

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC2  
<https://doi.org/10.5194/nhess-2021-55-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on nhess-2021-55**

Anonymous Referee #2

---

Referee comment on "Correlation of wind waves and sea level variations on the coast of the seasonally ice-covered Gulf of Finland" by Milla M. Johansson et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-55-RC2>, 2021

---

This paper assesses the role of sea level and wind waves in generating total water level events in the Gulf of Finland, which is covered by ice during certain times of the year, making it a more challenging but also very interesting analysis. I think the content of the manuscript is novel and deserves publication with NHESS. While most of the analysis is technically sound and well-presented there are some aspects which I think require a bit more work before the paper can be recommended for publication. I summarize these below split into one major general comment and several (mostly minor) specific comments.

General comment:

I find the copula analysis part of the paper pretty weak. The authors decide to only use three copulas, without providing convincing arguments for that selection other than pointing to previous literature. In the past especially people using Matlab ended up focusing on the Archimedean copulas as those were implemented and easy to use. However, the development of the MvCAT copula toolbox by Sadegh et al. (2017) has made it much easier to draw from a larger set of copulas. In R there are also many more copulas readily available to use. More important than using a larger set of copulas would be to show whether or not the copulas that are used at the moment are actually capable of capturing the dependence structure of the observations. There are many different goodness-of-fit tests available. Without any such tests blindly applying a random set of copulas does not provide relevant insight. I strongly encourage the authors to invest a little bit more time in strengthening this part of the analysis, as I find it to be an important component (otherwise it could just be left out but that puts a hole into the analysis).

## Specific comments

P1, l6 make clear in the abstract that  $H_s$  is used to represent waves

P1, l11 the one hour per 100 years sounds a bit strange; I understand the reasoning but isn't it just a 100-year event in the end?

P1, l18 It would be good to already mention here when defining total water levels that tides are negligible, not everyone will know that and wonder about the definition

P1, l21 "in many locations"

P2, l25 why not say "decimeters"?

P3, l29 consider replacing "pathed" by "inferred"

Fig. 1 mention in caption what the contours are

P4, l5 is the trend also statistically insignificant?

Sect. 2.2 mention somewhere early on why temperature is needed (it's not a typical variable one would use in most areas; recalling that it is used to determine ice-free periods would be helpful)

P6, l2 not sure if "long waves" is a good term here as it is basically reserved for long period waves (not travel distance)

P7, l16 I had a hard time following this definition, if  $H_s$  is the average of the 33% highest waves how is the highest single wave lower than that?

Eq. 3: using theta which is special to the Archimedean (and one parameter) copulas makes that an alteration of Sklar's theorem

P8, l19 first spell out and then introduce the abbreviation CCDFs

Fig. 3 when describing the results it would be helpful to recall that sim I has wave heights set to zero

P10, l5 "sea level variations are weaker"

Fig. 5 caption: the values for tau from the entire data are not in table 3 (the one for observations of 0.2 is included but just not rounded, but the 0.16 is not)

Fig. 7 & 8: make clear that it y-axis shows exceedance probability per hour (or translate values to exceedance probability per year)

Fig. 8 It would be good to show the convolution results from figure 7c as well for direct comparison (only the one for 1992 to 2019); also see general comment above about testing which (or any) of the copulas are actually a good choice

P15, l8 "is the highest"

P16, l6 "quite" is a better term than "peaceful"

P16, l6-8 that sentence wasn't clear to me please explain better what it's supposed to tell the reader and why it's relevant

P17, l22-25 In the discussion further up the authors correctly point to the fact that the observed data is way too short to infer information for longer return periods (I would have commented on that if I hadn't seen that remark); why not focus (or at least add) results for a more reasonable return period, such as 10 years or so? At least we know the results would be more robust.