The paper and related dataset are interesting and present significant effort in evaluating and monitoring flood events. However, the methods description and error assessment of the dataset are somewhere too vague and uncertain. The main concerns regarding various parts of the manuscript are presented below.

1) GPM, that is mentioned as the single source material for precipitation (section 2.2), has been operational since 2014. Also, precipitation series are necessary for the flood dataset construction (figure 1). However, obtained flood dataset is cover period from 2002 to 2016. It remains unclear how the dataset was obtained for 2002-2014.

2) As it was mentioned in the manuscript (section 2.1), the research used GRACE dataset that is based on set of spherical harmonic coefficients up to 40 degree and power. That is roughly equal to spatial resolution 20000/40=500 км or 5°. At the same time, GPM spatial resolution is 0.1°. In manuscript it was mentioned "we take the maximum values of the precipitation data under the GRACE grid coverage to further calculate the flood potential index and the number of extreme precipitation days". Such difference in spatial resolution between the datasets makes me wonder if the maximum is the best metric in this case.

3) The main disadvantage of this manuscript, which does not allow a full and clear assessment of the resulting data set, is absence of false alarm ratio or some other similar metric to understand how often received dataset falsely detect flood event.