth of forest land is changed. In the model parameter table, each land use type has given impervious, bare land, vegetation, and open water percentages. In the default parameter of the WetSpass model the vegetation area coverage for the bush and shrubland is 100%. This is based on the observation of the land use class type in the temperate zone (the Netherlands and Belgium). However, the Bush and Shrubland use in the Ethiopian (tropics) context is different: the vegetation is sometimes sparsely distributed, and is with some bare land component, during our field verification on the land use type, we tried to guess (using

parcels of land on which we do some measurement), and came up with about 10% is bare land and the rest 90% is consisting of vegetation. Similarly, adjustments have been made by Gebreyohannes et al. (2013) during their application of the model for Geba catchment, in northern Ethiopia. Similarly, the sub-afro-alpine vegetation land use type found in our area (consisting of about 0.3% of the total area coverage) is not present in the default land use classes of the WetSpass model. However, we made an equivalent with wet meadow land use type, and following a similar procedure we modified the land use percentage to 80% vegetation and 20% bare land (it was 100% vegetation for wet meadow land cover type in the model parameter table).

<< insert table1: the model land use parameter table here, attached in the supplement file>>

The steps followed in the calibration process, and the table for the calibrated parameters will be added in the revised version of the manuscript.

The modeling section of the methodology is more elaborately written in the revised version. However, to reduce the pages of the manuscript, all the detailed equations, and the empirical formulas developed and followed in the WetSpass model are better not to be included, we would rather cite the papers that developed the model first and improved further.

Uncertainties in recharge estimates

The authors mention in the introduction that it is important to take uncertainty in recharge estimation into account (line 47), and thus I was expecting a more elaborate analysis or discussion of the uncertainty in recharge estimation methods applied in this manuscript. As the WetSpass model is manually calibrated, no parameter uncertainties are available. It therefore remains unknown how uncertain these recharge estimates are. This could be addressed by a sensitivity analysis or a more elaborate uncertainty analysis. A discussion of the limitations of the different methods and the uncertainties of the recharge estimates at the end of section 4 would be a welcome addition to the current manuscript.

Reply: The WetSpass model has different parameters and needs input variables to give a reasonable output for the understanding of water balance components of a given river basin/watershed. Yes, to evaluate the efficiency of the WetSpass model applied in this study, the sensitivity analyses of all parameters and input variables would be a good addition of the paper. A simple sensitivity analysis is carried out and added in this revised version of the manuscript. Since the objective of doing this sensitivity analysis is to know the possible uncertainties that might be incorporated into the estimated water balance values, we did a sensitivity analysis for the calibration of the global parameters of the WetSpass model. These are alfa coefficients (it is a parameter adjusting soil moisture value), interception coefficients (adjusting interception by the plant canopy), LP coefficient (parameter adjusting evapotranspiration), rainfall intensity, and x coefficient (runoff delay factor). The other physical factors (called local parameters) such as the contribution of the slope (40%), landuse (30%) and soil (30%) in determining the balance components share, are originally developed from literature (experimental research). Thus, we kept them as they are (we have taken the default values of the model), and hence, we have not done a sensitivity analysis though their local variation might have caused some uncertainties in the estimated recharge and other water balance terms. Furthermore, we do not perform sensitivity analysis of the meteorological variables used in the model perhaps they might cause some uncertainties during measurement, interpolation, etc. They were fixed variables in our calibration similar to the physical parameters mentioned above.

Among the global model parameters calibrated in the model, the most sensitive global

model parameter is RF intensity. On average, for every 1 mm/hr increase of the intensity, there is a mean annual recharge increment of 38.2 mm. Fortunately, we measured the rainfall intensity at some of the meteorological stations that the BDU-IUC project (the project that sponsors this study) established (using the automatic rain gauges), and hence, we used the range of these data in our calibration.

<<insert Figure 1 attached in the supplementary file here>>

Figure 1. Mean annual groundwater recharge (mm) versus average rainfall intensity (mm/hr.) for the different possible intensity values. The model was running for these RF intensity values by keeping the other model parameters constant.

In general, except for the RF intensity (which highly depends on the rainfall characteristics of an area), the other global model parameters have given a possible range of values in the WetSpass model (LP coefficient and x coefficient: 0.1 to 1; and alfa coefficient and a interception 1 to 10). Hence, the sensitivity testing for values within these ranges has been performed.

For every variation in LP coefficient, there is a 4 mm increase or decrease of the annual recharge.

<<insert Figure 2 attached in the supplementary file here>>

Figure 2. Mean annual groundwater recharge (mm) versus LP coefficient for the different possible LP coefficient values.

On average, for every 1 value variation for alfa coefficient there is only about 2.3 mm variation in annual recharge rate.

<<insert Figure 3 attached in the supplementary file here>>

Figure 3. Mean annual groundwater recharge (mm) versus Alfa coefficient for the different possible Alfa coefficient values.

The variation in a interception is small and only for the initial values. For the relatively higher values, the recharge is almost not changing, and hence, the recharge can be considered as insensitive for the variation of a interception. This has been also shown with the low R2 value.

<<insert Figure 4 attached in the supplementary file here>>

Figure 4. Mean annual groundwater recharge (mm) versus a interception coefficient for the different possible a interception coefficient values.

The groundwater recharge is not changing for the possible changing intervals of the x-coefficient, and hence, the x-coefficient is the most insensitive parameter among all the global model parameters for the Lake Tana basin.

<<insert Figure 5 attached in the supplementary file here>>

Figure 5. Mean annual groundwater recharge (mm) versus x coefficient for the different possible x coefficient values.

The limitation and the sensitivity analysis which would bring uncertainties to the reported values have been added in the revised version.

Title

The title suggests a more general study on volcanic aquifers, while a case study is presented. I think it would be better if the title reflects the fact that it is a case study. Additionally, the title suggests that there are "point measurements" of groundwater recharge. In my view it be better to refer to these as 'point estimates', as they are empirical estimates from a recharge estimation method and not real measurements.

Reply: Thank you so much. Yes, indeed, they are point estimates. For example, the variables in the calculation of the WTF method are the change in the water level and the specific yield (Sy). The water level is based on fine-time scale measurement, however, there is estimation for Sy though it is also analyzed from slug and pumping tests (real measurements), estimation from Johnson (1967) literature compilation has been also used. Given the high aquifer variability (compared to the representativity of each test well), it is true the WTF values are also estimations. Thus, the topic is corrected accordingly.

Yes, indeed the methods are tested in the case study area, however, we do believe the evaluation part that recommends the effective application of the methods for aquifers lying at different topographical and hydrogeological conditions would make the study beyond being case study, and will attract global readers.

Line comments

L31: add a comma after floodplain

Reply: thank you, added.

L72: regionalize

Reply: accepted, and corrected accordingly.

L77: were all these studies in a specific study area? If so, good to mention that.

Reply: mentioned, and thus corrected accordingly.

L123: Perhaps refer to the GitHub repo: <https://github.com/WetSpass>

Reply: thank you so much. Cited accordingly.

L123: No reference to Wang et al (1996) required?

Reply: Wang et al (1996) is about the WetSpa model. Maybe it is a good idea to add it as WetSpa is the starting point and the basis for the development of WetSpass. Hence, it is added in the new revised version.

L132/135/139: Sentences introducing these equations would be nice.

Reply: thank you, added.!

L133: Should there not be a change in storage term?

Reply: Yes, in the WetSpass model, the storage term is neglected. There are two ways of incorporating change in storage to the model on a seasonal basis. The first instance is that the plant available soil moisture reservoir in the summer is assumed to be filled up while it can be depleted in the model in the winter (for areas like Lake Tana basin); the second

case is that in the model a different groundwater depth can be used for winter and summer. In our methodology, we followed the second option, we had prepared the groundwater table map for the summer and winter seasons using the time-series groundwater level monitoring data that we used for our study. This is explained in the methodology part of the manuscript. However, since we have Lake Tana in our study area, there is a possibility of surface water storage change. For example, the evaporation from the open lake water may cause the summation of annual AET + Runoff + Recharge to be greater than the annual RF i.e. $RF < AET + \text{runoff} + \text{groundwater recharge}$. Hence, there is a possibility that the storage term changes. However, change in storage in the WetSpa model is not as important as the other hydrological models, which can simulate water balance components at fine-time series (such as a daily or hourly). For such models, storage amount (variation) at each time-series highly affects the water balance components of the following time-series value. E.g. the storage change of day 1 affects the water balance values for day 2. But in the WetSpa model applied in this study (seasonal), there are only two time-series whose storage difference (especially the groundwater storage) by an independent water table map.

L142: Which evaporation equation was used?

Reply: the Penman equation is used for evaporation in WetSpa. It is discussed already in the result and discussion section.

L144: I think it would be good to state which physically-based equations (e.g., Darcy) were used and elaborate a bit on the model (e.g., finite differences).

Reply: In general, the WetSpa model is based on known hydrological equations (conceptually and physically meaningful equations), and also some empirical equations developed from works of literature. Yes, a better-elaborated explanation of the model background will be added in the revised version of the manuscript.

L147-156: This description is rather vague, what was done exactly? How were the values changed?

Reply: In the revised manuscript, detailed changed values of the land use parameters are explained, as it is already explained above. Basically, the changes are in the percentage of bare land and vegetation area percentage of a given land use type (bush and shrubland and sub-afro-alpine grassland), and also root depths of forest land (it is already explained in the paragraph).

L157-162: This could be elaborated and made more specific. How many parameters were calibrated, how many data points were used for calibration, was there a formal objective function used or only visual goodness of fit?

Reply: there is no formal objective function used, only the visual goodness of fit has been used. Basically, about five global parameters have been used for calibration (discussed above). The final optimal values (after intensive manual calibration trials) have been presented in a table in the revised manuscript. The goodness of fit was compared between the accumulated mean annual river discharge amount (simulated surface runoff +baseflow) vs the measured river discharge amount of the seven major rivers.

L164-164: "to validate the recharge estimates from WetSpa."

Reply: validation is changed by evaluation, the word might have been confusing.

L165: I was a bit confused here, perhaps change to "During model calibration".

Reply: thank you, changed accordingly.

L174: remove "our" or mention the operating organization

Reply: thank you. It is changed by the BDU-IUC project funding

L177: remove "relatively"

Reply: removed

L178-179: Add a reference for this statement

Reply; thank you, added.

L183: Can mostly be specified?

Reply: Yes, it has been done in the revised one.

L187-188: I think at least a reference would be appropriate here.

Reply: Ok, will be done.

L191: The naming "so-called first-class stations" seems inappropriate here, I suggest rephrasing, also in the Figures.

Reply: Ok, corrected accordingly!!

L198-201: How many monitoring wells were used? What period were the groundwater tables observed? Does the period overlap with the discharge data and the meteorological data? Is this the same data as that mentioned in L306?

Reply: Yes, they are the same data listed on L306. The monitoring wells were 65 in number. The measurements were taking place from 2017 to 2019: some using automatic water level measuring device (measuring every half an hour), and others were manually (per each day).

About data overlap: partly it overlaps with meteorological data. However, the meteorological data at most of the stations are long-term averages (2005 to 2018), generally, it is impossible to say the data exactly overlaps. However, since it is long-term average seasonal values that we are finally used as input, it will not be a major problem. The intention of the modeling work here is not primarily to see the long-term temporal variation of recharge and other water balance terms (but limited to mean seasonal variations). Similarly, the river discharge data used for calibration is not in the same time interval (2012-2013, mentioned on line 204). Nevertheless, as it is the average seasonal values that are applied, it will/would not have significant variation.

L205-206: One value per year, does that mean the model is calibrated to just a very few data points? How many rivers were in the dataset?

Reply: using 7 river gauging data at 7 major rivers of the basin (shown in Fig. 8). It is the seasonal values which are used for calibration (summer: June – September, and winter: October to May).

L231-232: Unclear what "appropriate" is here.

Reply: It is to mean code matching between the model parameter tables and land use

classes has been made so that the model reads the right land use.

L279: How many pumping and slug tests were done? What values for S_y were found, are these reported or available somewhere?

Reply: This pumping test analysis is our next paper (already prepared and will be submitted to a journal soon). A test on 31 hand-dug wells was successfully executed and analyzed (the statement is added in the new revised version). Yes, the results per aquifer type were reported already on lines 507 to 515 of the old version.

L282: Perhaps a separate subheading for the CMB.

Reply: Ok, accepted.

L306: So the recharge estimates from the WTF method are only representative for that period. Is this considered when comparing to the other methods?

Reply: As we have mentioned earlier, the recharge estimated in all of the methods is mean annual. It is true that the meteorological variables used in the WetSpass model, and perhaps the land use have variations over the years. However, the annual values of these meteorological variables used as input for WetSpass are more or less constant. However, day-to-day variations of these parameter values especially the rainfall is high and would have resulted in high daily variations of water balance components such as recharge. Unfortunately, the WetSpass cannot simulate such fine time-series variations except the long-term seasonal variations.

Comparing the long-term recharge amount simulated by the WetSpass and the WTF for these periods would not be much of a problem (as it is the mean annual which is being compared). Similarly, the CMB is calculated only for the groundwater and rainfall sampling years (2017-2018).

The effect of land use change is evaluated in this paper and found small.

Thank you, we will put this fact as one of the limitations of this research.

L325: Perhaps this could be combined with Fig. 2. with the hydrogeological setting?

Reply: Yes it is possible to combine. However, it would be too earlier to put it here for the readers to follow the methodology section.

L336: Where is 4.1? Also, I suggest starting with 4.3 (verification), before discussing the water balance components.

Reply: Thank you so much, we corrected both according to your suggestions.

L342: At this point, I do not know for what period WetSpass computes the recharge. Perhaps write: "The annual recharge over the period 20XX-20XX.."

Reply: The WetSpass model computes from 2005 to 2018. We added the phrase in the manuscript.

L348: minor typo.

Reply: corrected accordingly.

L351: Change "Next to" to "Apart from".

Reply: thank you, have done.

L360: minor typo

Reply: thank you, corrected accordingly.

L366: values

Reply: corrected.

L379: Exclude the lake in figure 7B. Perhaps add subplot with the precipitation. The legend in Fig.7a has a mistake in the values. Why was a continuous coloring scheme not used (e.g., from yellow to blue), this would make the figures much easier to interpret. In the figure caption, specify "long term" by mentioning the exact years.

Reply: thank you for the comments. It is not the lake which is mentioned in Fig.7b, rather the value. There is by far a higher AET value than in the other areas (since it is a water body where open water evaporation operates). That is why the contrast between the lake and the other land uses is higher, which causes the values at the lake to have a similar colour.

Yes, we will remove the lake in figure 7c. The mistake in the values of the figure 7a legend might be the 0 value (I guess) over the water body. WetSpass assumes that there is no recharge for the aquifers under the floor of the water body.

We have checked and corrected the maps according to the suggestions.

L407-447: I do not think this is a proper "Model verification", as the same data is used for calibration! We cannot verify a model using the same data that was used for calibration. Perhaps this section can be renamed to "Calibration results".

Reply: in the WetSpass model, there is no way that we insert calibration data (river discharge data) in the model itself during calibration. We were running the model (more than a hundred times) by changing the global model parameters, and we have been cross-checking the river discharge values of the basin. After doing it for several times, the optimal values (i.e. the closest values between the simulated and measured discharge data) were reached, and changing the values for the modeling parameters reached to a point that changing values no more improved the R^2 between measured and simulated values. However, this is mostly between the total annual river discharges of all the rivers together. Hence, comparing the discharge values of the simulated and the measured after extracting the values (in raster map) by each river catchment can be partly considered as validation. However, as you mentioned, it will give full confidence if we discuss it under calibration rather than validation. For validation, the one estimated by the WTF and CMB can be still kept on.

L427-429: A high R^2 was obtained after the calibration, which may be interpreted as that the model can explain a large part of the observed discharge variation. However, Fig. 8 also clearly shows a large systematic error between modelled and observed discharge, which I think could be more clearly stated. Contrary to what is stated in the text, the simulated discharge is always higher compared to the observations. This would be a good point to get back to in the discussion.

Reply: Thank you for your concern over the issue. As you stated, generally simulated values are higher than the observed ones. The possible reasons have been discussed in this section. The possible reason would be the assumption of the method that we followed to calculate the total river discharge: as it is already stated, we assumed the total

discharge is a summation of surface runoff (SR) plus groundwater recharge (GWR) that the model gives both in spatial maps and as a single total value. However, theoretically, river discharge can be lower than this summation when there is significant deep groundwater flow that flows through the subsurface (without emerging to the surface) i.e. part of the water goes without being measured by the river gauging station. In other words, baseflow is not always equal to recharge, there are cases that baseflow is much lower than the recharge. Hence, there is a probability that the river discharge value is less than the sum of SR and GWR. The real geological conditions for the different catchments are already mentioned with reference to previous studies in the discussion section.

In general, it does not mean that the WetSpass model overestimates the river discharge but it maybe because of the assumption of the baseflow amount to be equal to groundwater recharge in the total river discharge calculation of this paper. As the WetSpass model does not give total river flow amount as an output, we followed this procedure to equate it so that we can compare with the measured river discharge amount measured at the river outlets.

L448: Here the section "Model verification" starts in my opinion.

Reply: Ok, we edited accordingly.

L448-463: Is there a reason the estimates from the CMB method were not used here?

Reply: Because the CMB is done for the different river catchments (by subdividing the river catchments into eastern, southern, northern, and western catchments, based on the similarity and difference in the runoff characteristics) rather than calculating at each sampling point. We took the average chloride concentration sampled and analyzed for the different major river catchments (Table 4) (chloride for both groundwater and rainfall), for doing the calculation (as mentioned already in the methodology section). It is impossible to get effective precipitation (rainfall minus surface runoff at each sampling point, which is one of the important variables in the CMB calculation). Hence, it is not easy to extract values from the recharge raster map produced by WetSpass at each sampling point like for the WTF method. However, we have a discussion part about it in the section (comparing the result with other methods) "**summarizing groundwater recharge estimations by the different methods and comparing with other similar studies**".

In the validation section, we added a discussion with CMB as well in the new revised version.

L465: Figure 9 could be condensed/smaller.

Reply: accepted, we have done it.

L500-549: This section could be placed earlier, as the results described in this section were previously used to compare to WetSpass. Add "method" after WTF throughout the section.

Reply: Ok, thank you, rearranging has been done.

L506: "has been taken"

Reply: inserted

L526: Change "to catch up" to "to capture".

Reply: changed

L557: Perhaps the higher range of values could be reported here.

Reply: It is to mean if the CMB calculation had been made at each groundwater sampling point (if it was possible to do so), the spatial variation of recharge would have been more than what has been calculated and reported i.e. average values over each catchment.

L560-591: Just a suggestion. Perhaps the Authors can come up with a nice plot that visualizes the different recharge estimates from all the other studies, and those from this study.

Reply: yes, we will show it either in a table or cross-plot by summarizing the studies done at different sub-catchments in the Lake Tana basin.

L595: minor typo.

Reply: corrected

L612: Perhaps I misinterpret, but I added the percentages of runoff (29%), recharge (22%), and evapotranspiration (53%), and these do not add up to 100% (=104%). Why is this, are there model errors or changes in storage?

Reply: This is already discussed above as a reply to your general comments. This might have happened due to two reasons:

1) some of the recharged water has also transpired especially in the forest area (deep-rooted eucalyptus trees considered in the root depth parameter table (5.5 m depth), which is below the water level, especially for the summer season). The WetSpass adds this transpiration amount from the groundwater to the calculated total AET value, it is why the water table raster map is needed as an input file.

2) The AET at the open water (the Lake Tana and the Wetlands) is equal to the PET, which is higher than the RF, hence there might be some imbalance at, for example, the Lake Tana. This is also reported by different authors as a closure error.

L640: Some recommendation/implications for future studies and work could be added at the end.

Reply: Yes, recommendations suggesting how researchers and water management practitioners use this research output, and limitations of the study mainly method assumptions and limitations are added in the revised version.

Please also note the supplement to this comment:

<https://hess.copernicus.org/preprints/hess-2021-527/hess-2021-527-AC3-supplement.pdf>