

Reply to referee #2 comments

We thank Referee #2 for the comments and positive feedback to the basin-scale water-balance dataset and the manuscript (essd-2016-33). We will address the comments (in italic in the following) point-by-point in the following.

This paper presents a global data set of monthly water balance time series of up to 35 years in length for many of the largest river basins worldwide (BSWB v2016). The data are highly valuable for understanding large-scale hydrological dynamics and for validation of land surface and hydrological models. The article is clearly structured and describes in a concise and comprehensive way the methods used for generating the data set. Some reasonable analyses and figures are provided and illustrate the information content of the data and compare them to a previous version of the data set (BSWB v2011). The data themselves are easy to access and to read, and unambiguously described in the header of the data sets and with time specifiers. My only overall concern for publication in ESSD is that the data set BSWB v2011 using exactly the same methodology has already been published by Mueller et al. (2011) with data access on a website as indicated in the manuscript. While the data set BSWB v2016 presented here clearly is a substantial extension of the previous one (with regard to number of river basins (341 versus 37 basins), time period (1979–2015 versus 1989–2008) and a more recent ERA-Interim re-analysis version), it is beyond my scope to judge whether the updated data set BSWB v2016 merits publication in ESSD according to its publication standards.

As correctly stated by the Referee, BSWB v2011 and BSWB v2016 are consistent in terms of methodology, however not in terms of data sources. BSWB v2016 is not only based on a new version of ERA-interim, it also relies on a consistent runoff data base (the GRDC reference dataset). In contrast, BSWB v2011 was based on various heterogeneous data sources for runoff (in addition to GRDC also USGS and local sources with varying data formatting and quality checks). The use of one consistent runoff data source for BSWB v2016 ensures an enhanced consistency between the basins. Moreover, future updates are more easily feasible. These benefits of BSWB v2016 are mentioned now in Section 5:

“In addition, the BSWB v2016 datasets relies on a consistent runoff data base from GRDC (i.e., the GRDC reference dataset) while BSWB v2011 was based on various heterogeneous data sources for runoff (apart from GRDC also U.S. Geological Survey and local sources with varying data formatting and quality checks). By using the GRDC reference dataset, enhanced consistency between the basins can be ensured, and future updates of the BSWB dataset are more easily feasible by relying on one source for runoff only.”

Moreover, we more explicitly mention the increase in temporal and spatial coverage in the Conclusions:

“It extends the existing version of the dataset (Mueller et al., 2011) temporally and spatially (i.e., 1979–2015 vs. 1989–2008, 341 vs. 36 river basins).”

Some minor comments on the manuscript are as follows:

1) Lines 113-114: repetition of statement on critical basin size and of three references (already given in lines 92-93). Delete or re-formulate.

This sentence has been deleted and the following sentence was reformulated to:

“Our results roughly confirm the threshold for the critical domain size for water-balance computations (10^5 km^2 , see Section 4.1) as the imbalance decreases above this basin size.”

2) Line 115: “: : :the imbalance is diminished: : :”, better: “: : :the imbalance decreases: : :” ?

Has been changed accordingly.

3) Line 121ff: *“Thus, these drifts in TWS unlikely correspond to actual variations, though the latter can be important in some regions and could contribute to part of the signal (see above).”* The sentence is contradictory in itself. There are good physical reasons for the drift as mentioned above (groundwater withdrawal, groundwater discharge to ocean) so that it should not be stated that the drifts unlikely correspond to actual variations.

We reformulated this paragraph:

“... shows a drift in TWS. The likely reason for this drift are biases in the atmospheric moisture convergence data (Seneviratne et al., 2004), though actual drifts in TWS can be important in some regions and could contribute to part of the signal (see above). The errors in the runoff measurements are expected to be small, i.e., around 5% (Winter, 1981) for longer-term averages. As widespread information on natural sources of drifts in TWS is not available, the most appropriate procedure is to assume that the observed drifts are purely artificial and remove them by a high-pass filter.”

4) line 141: *‘their’ instead of ‘its’*

Has been changed accordingly.

5) Figure 1: expand caption: *“global coverage of river basins of the BSWB v2016 data set”, or similar.*

Has been expanded accordingly.