Comment on nhess-2022-74
Anonymous Referee #1


The MLMF method introduced by the authors seems to promise a nice way to generate good accuracy within a reasonable runtime for probabilistic hazards. This could be very useful for the nat haz community. However, as a traditionally educated hydrodynamic modeler and coastal engineer, it was very difficult for me to understand the details of the method presented. It would be helpful if the explanation could be extended or revised for understanding by a wider array of possible users. This could include something like "for each SFINCS run we calculate the mean and variance in water levels", "for each XBeach run we calculate mean and variance of water levels", "we do something with each of these means and variances to come up with a total mean and variance". With the explanation given so far, I have difficulty understanding how data from the various levels are used, how the various "samples" are used (there should be a better term for "samples"), and even how the two different models' results are used together. In general, the method should be accessible to hydrodynamic modelers to use, and understandable by such. Other comments follow.

Fig 1. I don't understand why the arrow at the left says "Level 1". Also, the term "number of samples" is confusing. Could be easier to understand as "number of runs" or "number of scenarios executed".

Line 163. What does := mean?

Please explain clearly the difference between Eq. 3 and Eq. 4.

It is difficult to tell whether sections 2.1, 2.2, and 2.3 are all just re-explaining what
Geraci et al. 2015 showed, or whether this is new material. Please clarify this in the text.

Why doesn't RMSE have a unit? Does RMSE mean relative RMSE, normalized to be unitless? This should be stated clearly.