

Review comments for the manuscript “How well are we able to close the water budget at the global scale?” by Fanny Lehmann, Bramha Dutt Vishwakarma, and Jonathan Bamber.

General

The aim of this paper is to provide a revised overview of the water budget closure on a global scale by evaluating 1694 combinations of P-ET-R. Detailed comments are given below.

Specific Concerns/Comments

1) Pages 8 and 9: The author mentioned that they applied the method of Long et al. (2014b) to estimate TWSC. However, the formula of Long et al. (2014b) shown below was different from that used in this manuscript. Please double check.

ble in Eq. (1). GRACE TWSC is computed as the backwards difference of TWSA (mm) whose reference is the mean gravity field for a calculation period (e.g., Jan 2003–Sep 2012 in this study):

$$dS/dt = \frac{TWSA(t) - TWSA(t-1)}{t}. \quad (2)$$

2) Page 9: It was stated that time-series of P, ET, and R also needed to be time-filtered by equation 3. In my opinion, P, ET, and R are fluxes, while TWS is a state variable. For one given month, P represents the total precipitation that occurs in that month. The use of equation 2 was likely due to that GRACE provides noisy monthly TWSA. It is kind of difficult for me to understand time-filtered P, ET and R. Please clarify.

3) It was concluded that TWS changes reconstructed from the water balance equation (P-ET-R) were more accurate than the long-term and monthly mean of GRACE time series in the corresponding basins. Was the GRACE-TWSC used as benchmark data in the manuscript? If so, I was confused by the conclusion that P-ET-R was more accurate than GRACE-TWSC. I may not understand it. Could the author explain it more?

4) Figure A12: “The top line of represents precipitations datasets.” I did not get it.

5) Page 24: “However, version 2.2 of this LSM, which assimilates GRACE data, performed poorly compared to its previous versions.” There was an explanation: “Since this dataset assimilates GRACE measurements and was validated against GRDC observations, this may reflect overfitting of runoff that is better constrained than evapotranspiration, therefore leading to unrealistic ET values.” Could the author explain it more? Generally, the following seems more easy to understand: when runoff became more accurate, ET would be improved accordingly.

6) Typo mistakes:

Pages 1, 21 and 23: The “Catchment Land Surface Model” was supposed to be CLSM as the abbreviation had appeared earlier.

“Fig. 2” and “Figure 2” were both used in the manuscript. Please unify the expression.