Interactive comment on "Precipitation Pattern in the Western Himalayas revealed by Four Datasets" *by* Hong Li et al.

Anonymous Referee #1

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

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. 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.

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 hydrol-

ogy 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).

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.