Coastal Engineering 135 (2018) 91–111

Contents lists available at ScienceDirect

Coastal Engineering

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

Modeling multivariate ocean data using asymmetric copulas

Yi Zhang [a,*], Chul-Woo Kim [a], Michael Beer [b], Huliang Dai [c], Carlos Guedes Soares [d]

[a] Department of Civil and Earth Resources Engineering, Graduate School of Engineering, Kyoto University, Japan
[b] Institute for Risk and Reliability, Leibniz Universität Hannover, Germany / Institute for Risk and Uncertainty, University of Liverpool, UK / Shanghai Institute of Disaster Prevention and Relief, Tongji University, China
[c] School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, China
[d] Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico, Universidade de Lisboa, Portugal

ARTICLE INFO

Keywords:
Ocean engineering
Joint distribution
Multivariate analysis
Asymmetric copula

ABSTRACT

Multivariate descriptions of ocean parameters are quite important for the design and risk assessment of offshore engineering applications. A reliable and realistic statistical multivariate model is essential to produce a representative estimate of the sea state for understanding the ocean conditions. Therefore, an advanced modeling of ocean parameters helps towards improving ocean and coastal engineering practices. In this paper, we introduce the concepts of asymmetric copulas for the modeling of multivariate ocean data. In contrast to extensive previous research on the modeling of symmetric ocean data, this study is focused on capturing asymmetric dependencies among the environmental parameters, which are critical for a realistic description of ocean conditions. This involves particular attention to both nonlinear and asymmetrically dependent variates, which are quite common for the ocean variables. Several asymmetric copula functions, capable of modeling both linear and nonlinear asymmetric dependence structures, are examined in detail. Information on tail dependencies and measures of asymmetric dependencies are exploited. To demonstrate the advantages of asymmetric copulas, the asymmetric copula concept is compared with the traditional copula approaches from the literature using actual environmental data. Each of the introduced copula models is fitted to a set of ocean data collected from a buoy at the US coast. The performance of these asymmetric copulas is discussed and compared based on data fitting and tail dependency characterizations. The accuracy of asymmetric copulas in predicting the extreme value contours is discussed.

1. Introduction

Offshore and coastal structures facilitate the exploitation of the vast ocean resource, which contributes significantly to technological and economic development. Compared with normal structures on land, offshore structures are bulky, expensive and in most cases constructed in a complex marine environment. The marine environment for offshore structures can be severe, adverse, varying and uncertain. It covers a broad area of climatic factors, which generally include tide, current, wind, waves, ice and other sometimes devastating events such as hurricanes. When addressing different environmental risks for the offshore engineering applications designers are usually required to estimate the environmental conditions at the ocean site, and usually a multivariate analysis is performed (Zhang and Cheng, 2016). For example, the environmental contour method developed by Winterstein et al. (1993) is popular for this purpose. As a basis to produce realistic results it requires a reliable multivariate environmental model for finding the maximum system response associated with a given exceedance probability. In this context the interaction among various environmental influences plays an important role. In practical applications, offshore and coastal structures can suffer from severe damages because of the occurrence of critical combinations of the ocean environmental variables which coexist in extreme weather events such as sea storms (Zhang and Lam, 2014, 2015). In turn, deficiencies in modeling their joint statistics may severely overestimate the safety and effectiveness of coastal and offshore structures, hence lead to unsafe design and consequently lead to expensive and unexpected catastrophes (Bitner-Gregersen, 2015; Zhang et al., 2017a,b). Particularly, the modeling of the joint distribution of wave height and wave period is normally a must in marine engineering applications since the sea state at a specific location primarily depends on these two ocean parameters simultaneously (DNV, 2014). However, the ocean climate system is an extremely complex system that contains many

* Corresponding author.
E-mail address: zhang.yi.25n@st.kyoto-u.ac.jp (Y. Zhang).

https://doi.org/10.1016/j.coastaleng.2018.01.008
Received 30 September 2016; Received in revised form 15 January 2018; Accepted 21 January 2018

more natural factors from both the ocean and the atmosphere. There are various sources of uncertainties and potential biases that influence the ocean conditions (Zhang and Cao, 2015). Specifically, the uncertainties related to the parameter dependencies are one of the most influencing factors. It was recognized that the understanding of the nonlinear dependency between ocean parameters remains one of the most difficult tasks, and the statistical modeling of the multivariate ocean data remains challenging due to their complicated relationships (Ewans and Jonathan, 2014).

Lots of attempts to cope with multivariate statistical analysis can be found in the literature, involving the use of a conditional distribution model (Lucas and Guedes Soares, 2015), a bivariate logistic model (Morton and Bowers, 1996), a Pareto distribution model (Muraleedharan et al., 2015) and so on. Clearly, the traditional conditional joint distribution model is only applicable to the multivariate problem to a certain extent. If the variables contain slightly more complex relationships such as non-constant correlation between pairs of random variables (e.g. the correlation coefficient value may change at extremes), the traditional joint statistical model is not appropriate any more. Therefore, many more advanced techniques have been employed in the multivariate analysis. Among all the developments, the application of copulas has been increasingly popular in recent years. Many initial studies have illustrated that application of copula theory can produce more realistic joint models for the ocean multivariate data. De Michele et al. (2007) have utilized copulas to characterize sea storms in terms of significant wave height, sea storm duration, sea storm inter-arrival (waiting) time, and sea storm direction. Antão and Guedes Soares (2014) have established copula based bivariate models for individual wave steepness and wave height. A similar copula model for description of water levels and waves are also presented by Masina et al. (2015). Montes-Iturrizaga and Heredia-Zavoni (2015) have proposed a formulation for expressing the environmental contours as functions of copulas and show that the dependence structure of sea state parameters can be well presented in this manner. Until recently, Jane et al. (2016) have employed the copula model to predict the wave height at a given location by considering the spatial dependence of the wave height at nearby locations. In contrast to the traditional joint model, a copula approach has the advantage that the dependency structure between the variables can be defined independently of the choice of the marginal distribution. This flexibility is highly desirable in modeling the environmental parameters as the natural factors often exhibit non-obvious dependencies. Moreover, it was also found that the copula model can save numerical effort when it is utilized to characterize the environmental loading in the offshore structural analysis. For instance, Zhang et al. (2015) have demonstrated an approach of using a copula model to characterize the sea load for the reliability analysis of a real jacket structure, which reduces the numerical effort by a factor of five. A practical guideline for using a copula in the design of coastal and offshore engineering applications can be found in Salvadori et al. (2014). Thorough guidelines involving the use of copulas in a structural approach are given in Salvadori et al. (2015). In general, from the recent advances in coastal engineering, it is now widely recognized that a copula approach is very efficient and powerful to model the statistical behavior of ocean dependent variables.

As exciting as the copula approach is, there are some obvious issues, which need to be addressed for a successful application. In former studies it was criticized that most families of parametric copulas can only model data having symmetric dependency (Genest and Favre, 2007). For example, the well established Archimedean copula families are all symmetric. If the data dependencies exhibit asymmetric behavior, the traditional copula model may no longer be adequate. Asymmetric Archimedean copulas are discussed in Grimaldi and Serinaldi (2006). Unfortunately, ocean data, fall into this category; they have been found as asymmetrically dependent in various previous studies. This is especially obvious for the sea state parameters, which are important for in engineering design (deWaal and van Gelder, 2005). Ignoring the asymmetric effects in the modeling of ocean data can be quite critical as it affects the estimates of the response statistics and eventually compromises the quality of the structural reliability assessment. A reason explaining the frequent (possibly unjustified) usage of symmetric (Archimedean) copulas might be that these are the ones provided by the Matlab package, the one traditionally used by maritime engineers. However, asymmetric copulas can remedy this problem. Asymmetric copulas can be constructed based on the families of symmetric copulas. This compounded procedure can significantly improve the fit (Jondeau, 2016). The modeling of the ocean data utilizing the asymmetric copulas has received much attention recently (Vanem, 2016). The well known Khoudraji-Liebscher family, introduced in (Durante and Salvadori, 2010; Salvadori and De Michele, 2010), gives the possibility to construct asymmetric copulas. The application of this family in a maritime context has been mentioned in Salvadori et al. (2014, 2015). De Michele et al. (2013) have also used it for the modeling of drought. Besides, the conditional mixture construction (Vine copulas), first introduced in maritime engineering by De Michele et al. (2007), also provides the possibility to construct asymmetric copulas starting from symmetric ones. However, the theoretical concepts and procedures of constructing an asymmetric copula have not yet been studied in detail. Despite this, it is recognized that there are many candidate asymmetric copulas in theory. These choices provide potent features and practical meaning in ocean and coastal engineering applications. This potential can readily be utilized once the applicability of asymmetric copulas for the modeling of ocean data has been verified and demonstrated. We aim to contribute to this development with the present real case study for demonstrating and highlighting the features, merits as well as limitations associated with asymmetric copulas.

The remainder of this paper is organized as follows. Section 2 presents a general literature review of the existing techniques in modeling multivariate ocean data. Section 3 presents the fundamental knowledge of copula theory and the basic dependence measure concepts. Basic concepts of asymmetry measure as well as the procedures of constructing asymmetric copula models are explained in detail in Section 4. Specific asymmetric copula models for ocean data are developed in Section 5 and compared against traditional parametric copula models based on collected, preconditioned ocean data. To understand the features of using asymmetric copulas in the ocean data modeling, a comparative study between symmetric and asymmetric copula models is presented in Section 6. The concluding remarks of this paper form Section 7.

2. Joint statistical models for ocean data

Among the probabilistic models available in the literature, the most commonly recommended model adopted in offshore engineering design codes is the conditional joint distribution model, which is widely applied to various kinds of ocean data (Burton et al., 2001; Jonathan and Ewans, 2011; Ernst and Seume, 2012). The most pertinent joint distribution model that is applied in ocean engineering is for the significant wave height and peak period, which characterize the spectrum of a sea state. For instance, Guedes Soares et al. (1988) and Bitner-Gregersen and Haver (1989) have demonstrated the use of a joint environmental model, which was constructed based on the combination of the marginal distribution of wave height and conditional distribution of the wave period. Later on, Ochi (1992) introduced a bivariate log-normal distribution in the modeling of the significant wave height and peak period. Generally, these conditional bivariate distribution models assume significant wave height follows a Weibull distribution while wave peak period follows a log-normal distribution whose model parameters are conditional on significant wave height. The primary reason for using such conditional distribution model is generally that the significant wave height is the most important parameter, which affects design conditions of ocean structures whereas other parameters have less influence. This also agrees well with the practice design code (DNV, 2010), which utilizes a bivariate conditional model for the wave height and wave period. This concept of conditional distribution models can also be extended to modeling