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 $F(x, y) = Prob(X \leq x, Y \leq y)$ 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:

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

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 the current Part I, 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):

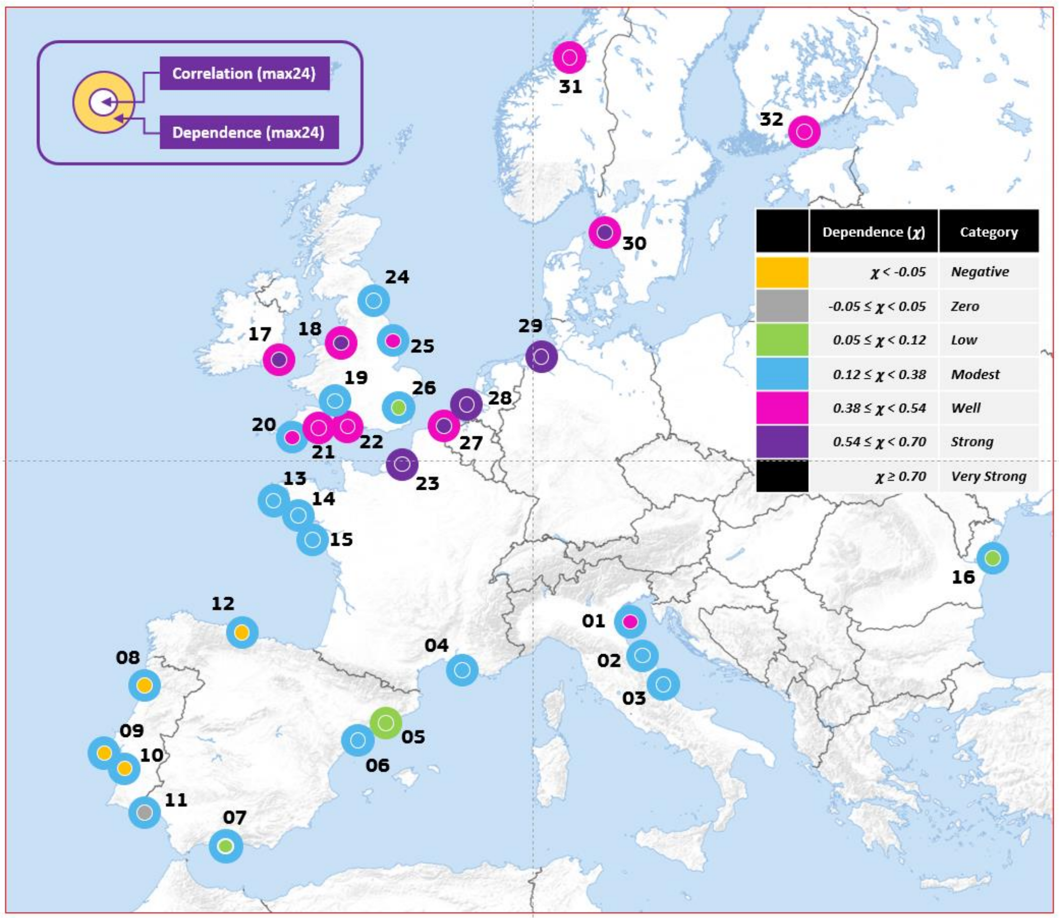
Horsburgh, K. J. and Wilson, C.: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, J. Geophys. Res., 112, C08003, doi:10.1029/2006JC004033, 2007.

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.

....

The new updated main text (manuscript) combined with the two new supplements (Statistical & Technical) has been uploaded as

nhess-2017-177-manuscript-version4

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

The paper addresses compound events defined by combined high surges and high wind waves along European coastlines, especially in estuaries/river mouths. Statistical methods are used to investigate joint probabilities of compound events and the statistical dependency, since flood risk is not a function of one parameter (storm surges with peak value and duration) but usually of more (e.g. wind waves, river runoff). Large scale weather systems can cause either high storm surges or high wind waves and further more high precipitation and river runoff/discharges. Two sets of almost 35-year hindcasts of storm surges and wave heights were used to analyse the correlation and statistical dependency. As expected the frequency of the occurrence of the top compound events in different coastal areas were found to be higher during the winter months. In the introduction the hydrological and meteorological conditions for high wind waves and extreme tidal surge events which can occur simultaneously with extreme precipitation events and high river flows (compound events) leading to increased flood risk is highlighted clearly. But the paper and the used methodology focused only on very few parameters. What is the background of the generalization? The subject of the paper is interesting yet a little confusing especially in the context of coastal engineering therefor the manuscript should be major improved. The paper and its structure is not easy to understand and the description of different data sets (and different time spans) of observed and modelled hindcast data is confusing (e.g. a lot of unusual abbreviations). The number of tables and especially the huge amount of data should be reduced as they are displayed in figures. The selected 32 stations at the end of the rivers or estuaries cover a wide variety of geographical areas and meteorological, oceanographical and hydrological (currents and tides) systems in coastal zones along European coasts. E.g. the tidal range varies from nearly zero to some meters and within the deterministic part of compound events in comparison to the stochastic part (surges, wind waves and river flow) of these compound events. Further discussion of the deterministic and the stochastic part of the compound events and the effects in the statistical analyses (dependency of different parameters) is recommended (page 41, line 26-30). In general I agree completely with reviewer # 1!

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

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

Comments (Ref2)

Com_Ref2_01: The description whether storm surge and/or wind waves are capable of reproducing extreme values is incomplete (e.g. river runoff?). It has to be explained, why river runoff is not taken into account!

This study is the first part (Part I: storm surge and wave height) of investigating how statistical dependence can act as modulator referring to the joint return period of compound events. It is clear that this is the case of surge and wave events, so, no river runoff was taken into account. For the preparation of Part II (storm surge and river discharge) and Part III (wave height and river discharge) the effect of runoff will be included and be given special emphasis. I truly believe that such a separate investigation (by parts) allows for a deeper and better understanding of the different components contributing to a compound coastal event. Further, a study including all three components would have become too lengthy and difficult for the reader to follow.

I will explain in more detail (in the Introduction) the reasoning behind this separate investigation of the different components contributing to coastal compound events. I will also refer to the preparation of Part II (storm surge and river discharge) and Part III (wave height and river discharge).

… ~~This study~~ The current work focuses on data preparation, parameter selection, methodology application and estimation of both correlation and statistical dependence between source variables. It also focuses on the prevailing (higher frequency) and dominant (higher intensity) low-level wind conditions over a set of preselected (top 80) extreme compound events. The critical time period during which such extremes take place is also analysed based on monthly frequency values of occurrence. The dependence analysis utilises 32 river ending points selected to cover a variety of geographical areas along European coasts. The variable-pairs presented in this report, which include enough information for calculations, are storm surge and wave height, relevant to most coastal flood defence studies. Two main time intervals were considered for the estimation of maximum values: the half-day interval (max12) and the one-day interval (max24) …

… This study represents the first part (i.e., Part I) of the investigation while Part II (storm surge and river discharge) and Part III (wave height and river discharge) are to follow. The reasoning behind such a separate investigation (by parts) is to allow the reader for a deeper and better understanding of the interaction between different components contributing to a compound coastal event.

Com_Ref2_02: In the context of the paper a very interesting problem is discussed where copula functions should be taken into account, so far only a simple approach for copula functions has been taken into consideration, the discussion of different copula functions within the scope of the addressed topic is to be considered, more references to copula functions could be helpful (e.g. Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E.

(2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nature Climate Change, 5(12), 1093-1097).

The study follows the methodology proposed by Coles et al. (2000) where the basic theory behind the utilisation of an optimal copula function refers to Nelsen (1998), Joe (1997) and Currie (1999). I agree that the inclusion of more references as the suggested one, i.e., Wahl et al. (2015) definitely helps the reader to get more insight in the use of copulas when joint probability methodologies are taken into account.

I will include the suggested reference (Wahl et al., 2015). In addition, I will include the extra references of Nelsen (1998), Joe (1997) and Currie (1999) in the main text.

References

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

- Currie, J.E., "Directory of coefficients of tail dependence," Department of Mathematics and Statistics Technical Report, ST-99-06, Lancaster University, 1999.

- Joe, H., Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.

- Nelsen, R.B., An Introduction to Copulas, Springer-Verlag, New York, 1998.

- Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nature Climate Change, 5(12), 1093-1097).

Com_Ref2_03: A point of criticism is that the meteorological conditions and oceanographic system were not sufficiently described and the temporal developments of surges and wind waves are also not clearly described. E.g. in fig 10 and 12 it is shown that for the Weser (RIEN 29) the dependence for the prevailing (highest frequency) and dominant (highest intensity) wind during the top 80 extreme compound events are caused by wind direction from WSW. This is completely in contrast to my experience and has to be explained (same for e.g. RIEN 3, 12,: : : 23, 24, : : :. and 32)! From my point of view, it would be advisable to consider a subarea, e.g. only the North Sea, and after a successful investigation of the statistical dependence then implicate other areas.

As a general comment: Directions falling in the WSW category do not count for the total percentage of the Top-80 events but besides this, there exists a logical explanation since the 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. 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 continues to exist longer than a day). 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 northwest (northern) direction. I have checked the validity of this during the second, third and even the fourth day of a 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 stress this unavoidable disagreement due to the veering of the wind and provide necessary explanations for such discrepancy.

... 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 ...

Com_Ref2_04: (Length of observations/hindcasts) As I understood the water level data/storm surge/wind waves: The 32 RIEN (Table 1, page 10) were selected mainly because of their proximity to tidal gauges, although many of them cannot be evaluated due to lack of long-term measurements. For most RIENs, there are no data from nearby open wave buoys. Only for the Rhine (RIEN 28) are the tide and sea data (without *data gaps) available from a nearby wave buoy for a period of 3 years. The validation of the combined hindcasts (tide and wind waves) was done on the basis of measured data at the Rhine (NL) was done on the tidal data at Hoek von Holland (HvH), wave buoy: Lichteiland (LiG) over a period of ~ 3 years on measurement data without gaps and comparison of daily and half-day maxima. The generation of the hindcast of storm surge data was done with Delft3D-Flow (according to Vousdoukas et al. (2016) and the generation of the hindcast of the wind waves data was done with ECWAM wave model (according to Bidlot et al. (2006), Bidlot (2012), ECMWF (2015), Philips (2017)), e.g. ~36 years, wind- and pressure fields from ERA-Interim (ERAI) (time resolution: 1 h, spatial resolution: 28x28 km, fixed water level, signif. wave height, max. wave height, mean wave period, mean wave direction and validation based on available records from 101 wave buoys throughout Europe + North Atlantic (1996-2015) (Fig. 2)) The overlapping period of the two hindcasts (~ 35 years) was used in statistical analysis.*

The methodology of the research (using the hindcast data sets and observed data) has to be explained more detailed and especially what that means for the interpretation of the results (for all 32 RIEN). A time series of observed water level and wave buoy of only 3 years and only for one station in the area at Hoek van Holland seems to me as being not sufficient and much too short for comparison/evaluation with the modelled (hind cast) data and the conclusions. There should much more field data (water level, surges, wind waves, river runoff) available around the 32 RIEN!

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.

Approx. 2.3 "extreme events" (at least 3 days between peaks) per year (total 80 top events) were chosen. It has to be explained more detailed why 2.3 "extreme events" where chosen and what that means for the interpretation of the results

Extreme value analysis can be carried out using two types of data series (Bezak et al., 2014), annual maximums (MA) or flows above a certain threshold (POT for Peak Over Threshold). The POT model used in this study can be composed of the Poisson, binomial and negative binomial distributions for modelling the annual number of events above threshold, and of exponential or generalized Pareto distributions for magnitudes of exceedances.

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). I will also add the relevant reference (Bezak et al., 2014) the one referring to 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).

Such a 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

References

Bezak, N., Brilly, M., and Sraj, M., 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. Hydrological Sciences Journal, 59 (5), 959-977.

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.

Improvements (Ref2)

Impr_Ref2_01: The number of tables and graphs should be reduced and more summarized.

I will reduce / convert tables and graphs to a more summarised form. List of changes in Figures and Tables are listed below:

- Table 1 will be moved in the new Technical Supplement to provide the reader with exact details of the selected RIENs (river ending points).

- New Table 1 (Details and Abbreviations of the main data sets) will be incorporated in the main text.

- Figure 9 (results based on statistical packages of R) will be moved in the new Statistical Supplement in Section 7 (Details and examples of the statistical packages used in the study)

- Table 3 will be moved in the new Statistical Supplement as new Table 4 (Section 3).

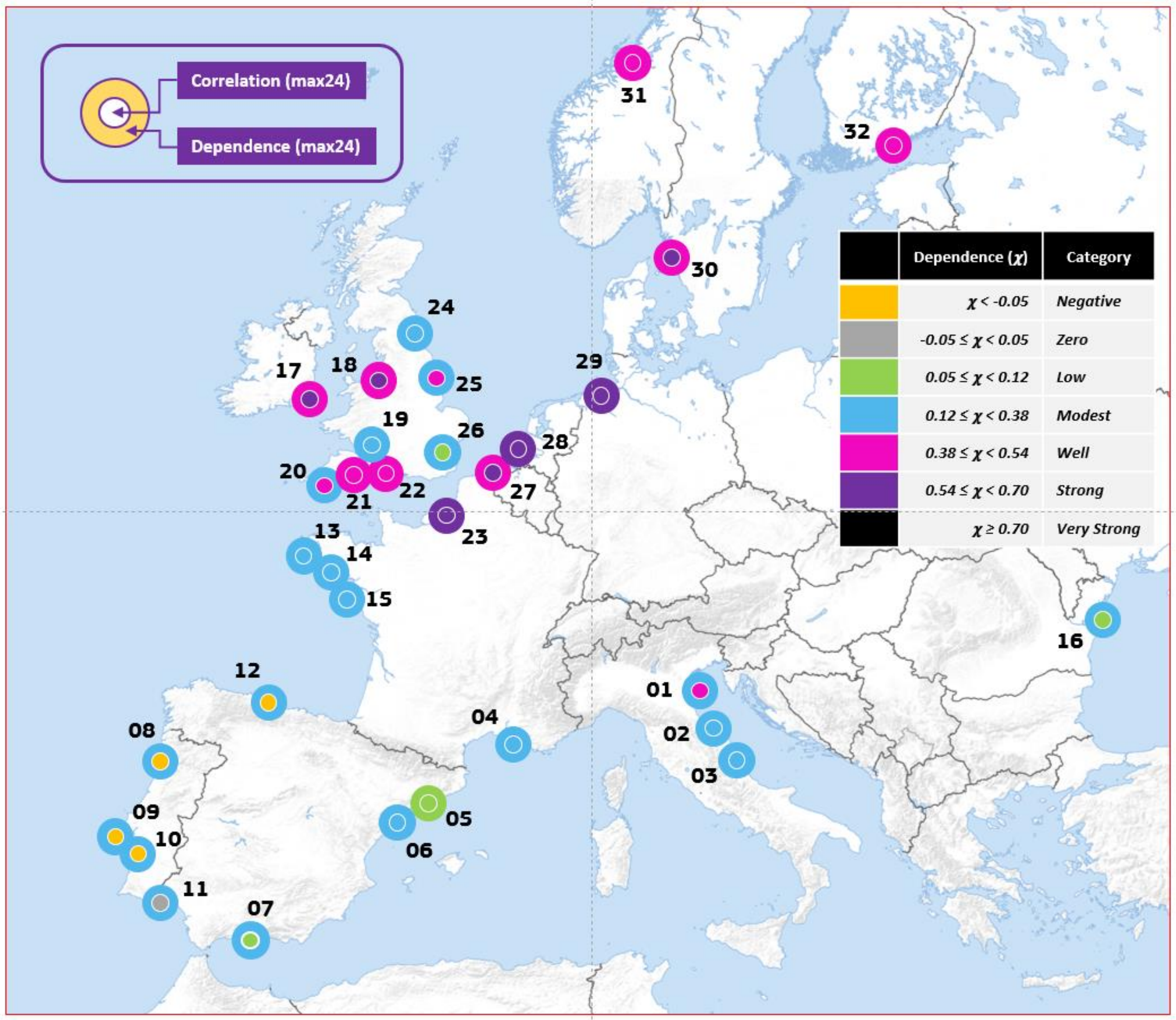
- Table 4 will be moved in the new Technical Supplement as new Table 6 (Section 3).

- Table 5 will be moved in the Technical Supplement as new Table 5 (Section 3).

- Table 6 will be moved in the Technical Supplement as new Table 7 (Section 3).

- Figure 12 will be skipped whereas old Table 7 will be kept as new Table 3 (in the main text) since it 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.

- Old Figure 10 (now new Figure 9 in the main text) will include the spatial distribution of both correlation and statistical dependence (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)

Impr_Ref2_02: The paper is not easy to understand for a wide diversified audience, the length of the paper is too long and has partly too much redundancy (e.g. table 1 and fig. 1).

I will reduce the length of the main paper by creating a separate Statistical Supplement and an additional Technical Supplement. These two new Supplements will help the reader to understand easier the main concept and findings of the current work. Redundant parts will be merged, shortened and improved.

Impr_Ref2_03: The pure agreement between hindcast and observation of daily maximum of storm surges in Fig. 4 has to be explained.

The pure agreement between hindcasts and observations is a clear indication of the model's (Delft3D-Flow) capability to simulate efficiently observations in hindcast mode having as input parameters (wind components and mean sea level pressure) from the ECMWF ERA-Interim reanalysis data set.

Indicative examples of such capabilities can be seen in Table 2 and Table 3 of Section 2 of the new Technical Supplement revealing that hindcasts above all were capable of identifying and resolving all seven (7) compound events (based on 98.5% percentile threshold) that took place during the common time interval of 1,114 days over HvH area of interest.

I will explain and stress this capability of Delft3D-Flow model of resolving daily maximum of storm surge observations in the main test referring also to Table 2 (Section 2) of the new Technical Supplement.

Impr_Ref2_04: Why are small storm surges, e.g. below 0.5 m are taken in to account?

In Figure 4, the capability of hindcasts to simulate correctly observations was done over the full range of observations, since it is important to show that model hindcasts are capable to perform well over any part of observations.

With the help of such models, it should be anticipated to have two validated sets of hindcasts resulting to the determination of the correct sign and strength of both correlation and statistical dependence.

I will point out that validation of both hindcast sets is done over the full spectrum of observations since the capability of the model to simulate correctly observations should refer to any part of the spectrum values.

Impr_Ref2_05: What is the definition of a storm surge?

Storm surge is the abnormal rise in seawater level during a storm, measured as the height of the water above the normal predicted astronomical tide https://oceanservice.noaa.gov/facts/stormsurge-stormtide.html.

Same wise the definition of significant wave height will be also included (see below).

In physical oceanography, the significant wave height (SWH or Hs) is defined traditionally as the mean wave height (trough to crest) of the highest third of the waves (https://en.wikipedia.org/wiki/Significant_wave_height).

I will include the definition of storm surge (and significant wave height) in the Introduction and provide the relevant (site) references.

Impr_Ref2_06: What is the reason to use the storm between 25th December 2012 and 24th January 2013? It is just an example chosen for demonstrating how a compound event looks like and how it is related to the prevailing synoptic conditions (Storm Emil). Further, it is an example of a compound event that lasts for three consecutive days (from 4 to 6 January 2012) as shown in Table 2 and Table 3 of the new Technical Supplement. During de-clustering this event will be counting only once and it will refer to its first date that this event took place (4 January 2012). I will point out the concept of this multi-purpose demonstrating example and give emphasis in the de-clustering concept.

Impr_Ref2_07: The pure agreement between hindcast and observation of daily maximum of the significant wave height in Fig. 6 has to be explained. As in the previous case (Imp_02_03), the pure agreement between hindcasts and observations is a clear indication of the model's (ECMWF / ECWAM) capability to simulate efficiently observations in hindcast mode having as input parameters (wind components) from the ECMWF ERA-Interim reanalysis data set. Once again, indicative examples of such capabilities can be seen in Table 2 and Table 3 of Section 2 of the new Technical Supplement revealing that hindcasts above all were capable of identifying and resolving all seven (7) compound events (based on 98.5% percentile threshold) that took place during the common time interval of 1,114 days over HvH area of interest. I will explain and stress this capability of ECMWF / ECWAM model of resolving daily maximum of significant wave height observations.

Impr_Ref2_08: Fig. 8: The fairly pure agreement (chi) of the statistical dependence (chi) of storms surge and significant wave height between observation and hindcasts has to be explained. The fairly pure agreement between chi values estimated by observations (of surge and waves) and hindcasts (of surge and waves) is a clear indication that hindcasts were found capable of resolving and estimating both the correct type and strength of correlation and dependence between source variables. I will point out the capability of the hindcasts to resolve and estimate the correct type and strength of correlation and dependence and stress the significance of such an agreement between dependence values estimated from observations and hindcasts.

Impr_Ref2_09: Fig. 9: For the lower and higher quantiles the chi plots have to be explained and discussed.

Values of dependence in the area of lower and higher quantiles seem (and somehow expected) to be quite unstable due the sparse of data.

I will explain and stress the behaviour of chi in lower and higher percentiles. Emphasis will be given on the stability of chi (graph) curves by identifying the area that dependence is clearly converging to a specific value (with no abrupt fluctuations).

Impr_Ref2_10: Fig. 10: I do not find the category dependence "negative" and "zero".

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 new Table 4 and Table 5 of the new Technical Supplement) valid for both max12 and max24 configurations.

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

I will produce the new combined Figure 9 (in place of the 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 old relevant Table 7 (now new Table 3 in the main text).

Impr_Ref2_11: Symbol and wind N to NNW is not necessary.

I will keep the 16 main and secondary directions of the wind (N – NNE – NE – ENE – E – ESE – SE – SSE – S – SSW – SW – WSW – W – WNW – NW – NNW).

Impr_Ref2_12: The description of tables and figures should be improved.

I will improve the description of both tables and figures accordingly. This will be also applied for the new updated Tables and Figures. A full description of the updated Tables and Figures is contained in author's reply to Impr_Ref2_01 comment (in improvements suggested by Ref 02 comments).

Suggestions (Ref2)

Sugg_Ref2_01: I do not find a clear definition of highest intensity, page 34, row 2 and page 41, row 9, does it mean only the dominant wind? Direction and/or speed?

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 and add the 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.

Sugg_Ref2_02: I do not find a clear definition of negative bias: Systematically underestimated parameter?

Bias is the difference between the mean of the forecasts and the mean of the observations. It could be expressed as a percentage of the mean observation. Also known as overall bias, systematic bias, or unconditional bias (http://www.cawcr.gov.au/projects/verification/).

I will provide the definition and include the relevant (site) reference.

Minor Improvements (Ref2)

Mimp_Ref2_01: page 2 row 18 "This is"

I will correct it.

Mimp_Ref2_02: page 5/6 row 19/1 "Matlab"

I will correct it.

Mimp_Ref2_03: page 7 row 22 "also uses"

I will correct it.

Mimp_Ref2_04: page 8 row3 "… Good (1994)"

I will correct it.

Mimp_Ref2_05: page 14 row 14 "to the"

I will correct it.

Mimp_Ref2_06: page 16 rows 10-14 "Storm Emil" as well as page 18 rows 1 and 30

I will correct it.

Mimp_Ref2_07: p.42, row 10: providing "us"?

I will delete the word "us".

....

The new updated main text (manuscript) combined with the two new supplements (Statistical & Technical) has been uploaded as

nhess-2017-177-manuscript-version4

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

Thomas I. Petroliagkis

Joint Research Center, Ispra, I-21027, Italy

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

Abstract. The possibility of utilising statistical dependence methods in coastal flood hazard calculations is investigated, since flood risk is rarely a function of just one source variable but usually two or more. Source variables in most cases are not independent as they may be driven by the same weather event, so their dependence, which is capable of modulating their joint return period, has to be estimated before the calculation of their joint probability. Dependence and correlation may differ substantially from one another since dependence is focused heavily on tail (extreme) percentiles. The statistical analysis between surge and wave is performed over 32 river ending points along European coasts. Two sets of almost 35-year hindcasts of storm surge and wave height were ~~adapted~~ adopted and results are presented by means of analytical tables and maps referring to both correlation and statistical dependence values. Further, the top 80 compound events were defined for each river ending point. Their frequency of occurrence was found to be distinctly higher during the cold months while their main low-level flow characteristics appear to be mainly in harmony with the transient nature of storms and their tracks. Overall, significantly strong values of positive correlations and dependencies were found over the Irish Sea, English Channel, south coasts of the North Sea, Norwegian Sea and Baltic Sea, with compound events taking place in a zero-lag mode. For the rest, mostly positive moderate dependence values were estimated even if a considerable number of them had correlations of almost zero or even negative value.

1 Introduction

In the coastal and inter-tidal zones, high waves and extreme tidal surge events can occur simultaneously with extreme precipitation events and high river flows, leading to increased flood severity, duration or frequency as highlighted in Svensson and Jones (2002, 2003, 2004a, 2004b, 2005), Hawkes and Tawn (2000), Hawkes et al. (2005). These interactions are generally referred to as coincident or compound events (IPCC, 2012). In the current Part I, compound events of surge and wave are those events that coincidently are above a certain upper percentile criterion (representing a critical threshold). A key component of any coincident event assessment is to understand the historical relationships between the different factors that may lead to a compound flood event. However, assumptions are often made regarding how these different factors and variables coincide or combine, typically leading to either an under- or over-estimation of the probability of flooding (Coles et al., 2000). In reality, while some events may indeed occur independently from one another, others involve an interaction, or may have compounding consequences when they occur simultaneously, and need to be treated as partially dependent for the estimation of their joint probability or joint return period (https://www.niwa.co.nz/natural-hazards/faq/what-is-a-return-period).

Joint probability values provide the likelihood of source variables taking high values simultaneously resulting in a situation where flooding may occur. Acceptance of joint probability methods has been relatively sparse so far mainly due to the lack of information on dependence among source variables and the intrinsic difficulty in usage and interpretation of the methods as pointed out in Australian Rainfall & Runoff Project 18 (2009), Bevacqua et al., (2017). The main concept of dependence as presented by Reed (1999) refers to the tendency for critical values of source variables to occur at the same time resulting in an increase in frequency of an extreme event. ~~This~~ This is because dependence is able of modulating the joint return period as documented in Hawkes (2004), Meadowcroft et al. (2004), White (2007), Australian Rainfall & Runoff Project 18 (2009).

The method for estimating the probability of extreme values from a single variable has been well understood and documented (Coles, 2001). Such probability is usually expressed in the form of a return period. In a similar way, the (joint) probability of two variables producing high or extreme values together, assuming to be fully independent or fully dependent, is also considered straightforward as explained in Defra TR0 Report (2003). On the other hand, examples of coincident flood event studies, which incorporate a measure of the relationship between the input variables, are generally limited due to the complexity of the broader coincident events problem (Bevacqua et al., 2017). Assessing the probability of flooding from the joint occurrence of high waves and high sea level values for instance is not an easy process, as high waves and storm surge tides may be attributed to the same prevailing storm system; thus, independence cannot and should not be assumed. Further, it is more complicated to estimate such conditional (joint) probabilities than those referring to totally independent events (http://onlinestatbook.com/2/probability/basic.html). However, some approachable and user-friendly methods seem to exist for quantifying the statistical dependence between the input variables as noted in Hawkes (2004), White (2007) and applied by Zheng et al. (2013 & 2014) and Klerk et al. (2015).

In the case of independent events, the chance of one event occurring is not changed by the occurrence of the other event. However, if the occurrence of one event is dependent on the occurrence of a second event then the events are termed conditional even if their correlation might be equal to zero. It should be stressed that correlation and dependence might differ substantially from one another. Two source variables may have low correlation but there may exist considerable statistical dependence between them referring to their upper percentiles where actually extremes reside. Further, it should be well established by now that correlation coefficients measure the degree of straight line or linear relationship only and that there are situations in which correlations are zero but where strong nonlinear relationships exist among variables (Drouet Mari and Kotz, 2004).

Assuming independence between input variables might underestimate considerably the likelihood of flooding resulting in higher risk for the coastal community, since the conditional probability of both events occurring at the same time is different from the product of their individual probabilities (Blank, 1982). Similarly, assuming total dependence could be too conservative (Beersma and Buishand, 2004). What someone should anticipate is the fact that dependence is likely to occur when different processes are linked to some common weather (forcing) conditions. It may also arise when the same process is studied at different spatial locations or over different periods (Coles et al., 2000). In an estuarine or riverine area, an example would be a storm accompanied by high winds and intense precipitation phenomena. For such cases where two (or more) variables, capable of producing high-impact events, are not totally independent or totally dependent, but may be partially dependent, probabilistic approaches are limited in both their reliability and scope (White 2007).

In this work, the possibility of utilizing statistical dependence methods in coastal flood hazard calculations is investigated, since an estimation of the joint probability (joint return period) is necessary for the calculation of compound flood hazard in a coastal area. Such an approach points to taking into account the variability and exact nature of extreme conditions. The basic idea behind joint probability theory is to identify extreme data within each of the input variables and statistically correlate their linkages and risk of simultaneous occurrence. Therefore, it seems quite important to find an appropriate way to undertake this task. Understanding such risks, created by the combination of extreme events, is crucial for the design of adequate and cost effective river and coastal defences as well as for the true estimate of flood risk as highlighted in Merz et al. (2009), Australian Rainfall & Runoff Project 18, (2009).

~~This study~~ The current work focuses on data preparation, parameter selection, methodology application and estimation of both correlation and statistical dependence between source variables. It also focuses on the prevailing (higher frequency) and dominant (higher intensity) low-level wind conditions over a set of preselected (top 80) extreme compound events. The critical time period during which such extremes take place is also analysed based on monthly frequency values of occurrence. The dependence analysis utilises 32 river ending points selected to cover a variety of geographical areas along European coasts. The variable-pairs presented in this report, which include enough information for calculations, are storm surge and wave height, relevant to most coastal flood defence studies. Two main time intervals were considered for the estimation of maximum values: the half-day interval (max12) and the one-day interval (max24).

This study represents the first part (i.e., Part I) of the investigation while Part II (storm surge and river discharge) and Part III (wave height and river discharge) are to follow. The reasoning behind such a separate investigation (by parts) is to allow the reader for a deeper and better understanding of the interaction between different components contributing to a compound coastal event.

In Sect. 2, the concept and implications of estimating statistical dependence are documented, while in Sect. 3, the data and methods used are presented. Results are shown in Sect. 4, while discussion and conclusions are contained in Sect. 5.

2 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~~ as the one ~~has been~~ documented ~~by~~ in Buishand (1984) ~~and Coles et al. (2000)~~. A brief description of ~~the~~ this method based on Coles et al. (2000) is ~~given below~~ contained in the Statistical Supplement while the basic theory behind the utilisation of an optimal copula function refers to Nelsen (1998), Joe (1997), Currie (1999), Wahl et al. (2015).

2.1 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) = 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:

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

where F_x and F_y can be any marginal distributions. The copula C contains complete information about the joint distribution of X and Y and it is invariant to marginal transformation. It follows that the dependence measure χ(u) for a given threshold u can be given by:

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

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):

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

and

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

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.

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 refers to the statistical dependence of asymptotically independent variables whereas chi (χ) refers to the statistical dependence of asymptotically dependent ones. The latter 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).

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 areas 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) 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).

2.2 Selection of critical thresholds

For selecting a threshold u (referring to a critical percentile) as required in Eq. 5, 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:

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

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

$$T_\alpha = 1 / (1 - \alpha) \qquad (7)$$

For a transformation from annual maximum to POT series (see details and scope in the previous Sect. 2.1), 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, as:

$$\alpha = \exp(-\lambda (1 - p)) \qquad (8)$$

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

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

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, an estimation of α is possible as given by Eq. 10:

$$\alpha = \exp(-(i - 0.5) / N) \qquad (10)$$

Therefore, going after the magnitude of x in Eq. 6 it is equivalent to trying to define the magnitude of the POT element with rank i in Eq. 10 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).

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

Studying Eq.11 closely it becomes obvious that dependence is capable of substantially modulating the joint return period. For details and potential limitations of Eq. 11, 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 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 is given in Sect. 4.2 for the river ending point of Rhine (NL). Some limitations of Eq. 11 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.

2.3 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.

2.4 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.

3 Data and methodology

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. These points were selected mainly for their proximity to tide gauge recorders although not many observations were found suitable to be exploited due to the lack of long-period coincident wave (buoy) observations in the near-by area. 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 ~~all~~ RIEN (RIver ENding) points used in the study is shown in Fig. 1. Additional details can be found in Table 1 of the Technical Supplement containing the exact location (lat, lon) of all RIEN points.

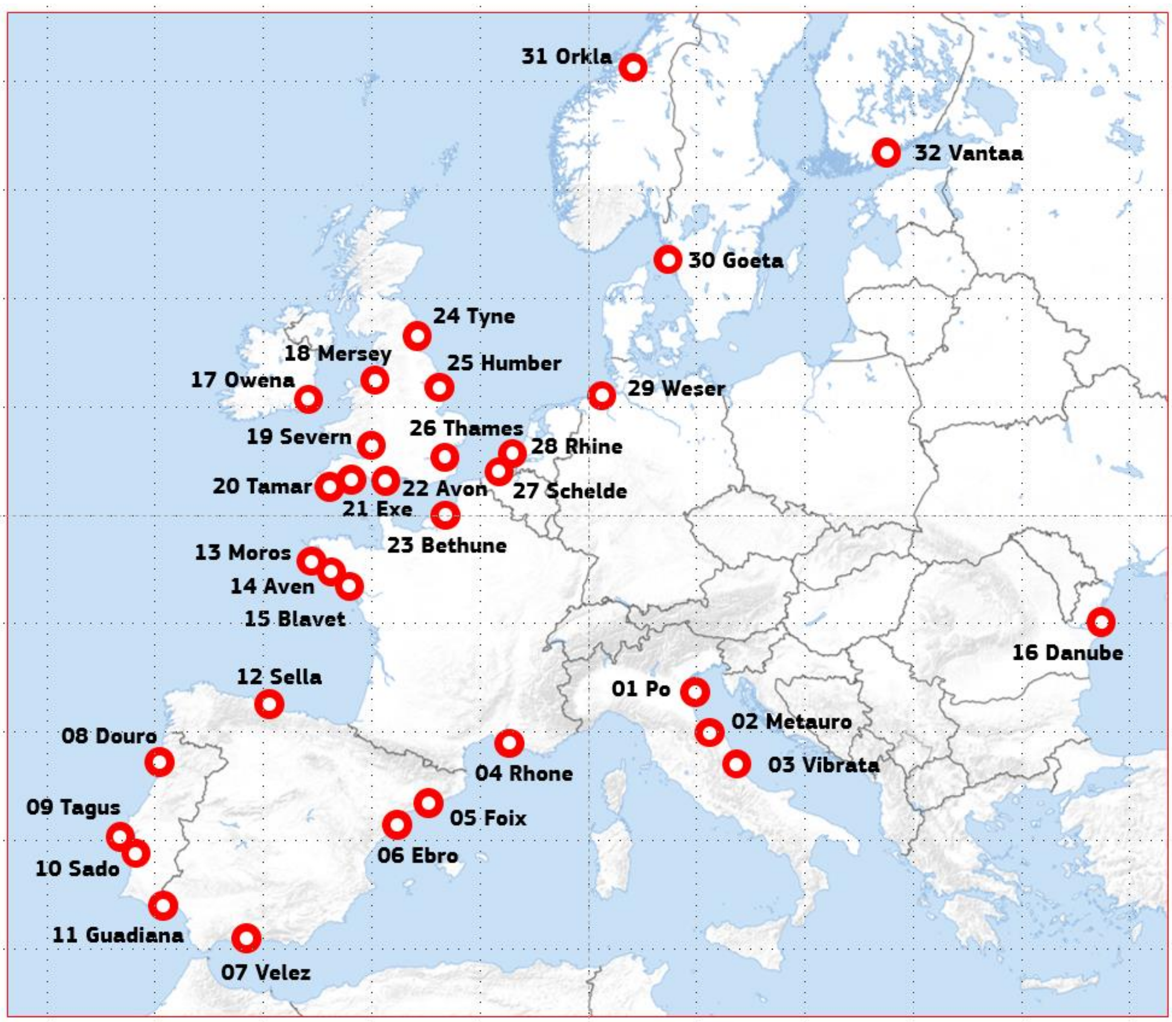


Fig. 1. Positions of the 32 RIEN (river ending) points used in the study. Names refer to river names.

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).

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

	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

3.1 Storm surge hindcasts

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 (Delft3DFlow) uses as input both ERA-Interim (Dee et al., 2011) wind and pressure fields.

For assessing dependence, a full set of coincident observation data is needed over a relatively long time period (at least five years) for the primary variables, as pointed out in Defra TR2 Report (2005). Such a demanding requirement is too difficult if not impossible to be fulfilled only by observational data over all 32 RIEN points, so, the methodology of simulating data observations by modelling (hindcasts) was applied resulting in a set of long period model simulations (hindcasts) for the two primary variables (surge and wave).

For storm surge hindcasts, the Delft3D-Flow hydrodynamic module of the open source model Delft3D (Deltares, 2014) was used to compile storm surge time-series due to the combined effect of the wind and the atmospheric pressure gradient. The model (Delft3D-Flow) has been used successfully in similar applications (hindcasts) in the past (Sembiring et al., 2015). In our case, long-term storm surge series for the 32 RIEN points were compiled from a similar hindcast set to that used in Vousdoukas et al. (2016). This was obtained by forcing the Deflt3D-Flow module by 6-hourly wind and pressure fields retrieved from the ECMWF ERA-Interim reanalysis data set (Dee et al., 2011). The ERA-Interim (ERAI) is a global atmospheric reanalysis from 1979 to present. ERAI’s main products include global atmospheric and surface parameters from 1 January 1979 to present, at T255 spectral resolution (~75 x 75 km) on 60 vertical levels.

Storm surge hindcasts (1 January 1980 to 30 November 2014) span a total interval of 12,753 days (~35 years), having a time separation of 3 hours with a spatial resolution of about 0.2 degrees (~25 x 25 km) along the European coastline and the NE Atlantic Ocean areas. Hindcast storm surge levels were validated with measurements from 110 tide gauges from the JRC Sea Level Database (http://webcritech.jrc.ec.europa.eu/SeaLevelsDb). Details can be found in Vousdoukas et al. (2016). The relative rms (root mean squared) error for more than 105 stations was found to be less than 20% and for more than 60 stations less than 15%. More specifically, the validation performance of hindcasts is contained in the scatter plot of Fig. 4 (Vousdoukas et al., 2016) for both RMS error in m (a) and as a percentage of the SSL (Storm Surge Level) range (b) for all the available tidal gauge stations. Further, even if there were cases where some extreme storm surge levels were underestimated by the hindcasts, the overall model performance is considered to be satisfactory.

3.2 Wave height hindcasts

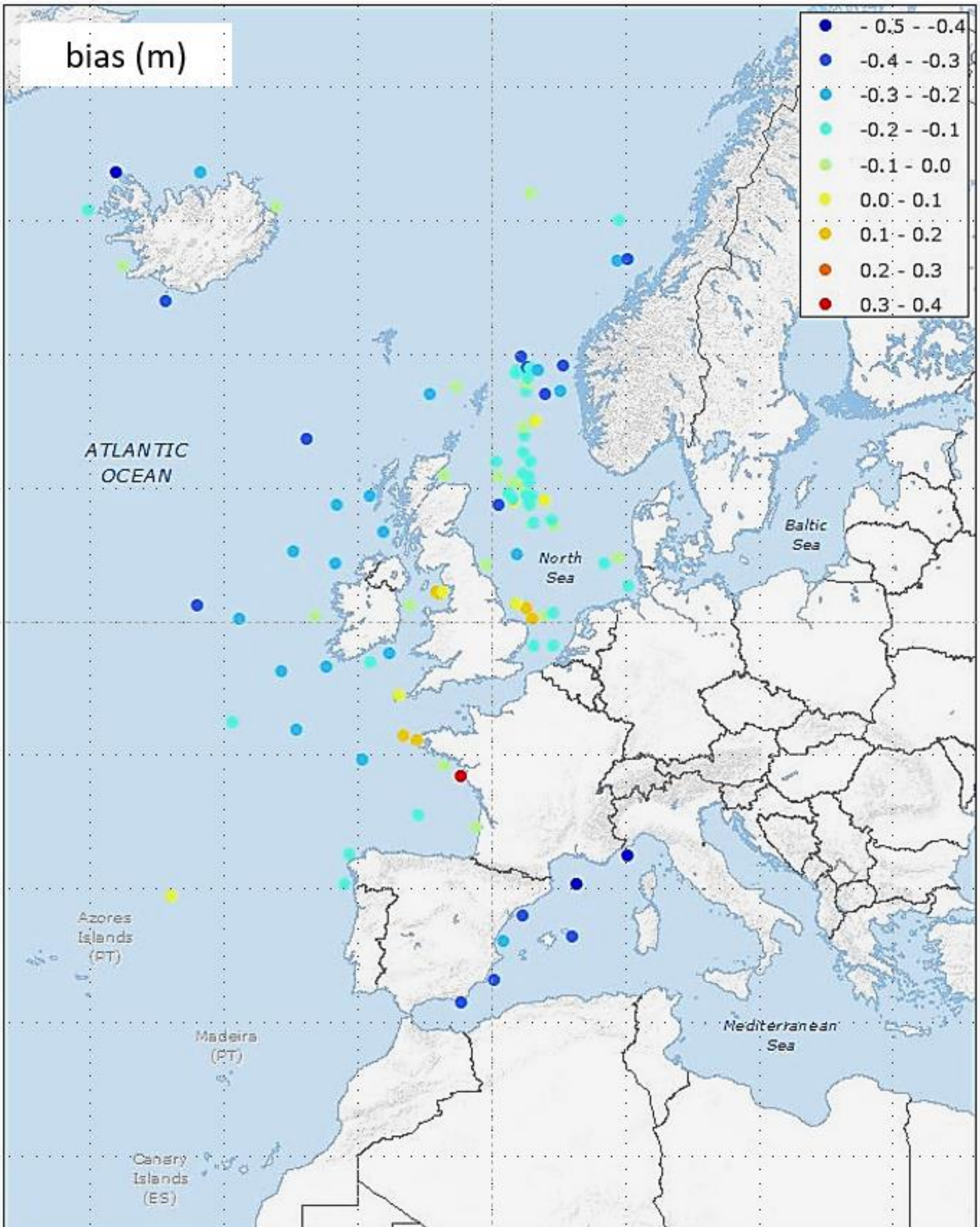
In many applications, a selection of heights and periods of the higher waves in a wave train seem to be of practical significance. For this reason, the average height of the highest one-third of the waves, after eliminating the ripples and waves of height less than one foot is considered as a useful statistical measure. This average is commonly named as "significant wave height" (Sverdrup and Munk, 1946) and it is utilised in the current study.

As in the case of storm surge, global fields of one3-hourly (significant) wave data were assembled by utilizing a set of hindcasts produced with the latest stand-alone version of ECWAM wave model (for details see Bidlot et al., 2006, Bidlot, 2012 and ECMWF, 2015, Philips et al., 2017). The ECWAM model was run on a 0.25 degrees lat-lon global grid (~28 x 28 km) with fixed water depth (mean bathymetry, i.e., no surges or tides) being forced by neutral wind fields (as forcing terms) extracted from the ERAI reanalysis. Due to such resolution limitations, the model may not represent the best source of wave data for a particular single coastal location, but it does offer consistent coverage over the area of this study within an acceptable degree of accuracy. 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 wave in situ observations data provided by Dr Jean-Raymond Bidlot (ECMWF) and used for validation and compiling Fig. 2). For more details on wave validation and verification data, see https://www.ecmwf.int/en/newsletter/150/meteorology/twenty-one-years-wave-forecast-verification.

For each RIEN point, 3-hourly hindcast wave data time series for the period of 1 January 1979 to 31 December 2015 were assembled (13,149 days) by considering the closest model grid (sea) point, with no missing records. The resulted records consist of significant and maximum wave height values, mean wave period and mean wave direction. The ECWAM model has been configured in its CY41R1 parametrization cycle employing 30 frequencies and 36 directions for the wave spectra (ECMWF, 2015).

Validation of wave hindcasts was made utilising a set of available data collected from 101 buoys over European and NE Atlantic sea areas during the period from 1996 to 2015. The exact position of the buoys used in validation is shown in Fig. 2. Both bias and rms error scores were considered and results are also shown in Fig. 2, for bias (upper panel) and for rms error (lower panel). Both scores (bias & rms error) suggest that the model's performance was satisfactory, although bias is lagging slightly in quality compared to rms error mainly due to the weak ERA-iInterim winds that seems to affect the bias more than rms error.

It should be noted that the maximum common time interval of 12,753 days (~35 years) for surge and wave variables was considered for the statistical (dependence) analysis over the 32 RIEN points of this study. Further, as already stated above, both sets of hindcasts had already been validated (Vousdoukas et al., 2016, Philips et al., 2017), so, emphasis was given in demonstrating the methodology of estimating the type and strength of statistical dependence. Such an investigating approach was performed initially over the ending point of Rhine River (NL) with very satisfactory results (see next Sect. 3.3) while the same approach (of estimating statistical dependence) was adopted for the rest of ending points (RIENs) of the study.


bias (m)
- 0.5 - -0.4
-0.4 - -0.3
-0.3 - -0.2
-0.2 - -0.1
-0.1 - 0.0
0.0 - 0.1
0.1 - 0.2
0.2 - 0.3
0.3 - 0.4
ATLANTIC OCEAN
North Sea
Baltic Sea
Azores Islands (PT)
Madeira (PT)
Canary Islands (ES)
Mediterranean Sea

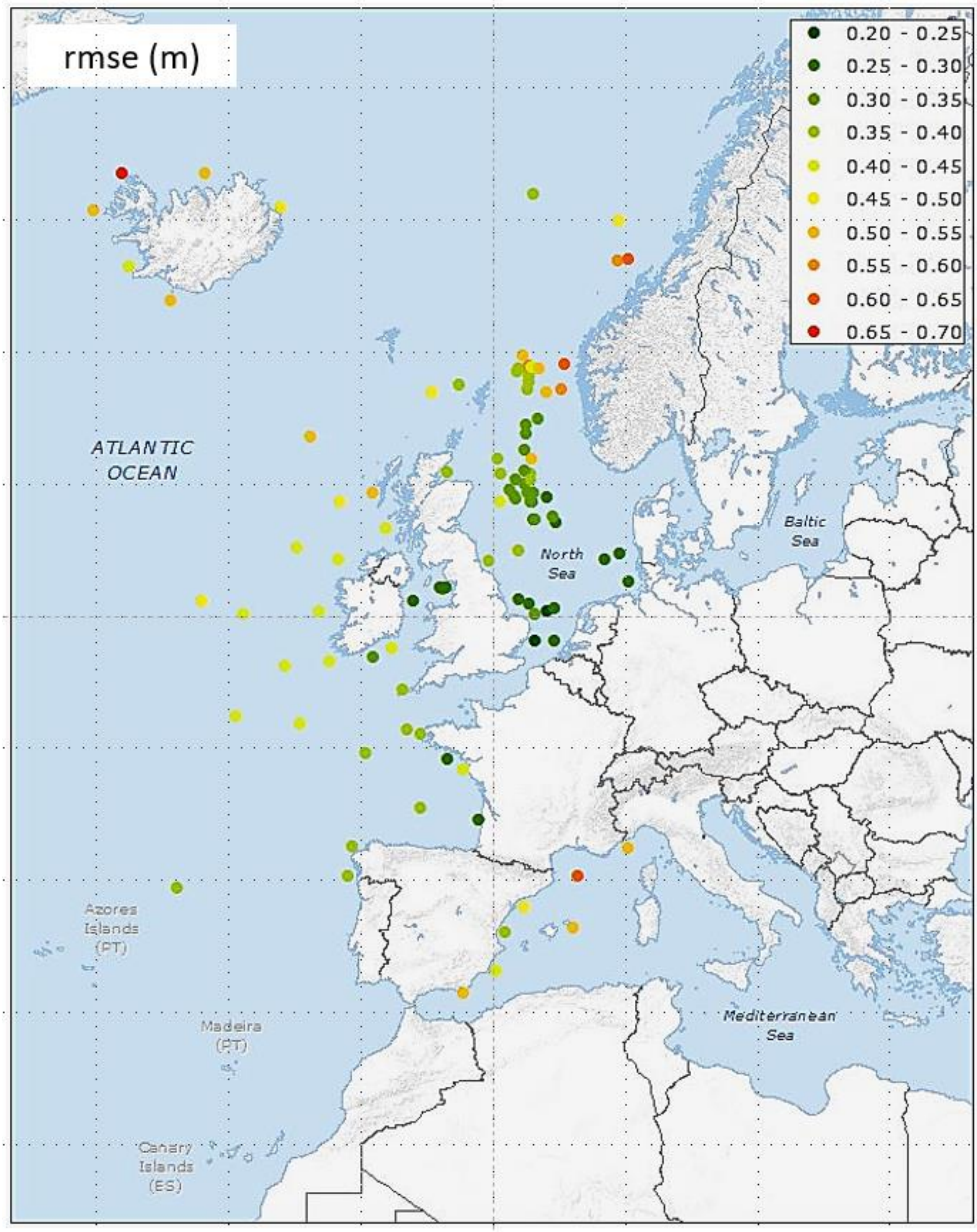


Fig. 2. Bias (upper panel) and rms error (lower panel) values (m) for wave hindcasts during 1996 to 2015.

3.3 Local joint validation for the RIEN of Rhine River (NL)

A joint validation of surge and wave hindcasts utilising a relatively long series of surge and wave observations close to Rockanje (RIEN of Rhine River) was performed for testing the quality of hindcasts during the common period of observation records. For this task, storm surge observations close to Rockanje were downloaded from MATROOS (Deltares Multifunctional Access Tool foR Operational Oceandata Services database - http://noos.deltares.nl/) database. Such surge observations were recorded by the near-by Hook van Holland (HvH) tide gauge recorder positioned at about 15 km northeast of Rockanje (as depicted in Fig. 3). Referring to observation parameters, sea level is the recorded still (i.e. in the absence of waves) water level, and surge is the difference between sea level and predicted tide for that time and location. Similarly, a set of significant wave height observations were retrieved from the close-by wave buoy platform of Lichteiland Goeree I (LiG) moored in North Sea at about 55 km northwest of Rockanje (see details in Fig. 3).

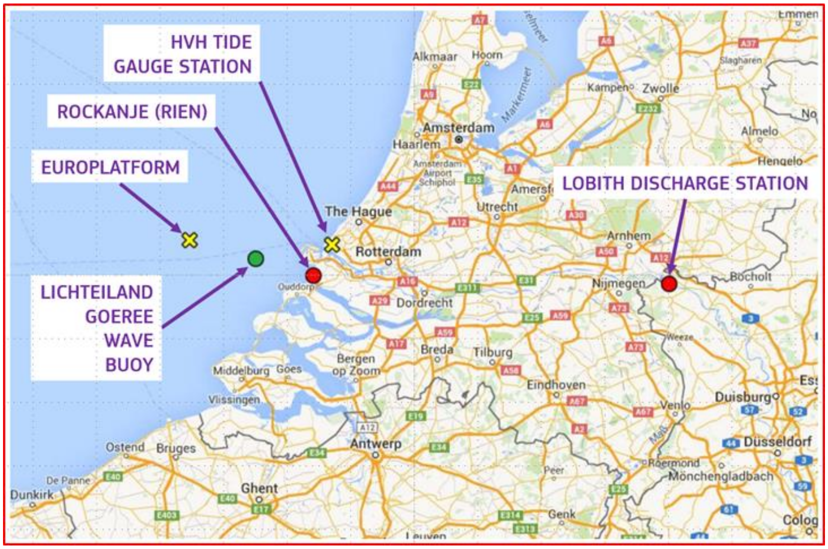


Fig. 3. Position of HvH tide gauge station, Rockanje RIEN and Lichteiland Goeree I (LiG) wave buoy (NL).

A common time interval of 1,114 days (from 22 September 2010 until 9 October 2013) with no gaps was selected for validating surge (SUR) and wave (WAV) hindcast data sets. Such surge and wave hindcasts were different ~~from~~ to the ones referring to Rockanje RIEN point since they were performed as close as possible to the exact positions of HvH tide gauge recorder (SUR) and LiG wave buoy (WAV) for obvious reasons. Both types of observation, i.e., surge, over HvH and wave height, over LiG were made on an hourly basis, so, daily (max24-hour) and half-day (max12-hour) maximum values were calculated. In harmony with observations, 3-hourly based storm surge and ~~1~~3-hourly wave hindcasts were transformed to daily (max24) and half-day (max12) maximum hindcast values. Daily maximum (max24) levels of SUR hindcasts for HvH were compared against daily maximum observations measured by HvH tide gauge over the reference period of 1,114 days. The closest storm surge point used in our analysis was situated ~20 km to the north of HvH (North Sea).

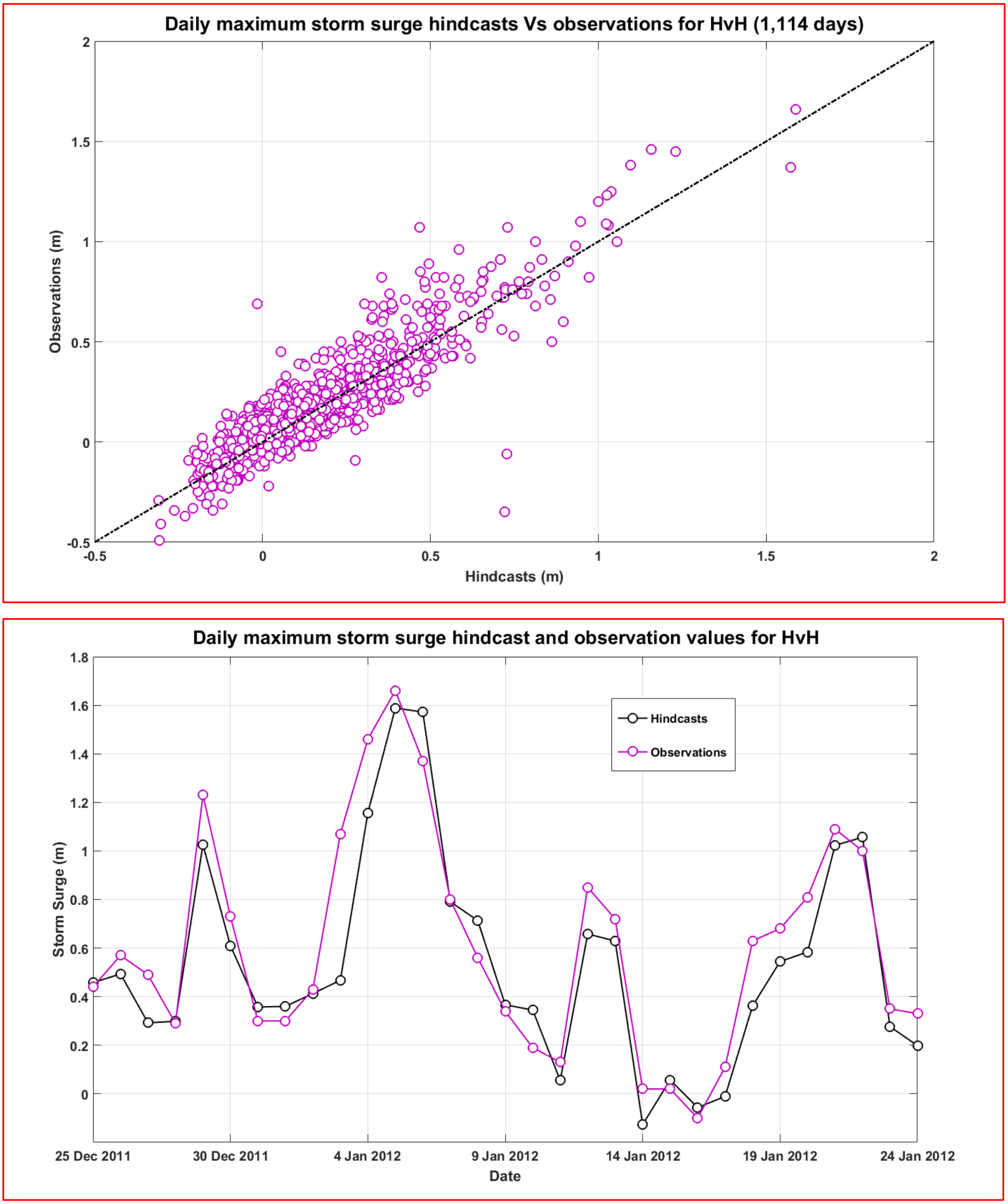
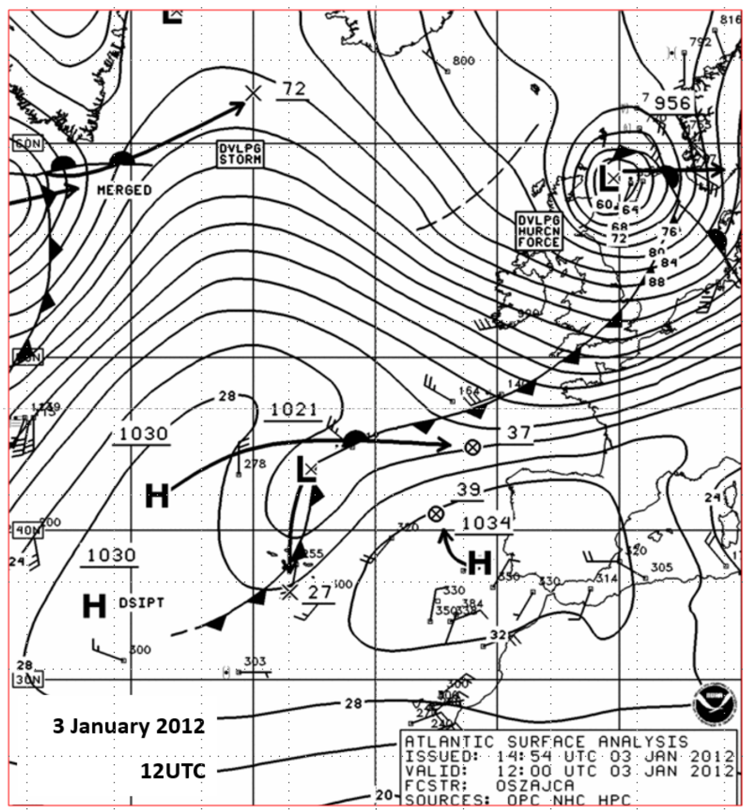
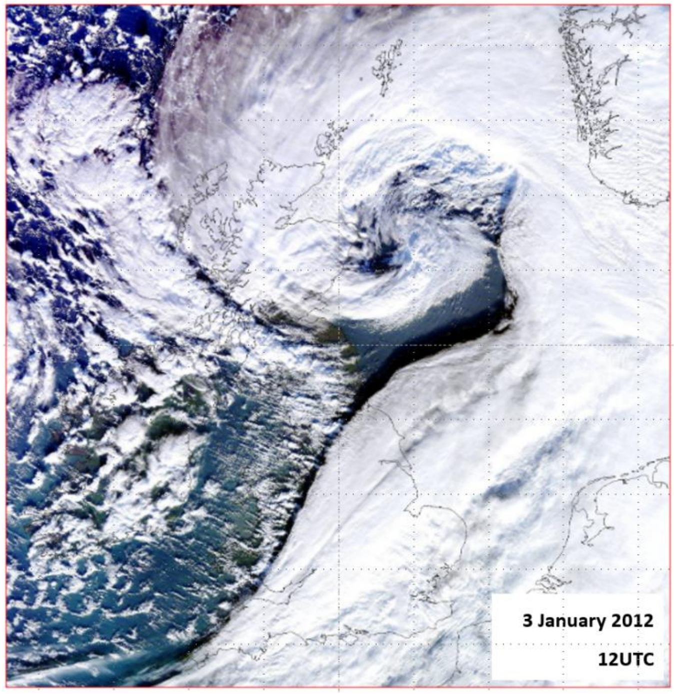


Fig. 4. Scatterplot of surge hindcasts against observations for HvH (upper panel) and a subsection of hindcast and observation values during 2~~6~~5 December 2011 to 2~~9~~4 January 2012 (lower panel).

Fig. 4 (upper panel) contains the scatterplot of surge hindcasts against observations in max24 mode. It appears that SUR hindcasts are in most cases lower than their corresponding observation values. This difference might be attributed to the relatively low temporal and spatial resolution of ERA-Interim forcing terms, although hindcasts overall were found capable of coping well with both the timing and magnitude of extremes.

An example of such an extreme taken place on 5 January 2012 is shown in Fig. 4 (lower panel) that contains a subsection of SUR hindcasts plotted together with surge observations over HvH. This extreme was selected as a multi-purpose demonstrating example (see also Fig. 6 for corresponding maxima of WAV values during the same period and Sect. 4.5 where the de-clustering technique of Peaks-Over-Threshold approach is explained). It is evident that although SUR hindcasts have a tendency to underestimate observations, the model simulations seem able to resolve both the magnitude and the duration of storm events relatively well. The storm surge peak of 5 January 2012 was found to be linked to a very intense extratropical cyclone ~~(Ulli / Emil Storm)~~ (Storm Ulli / Emil) affecting the greater area of the North Sea. The position and details of the storm are contained in the surface weather map of 12UTC of 3 January 2012 shown in Fig. 5 (upper panel). The corresponding (12UTC) satellite picture capturing ~~Emil storm~~ Storm Emil is contained in the same Fig. 5 (central panel).





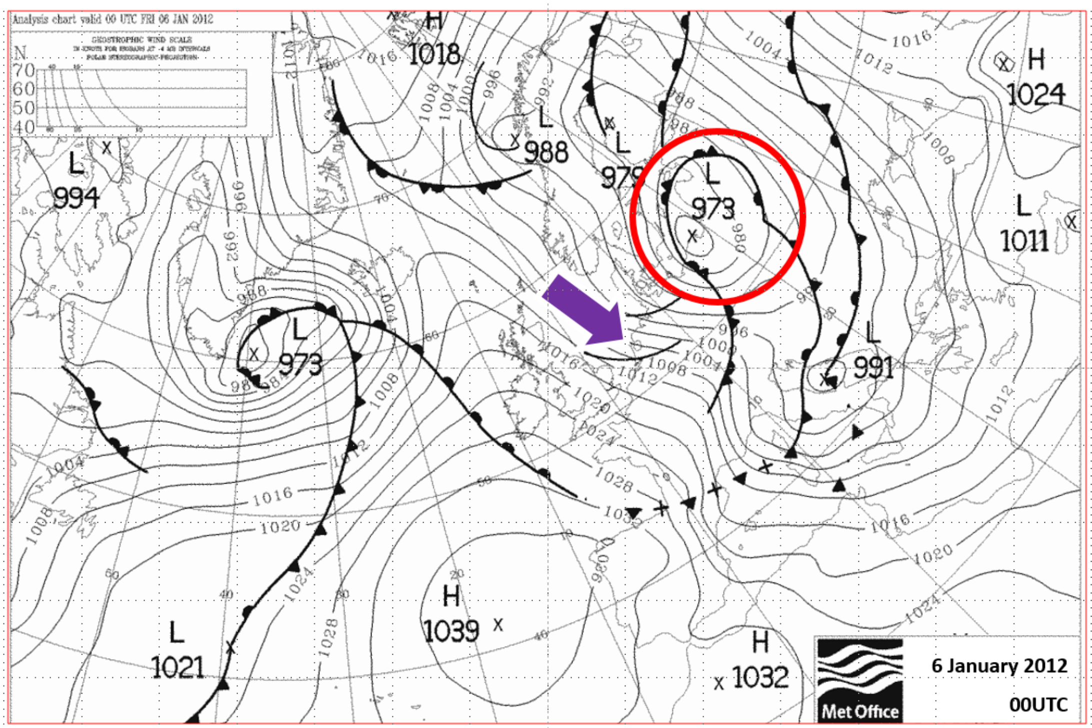


Fig. 5. Surface map (upper panel) and its corresponding satellite image (middle panel) valid for 3 January 2012 12UTC. Surface weather map in the late hours (midnight) of 5 January 2017 (lower panel).

Studying closely both surge hindcast and observation values referring to ~~Emil Storm~~ Storm Emil (lower panel of Fig. 4) it seems that hindcasts could resolve and simulate well both the phase and the magnitude of such an extreme event. It is important to realise that such events are linked to intense pressure gradients such as those clearly seen in Fig. 5 (lower panel) prevailing during the late-night hours of 5 January 2012. Besides the intensity of pressure gradients, the orientation of isobars that was almost vertical to the coasts of Holland (as indicated by a ~~blue~~ purple arrow) it was also contributing strongly to the extremity of the event.

Overall, storm surge hindcasts were found to have a negative bias (defined as difference between the mean of hindcasts and the mean of observations; see details in http://www.cawcr.gov.au/projects/verification/) of about -2.65 cm. It was also found that hindcast and observation values exhibit a very strong correlation reaching a value of ~0.90, while slightly lower correlation was found for max12 (0.88).

Similarly, daily maximum (max24) values of significant wave height (WAV) were compared against daily maximum values of observations measured at the LiG wave buoy platform as shown in Fig. 6 (upper panel). The closest wave model point used in our analysis was suited ~9.5 km northwest of the position of the LiG platform. As in the storm surge case, WAV hindcasts are in most cases lower than their corresponding observation values.

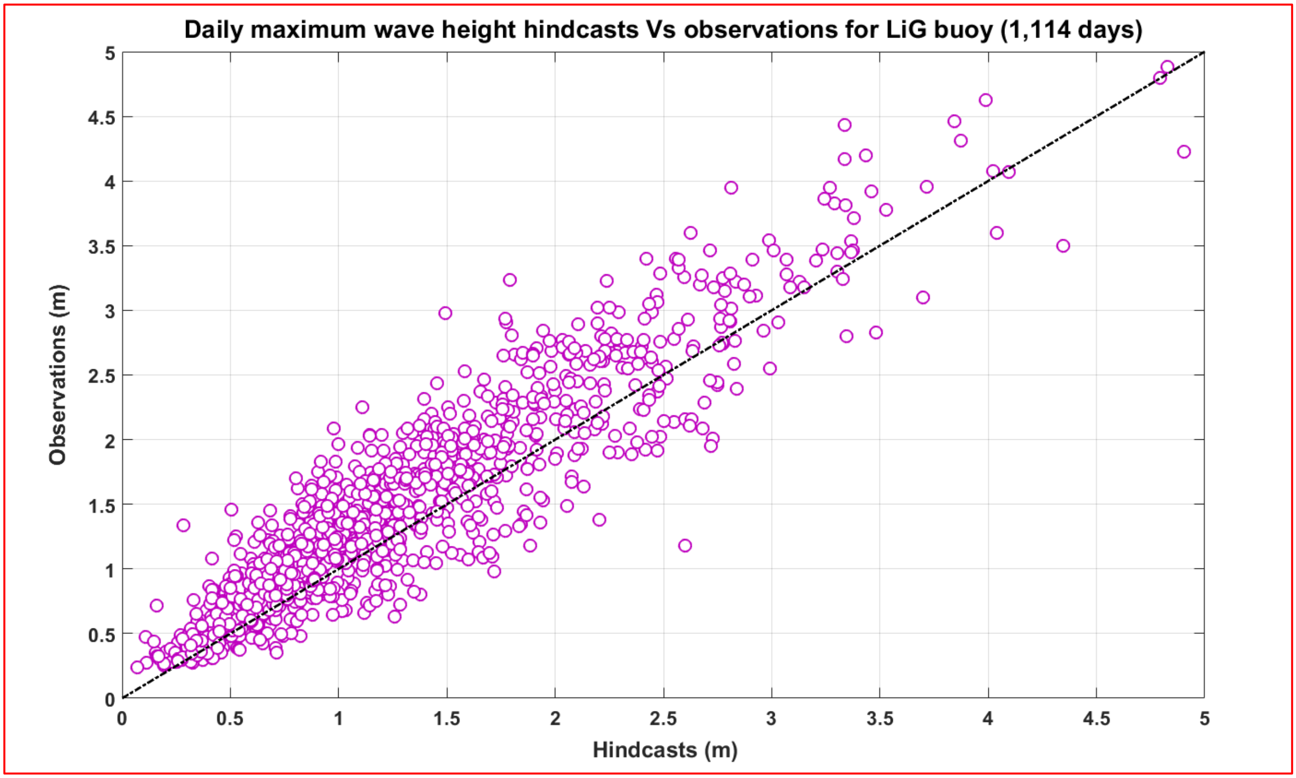
As mentioned already, this might be due to the smoothness of ERAI forcing terms not possessing the required resolution to resolve the exact magnitude of wind components. The systematic bias of wave hindcasts was found to be relatively small, being equal to 20.30 cm. Wave hindcasts were found to exhibit a very strong correlation value (~0.92) to observations, while similar (slightly lower) correlation was found for max12 data pairs (~0.90).

As in the case of SUR hindcasts, Fig. 6 (lower panel) contains a subsection of time series of both WAV hindcasts and observations (with dates being identical to the previous SUR case). WAV hindcasts appear to be lower than observations and besides the unavoidable smoothness of ERAI fields, another explanation for such deviation might be the proximity of the LiG buoy to the coast. It is well known that enclosed areas and near-shore locations are indeed much more difficult to model (http://www.ecmwf.int/en/newsletter/150/meteorology/twenty-one-years-wave-forecast-verification).

Referring to the example (SUR) already presented in the lower panel of Fig. 4, the storm surge extreme of 5 January 2012 was found to be in harmony with the significant wave height extreme observed during the same max24 interval (shown in the lower panel of Fig. 6). This is also an indication that storm surge and wave compound extremes are linked to the same weather system (~~Emil storm~~ Storm Emil) with a clear tendency to take place in a zero-lag time mode over the south coasts of the North Sea.

The obvious agreement between hindcasts and observations (Fig. 4) is a clear indication of the model's (Delft3D-Flow) capability to simulate efficiently observations (over the whole spectrum of observations) in hindcast mode having as input parameters (wind components and mean sea level pressure) from the ECMWF ERA-Interim reanalysis data set. Same wise the obvious agreement between hindcasts and observations (Fig. 6) is a clear indication of the model's (ECMWF / ECWAM) capability to simulate efficiently observations in hindcast mode having as input parameters (wind components) from the ECMWF ERA-Interim reanalysis data set.

Indicative examples of such capabilities can be seen in Table 2 and Table 3 of Section 2 of the Technical Supplement revealing that hindcasts above all were capable of identifying and resolving all seven (7) compound events (based on 98.5% percentile threshold) that took place during the common time interval of 1,114 days over HvH area of interest.



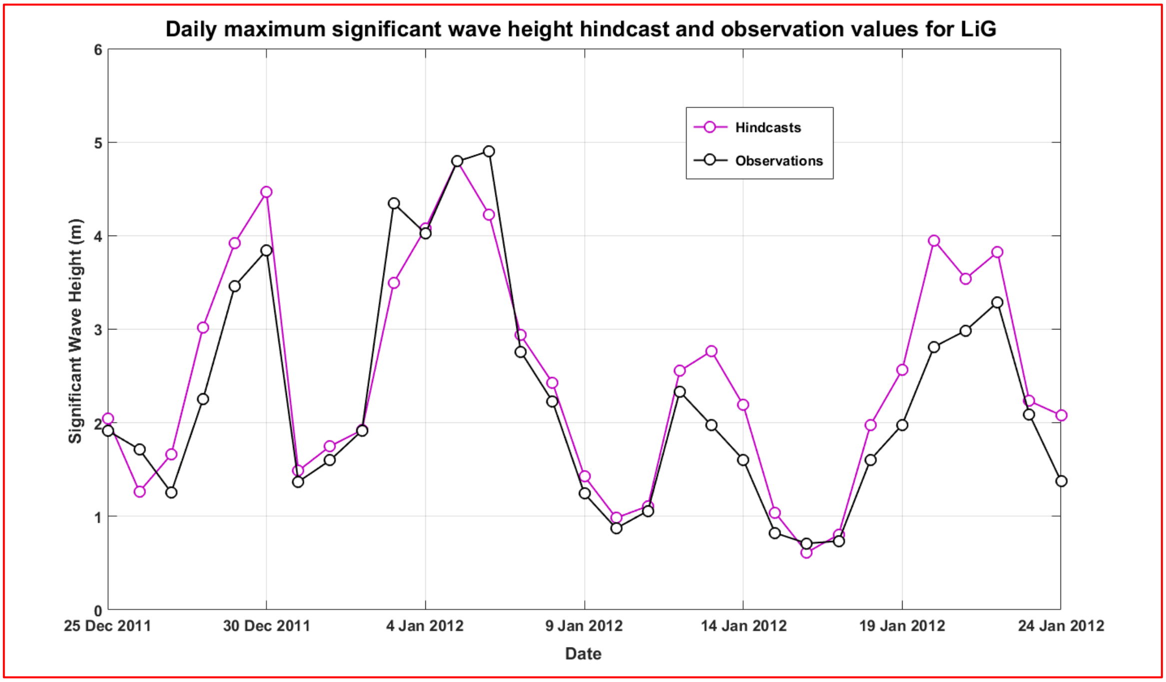


Fig. 6. Scatterplot of wave hindcasts against observations (upper panel) and a subsection of hindcast and observation values during 2~~6~~5 December 2011 to 2~~9~~4 January 2012 (lower panel).

Since this study is focused over maxima taken place over 12- and 24-hours based on 3-hour set of hindcast values, 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. Details and examples of the capability of hindcasts to identify and resolve compound events of surge and waves are contained in Sect. 2 of the Technical Supplement.

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 revealed that both storm surge and their corresponding wave height hindcasts were able to capture (resolve) almost all of the 12- and 24-hour extremes (not necessarily compound ones) on the same (correct) day but with a weaker intensity (i.e., with a correct footprint of lesser intensity).

4 Results

The main tools for estimating statistical dependence (χ) are briefly summarised in the next section (Sect. 4.1). Besides the ability of surge and wave hindcasts to simulate correctly observations over HvH tide gauge and LiG wave buoy platforms respectively (as analysed in Sect. 3.3), their potential for resolving the correct type and strength of both correlation and dependence between primary variables in a joint (compound) mode environment is investigated over a common period of

1,114 days in Sect. 4.2. Referring to the full span of hindcasts, analytical maps and tables have been assembled ~~referring to~~ containing 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). These results are presented in Sect. 4.3 (southern European areas) and Sect. 4.4 (northern European areas). An evaluation of the low-level flow during the top 80 extreme compound events utilising wind rose diagrams is contained in Sect. 4.5. The critical period (of the year) for such high-impact events to take place is also assessed by considering monthly frequencies of occurrence.

4.1 Main tools for estimating statistical dependence

The main tools for assessing dependence between surge and wave has been a set of ~~matlab~~ Matlab routines (mat_chi) for estimating the asymptotic behaviour of statistical dependent variables. Other ~~matlab~~ Matlab routines such as mat_chibar (see details and examples 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~~ 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 (see details in the Statistical Supplement). ~~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~~ while 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.

4.2 Validation of hindcasts in "compound" mode

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. Daily maximum (max24) values of storm surge observations collected at HvH station for the common time period (obs_com) are plotted against corresponding significant wave height observations recorded at the LiG buoy as shown in Fig. 7 (upper panel). Observations of surge and wave seem to be well correlated with a coefficient reaching a value of 0.70. In hindcast mode (hind_com), the exact correlation value (0.70) was found between surge and wave hindcasts during the selected common interval. Surge and wave values in hind_com are plotted in Fig. 7 (lower panel). In both obs_com and hind_com the maximum values of correlations were achieved in zero lag mode (i.e., during the same 24-hour interval).

Table 1. Details and abbreviations of 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)

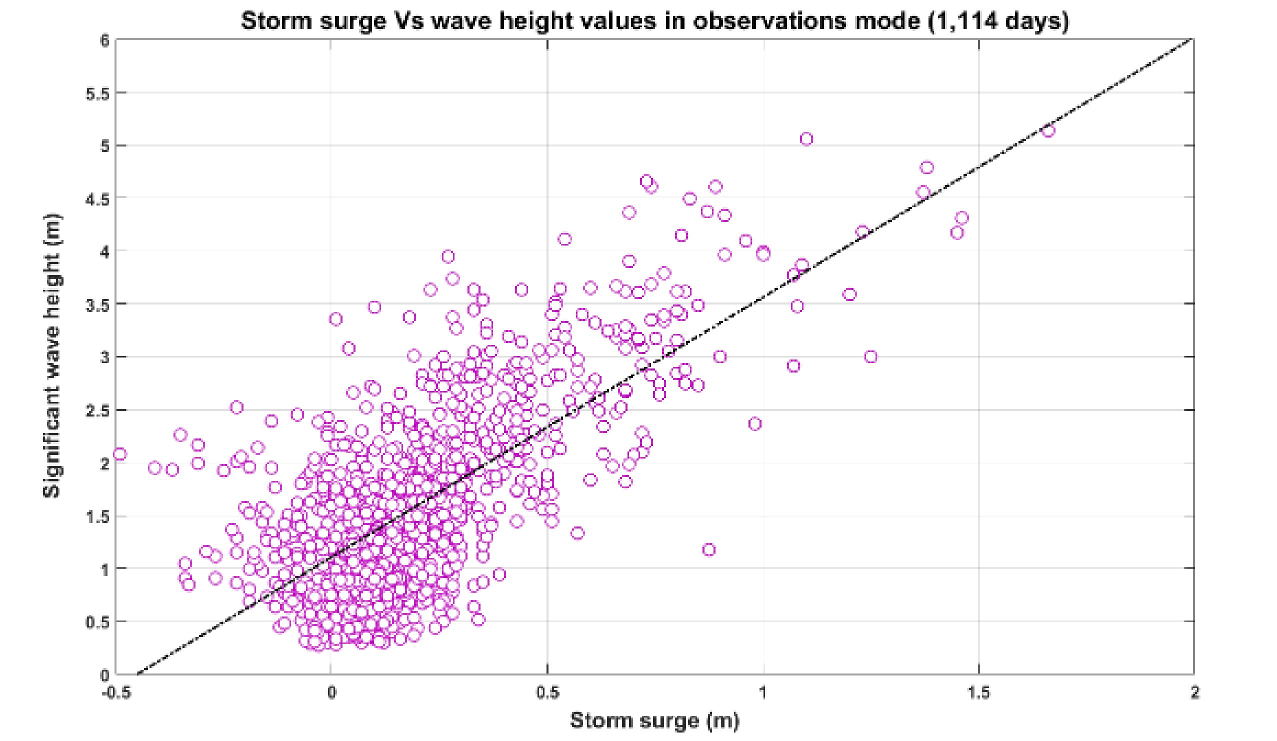
Similar ~~(slightly lower)~~ results were found in max12 case with a (slightly lower) correlation value reaching 0.69 in zero-lag mode. A slightly higher correlation value (0.73) was found in zero-lag mode when the total 12,753 daily data pairs of hindcasts (hind_tot) were used. This deviation should not be considered significant since different number of data pairs was used. Similar (slightly lower) values of correlation were found for the max12 data pairs (0.71). 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.

As in the case of correlations, the capability of hindcasts to resolve correctly the statistical dependence between surge and wave focusing on the upper (extreme) percentiles is investigated by inter-comparing dependencies estimated in obs_com and in hind_com (1,114 days). Fig. 8 shows the full range of χ values for all different types of data pairs considered in this study. The Peak-Over-Threshold (POT) methodology (see details in the Statistical Supplement) was applied for a minimum three-day separation of extremes and an optimal selection of threshold was made not allowing more than ~2.3 events per year to exceed it, ~~(see details in Sect. 2.2 & Sect. 4.1). This setup was found to produce~~ resulting in quite stable χ graphs over a wide range of percentile values as shown in Fig. 8 (for obs_com, hind_com and hind_tot ~~also~~), whereas due to the sparse of data pairs, values of dependence in the area of lower and higher quantiles appear to be quite unstable (i.e., with abrupt fluctuations).

Table 2. Dependencies between surge and wave values for observations (obs_com) and hindcasts (hind_com) in common interval and hindcasts over the total period (hind_tot). POT thresholds are shown in parentheses.

~~*obs_com*~~ **obs_com**	~~*hind_com*~~ **hind_com**	~~*hind_tot*~~ **hind_tot**
0.5850 (98.2%)	0.5840 (98.3%)	0.5629 (98.0%)

Similar (strong) values of χ between surge and wave were found for all three configurations (obs_com, hind_com and hind_tot) in zero-lag mode as shown in Table 2. Similar results (for all three configurations) were also found when the same data pairs were considered for running chiplot of R (as shown in Fig. 1 of the Statistical Supplement). Similar but slightly lower values were found for max12 case (in zero-lag mode).



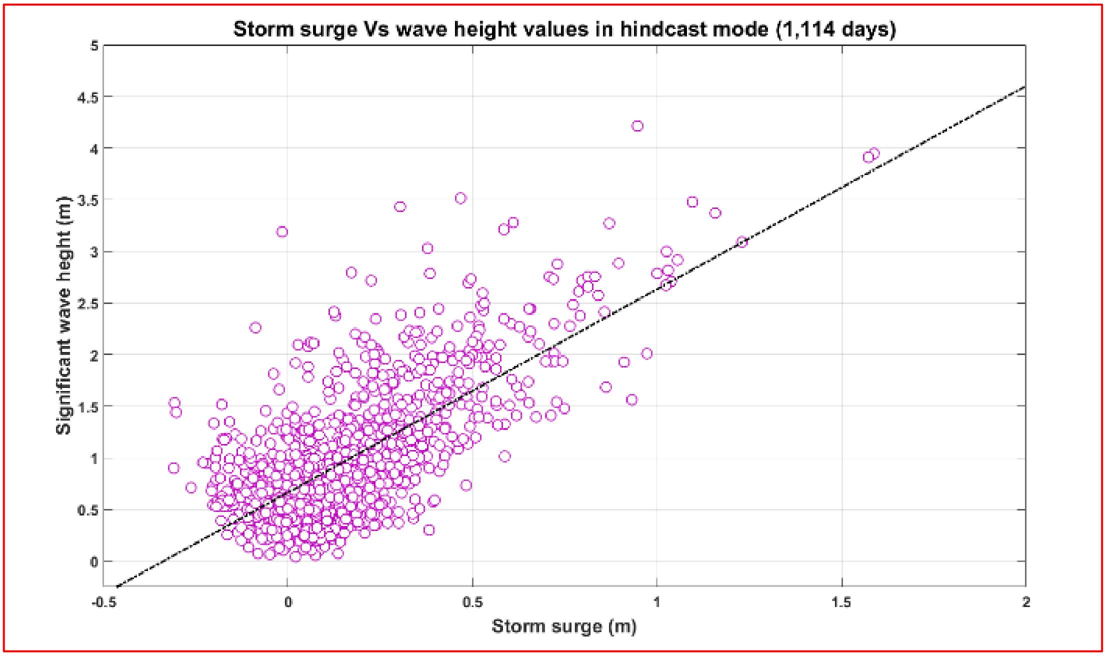


Fig. 7. Scatterplots of storm surge against significant wave height in obs_com (upper panel) and hind_com (lower panel) for the common time interval of 1,114 days.

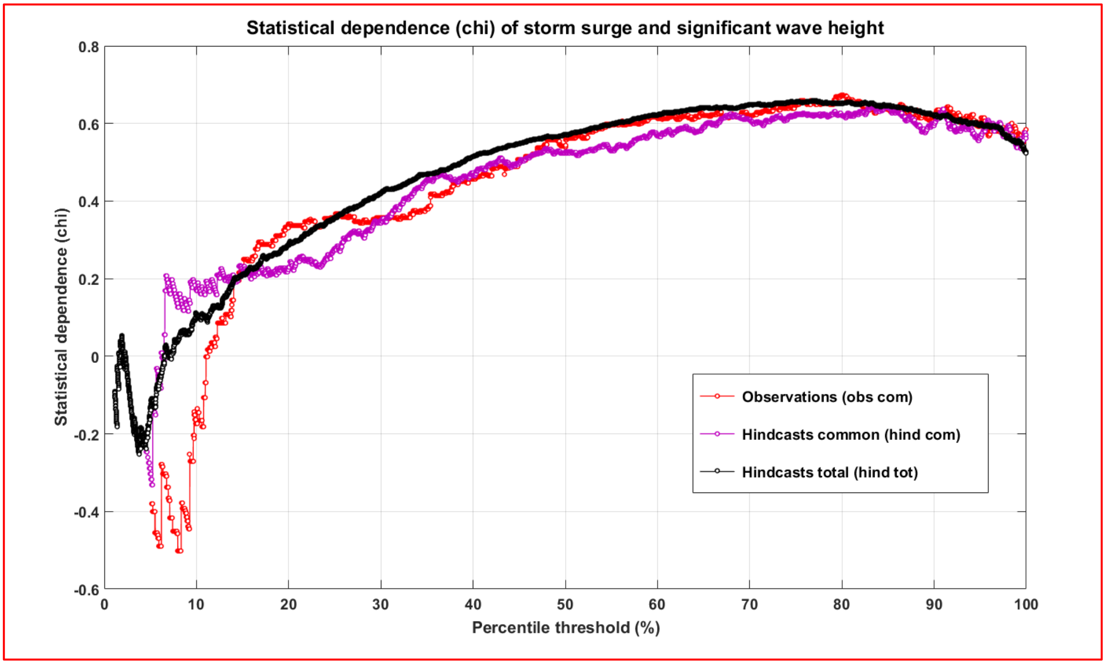
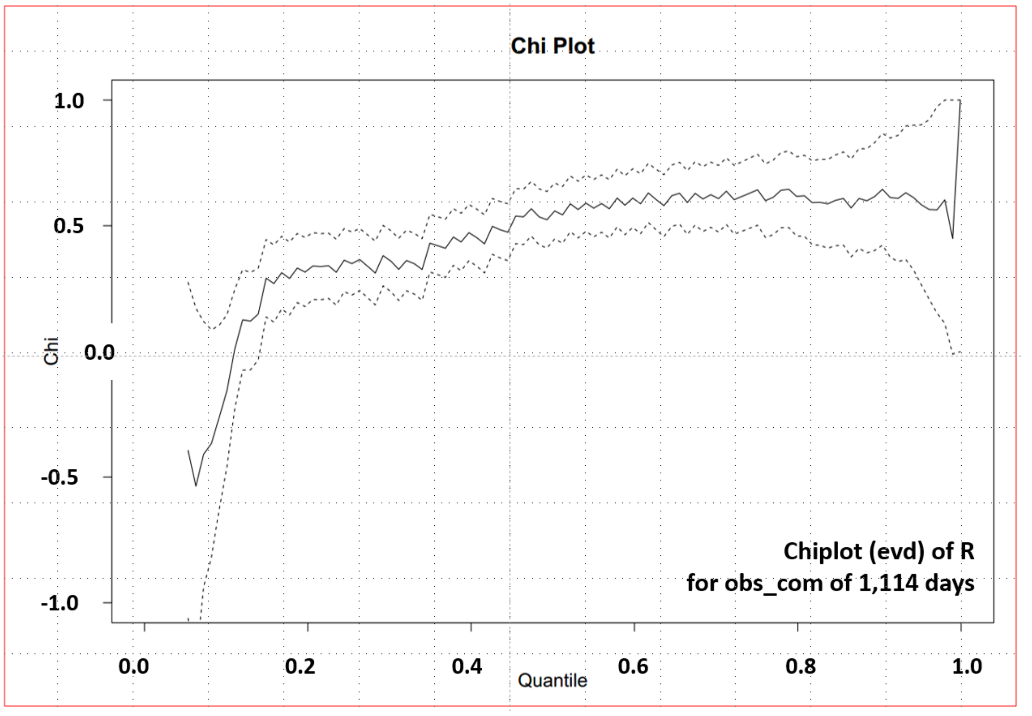
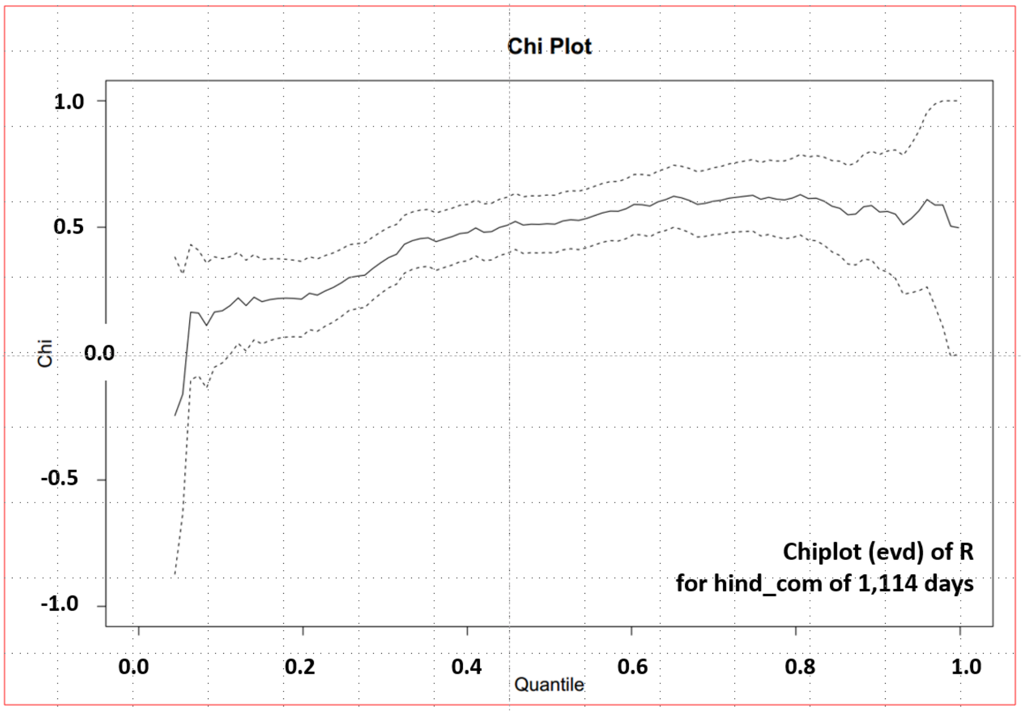


Fig. 8. Statistical dependence (χ) of storm surge (HvH) and significant wave height (LiG) max24 values in common obs_com & hind_com (1,114 days) and ~~in~~ total hind_tot (12,753 days) mode.


Chi Plot
1.0
0.5
0.0
-0.5
-1.0
Chi
0.0
0.2
0.4
0.6
0.8
1.0
Quantile
Chiplot (evd) of R
for obs_com of 1,114 days


Chi Plot
1.0
0.5
0.0
-0.5
-1.0
Chi
0.0
0.2
0.4
0.6
0.8
1.0
Quantile
Chiplot (evd) of R
for hind_com of 1,114 days

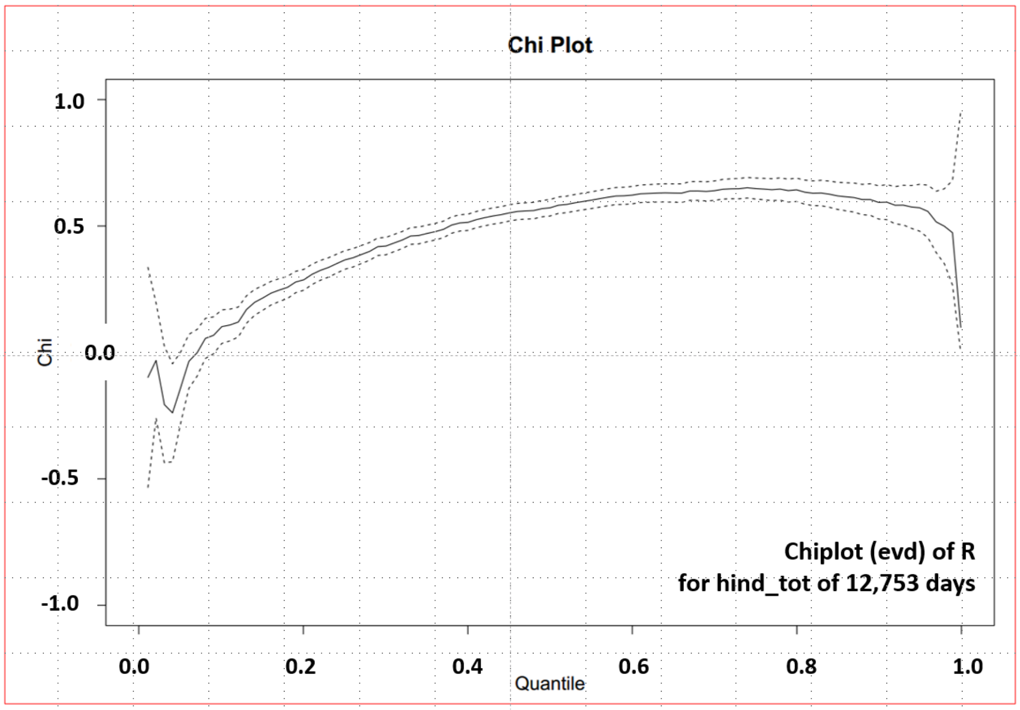


Fig. 9. Estimated χ values between surge (HvH) and wave (LiG) max24 values in obs_com (upper panel) & hind_com (middle panel) and in hind_tot (lower panel) mode by chiplot routine of evd module (R).

It should be pointed out that the real (correct) statistical dependence is estimated by utilising the formula of Eq. 4 of Statistical Supplement 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. It should be also 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.

The importance and implications of such high values of dependence can be demonstrated with an example as the one presented in Sect. 7 of the Statistical Supplement by considering the total hindcast (hind_tot) series for surge (HvH) and wave (LiG). ~~Utilising the 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. This value 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).~~

It should be pointed out that the real (correct) statistical dependence is estimated by utilising the formula of Eq. 4 in the Statistical Supplement 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. It should be also 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.

Overall, considering the complexity of all physical drivers behind such dependencies that are focused intentionally on the upper percentiles where extremes reside, hindcasts of storm and wave seem to perform quite well in their ability to simulate observations and to resolve correctly the type and strength of both correlation and statistical dependence.

4.3 Correlations and dependencies for southern coastal areas

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 ~~3~~ 4 (based on ~~matlab~~ Matlab routines) and Table 4 6 (based mainly on R routines) of the Technical Supplement. 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. The different categories of correlation and dependence used later in the text refers to the categorisation adapted by Defra TR1 Report (2005), shown also in Fig. ~~10~~ 9.

In Table ~~3~~ 4 (Technical Supplement), 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. In max12 mode, low ($0.05 \leq \chi < 0.12$), or even negative correlations were found over most coastal areas with the exception of Adriatic Sea RIENs and the RIEN of Aven River (belonging to a higher category), whereas moderate ($0.12 \leq \chi < 0.38$) values of dependence (max12) were estimated for most of those RIEN points. Such differences do not come as a surprise since dependence is focusing selectively on the upper (extreme) percentiles and not on the full range of data pairs, meaning that surge and wave may have a considerable statistical dependence capable of modulating joint return period even if correlation in some cases is remarkably low (Drouet Mari and Kotz, 2004). Higher correlations ~~(Table 3)~~ were found in max24 mode compared to max12, although low (close to zero) and even negative values were estimated locally over the Balearic Sea, Alboran Sea, North & West Iberian and Black Sea.

Table 3. 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 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 4.** As in Table 3, 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~~

Referring to dependence, with the exception of Foix RIEN point (belonging to the low category), the rest of the "comb" dependencies (last column of Table ~~4~~ 6 of the Technical Supplement) fall into the moderate category ($0.12 \le \chi < 0.38$). Besides dependence (chi), chibar values were estimated. Significance values, lower and upper confidence interval values of χ were calculated also ~~(Table 3)~~. In Table ~~4~~ 6 (Technical Supplement), a set of R values is shown based on chiplot (extRemes module) and taildep (evd module) routines. Relatively small differences were found in estimations of dependence based on ~~matlab~~ Matlab and R routines. Such differences may be attributed to the methodology for selecting critical percentile thresholds and how to identify and confine POT extremes in every case but nevertheless, in almost all cases both ~~matlab~~ Matlab and R routine estimations were found to belong in the same category. In addition, except for Foix RIEN both taildep & chiplot estimations of dependence fall well inside the confidence (95%) intervals estimated by mat_chi routines.

Extensive lag tests were made for both correlation and dependence revealing that the maximum strength of almost any compound surge-wave event tends to take place during the same max24 or max12 period (zero-lag mode). Exceptions were found for Rhone, Ebro and Danube RIENs with one-day (2 half-day interval in max12 case) lag interval revealing that surge values were (slightly) higher correlated / dependent with wave values of the previous day. Further, the agreement of dependence values (among ~~matlab~~ Matlab and R routines) became more pronounced, as the signal (value of dependence) got stronger. The ensemble mean (comb) value of χ ~~(Table 4)~~ is used hereafter for defining the category of dependence (max24) in all relevant text and maps.

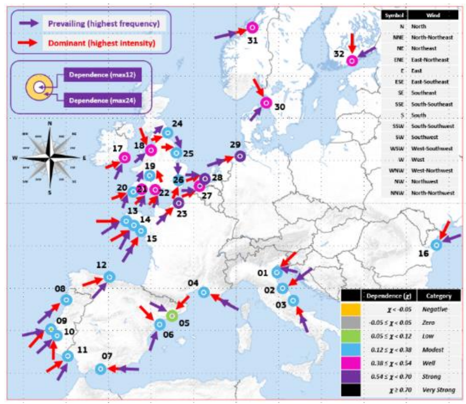
4.4 Correlations and dependencies for northern coastal areas

In Table 5 (Technical Supplement), correlation (corr) and dependence (chi) values for both max12 and max24 intervals are presented together with chibar, critical threshold (thrs), significance (sig) and values of 95% confidence levels (max24). Distinctly higher values were found from those over southern areas for both max12 and max24 cases. All values were achieved in zero lag mode (with the exception of Thames and Goeta RIENs reaching their highest values with one-day lag.

Apart from Thames RIEN (having negative correlation), all max12 correlations fall in the moderate category and above (corr ≥ 0.12) with the top maximum value (0.59) of Bethune (RIEN) belonging to the strong category ($0.54 \leq$ corr < 0.70). Even higher correlation values were found in max24 mode with almost all values falling in the "well" category and above (corr $\geq$ 0.38). Correlations belonging to the "strong" category were estimated for a considerable number of RIENs over Irish Sea, English Channel, North Sea and Baltic Sea.

Contrary to findings over southern areas, smaller differences between correlation and dependence values were found over the northern areas for both max12 and max24 cases. Significantly high values of dependence belonging to the "well" category and above ($\chi \geq 0.38$) between surge and wave were found over the Irish Sea, English Channel, North Sea, Norwegian Sea and Baltic Sea in zero-lag mode (except for Goeta RIEN with one day lag time). Besides Bethune RIEN in the English Channel, having a strong dependence (0.65) in max24 mode, strong dependencies were also found for Rhine (0.54) and Weser (0.55) RIENs. Such findings suggest that over the south coasts of the North Sea, when a surge extreme event is anticipated, probabilities are quite high for an extreme wave event to take place at the same time (as a compound event).

As in Table ~~4~~ 6 (Technical Supplement), a set of dependence values for northern coastal areas based on R routines is shown in Table ~~6~~ 7 (Technical Supplement). Once more, small differences were detected in estimations of statistical dependence between ~~matlab~~ Matlab and R routines but in ~~most~~ almost all cases both ~~matlab~~ Matlab and R routine estimations were found to fall in the same category. In addition, both taildep & chiplot estimations of dependence fall well inside the confidence (95%) intervals estimated by mat_chi routines. As in Sect. 4.3, an ensemble mean value of chi contained in the ~~(~~last column of Table ~~6)~~ 7 (Technical Supplement) is considered as a reference value of dependence (max24).



~~(Old) Fig. 10~~

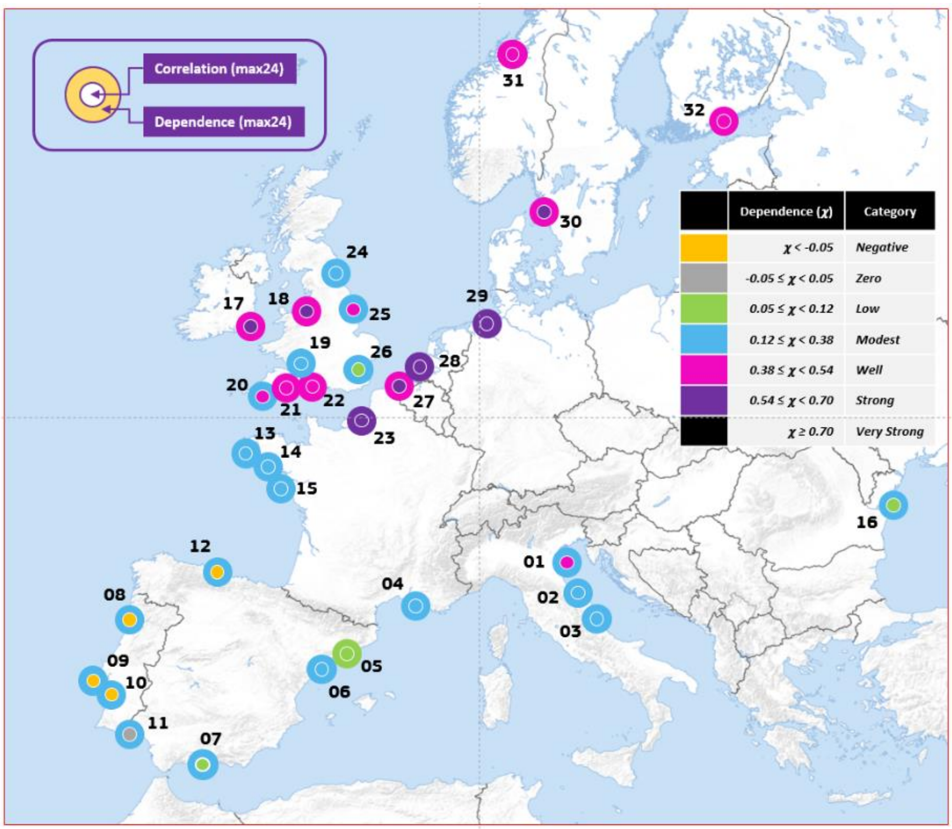


Fig. ~~10~~ 9. ~~Dependence (χ) values for max12 (inner circle) and max24 (outer circle). The prevailing and dominant winds during the Top-80 extreme compound mode are also shown with a blue and red arrow respectively.~~ Correlation (corr) and dependence (chi) values valid for max24 interval.

Table 5. As in Table 3 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

~~**Table 6.** As in Table 5, based mainly on R routines~~

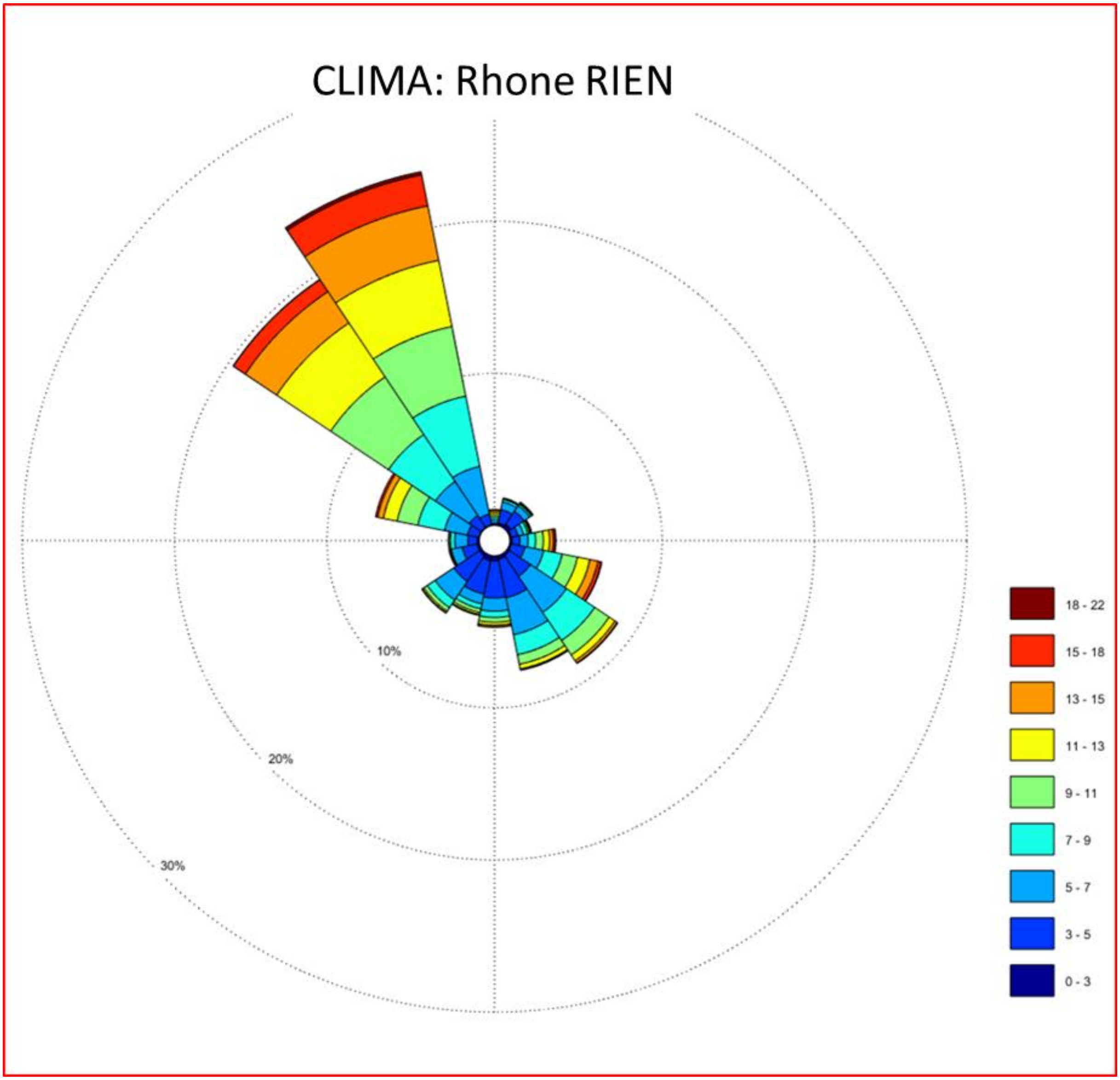
			R				MAT	ENS
	RIEN	sea	lower	upper	chiplot	taildep	mat_chi	comb
17	Owena	IRS	0.26	0.52	0.39	0.40	0.45	0.41
18	Mersey	IRS	0.26	0.48	0.38	0.38	0.43	0.40
19	Severn	BRC	0.16	0.32	0.24	0.30	0.30	0.28
20	Tamar	ENC	0.21	0.41	0.31	0.34	0.35	0.33
21	Exe	ENC	0.25	0.46	0.36	0.38	0.40	0.38
22	Avon	ENC	0.33	0.57	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

Results referring to the ensemble (comb) value of χ and correlation for ~~both max12 and~~ max24 cases over all RIEN points are shown in Fig. ~~10~~ 9. The categorisation applied in this study (shown graphically as an enclosed table in Fig. ~~10~~ 9) is similar to the one introduced by Defra TR1 Report (2005). ~~The prevailing (blue arrows) and dominant (red arrows) winds during the top 80 extreme compound events over each RIEN are also shown in Fig. 10 (see details in Sect. 4.5).~~

Lastly, a full set of lag tests was made for both correlation and dependence. It was found that 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 Thames (UK), Goeta Aelv (SE) RIEN points with one-day lag (2 half-days in the case of max12).

4.5 Wind rose diagrams assessing the low-level flow characteristics during critical compound events

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.



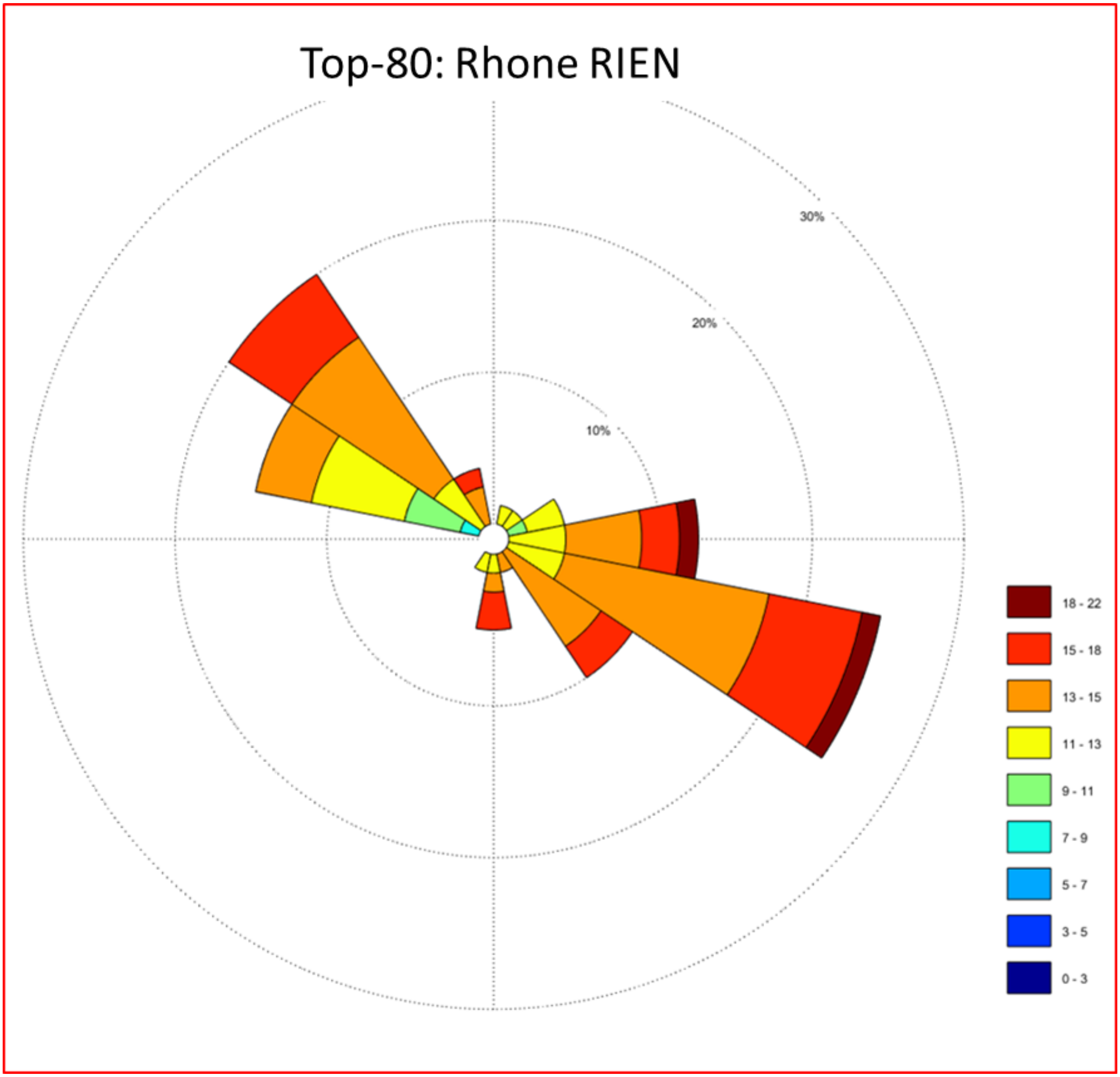


Fig. ~~11~~ 10. Statistical "clima" average (upper panel) and Top-80 extreme compound (lower panel) daily maximum wind roses for Rhone River RIEN.

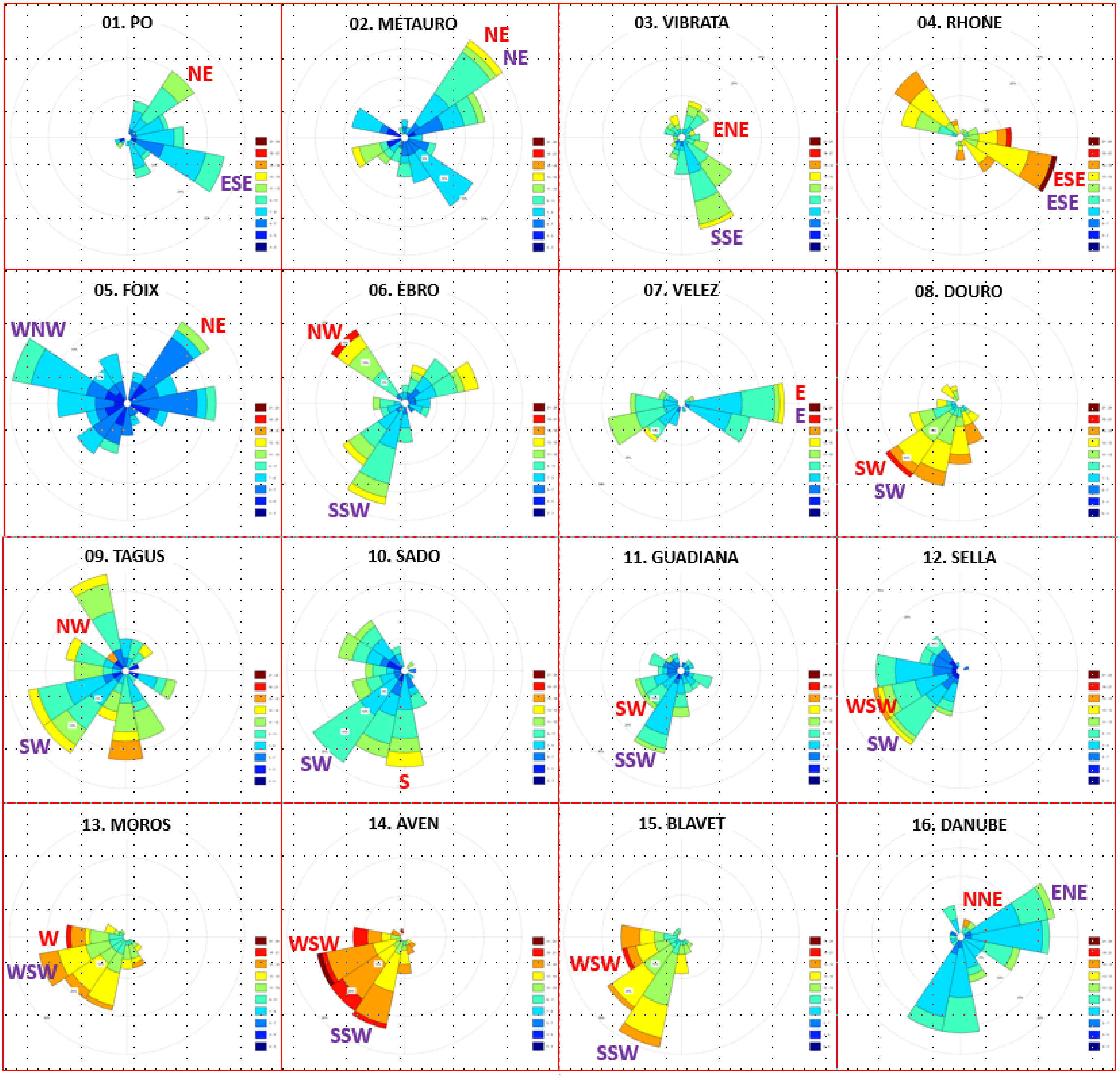
Extreme compound surge and wave events are unavoidably linked to severe weather conditions. These conditions include very strong winds and low atmospheric pressure that is caused mainly by intense storms. Focusing on the low-level circulation, a set of wind rose diagrams was compiled for all RIEN points utilising ERAI reanalysis winds spanning over the total period of 12,753 days.

ERAI winds are referring to the four main synoptic hours (00 – 06 – 12 & 18 UTC) of reanalysis. From such 4-term (daily) sets the maximum speed was estimated and kept together with its corresponding direction to be used in the wind rose diagrams. Wind roses are an information packed plot providing frequencies of wind direction and speed. A wind rose diagram can quickly indicate both the prevailing wind referring to the principal or most common wind direction (having the highest percentage of occurrence) and the dominant wind, indicating the direction of the highest wind speed. Examples of wind roses are given in Fig. ~~11~~ 10 referring to daily maximum winds for the River Rhone RIEN during 12,753 days that may be taken as “clima” conditions (upper panel) and during the Top-80 compound events (lower panel) that may be considered as extreme compound mode conditions.

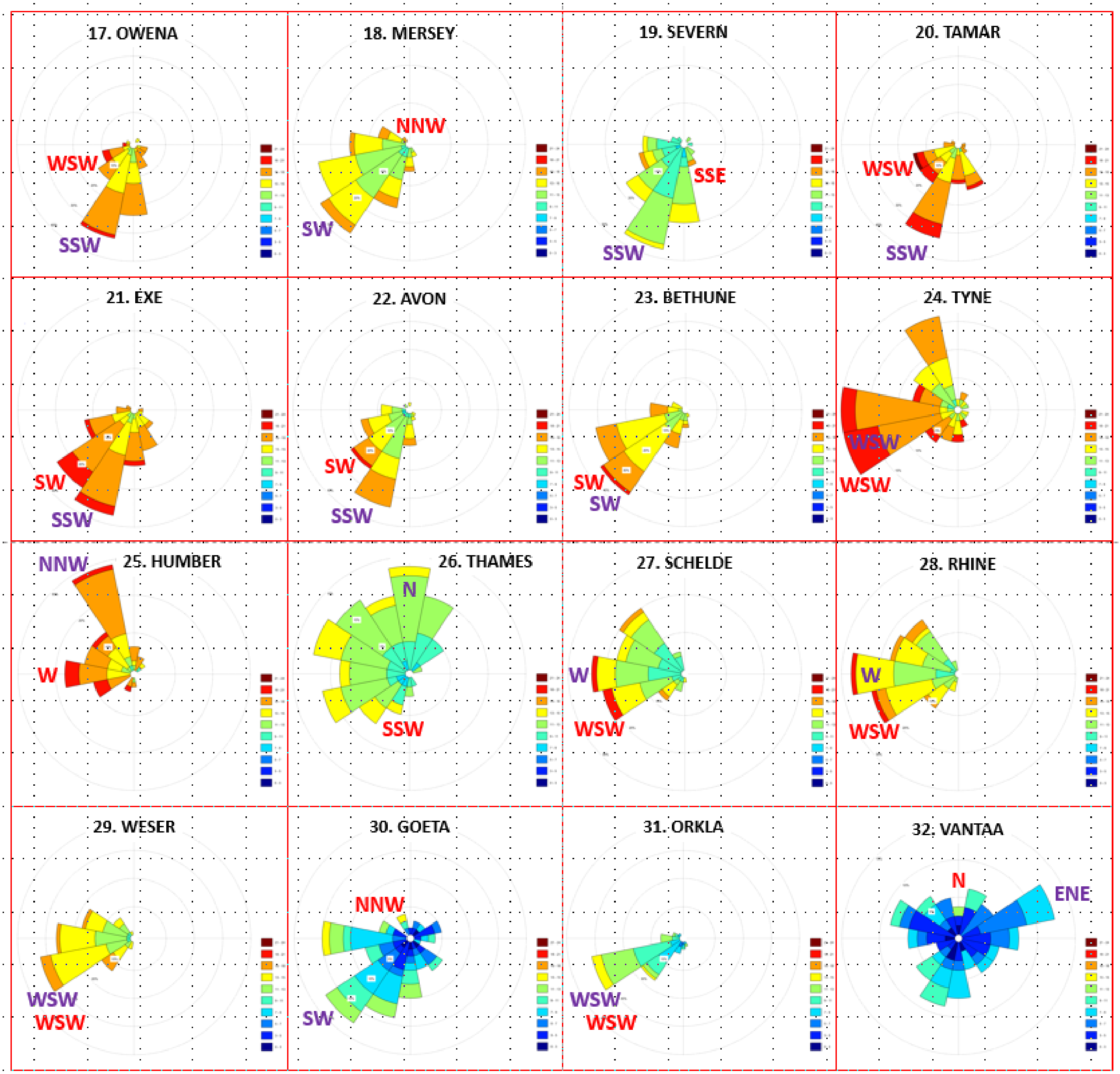
From Fig. ~~11~~ 10 (upper panel) it is obvious that the clima prevailing (highest frequency) wind is north-northwest (NNW), a local type of wind named “Mistral” (www.cs.mcgill.ca/~rwest/wikispeedia/wpcd/wp/w/Wind.htm). The dominant wind (highest intensity) also is of a similar type (Mistral) blowing from a northwest (NW) direction. Mistral is a strong northerly wind blowing over the Gulf of Lion (GoL) and Rhone Valley. The air is usually dry, bringing bright and clear weather with freezing temperatures to the south of France. The Mistral often reaches gale force especially in winter and is capable of raising heavy sea conditions in a short space of time.

The same type of diagram was produced for the Top-80 compound events (lower panel of Fig. ~~11~~ 10) revealing a quite different story. The prevailing wind does not belong to the Mistral “family” since it clearly comprises southeast components of another local wind named “Marin” (www.cs.mcgill.ca/~rwest/wikispeedia/wpcd/wp/w/Wind.htm). Marin is a strong wind in the area of GoL blowing from south-easterly directions, and is next in frequency and importance to the Mistral wind. It is generally warm, moist and cloudy, with rain and heavy weather, and is associated with depressions (storms) that enter the GoL area from the west or south-west after traversing southern France and northern Spain.

The implication of such findings is that although the prevailing and dominant wind in clima mode is of the Mistral type, most of the Top-80 extremes take place under Marin conditions in a relatively stronger wind environment (compared to Mistral conditions). On the other hand, Mistral conditions are also found to be responsible for a considerable percentage of Top-80 events accompanied by winds of lesser intensity (compared to the Marin ones). Similar detailed wind (clima & Top-80) roses were produced for the rest of the RIEN points. ~~Top-80 wind roses are shown in Fig. 12 using a common wind speed scale.~~ Distinct differences between southern ~~(upper panel)~~ and northern ~~(lower panel)~~ coastal areas are once more pronounced revealing relatively stronger intensity flow characteristics over the northern areas.


01. PO
NE
ESE
02. METAURO
NE
NE
03. VIBRATA
ENE
SSE
04. RHONE
ESE
ESE
05. FOIX
WNW
NE
06. EBRO
NW
SSW
07. VELEZ
E
E
08. DOURO
SW
SW
09. TAGUS
NW
SW
10. SADO
SW
S
11. GUADIANA
SW
SSW
12. SELLA
WSW
SW
13. MOROS
W
WSW
14. AVEN
WSW
SSW
15. BLAVET
WSW
SSW
16. DANUBE
NNE
ENE

5



~~**Fig. 12.** Critical flows as captured by Top-80 extreme compound wind roses over southern areas (upper panel) and northern areas (lower panel). Blue colour lettering refers to prevailing whereas red refers to dominant winds.~~

Details of clima and Top-80 flow characteristics are contained in Table ~~7~~ 3. 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 3 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 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 than a day). 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 becoming more perpendicular to the coast.

Table ~~7~~ 3. Prevailing and dominant winds in clima and Top-80 extreme compound mode.

		clima		**Top-80**				**clima**		**Top-80**	
	RIEN	**prev**	**domi**	**prev**	**domi**		**RIEN**	**prev**	**domi**	**prev**	**domi**
1	Po	NE	ENE	ESE	NE	17	Owena	SSW	WSW	SSW	WSW
2	Metauro	NE	NE	NE	NE	18	Mersey	WSW	W	SW	NNW
3	Vibrata	NNE	SSE	SSE	ENE	19	Severn	WSW	SW	SSW	SSE
4	Rhone	NNW	NW	ESE	ESE	20	Tamar	WSW	WSW	SSW	WSW
5	Foix	S	NNW	WNW	NE	21	Exe	SW	WSW	SSW	SW
6	Ebro	NW	NW	SSW	NW	22	Avon	SW	SW	SSW	SW
7	Velez	ESE	E	E	E	23	Bethune	WSW	SW	SW	SW
8	Douro	NNW	NNW	SW	SW	24	Tyne	WSW	SSE	SW	WSW
9	Tagus	NNW	NW	SW	NW	25	Humber	SW	W	NNW	W
10	Sado	NNW	NW	SW	S	26	Thames	SW	W	N	SSW
11	Guadiana	NNW	W	SSW	SW	27	Schelde	WSW	WSW	W	WSW
12	Sella	NE	WSW	SW	WSW	28	Rhine	WSW	WSW	W	WSW
13	Moros	SW	SSW	WSW	W	29	Weser	WSW	WSW	WSW	WSW
14	Aven	SW	WSW	SSW	WSW	30	Goeta	WSW	W	SW	NNW
15	Blavet	SW	WSW	SSW	WSW	31	Orkla	SSE	WSW	WSW	WSW
16	Danube	NNW	NNE	ENE	NNE	32	Vantaa	SW	SSE	ENE	N

Besides wind roses, the critical time period of the Top-80 events was investigated. For instance, in the case of the Rhone River, most Marin (east-southeast flow) and Mistral (north-west flow) Top-80 extreme compound events took place during the cold period of the year. Such a critical period was confined from October to March containing 91% of all Top-80 compound events. Similarly, the critical period of Top-80 events was calculated for the rest of the RIEN points based on monthly frequencies of occurrence (Table ~~8~~ 4). This critical interval comprised mostly cold months. There were even cases such as for the RIEN of Rhine (NL) and Schelde (BE) where all (100%) Top-80 compound events took place during the cold period (September to April). During these critical intervals (Table ~~8~~ 4) there appears to exist a clear tendency for the northern extreme compound events to take place mostly with south-western components of stronger wind intensity (compared to southern events). This tendency for both prevailing and dominant winds to be clustered around the south-western quadrant is more pronounced over the Irish Sea, English Channel, North Sea and Norwegian Sea.

Table ~~8~~ 4. Critical period and percentage of occurrence for Top-80 compound events.

		Top-80				Top-80	
	RIEN	**per**	**%**		**RIEN**	**per**	**%**
1	Po	Oct – Mar	91	17	Owena	Oct – Mar	93
2	Metauro	Oct – Mar	88	18	Mersey	Oct – Mar	96
3	Vibrata	Oct – Mar	91	19	Severn	Sep – Apr	91
4	Rhone	Oct – Mar	91	20	Tamar	Sep – Apr	94
5	Foix	Sep – Apr	94	21	Exe	Sep – Mar	91
6	Ebro	Oct – Apr	88	22	Avon	Oct – Mar	93
7	Velez	Oct – May	98	23	Bethune	Oct – Mar	93
8	Douro	Oct – Apr	88	24	Tyne	Oct – Mar	96
9	Tagus	Oct – Apr	94	25	Humber	Oct – Apr	98
10	Sado	Oct – Apr	97	26	Thames	Oct – Apr	91
11	Guadiana	Oct – Apr	93	27	Schelde	Sep – Apr	100
12	Sella	Sep – Apr	93	28	Rhine	Sep – Apr	100
13	Moros	Sep – Apr	94	29	Weser	Oct – Apr	97
14	Aven	Sep – Apr	91	30	Goeta	Sep – Mar	98
15	Blavet	Sep – Apr	93	31	Orkla	Sep – Mar	95
16	Danube	Nov – Apr	91	32	Vantaa	Sep – Jun	98

The validity of such findings is briefly investigated. For the Irish Sea, extreme surge conditions especially in its eastern side are generated by south-westerly to westerly winds as documented by Brown et al. (2010). For the English Channel, this south-western signature is compatible with the path of (extra-tropical) storms that tend to generate large surges (Henderson and Webber, 1977). For the North Sea such south-western preference seems to partly contradict the fact that the largest wave events occur in the central North Sea when a low-pressure system is situated over southern Scandinavia (such as the one shown in Fig. 5) giving rise to a long northerly fetch associated with strong northerly winds. An obvious explanation could be that southerly-wind events can also create large wave heights despite their limited fetch, since southerly-wind events are associated with the existence of zonal jets (embedded in extratropical cyclones) that intensify rapidly in the left exit region of the jet stream as indicated by Bell et al. (2017). Besides this, depths in the southern North Sea are only about 40m on average adding to the fact that wind stress is particularly effective in piling up water against the coast in the shallow water as the effect is inversely proportional to water depth (Wang et al., 2008).

It should be kept in mind that besides the prevailing and dominant wind directions responsible for most compound extremes there still exist additional critical directions linked to extremes. For instance in the area of German Bight (southern North Sea), northwest wind components (visible in the wind rose for Weser RIEN) have been identified as having a significant link to both surge and wave extremes (Staneva et al., 2016).

Lastly, for the Norwegian Sea, observations seem to fully support our findings as documented in an earlier work of Gjevik and Røed (1976) showing that large storm surges are caused by strong south-westerly winds acting along a large section of the Norwegian coasts.

Overall, the low-level flow characteristics (prevailing and dominant winds) appear to be first in harmony with the transient nature of (extra-tropical) storms and their footprints (storm tracks). This seems to be consistent with similar findings (even if they apply for different pair of variables) in Defra TR1 & TR3 Reports and in Svensson and Jones (2002, 2004a, 2004b, 2005) documenting that (storm) surge and (river) flow dependence appears to be largely influenced by the storm track of the depressions although it should be kept in mind that a thorough understanding of all factors leading to such compound events is above the scope of this study.

5 Discussion and conclusions

The possibility of utilizing statistical dependence methods in coastal flood hazard calculations is investigated, since flood risk is rarely a function of just one source variable but usually two or more. Source variables in most cases are not independent as they may be driven by the same weather event, so their dependence (χ), which is capable of modulating their joint return period, has to be estimated before the calculation of their joint probability. The source variable-pairs presented here, are storm surge and wave height, and their correlation and dependence were assessed over 32 river ending (RIEN) points along European coasts. It should be noted that correlation and dependence may differ substantially from one another. This is because correlation is estimated over the full range of percentiles whereas dependence is focused on the upper (extreme) percentiles.

In the absence of widespread coincident long-term measurements of surge and wave, a set of ~35-year (12,753 days) hindcasts was compiled. Storm surge hindcasts were performed by utilising the hydrodynamic model Delft3D-Flow while wave hindcasts were generated with ECWAM wave (stand-alone) model. Although in some cases extreme surge and wave hindcast levels were underestimated, the overall performance of both surge and wave hindcasts is considered satisfactory. Further, a joint validation in "compound mode" was made over the area of Hook van Holland (HvH) taking into account real measurements of both tides and waves. Overall, hindcasts for the common period of observations (1,114 days) were found capable of resolving and estimating both the correct type and strength of correlation and dependence between source variables.

A necessary split of results revealed distinct differences between southern and northern coastal European areas since significantly higher values of correlation and dependence were found over northern areas with compound events taking place on the same max12 (during half a day) or max24 (daily) interval in a zero-lag mode. Results are presented by means of analytical tables and maps for each RIEN point and can be used to calculate the joint return period by inserting the value of dependence (χ) in a simple formula (Eq. ~~11~~ 12 of the Statistical Supplement) containing the individual return periods of source variables as documented in Hawkes (2004), Meadowcroft et al. (2004), White (2007) and Australian Rainfall & Runoff Project 18 (2009), Petroliagkis et al. (2016). Some limitations of Eq. ~~11~~ 12 (Statistical Supplement) could be overcome if a more complete formula is used such as Eq. 2.15 for instance taken from White's thesis (2017) but this is beyond the scope of the current study.

Overall, significant correlations and dependencies ranging from "well" and above (≥ 0.38) categories were found over the Irish Sea, English Channel, south coasts of the North Sea, Norwegian Sea and Baltic Sea in a zero-lag mode. Over these areas, dependencies reaching locally up to 0.65 (Bethune RIEN) stress the fact that when the first variable (surge) has an extreme value there exists a high probability that the other one (wave) will also produce an extreme level. For the rest of the RIENs mostly positive moderate ($0.12 \leq \chi < 0.38$) dependence values were estimated although a considerable number of them had correlations that were almost zero or even negative. This does not come as a surprise since even in cases of very low correlation there may exist a considerable amount of tail dependence (Drouet Mari and Kotz, 2004).

An effort for inter-comparing our results with previous studies was made although there were very few relevant journal papers (to our knowledge) focusing on correlations and dependencies over such a wide range of coastal areas. A relevant study (thesis) by Kergadallan (2016) for the coasts of France has documented that the surge wave dependence is medium along the Mediterranean coasts whereas dependence values are more important along the English Channel and the Atlantic coasts, which seems consistent with our findings.

Other relevant references pointed to a series of U.K. Defra / Environmental Agency Institute Reports (2003 [Defra TR0], 2005 [Defra TR1], 2005 [Defra TR2] & 2005 [Defra TR3]), hereafter referenced as TRx Reports. This set of Reports (TRx) though, refer to a different measure of dependence constituting a “special” correlation coefficient ρ, above a chosen threshold (90%). Over U.K. coasts, such values of ρ were positive as our set of χ values but considerable higher. Such differences could be attributed partly to the fact that χ values were estimated by considering a quite different (POT) threshold from the one (90%) used in ρ estimations. It could be also attributed to the different nature (methodology of estimation) between χ and ρ, since it is clearly mentioned in TR1 Report that different statistical models are underlying χ and ρ values that could cause considerable distortion when converting from one parameter to the other.

Above all, it appears that such values of ρ (coming from TRx Reports) should not be considered as reliable statistical dependence (χ) values as they point to overestimated levels. In support of this, we refer to the methodology of estimating statistical dependence $\chi(u)$ by Coles (2001) utilising a set of reference data for surge and wave over the Port of Newlyn (Cornwell, U.K.). Results taken from Figure 8.11 (Coles, 2001) suggest a dependence value ~0.35 as $\chi(u)$ clearly tends to this value for the upper percentiles. This is very close to our estimation for the RIEN of Tamar River (0.34) and significantly different from the value found in Table 4.4 of TR1 Report suggesting a value of ρ higher than 0.60.

A further investigation into the low-level flow characteristics of extreme compound events was made. First, a set of 10-metre wind roses was compiled utilising ERAI wind terms over the total period of 12,753 days. These winds are referring to the four main synoptic hours (00 – 06 – 12 & 18 UTC) based on which the daily maximum speed and its corresponding direction were defined and used for producing a set of “clima” wind roses for all RIENs. Based on such clima wind roses the estimation of the prevailing (highest frequency) and dominant (highest intensity) winds was possible. In addition, the 80 most extreme (Top-80) compound events were defined by applying POT (Peaks-Over-Threshold) methodology and allowing a maximum number (~2.3) of compound events on annual basis. A set of wind roses in such extreme mode was assembled revealing distinct differences between clima and Top-80 events in many cases (Table ~~7~~ 3). For instance, in the case of the River Rhone (RIEN), the clima prevailing average conditions were of Mistral (north-western) type whereas the top extremes (Top-80) were mostly of Marin (south-eastern) type conditions.

The critical period of Top-80 events was also estimated based on monthly frequency values of occurrence. This critical interval comprised mostly cold months (Table ~~8~~ 4). There were even cases such as for Rhine RIEN (NL) and Schelde RIEN (BE) where all (100%) Top-80 compound events took place during the cold period (September to April).

Detailed wind roses (Top-80 mode) were produced for the rest of RIEN points as shown in Fig. 12 using a common wind speed scale. It seems that there is a clear tendency for the northern extreme compound events to take place mostly with south-western components of stronger wind intensity (compared to the southern ones) with emphasis during the cold months. This appears mainly to be in harmony with the transient nature of winter storms and their storm tracks as already indicated in Svensson and Jones (2004a & 2004b) in a similar analysis for surge and discharge compound events around Britain.

Besides the relevant link between transient storm systems and compound events, the morphological and topographical characteristics of RIEN areas appears to play a significant role in the genesis and evolution of such extremes. For instance, in addition to the local circulation systems such as the Mistral and Marin winds in the case of Rhone RIEN, a similar pattern was seen with the Bora (north-eastern) and Sirocco (south-eastern) winds providing the main dominant and prevailing (respectively) flows during the Top-80 compound events over Po RIEN (North Adriatic Sea).

Further, not all prevailing and dominant directions contained in Table 3 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 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. 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. 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 becoming more perpendicular to the coast.

This work has been a first step of studying and investigating joint probabilities and return periods of compound events in a relatively low-resolution environment. Having this in mind, results referring to dependence estimations should be considered valid for coastal areas up to a certain distance (a few kilometres) away from the shoreline. Nevertheless, maps and tables can be used to get a valuable indication of the possibility for a combined (compound) hazard based on how the source variables are related (though statistical dependence) over various coastal areas of Europe.

A thorough estimation of the design conditions at the coastal zone would require including more primary and proxy variables in a higher resolution environment. For instance, in addition to the significant wave height, the maximum wave height or/and the period or/and the direction of waves should be also considered. Another important point here is the effect of seasonal circulation and water-mass distribution (currents & tides) besides the prevailing weather system and atmospheric circulation contained in relevant weather maps.

Acknowledgements

Dr Jean-Raymond Bidlot (ECMWF) is to be gratefully thanked for providing ~~us~~ the set of wave hindcasts and in situ wave observation data besides valuable guidance and suggestions. Evangelos Voukouvalas (JRC) is to be thanked for providing us the set of storm surge hindcasts. A long list of JRC colleagues should be also thanked for their invaluable help and support during ~~the writing of this study~~ the Exploratory Research Project Coastal-Alert-Risk (CoastAlRisk) of the JRC (Joint Research Center). The CoastAlRisk Project (2015-2016) had been an initial effort of developing the first global integrated coastal flood risk management system with emphasis on such compound events, by linking satellite monitoring, coupled wave, tide and surge forecasting, inundation modelling and impact analysis.

References

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

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.

Bell, R.J., Gray, S.L. and Jones, O.P.: North Atlantic storm driving of extreme wave heights in the North Sea, J. Geophys. Res. Oceans, 122, doi:10.1002/2016JC012501, 2017.

Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), Hydrol. Earth Syst. Sci., 21, 2701-2723, https://doi.org/10.5194/hess-21-2701-2017, 2017.

Bidlot, J.-R., Janssen, P. and Abdalla, S.: Impact of the revised formulation for ocean wave dissipation on ecmwf operational wave model. ECMWF Techn. Memo. No. 509, ECMWF, Reading, U.K., 27 pp, 2006.

Bidlot, J.-R.: Present status of wave forecasting at ECMWF. Proceeding from the ECMWF Workshop on Ocean Waves, 25-27 June, 2012.

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

Brown, J.M., Souza, A.J. and Wolf J.: An investigation of recent decadal-scale storm events in the eastern Irish Sea, J. Geophys. Res., 115, C05018, doi:10.1029/2009JC005662, 2010.

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.

Currie, J.E.: ‘‘Directory of coefficients of tail dependence,’’ Department of Mathematics and Statistics Technical Report, ST-99-06, Lancaster University, 1999.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553– 597, doi:10.1002/qj.828, 2011.

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.

Deltares Delft3D-FLOW: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. User Manual. Deltares, Delft, The Netherlands, 2014.

Drouet Mari, D. and Kotz, S.: Correlation and Dependence. Imperial College Press, London, 2004.

ECMWF: IFS DOCUMENTATION – Cy41R1, Operational implementation 12 May 2015, PART VII: Wave Model, ECMWF IFS documentation, 2015.

Gjevik, B. and Røed, L.P.: Storm surges along the western coast of Norway. Tellus 28, 2, 166–182, 1976.

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. and Tawn, J.A.: Joint probability of waves and water levels: JOIN-SEA: A rigorous but practical new approach. Internal Document No. SR 537, HR Wallingford with Lancaster University, UK. (Originally dated Nov. 1998, re–issued with minor adjustments in final form May 2000, http://eprints.hrwallingford.co.uk/701/1/SR537-JOINSEA-Probablility-waves-water-HRWallingford.pdf), 2000.

Hawkes, P.J., Svensson, C. and Surendran S.: The joint probability of pairs of variables relevant to flood risk: Dependence mapping and best practice. Defra Flood and Coastal Management Conference, University of York (http://eprints.hrwallingford.co.uk/188/1/HRPP345_The_joint_probability_of_pairs_of_variables_relevant_to_flood_risk_dependence_mapping_and_best_practice.pdf), 2005.

Henderson, G. and Webber, N.B.: Storm surges in the UK south coast. Dock and Harbour Authority 57(678): 21–22, 1977.

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.

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.

Joe, H.: Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.

Kergadallan, X.: Estimation of extreme marine sea levels with and without wave component along the French coasts. Mécanique des fluides [physics.class-ph]. Université Paris-Est, Français, 2015.

Klerk, W.J., Winsemius, H.C., Verseveld, W.J. van Bakker, A.M.R. and Diermanse, F.L.M.: The co-incidence of storm surges and extreme discharges within the Rhine–Meuse Delta, Environ. Res. Lett., 10, 035005, 2015.

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.

Merz, B., Elmer, F. and Thieken, A.H.: Significance of “high probability/low damage” versus “low probability/high damage” flood events, Nat. Hazards Earth Syst. Sci., 9, 1033–1046, http://www.nat-hazards-earth-syst-sci.net/9/1033/2009/, 2009.

Nelsen, R.B.: An Introduction to Copulas, Springer-Verlag, New York, 1998.

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.

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

Reed, D.W.: Flood Estimation Handbook, Vol. 1: Overview. Institute of Hydrology, Wallingford, UK, 1999.

Sembiring, L., van Ormondt, M., van Dongeren, A. and Roelvink, D.: A validation of an operational wave and surge prediction system for the Dutch coast. Nat. Hazards Earth Syst. Sci. 15:1231-1242, 2015.

Staneva, J., Wahle, K., Koch, W., Behrens, A., Fenoglio-Marc, L. and Stanev, E.V.: Coastal flooding: impact of waves on storm surge during extremes – a case study for the German Bight, Nat. Hazards Earth Syst. Sci., 16, 2373-2389, doi:10.5194/nhess-16-2373-2016, 2016.

Svensson, C. and Jones, D.A.: Dependence between extreme sea surge, river flow and precipitation in eastern Britain. Int. J. Climatol., 22 (10), 1149–1168, 2002.

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.

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.

Svensson, C. and Jones, D.A.: Climate change impacts on the dependence between sea surge, precipitation and river flow around Britain. In Proceedings from the Defra 40th Flood & Coastal Management Conference, York, UK, July 2005, 6A.3.1–6A.3.9, 2005.

Sverdrup, H.U. and Munk, W.H.: Empirical and theoretical relations between wind, sea, and swell: Trans. Amer. Geophys. Union, vol. 27, pp. 823-827, 1946.

Thomas, D.G.: Dictionary of physical geography. Blackwell, 2000.

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.

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E.: Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nature Climate Change, 5(12), 1093-1097, 2015.

Wang, S., McGrath, R., Hanafin, J., Lynch, P., Semmler, T., Nolan, P.: The impact of climate change on storm surges over Irish waters. Ocean Modelling, 25(1–2), 83-94, 2008.

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.

Zheng, F., Westra, S. and Sisson, S.A.: Quantifying the dependence between extreme rainfall and storm surge in the coastal zone J. Hydrol. 505 172–87, 2013.

Zheng, F., Westra, S., Leonard, M. and Sisson, S.A.: Modelling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. Water Resources Research, 50(3), 2050–2071, 2014.

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 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 (Beersma and Buishand, 2004), 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. ~~In reality,~~ Further, as variables reach their extreme values, ~~a~~ special methodolog~~y~~ies of estimating statistical dependence could be utilised ~~This methodology~~ as the one ~~has been~~ documented ~~by~~ in Buishand (1984) ~~and Coles et al. (2000)~~.

A brief description of ~~the~~ this methodology based on Coles et al. (2000) is ~~given below~~ described below (Sect. 2) while the basic theory behind the utilisation of an optimal copula function refers to Nelsen (1998), Joe (1997), Currie (1999), Wahl et al. (2015).

2 Estimation of statistical 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:

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

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) = 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:

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

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 χ(u) for a given threshold u can be given by:

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

Taken into account the upper limit of the observations (previously defined as z^* in Eq. 1), the dependence measure χ(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):

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

and

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

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:

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

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

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

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:

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

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:

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

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:

$$\alpha = \exp\left(-(i - 0.5) / N\right) \qquad \text{~~(10)~~ (11)}$$

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. 12 following White (2007), Australian Rainfall & Runoff Project 18 (2009).

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

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). Further, ~~S~~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.

It should be kept in mind that the significance level of 5% represents the probability of rejecting the null hypothesis when it is true. In simple words, it indicates a 5% risk of concluding that a difference exists capable of rejecting the null hypothesis (the population mean equals to the hypothesized mean) when there is no actual difference.

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.

To draw the distinction between significance (previous Sect. 4) and confidence levels it should be noted that a confidence interval is a range of values that is likely to contain an unknown population parameter (in our case the statistical dependence) whereas the significance represents the probability of rejecting the null hypothesis when it is true. It follows that if a random sample is drawn many times, a certain percentage of the confidence intervals will contain the population mean. That is the reason behind the usage of confidence intervals for bounding the mean or standard deviation.

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

Extreme value analysis can be carried out using two types of data series (Bezak et al., 2014), annual maximums (MA) or flows above a certain threshold (POT) for Peak Over Threshold. The POT model used in this study can be composed of the Poisson, binomial and negative binomial distributions for modelling the annual number of events above threshold, and of exponential or generalized Pareto distributions for magnitudes of exceedances.

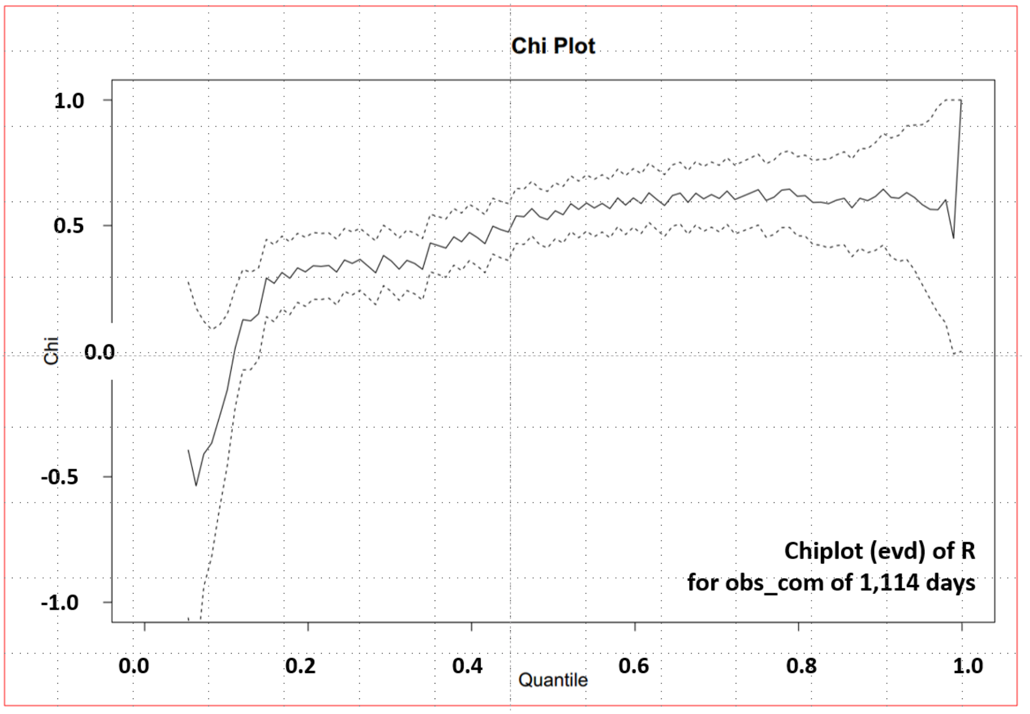
Since values of dependence (χ) can be estimated for any lower or upper threshold, initial trials were performed studying the behaviour of χ over a wide range of thresholds. Findings were similar to those contained in Defra TR3 Report (2005), justifying the selection of an optimal threshold for "alpha" (α) equal to 0.1 corresponding to an annual maximum being exceeded in 9 out of 10 years (see details in Sect. 3). 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.

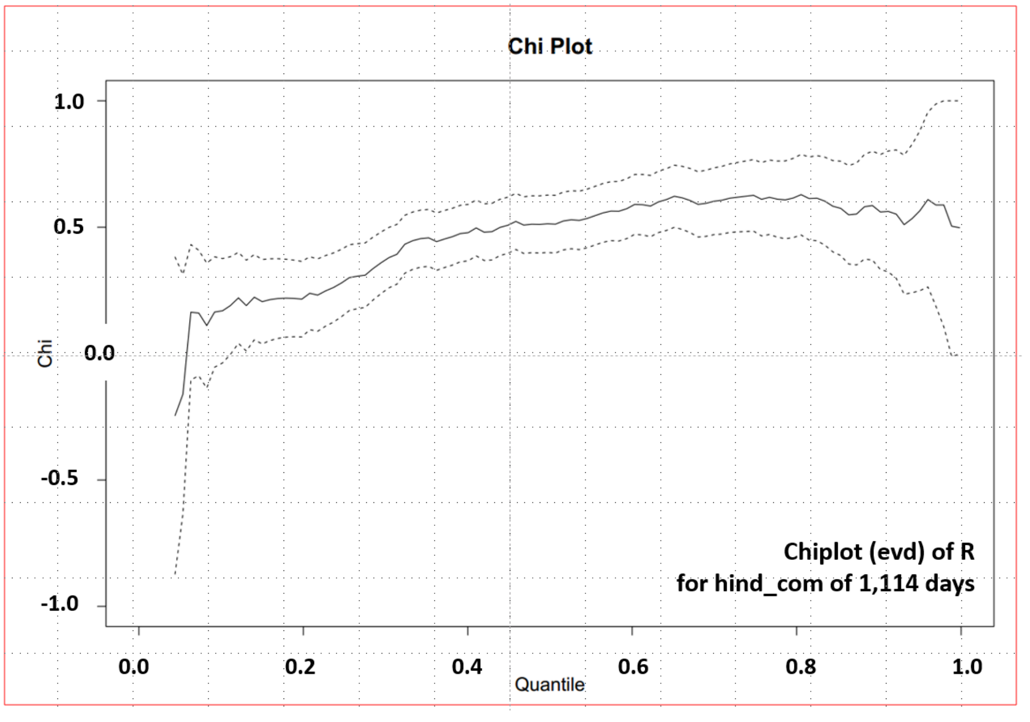
Such annual threshold of ~2.3 events corresponds to the top 80 (Top-80) compound events taking place during any (POT separated) day of the total 12,753 days and it was dictated mainly by two factors: the threshold had to be low enough to allow a sufficient number of data points to exceed it for estimating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

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 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. As mentioned above (Sect. 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.


Chi Plot
1.0
0.5
0.0
-0.5
-1.0
Chi
Chiplot (evd) of R
for obs_com of 1,114 days
0.0
0.2
0.4
0.6
0.8
1.0
Quantile


Chi Plot
1.0
0.5
0.0
-0.5
-1.0
Chi
Chiplot (evd) of R
for hind_com of 1,114 days
0.0
0.2
0.4
0.6
0.8
1.0
Quantile

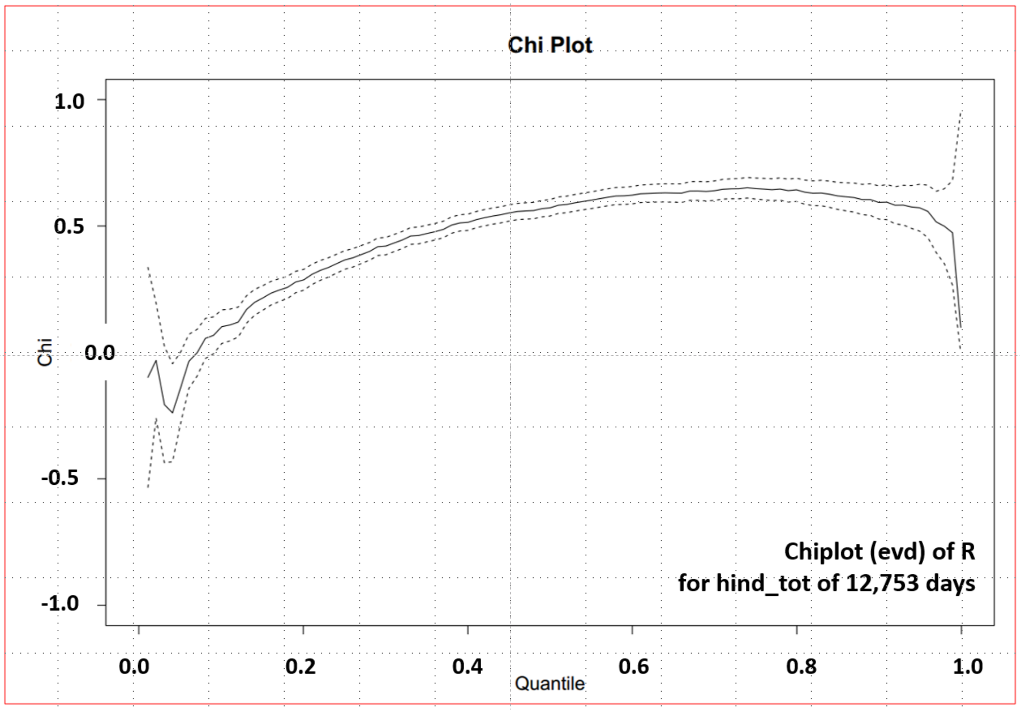


Fig. 1. Estimated χ values between surge (HvH) and wave (LiG) max24 values in c (upper panel) & hind_com (middle panel) and in hind_tot (lower panel) mode by chiplot routine of evd module (R).

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. 12, 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.

Bezak, N., Brilly, M., and Sraj, M.: Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. Hydrological Sciences Journal, 59 (5), 959-977, 2014.

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.

Currie, J.E.: ''Directory of coefficients of tail dependence,'' Department of Mathematics and Statistics Technical Report, ST-99-06, Lancaster University, 1999.

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.

Hazen, A., 1914: Storage to be provided in impounding reservoirs for municipal water supply. Trans. Amer. Soc. Civ. Eng. Pap., 1308 (77), 1547–1550.

Joe, H.: Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.

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.

Nelsen, R.B.: An Introduction to Copulas, Springer-Verlag, New York, 1998.

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.

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E.: Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nature Climate Change, 5(12), 1093-1097, 2015.

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 observations (>0.95m) and hindcasts (>0.89m) for HvH tide gauge station over the common time interval of 1,114 days. Compound events of surge and wave (i.e., both surge & wave above critical threshold) are marked by orange 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. Extreme wave height (above 98.5% percentile) values for observations (> 4.07m) and hindcasts (>3.38m) for LiG wave buoy station over the common time interval of 1,114 days. Compound events of surge and wave (i.e., both surge & wave above critical threshold) are marked by orange shade.

#	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 and northern RIEN points are presented analytically in Table 4 and Table 5 based on ~~matlab~~ Matlab routines. 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 and northern 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 in the main text (and in Figure 9) refers to the categorisation adapted by Defra TR1 Report (2005) and TR3 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.