Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2017-177-RC1, 2017

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Estimations of statistical dependence as joint return period modulator of compound events. Part I: storm surge and wave height" *by* Thomas I. Petroliagkis

Anonymous Referee #1

Received and published: 21 August 2017

——— General Comment

This paper addresses an important issue: the probability of marine storms characterized by the simultaneous presence of high waves and large storm surges. On the basis of two hindcast studies (one for waves and another for surges) it describes the dependence and the correlation between the two components of marine storminess at 32 points, which are located in correspondence of river mouths along the coastline of north and south Europe.

I find the subject interesting and results potentially worth to be published. However, I recommend that the author improves his manuscript. Some, hopefully helpful, sugges-

Printer-friendly version

Discussion paper

tions are in the specific comments here below.

In fact, the paper needs major improvements for being publishable. It relies on intense/extreme events simulated by hindcast studies without providing sufficient information on their validation. The description of the statistical method should be more precise The presentation of results should be improved by optimizing tables and figures. Causes of spatial variation of dependent and correlation should be better discussed. Parts containing details in a form similar to a technical report should be removed from the main body of the text.

————Specific comments

1) Validation of hincasts and its presentation

No sufficient attention is paid to assessing whether storm surge and wave simulations are capable of reproducing extremes values. Correlation and bias are not representative in this sense. Further, the presentation is strongly asymmetric between waves and surges, with a discussion of percent errors for surge and absolute errors for waves. Further, there is no information on the spatial distribution of storm surge errors. The paper needs to present information on the spatial distribution of percent errors in reproduction of high storm surges and waves (possibly of their extremes). In general, I suggest to use maps with percent errors, which are much more effective than tables to present such information. Without this it is difficult to estimate how realistic conclusions are. Errors in timing are important and are not discussed. In my view, the statement in the conclusions “the overall performance of both surge and wave hindcasts is considered satisfactory” is not documented in the results

The local validation of maxima at the Rhine River ending point is very convincing. It is anyway not clear whether such good performance of the models can be extended to other selected stations. Is this validation possible in other stations in other parts of the domain so that reader can be convinced that results in terms of correlation and dependence are convincing across the domain?

Section 4.2 line 16-17 the statement "Overall, it seems that hindcasts in this case were able of resolving and estimating both the correct type and strength of correlation between source variables." Could this be better enlightened at least for the Rhine station where data are available. How do we assess what is the real correct statistical dependence between surge and waves?

2) Description of statistical methods.

The description of the method should be clear also to a reader not familiar with the involved statistical methods. Some details appear confusing. Eventually, if clarifying them requires too much text, I suggest the author to publish it in the supplementary material. Here is a list of points that I recommend to clarify.

Line 1 page 5 writes that a transformation is adopted (please describe it) to produce identical marginal distribution. Line 4 writes that a copula function is used to diminish the effect of different marginal distributions. The two statement do not appear consistent to me.

In eq(1) the dependence chi is defined for z* (upper limit of the observations), while in eq (3) is defined for any generic level u. Please explain this apparent inconsistency

The derivation of eq.(3) does not appear straightforward to me. Please add a reference

In eqs.(3-5) the relation between U, V, u and X, Y, x* is not provided in the text.

The way in which chiˆbar (statistical dependence of asymptotically independent variables) is computed is not given. Distinction between chi and chiˆbar is not well explained

The concept of asymptotic dependence is not explicitly stated

It is not described how correlation is computed. Is it correlation between time series of hourly (or 3-hourly or 6-hourly) values of surge levels and wave heigh? Is correlation between the sequence of daily maxima? Between the sequence of maxima in 12 hours long windows?

Provide a precise definition of definition of compound events as adopted in this study

Clarify the criterion leading to the selection of top 80 evennts

In Chapter 2.2, after the discussion, I cannot find the information on the values of alpha and u actually used in this study.

page 4 lines 16 The statement "hydro-meteorological analyses based on real data often lead to an assessment of complete independence that could result to an underestimation of the joint probability of concurrent extreme events" is written in an ambiguous form. Please explain how joint probability is underestimated if data are "real" and the analysis is correct.

I am confused by section 2.2 (which I fail to follow concerning the selection of the chi value) and section 2.4. Establishing a confidence interval (section 2.3) should be sufficient for assessing the significance of the computed dependence values. Is here a duplication of information?

3) Spatial variations of dependence and correlation, interpretation of results

The discussion of the spatial distribution of correlation and dependence and explanation for the differences is rather inconclusive. The author writes that "dependence is likely to occur when different processes are linked to some common weather (forcing) conditions" but no convincing investigation is made on that respect. Lack of dependence could for instance be explained by a substantial contribution of inverse barometer effect to storm surges, but there is no mention of this in the paper.

Figures with the spatial distribution of correlation and dependence would be very useful. I suggest to replace the corresponding tables with maps

Section 4.5 des not provide interesting interpretation of results. Interpretation of results in term of understanding factors leading to compound events is not provided Further annotation in figure 12 is not readable . Interpretation of results at the Rhone river mouth does not account for the possibility that many surge events are produced by inverse barometric effect and not by winds.

At some stations, wind during compound events is blowing offshore. Local high waves are unlikely caused by those winds

Actual definition of prevailing and dominant wind is not clear to me (page 3, line 29-30)

4) Parts to be removed from the main body the text

A part of the paper is devoted to differences between the results produced by two software packages: R and Mathlab. Lines such as 19-26 at page 5 are interesting in a technical report, but of limited interest for a scientific paper. The cause of differences is not discussed and it is not clear whether it has a scientific relevance. Lines 16-18 at page 6 write that "Relatively small differences among various estimates made by chiplot of evd (R), taildep of extRemes (R) and mat_chi (matlab) were found. This most probably is due to the unavoidable dissimilarities between the criteria being imposed on data pairs when applying POT methodology (selection of different critical thresholds)". Continuing along this comment... Table3 and 4(and analogously 5 and 6) are presented as a comparison between packages, which is correct in a technical report but not in a scientific paper. I suggest to skip this discussion or eventually use the possibility of providing supplementary material for explaining technical differences between software packages and how they are used.

——— Other points and technical corrections:

Table7 Is redundant with respect figure 7

Figure 8 wind rose and related annotation in this figure redundant in my opinion I failed to find the " Defra/Environment Agency R&D Technical Report FD2308/TR3 on-line. I recommend the web link for downloading this and other technical reports to be provided in the reference list

Page 7, line 6 graphically or empirically?

Abstract line 14 adapted or adopted?

Lines 13-14 ref to personal communication (which cannot be properly documented) looks useless here

page 11, line 22-23 refer to ”personal communication”, which I think is not suitable in this form

Table 1 is not needed in the main body of the text Figure1 provides the same information

I do not find a clear explanation on which data are grouped under the lable hind_com,obs_com and Hind_tot . One can guess but a clear description should be given in data and method.

Results section contain description of tools (lines 12-18, page 18) . This should be moved to section 2 or 3, or (preferably in my view) removed or transferred to a supplement.

Fig.10 I cannot see the negative and zero dependence values that are mentioned in the text (page40, line 15). . .

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-177, 2017.