Review of “Synergy between satellite observations of soil moisture and water storage anomalies for global runoff estimation”

We thank the reviewers and the Editor for their supportive review. In the revised version of the manuscript the following relevant changes have been made:

- the title has been changed in “Synergy between satellite observations of soil moisture and water storage anomalies for runoff estimation” (see comment#2 of RC2 or comment#3 of RC3).
- To avoid any misunderstanding, the STREAM model has been defined as a conceptual hydrological model instead of a data-driven model (see comment#4 of RC2).
- The section “code availability” has been modified specifying that “STREAM model code version 1.3 has been stored in Zenodo (https://zenodo.org/record/4744984, doi: 10.5281/zenodo.4744984)”. The authors have included the name of the model and the version throughout the revised manuscript (see EC comments).
- In Zenodo the authors provided a tutorial (a markdown “README.rst” file), describing the input and output files as well as the codes to be run for the simulation of river discharge and runoff. Moreover, in the “README.rst” file the authors specified that STREAM codes are distributed through M language files that can be run through different interpreters such as the free Octave (I tried to run the code by using Octave and it works).
- References and details on the equations have been added according to the reviewer’s suggestions (see comment#8 of RC2).
- A Table (Table 1A) has been added in appendix to better describe the STREAM parameters and their range of variability (see comment#10 of RC2).

EC: Comments and replies

1. CEC1: Dear authors,

   You can not publish the code of the model after the acceptance of the paper. It must be available during the review process. Therefore, please, make it available in a permanent repository—for example, Zenodo. As soon as you have done it, post a comment to your manuscript, including the model version number and the link to the Zenodo repository and corresponding DOI.

   Also, please, in the next step of the review process, be sure that you include the name of the model and the version that you use in the title of the manuscript.

   Best regards,

   Juan A. Añel

   AC: The STREAM model code version 1.3 has been stored in Zenodo (https://zenodo.org/record/4744984, doi: 10.5281/zenodo.4744984). In particular, following your suggestion, in the "README.rst" file we specified that STREAM codes are distributed through M language files that can be run through different interpreters such as the free Octave (I tried to run the code by using Octave and it works). Details are given in the revised manuscript at the section "Code availability".

2. EC2: For purposes of this paper, I would recommend that you use the DOI of the version of record. This is in order to ensure that the model described in this manuscript matches the source code and repository, and that future changes may be made without breaking that link.

   AC: The authors thank the editor for the suggestion. In the revised version of the paper, the DOI of the version of record has been used. Details are given in the section "Code availability" of the manuscript as follows:

   "The STREAM model version 1.3 , with a short user manual, is freely downloadable in Zenodo (https://zenodo.org/record/4744984, doi: 10.5281/zenodo.4744984)."

3. EC2: In addition, GMD requires a user's manual as part of the manuscript. I see here (\url{https://github.com/IRPIhydrology/STREAM}) that you have a small ``tutorial''. However, while this describes variables (quite useful), it does not actually give instructions for how to run the model, which is more of what I would expect when I see a ``tutorial''. I would recommend that you move all of this material to the README, and I ask that you generate a more full-featured user's guide as documentation.

   AC: In agreement with the Editor suggestion, on Github (https://github.com/IRPIhydrology/STREAM) and on Zenodo (https://zenodo.org/record/4744984, doi: 10.5281/zenodo.4744984), the authors have provided an updated document "README.rst", describing the input and output files as well as the codes to be run for the simulation of river discharge and runoff.

4. EC2: Thank you for the updated tutorial. For the purposes of presentation -- though beyond the scope of what I request as an editor -- you may want to consider using the README markdown file instead of your own txt file. This allows for better formatting and increased readability.

AC: Following your suggestion, on Github (https://github.com/IRPIhydrology/STREAM) and on Zenodo (https://zenodo.org/record/4744984, doi: 10.5281/zenodo.4744984) a markdown file “README.rst” has been added.

RC1: Comments and replies

1. R1: The paper is very thorough in presenting the novel and useful STREAM model. The paper is well-organized, clearly presented, and contains all necessary information for the reader to conceptualize and understand the design and validation. I am very interested in any future work applying this model to other global basins and its performance with respect to land surface models that require substantially more parameterizations and computational load.

   AC: The authors are thankful to the reviewer for their assessment of our paper. Really appreciated!

RC2: Comments and replies

1. RC2: This paper presents a simple data-driven model, call STREAM, to estimate global runoff using satellite observations of precipitation, soil moisture, and total water storage anomalies (TWSA). The structure of the model is simple but clear — precipitation and soil moisture are used to estimate surface quick flow while TWSA is used to compute underground slow flow. It is shown that the model can be used to estimate runoff at a basin scale after careful calibrations, which is evidenced by the validation over five calibrated sections in Figure 4.

   AC: The authors are thankful to the reviewer for their assessment of our paper. We have provided a point-by-point reply to each of the comments in the sequel.

2. RC2: However, I doubt very much whether the model can be used for the global runoff estimation since 8 parameters need to be well calibrated based on observed river discharge, which will, to a large extent, limit its application on a global scale.

   AC: The intent of the paper is to describe a model that could be used for the estimation of river flow (and runoff) worldwide. However, as correctly stated by the reviewer, the model results are only shown for the Mississippi River basin and the "Global" in the title may not be appropriate. Consistent with the reviewer's assertion, the article and model should be evaluated for their ability to estimate river flow and runoff in the Mississippi River Basin, not globally. In the revised version of the manuscript, the term "Global" from the title has been removed. The new title is: "Synergy between satellite observations of soil moisture and water storage anomalies for runoff estimation".

   The possibility to regionalize the model parameters for the estimation of river discharge at global scale is a topic beyond the scope of this article. However, the authors are working on the regionalization of the model parameters. Preliminary results, which have been shown at the EGU 2021 conference, demonstrate the possibility to link model parameter values to the basin characteristics, still obtaining satisfactory results. Future studies will address the problem of the regionalization of the model parameters. This aspect has been specified in the revised version of the manuscript (see Lines 473-474):

   "A more in-depth investigation about the model calibration procedure and the regionalization of the model parameters will be carried out in future studies."

3. RC2: For example, the validation results over the gauged sections not used in the calibration phase do not show very good performance of the model as the difference between simulations and observed river discharge may go beyond 1000 m3/s in most sections. The authors attribute this difference to the presence of dams, but this may also happen in sections without dams such as sections 3 and 7.

   AC: As any hydrological model calibrated against observed data, it is expected that the best performances of STREAM model can be obtained over the calibrated gauging sections whereas the performances will decrease over the gauging sections not used for the calibration. However, gauging section 3 and section 7 cannot be taken as reference to understand if the STREAM model is suitable for reproducing river discharge over not calibrated sections as they are affected by local characteristics (section 7 is located over the Rock river (near Joslin), a relatively small tributary of Mississippi river, 23'000 km^2, if compared with GRACE resolution, 160'000 km^2) and by the presence of an important dam (section 3).

The manuscript has been modified accordingly (see Lines 460-465):

"In particular, over sections 3 (influenced by the presence of dams in section 1 and 2) and 7 (located over the Rock river a relatively small tributary of Mississippi river see Table 1), the STREAM model overestimates the observed river discharge highlighting that the model parameters estimated for river section 4 and 6, respectively, are not suitable to accurately reproduce river discharge for river sections 3 and 7 (see Figure 3 and Figure 5)."

and (see Lines 469-473):

"Although it is expected that the performances of STREAM model, as any hydrological model calibrated against observed data, can decrease over the gauging sections not used for the calibration, the findings obtained above, raises doubts about the robustness of model parameters and whether it is actually possible to transfer model parameters from one river section to another with different interbasin characteristics. A more in-depth investigation about the model calibration procedure and the regionalization of the model parameters will be carried out in future studies."

Moreover, we underline that a decrease in the model performance is obtained for all hydrological models. We are preparing a second paper showing the comparison between STREAM and other global hydrological model performances, and STREAM model is working similarly (and even better) than other models (much more complex and with a much larger number of parameters). Preliminary results have been shown at EGU conference in 2020 (https://doi.org/10.5194/egusphere-egu2020-13718).

4. RC2: On the other hand, this paper highlights the use of three satellite observations of precipitation, soil moisture, and TWSA. However, these three components are highly correlated with each other. For example, soil moisture can be used to estimate rainfall through the SM2RAIN algorithm [1]. Another example is that, on a regional scale, TWSA is very synchronous with soil moisture [2]. Accordingly, the synergy between precipitation, soil moisture, and TWSA, to me, shall be very limited.

   AC: The reviewer is right that precipitation, soil moisture and TWSA are in some way correlated but this does not represent a problem for our approach. The proposed STREAM model is a conceptual hydrological model where the inputs contribute to the different runoff components according to specific laws. Specifically, soil moisture and precipitation contribute to the quick component of runoff (daily time scale) while TWSA contributes to the slow component (monthly scale). The differences in the temporal (and spatial) scale of the input data allow us to use the different input consistently and to optimize their synergy for runoff estimation.

   This aspect has been better specified in the text (see Lines 236-250):

   "*The concept behind the STREAM model is that river discharge is a combination of hydrological responses operating at diverse time scales (Blöschl et al., 2013; Rakovec et al., 2016). In particular, river discharge can be considered made up of a slow-flow component, produced as outflow of the groundwater storage and of a quick-flow component, i.e. mainly related to the surface and subsurface runoff components (Hu and Li, 2018). While the high spatial and temporal (i.e., intermittence) variability of precipitation and the highly changing land cover spatial distribution significantly impact* *the variability of the quick-flow component (with scales ranging from hours to days and meters to kilometres depending on the basin size), slow-flow river discharge reacts to precipitation inputs more slowly (i.e., months) as water infiltrates, is stored, mixed and is eventually released in times spanning from weeks to months. Therefore, the two components can be estimated by relying upon two different approaches that involve different types of observations. Based on that, within the STREAM model, satellite soil moisture, precipitation and TWSA will be used for deriving river discharge and runoff estimates. The first two variables are used as proxy of the quick-flow river discharge component while TWSA is exploited for obtaining its complementary part, i.e., the slow-flow river discharge component."*

Likely, the misunderstanding could have been generated as in the text the STREAM model is defined as a "data-driven model" to indicate that the model is mainly based on the contribution of the input data rather than of complex equations and processes. To avoid any misunderstanding, in the revised version of the manuscript the model has been defined as conceptual hydrological model and "data-driven" has been removed.

5. RC2: For these reasons, I suggest rejecting this paper as is.

   AC: We hope that with these new explanations the reviewer might reconsider their decision.

   RC2: [1] Luca Brocca et al (2015). Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of the SM2RAIN algorithm. HESS. [2] A Geruo et al (2017). Satellite-observed changes in vegetation sensitivities to surface soil moisture and total water storage variations since the 2011 Texas drought. ERL.

Minor comments:

6. RC2: As the experiments are only conducted over the Mississippi river basin, the word "GLOBAL" used in the title may not be suitable.

   AC: The title of the revised manuscript has been modified as: "Synergy between satellite observations of soil moisture and water storage anomalies for runoff estimation".

7. RC2: In line 122, please add necessary references regarding SMAP and GPM.

   AC: The references for SMAP and GPM have been added to the revised manuscript.

8. RC2: Please add some necessary references to Eqs. 1 and 4.

   AC: The references for equations 1 and 4 have been added to the revised manuscript. In the manuscript you can read (see Lines 278-280):

   "Consequently, the quick runoff response, Qfu from the first storage can be computed through equation following the formulation proposed by Georgakakos and Baumer (1996), as in equation (1):"

   and Lines 299-301:

"For that, the slow runoff response Qsl, from the second storage, can be computed following the formulation proposed by Famiglietti and Wood (1994), through equation (4) as follows:"

9. RC2: The statements in lines 295-296 are slightly in conflict with the statements in lines 306-308. As I know, TWSA can partly include information on soil moisture.

   The two statements have been slightly modified to underline that TWSA can partly include information on soil moisture. However, by considering the different time scales at which the quick and slow runoff responses act, in the STREAM model we are making the assumption that they are independent. This aspect has been specified in the revised version of the manuscript. In particular, the statement in Lines 295-295 (see Lines 298-299 in the new manuscript) has been modified as in the following:
   "*Indeed, the time scale of slow runoff response is typically in the range of seasons to years and it can be assumed almost independent upon the water that is contained in that upper storage*".

   The statement in Lines 306-308 has been modified as follows (see Lines 309-312 of the new manuscript):

   "Note that we made the hypothesis that soil moisture and TWSA observations are independent (whereas in the reality soil moisture can be responsible both for the generation of the quick flow part (mainly) and for the slow flow contribution) given the different temporal (and spatial) scales at which the quick and slow runoff responses act."

10. RC2: In line 303, what are the ranges of beta and m values?

    AC: A Table (Table 1A) describing the STREAM parameters, the module to which they belong, the variability range and unit has been added in the Appendix of the revised manuscript.

11. RC2: In line 344, the meaning of the Horton-Strahler order is not clear.

    AC: The sentence has been modified as (see Lines 349-350): *"by considering only rivers with order greater than 3 (according to the Horton-Strahler rules, Horton, 1945; Strahler, 1952)"* and references have been added to the revised manuscript.

12. RC2: In lines 444-445, the authors mention that the performance of model in section 3 is not bad. However, as I checked from Figure 5, the difference between simulated and observed discharge can go beyond 8000 m3/s.

    AC: On overall, the performance of the model in section 3 of Figure 5 has been defined as "not bad" because the KGE value obtained for the 2003-2016 period, between observed and simulated river discharge, is equal to 0.48. However, by looking at the time series it can be noted that sometimes the STREAM model largely overestimates the observed river discharge. This issue could be due to:

    1) observed river discharge influenced by the presence of upstream dams. Indeed, observed river discharge time series from section 1, 2 and 3 are quite similar each other and clearly influenced by the reservoir management (see e.g. flood peak on 2011 or the time series between 2012-2014); however the STREAM model does not include specific module for modelling reservoirs; 2) the model parameters are estimated for a different section, so the basin characteristics could have an impact on the model results.

This aspect has been specified in the revised manuscript (see Lines 459-464) as follows:

"Positive KGE values are obtained over river sections 3, 7 and 8. In particular, over sections 3 (influenced by the presence of dams in section 1 and 2) and 7 (located over the Rock river a relatively small tributary of Mississippi river see Table 1), the STREAM model overestimates the observed river discharge highlighting that the model parameters estimated for river section 4 and 6, respectively, are not suitable to accurately reproduce river discharge for river sections 3 and 7 (see Figure 3 and Figure 5)."

RC3: Comments and replies

1. RC3: The authors present a model 'STREAM' that is used to derive river discharge and runoff. The STREAM model is conceptually and computationally simple, and uses inputs of precipitation, total water storage, soil moisture as well as air temperature to provide estimates of global runoff. The results are tested for the Mississippi River basin in the United States and indicate good agreement. Enhancing the ability to model distributed runoff has important applications for hydrology. However, further justification of the methods used and applicability to other climate regimes and regions is needed.

   AC: The authors are thankful to the reviewer for their assessment of our paper. We have provided a point-by-point reply to each of the comments in the sequel. Additional details and justifications of the method used will be added in the revised version of the manuscript.

2. RC3: For example, the authors should comment on the sensitivity of the model to the hydrological inputs of precipitation, soil moisture and total water storage anomaly. The authors should comment on the contribution of using these inputs, and whether results are improved or not by using all three. Otherwise it is not immediately clear to the reader the contribution of each to the estimation of runoff.

   AC: In a preliminary analysis, we have tested the sensitivity of STREAM model to the different hydrological inputs of precipitation, soil moisture and total water storage anomaly (not shown for brevity). In particular, by running the STREAM model with different input configurations (e.g., by using TRMM3B42 or CPC data for precipitation, ESA CCI or ASCAT data for soil moisture, TWSA or soil moisture data to simulate the slow-flow river discharge component), we found that STREAM results are more sensitive to soil moisture data rather than to precipitation input. In addition, by running STREAM model with soil moisture data as input to simulate the slow-flow river discharge component (i.e. without using TWSA data) we found a deterioration of the model results.

   A sentence to explain the contribution of the hydrological inputs to the STREAM results has been added in the revised version of the manuscript (see Lines 498-511):

   "This is an important result of the study as it demonstrates, on one hand, that the model structure is appropriate with respect to the data used as input and, on the other hand, the great value of information contained into TWSA data that, even if characterized by limited spatial/temporal resolution, can be used to simulate runoff and river discharge at basin scale. This finding has been also confirmed by a preliminary sensitivity analysis in which the STREAM v1.3 model has been run with different hydrological inputs of precipitation, soil moisture and total water storage anomaly (not shown here for brevity). In particular, by running the STREAM model with different input configurations (e.g., by using TMPA 3B42 V7 or Climate Prediction Center (CPC) data for precipitation, ESA CCI or Advanced SCATterometer (ASCAT) data for soil moisture, TWSA or soil moisture data to simulate the slow-flow river discharge component), we found that STREAM results are more sensitive to soil moisture data rather than to precipitation input. In addition, by running STREAM model with soil moisture data as input to simulate the slow-flow river discharge component (i.e. without using TWSA data) we found a deterioration of the model results."

3. RC3: The authors also only test their model in the Mississippi River basin, however it would be interesting and informative to address the performance of this model in different regions including more arid basins, snow-dominated, lots of topography, heavily managed, etc. The study indicates it is a 'global' model so more discussion of its applicability worldwide is needed.

   AC: The intent of the paper is to describe a model that could be used for the estimation of river flow (and runoff) worldwide. However, as correctly stated by the reviewer, the model results are only shown for the Mississippi River basin and even if the model could be applied at global scale, the "Global" in the title may not be appropriate for this manuscript. For that, in the revised version of the manuscript, "Global" in the title has been removed.

   Concerning the applicability of the model to other climate regimes and regions, the authors are preparing a new manuscript where STREAM model has been tested on 5 pilot basins (Mississippi, Amazon, Niger, Danube and Murray-Darling) across the world with good model performances. Preliminary results have been shown at EGU conference in 2020 (https://meetingorganizer.copernicus.org/EGU2020/EGU2020-13718.html) and 2021 (https://meetingorganizer.copernicus.org/EGU21/EGU21-14175.html).

   More specific comments on the applicability of the model in different regions including more arid basins, snow-dominated, lots of topography, heavily managed will be addressed in a future manuscript. This aspect has been specified in the revised version of the manuscript (see Lines 594-599):

   "The application of the STREAM v.1.3 model on a larger number of basins with different climatic- physiographic characteristics (e.g., including more arid basins, snow-dominated, lots of topography, heavily managed) will be object of future studies and it will allow to investigate the possibility to regionalize the model parameters and overcome the limitations of the automatic calibration procedure highlighted in the discussion section."

Specific comments:

4. RC3: Line 208 - Can the authors provide the depth of the soil moisture used in this study.

   AC: Satellite soil moisture observations obtained by the ESA CCI soil moisture product refer to first centimeters of soil (2-3 cm). However, in the STREAM formulation we used a root-zone soil moisture product derived by the surface ESA CCI satellite soil moisture product, by applying the exponential filtering approach in its recursive formulation as in Albergel et al. (2009). For that, the root zone soil moisture used in the STREAM model is referred to the first layer of the model, whose depth varies approximatively from 5 to 30 cm.

   This aspect has been specified in the manuscript (see Lines 205-208):

   "Soil moisture data have been taken from the European Space Agency Climate Change Initiative (ESA CCI) Soil Moisture project (https://esa-soilmoisture-cci.org/) that provides a surface soil moisture product (referred to first 2-3 centimeters of soil) continuously

updated in term of spatial-temporal coverage, sensors and retrieval algorithms (Dorigo et al., 2017)."

and Lines (283-286):

"SWI is the Soil Water Index (Wagner et al., 1999), i.e., the root-zone soil moisture product referred to the first layer of the model (representative of the first 5-30 centimeters of soil), derived by the surface satellite soil moisture product, θ, by applying the exponential filtering approach in its recursive formulation (Albergel et al., 2009):"

5. RC3: Line 262 - Given that the authors only validate in the Mississippi basin, can they comment on how different climate regimes could impact the accuracy of the modeled runoff in particular more snow-dominated basins. Can the authors comment on the validation of the snow module.

   AC: The Mississippi basin contains different climatic regimes, particularly if we consider the differences between the eastern and western parts of the basin. The authors are preparing a new manuscript where STREAM model has been tested on 4 additional basins (Amazon, Niger, Danube and Murray-Darling) thus exploring a larger variability of climatic regimes, as well as soil and land use characteristics.

   The impact of snow on runoff is difficult to evaluate at the scales of the grid considered in the study (25 km). Mountainous areas which are mainly located in the north western and eastern part of the Mississippi basin (i.e., the Rocky Mountains and Appalachian Mountains) receive significant precipitation as snow during the winter period and might provide a significant portion of snowmelt runoff during the warm season. While a comparison with point stations would be not meaningful at the working scale, the river discharge simulations at section11 do show relatively good agreement with observations during both calibration and validation periods suggesting the ability of the proposed approach to simulate snowmelt contribution.

6. RC3: Line 300 - Can the authors provide additional information on how they resolve the differences in spatial and temporal scale between the various input data sets provided. In particular, the coarser scale of the GRACE data.

   AC: Concerning the spatial scale, air temperature, soil moisture, precipitation and GRACE data have been resampled over the precipitation spatial grid at 0.25° resolution through a bilinear interpolation. For the temporal scale, air temperature soil moisture and precipitation data are available at daily time step, while monthly GRACE data have been linearly interpolated at daily time step. Major details on how the differences in spatial and temporal scale between the various input data sets is resolved will be added in the revised version of the manuscript (see Lines 360-367).

   "*If characterized by different spatial/temporal resolution, these datasets need to be resampled over a common spatial grid/temporal time step prior to be used as input into the model. To run the STREAM v1.3 model over the Mississippi river basin, input data have been resampled over the precipitation spatial grid at 0.25° resolution through a bilinear interpolation. Concerning the temporal scale, T_{air}, soil moisture and precipitation data are available at daily time step, while monthly TWSA data have been linearly interpolated at daily time step. For each of the 53 Mississippi subbasins, the resampled precipitation, soil moisture, T_{air} and TWSA data have been extracted."*

7. RC3: Line 462 - can the model be run with same input precipitation as GRUN for the validation purposes? Or can the authors comment on precipitation differences between either product .

   AC: The STREAM model could be run with the same input precipitation as GRUN to compare the runoff maps obtained by the two approaches. However, as it is beyond the purpose of the paper, the authors have added only a comment in the revised version of the manuscript to underline that the precipitation estimates used in GRUN and STREAM are very similar in terms of spatial pattern and mean annual rainfall amount. In particular, both the Global Soil Wetness Project Phase 3 (GSWP3) and the TRMM 3B 42 precipitation products clearly identify two distinct zones in the Mississippi basin, a western dry and an eastern wet area.

   This aspect has been specified in the revised version of the manuscript (see Lines 479-485):

   "*Although the two datasets consider different precipitation inputs, the two models agree in identifying two distinct zones in terms of runoff, i.e., the western dry and the eastern wet area. This two distinct zones can be clearly identified also in the GSWP3 and TMPA 3B42 V7 precipitation maps (not shown here) used as input in GRUN and STREAM, respectively, stressing that STREAM runoff output is correctly driven by the input data. However, likely due to the calibration procedure, the STREAM runoff map appears patchier with respect to GRUN and discontinuities along the sub-basin boundaries (identified in Figure 3) can be noted*."

8. RC3: Line 520 - Does using GRACE data for water storage (which captures both human and natural processes) address this? GRACE can indicate human activity and water extraction practices, which I think could help improve purely 'natural' estimates of runoff.

   AC: The added value of satellite observations in the STREAM model is the possibility to capture processes and human activities not modelled but directly observed by satellite. However, we think that GRACE could not be appropriate for capturing reservoir management and for that, in the current development of STREAM model, we are including additional modules for simulating the presence of reservoirs and diversions along the river that can be relevant in several basins/regions.