Coastal Engineering 88 (2014) 1–14

Contents lists available at ScienceDirect

Coastal Engineering

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

Practical guidelines for multivariate analysis and design in coastal and off-shore engineering

G. Salvadori [a,*,1], G.R. Tomasicchio [b], F. D'Alessandro [b]

[a] *Dipartimento di Matematica e Fisica, Università del Salento, Provinciale Lecce-Arnesano, P.O. Box 193, I-73100 Lecce, Italy*
[b] *Dipartimento di Ingegneria dell'Innovazione, Università del Salento, Provinciale Lecce-Arnesano, P.O. Box 193, I-73100 Lecce, Italy*

ARTICLE INFO

Article history:
Received 22 January 2013
Received in revised form 15 January 2014
Accepted 20 January 2014
Available online 14 February 2014

Keywords:
Multivariate copula
Multivariate return period
Multivariate coastal and off-shore design
Multivariate breakwater design
Randomization (jittering)
Sea storm

ABSTRACT

A frequent statistical problem in many coastal and off-shore engineering situations is to estimate the probability of structural failure expressed in terms of Return Period and Design Quantile. Usually, only the univariate approach is carried out to quantify the risk of failure. However, coastal and off-shore structures typically fail because of the occurrence of a critical combination of all the variables at play in a single sea storm: thus, it may be important to consider the joint occurrence of dangerous conditions. The present manuscript provides practical guidelines in order to carry out a sensible multivariate analysis of the available data, including a randomization procedure to cope with repeated observations. In addition, suitable strategies for performing multivariate design are presented and discussed. A practical case study is used to show the application of the techniques illustrated throughout the paper, and a preliminary rubble mound breakwater design is also carried out.

1. Introduction

A frequent statistical problem in many coastal and off-shore engineering situations is the estimate of the probability of structural failure expressed in terms of Return Period (hereinafter, *RP*) and Design Quantile (hereinafter, *DQ*). The traditional definition of RP is as “the average time elapsing between two successive realizations of a prescribed event”, which clearly has a statistical base. Equally important is the related concept of DQ, generally defined as “the value(s) of the variable(s) characterizing the event associated with a given RP”.

The importance of the concept of RP in coastal and off-shore engineering (and, more generally, in civil engineering) is well known, since it is used for designing and sizing structures, for the identification of dangerous events, for rational decision making, and for risk assessment (for a review, see Singh et al., 2007, and references therein). The inspiring principle of this work is that the RP is used to design and size the structure of interest: thus, the corresponding failure region depends upon the chosen RP. Clearly, other criteria can be adopted.

Univariate frequency analysis has long been carried out in past decades, both in coastal and off-shore applications: among others, see (Ferreira and Guedes Soares, 1998; Ferreira and Guedes Soares, 2000; Goda, 1988; Guedes Soares and Scotto, 2001; Guedes Soares and Scotto, 2004; Haver, 1985; Krogstad, 1985; Kuwashima and Hogben, 1986; Mathiesen, 1994; Petrov et al., 2013; Smith, 1988), and references therein. Note that, in the univariate case, the DQ is usually identified without ambiguity (Chow et al., 1988): essentially, it is enough to invert the probability distribution at play. In many coastal and off-shore structural design problems, univariate theory is usually applied in order to quantify the risk of failure due to (extreme) sea conditions: frequently used models are the Generalized Extreme Value distribution, the Generalized Pareto distribution, and the Weibull distribution, although quite a few further different choices are available (see, e.g., the references cited above).

However, in general, several are the (dependent) variables which characterize sea storms: for instance, the significant wave height, the storm duration, the storm inter-arrival time, the peak wave period, the water level, and the direction, which may represents key variables when dealing with coastal and off-shore dynamics, and durability/reliability/fatigue assessment (Boccotti, 2000).

Traditionally, coastal structures are designed mainly for a characteristic value of significant wave height, whereas off-shore structures (de Waal and van Gelder, 2005; Tomasicchio et al., 2012; Wist et al., 2005) and/or sandy coastal bodies (beaches, dunes) (D'Alessandro

---

* Corresponding author. Tel.: +39 0832 29 7584; fax: +39 0832 29 7594.
E-mail addresses: gianfausto.salvadori@unisalento.it (G. Salvadori), roberto.tomasicchio@unisalento.it (G.R. Tomasicchio), felice.dalessandro@unisalento.it (F. D'Alessandro).

[1] Helpful discussions with C. Sempi (Università del Salento, Italy) and F. Durante (Free University of Bozen-Bolzano, Bolzano,Italy) are acknowledged. The research was partially supported by the Italian M.I.U.R. via the project “Metodi stocastici in finanza matematica” (PRIN 2010). Also the support of the CMCC—Centro Euro-Mediterraneo sui Cambiamenti Climatici—Lecce (Italy) is acknowledged.

0378-3839/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.coastaleng.2014.01.011

et al., 2012; Tomasicchio et al., 2011a; Tomasicchio et al., 2011b; van Gent et al., 2008) depend very much on a characteristic wave period, which is the variable that governs their dynamic amplification, and thus their responses. However, in practical applications, coastal and off-shore structures may suffer from severe damages because of the occurrence of critical combinations of the variables which coexist in a single sea storm. In turn, the lack of knowledge concerning their joint statistics may severely limit the effectiveness of coastal, port and off-shore structures protection, and can lead to expensive and inappropriate decisions (Li et al., 2008). As a consequence, it may be important to consider the joint occurrence of combined conditions: among others, see (Corbella and Stretch, 2012; De Michele et al., 2007; Dong et al., 2013; Ferreira and Guedes Soares, 2002; Jonathan et al., 2010; Repko et al., 2005), and references therein.

For the sake of illustration, in the following we shall mainly concentrate on the coastal framework: thus, the (multivariate) analyses will involve the variables significant wave height and duration of sea storms. However, the same procedures outlined in this paper can straightforwardly be used to deal with off-shore problems, by jointly considering more relevant variables like, e.g., the period, the significant wave height, the sea storm direction, the sea storm duration, the water level, and so on.

In Hawkes et al. (2002) and (Heffernan and Tawn, 2004), semi-empirical and, respectively, conditional procedures were proposed for working with joint extremes. However, such techniques are slow in being taken up in engineering practice, due to several reasons (Li et al., 2008): on the one hand, the costs to obtain sufficient data for complex methods; on the other hand, the costs of staff training on techniques that are not an industry standard. Nevertheless, there is a growing requirement amongst engineers, researchers and practitioners, to be able to quantify the uncertainty associated with multivariate design conditions.

The identification problem of design events in a multivariate context is of fundamental importance but, at the same time, is of troublesome nature. In addition, also the related problem concerning the construction of a multivariate notion of RP is rather tricky, since different definitions are possible. Recently, several efforts have been spent on these issues: see, e.g., (Belzunce et al., 2007; Chaouch and Goga, 2010; Chebana and Ouarda, 2009; Chebana and Ouarda, 2011; De Michele et al., 2013; Salvadori et al., 2011; Salvadori et al., 2013a; Serfling, 2002), and also (Gräler et al., 2013) for a thorough review and comparison of procedures. Here we address the following crucial question: “How is it possible to calculate, in a probabilistically well-founded and consistent way, the return periods and the critical design conditions in a multivariate context?”

The present manuscript provides practical guidelines in order to carry out a sensible multivariate analysis of the available data. In addition, suitable strategies for performing multivariate design are presented and discussed, including a randomization procedure to cope with repeated observations. The proposed approach uses the Theory of Copulas: indeed, recent advances in mathematics (see, e.g., (Joe, 1997; Nelsen, 2006; Salvadori et al., 2007)) show how copulas may represent an efficient tool to investigate the statistical behavior of dependent variables.

The paper is organized as a sequence of successive STEPS to be performed. In Section 2 a preliminary data survey is carried out. In Section 3 a randomization procedure is outlined, to be used when the data base contains repeated observations. In Section 4 we present the STEPS 1, 2, and 3 that should be carried out for performing a sensible univariate/multivariate analysis and fit of the available data. In Section 5 we present the STEPS 4 and 5 that should be carried out for performing a wise multivariate frequency analysis of the available data, including some strategies for multivariate design. In Section 6, as an illustration, a preliminary rubble mound breakwater design is carried out. Finally, in Section 7 suitable conclusions are drawn.

2. Preliminary data analysis

The present paper is of methodological nature, and we feel essential to show the usage of the techniques outlined in this work by considering a certified data base. The data investigated in the present study have been collected at the Alghero wave buoy (Sardinia, Italy), for a period of about 19 years: from July 1st, 1989, to April 5th, 2008. This wave buoy, located at an Italian extremely exposed sea area, is a part of the Italian Sea Wave Measurement network, monitored by ISPRA (see `www.isprambiente.gov.it` and `www.idromare.com`). The data set includes observations of the following variables: the significant wave height, the peak period, the wave direction, and the water temperature—for further details see (Arena et al., 2001; Gencarelli et al., 2006; Piscopia et al., 2002). For the ease of illustration, in the following we shall concentrate on the analysis of two variables only, i.e. the significant wave Height H (in meters), and the sea storm Duration D (in hours).

Following (Piscopia et al., 2002), we adopt a standard criterion in order to (automatically) select independent and homogenous sea storms. Practically, we assume that a storm starts when H crosses upwards the threshold 4 m, and ends when H persists below the same level for at least 24 h. Clearly, different thresholds can be chosen, depending upon the target of the data analysis. Thus, $N = 301$ sea storm events are extracted from the available data base. The selection algorithm guarantees that successive storms are independent (Piscopia et al., 2002), as we shall assume hereinafter. We shall also take for granted that the storms are homogenous, viz. statistically identically distributed.

We stress that a preliminary survey of the data base should always be carried out, in order to fix possible anomalies due to software bugs and/or hardware/instrumental limitations.

3. Randomization

Before proceeding with the analyses, it is important to realize that, in principle, both H and D describe continuous phenomena (viz., a length and a time). Unfortunately, due to a limited instrumental resolution, the available measurements may be a discretized version of the actual continuous values of these variables. For instance, H may be a multiple of the basic resolution equal to 10 cm, and D may be a multiple of 3 h. Thus, for example, two different wave heights (say, $H_1 = 4.01$ m and $H_2 = 4.09$ m) are both recorded as a common height value $H = 4$ m, and similarly for D. As an illustration, in Fig. 1 we show a zoom of the

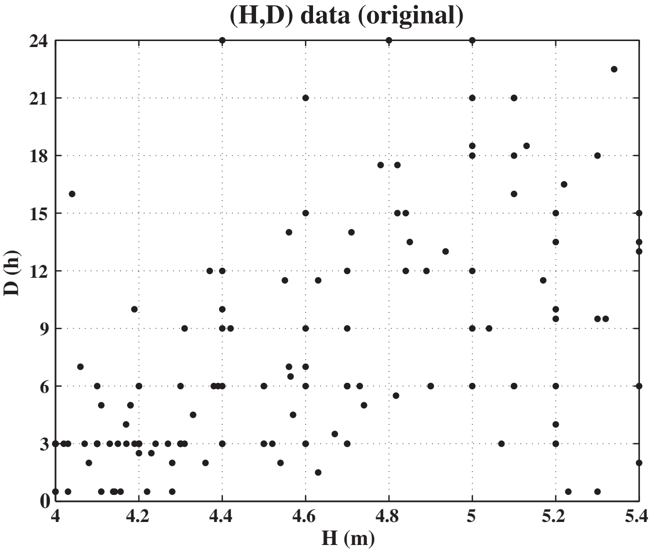

Fig. 1. Zoom of the original (non-randomized) data in the (H,D) region [4,5.4]m × [0,24]h.