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

Thomas I. Petroliagkis
thomas.petroliagkis@ec.europa.eu

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, suggestions 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.

I truly thank the reviewer for his/her comments on the manuscript.

Next, I will address all referee's comments specifically.

Specific Comments (Ref1)

(1) Validation of hindcasts 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.

As a general comment:

The lack of observations (and especially those in compound mode) make hindcasts necessary. The scope of this paper has not been to re-validate or further validate the hindcasts of storm surge neither those of (significant) wave height. The validation of storm surge hindcasts have already been performed in Vousdoukas et al. (2016) whereas the validation of wave height hindcasts in Philips et al. (2017) in which the validation of the new ECMWF ERA5 reanalysis wave data set (including such long time series of waves) is documented. Both studies are falling in the category of peer-reviewed papers. Additional details referring to these relevant validations are provided below (by parts). Based on such validated data (hindcasts) over the main European coasts the statistical dependence analysis has been performed for 32 points of interest. Further, the study is mainly demonstrating the possibility of utilizing joint probability methods in coastal flood management by considering and putting emphasis on the statistical dependence between source variables. It also demonstrates that dependence, which is capable of modulating their joint return period, has to be estimated before the calculation of their (final) joint probability.

In brief: the main idea of this study has been to adopt a set of two already validated hindcasts (of surge and wave) and investigate over extreme compound events and their joint probability by utilising the so-called statistical dependence.

In addition, two new Supplements (Technical and Statistical) have been compiled and submitted providing explanations and clarifications to both technical and statistical issues.

References

- Phillips, B.T., Brown, J.M. and Bidlot, J.-R. and Plater, A.J.: Role of Beach Morphology in Wave Overtopping Hazard Assessment. J. Mar. Sci. Eng. 5(1), 2017

- Vousdoukas, M.I., Voukouvalas, E., Annunziato, A., Giardino, A. and Feyen, L.: Projections of extreme storm surge levels along Europe. Clim. Dyn. 47(9): 3171-3190, doi:10.1007/s00382-016-3019-5, 2016.

Valid_Ref1_01: 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.

As a general comment:

when low resolution models are used (as in this case) for reproducing time series of significant weather parameters, extremes cannot be captured with their exact (high-impact) value but in most cases only their footprints can be resolved (as extremes of a lesser value). A previous example can be seen in Petroliagis and Pinson (2012) where the footprints of extreme wind speed values over Bremen airport are captured by ERA-Interim (as footprint spikes) but they are considerably underestimated. In a similar approach, the scope of the study is to take (at least) into account such spikes (footprints) of extremes and study the statistical dependence of these spikes of storm surge and (significant) wave height.

Such footprints of extremes (resolved by hindcasts) can be found in Table 2 (Technical Supplement) where the 98.5% percentile extremes of storm surge observations are compared to their corresponding hindcast values (falling in the same 98.5% category). It becomes obvious that although hindcasts could not resolve the exact extremity of events at least their footprints were well captured. In a similar way in Table 3 (Technical Supplement) the footprints of significant wave height observation extremes are resolved by their corresponding hindcast (less intense) values.

It is important to point out that hindcasts above all were capable of identifying and resolving all seven (7) compound events that took place during the common time interval of 1,114 days.

On the same track, the set of storm surge hindcasts used in the current paper was already validated against 110 tidal gauge stations as described in Vousdoukas et al. (2016) reference paper. Vousdoukas et al. (2016) utilised both RMSE and relative (%) RMSE metrics. Overall, the model showed to reproduce satisfactory the measurements as shown in examples given in Figure 3 (Vousdoukas et al., 2016) over four tide-gauge stations in various coastal points of European coasts (Saint-Nazaire in France, Millport in UK, Hirsthals in Denmark and Rorvik in Norway). Studying closely Figure 3 it becomes obvious that hindcasts were able to simulate quite well the available set of observations capturing also efficiently local extremes. Further, the period of validation (2008-2014) had been characterized by an increased marine storm activity including high impact events as mentioned in Bertin et al. 2014; Breilh et al. 2013; Met Office and Centre for Ecology and Hydrology 2014; Vousdoukas et al. 2012.

Referring to the suggestion of using percent maps a new reference in text will be made pointing to Figure 4 (Vousdoukas et al., 2016) scatter plot showing RMS error in m (a) and as a percentage of the SSL (Storm Surge Level) range (b) for all the available tidal gauge stations.

Concerning the validation of wave hindcasts, the set used in the study is considered as a validated set with further details to be provided in Philips et al. (2017). The data are based on a dedicated re-run of the European Centre for Medium-Range Weather Forecasts

(ECMWF) ECWAM Wave Model (ECMWF, 2016) Cycle 41R1 at 28-km resolution. The model is forced by a six hourly ERA-interim (Dee et al., 2011) wind field with no wave data assimilation. The effect of water level change and surface current due to tides and surge is neglected. This global hindcast set has been produced in preparation of the ECMWF next reanalysis (ERA5).

I will add in the main text a reference to Figure 4 (Vousdoukas et al., 2016) scatter plot showing RMS error in m (a) and as a percentage of the SSL range (b) for all available tidal gauge stations. This reference will be in harmonisation with Figure 2 (current study) that is referring to the validation of wave hindcasts (RMSE values).

References

- Bertin X., Li K., Roland A., Zhang Y.J., Breilh J.F., Chaumillon E.: A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. Coastal Eng 94:80–89, 2014.

- Breilh J.F., Chaumillon E., Bertin X., Gravelle M.: Assessment of static flood modeling techniques: application to contrasting marshes flooded during Xynthia (western France). Nat Hazards Earth Syst Sci 13:1595–1612, 2013.

- Met Office, Centre for Ecology & Hydrology: The recent storms and floods in the UK. p 29, 2014.

- Petroliagis, T. I. and Pinson, P.: Early warnings of extreme winds using the ECMWF Extreme Forecast Index, Meteorol. Appl., 21, 171–185, 10 doi:10.1002/met.1339, 2014.

- Phillips, B.T., Brown, J.M. and Bidlot, J.-R. and Plater, A.J.: Role of Beach Morphology in Wave Overtopping Hazard Assessment. J. Mar. Sci. Eng. 5(1), 2017

- Vousdoukas M.I., Almeida L.P., Ferreira Ó.: Beach erosion and recovery during consecutive storms at a steep-sloping, mesotidal beach. Earth Surf Process Landforms 37:583–691, 2012.

- Vousdoukas, M.I., Voukouvalas, E., Annunziato, A., Giardino, A. and Feyen, L.: Projections of extreme storm surge levels along Europe. Clim. Dyn. 47(9): 3171-3190, doi:10.1007/s00382-016-3019-5, 2016.

Valid_Ref1_02: Errors in timing are important and are not discussed.

This study is focused over maxima taken place over 12- and 24-hours based on 3-hour set of hindcast values. These kind of (timing) errors were investigated over Rhine River (NL) ending point and the overall conclusion has been that hindcasts were able to pick up similar (to observations) maxima during both the 12-and 24-hour intervals.

An extra investigation based on extreme values of observations (during the common time interval of 1,114 days) exceeding a variety of percentile values (for the RIEN of Rhine River) showed that both storm surge and their corresponding wave height hindcasts were able to capture almost all of the 24-hour extremes on the same (correct) day but with a weaker intensity (i.e., with a correct footprint of lesser intensity).

I will include the results of this latter investigation concerning various percentile extremes in the Technical Supplement and I will add a relevant reference (to the Technical Supplement) in the main text.

Valid_Ref1_03: 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.

As already stated above, both sets of hindcasts had already been validated (Vousdoukas et al., 2016, Philips et al., 2017). Emphasis was given if these two sets were suitable to allow someone to go the extra step of resolving correctly the type and strength of both correlation and statistical dependence. Such an investigation was performed over the ending point of Rhine River (NL) with very satisfactory results. The same approach (of estimating statistical dependence) was adopted for the rest of ending points of the study.

I will point out and stress (in the Introduction) the fact that both sets of hindcasts are considered to be (already) validated and provide the reader with the relevant references.

Valid_Ref1_04: 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?

Although the results of this study are based on already validated hindcasts of surge and wave, it is not straightforward how these hindcasts could guarantee for the exact (correct) estimation of both correlation and dependence between source variables (in places other than the Rhine River ending point) but nevertheless, the results of this study represent the first step on this direction.

Further, I agree that such specific type of validation (referring to correlation and dependence estimation) should be extended to other ending points of the study by utilising appropriate sets of observations. I am afraid this could be proved quite difficult if not impossible due to the necessity of long-period co-existing (real-time) observations of surge and wave over the areas of interest.

For time being, the study is mainly demonstrating the possibility of utilizing joint probability methods in coastal flood management by considering and putting emphasis on the statistical dependence between source variables. It also demonstrates that dependence, which is capable of modulating their joint return period, has to be estimated before the calculation of their (final) joint probability.

Valid_Ref1_05: 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?

The real (correct) statistical dependence is estimated by utilising the formula of Equation 3 over a long set of real data (observations) of storm surge coming from a tide gauge and real data of wave height coming from a close by wave buoy. The tide gauge and wave buoy have to be relatively close for obvious reasons. Usually the tide gauge is in the vicinity of the port while the wave buoy is suited some kilometres offshore in front of the port.

Besides observations (that are limited in time length) hindcasts can be used as in our case. Storm surge hindcasts were compiled by the Delft3D-Flow hydrodynamic model pinpointing the position of various tide gauges whereas wave height hindcasts were made by another (wave) model (ECWAM of ECMWF) pinpointing the position of relevant close by wave buoys.

It should be evident by now that even if hindcasts might be missing the exact magnitude of the extremes mainly due to the limited (model) resolution the most important issue here is their ability to resolve and estimate the correct value of both correlation and dependence as it is estimated over real data (observations).

In the case of the RIEN of the Rhine River, the high level of agreement between the dependence estimated utilising (surge and wave) observations and the one utilising (surge and wave) hindcasts, points to the direction that hindcasts are capable of resolving both the correct type and strength of dependence between the source variables.

I will stress this point (how we access the real correct statistical dependence) in the main text by presenting the concept behind estimating similar (if not almost the exact) dependence values by utilising both observation and hindcast sets of data.

(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.

As a general comment:

A Statistical Supplement has been compiled clarifying missing or confusing details.

Stat_Ref1_01: 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.

The transformation refers to the separately ranking of observations and the division of each rank by the total number of observations. It is considered as a trivial methodology of obtaining identical marginal distributions with Uniform [0, 1] margins. The utilisation of the copula C function does exactly this. At the same time, copula C contains the complete information about the joint distribution of X and Y.

I will rewrite the paragraph providing the required information for consistency with additional clarifying details that will be available in the new Statistical Supplement.

... For obtaining identical marginal distributions, each set of observations is ranked separately and each rank is then divided by the total number of observations resulting in a data transformation with Uniform [0, 1] margins. At this point, it is convenient to consider the bivariate cumulative function F(x, y) = Prob(X ≤ x, Y ≤ y) that describes the dependence between X and Y completely. The effect of different marginal distributions can be diminished by assuming the copula function C in the domain [0, 1] x [0, 1] such as:

$$F(x,y) = C\{F_x(x), F_x(y)\} \qquad (2)$$

where F_x and F_y can be any marginal distributions. Such utilisation of the copula function has the same effect as if observations were ranked separately and divided by the total number of observations. In addition, ~~The~~ the copula C contains the complete information about the joint distribution of X and Y and it is invariant to marginal transformation ...

Stat_Ref1_02: 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.*

In eq (1), z* represents the upper limit of the observations but after the data transformation to Uniform [0, 1] margins, this upper limit is equal to (becomes) 1. For completeness eq (4) is added providing the (final) estimation for statistical dependence (chi)

$$\chi = \lim_{u \to 1} \chi(u) \qquad (4)$$

I will rewrite the relevant paragraph to provide the required explanation and clarifying details in the new Statistical Supplement.

... Taken into account the upper limit of the observations (previously defined as z^* in Eq. 1 but now being equal to 1), the dependence measure $\chi(u)$ will be given by:

$$\chi = \lim_{u \to 1} \chi(u) \qquad (4) \ldots$$

A necessary update to the numbering for the rest of equations will be applied ...

Stat_Ref1_03: The derivation of eq (3) does not appear straightforward to me. Please add a reference.

The main reference of Coles et al. (2000) is mentioned in line 21 page 4. Subsection 2.1 contains only a brief description of the methodology that is described in details in Coles et al. (2000).

I will include the relevant reference at an earlier point in the new Statistical Supplement (Section 2).

… Details of deriving Eq. 3 can be found in Coles et al. (2000). Based on Eq. 3, a set of χ values can be evaluated at different quantile levels u ~~(for details see Coles et al., 2000)~~. The selection of a particular level u corresponds to threshold levels (x^*, y^*) for the two different data series.

Stat_Ref1_04: In eqs. (3-5) the relation between U, V, u and X, Y, x is not provided in the text.*

I will provide details of the relation between all mentioned terms in eqs ~~(3-5)~~ (3-6) by rewriting and incorporating relevant statements in the new Statistical Supplement.

… In addition, ~~The~~ the copula C contains the complete information about the joint distribution of X and Y and it is invariant to marginal transformation. This means that C can be described as the joint distribution function of X and Y. Further, X and Y are transformed to the new variables U and V with Uniform [0, 1] margins. It follows that the dependence measure $\chi(u)$ for a given threshold u can be given by …

Stat_Ref1_05: The way in which chi_bar (statistical dependence of asymptotically independent variables) is computed is not given. Distinction between chi and chi_bar is not well explained.

The calculation of chi_bar is clearly mentioned in line 1 page 6. It refers to the methodology described in Coles et al. (2000). More details on chi_bar and examples of how differs from chi are given in Coles (2001).

I will rewrite the relevant paragraph and add references with examples in the new Statistical Supplement.

… Chibar (chi_bar) parameter refers to the statistical dependence of asymptotically independent variables whereas chi (χ) refers to the statistical dependence of asymptotically dependent ones. Details on the estimation of chibar are documented in Coles et al. (2000) whereas examples and how to utilise ($\bar{\chi}$) can be found in Coles (2001). The ~~latter~~ class of asymptotic dependence appears to be the case in Literature, having reached a consensus that there is strong, although not overwhelming, evidence for asymptotic dependence between wave height and surge ...

References

- Coles, S.G., Heffernan, J. and Tawn, J.A.: Dependence measures for extreme value analyses. Extremes, 2, 339-365, 2000.

- Coles, S.G.: An Introduction to Statistical Modelling of Extreme Values. Springer Series in Statistics. Springer Verlag London. 208p, 2001

Stat_Ref1_06: The concept of asymptotic dependence is not explicitly stated.

The concept of asymptotic dependence (chi) is stated with adequate details in the main reference of Coles et al. (2000).

In summary, chi is on the scale [0, 1] with the set of values (0, 1] corresponding to asymptotic dependence whereas the measure chibar falls within the range [-1, 1] with the set of values [-1, 1) corresponding to asymptotic dependence. That is why the complete pair of chi and chibar is required as a summary of extremal dependence:

- chi > 0 & chibar = 1 reveals asymptotic dependence, in which case the value of chi determines a measure of strength of dependence within the class.

- chi = 0 & chibar < 1 reveals asymptotic independence, in which case the value of chibar determines the strength of dependence within the class.

Based on the main reference of Coles et al. (2000), I will incorporate the main concept behind asymptotic dependence in the new Statistical Supplement.

... The ~~latter~~ class of asymptotic dependence appears to be the case in Literature, having reached a consensus that there is strong, although not overwhelming, evidence for asymptotic dependence between wave height and surge (Wadsworth et al., 2017).

The concept of asymptotic dependence (χ) is stated with adequate details in Coles et al. (2000). In brief, χ is on the scale [0, 1] with the set (0, 1] corresponding to asymptotic dependence whereas the measure chibar ($\bar{\chi}$) falls within the range [-1, 1] with the set [-1, 1) corresponding to asymptotic independence. That is why the complete pair of χ and $\bar{\chi}$ is required as a summary of extremal dependence:

- $\chi > 0$ & $\bar{\chi} = 1$ reveals asymptotic dependence, in which case the value of χ determines a measure of strength of dependence within the class

- $\chi = 0$ & $\bar{\chi} < 1$ reveals asymptotic independence, in which case the value of $\bar{\chi}$ determines the strength of dependence within the class ...

Stat_Ref1_07: 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 height?

Both correlation and dependence estimations refer to maximum values during 12- or 24-hour time intervals. This is mentioned in Section 1 (Introduction) where the definition of max12 (maxima over a time interval of 12 hours) and max24 (maxima over a time interval of 24 hours) are introduced for the first time.

I will introduce and stress appropriately the definition of both max12 and max24 intervals in the Results Section.

… Referring to the full span of hindcasts, analytical maps and tables have been assembled ~~referring to~~ containing to both correlation and dependence values between surge and wave over the 32 RIEN points considered in this study. Both correlation and dependence values were estimated over maximum values of surge and wave during 12- and 24-hour intervals (labelled as max12 and max24 respectively) …

Stat_Ref1_08: Is correlation between the sequences of daily maxima? Between the sequence of maxima in 12 hours long windows?

As mentioned above (Stat-Ref1_08), both types of correlations have been estimated. Correlation values over daily maxima are referred as max24 whereas correlations over 12-hour interval (half-day) maxima are referred as max12. This separation has been kept for both correlation and dependence estimations throughout the paper.

Stat_Ref1_09: Provide a precise definition of definition of compound events as adopted in this study.

Compound events of surge and wave are those events that coincidently are above a certain joint upper percentile criterion (here playing the role of a critical threshold).

I will add this definition in Section 1 (Introduction) for clarity reasons.

… These interactions are generally referred to as coincident or compound events (IPCC, 2012). In brief, compound events of surge and wave are those events that coincidently are above a certain upper percentile criterion (representing a critical threshold) …

Stat_Ref1_10: Clarify the criterion leading to the selection of top 80 events.

The selection is defined by the parameter alpha (a) representing the annual maximum non-exceedance probability taken equal to 0.1 following Defra TR3 Report suggestions. Such a value (0.1) of alpha corresponds to ~2.3 compound POT (Peaks-Over-Threshold) events per year exceeding the corresponding optimal selected percentile threshold (the one providing ~2.3 compound events).

Since both surge and wave time series are almost 35 years long this points to ~80 (~2.3 x 35) events over the total time period.

I will add a more detailed explanation (Section 6) in the new Statistical Supplement taking into consideration the basic guidelines documented in Defra TR3 Report (2005).

6 Selection of criterion thresholds resulting in the consideration of top-80 events

Since values of dependence (χ) can be estimated for any lower or upper threshold, initial trials were performed studying the behaviour of χ over a wide range of thresholds. Findings were similar to those contained in Defra TR3 Report (2005), justifying the selection of an optimal threshold for "alpha" (α) equal to 0.1 corresponding to an annual maximum being exceeded in 9 out of 10 years (see Sect. 2.2 of the main text for details). This value (0.1) of alpha was considered for both mat_chi and mat_chibar routines when utilising POT (Peaks-Over-Threshold) methodology resulting in an annual maximum of ~2.3 compound events.

Such an annual threshold of ~2.3 events corresponds to the top 80 (Top-80) compound events taking place during any (POT separated) day of the total 12,753 days and it was dictated mainly by two factors: the threshold had to be low enough to allow a sufficient number of data points to exceed it for estimating dependence reliably, while being high enough for the data points to be regarded as extremes.

References

- Defra TR3 Report by Svensson, C. and Jones, D.A.: Joint Probability: Dependence between extreme sea surge, river flow and precipitation: a study in south and west Britain. Defra/Environment Agency R & D Technical Report FD2308/TR3, 62 pp. + appendices (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3430_TRP_pdf.sflb.ashx), 2005.

Stat_Ref1_11: In Chapter 2.2, after the discussion, I cannot find the information on the values of alpha and u actually used in this study.

The old Section 2.2 (now Section 3 in the new Statistical Supplement) contained the main theoretical concept behind the alpha (the annual maximum non-exceedance probability) and u parameter (percentile) threshold values. It was Section 4.1 (lines 20 to 28 of page 20) that an extensive explanation about the selection of alpha (α) value being equal to 0.1 was documented. Now lines 20 to 28 (of page 20) have moved in Section 6 of the new Statistical Supplement. Based on such predefined alpha value the selection of an optimal threshold percentile (u) is straightforward. Further, alpha (α) is capable of modulating the (optimal) percentile threshold (u) in such way to allow ~2.3 compound events (of 80 in total) to take place on a yearly basis.

Such information is now contained in Section 6 of the new Statistical Supplement being in harmony with Stat_Ref1_10 (see previous comment).

Stat_Ref1_12: 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 will rewrite the statement and add required details clarifying ambiguous terms. Examples of under- and over-estimating joint probabilities are also included in the new Statistical Supplement (Section 7).

... Similarly, if the extreme observations of one variable exceed a given threshold but the other variable produces lower observations than would normally be expected, this indicates negative dependence (χ = -1). In practice, ~~hydro-meteorological analyses based on real data often lead to an assessment of complete independence that could result to an under-estimation of the joint probability of concurrent extreme events, whereas, an assumption of complete dependence could result to an over-estimation of joint probabilities~~ ... in tidal and estuarine environments, assessing the probability of flooding from the joint occurrence of both high storm surge and high wave values is not an easy process, as high surges and waves might be related to the same prevailing meteorological conditions, thus independence cannot and should not always be assumed. For instance, if we assume independence between input variables, this might underestimate considerably the likelihood of flooding (estimated by the product of their individual probability) resulting in higher risk for the coastal community. Similarly, assuming total dependence could be too conservative ...

Stat_Ref1_13: 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?

The old Section 2.2 (now Section 3 of the Statistical Supplement) was referring to the selection of an optimal percentile threshold (u) based on the annual maximum non-exceedance probability alpha (a). Then the estimation of dependence (χ) was straightforward as described analytically in the old Section 2.1 (now Section 2 of the new Statistical Supplement). Further, the old Section 2.3 (now Section 4 of the new Statistical Supplement) was referring to the estimation of 5% significance level using a permutation method whereas the old Section 2.4 (now Section 5 of the new Statistical Supplement) was referring to the estimation of confidence levels. All related values (significance and lower &

upper confidence levels) are now contained in Table 4 and Table 5 (of the new Technical Supplement) following a similar approach as the one documented in TR1 Defra Report.

I will move for clarity reasons the old Section 2.1, Section 2.2, Section 2.3 and Section 2.4 to the new Statistical Supplement whereas I will move the old Table 3 and Table 5 to the new Technical Supplement stressing the difference between assessing the significance and confidence intervals based on the methodology documented in the relevant reference (TR1 Defra Report).

References

- Defra TR1 Report by Hawkes, P.J. and Svensson, C.: Joint probability: dependence mapping & best practice. R & D Final Technical Report FD2308/TR1 to Defra. HR Wallingford and CEH Wallingford, U.K. (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3428_TRP_pdf.sflb.ashx), 2005.

(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.

As a general comment:

Storm surge is an abnormal rise of water generated by a storm, over and above the predicted astronomical tide values (http://www.nhc.noaa.gov/surge/faq.php). In observations mode, storm surge is calculated as a residual by subtracting harmonic tidal predictions from the observed sea level (Horsburgh and Wilson, 2007). Such "residual" may contain surge, tide-surge interaction, harmonic prediction errors and timing errors. Tide-surge interaction, harmonic prediction errors and timing errors are not taken into consideration in this study. On the other hand (e.g. in hindcast mode) a similar "residual" refers to the genuine meteorological contribution to sea level that represents the storm surge term. It should pointed out that the effect of wind and atmospheric pressure (inverse barometric effect) are contained in both the "residual" and storm surge terms. Based on this, it becomes clear that all data (storm surge) sets used in the study contain the effect of the inverse barometric effect besides the effect due to wind. This is the reason why the dedicated model (Delft3D-Flow) uses as input both ERA-Interim wind and pressure fields.

I will add the work of Horsburgh and Wilson (2007) in the list of References (shown below):

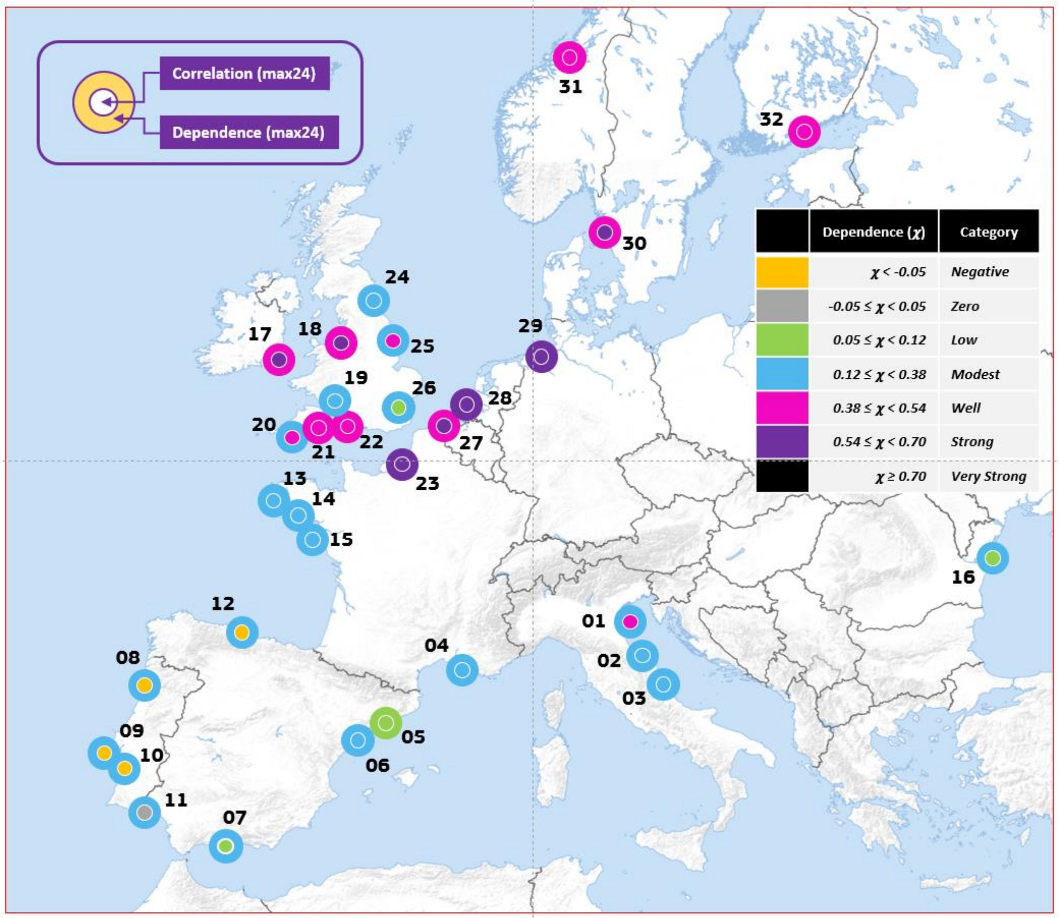
Horsburgh, K. J. and Wilson, C.: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, J. Geophys. Res., 112, C08003, doi:10.1029/2006JC004033, 2007.

Below I am addressing specifically the points.

Spat_Ref1_01: Figures with the spatial distribution of correlation and dependence would be very useful. I suggest to replace the corresponding tables with maps

Since the spatial distribution of both correlation and dependence should be displayed, a necessary change in the old Figure 10 (now Figure 9 in the main text) has been made to incorporate both correlation and dependence values using the same seven (7) relevant / reference categories. The exact values of correlation and dependence contained in the old Tables 3 to 6 have moved in the new Technical Supplement as Tables 4 to 7.

I will include the spatial distribution of both correlation and statistical dependence in the new Figure 9 (old Figure 10) shown below utilising seven (7) relevant / reference categories. Prevailing and dominant winds are to be left out of this new Figure 9 for more clarity. Their exact details contained in the old Table 7 can be found in the new Table 3.



New Figure 9 (Old Figure 10)

Spat_Ref1_02: Section 4.5 does 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.

In Section 4.5 an effort has been made to assess the low-level flow characteristics during critical compound events and not to provide a thorough explanation of the exact conditions leading to such compound events (that is beyond the scope of the paper). It also seems logical for someone to expect that in cases of surge events driven by winds this should be well captured in the corresponding wind roses relating to such prevailing and dominant (climatological) winds. Referring to the inverse barometric effect (as explained previously), all hindcast (storm surge) sets used in the study contain the effect of the inverse barometric effect besides the effect due to wind. This is the reason why the model (Delft3D-Flow) used for the production of hindcasts had as input both ERA-Interim wind and pressure fields.

I will add and explain accordingly that a thorough understanding of all factors leading to a compound event is above the scope of this study and I will skip Figure 12 (since main characteristics of both prevailing and dominant winds were contained in the old Table 7 (now in the new Table 3).

Spat_Ref1_03: At some stations, wind during compound events is blowing offshore. Local high waves are unlikely caused by those winds.

Combined events had to be de-clustered if they lasted longer than 24 hours. This means that a compound event lasting more than one day had to be counted as one (1) event even if this event could have lasted for a few days due to an approaching storm (barometric low). An example of such a compound event lasting for three consecutive days can be seen in Table 2 and Table 3 of the new Technical Supplement (referring to the time interval between 2 to 4 January 2012). After de-clustering this event will count only once and it will refer to its first date (4 Jan 2012) since after the necessary de-clustering all cases of compound events are referring to the first day of the event (the first day that both storm surge and wave height found to be above a predefined critical threshold). With such an approach, a compound event is considered only once and no other (another) event is taken into account for the next three days (even if the same event of day 1 continues to exist). Both prevailing and dominant directions are referring to the time of maximum daily wind intensity and if we consider the most common case of an approaching barometric low (storm) from the west the wind in the beginning is more WSW whereas with the passage of the storm tends to veer to a more northwest (northern) direction. I have checked the validity of this during the second, third and even the fourth day of an extended compound event and such a distinct veering is true.

Another important point is that not only an incoming onshore perpendicular wind leads to a significant storm surge or even to compound event. As an example Mistral (of north direction) that is heading to the open sea – Marin (of south direction) that is heading toward the coast of Marseille are capable of producing extreme storm surge events of equal intensity (during distinct periods of rough seas) meaning that there exist other directions as well besides the ones blowing perpendicular to the coast relating to extremes as well.

I will refer and stress this unavoidable disagreement due to the veering of the wind and provide necessary explanations for such discrepancies in the main text (as analysed above).

… Details of clima and Top-80 flow characteristics are contained in Table 7. A possible exploitation of such information referring to both prevailing and dominant low-level flow characteristics should be considered significant and kept in mind when such extreme events possibly driven by intense storm outbreaks are anticipated over the area of interest (in forecast mode) …

… Not all prevailing and dominant directions contained in Table 7 fall in the perpendicular onshore category. Especially for the RIEN points of the south North Sea, wind directions appear to be more SWS instead of rather more northerly directions and this is because combined events had to be de-clustered. This means that a compound event lasting more than one day had to be counted as one (1) event even if this event could have lasted for a few days. After this necessary de-clustering all cases of compound events, are referring to the first day of the event (the first day that both storm surge and wave height found to be above a predefined critical threshold). With such an approach, a compound event is considered only once and no other (another) event is taken into account for the next three days (even if the same event continues to exist). Both prevailing and dominant directions are referring to the maximum daily intensity and if we consider the most common case of an approaching barometric low (storm) the wind in the beginning is more WSW whereas with the passage of the storm tends to veer to a more north-western (northern) direction …

Spat_Ref1_04: Actual definition of prevailing and dominant wind is not clear to me (page 3, line 29 30).

Prevailing Wind is the most common wind direction over an area, i.e., the direction of wind with the highest frequency (AMS, 2017), whereas Dominant Wind is the direction of the strongest wind that might blow from a different direction than the prevailing wind, i.e., from a less common direction (Thomas, 2000). The periods most frequently used for the estimation of prevailing and dominant winds are the observational day, month, season, and year. Methods for determination vary from a simple count of periodic observations to the computation of a wind rose.

I will provide definitions of both prevailing and dominant wind (as presented above) and add relevant references.

References

- AMS (American Meteorological Society) Glossary: Prevailing Wind. Glossary of Meteorology (Available online at http://glossary.ametsoc.org/wiki/Prevailing_wind_direction), 2017.

- Thomas, DG. 2000. Dictionary of physical geography. Blackwell.

(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 Matlab.

As a general comment:

Focusing over various software packages has been mainly for the interest of the reader getting a feeling about the capabilities and limitations (differences) of various statistical packages used to estimate statistical dependence. Most of these explanations on differences have moved to the new Technical Supplement and Statistical Supplement.

Below I am addressing specifically the points.

Parts_Ref1_01: Lines such as 19-26 at page 5 are interesting in a technical report, but of limited interest for a scientific paper.

I will move lines 19-26 of page 5 to the new Statistical Supplement. They are now included in Section 7 (Details and examples of statistical packages used in the study).

Parts_Ref1_02: 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)".

Lines 16-18 of page 6 can be moved to the Statistical Supplement (due to their minor scientific relevance) together with the possible explanation causing differences between various statistical packages. It is not in the scope of this study to investigate further such (technical) differences.

I will move lines 16-18 of page 6 to the new Statistical Supplement. They will be included in Section 7 (Details and examples of statistical packages used in the study).

Parts_Ref1_03: 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.

Old Table 3, Table 4, Table 5 and Table 6 can be moved to the new Technical Supplement (as Table 4, Table 6, Table 5 and Table 7 respectively). The relevant discussion points relating to the main characteristics among various dedicated statistical packages are to move to the new Statistical Supplement.

The main results referring to both correlations and dependence (contents of the new Table 4, 5, 6 & 7 of Technical Supplement) are now contained in the new Figure 9 (old Figure 10) in graphical mode.

I will move Table 3, Table 4, Table 5 and Table 6 to the new Technical Supplement. I will make all necessary changes to old Figure 10 (new Figure 9) to include the main results referring to both correlation and dependence values.

(5) Other points and technical corrections:

Points_Ref1_01: Table7 is redundant with respect figure 7.

Most probably meant Figure 12 (instead of figure 7) since Figure 12 refers to the main elements of Table 7 (in graphical mode).

I will skip Figure 12 and keep old Table 7 (new Table 3) in the main text that contains all relevant information of prevailing and dominant winds that was graphically presented in Figure 12 (upper and lower panels) over the selected 32 RIEN points.

Points_Ref1_02: Figure 8 wind rose and related annotation in this figure redundant in my opinion.

Most probably meant Figure 11 (wind roses over the ending point of river Rhone). I trust that this set of the two wind roses (in "clima" and in "Top-80" extreme mode) are necessary for the reader to get a feeling of the difference between prevailing and dominant wind as captured in a wind rose diagram.

Further, after skipping Figure 12, I strongly believe that at least an example of a wind rose diagram should remain for explanatory and demonstrating reasons to the reader.

Taken into account the deletion of Figure 12, I will keep old Figure 11 (new Figure 10 in the main text) as an example of wind rose diagram and reference point of how to differentiate prevailing from dominant wind conditions.

Points_Ref1_03: 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.

I will include the web links referring to all Technical Reports contained in the list of references as shown below:

- Australian Rainfall & Runoff Project 18: Coastal Processes and Severe Weather Events: Discussion Paper, Water Technology report to Australia Rainfall & Runoff (2009) referring to the report of Department of Science, IT, Innovation and the Arts – Science Delivery (October 2012) "Coincident Flooding in Queensland: Joint probability and dependence methodologies" (https://www.longpaddock.qld.gov.au/coastalimpacts/inundation/coincident_flood_technical_review.pdf), 2009.

- Defra TR0 Report by Hawkes, P.J.: Extreme water levels in estuaries and rivers: the combined influence of tides, river flows and waves. R & D Technical Report FD0206/TR1 to Defra. HR Wallingford, U.K., (http://randd.defra.gov.uk/Document.aspx?Document=FD0206_5270_TRP.pdf), 2003.

- Defra TR1 Report by Hawkes, P.J. and Svensson, C.: Joint probability: dependence mapping & best practice. R & D Final Technical Report FD2308/TR1 to Defra. HR Wallingford and CEH Wallingford, U.K. (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3428_TRP_pdf.sflb.ashx), 2005.

- Defra TR2 Report by Hawkes, P.J.: Use of joint probability methods in flood management: a guide to best practice. R & D Technical Report FD2308/TR2 to Defra. HR Wallingford, U.K. (http://www.estuary-guide.net/pdfs/FD2308_3429_TRP.pdf), 2005.

- Defra TR3 Report by Svensson, C. and Jones, D.A.: Joint Probability: Dependence between extreme sea surge, river flow and precipitation: a study in south and west Britain. Defra/Environment Agency R & D Technical Report FD2308/TR3, 62 pp. + appendices (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3430_TRP_pdf.sflb.ashx), 2005.

- Hawkes, P.J.: Use of joint probability methods for flood & coastal defence: a guide to best practice. R&D Interim Technical Report FD2308/TR2 to Defra. HR Wallingford, U.K. (http://www.estuary-guide.net/pdfs/FD2308_3429_TRP.pdf), 2004.

- Hawkes, P.J.: Use of joint probability methods for flood & coastal defence: a guide to best practice. R&D Interim Technical Report FD2308/TR2 to Defra. HR Wallingford, U.K. (http://www.estuary-guide.net/pdfs/FD2308_3429_TRP.pdf), 2004.

- IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp (https://www.ipcc.ch/pdf/special-reports/srex/SREX_Full_Report.pdf), 2012.

- Petroliagkis, T.I., Voukouvalas, E., Disperati, J. and Bidlot, J.: Joint Probabilities of Storm Surge, Significant Wave Height and River Discharge Components of Coastal Flooding Events, JRC Technical Report EUR 27824 EN, doi:10.2788/677778, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC100839/lbna27824enn.pdf, 2016.

- Svensson, C. and Jones, D.A.: Dependence between extreme sea surge, river flow & precipitation: a study in south & west Britain. R&D Interim Technical Report FD2308/TR3 to Defra. CEH Wallingford, UK (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_1135_INT_pdf.sflb.ashx), 2003.

Points_Ref1_04: Page 7, line 6 graphically or empirically?

The correct one is graphically using Hazen's (1914) formula (a reference that to be added in the list of References).

I will keep the (correct) term "graphically" and I will add the relevant paper (Hazen, 1994) to the list of references.

References

- Hazen, A.: Storage to be provided in impounding reservoirs for municipal water supply. Trans. Amer. Soc. Civ. Eng. Pap., 1308 (77), 1547–1550, 1994.

Points_Ref1_05: Abstract line 14 adapted or adopted?

The correct one is "adopted".

I will change the term to the correct one "adopted".

Points_Ref1_06: Lines 13-14 ref to personal communication (which cannot be properly documented) looks useless here.

I will skip the reference (personal communication) whereas I will refer to the concept of the stability of graph curve as an important issue of how to determine efficiently (as reliable as possible) the statistical dependence.

Emphasis was also given on the stability of χ (graph) curves ~~as strongly recommended by Prof Pieter Van Gelder of Delft University, Nederlands (personal communication, 2016)~~ identifying areas that dependence was clearly converging to a specific value (with no abrupt fluctuations).

Points_Ref1_07: Page 11, line 22-23 refer to "personal communication", which I think is not suitable in this form.

I will skip the reference (personal communication) whereas I will refer to the necessary details of data validation as explained below:

The reason is that even if model resolution does not seem capable of simulating local coastal topographical details, the main characteristics of the large-scale wave evolution are expected to be captured (~~Jean Bidlot, ECMWF, personal communication 2017,~~ based on the validation data used for compiling Fig. 2).

Points_Ref1_08: Table 1 is not needed in the main body of the text Figure1 provides the same information.

Table 1 in contrary to Figure 1, contains the exact names of the RIEN (River Ending) points (with lat / lon) whereas in Figure 1 only the names of the rivers and in most cases these names are different from the names of the RIENs.

Since such topographical details will help the reader to locate easier and as close as possible the points of interest (RIENs), Table 1 could be move to the Technical Supplement.

I will move Table 1 to the Technical Supplement. It will be referenced as shown below:

Additional details can be found in Table 1 of the Technical Supplement containing the exact location (lat, lon) of RIEN points.

Points_Ref1_09: I do not find a clear explanation on which data are grouped under the label hind_com, obs_com and hind_tot. One can guess but a clear description should be given in data and method.

I will add a clear description of data (as new Table 1). See below for details:

First, the (Pearson) correlation between the two source variables (surge & wave) in observations mode is estimated while the same type of correlation is calculated in hindcast mode (see details in Table 1) for inter-comparison.

(New) Table 1. Details and abbreviations of main data sets used in the study.

obs_com	Observations during the common period (1,114 days)
hind_com	Hindcasts during the common period (1,114 days)
hind_tot	Hindcasts during the total period (12,753 days)

Points_Ref1_10: Results section contains 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.

Main parts of Section 4.1 (as most of the technical details contained in the old Section 2.1) could be moved to the new Statistical Supplement.

I will move main parts of Section 4.1 (and most of the technical parts of the old Section 2.1) to the new Statistical Supplement. See below details (referring to changes of Section 4.1 in the main text):

4.1 Main tools for estimating statistical dependence

The main tools for assessing dependence between surge and wave has been a set of matlab routines (mat_chi) for estimating the asymptotic behaviour of statistical dependent variables. Other Matlab routines such as mat_chibar (see details in ~~Sect. 2.1~~ the Statistical Supplement) for assessing the asymptotic behaviour of statistical independent variables were also used and main findings are contained in ~~Tables 3 and Table 5~~ Table 4 and Table 5 of the Technical Supplement). Besides matlab functions additional routines from the statistical package R, namely "taildep" of module extRemes and "chiplot" of module evd (Extreme Value Distributions) were used for estimating and inter-comparing χ values. ~~Utilising for instance the chiplot routine a detailed plot of χ is possible based on a wide range of percentile values. Chiplot can also provide pre-selected confidence intervals in harmony with those considered in matlab routines.~~

~~Since values of χ can be estimated for any lower or upper threshold, initial trials were performed studying the behaviour of χ over a wide range of thresholds. Findings were similar to those contained in Defra TR3 Report (2005), justifying the selection of an optimal threshold for "alpha" (α) equal to 0.1 corresponding to an annual maximum being exceeded in 9 out of 10 years (see Sect. 2.2 for details). This value (0.1) of alpha was considered for both mat_chi and mat_chibar routines when utilising POT (Peaks-Over-Threshold) methodology resulting in an annual maximum of ~2.3 compound events. Such an annual threshold of ~2.3 events corresponds to the top 80 (Top-80) compound events taking place during any (POT separated) day of the total 12,753 days and it was dictated mainly by two factors: the threshold had to be low enough to allow a sufficient number of data points to exceed it for estimating dependence reliably, while being high enough for the data points to be regarded as extremes. Lastly, this threshold (~2.3 events) also proved optimal for providing quite stable dependence graphs. A full set of lag tests was performed for both correlation and dependence.~~ An optimal threshold of ~2.3 events on a yearly basis was found to provide quite stable dependence graphs (see details in the Statistical Supplement). ~~It was found that the~~ The maximum strength of almost any compound (surge and wave) event tends to take place during the same 24-hour (max24) time or during the same 12-hour (max12) period corresponding to zero-lag mode. Exceptions were found for Rhone, Ebro, Danube, Thames and Goeta RIEN points with one-day lag (2 half-days in case of max12), suggesting that storm surge values were (slightly) higher correlated with wave height values of the previous day. Results in Tables and Figures refer to zero-lag values.

Points_Ref1_11: Fig.10 I cannot see the negative and zero dependence values that are mentioned in the text (page40, line 15).

The old Figure 10 (new Figure 9) contained only dependence values (no correlations). Zero and negative values refer to a certain number of correlations contained in the old Table 3 and Table 5 (now moved to the new Technical Supplement as Table 4 and Table 5 respectively) valid for both max12 and max24 configurations.

In the new Figure 9 (old Figure 10) that now contains both correlations and dependence (max24) values, zero correlations are marked by a grey colour whereas negative correlations by a yellow one.

I will compile the new Figure 9 (old Figure 10) containing both correlation and dependence values. For more clarity, the prevailing and dominant components will be skipped since they are also presented analytically in the relevant old Table 7 (new Table 3) in the main text.

Statistical Supplement of

Estimations of statistical dependence as joint return period modulator of compound events. Part I: storm surge and wave height.

Thomas I. Petroliagkis

Correspondence to: Thomas I. Petroliagkis (thomas.petroliagkis@ec.europa.eu)

Contents

1 Statistical dependence (χ)

The main concept of the so-called dependence measure χ (chi) is related to two or more simultaneously observed variables of interest – such as in our case storm surge and wave height – known as observational pairs. If one variable exceeds a certain extreme (high-impact) threshold, then the value of χ represents the risk that the other variable will also exceed a high-impact threshold as explained in Hawkes (2004), Svensson and Jones (2004a & 2004b), Petroliagkis et al. (2016).

Following Coles et al. (2000), if all of the extreme observations of two variables exceed a given threshold at the same time, this indicates total dependence ($\chi = 1$). If the extreme observations of one variable exceed a given threshold but the second variable does not, this indicates total independence ($\chi = 0$). Similarly, if the extreme observations of one variable exceed a given threshold but the other variable produces lower observations than would normally be expected, this indicates negative dependence ($\chi = -1$). In practice, hydro-meteorological analyses based on real data often lead to an assessment of complete independence that could result to an under-estimation of the joint probability of concurrent extreme events, whereas, an assumption of complete dependence could result to an over-estimation of joint probabilities (Beersma and Buishand, 2004). In reality, as variables reach their extreme values, a special methodology of estimating statistical dependence could be utilised. This methodology has been documented by Buishand (1984) and Coles et al. (2000). A brief description of the method based on Coles et al. (2000) is given below.

2 Estimation of dependence (χ)

For bivariate random variables (X, Y) with identical marginal distributions, the dependence measure (χ) can estimate the probability of one variable being extreme provided that the other one is extreme:

$$\chi = \lim_{z \to z^*} \Pr(Y > z \mid X > z) \qquad (1)$$

where z* is the upper limit of the observations of the common marginal distribution.

For obtaining identical marginal distributions, each set of observations is ranked separately and each rank is then divided by the total number of observations resulting in a data transformation with Uniform [0, 1] margins. At this point, it is convenient to consider the bivariate cumulative function $F(x, y) = \mathrm{Prob}(X \leq x, Y \leq y)$ that describes the dependence between X and Y completely. The effect of different marginal distributions can be diminished by assuming the copula function C in the domain [0, 1] x [0, 1] such as:

$$F(x, y) = C\{F_x(x), F_y(y)\} \qquad (2)$$

where F_x and F_y can be any marginal distributions. Such utilisation of the copula function has the same effect as if observations were ranked separately and divided by the total number of observations. In addition, ~~The~~ the copula C contains the complete information about the joint distribution of X and Y and it is invariant to marginal transformation. This means that C is invariant to marginal transformation and it can be described as the joint distribution function of X and Y. Further, X and Y are transformed to new variables U and V with Uniform [0, 1] margins. It follows that the dependence measure $\chi(u)$ for a given threshold u can be given by:

$$\chi(u) = 2 - \frac{\ln \Pr(U \leq u, V \leq u)}{\ln P(U \leq u)} \quad \text{for } 0 \leq u \leq 1 \qquad (3)$$

Taken into account the upper limit of the observations (previously defined as z^* in Eq. 1), the dependence measure $\chi(u)$ will be given by:

$$\chi = \lim_{u \to 1} \chi(u) \qquad (4)$$

Details of deriving Eq. 3 can be found in Coles et al. (2000). Based on Eq. 3, a set of χ values can be evaluated at different quantile levels u ~~(for details see Coles et al., 2000)~~. The selection of a particular level u corresponds to threshold levels (x*, y*) for the two different data series. For applying Eq. 3, the number of appropriate observation-pairs (X, Y) is counted for estimating the numerator and denominator terms (~~Eq. 4 & Eq. 5~~ Eq. 5 & Eq. 6):

$$P(U \leq u, V \leq u) = \frac{\text{Number of } (X, Y) \text{ such that } X \leq x^* \text{ and } Y \leq y^*}{\text{Total number of } (X, Y)} \qquad \cancel{(4)}\ (5)$$

and

$$\ln P(U \leq u) = \frac{1}{2} \ln\left[\frac{\text{Number of } X \leq x^*}{\text{Total number of } X} \cdot \frac{\text{Number of } Y \leq y^*}{\text{Total number of } Y} \right] \qquad \cancel{(5)}\ (6)$$

In this study, a set of routines (mat_chi) based on ~~matlab~~ Matlab software were coded following Eq. 3 to ~~5~~ 6 for estimating χ. Additional modules and routines based on the integrated statistical package R were also used for estimating dependence terms and inter-comparing various parameters. Emphasis was given on the routine "taildep" of the module "extRemes" (https://cran.r-project.org/web/packages/extRemes/extRemes.pdf) that is capable of estimating χ values when a critical percentile (extreme) threshold is considered. Another "powerful" routine capable of providing a variety of dependence graphs and plots (besides single estimated values of χ) has been the routine "chiplot" of the module "evd" (Extreme Value Distributions) of R (https://cran.r-project.org/web/packages/evd/evd.pdf). The routine chiplot is also capable of providing confidence intervals at any preselected level.

Besides estimating values of χ, similar routines (mat_chibar) were coded in ~~matlab~~ Matlab following Coles et al. (2000) for calculating the "sister" attribute of χ, namely chibar ($\bar{\chi}$). Chibar (chi_bar) parameter refers to the statistical dependence of asymptotically independent variables whereas chi (χ) refers to the statistical dependence of asymptotically dependent ones. Details on the estimation of chibar are documented in Coles et al. (2000) whereas examples and how to utilise ($\bar{\chi}$) can be found in Coles (2001). The ~~latter~~ class of asymptotic dependence appears to be the case in Literature, having reached a consensus that there is strong, although not overwhelming, evidence for asymptotic dependence between wave height and surge (Wadsworth et al., 2017).

The concept of asymptotic dependence (χ) is stated with adequate details in Coles et al. (2000). In brief, χ is on the scale [0, 1] with the set (0, 1] corresponding to asymptotic dependence whereas the measure chibar ($\bar{\chi}$) falls within the range [-1, 1] with the set [-1, 1) corresponding to asymptotic independence. That is why the complete pair of χ and $\bar{\chi}$ is required as a summary of extremal dependence:

- $\chi > 0$ & $\bar{\chi} = 1$ reveals asymptotic dependence, in which case the value of χ determines a measure of strength of dependence within the class
- $\chi = 0$ & $\bar{\chi} < 1$ reveals asymptotic independence, in which case the value of $\bar{\chi}$ determines the strength of dependence within the class.

For estimating both χ and $\bar{\chi}$ parameters, the general POT (Peaks-Over-Threshold) methodology was followed. Such an approach (POT) is considered as giving a more accurate estimate of the probability distribution than using the annual maximum series (see details in Stedinger et al., 1993). Applying POT as described in detail in Defra TR1 Report (2005), the selection of an optimal threshold for the data pairs (~2.3 events per year) was adopted as suggested in Defra TR3 Report (2005). Care was taken to force two POT extreme compound events not occurring on consecutive days, but separated by at least three days from each other. Emphasis was also given on the stability of χ (graph) curves ~~as strongly recommended by Prof Pieter Van Gelder of Delft University, Nederlands (personal communication, 2016)~~ identifying the area that dependence was clearly converging to a specific value (no abrupt fluctuations).

Relatively small differences among various estimates made by chiplot of evd (R), taildep of extRemes (R) and mat_chi (~~matlab~~ 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).

3 Selection of critical thresholds

For selecting a threshold u (referring to a critical percentile) as required in Eq. ~~5~~ 3, it seems appropriate to transform the Uniform distribution to an annual maximum non-exceedance probability scale (Defra TR3 Report, 2005). Then the annual maximum non-exceedance probability (α) is defined as:

$$\alpha = \text{Prob (Annual maximum} \leq x) \qquad \cancel{(6)}\ (7)$$

where x is the magnitude of the source variable. Such non-exceedance probability relates to the return period, T_α, as:

$$T_\alpha = 1 / (1 - \alpha) \qquad \cancel{(7)}\ (8)$$

For a transformation from annual maximum to POT series (see details and scope in the previous Sect. ~~2.1~~ 2), we define the "new" non-exceedance probability, the so-called p, referring to a rate of λ events per year, relating to the annual maximum of Eq. ~~6~~ 7, as:

$$\alpha = \exp(-\lambda(1-p)) \qquad \cancel{(8)}\ (9)$$

where 1-p is the "new" exceedance probability of the POT series. The term (1 – p) can be estimated graphically leading to Equation ~~9~~ 10:

$$\lambda(1-p) = (N_e / N) * (i - 0.5) / N_e = (i - 0.5) / N \qquad \cancel{(9)}\ (10)$$

where i, represents the rank of the independent POT events, N_e is the number of POT events while N represents the number of years (see details in Defra TR3 Report, 2005). The independence criterion of two POT events to be separated by at least three days (six half-day intervals in the max12 case) was applied for all river ending points. Combining ~~Eq. 8 and Eq. 9~~ Eq. 9 and Eq. 10, an estimation of α is possible as given by Eq. ~~10~~ 11:

$$\alpha = \exp(-(i - 0.5) / N) \qquad \cancel{(10)}\ (11)$$

Therefore, going after the magnitude of x in Eq. ~~6~~ 7 ~~it~~ is equivalent ~~to~~ as trying to define the magnitude of the POT element with rank i in Eq. ~~10~~ 11 for the same maximum non-exceedance annual probability, alpha (α). After the selection of an optimal threshold (u) based on alpha (α), the estimation of χ is straightforward (Eq. 3). The main idea here is to use χ in a relatively simple formula that ~~uses also~~ also uses as input the individual return periods T_X and T_Y for estimating the joint return period ($T_{X,Y}$), like the formula described by Eq. 11 following White (2007), Australian Rainfall & Runoff Project 18 (2009).

$$T_{XY} = \sqrt{T_X * T_Y / \chi^2} \qquad \cancel{(11)}\ (12)$$

Studying Eq.~~11~~ 12 closely it becomes obvious that dependence is capable of substantially modulating the joint return period. For details and potential limitations of Eq. ~~11~~ 12, see discussions in White (2007), Hawkes (2004), Meadowcroft et al. (2004), Australian Rainfall & Runoff Project 18 (2009). Furthermore, in cases of totally dependent variables, Eq. ~~11~~ 12 yields the common individual return period of source variables as an estimation of the joint return period. An example of how to utilise the formula of Eq. ~~11~~ 12 is given in Sect. 4.2 of the main text for the river ending point of Rhine (NL). Some limitations of Eq. ~~11~~ 12 could be overcome if a more complete formula is used such as Eq. 2.15 for instance taken from White's thesis (2007) but this is above the scope of the current study.

4 Significance

The values of dependence (χ) corresponding to the 5% significance level were estimated using a permutation method as described by ~~Good, 1994~~ Good (1994). As in Defra TR3 Report (2005), 199 permutations of the data were made for each surge-wave pair and a new value of χ was calculated each time. All 199 values of χ were subsequently ranked in descending order and the 5% significance level was defined by selecting the 10th largest value representing the 95% point of the null distribution (the hypothetical distribution occurring if data-pairs were indeed independent). Care was taken to preserve the seasonality since permutation of data was performed by randomly reshuffling intact blocks of one year time period.

5 Confidence intervals

For the estimation of confidence intervals, a well-tested bootstrapping method was applied similar to the permutation method already used for estimating significance (for details see Defra TR3 Report, 2005). This bootstrapping resulted in the generation of many new data-sets (resamples). The original sample of observation-pairs was used as the main (reference) distribution from which the resamples were chosen randomly. A large number of data sets were generated for calculating χ for each of these new data sets. This provided a sample of what would occur for a range of situations. Seasonality was kept intact by sampling in blocks of one year, rather than using individual observation-pairs. The balanced resampling as documented by Fisher (1993) was applied ensuring that each year occurs equally often overall among the total number of bootstrap samples. In total, 199 bootstrap samples of the data were made for each station-pair and a new χ value was calculated each time. The 199 values were subsequently ranked in descending order and the 10 and 190 largest values were accepted as determining the 90% confidence interval.

6 Selection of critical thresholds resulting in the consideration of top-80 events

Since values of dependence (χ) can be estimated for any lower or upper threshold, initial trials were performed studying the behaviour of χ over a wide range of thresholds. Findings were similar to those contained in Defra TR3 Report (2005), justifying the selection of an optimal threshold for "alpha" (α) equal to 0.1 corresponding to an annual maximum being exceeded in 9 out of 10 years (see Sect. 3 of the accompanying Statistical Supplement). This value (0.1) of alpha was considered for both mat_chi (χ) and mat_chibar ($\bar{\chi}$). routines when utilising POT (Peaks-Over-Threshold) methodology resulting in an annual maximum of ~2.3 compound events.

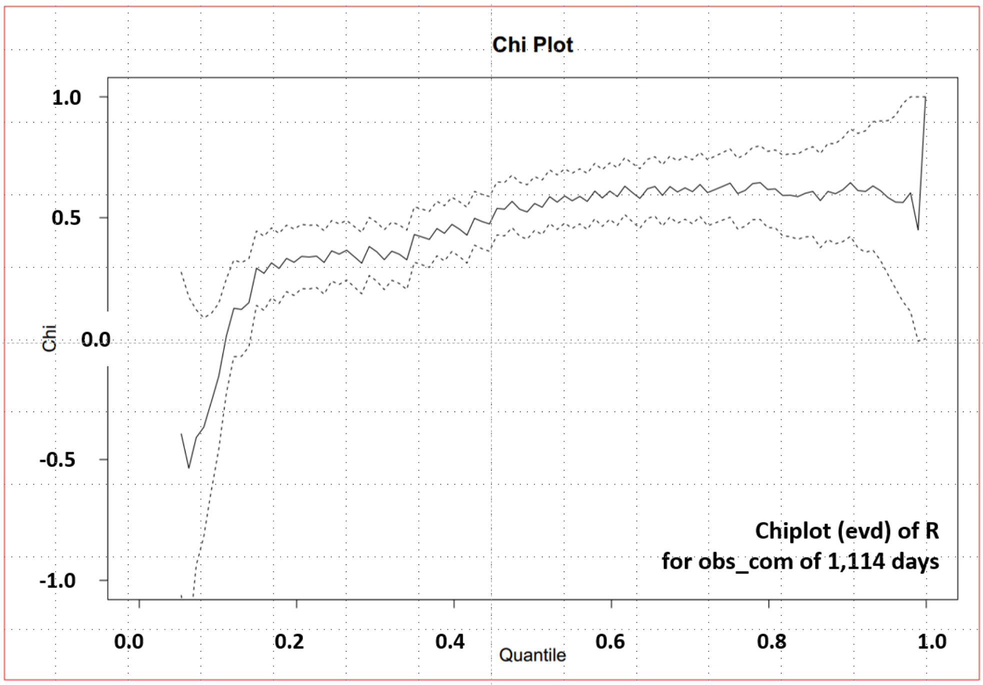
Such an annual 25 threshold of ~2.3 events corresponds to the top 80 (Top-80) compound events taking place during any (POT separated) day of the total 12,753 days and it was dictated mainly by two factors: the threshold had to be low enough to allow a sufficient number of data points to exceed it for estimating dependence reliably, while being high enough for the data points to be regarded as extremes.

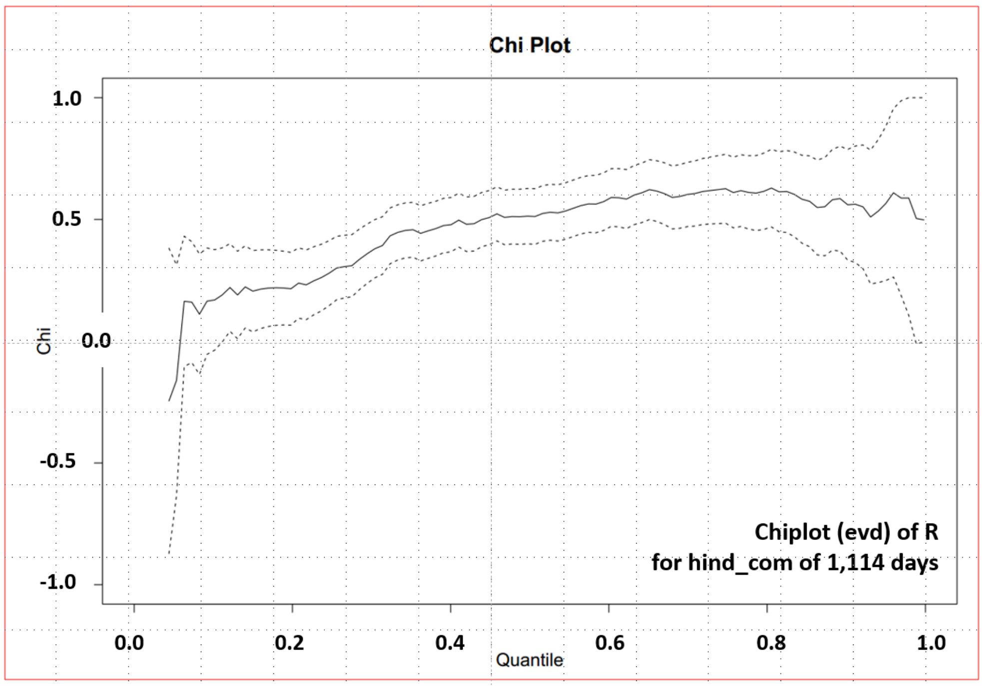
7 Details and examples of the statistical packages used in the study

In this study, a set of routines (mat_chi) based on ~~matlab~~ Matlab software were coded following Eq. 3 to 5 for estimating χ. Additional modules and routines based on the integrated statistical package R were also used for estimating dependence terms and inter-comparing various parameters. Emphasis was given on the routine "taildep" of the module "extRemes" (https://cran.r-project.org/web/packages/extRemes/extRemes.pdf) that is capable of estimating χ values when a critical percentile (extreme) threshold is considered.

Another "powerful" routine capable of providing a variety of dependence graphs and plots (besides single estimated values of χ) has been the routine "chiplot" of the module "evd" (Extreme Value Distributions) of R (https://cran.r-project.org/web/packages/evd/evd.pdf). The routine chiplot is also capable of providing confidence intervals at any preselected level. As mentioned above (Section 2) relatively small differences among various estimates made by chiplot of evd (R), taildep of extRemes (R) and mat_chi (~~matlab~~ Matlab) were found and this most probably is due to the unavoidable dissimilarities between the criteria being imposed on data pairs when applying POT methodology.

Examples of estimated statistical dependence (χ) values between surge (HvH) and wave (LiG) max24 values in obs_com (upper panel), hind_com (middle panel) and in hind_tot (lower panel) mode by chiplot routine of evd module (R) are given in Fig. 1.





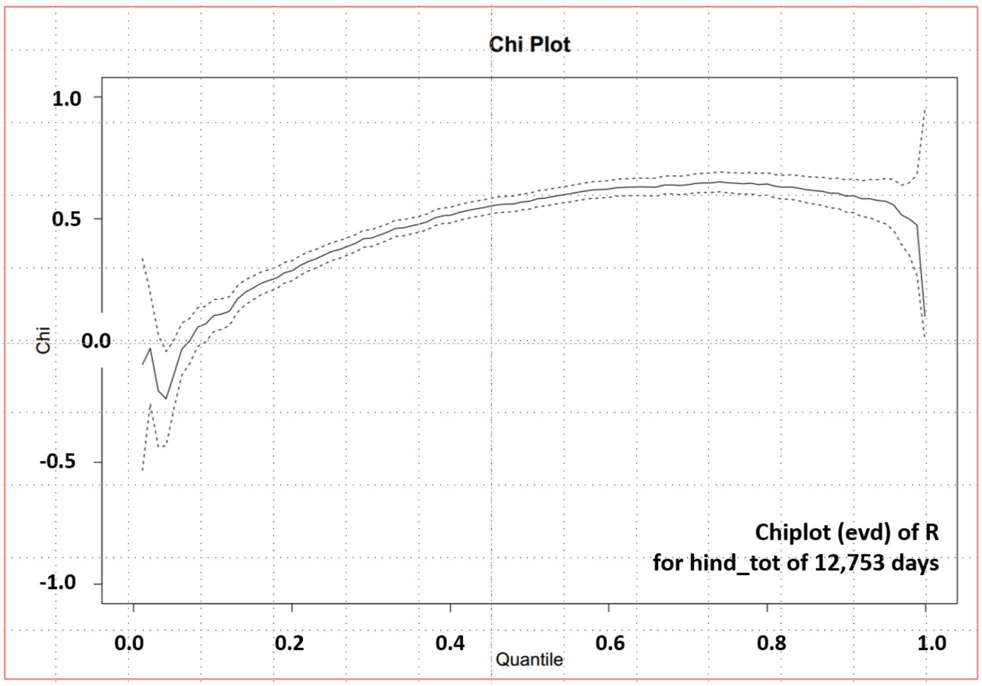


Fig. 1. Estimated χ values between surge (HvH) and wave (LiG) max24 values in c (upper panel) & hind_com (middle panel) and in hind_tot (lower panel) mode by chiplot routine of evd module (R).

Studying closely Fig. 1 it becomes obvious that considerable high values of dependence are estimated over all three (obs_com, hind_com & hind_tot) modes. The importance and implications of such high values of dependence can be demonstrated with an example by considering the total hindcast (hind_tot) series for surge (HvH) and wave (LiG). Utilising the ~~matlab~~ Matlab function "gevfit" an estimation of the return levels having a 100-year return period for surge and wave height variables was made (1.78 and 6.05 metres respectively). Inserting the common return period value (100-year) together with the estimated χ value (0.56) in Eq. 11, the Joint Return Period (JRP) of such a compound event (surge ≥ 1.78 metres and significant wave height ≥ 6.05 metres) was estimated at ~179 years.

Such a value (~179 years) is significantly different from the value of 10,000 years representing the estimated JRP assuming that surge and wave variables were totally independent. In a case like this (of independent events), the dependence would have been equal to zero and the JRP would be given by the product of their individual probabilities (Blank, 1982).

References

Australian Rainfall & Runoff Project 18: Coastal Processes and Severe Weather Events: Discussion Paper, Water Technology report to Australia Rainfall & Runoff (2009) referring to the report of Department of Science, IT, Innovation and the Arts – Science Delivery (October 2012) "Coincident Flooding in Queensland: Joint probability and dependence methodologies" (https://www.longpaddock.qld.gov.au/coastalimpacts/inundation/coincident_flood_technical_review.pdf), 2009.

Beersma, J.J. and Buishand, T.A.: Joint probability of precipitation and discharge deficits in the Netherlands. Water Resour Res. doi:10.1029/2004WR003265, 2004.

Blank, L.: Statistical Procedures for Engineering, Management, and Science, McGraw Hill, 1982.

Buishand, T.A.: Bivariate extreme-value data and the station-year method. Journal of Hydrology, 69, 77-95, 1984.

Coles, S.G., Heffernan, J. and Tawn, J.A.: Dependence measures for extreme value analyses. Extremes, 2, 339-365, 2000.

Coles, S.G.: An Introduction to Statistical Modelling of Extreme Values. Springer Series in Statistics. Springer Verlag London. 208p, 2001.

Defra TR1 Report by Hawkes, P.J. and Svensson, C.: Joint probability: dependence mapping & best practice. R & D Final Technical Report FD2308/TR1 to Defra. HR Wallingford and CEH Wallingford, U.K. (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3428_TRP_pdf.sflb.ashx), 2005.

Defra TR3 Report by Svensson, C. and Jones, D.A.: Joint Probability: Dependence between extreme sea surge, river flow and precipitation: a study in south and west Britain. Defra/Environment Agency R & D Technical Report FD2308/TR3, 62 pp. + appendices (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3430_TRP_pdf.sflb.ashx), 2005.

Fisher, N.I.: Some modern statistical techniques for testing and estimation. Chapter 8 in Statistical analysis of circular data, pp. 199-218. Cambridge: Cambridge University Press, 1993.

Good P.: Permutation tests. New York: Springer, 1994.

Hawkes, P.J.: Use of joint probability methods for flood & coastal defence: a guide to best practice. R&D Interim Technical Report FD2308/TR2 to Defra. HR Wallingford, U.K. (http://www.estuary-guide.net/pdfs/FD2308_3429_TRP.pdf), 2004.

Meadowcroft, I., Hawkes, P.J. and Surendran, S.: Joint probability best practice guide: practical approaches for assessing combined sources of risk for flood and coastal risk managers. In Proceedings from the Defra 39th Flood & Coastal Management Conference, York, UK, July 2004, 6A.2.1– 6A.2.12, 2004.

Petroliagkis, T.I., Voukouvalas, E., Disperati, J. and Bidlot, J.: Joint Probabilities of Storm Surge, Significant Wave Height and River Discharge Components of Coastal Flooding Events, JRC Technical Report EUR 27824 EN, doi:10.2788/677778, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC100839/lbna27824enn.pdf, 2016.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou E.: Frequency analysis of extreme events. In Handbook of Hydrology (ed. D R Maidment), pp. 18.1-18.66. London: McGraw-Hill, 1993.

Svensson, C. and Jones, D.A.: Dependence between sea surge, river flow & precipitation in south & west Britain. Hydrol. Earth Sys. Sci., 8, 973–992, 2004a.

Svensson, C. and Jones, D.A.: Sensitivity to storm track of the dependence between extreme sea surges and river flows around Britain. In Hydrology: Science and Practice for the 21st Century, Vol. 1. Proc. from the British Hydrological Society’s international conference, London, UK, July 2004, 239a–245a (addendum), 2004b.

Wadsworth, J.L., Tawn, J.A., Davison, A.C. and Elton, D.M.: Modelling across extremal dependence classes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79: 149–175. doi:10.1111/rssb.12157, 2017.

White, C.J.: The use of joint probability analysis to predict flood frequency in estuaries and tidal rivers, PhD Thesis, School of Civil Engineering and the Environment, University of Southampton, p343, 2007.

Technical Supplement of

Estimations of statistical dependence as joint return period modulator of compound events. Part I: storm surge and wave height.

Thomas I. Petroliagkis

Correspondence to: Thomas I. Petroliagkis (thomas.petroliagkis@ec.europa.eu)

Contents

1 Details of RIEN (RIver ENding) point positions

The current statistical (dependence) analysis is focused over 32 river ending points that have been selected to cover a variety of riverine and estuary areas along European coasts. The sea areas used in the study refer to the Mediterranean Sea (central and north Adriatic Sea, Balearic Sea, Alboran Sea and Gulf of Lion), West Iberian, North Iberian, Bay of Biscay, Irish Sea, Bristol Channel, English Channel, North Sea, Norwegian Sea, Baltic Sea and Black Sea. A map showing the position of RIEN (RIver ENding) points used in the study is shown in Fig. 1 (of the main text). Additional details can be found in Table 1 (current Technical Supplement) containing the exact location (lat, lon) of all RIEN points

Table 1. Positions (lat, lon) of 32 RIEN points used in the study. Names refer to river ending areas.

	RIEN	lat	lon		RIEN	lat	lon
1	Po Della Pila	44.96	12.49	17	Muir Eireann	52.65	-6.22
2	Madonna Del Ponte	43.83	13.05	18	Wallasey	53.44	-3.04
3	Martinsicuro	42.84	13.93	19	Severn Bridge	51.61	-2.65
4	Aries	43.34	4.84	20	Fort Picklecombe	50.34	-4.17
5	El Foix	41.20	1.67	21	Exmouth	50.62	-3.42
6	Illa de Buda	40.71	0.89	22	Christchurch District	50.72	-1.74
7	Rio De Velez	36.72	-4.11	23	Dieppe	49.91	1.09
8	Matosinhos	41.18	-8.71	24	South Tynesid	55.01	-1.43
9	Carcavelos	38.69	-9.26	25	Spurm Point	53.57	0.11
10	Setubal	38.53	-8.89	26	Sheerness	51.45	0.74
11	San Bruno	37.18	-7.39	27	Western Scheldt	51.43	3.55
12	Punta Del Arenal	43.47	-5.07	28	Rockanje	51.87	4.01
13	Concarneau	47.86	-3.92	29	Wurster Arm	53.65	8.14
14	Riviere De Belon	47.81	-3.72	30	Kattegat	57.77	11.76
15	Larmor-Plage	47.71	-3.38	31	Trondheimsfjord	63.32	9.82
16	Musura Bay	45.22	29.73	32	Vanhankaupunginselka	60.24	24.99

2 Capability of hindcasts to identify and resolve compound events of surge and wave.

As already mentioned long-period water level data coinciding with wave observations directly or very close to the exact sites of interest (RIEN points) were not available with the exception of the Rhine River (RIEN). For this RIEN, concurrent (close-by) observations with no gaps of sea level, astronomical tide, storm surge, and wave height from a close-by wave buoy were available for a period of about 3 years (1,114 days).

In Table 2, extreme storm surge (above 98.5% percentile) values for both observations and hindcasts for HvH tide gauge station over the common time interval of 1,114 days are shown. Same way extreme significant wave height (above 98.5% percentile) values for both observations and hindcasts for LiG wave buoy station over the common time interval are contained in Table 3.

Table 2. Extreme storm surge (above 98.5% percentile) values for both observations and hindcasts for HvH tide gauge station over the common time interval of 1,114 days. Compound events of surge and wave are marked by yellow shade.

#	Date	Observations	hindcasts
1	12 Nov 2010	1.38	1.10
2	4 Feb 2011	1.20	1.00
3	27 Nov 2011	1.25	1.04
4	28 Nov 2011	0.98	0.93
5	3 Dec 2011	1.08	1.03
6	7 Dec 2011	1.10	0.95
7	9 Dec 2011	1.45	1.23
8	29 Dec 2011	1.23	1.03
9	3 Jan 2012	1.07	0.47
10	4 Jan 2012	1.46	1.16
11	5 Jan 2012	1.66	1.59
12	6 Jan 2012	1.37	1.57
13	21 Jan 2012	1.09	1.02
14	22 Jan 2012	1.00	1.07
15	30 Jan 2013	1.07	0.73
16	10 Sep 2013	0.96	0.59

Compound events of surge and wave are marked by orange shade (in both Table 2 & 3) based on joint observations of storm surge and significant wave height. It becomes obvious that hindcasts were able to resolve all seven (7) compound events that took place during the common time period of 1,114 days.

Table 3. As in Table 2, but for significant wave height for LiG wave buoy station.

#	Date	Observations	hindcasts
1	12 Nov 2010	4.79	3.99
2	14 Jul 2011	4.61	3.34
3	7 Oct 2011	4.34	3.34
4	7 Dec 2011	5.06	4.83
5	8 Dec 2011	4.49	3.87
6	9 Dec 2011	4.17	3.53
7	24 Dec 2011	4.37	3.27
8	29 Dec 2011	4.18	3.46
9	30 Dec 2011	4.66	3.84
10	4 Jan 2012	4.31	4.02
11	5 Jan 2012	5.14	4.79
12	6 Jan 2012	4.55	4.90
13	20 Jan 2012	4.15	2.81
14	31 Aug 2012	4.11	3.24
15	24 Sep 2012	4.61	3.43
16	25 Nov 2012	4.36	4.09

An extra investigation based on extreme values of observations (during the common time interval of 1,114 days) exceeding a variety of percentile values (for the RIEN of Rhine River) showed that both storm surge and their corresponding wave height hindcasts were able to capture almost all of the 24-hour extremes on the same (correct) day but with a weaker intensity (i.e., with a correct footprint of lesser intensity).

3 Analytical values of correlation and statistical dependence based on Matlab routines.

A necessary split of results had to be made for a better and easier visualisation due to the relatively large amount of RIEN points to fit in one single Table. This split also revealed the distinct differences between southern and northern coastal European areas. Details of both correlations and dependencies found over southern RIEN points are presented analytically in Table 4 and Table 5 based on matlab routines. In Table 4 and Table 5, correlation (corr) and dependence (chi) values for both max12 and max24 intervals are presented together with critical threshold (thrs), significance (sig) and 95% confidence level (lower & upper) max24 values. Referring to correlation values, a large amount of variability is evident in both max12 and max24 modes

Table 4. Correlation and statistical dependence values for storm surge and significant wave heights over Mediterranean (ADR: Adriatic Sea – GOL: Gulf of Lion – BAL: Balearic Sea – ALB: Alboran Sea), West and North Iberian coasts (WIB & NIB), Bay of Biscay (BOB) and Black Sea (BLK) based on ~~matlab~~ Matlab routines.

			max12			max24						
	RIEN	**sea**	**corr**	**thrs**	**chi**	**corr**	**thrs**	**chi**	**chibar**	**sig**	**lower**	**upper**
1	Po	ADR	0.26	97.4	0.28	0.39	97.1	0.29	0.43	0.02	0.21	0.37
2	Metauro	ADR	0.23	96.8	0.26	0.35	95.7	0.22	0.30	0.05	0.03	0.35
3	Vibrata	ADR	0.23	96.6	0.35	0.37	96.5	0.32	0.36	0.04	0.23	0.37
4	Rhone	GOL	0.08	94.6	0.20	0.13	93.8	0.21	0.17	0.04	0.13	0.30
5	Foix	BAL	0.09	92.2	0.03	0.10	91.2	0.03	0.05	0.03	0.00	0.08
6	Ebro	BAL	0.04	94.7	0.19	0.12	94.5	0.22	0.22	0.03	0.10	0.30
7	Velez	ALB	0.02	93.9	0.19	0.06	93.1	0.11	0.13	0.04	0.05	0.17
8	Douro	WIB	-0.18	97.0	0.30	-0.06	95.7	0.30	0.30	0.05	0.11	0.38
9	Tagus	WIB	-0.30	94.3	0.05	-0.22	93.7	0.14	0.16	0.03	0.09	0.22
10	Sado	WIB	-0.26	94.9	0.10	-0.19	93.9	0.13	0.17	0.03	0.06	0.21
11	Guadiana	WIB	-0.04	95.9	0.22	0.03	95.7	0.28	0.29	0.02	0.15	0.36
12	Sella	NIB	-0.25	93.2	0.10	-0.17	86.2	0.14	0.07	0.05	0.07	0.19
13	Moros	BOB	0.07	96.2	0.32	0.22	96.2	0.30	0.34	0.03	0.17	0.39
14	Aven	BOB	0.13	97.0	0.34	0.25	96.7	0.35	0.39	0.01	0.23	0.42
15	Blavet	BOB	0.11	96.5	0.33	0.25	96.7	0.34	0.39	0.02	0.22	0.40
16	Danube	BLK	-0.01	96.7	0.21	0.09	96.3	0.24	0.35	0.05	0.07	0.38

Table 5. As in Table 4 but for Irish Sea IRS), Bristol Channel (BRC), English Channel (ENC), North Sea (NRS), Norwegian Sea (NOS) and Baltic Sea (BAS). Owena stands for Owenavarragh RIEN (IE) while Goeta is Goeta Aelv RIEN (ES).

			max12			**max24**						
	RIEN	**sea**	**corr**	**thrs**	**chi**	**corr**	**thrs**	**chi**	**chibar**	**sig**	**lower**	**upper**
17	Owena	IRS	0.50	98.4	0.46	0.59	97.9	0.45	0.53	0.05	0.30	0.55
18	Mersey	IRS	0.45	98.2	0.43	0.56	97.4	0.43	0.48	0.03	0.29	0.52
19	Severn	BRC	0.19	96.1	0.29	0.30	94.9	0.30	0.24	0.04	0.22	0.35
20	Tamar	ENC	0.28	97.8	0.35	0.39	96.9	0.35	0.41	0.02	0.24	0.49
21	Exe	ENC	0.31	97.9	0.38	0.41	97.1	0.40	0.43	0.03	0.29	0.54
22	Avon	ENC	0.37	98.1	0.44	0.50	97.9	0.48	0.55	0.04	0.35	0.58
23	Bethune	ENC	0.59	99.1	0.62	0.68	98.8	0.64	0.77	0.02	0.55	0.73
24	Tyne	NRS	0.14	91.7	0.31	0.28	94.5	0.26	0.21	0.05	0.10	0.39
25	Humber	NRS	0.18	97.3	0.35	0.38	96.6	0.35	0.37	0.04	0.20	0.49
26	Thames	NRS	-0.10	92.6	0.22	0.06	92.7	0.22	0.11	0.05	0.11	0.31
27	Schelde	NRS	0.31	97.6	0.54	0.54	97.5	0.53	0.50	0.01	0.45	0.61
28	Rhine	NRS	0.52	98.5	0.57	0.67	98.0	0.56	0.57	0.03	0.41	0.64
29	Weser	NRS	0.56	99.0	0.58	0.65	98.5	0.56	0.69	0.02	0.42	0.63
30	Goeta	NRS	0.43	97.2	0.53	0.55	96.8	0.51	0.39	0.05	0.44	0.61
31	Orkla	NOS	0.35	97.6	0.46	0.46	97.0	0.41	0.43	0.03	0.33	0.50
32	Vantaa	BAS	0.30	97.0	0.43	0.44	96.9	0.44	0.42	0.03	0.36	0.50

4 Analytical values of correlation and statistical dependence based mainly on R routines.

Details of both correlations and dependencies found over southern RIEN points are presented analytically in Table 6 and Table 7 based mainly on R routines.

Table 6. As in Table 4, but based mainly on R (chiplot & taildep) routines. Ensemble mean (comb) values of dependence are also shown (last column).

			R				MAT	ENS
	RIEN	**sea**	**lower**	**upper**	**chiplot**	**taildep**	**mat_chi**	**comb**
1	Po	ADR	0.13	0.34	0.23	0.27	0.29	0.26
2	Metauro	ADR	0.08	0.26	0.17	0.22	0.22	0.20
3	Vibrata	ADR	0.13	0.32	0.23	0.36	0.32	0.30
4	Rhone	GOL	0.06	0.21	0.14	0.22	0.21	0.19
5	Foix	BAL	0.01	0.13	0.07	0.16	0.03	0.09
6	Ebro	BAL	0.14	0.30	0.22	0.28	0.22	0.24
7	Velez	ALB	0.03	0.18	0.10	0.16	0.11	0.12
8	Douro	WIB	0.17	0.33	0.26	0.31	0.30	0.29
9	Tagus	WIB	0.07	0.21	0.14	0.22	0.14	0.17
10	Sado	WIB	0.08	0.21	0.14	0.21	0.13	0.17
11	Guadiana	WIB	0.19	0.34	0.27	0.32	0.28	0.29
12	Sella	NIB	0.05	0.19	0.12	0.18	0.14	0.15
13	Moros	BOB	0.14	0.32	0.23	0.28	0.30	0.27
14	Aven	BOB	0.18	0.37	0.27	0.31	0.35	0.31
15	Blavet	BOB	0.17	0.36	0.27	0.30	0.34	0.30
16	Danube	BLK	0.13	0.32	0.23	0.26	0.24	0.24

Table 7. As in Table 5, but based mainly on R (chiplot & taildep) routines. Ensemble mean (comb) values of dependence are also shown (last column).

			R				MAT	ENS
	RIEN	**sea**	**lower**	**upper**	**chiplot**	**taildep**	**mat_chi**	**comb**
17	Owena	IRS	0.26	0.52	0.39	0.40	0.45	0.41
18	Mersey	IRS	0.26	0.48	0.38	0.38	0.43	0.40
19	Severn	BRC	0.16	0.32	0.24	0.30	0.30	0.28
20	Tamar	ENC	0.21	0.41	0.31	0.34	0.35	0.33
21	Exe	ENC	0.25	0.46	0.36	0.38	0.40	0.38
22	Avon	ENC	0.33	0.57	0.45	0.46	0.48	0.46
23	Bethune	ENC	0.49	0.80	0.64	0.66	0.64	0.65
24	Tyne	NRS	0.11	0.27	0.19	0.26	0.26	0.24
25	Humber	NRS	0.20	0.40	0.30	0.33	0.35	0.33
26	Thames	NRS	0.08	0.22	0.15	0.25	0.22	0.21
27	Schelde	NRS	0.36	0.58	0.47	0.48	0.53	0.49
28	Rhine	NRS	0.41	0.64	0.52	0.54	0.56	0.54
29	Weser	NRS	0.40	0.67	0.55	0.54	0.56	0.55
30	Goeta	NRS	0.35	0.53	0.44	0.46	0.51	0.47
31	Orkla	NOS	0.25	0.45	0.35	0.38	0.41	0.38
32	Vantaa	BAS	0.27	0.48	0.37	0.40	0.44	0.40

For the analysis of results, the ensemble mean value of χ (by averaging mat_chi, chiplot and taildep values) is taken as a reference value (contained in the last column of Table 6 and Table 7). The different categories of correlation and dependence used later in the text (and in Figure 9) refers to the categorisation adapted by Defra TR1 Report (2005).

References

Defra TR1 Report by Hawkes, P.J. and Svensson, C.: Joint probability: dependence mapping & best practice. R & D Final Technical Report FD2308/TR1 to Defra. HR Wallingford and CEH Wallingford, U.K. (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3428_TRP_pdf.sflb.ashx), 2005.

Defra TR3 Report by Svensson, C. and Jones, D.A.: Joint Probability: Dependence between extreme sea surge, river flow and precipitation: a study in south and west Britain. Defra/Environment Agency R & D Technical Report FD2308/TR3, 62 pp. + appendices (http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2308_3430_TRP_pdf.sflb.ashx), 2005.