Coastal Engineering 122 (2017) 44–59

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: www.elsevier.com/locate/coastaleng

An event-based approach for extreme joint probabilities of waves and sea levels

Franck Mazas*, Luc Hamm

ARTELIA, 6 rue de Lorraine, 38130 Echirolles, France

ARTICLE INFO

Keywords:
Joint probabilities
Event
POT-JPM
Extreme-value copula
Upper tail dependence coefficient
Wave height
Sea level
JOIN-SEA
Chi-plot

ABSTRACT

A methodology for determining extreme joint probabilities of two metocean variables, in particular wave height and sea level, is presented in the paper. This methodology focuses in particular on the sampling of the time series, which should be based on the notion of event: either the event generating the variables whose joint probabilities are wanted (such as a storm generating waves and surges) or the event that is a result of the combination of these variables (such as a beach erosion event generated by waves at high sea level). A classification is proposed for multivariate analyses in order to help the choice of the sampling method. The dependence between the variables is analysed using tools such as the chi-plot, of which an enhanced presentation is proposed, then is modelled by extreme-value copulas (Gumbel-Hougaard, Galambos and Hüsler-Reiss) estimated by Canonical Maximum Likelihood or by the upper tail dependence coefficient. Joint return periods are then computed. A comparison is made with a simulation from the JOIN-SEA software on a dataset of wave height and sea levels offshore Brest, France. Then the bivariate methodology is extended to a multivariate framework. The distribution of sea level is determined by an indirect approach (extrapolation of extreme surges then convolution with the astronomical tide) and the dependence is analysed between the wave height and the surge component only. A bidimensional convolution between the joint distribution of wave height and surge and the distribution of the astronomical tide yields the joint distribution of wave height and sea level. The application of this method to the dataset of Brest and its comparison with the bivariate approach are finally discussed.

1. Introduction

The estimation of extreme environmental variables has been widely studied in the univariate cases. In the field of coastal engineering, the methodologies have been progressively enhanced and now provide reliable estimates of extreme wave heights, storm surges, wind speed…, provided the quantity of data is sufficient. Looking back, a gradual convergence towards the so-called "GPD-Poisson" model appears. Fitting a Fisher-Tippett distribution (now known as GEV distribution) to a sample of annual maxima was among the first popular methodologies. In the mid-1990s, the IAHR Working Group on Extreme Wave Analysis recommended using Peaks-Over-Threshold (POT) declustering along with a Weibull distribution estimated by maximum likelihood [50]. A few years later, the logics of the POT declustering was pushed a step further by fitting a Generalized Pareto Distribution (GPD) to the peak excesses, since the law of the exceedances over a threshold u asymptotically tends to a GPD when u is high enough [55]. The GPD-Poisson model is now widely recommended [36] for univariate extremes and a detailed description can be found in Coles [9].

Along with fellow researchers, the authors proposed several improvements within this general framework, based upon their daily practice of determining design waves and sea levels for coastal engineering projects. First, Mazas and Hamm [52] advocated a multi-distribution approach in order to deal with their experience that the GPD often has a tendency to under-estimate the quantiles such as 100-year H_s if u is not high enough. This approach considers additional distributions to the GPD (e.g. the Weibull, Exponential or Gamma distributions), following other authors (e.g. Gōda and Kudaka [27] or Van Vledder et al. [72]). Second, Mazas et al. [51] investigated the behaviour of the Maximum Likelihood Estimator (MLE) when u varies, concluding that the conditions for the MLE to be valid ([47], chapter 6) are not met and proposing to use the L-moments estimator [39] instead along with a location parameter μ (3-parameter GPD). Last, Bernardara et al. [2] provided a deeper justification of the recommendation by Mazas and Hamm [52] for a two-step framework for over-threshold modelling. The first step consists of event identification and characterization, based on physical considerations, in order to set up an i.i.d. (independent and identically distributed) sample from the time

---

* Corresponding author.

http://dx.doi.org/10.1016/j.coastaleng.2017.02.003
Received 24 May 2016; Received in revised form 30 January 2017; Accepted 5 February 2017

series. The POT declustering that selects peaks over a physical threshold u_p is the most frequent method. The second step is a statistical optimization aiming at determining the optimal threshold u_s above which the peak excesses should be fitted by a proper statistical distribution such as the GPD.

However, many coastal phenomena, such as overtopping, beach erosion, load on a structure, coastal flooding or others are the result of the combined actions of two physical processes or more such as sea level, waves, currents or winds. The most common cases are the combination of high sea levels and large wave heights and the joint occurrence of a storm surge along with a high astronomical tide level. Thus the univariate extreme models must be extended to a bivariate and more generally multivariate dimension.

However multivariate analyses may be quite different as regards sampling, dependence modelling, output… Generally speaking, three types of multivariate cases may be distinguished and we propose the following classification:

- Type A: a metocean process described by several parameters (e.g.: a sea state described by its significant wave height H_s, its peak period T_p, its peak direction θ_p, its directional spreading…);
- Type B: a metocean process that can be broken down in several elementary processes (e.g. a sea state made of a swell system and a wind sea system; or the sea level made of a mean sea level, the astronomical tide, the meteorological surge, the wave set-up…);
- Type C: the joint occurrence of several distinct metocean processes (e.g. waves, sea level, wind, current).

In the literature, the joint probabilities of tide and surge are among the soonest to have drawn deep attention, which is not surprising considering the stakes in safety associated with coastal flooding in macrotidal environments: this requires a Type B analysis. A methodology was built and progressively enhanced for determining extreme sea levels by what is called an indirect approach [31], i.e. a separate analysis of the deterministic astronomical tide component T and of the stochastic meteorological surge component S. Pugh and Vassie [57] introduced the Joint Probability Method (JPM), extended to the Revised Joint Probability Method (RJPM) by Tawn and Vassie [70] who introduced a GEV for fitting extreme surges and the extremal index. Tawn [69], Dixon and Tawn [16] then refined the RJPM by accounting for tide-surge dependence.

Indeed, the dependence that may exist between the different variables is a key issue of multivariate analysis. This dependence may be of different types: it may exist because large values of metocean processes such as waves, surges and winds are often generated together by a larger scale process (a storm), or because of interactions between two metocean processes: e.g. onshore wave height or surge magnitude depend on the water depth and thus on the water level.

The modelling of the dependence is at the core of the models built for Type C analyses: estimating the joint occurrence of two (or more) distinct metocean processes such as waves and sea levels. Tawn [67] was amongst the first to apply the parametric models for the dependence function discovered by Gumbel [30], particularly the logistic model, in the field of coastal engineering. The word "copula" was not present yet, though it was used by Sklar [62] as soon as the late 1950s, but here we have a mathematical object that links the distributions of the variables *via* their dependence. The logistic model is now widely known as the Gumbel-Hougaard copula.

In 1994, the UK Ministry of Agriculture, Fisheries and Food (MAFF, now DEFRA) started the funding of joint research projects at HR Wallingford and Lancaster University. A complete methodology for determining the joint probability of waves and sea levels was established ([10,32–34,54], among other references) and the software JOIN-SEA was developed and spread. Joe et al. [41], Zachary et al. [78] and De Haan and De Ronde [13] also applied these "dependence functions" for multivariate environmental extremes. Last, Heffernan and Tawn [37] proposed a semi-parametric approach which applies whether the variables are asymptotically dependent or asymptotically independent and is suitable for highly multidimensional analyses.

During the last decade or so, the copula functions spread into the field of coastal engineering and were used for analysing the combination of two or more variables: wave heights and periods [14]; storm surges and wind waves [15,73]; wave height, wave period and storm duration [12]; wave height, wave period, storm duration and storm surge [48].

Sampling and dependence are closely linked and here it is assumed that a sampling approach based on the notion of event is relevant for a thorough understanding of the physics. Callaghan et al. [5] highlight the interest of working with meteorological events, defined by the peak wave and sea level conditions but also by their duration and the spacing between two successive events, in order to understand the occurrence of beach erosion events. Li et al. [48] also consider events (storms) for the assessment of coastal flooding hazard in the Netherlands. Mazas et al. [53] proposed a "POT-JPM" approach for determining extreme sea levels (Type B analysis), distinguishing the distribution of events (surge events and sea level events) from the distribution of sequential values (e.g. hourly surges or high tide sea levels). The POT-JPM approach also provides confidence intervals for extreme sea levels (the uncertainty coming from the extrapolation of extreme surges) and can account for tide-surge dependence.

On February 2010, 28th, the storm Xynthia wreaked havoc on the French Atlantic coastline, causing more than 30 casualties because of coastal flooding in inhabited low areas and reminding coastal engineers the terrible effect of the combination of a storm surge occurring at the high water of a spring tide. But an in-depth analysis also shown the role played by the waves, that increased the ocean roughness and whose breaking added a set-up component [3]. As a consequence, an accurate estimation of such an event requires considering waves and sea levels on the one hand, and a separate analysis of storm surge and astronomical tide on the other hand.

The present paper is the result of the attempts by the authors to extend their event-based framework to the joint occurrence of two distinct metocean processes (Type C analysis), while applying an indirect approach for determining extreme sea levels. It should then be understood as a continuity of the work presented in the aforementioned references.

In Section 2, a bivariate methodology is presented, which can be applied for determining extreme joint probabilities of waves and sea levels (Type C), storm duration and peak H_s, storm peak H_s and T_p (Type A)… While following a classical four-step framework, the different stages are made consistent with previous publications and a particular focus is made on data sampling. Different methods are discussed, among which high tide sampling, bivariate threshold and the use of a univariate response function that allows accounting for covariates.

In Section 3, this methodology is extended to incorporate an indirect approach for determining extreme sea levels (separate analysis of tide and surge, then recombined by convolution). The main interest consists in the dependence that is modelled between the waves and the meteorological surge, instead of the total sea level.

In Section 4, the dataset of the case study of Brest, France is presented and applied first to the bivariate methodology with comparison with JOIN-SEA results, kindly provided by Dr Peter Hawkes (HR Wallingford), then to the multivariate methodology. The differences between the methodologies and the extension to other applications are discussed in Section 5.

2. Bivariate methodology

In the U.K., current advanced engineering practice largely relies on the approach implemented in the JOIN-SEA software [34]. Despite the practical difficulties in routinely using this software, which thus is not