__**Response to Referee 1:**__

In this paper, a non-stationary approach is applied for the estimation of the probability of failure of infrastructures in two locations in the Upper Truckee River Basin (US). The approach uses climate scenarios as input to determine the expected (and range of) changes in precipitation and temperature. The results show that, based on the assumptions made, the probability of failure of infrastructures increases considerably with time, from now to the end of the century.

The paper is well written and interesting. I like the fact that the evolution of probability of failure is investigated, instead of the change in return period. However I have the following concerns, which I believe should be addressed/discussed before publication in HESS:

Response: We thank the reviewer for their thoughtful review. We have addressed all of the specific comments below.

1) What is the novelty of the paper? Non-stationary models for flood hazard are not new and nor is the use of the probability of failure in climate change studies (i.e., the "design life level" of Rootz and Katz, 2013).

Response: We appreciate the reviewers comment and would like to note that we are not claiming that this is a new method. Rather the purpose of this work is to demonstrate how this approach can be applied. In their both their 2013 and 2014 papers, Salas and Obeysekera comment that, although non-stationary flood methods have been well documented there has been 'insufficient' attention paid in the water resources engineering literature. Our goal with this paper was to demonstrate the application of these methods for regional decision making. We have modified the end of the introduction to make this point more clearly as follows:

> ***"To address this issue, Rootzén and Katz [2013] introduced the concept of design life level to calculate the risk of a given flood magnitude occurring over a specified time period. Salas and Obeysekera [2014] further demonstrated the relevance of this technique to hydrologic community using flood frequency examples. However, this methodology has yet to receive widespread attention within the hydrologic community. Here, we present a non-stationary flood frequency assessment for the Upper Truckee River Basin (UTRB) using contemporary downscaled climate projections and the non-stationary design life level technique introduced by Rootzén and Katz [2013] to quantify flood risk (Note that following the convention of Rootzén and Katz [2013] we use the term flood risk in the non-technical sense to refer to the probability of an extreme event occurring and not as a quantification of expected losses). While the methodology used for this analysis is previously established, this paper provides the first end-to-end demonstration of non-stationary GEV analysis coupled with contemporary downscaled climate projections (specifically, downscaled climate projections from the Coupled Model Intercomparison Project Phase-5 (CMIP-5)), to quantify how the flood risk profiles may evolve in the Truckee river basin over the next century. The intent of this work is 1) to investigate potential flood risk changes over time in the Truckee basin and 2) to demonstrate the applicability of non-stationary techniques in a regional flood analysis to make these tools more accessible to the hydrologic community." (Revised Manuscript, Page 6, Lines 11-28)***

2) The results are conditioned to strong assumptions and there is no explicit uncertainty analysis in the paper (e.g., Steinschneider et al., 2012, provide a framework for that). Prediction bounds are plotted in the figures, but they just show the range of variability of climate model inputs once propagated through the hydrologic models (VIC + non-stationary GEV). In my opinion, it would have been more interesting to analyse if, based on the observations in the last decades, the use of non stationary flood-frequency models gives results significantly different from those obtained with stationary models. To do so, the uncertainty associated to the use of both approaches should be quantified: Fig. 6 could contain the stationary models results with confidence bounds + the non-stationary model results for the observation period with confidence bounds that account not only for the variability of the climate models, but also for the uncertainty in the estimation of (VIC+ non-stationary GEV) model parameters. It would be very interesting to see how the two ranges of estimates differ.

Response: We appreciate the reviewer's suggestion and agree that our approach only encompasses uncertainty in the climate model inputs. We agree that a rigorous analysis of different sources of uncertainty would be interesting however this outside the goals of our current study. Recent analysis by Elsner et al. (2014) did investigate uncertainties in historical forcings and their impact on VIC simulations. In response to this comment, we have added the following clarification to section 4.3 of the manuscript:

> ***"Here we use the interquartile range, as opposed to the 5th and 95th percentile, to focus on the central tendencies of each time period and not the variability between projections. Note that the ranges presented here express the variability between climate models. Uncertainty in the historical data sources used for calibration and in the parameters of the VIC model are not investigated directly here. For more detailed analysis on uncertainty in VIC simulations the reader is referred to Elsner et al. [2014]." (Revised Manuscript, Page 17, Lines 17-23)***

3) The Authors use the wording "flood risk" to refer to the probability of failure. Even though in engineering books "risk" and "probability of failure" are used interchangeably, "flood risk" is widely accepted in the literature as product of hazard (probability of flooding) and consequences (see e.g. Plate, 2002, among many). Since this paper looks at hazard only, I would strongly suggest to change the wording in it (including the title).

Response: We use the term 'risk' to be consistent with the design life literature we are citing (e.g. Salas and Obeysekara, 2014). However we acknowledge that the term risk can be used several ways. In response to this comment we have added the following clarifications to the manuscript:

> ***"The resulting exceedance probabilities are combined to calculate the probability of a flood of a given magnitude occurring over a specific time period (referred to as flood risk) using recent developments in design life risk methodologies." (Revised Manuscript, Page 2, Lines 10-12)***
>
> ***"Here, we present a non-stationary flood frequency assessment for the Upper Truckee River Basin (UTRB) using contemporary downscaled climate projections and the non-stationary design life level technique introduced by Rootzén and Katz [2013] to quantify flood risk (Note that following the convention of Rootzén and Katz [2013] we use the term flood risk to refer to the probability of an extreme event occurring and not as a quantification of expected losses)." (Revised Manuscript, Page 6, Lines 15-20)***

> ***"This concept is easily extended to flood risk (here defined as the probability of a flood of a given magnitude occurring, not expected loses)." (Revised Manuscript, Page 13, Lines 30-31)***

4) The references in the paper are biased toward US, while relevant literature exists abroad. As a suggestion, since I am European, the Authors could refer to some of the many studies cited in Hall et al. (2013) about flood changes in Europe (and scenario approaches).

Response: We thank the reviewer for pointing out this reference. We focused our introduction on studies in the Western US, as this is where our study area is. However, we acknowledge that there is also much to be learned from studies of other regions of the world. In response to this comment we have added the following reference to the introduction:

> **"Difficulty in diagnosing flood trends is not unique to the Western US; a literature review of historical flood studies across Europe also found spatial variability in flood trends [Hall et al., 2014]."** ***(Revised Manuscript, Page 4, Lines 9-11)***

Also, through Hall et al. we found Merz et al. [2012] which provides a useful discussion on the drivers of change. We also added the following later in the introduction:

> ***"Merz et al. [2012] note that attributing changes in flood hazard is complicated by the complex array of drivers that can include; land cover change and infrastructure development as well as natural climate variability and change. Here we set aside the impacts of development and management practices and focus on the role of climate change." (Revised Manuscript, Page 4, Lines 24-26)***

<u>Specific comments:</u>

Page 5079, line 11: I do not agree with the wording "additional non-stationarity". It does not make sense unless stationarity is defined (see e.g., Koutsoyiannis, 2006; Montanari and Koutsoyiannis 2012). Under my understanding, stationary models can cope with long-term climate oscillations (see e.g., Koutsoyiannis, 2011).

Response: We agree with the reviewer that this sentence was unclear and have changed it to read:

> ***"Antropogenic climate change has the potential to amplify natural climatic variability throughout the interconnected climate and hydrologic systems" (Revised Manuscript, Page 13, Lines 13-15)***

Page 5083, line 7: being HESS international (and European) international unit system (e.g., km instead of miles) should be used. This comment applies to the all paper.

Response: We have changed all units to SI.

Page 5088, Eqs. (2) and (3): one line could be added to motivate why the shape parameter xi is considered stationary.

Response: We have added the following clarification following equations 2 and 3:

> ***"In keeping with previous studies the shape parameter, which is the most difficult to estimate, is assumed constant [e.g.* Obeysekara and Salas, 2014; Salas and Obeysekera, 2013; Towler et al., 2010*]." (Revised Manuscript, Page 12, Lines 3-6)***

Page 5088, line 18: are the GEV distributions fitted to simulated streamflows only? The Authors should add a line here to motivate why the observed streamflows are not used here. I see that Fig. 2 and 3 include observed flows and provide a kind of validation of the procedure.

Response: Unfortunately we only have unregulated flow estimates for six flood events which would not be enough to fit the GEV distributions alone. We decided not to mix simulated and 'observed' flow datasets in order to be consistent. The unregulated flows are estimated by adjusting observed gauge flows for water resources management, i.e. water deliveries to meet agricultural water demands and reservoir storage changes. There are fundamental differences in the underlying processes (and assumptions) to estimate unregulated flows from gauge data and to simulate flows using a hydrologic model. Hence, it is not appropriate to combine flows from these two disparate sources. In response to this comment we have added the following clarification to the text.

> ***"Although, there are some unregulated historical flow estimates, the available dataset only covers six storms. Therefore, to be consistent we fit our model only to the simulated flows." (Revised Manuscript, Page 12-13, Lines 28 & 1-2)***

Page 5089, Section 3.2: the section discusses "flood hazard", not "flood risk". The same applies to the rest of the paper, specially to Section 4.3 and related figures (e.g., y-axis in Figs. 5-8 should not be "risk" but "probability")

Response: Please refer to the response to comment 3. We have decided to use the term 'risk' to be consistent with the terminology of the design life level we apply. However, to avoid confusion we have added clarification throughout the text. Also in response to this comment we have changed the captions of the figures in question.

Additional References:

Hall, J., et al. (2013) Understanding flood regime changes in Europe: a state of the art assessment, Hydrol. Earth Syst. Sci. Discuss., 10, 15525-15624, doi:10.5194/hessd-10-15525-2013.

Koutsoyiannis, D. (2006), Nonstationarity versus scaling in hydrology, J. Hydrol., 324, 239-254.

Koutsoyiannis, D. (2011), Hurst-Kolmogorov dynamics and uncertainty, J. Am. Water Resour. Assoc., 47, 481-495.

Montanari, A., and D. Koutsoyiannis (2012), A blueprint for process-based modeling of uncertain hydrological systems, Water Resour. Res., 48, W09555, doi:10.1029/2011WR011412.

Plate, E.J. (2002) "Flood risk and flood management." Journal of Hydrology 267.1: 2-11.

Steinschneider, S., A. Polebitski, C. Brown, and B. H. Letcher (2012), Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change, Water Resour. Res., 48, W11525, doi:10.1029/2011WR011318.