General Comments

(This review is based on the revised manuscript uploaded by the authors on January 4th 2017.)

The authors present an analysis of the water balance of 18 catchments on the Tibetan Plateau using multiple datasets. The basis of their approach is to consider actual evapotranspiration (ET) as the main unknown in the water balance and estimate this using data on the other components. This work is interesting and potentially a useful contribution to our understanding of the hydrology of the Tibetan Plateau. However, the paper suffers from several major limitations in its current form, as described below.

1. Although uncertainty is briefly addressed in Section 3.4, I do not feel that this issue is considered in enough depth. As the authors acknowledge, there are large uncertainties associated with the datasets used in the study, as well as the method for estimating ET. I feel that these uncertainties should be incorporated into the analysis rather than addressed in general terms afterwards (Section 3.4). Otherwise the significance of the conclusions is unclear, i.e. whether the results are really an artefact of data limitations (see points below). This is my most significant concern and I feel that a fair amount of additional work and revision could be required to address this. (One variant of this approach could be to examine the errors from attempting to close the water balance with published data products, rather than forcing closure by considering ET as the residual term.)

2. It is not clear to me whether trend analysis is fully justified by the data. This is particularly the case for ET, as the authors use an approximate method to extend their ET estimates prior to the period of available GRACE data. As ET is also resolved as the residual in the water balance, any errors in the other components could be compensated for in ET estimates, which could presumably affect trend analysis. It may even be worth focusing on annual and seasonal water balance estimation in this paper and leaving the trend analysis for a more considered treatment in a separate manuscript.

3. Some of the main conclusions are stated in the abstract and summary as precipitation being the main contributor to runoff and snow water equivalent (SWE) being higher from late autumn to spring. These conclusions seem fairly basic and general. I think it may be possible to draw more substantial and specific conclusions from the analysis presented. This would stem from more focused discussion of results in Section 3.

4. The standard of English needs to be improved throughout the manuscript. While the meaning is usually (but not always) clear, there are a lot of grammatical errors (far too many to list). I suggest the authors enlist the help of someone with native-level proficiency to carefully revise the text.

Specific Comments

(Line numbers refer to the revised manuscript uploaded by the authors on January 4th 2017.)

Table 1 – It is not clear what data sources are behind all of the summary statistics for the catchments listed in this table – in particular for precipitation, temperature, NDVI, LAI and snow cover (imprecise reference). In addition, should the "Second Glacier Inventory of China" appear in Table 2 and/or the reference list? Perhaps it would make sense to introduce the catchment characteristics after the section on data rather than before.

Figure 1 – I wonder if the solid shading of catchments is really needed.

Section 2.1 – Would it be better to introduce the uncertainties associated with each dataset here? This should include observational uncertainty (e.g. precipitation undercatch (including snowfall) and discharge), as well as all model and remote sensing datasets.

Line 178-181 – What additional processing of the GRACE data products was done (regarding the glacier isostatic adjustment correction and destriping)?

Equation 2 – I am not sure that this equation is helpful. It may be better just to retain the statement in the preceding paragraph that glacier melt is a component of runoff. At the very least Q would need to

be defined differently from Equation 1 (i.e. introducing a term for non-glacial runoff), otherwise the two equations are not consistent (unless glacier melt equals zero). More generally in Section 2.2.1, I am not clear why the glacier melt contribution to catchment runoff needs to be estimated at all for the ET calculations (as observed runoff at the catchment outlets includes glacier melt).

Table 3 – As this is a very approximate estimation of glacier contribution to discharge, quoting percentages to two decimal places seems too precise.

Section 2.2.1 – I am struggling to fully understand all of the bias-correction procedure, particularly the rationale behind the second step. It would be useful if the authors could clarify this please. In addition, "m" and "a" are used in Equations 3 and 4 but I am not sure that they are defined anywhere. The uncertainties arising from extending the ET series back prior to the GRACE data period should also be considered.

Line 293-294 – The variable X should be defined.

Line 309-313 – It may be more useful to evaluate the VIC flow results in terms of biases and consistency of anomalies (monthly and annual) relative to observed discharge rather than Nash Sutcliff Efficiency (NSE), as the focus of the study is on water budgets. From Figure 2 it looks like peak flows are underestimated during "wetter" years. In addition, why does runoff simulated by VIC appear to drop to zero during the low flow season?

Line 313-317 – What data were used to force the VIC model (i.e. is there any circularity in this comparison of precipitation datasets)? The uncertainties in observation-derived and TRMM-estimated gridded precipitation products should also be acknowledged (i.e. consistency may be encouraging but neither represents "absolute truth").

Section 3.1 – In general, it would be useful to see how the different datasets look in a little more detail. For example, what magnitude of storage changes is present according to the GRACE data products? This would have big impacts on the ET calculations, so the authors need to demonstrate the credibility of GRACE data with reference to other studies (and with reference to our understanding of Tibetan Plateau hydrology to date).

Line 319-331 – Evaluating different ET products with reference to ET calculated as a residual of the water balance depends very heavily on the uncertainties/accuracy of the datasets underpinning the other terms of the water balance. The selection/rejection of ET products for further analysis in this section/paragraph is done without reference to uncertainties – so do we really know that these are the better products?

Figure 3 – This might need more clarity on what time scales are used in the analysis underpinning the figure.

Line 333-353 – Could more be made of the discussion of Figure 4? The differences in catchment properties and their relationships to climatic influences are interesting.

Figure 5 – It might be easier to interpret this figure if both the primary and secondary vertical axes used the same range (i.e. so precipitation can be compared with ET and runoff). Indeed displaying the data as a bar chart could be preferable (e.g. at its simplest, separate bars for precipitation, ET, runoff and implied storage change).

Figure 7 – Some of the VIC SWE estimates look unrealistic for some of the catchments and the range of scales on the secondary vertical axis complicates interpretation. I am not sure whether conclusions on SWE can really be drawn from this dataset.

Section 3.2 – I am not sure if full use is made of the figures and their underpinning analysis in this section generally. I think more focused discussion of the results of the water balance (annual and seasonal) should be possible.

Section 3.3 – Some of the discussion in this section seems speculative, particularly regarding the relationships of calculated trends with climate indicators. This is not a simple subject and I suggest

that this section should be worded more carefully and discuss the drivers of trends in less definitive terms.

Section 3.4 – As discussed above, I do not feel that this is a sufficient treatment of uncertainty (see general comments).

In addition, some of the references in the text appear to be misspelt (e.g. line 66 Immerzeel, line 99 Harris). I suggest that all references are carefully checked.