Reply to reviewers' comments

Dear Editor and Reviewers:

We thank the handling Editor and the reviewers for your comments and suggestions concerning our manuscript entitled "Precipitation Pattern in the Western Himalayas revealed by Four Datasets" (Manuscript Number: hess-2017-296). These comments are all valuable and very helpful not only for improving this paper but also beneficial for our research in general. We have carefully studied these comments and will address them in making revisions. The point-by-point responses to each of the comments are presented as follows.

Anonymous Referee #1

Received and published: 4 July 2017

Major comment 1: The methodology of attempting to distinguish precipitation trends from the four types of dataset is not scientifically valid. Firstly, there is no attempt to calculate the inter-annual variability of precipitation for each of the datasets. Therefore, it is impossible to tell whether the changes between the 2003-2007 and 1981-1985 periods are meaningful. There was also no justification for why these periods were even chosen. It is also impossible to tell whether choosing five years for each period is large enough to capture the precipitation representative of the 1980's and 2000's, i.e. these periods could easily have been anomalous. There is also no physical justification given as to whether the e.g. increase in summer precipitation during these two periods is physically consistent with either dynamic or thermodynamic large-scale changes, such as in response to the observed weakening of summer monsoonal precipitation (e.g. Bollasina et al. 2011). The differences also aren't quantified, and are referred to in the Abstract as 'an increase in summer and a decrease in winter with large variations'. The fact that there are large differences between the datasets is also a cause for concern, with the differences between datasets being possibly greater than the magnitude of the differences between the two periods.

Reply: Thanks for the reviewer's comments. There are three questions and we answer each individually.

For first question, we will follow the reviewer's advice, we will calculate and compare the inter-annual variability of precipitation for each of the datasets, a discussion of which will be included in the revision.

For second question, we agree that it is not clear why the 2003-2007 and 1981-1985 periods are chosen, and whether they are representative. As we stated in the manuscript that there are large differences among the datasets with respect to seasonality and spatial pattern. Such differences have implications for hydrology and glaciers. In the revised version we will follow reviewer's advice to provide more in-depth discussion on this aspect. We will make more comprehensive analysis (annual and seasonal variability), and extend the analysis to the whole study period with moving time-slices. Such analysis will give robust results and uncertainty existing in trend analysis.

For third question, we will look at some other variables, such as wind speed and air pressure in the ERA-interim to find possible causes. We will do a more thorough literature study on the changes in the Western Himalayan Region and to find if the changes will continue in the future.

Secondly, the authors showed that the linear precipitation trend was insignificant, which surely contradicts their claim that precipitation patterns have changed with time.

Reply: Sorry that we failed to state that clear enough in the original version. Linear trend for yearly total precipitation is not significant, but precipitation patterns as reflected by seasonality and spatial variability have changed more. We will provide more explanation on this aspect in the revised version.

Thirdly, Figure 6 shows little agreement in the trends during 1981-2007 for the four datasets, with large differences in magnitude as well as even differences in sign.

Reply: This is true. As we mentioned in the reply of previous comment, there are large differences among the datasets with respect to seasonality and spatial pattern. Such differences have implications for hydrology and glaciers. We will put more effort in the revised version to provide more in-depth discussion on this aspect. Additionally, we will do trend analysis at seasonal scales for more robust results. The overlap period of the four datasets is 27 years. However, the ERA-interim and the IMD datasets cover respectively the periods of 1979-2016 and 1951-2007. We will look at the long-term changes of these two datasets to see if the trends of the study period agree with the long term trends.

Bollasina et al., Anthropogenic aerosols and the weakening of the South Asian Summer Monsoon, Science, 2011.

Major comment 2: The description of the datasets is poor and overly generalised, and does not focus enough on the study regions. For example, the description of the IMD dataset does not say explicitly how many stations are used in the study region, and what altitudes. Instead, vague language such as ‘less stations near the borders of India and the in the northern part’ are used. This is insufficient information to make any robust judgement of the veracity of the data. This type of vague description is continued for APHRODITE. For example, in the description of APHRODITE, evidence of its representativeness of precipitation distribution is given by the claim that it is better than the MRI/JMA AGMC model – when in fact it is data that should be used to ground-truth models, and not the other way around. My understanding of both these datasets is that due to the sparcity of gauge measurements in the Himalayas, and particularly the lack of measurements at high altitudes, that these datasets are highly biased. In the description of ERA-Interim, it is stated that ‘the spatial resolution of the ERA-interim dataset is limited in representing the spatial variability’. If this is not representative of precipitation, then why is it being used? Moreover, a vague statement that ‘precipitation is adjusted based on GPCP v2.1. before release’ is included. What is GPCP data? How does this affect the representation of precipitation over the Himalayas in ERA-interim? None of these questions are answered.

Reply: Thanks for the professional and constructive comments. We agree and we apologize for the vague description and we will add more details in the revised version on the description of all the datasets used in the study.

Finally, it was odd that the WRF model run used the configuration of EURO-CORDEX, rather than that recommended by e.g. papers by Maussion et al. (2011) or Collier and Immerzeel (2015) which focused on the Himalayas. This unfortunately gives the impression that the authors were using the model has a 'black box', and had little understanding of regional atmospheric modelling. This is reinforced by statements such as 'The ERA-Interim and WRF datasets are products with different dynamical models' and referring to both of these as 'the products from dynamic models do not suffer from an undercatch', which suggests that the authors aren't properly aware of the considerable differences between reanalysis products and numerical weather prediction modelling. The claim that the model run was not optimised was 'due to the complex orography' is unfounded, as studies such as Maussion et al., Collier et al. have shown that the choice of model grids and physics parameterisations is critical. There are also no details as to the spatial resolution of the WRF model, and claims such as 'the climate model has been proved to produce the regional precipitation at a fine scale' are for models running at kilometre scale for small regions around 100 km in size (See Collier and Immerzeel).

Collier and Immerzeel, High resolution modelling of atmospheric dynamics in the Nepalese Himalayas, JGR, 2015.

Reply: Thanks for the comment, and sorry that we did not state it clear enough. In the revised version we will provide more discussion concerning the model configuration and also explain the difference with the nice work of Maussion et al. (2011). The WRF model runs at 0.15◦×0.15◦ grid. Maussion et al. (2011) indeed did a very good comparison of model configuration at the Tibet Planet (TiP) area, which is around 10 degree eastern of our study area. However, they conclude, "Our study reveals that there is nothing like an optimal model strategy applicable for the high-altitude TiP, its fringing high-mountain areas of extremely complex topography and the low-altitude land and sea regions from which much of the precipitation on the TiP is originating. The choice of the physical parameterization scheme will thus be always a compromise depending on the specific purpose of a model simulation. Our study demonstrates the high importance of orographic precipitation, but the problem of the orographic bias remains unsolved since reliable observational data are still missing". Our WRF configuration has also been used in other projects, which focus on the western Asia and our results show that this configuration is reliable at the study area. We use the same microphysics and land surface as their reference experiments and Li et al. (2017) has tested the microphysics, cumulus and land surface scheme. This part of the text will be largely rewritten in accordance with reviewer's comment and advice.

Li, L., Gochis, D. J., Sobolowski, S., & Mesquita, M. D. S. (2017). Evaluating the present annual water budget of a Himalayan headwater river basin using a high-resolution atmosphere-hydrology model. *Journal of Geophysical Research: Atmospheres*, *122*(9), 4786–4807. article. https://doi.org/10.1002/2016JD026279

Major comment 3: The Abstract begins by saying that 'data scarcity is the biggest problem . . . in the Himalayas' and that 'high quality precipitation data are difficult to obtain'. Yet the paper never properly addresses which of the four datasets is, despite their deficiencies, best able to represent Himalayan precipitation patterns. This might have been a worthwhile objective. Indeed, the abstract states that 'all the datasets can give a good overview of the precipitation'. How can this be possible when one of your datasets is ERA-Interim and another is WRF-based downscaling of ERA-Interim? It is unclear how this conclusion is reached, other than the broad generalisation that all the datasets show a wetter summer compared to winter. Moreover, does this mean that all datasets would give broadly the same answers if they were used as input to hydrology models? Additionally, many of the findings are well known, such as 'the highest precipitation locates at the foothill of the mountains and stretches from southeast to northwest'. Some of the results seemed distinctly unoriginal. The authors cite the Bookhagen and Burbank (2006) study, which did a very thorough job of describing precipitation characteristics in the Himalayas. I was unsure whether one of your aims was to show which datasets could recreate their findings? Also, any results which claimed to show something original, such as changes in precipitation, were highly flawed (see comment above).

Reply: We agree with reviewer's concern, and we apologize that we failed to address our important objectives clear enough, i.e. to look at differences among datasets and their implications on hydrology and glaciers. To achieve this, we will add a new section, in which we will select available catchments in The Global Runoff Data Centre (GRDC, http://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html;jsessionid=0F83978153B3C02214DBA224F2084914.live21302) and the Integrated Hydrological Data Book (CWC, http://www.cwc.nic.in/ISO_DATA_Bank/ISO_Home_Page.htm) within the study area. We will analyze runoff data and actual evaporation from MODIS. We will answer the question like which datasets are more reliable in water balance assessment and how much is uncertainty.

Major comment 4: The manuscript is poorly organised, and lacking depth and understanding of the topic. For example, much of the results section is filled with material which should have been in either sections 1 or 2. The authors cite the study of Li et al. (2016) as indirectly proving that the WRF model can realistically simulate precipitation as it was able to force a hydrological model which was able to simulate discharge values. However, possibly the hydrology model was tuned to get this result? Moreover, the WRF model is highly sensitive to choice of physics and model setup (as the study by Maussion et al. (2011) shows), so much more details should have provided of how your model setup agrees with that of Maussion et al. This again illustrates that the authors are not suitably experienced in modelling to have included the WRF output.

Reply: We will reorganize the structure of the paper, into Introduction, Study area and data, Methodology, Results and Discussion, Conclusions. In the revised version, the Results will be divided into two sub-sections. One is on precipitation data, which consists of the current sections 1 and 2. Another is on implications about impacts, which consists of the current section 3 and a new section, which is usage of runoff and actual evaporation of selected catchments. The WRF-hydro is physically based land-surface process model. The tuning is used to find correct values for parameters, which is due to lack of measurements of vegetation, soil and river channel characteristics. In the revised version, we will give more details of the WRF model setting and will provide the configuration file as support material.

Anonymous Referee #2

Received and published: 5 July 2017
Precipitation data are a key input in hydrologic modeling and the present paper compares four precipitation datasets which include data obtained by ground based measurements, interpolation, and reanalysis data. The paper addresses an important issue faced by hydrologists. I have the following comments on the paper.

Reply: Thanks for reviewer's positive evaluation in general, and the specific comments which are detailed below.

1. The strengths and weaknesses of the four type of gridded precipitation datasets explained in lines 15 to 30 on page 2 can be better explained by means of a table. Each row of this table may correspond to a particular dataset and the columns could the how the data is obtained, its strengths, and weaknesses.

Reply: We thank the reviewer for this good suggestion and we will add a table following the advice of the reviewer.

2. In this paper, IMD dataset at 1 degree grid has been used. Currently, data at 0.25 degree resolution are also available.

Reply: This research is based on 1-degree grid data and we obtained this dataset from our Indian research partner. This resolution is comparable with other datasets used in the study. Thanks for introducing the 0.25° IMD dataset, which is not available for the authors at this moment and will be considered in the future study.

3. There is a view that the precipitation data obtained from instrumented stations does not reveal the actual values over a catchment in Western Himalaya because the network of stations does not have the desired density and most stations are located in valleys. Thus, the actual precipitation in the hill tops is not known.

Reply: The actual precipitation is largely unknown at the hill tops. Some methods are available to partly overcome this problem. One is to interpolate the ground observations with elevation and undercatch correction. Numerical models are also valuable to estimate precipitation at high mountains, especially where observations are not reliable or not available. This manuscript is an attempt to compare four datasets produced by different methods.

4. Authors mention that the four datasets are similar in terms of spatial and temporal variation but there is very large variation in absolute values from 497 to 819 mm/year. Given this, a reader would expect clear view from the authors: a) what is their assessment of mean annual

rainfall, and b) which dataset(s) can be used in applications such as water yield assessment, flood forecasting, climate change impact assessment, and so on.

Reply: We agree these questions are very important and we must give a clear view. To achieve this, we will add a new section, in which we will select available catchments in The Global Runoff Data Centre (GRDC, http://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html;jsessionid=0F83978153B3C02214DBA224F2084914.live21302) and the Integrated Hydrological Data Book (CWC, http://www.cwc.nic.in/ISO_DATA_Bank/ISO_Home_Page.htm) within the study area. We will analyze runoff data and actual evaporation from MODIS. We will try to answer the question like which datasets are more reliable in water balance assessment and how much is uncertainty.

Anonymous Referee #3

Received and published: 6 July 2017

This paper is dealing with an interesting data challenge in Himalayan region through comparison of four globally available gridded precipitation data sets. Though this work is interesting and publishable in terms of regional importance of Himalaya, illustrations in present format are not very strong.

Reply: Thanks for reviewer's positive evaluation in general, and we appreciate very much the constructive comments which are detailed below.

Major demerits are 1. This paper looks like a quick dissemination with limited analysis from authors (some figures just direct illustration of latitude wise and seasonal raw data).

Reply: We will add more details in response to specific comments and a new section about uncertainty based on runoff and evaporation. More details of our response are given below.

2. Lack of appropriate inter-comparison technique which ensures comparability of spatial patterns with different grid size.

Reply: The precipitation datasets are at different grid size and it is quite challenging. The results show in forms of spatial maps, individual grids and regional mean. We will additionally use histogram as well as different quantiles to discuss in more details of the results.

3. Lack of detailed discussions on spatial patterns and issues of scale.

Reply: Agree. We will add more details about the spatial pattern. More details about issues of scale will be added in the section of spatial pattern as well as in a new section about uncertainty based on runoff and evaporation (MODIS) data for available catchments in The Global Runoff Data Centre (GRDC, http://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html;jsessionid=0F83978153B3C02214DBA224F2084914.live21302) and the Integrated Hydrological Data Book (CWC, http://www.cwc.nic.in/ISO_DATA_Bank/ISO_Home_Page.htm).

4. Lack of verification indicators (example: POD, TS, FAR, FBI etc.) in comparison with available rain gauge data or IMD data as reference.

Reply: Thanks for the comment and suggestion. Some of the verification indicators as used in Sikder & *Hossain (2016*), i.e., POD: the probability of detection. TS: threat Score. FAR: the false alarm ratio. FBI: the frequency bias index, will be used in the revised version. These indexes were used to assess forecasting results with observations. Here we do not have "true observations"- rain gauge data. The IMD dataset will be used as reference although it is interpolated by the India Meteorological Department from rain gauge data. During the period of 1981-2007, the average number of stations per grid point (1°×1°) varies from 0.2 to 4.4. Fewer stations near the borders of India and in the northern part of the study area and observations are available near the latitude of 35.5°N and its north.

Sikder, S., & Hossain, F. (2016). Assessment of the weather research and forecasting model generalized parameterization schemes for advancement of precipitation forecasting in monsoon-driven river basins. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1210–1228. article. http://doi.org/10.1002/2016MS000678

Some specific comments are given below:

As this work is dealing with comparison gridded data sets with varying spatial resolution and other features, the discussion needs to be strengthened incorporating these uncertainty aspects and how reliable/meaningful these absolute precipitation values are, which are used for comparison.

Reply: Thanks for the comment. To address this issue, in the revised version we will add a new section, in which, we will select available catchments in The Global Runoff Data Centre (GRDC, http://www.bafg.de/GRDC/EN/01_GRDC/grdc_node.html;jsessionid=0F83978153B3C02214DBA224F2084914.live21302) and the Integrated Hydrological Data Book (CWC, http://www.cwc.nic.in/ISO_DATA_Bank/ISO_Home_Page.htm) within the study area. We will analyze runoff data and actual evaporation from MODIS. We will answer the question like which datasets are more reliable and how much is uncertainty.

Need some consistency in acronyms used in this text (e.g.: authors have used both AGMC and AGCM)

Reply: This is a typos. We will revise it in the revised version.

Bit more clarity is needed in the description of seasons in the study region (e.g.: in Page 3 authors have considered November-April period as winter, and May to October as summer).

Reply: We apologize for this ambiguity. Here it refers to summer monsoon season and winter monsoon season. The analysis is based on northern meteorological seasons (spring: March to May; summer: June to August; Autumn: September to November; Winter: December to February). We will revise it in the revised version.

More clarity and justifications through references are required to strengthen "do not suffer from the undercatch problem" explained in the page number 3.

Reply: Thanks. In the revised version we will clarify it. Undercatch means the amount of precipitation is not measured by gauges mostly caused by wind turbulence especially for snow. The undercatch has a significant effect for unshielded and single gauges at high latitude and high altitude. The undercatch applies to most types of precipitation observations and their products. Here we meant that Regional climate models simulate precipitation based on mathematical equations and the results are considered not affected by undercatch problem.

Better clarity is needed on the selection of 5 year time slices (1981-1985, 2003-2007) for comparison.

Reply: We will clarify that the trend is analyzed by the Mann-Kendall test. It shows winter became drier and summer became wetter during the period of 1981 to 2007. However, few (May by the WRF dataset; June by the IMD and ERA-interim datasets) of them are statistically significant at the 95% confidence level (Figure 6 and Page 7). We believe the trends will have negative impact on mass balance of glaciers and comparison between the first and the last five years will give a quantitative picture of the impact. But in the revised version we will use the whole period to quantify the changes.

Why seasonal comparison is limited to selected two months (Page 6) only?

Reply: The seasonal comparison is shown in Figure 3 whereas these months are purposefully selected to have a better visual impression of spatial distribution and interaction with topography. JA (July and August) and ND (November and December) have respectively highest and lowest precipitation and these months clearly show the changes as illustrated in Figure 5.

Page 2 Paragraph 2. This part of text emphases more on to a specific project and associated difficulties- it appears to limit the scope of this work. It would be more appropriate if you define these site-specific difficulties and rationale of this paper through proper references from Himalayan region than describing it as a 'INDICE project' related issue.

Reply: We agree with the reviewer and will revise the manuscript accordingly.

I am doubtful about the usefulness of west-to-east latitude based precipitation comparisons for hydrologists in the region. It would be more useful for the hydrologic research community if you could include precipitation PDFs comparisons of larger river basins (Indus River and the upper Ganges River) in this study domain

Reply: This figure aims to show interactions between atmosphere and topography in JA (July and August, high precipitation months) and ND (November and December, low precipitation months). As shown in the figure, the precipitation changes at different locations and the changes vary among the four datasets. We agree with the reviewer that precipitation PDFs for catchments are very useful for hydrologists and we will add such analysis in the revised version.

Titles of Figure 7 and 8: you need to clearly write the details of data points with different colors.

Reply: In Figure 7, the blue color shows the period 1981-1985 and the orange shows the period 2003-2007. These are shown as in the legend. In Figure 8, the colors differentiate between below or above zero. We will add more details in the revised version.