Authors' replies to Reviewers comments for NHESS-2021-236

Once again we would like to thank Dr. Philip Ward and the two anonymous Reviewers for their helpful comments and suggestions, which further improved the manuscript.
Below we provide our discussion of Reviewer's comments (in blue italic font) and describe the changes addressing them (in black font).

Reviewer #1:

"Extreme Coastal water level" is in the title but it is never mentioned in the abstract. In the abstract, the Authors refer to "extreme sea level" and "total water level". Moreover, projections are done on total water level (e.g., Section 2.2.6 and Fig. 7 – revised manuscript).

Thank you for pointing out this inconsistency in the terminology used in the abstract. We have modified this section according to this comment, which it now reads: "Accurate estimates of the probability of extreme sea levels are pivotal for assessing risk and for designing coastal defense structures. This probability is typically estimated by modelling observed sea-level records using one of a few statistical approaches. In this study we comparatively apply the Generalized Extreme Value (GEV) distribution, based on Block Maxima (BM) and Peaks-Over-Threshold (POT) formulations, and the recent Metastatistical Extreme Value Distribution (MEVD) to four long time series of sea-level observations distributed along European coastlines. A cross-validation approach, dividing available data into separate calibration and test sub-samples, is used to compare their performances in high-quantile estimation. To address the limitations posed by the length of the observational time series, we quantify the estimation uncertainty associated with different calibration sample sizes, from 5 to 30 years. We study extreme values of the coastal water level – the sum of the water level setup induced by meteorological forcing and of the astronomical tide – and we find that the MEVD framework provides robust quantile estimates, especially when longer sample sizes of 10-30 years are considered. However, differences in performance among the approaches explored are subtle, and a definitive conclusion on an optimal solution independent of the return period of interest remains elusive. Finally, we investigate the influence of end-of-century projected mean sea levels, on the probability of occurrence of extreme total water levels (the sum of the instantaneous water level and the increasing mean sea level) frequencies. The analyses show that increases in the value of total water levels corresponding to a fixed return period are highly heterogeneous across the locations explored".

We believe that the use of "total water level" in the title, before an adequate definition of the term (given in Section 2.2.1 and, very succinctly, in the abstract) would be confusing to the reader.

In the abstract, the Authors say that "the MEVD estimates outperform the traditional methods" (Line 10) while from the conclusion and the responses to reviewers it seems not possible to draw such a strong statement.

In the revised manuscript, we have attenuated this conclusion. In particular, we have replaced this statement with: "We study extreme values of the coastal water level – the sum of the water level setup induced by meteorological forcing and of the astronomical tide – and we find that the MEVD framework provides robust quantile estimates, especially when longer sample sizes of 10-30 years are considered. However, differences in performance among the approaches explored are subtle, and a definitive conclusion on an optimal solution independent of the return period of interest remains elusive".

Section 3.3. and discussion about the effect of changes in total water level as a function of mean sea level. I do appreciate the idea of trying to quantify the changes in Tr. However, changes in exceedance probability (so return period) depend mostly on the shape of the distribution (the tail) especially for high quantile while here the focus is on the frequency of the observations.

We interpret this comment as asking to clarify the rationale for the discussions in Section 3.3. This section seeks to examine how events with different *Tr* may change under sea-level rise at different sites. We believe that Eq. 9 serves this purpose. We have now further developed the discussion by highlighting how changes are site-dependent (lines 465-466): “Finally, Eq. 9 shows that percentage changes in *Tr* are highly site-dependent through the shape of *f(z-msl)*.”

It is important that the Authors specify the value of msl used in Figure 7. The caption says “The green curve represents the estimates obtained with the observed record; the blue and red curves represent the estimates obtained with the projected sea-level rise (SLR) until 2100 with RCP4.5 (blue) and RCP8.5”. Does it mean that msl is equal to mean sea level in the year 2100? The return level will differ significantly depending on msl assumed.

We agree with this suggestion. As reported from lines 314 to 316, “2) we estimate the future probability of extreme total water levels by translating extreme level quantile estimates upward according to location-specific projections of mean sea level in the year 2100 […]”). Therefore, the *msl* is computed with respect to the year 2100. In the revised caption of Figure 7 we have changed “until 2100” with “in the year 2100”.

Reviewer #2:

As general comment: you state that the Poisson hypothesis is a limit of the POT-GPD approach, though it is known that the frequency of occurrence of the events does not affect significantly peak estimates (see e.g. Onoz & Bayazit, 2001). As such, the fact that the MEVD approach does not rely on such hypothesis really yields an advantage over the POT-GPD? please comment in the paper.

We agree with this Reviewer that the literature shows that assumptions for the occurrence process that are different from the Poisson distribution lead to small differences in estimated quantiles. The relaxation of the Poisson assumption is not the main advantage afforded by the MEVD, and we have de-emphasized this point in the discussion at line 73 by removing "(while, e.g. POT-GPD assumes a Poisson occurrence process)". Önöz and Bayazit, (2001) were already cited in the manuscript.

Page 8, line 202: please be more specific on the test used to select GPD also in view of the fact that you discuss the results at page 13, line 349.

Lines 211-213 now report: “Based on the comparative evaluation of the performance of these distributions, e.g. using diagnostic quantile-quantile scatter plots, the Generalized Pareto distribution emerged as the best model for the “ordinary” coastal water level values”.

As you title Section 3.1 "Mann-Kendall trend analysis", you should at least mention the test in the Methods (perhaps in Section 2.2.1?).

In the Methods section, we have added a brief introduction to the following Section 3.1 (lines 139-142): “[…] are identified and their values constitute the basis for subsequent analyses of (i) log-term trends study of maximum yearly departures from the average mean sea level (two-tail Mann-Kendall test, Mann (1945)), and (ii) statistical inference of past coastal flooding events and their potential future changes”.

Please review the grammar and some terminology. A few examples: […].

Thanks, we have revised terminology and grammar throughout the manuscript.