

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. 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. 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 m<sup>3</sup>/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.

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

[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:

1. As the experiments are only conducted over the Mississippi river basin, the word “GLOBAL” used in the title may not be suitable.
2. In line 122, please add necessary references regarding SMAP and GPM.
3. Please add some necessary references to Eqs. 1 and 4.
4. 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.
5. In line 303, what are the ranges of beta and m values?
6. In line 344, the meaning of the Horton-Strahler order is not clear.

7. 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 m<sup>3</sup>/s.