First of all we want to thank the reviewer for his/her valuable and thoughtful comments. Following, we will reply to each of the comments made.

**General comments:**

This paper is a very interesting analysis of the relevant processes and triggering factors that lead to widespread extreme floods in a large region, e.g. in Germany. Due to very detailed (in time and space) comprehensive data sets on precipitation, whether types and river flows, it was possible to analyze the initial and the actual conditions of rainfall and runoff in a homogeneous way all over the time of more than 50 years. This was done, I think, for the first time in this extensive manner. The comparison of three large flooding events (1954 2002, 2013) lead to the result, that these three cases, different main reasons were responsible for the development of widespread floods over Germany and their neighboring countries.
Specific comments:

Referee Comment:

Soil moisture/wetness: several times (among others in the abstract) the term soil moisture or wetness is used in a way that the reader thinks, there are some data on soil moisture. In the text however, it is mentioned, that API is used as a proxy for soil wetness. I think we must be careful. API is indeed a proxy, but only for the potential of initial conditions of soil moisture conditions. The quality of this proxy depends a lot on soil structure and soil depth as well as on land use. And additionally it depends on the length of the API-period. Did you perform any sensitivity analysis on this?

Authors’ response:

We agree that any confusion concerning the use of soil moisture data has to be avoided. We checked the text and changed this accordingly.

Certainly, soil and land use characteristics control hydrological processes relevant for flood generation. In our study we use API as a proxy for wetness conditions before the onset of large scale floods in Germany. Depending on the site specific soil characteristics antecedent precipitation might exceed soil water storage capacity. However, in the context of our study we do not target at a comparison of regional differences in flood generation throughout Germany. Instead we evaluate the wetness situation before the historic large scale floods. API represents accumulated precipitation in the time period before the onset of the flood and as such indicates potential wetness in the catchment. For a relative comparison of different large scale flood events this information is sufficient.

We think the following sentence in 2.2.3 Antecedent precipitation makes this clear:

“The API is used as a proxy for wetness conditions in a catchment in the period before the event precipitation.”

Within the sensitivity analysis we used different length of API periods (30 days and 15 days). The results show that the implications of this variation for the ranking of the wetness indices for the set of flood events are rather low.

Referee Comment:

Initial hydraulic load: This is certainly a very important factor. But I feel it is somehow redundant/dependent to API. Did you check correlations?
**Authors’ response:**

Intuitively high API should be related to increased streamflow in the river network. On a small scale this relation should be stronger than on a larger scale where additional interfering factors become important, e.g. hydraulic processes such as translation and retention of discharge in the river network. Further, the initial hydraulic load index I accounts for seasonal variation of flow levels in the river network as for instance increased discharge due to snowmelt which is not included in the API based proxy and the wetness index W. The correlation between I and W indices among the large-scale flood event set is low (r = 0.25).

**Referee Comment:**

Drainage basins surface and time resolution: It is mentioned, that drainage basins are used from 500 km² upwards. Unfortunately here is no overview on the distribution of drainage basins surfaces. At the other hand the Authors’ use daily means of discharge in the analysis. In my view, flood peaks cannot be detected/assessed in basins smaller that about 3-5000 km² with daily runoff resolution.

**Authors’ response:**

The set of gauges has been adopted from Uhlemann et al., (2010). Therein the reasoning for the selection of gauges and also information about the distribution of drainage basins is provided. We have included some details on the distribution of drainage basins in the paper.

Concerning the detectability of flood peaks in time series of daily mean discharges we do not fully agree with the reviewer because catchment size is not the only relevant factor in this regard. Also other characteristics as for instance topography and hydraulic regime play an important role. Records of daily mean discharges from gauges with similar basin areas have been successfully used for several flood related studies in Germany, e.g. Uhlemann et al. 2010, Beurton and Thieken 2009, Petrow and Merz 2009


Technical corrections

Referee Comment:

Chapter 2: This chapter should be reorganized. 2.1 should be “Data”, following by 2.1.1 etc., and then (as it is) 2.2 “Methods”. Now, in 2.1 and 2.3, both are dealing with hydrological data (floods)

Authors’ response:

We changed the organization of chapters accordingly.

Referee Comment:

Page 8130, line 3 ff: which classification was finally used by the Authors’?

Authors’ response:

We decided to delete this paragraph with the large-scale weather patterns because it provides no useful insight.

Referee Comment:

Chapter 2.4.2: How the analysis on precip. was done for the 1954 event?

Authors’ response:

For this event (and the 2013) we also used the REGNIE data; we changed the text to clarify this.

Referee Comment:

Page 8136, line 17: Do you have references for this statement?

Authors’ response:

During the changes made to the manuscript this sentence has been deleted.

Referee Comment:

Page 8137, line10: “Highest precip. …” Compared with what? I cannot see this on fig 3a

Authors’ response:

This is related to the time period of 30-days shown. In the Elbe catchment, the areal rainfall 3 days ahead is approx. 75 mm; this is the highest value during the whole period. We reformulated this sentence to make it clearer.
Referee Comment:

Chapter 3.3.1: API: can we really compare API between regions or between different events. It might be a problem, that water storage capacity in the soil is smaller than API, so we compare high API that is no realistic and therefore not relevant.

Authors’ response:

Please refer to the specific comment above.