EGUSPHERE-2023-3039
Response to referees

Response to reviewer # 1

We thank the referee for having reviewed the paper and for having provided comments. NHESS/Copernicus mentions that a revised manuscript should not be prepared at this stage. We include below our answers to each comment, as well as to how we would change the manuscript in the revision. The comments are replicated integrally and our answers are written in **blue and bold** font immediately underneath.

****

This manuscript contributes by modelling flood occurrences and their impacts using statistical and machine learning methods. The paper demonstrates high-quality research based on the methodology's effectiveness in modeling floods and predicting not only affected areas but also the displacement of populations. It presents important characteristics related to flood occurrence and impact that will inspire future research, especially when extrapolating the methodology to other regions. The paper is well-presented overall, containing crucial information that is carefully provided.

> **Thanks very much for the positive review.**

Additional comments:

1. Check for spacing between numbers and units throughout the paper to correct numerous typos in this regard.

   We indeed found instances of such typos at L158, L176 and elsewhere.

   ⇒ We will double check the manuscript to correct these typos. If the paper is accepted for publication, the typesetting stage shall correct any of these remaining typos as well.

2. Provide a more detailed explanation of what is meant by "Modelling flood occurrence is akin to a classification problem." Be specific.

   We defined flood occurrence over L188-L190 as the intersection of a DFO polygon with a PL8 watershed. There are therefore two possible outcomes: flood (intersects) or no flood (does not intersect). This is different from hydrological approaches that typically define flooding using discharge or runoff and whose flood frequency is expressed in return periods.

In data science or machine learning, there is a classification problem when the variable we aim to explain is categorized in 2 or more classes. In our case, flood or no flood are 2 classes. As such, flood occurrence as defined in the manuscript is a classification problem. Moreover, usage of the word "akin" is not necessary in the quoted sentence.

⇒ We suggest explicitly stating that the variable we aim to explain has 2 classes (flood occurrence or no flood occurrence) which yields a classification problem. We would therefore modify the first sentence of Section 3.2 accordingly.

3. Represent "s" as a sub-index in "βs".

   The "s" was meant to be the plural form of β as in betas.

   ⇒ We suggest using formal mathematical notation to define the betas along the lines of (using LaTeX) : and $\beta_0, \beta_1, \beta_2, …$ are the coefficients of the model.

4. L497-L500. Reconsider the phrasing of the conclusion regarding what-if scenarios percentages. This is not very clear in the conclusions as it is in Section 5.2.

   ⇒ We suggest replacing the current sentence with (changes in italics):
   "What-if scenarios and sensitivity analyses are also useful to understand the impact of a % increase in population displaced for scenarios where *the average precipitation is shocked by +10%, +25% or +50% above its climatology*."