International Journal of Offshore and Coastal Engineering
Vol.7 | No. 2 | pp. 88 – 93 | November 2023
e-ISSN: 2580-0914

Submitted: July 06, 2023 | Revised: August 31, 2023 | Accepted: October 25, 2023

Bivariate Archimedean Copulas to Solve Complex Dependency in Marine Engineering Problems

Adhitya Ryan Ramadhani[a,*] and Waskito Pranowo[b]

[a)] Department of Mechanical Engineering, Universitas Pertamina, Jakarta Selatan, Indonesia
[b)] Department of Geophysical Engineering, Universitas Pertamina, Jakarta Selatan, Indonesia
*Corresponding author: adhitya.rr@universitaspertamina.ac.id

ABSTRACT
Climate change has made offshore installations suffer severe consequences from extreme marine hazards. These offshore installations were designed to withstand extreme natural hazards. Marine natural hazards commonly occur interdependently. Thus, a multivariate model to capture this complex dependence structure is necessary. Practically, modelling marine natural hazards is usually assumed to be mutually independent or correlated by a multivariate Gaussian distribution. However, this biased assumption is not relevant to capture the real dependence structure between marine parameters. Copula functions are used to model the complex dependence structure between marine parameters. A simplified case study is selected to illustrate the modelling between wave height and wind speed. Results are compared with the traditional joint probability approach to demonstrate the advantage of copula functions. The use of copula functions provides a better result to model any complex dependence structure between correlated variables.

Keywords: *Dependence modelling, Copula functions, Offshore structures, Probabilistic method, Marine engineering.*

1. INTRODUCTION

Climate change has made offshore installations suffer severe consequences from extreme marine hazards. A common natural hazard affecting offshore installations in harsh environments is hurricanes. Hurricane Katrina caused 44 offshore oil and gas installations to be destroyed and 21 others to be damaged, while 69 were destroyed and 32 were damaged in the aftermath of Hurricane Rita [1]. Due to their location, offshore installations are vulnerable to other natural hazards, such as storms, earthquakes, and ice load [2]. Thus, it is necessary to understand the characteristics of marine natural hazards to reduce the risk associated with marine extreme events.

The marine environment is a complex system with many sources of uncertainties. Generally, two types of uncertainties areound in a complex engineering system: aleatory and epistemic [3,4]. Aleatory uncertainties deal with randomness in a system's characteristics, while epistemic uncertainties are usually caused by a lack of information about the observed system [5]. These two types of uncertainties occur concurrently in the engineering system. Updating and collecting information is sufficient to treat epistemic uncertainties, while aleatory uncertainties require more appropriate approaches to deal with [6]. The most commonly mistreated epistemic uncertainty in engineering problems is dependence between correlated variables. The assumption of independence when investigating marine environmental parameters should be eliminated. Thus, a multivariate approach that can consider dependence between variables is required.

There have been many attempts to model multivariate variables in marine engineering. The most common approach is the conditional joint distribution model. Lucas & Guedes Soares (2015) compared the conditional modelling approach, the Plackett model, and Box-Cox transformations to identify which method performed best. Muraleedharan et al. (2015) used the generalized Pareto and three-parameter Weibull distribution to model the average conditional exceedance of wave peak periods. In addition to the conditional joint probability distribution, Nataf transformation is also a widely used method to deal with multivariate analysis. Sinsabvarodom et al. (2020) used Nataf transformation to deal with bivariate marine parameters and their dependence structure. Xiao (2014) calculated the equivalent correlation coefficients for two variables using the Nataf transformation. Although conditional joint distribution and Nataf transformation have been widely used in marine engineering problems, they can also be applied to follow a certain type of distribution function. A more flexible and robust multivariate model is necessary to capture the dependence between marine variables without following a certain type of marginal distribution.

Copula functions are powerful tools for performing multivariate analysis. The studies of copula functions started in the financial industry [11-13], and then started to be implemented to model risks in the process industry [14-16]. Copula functions also attracted attention to model the multivariate analysis of hydrological variables, such as rainfalls and floods [17,18]. From these studies, copula functions were concluded to be flexible in capturing dependence structure between correlated variables. Copula functions also do not require a certain type of marginal distribution. Due to these advantages, copula functions have started gaining attention in modelling marine environmental variables [19-23].

This paper investigates copula functions to model bivariate marine ocean parameters with symmetric dependence structure. The results are compared with other traditional and conditional methods to study the significance of copula functions.

2. METHODOLOGIES

Figure 1 shows the overall framework used to model bivariate marine variables using copula functions.

2.1 Copula Functions

A Copula is a function that joins marginal distributions to create a multivariate model [24]. Sklar's Theorem defined copula as:

Let H be an n-dimensional distribution function with marginal distribution $F_1, F_2, \ldots, F_n$, then there exists a copula C:

$$H(x_1, x_2, \ldots, x_n) = C\big(F_1(x_1), F(x_2) \ldots, F(x_n)\big) \quad (1)$$

The joint conditional probability distribution can be estimated using:

$$f_{X_2|x_1} = \frac{f_{X_1X_2}(x_1, x_2)}{f_{X_1}(x_1)} = C_{X_1X_2}\left(F_{X_1}(x_1), F_{X_2}(x_2)\right). f_{X_2}(x_2) \quad (2)$$

Due to their flexibility, Copula functions do not require a specific marginal distribution. Copula function transforms a random variable into a uniform cumulative distribution function [25].

A wide range of copula families is available. One that has been widely known is the Archimedean family. In this copula families, Clayton, Gumbel, and Frank copulas are used in this paper. These three common Archimedean copula families and their parameters are presented in Table 1.

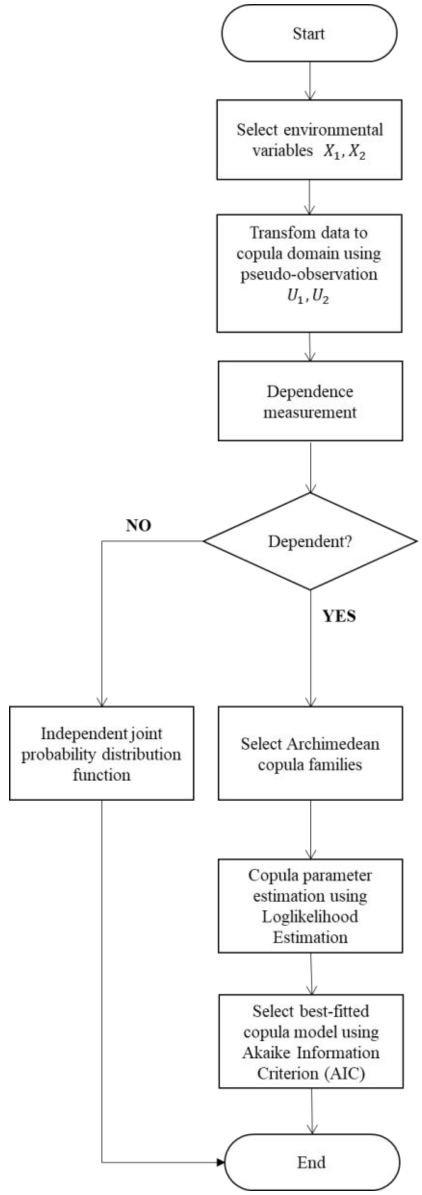

Figure 1. Research framework for copula modelling

Table 1. Common Archimedean copulas [25]

Copula	$C_\gamma(u_1, u_2)$	**Generating function** $\phi_\gamma(t)$	$\gamma \in$
Clayton	$(u_1{}^{-\gamma} + u_2{}^{-\gamma} - 1)^{\frac{-1}{\gamma}}$	$\frac{\gamma}{\gamma+2}$	(1.∞)
Gumbel	$exp\left\{-[(-lnu_1)^\gamma + (-lnu_2)\gamma]^{\frac{1}{\gamma}}\right\}$	$1 - \frac{1}{\gamma}$	[1,∞)
Frank	$\frac{-1}{\gamma} ln\left(1 + \frac{(e^{-u_1\gamma} - 1)(e^{-u_2\gamma} - 1)}{e^{-\gamma} - 1}\right)$	$1 - \frac{4}{\gamma}\big(1 - D_1(\gamma)\big)$ Where $D_1(\gamma) = \frac{1}{\gamma}\int_0^\infty \frac{tdt}{exp(t) - 1}$	$(-\infty, \infty)$

2.2 Dependence Measurement

Three different methods are used to measure the dependence of correlated variables. Linear Pearson correlation is used to account for linearity and estimated using this equation:

$$\rho(X,Y) = \frac{CoV(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \quad (3)$$

Where X, and Y are the observed data.

While Spearman's rho (ρ_S) and Kendall's tau (τ) are used to illustrate nonlinearity and estimated using these respective equations [27] :

$$\rho_s = 12\int\int uv\, dC(u,v) - 3 \quad (4)$$

$$\tau = 4\int C(u,v)\, dC(u,v) - 1 \quad (5)$$

Where u and v are the transformed data in the copula domain.

2.3 Estimation of Copula Parameter

Maximum Log-Likelihood Estimation (MLE) is used to estimate the copula parameter. The log-likelihood function $L(\theta)$ is expressed as [28] :

$$L(\theta) = \sum_{i=1}^{N} lnc(u_1, u_2; \theta) \quad (6)$$

$L(\theta)$ is then maximized to get the estimated parameter θ

2.4 Selection of the Best-fit Copula

Akaike Information Criterion (AIC) is used to select the best-fit copula function among all the compared Archimedean Copula families. AIC scores can be estimated using this equation:

$$AIC = -2L(\theta) + 2k = -2\sum_{i=1}^{N} lnc(u_{1i}, u_{2i}; \theta) + 2k \quad (7)$$

Where k is the unknown number of parameters estimated using MLE. The lowest AIC score indicates the best-fit model [28,29].

3. RESULTS AND DISCUSSION

Wind speed (m/s) and wave height (m) are selected as the observed marine parameters in this paper. A data set was obtained from the mouth of Placentia Bay, Newfoundland and Labrador, Canada [30]. The data were recorded hourly from January 1st, 2013, to December 31st.

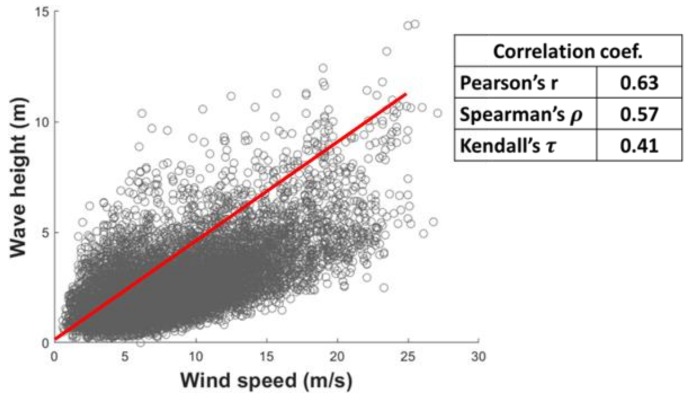

Figure 2. Scatter plot and dependence measurements for wind speed and wave height

Figure 2 shows that wind speed and wave height correlate positively and powerfully. From the scatter plot, it can also be seen that this environmental variables pair has a slightly nonlinear relationship. Copula functions are also able to deal with nonlinearity [25]. In this paper, only symmetric dependence between correlated variables is considered. Thus, copula functions significantly capture this complex dependence structure between variables without worrying about their marginal distributions.

Copula functions from the Archimedean families are used to model the dependence structure between wind speed and wave height. The correlated bivariate model is used to construct Clayton, Gumbel, and Frank copulas.

Table 2 presents the copula parameter for all Archimedean copulas selected in this paper. Akaike Information Criterion (AIC) selects the best-fitted copula function to construct the bivariate model of wind speed and wave height. The lowest AIC score indicates the best-fitted copula. Based on this AIC score, the Gumbel copula has the lowest score, which indicates the best-fitted copula to model wind speed and wave height, with copula parameter $\gamma = 1.6352$.

Table 2. Best-fitted Archimedean copulas and their parameters

Copula	Copula Parameter	$L(\theta)$ ($\times 10^5$)	AIC ($\times 10^5$)
Clayton	$\gamma = 1.0049$	-3.3290	6.6581
Gumbel	$\gamma = 1.6352$	-3.3127	6.6255*
Frank	$\gamma = 4.2656$	-3.3168	6.6337

*Lowest AIC score indicates the best-fitted copula

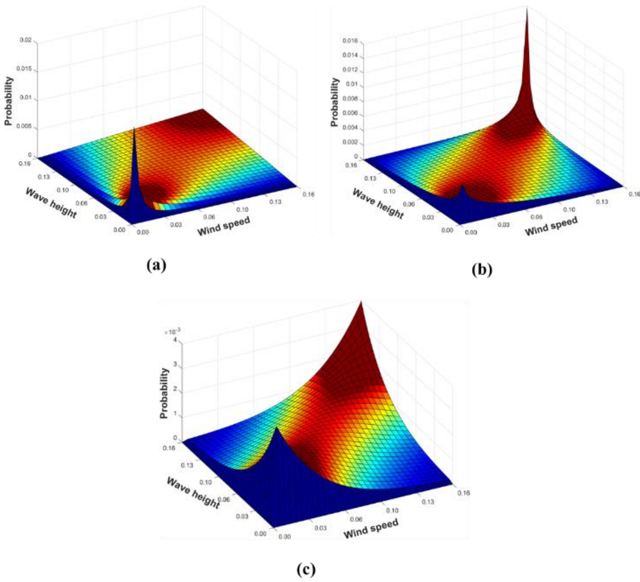

Figure 3. The probability distribution function for wind speed and wave height modelled using (a) Clayton copula, (b) Gumbel copula, (c) Frank copula

Figure 3 shows the probability distribution functions constructed using copulas. Clayton is best fitted to model correlated variables that have strong lower tail dependence. In comparison, Gumbel is best fitted to model correlated variables with strong upper-tail dependence. Frank copulas are used to model correlated variables with no stronger tail dependence on each side. Thus, from these characteristics, copula functions can consider all types of tail dependence that are usually neglected when performing multivariate analysis. From Figure 3, as Gumbel is the best-fitted copula for the selected environmental parameters, it can be seen that the bivariate Gumbel model shows a strong correlation at the upper tail. We can also see from Figure 2 that there are plotted data in the upper tail that should be considered. Failure to consider these data sets will lead to misinterpretation of capturing the real dependence structure of marine environmental data.

Error values and distribution fittings are presented to identify which model is best fitted to construct bivariate models for wind speed and wave height. The Root-Mean-Square error (RMSE) and the mean absolute error are used to compare the results. The bivariate models compared are the Gumbel copula, independent joint probability function, and conditional joint probability function. Figure 4 shows the distribution fittings of these models. From this figure, the Gumbel copula provides the most negligible error value and shows the best fit compared to the other two models. Thus, copula models, in this case, Gumbel copula, are found to be the most appropriate bivariate model to capture the actual complex dependence structure between wind speed and wave height. This result also agrees with the findings from other related studies [25,31,32].

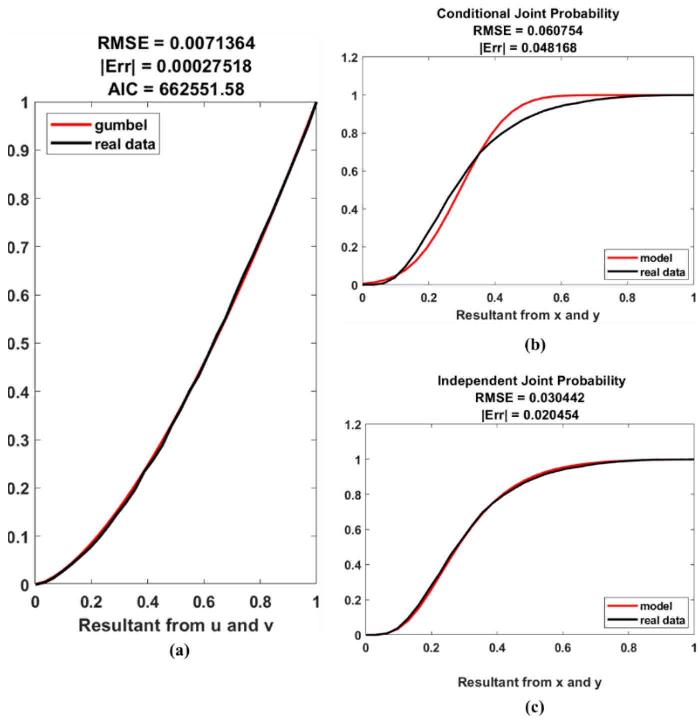

Figure 4. Error values for (a) Gumbel copula, (b) Conditional Joint Probability, (c) Independent Joint Probability

4. CONCLUSION

Uncertainties are very common in a complex engineering system. The most commonly neglected epistemic uncertainty is the dependence model between correlated variables. This paper investigates the application of copula functions to construct bivariate models between two marine environmental variables. The Archimedean copula families are selected, and wind speed and wave data are collected from an available resource.

Wind speed and wave height are best modelled using Gumbel copulas, providing the least Akaike Information Criterion (AIC) score. These two marine variables show a positive and strong correlation. Nonlinearity is also spotted in the constructed scatter plot. Copula functions are concluded to capture all types of complex dependence structures. Compared to the other two commonly used models, a comparison is also provided to conclude that the Gumbel copula is best fitted to capture the complex dependence structure between wind speed and wave height. Copula models are then concluded to be a powerful tool to capture complex dependence structures between variables and can be implemented further in engineering problems. Copulas can be a very useful input to assess an engineering system's risk and reliability, which usually shows dependence among the observed variables.

Future works will consider other copula families and asymmetric dependence structure between correlated variables.

REFERENCES

[1] Cruz AM, Krausmann E (2008) Damage to offshore oil and gas facilities following hurricanes Katrina and Rita: An overview. J Loss Prev Process Ind 21:620–626
[2] Wilson JF (2002) Dynamic Offshore Structure. John Wiley & Sons, New Jersey
[3] McKeand AM, Gorguluarslan RM, Choi S-K (2021) Stochastic analysis and validation under aleatory and epistemic uncertainties. Reliab Eng Syst Saf 205:107258
[4] Reilly AC, Baroud H, Flage R, Gerst MD (2021) Sources of uncertainty in interdependent infrastructure and their implications. Reliab Eng Syst Saf 213:107756
[5] Chen C, Yang M, Reniers G (2021) A dynamic stochastic methodology for quantifying HAZMAT storage resilience. Reliab Eng Syst Saf 215:107909
[6] Chen Y, Li S, Kang R (2021) Epistemic uncertainty quantification via uncertainty theory in the reliability evaluation of a system with failure Trigger effect. Reliab Eng Syst Saf 215:107896
[7] Lucas C, Guedes Soares C (2015) Bivariate distributions of significant wave height and mean wave period of combined sea states. Ocean Engineering 106:341–353
[8] Muraleedharan G, Lucas C, Martins D, Guedes Soares C, Kurup PG (2015) On the distribution of significant wave height and associated peak periods. Coastal Engineering 103:42–51
[9] Sinsabvarodom C, Chai W, Leira BJ, Høyland K V., Naess A (2020) Uncertainty assessments of structural loading due to first year ice based on the ISO standard by using Monte-Carlo simulation. Ocean Engineering 198:106935
[10] Xiao Q (2014) Evaluating correlation coefficient for Nataf transformation. Probabilistic Engineering Mechanics 37:1–6
[11] Zhi B, Wang X, Xu F (2022) Managing inventory financing in a volatile market: A novel data-driven copula model. Transp Res E Logist Transp Rev 165:102854
[12] Fang Y, Madsen L (2013) Modified Gaussian pseudo-copula: Applications in insurance and finance. Insur Math Econ 53:292–301
[13] Boako G, Tiwari AK, Ibrahim M, Ji Q (2019) Analyzing dynamic dependence between gold and stock returns: Evidence using stochastic and full-range tail dependence copula models. Financ Res Lett. https://doi.org/10.1016/j.frl.2018.12.008
[14] Hashemi SJ, Ahmed S, Khan FI (2015) Correlation and Dependency in Multivariate Process Risk Assessment. IFAC-PapersOnLine 48:1339–1344
[15] Kabir S, Taleb-Berrouane M, Papadopoulos Y (2019) Dynamic reliability assessment of flare systems by combining fault tree analysis and Bayesian networks. Energy Sources, Part A: Recovery, Utilization and Environmental Effects. https://doi.org/10.1080/15567036.2019.1670287
[16] Taleb-berrrouane M (2019) Dynamic Corrosion Risk Assessment in the Oil and Gas Production and Processing Facility (Doctoral dissertation, Memorial University of Newfoundland). Memorial University of Newfoundland
[17] Jiang C, Xiong L, Yan L, Dong J, Xu C-Y (2019) Multivariate hydrologic design methods under nonstationary conditions and application to engineering practice. Hydrol Earth Syst Sci 23:1683–1704
[18] Salleh N, Yusof F, Yusop Z (2016) Bivariate copulas functions for flood frequency analysis. p 060007
[19] De Michele C, Salvadori G, Passoni G, Vezzoli R (2007) A multivariate model of sea storms using copulas. Coastal Engineering 54:734–751
[20] Ramadhani A, Khan F, Colbourne B, Ahmed S, Taleb-Berrouane M (2021) Environmental load estimation for offshore structures considering parametric dependencies. Safety in Extreme Environments. https://doi.org/10.1007/s42797-021-00028-y
[21] Ramadhani A, Khan F, Colbourne B, Ahmed S, Taleb-Berrouane M (2022) Resilience assessment of offshore structures subjected to ice load considering complex dependencies. Reliab Eng Syst Saf 222:108421
[22] Pranowo W, Ramadhani A (2022) Geostatistics for Risk Level Mapping: Synthetic Data Example. Proceedings of the 1st International Conference on Contemporary Risk Studies, ICONIC-RS 2022, 31 March-1 April 2022, South Jakarta, DKI Jakarta, Indonesia. https://doi.org/10.4108/eai.31-3-2022.2320971
[23] Pranowo W, Ramadhani AR (2023) A Comparison of Python-Based Copula Parameter Estimation for Archimedean-Based Asymmetric Copulas. SN Comput Sci 4:207
[24] Nelsen RB (2006) An introduction to copulas, 2nd ed. Springer
[25] Zhang Y, Kim C-W, Beer M, Dai H, Soares CG (2018) Modeling multivariate ocean data using asymmetric copulas. Coastal Engineering 135:91–111
[26] Genest C, Rémillard B, Beaudoin D (2009) Goodness-of-fit tests for copulas: A review and a power study. Insur Math Econ 44:199–213
[27] Hashemi SJ, Ahmed S, Khan FI (2015) Correlation

and Dependency in Multivariate Process Risk Assessment. IFAC-PapersOnLine 48:1339–1344

[28] Lü T-J, Tang X-S, Li D-Q, Qi X-H (2020) Modeling multivariate distribution of multiple soil parameters using vine copula model. Comput Geotech 118:103340

[29] Li D, Tang X (2014) Modeling and simulation of bivariate distribution of shear strength parameters using copulas. In: Risk and Reliability in Geotechnical Engineering. CRC Press, London, pp 77–128

[30] ERDDAP (2022) Mouth of Placentia Bay Buoy. https://www.smartatlantic.ca/erddap/tabledap/SMA_MouthofPlacentiaBayBuoy.html. Accessed 28 Aug 2021

[31] Zhang Y, Gomes AT, Beer M, Neumann I, Nackenhorst U, Kim C-W (2019) Modeling asymmetric dependences among multivariate soil data for the geotechnical analysis – The asymmetric copula approach. Soils and Foundations 59:1960–1979

[32] Amini A, Abdollahi A, Hariri-Ardebili MA, Lall U (2021) Copula-based reliability and sensitivity analysis of aging dams: Adaptive Kriging and polynomial chaos Kriging methods. Appl Soft Comput 109:107524