Modeling asymmetric dependences among multivariate soil data for the geotechnical analysis - the asymmetric copula approach

Yi Zhang[a,d], António Topa Gomes[b,c], Michael Beer[c], Ingo Neumann[d], Udo Nackenhorst[e], Chul-Woo Kim[f]*

[a] Department of Civil Engineering, Tsinghua University, Beijing, China
[b] Construct, Faculty of Engineering (FEUP), University of Porto, Portugal
[c] Institute for Risk and Reliability, Leibniz Universität Hannover, Germany
[d] Geodetic Institute, Leibniz Universität Hannover, Germany
[e] Institute of Mechanics and Computational Mechanics, Leibniz Universität Hannover, Germany
[f] Department of Civil and Earth Resources Engineering, Graduate School of Engineering, Kyoto University, Japan

Abstract

Multivariate information of soil parameters is quite important for the design and risk assessment of geotechnical engineering problems. It is necessary to have an accurate and realistic statistical multivariate model for representing the soil properties and thus evaluating the soil conditions. Thus, advanced multivariate modeling of soil parameters could help to improve the geotechnical engineering practice. In this paper, the asymmetric copulas are introduced to model the geotechnical soil data. Compared to extensive previous research on the use of symmetric copulas on the modeling of engineering data, this study is focusing on capturing asymmetric dependencies among the natural soil parameters, which are critical for engineering design. A copula-based multivariate probabilistic model is built based on a set of collected samples from a granite residual soil from Portugal. Several asymmetric copula functions, capable of capturing nonlinear asymmetric dependence structures, are tested and analyzed. The fundamental information on tail dependencies and measures of asymmetric dependencies are also exploited. To demonstrate the advantages of asymmetric copulas, its concept is compared with the traditional copula approaches for modeling site soil data. The performance of these asymmetric copulas is discussed and compared based on data fitting and extreme value characterizations.

Keywords: geotechnical analysis, asymmetric copula, soil properties, joint distribution, multivariate analysis

1. Introduction

Geotechnical engineering problems involve frequently multivariate data analysis. To consider multiple variables in a geotechnical design, a multivariate probabilistic model is usually required. This enables an application of well-developed joint statistical models to represent and, eventually, to evaluate uncertain results of the problem due to geotechnical random parameters. In this context, the dependencies among various soil parameters play an important role. Deficiencies in modeling their joint relationship may largely contribute to wrongly estimate the failure probability of geotechnical structures, hence may lead to expensive engineering loss (Angeli et al. 2000; Harris et al. 2008).

In real practice, the soil parameters are often observed to be dependent. For instance, the test results for the soil such as standard penetration test (SPT) and piezocone test (CPTU) tend to be physically related. However, the question is about how to define this relationship between the soil data. The definition of "dependencies" in this context can have various meanings. When addressing different dependencies for the soil parameters, the typical concept of correlation is commonly used to construct the joint distribution models. The applicability of this concept may be problematic when the dependencies are not perfectly linear. Many former works have addressed this issue (Vanapalli et al., 1996; Robertson, 2009; L'Heureux & Long, 2017). Still, many multivariate models have been developed adopting this concept (Yan et al., 2009; Sideri et al., 2014; Zhu et al., 2017).

It should be noted that in most cases, in geotechnical engineering practice, the joint cumulative distribution function (CDF) or joint probability density function (PDF) is often unknown due to limited data

* Corresponding author. Tel.: +81-075-383-3421
E-mail address: zhang_yi87@163.com; (Y. Zhang).

from field tests, laboratory tests or other resources (Beer et al, 2013; Li et al., 2012). Nevertheless, in recent years several works were published with presentation of multivariate information (Santoso et al., 2013; Zhang et al., 2018; Tang & Phoon, 2018). The most popular studies are related to clay parameters (Phoon & Kulhawy, 1999) or regarding the Mohr Coulomb failure envelope, and the negative correlation between cohesion, *c'*, and the friction angle, ϕ' (Phoon & Kulhavy, 1999; Duncan, 2000; Forrest & Orr, 2010; Tang et al., 2013; Zhang et al., 2018). Although Tang et al. (2013) and Li et al. (2015) investigated the influence of different copulas on the probability of failure of some simple geotechnical structures, examples applied to real data continue to be relatively scarce. From a geotechnical point of view, the topic attracts more attention is to achieve consistency between geotechnical and structural-based design (Phoon et al., 2016).

In contrast to the traditional joint model, the copula model has shown its advantage and attracted significant attention from many geotechnical engineering researchers (Wu, 2013, Tang et al., 2015). The key feature of a copula approach is its flexibility in modeling the dependence structure, which can be separated from the modeling of individual behavior. Such prominent characteristic is highly desirable in geotechnical engineering as most soil data exhibit non-obvious dependencies. Moreover, it was also found by utilizing the copula model, that the accuracy of reliability analysis of a geotechnical engineering problem can be largely improved (Li et al., 2015). In general, from the recent advances in geotechnical engineering, it is now widely recognized that the copula model is a very accurate and efficient tool in modeling multivariate soil data. However, there are various types of complicated dependencies and potential biases that could affect the quality of a multivariate model. Specifically, the uncertainties related to asymmetric dependencies are one of the most influencing factors. It was realized that an accurate modeling of the asymmetric dependences for soil data is still one of the most difficult tasks, and the statistical modeling of the multivariate soil data remains quite challenging. Fortunately, asymmetric copulas which were developed only recently provide a feasible solution to this problem (Kazianka & Pilz, 2010). The use of asymmetric copulas can significantly improve the functionality of traditional copula approaches in fitting the asymmetrically dependent variables. Nevertheless, the modeling of soil data using the asymmetric copula has never been studied in detail. The theoretical concepts and procedures of how to construct a reliable asymmetric copula for soil data have not yet been investigated. Therefore, this work aims to close this gap providing a real case study for demonstrating and highlighting the merits, as well as limitations, regarding the use of asymmetric copulas.

This paper is divided into seven sections. A general literature review of the existing techniques and former works on the modeling of multivariate soil data is presented in Section 2. Section 3 then reviews the fundamental copula theory and highlights the issues of basic dependence measures. Section 4 explains the detailed information of asymmetry measures as well as the procedures of constructing asymmetric copula models. A set of soil data is then analyzed through the use of asymmetric copulas. Section 5 provides the detailed information of the collected soil data. A comparative study between symmetric and asymmetric copula approaches for modeling the collected soil data is presented in Section 6. This includes the discussion on the quality of model fitting, tail dependence characterization and extreme value prediction. The final concluding remarks are summarized in Section 7.

2. Literature review of multivariate distributions for soil parameters

The variability of soil parameters is admittedly higher than for the remaining construction materials. Additionally, it presents local characteristics, creating obstacles to the generalization of results. In any case, since the 90s, efforts have been done to estimate the variability of design soil parameters, in order to develop a sound Reliability-Based Design (Duncan, 2000; Baecher & Christian, 2003; Forrest & Orr, 2010). Initially, the characterization of the variability of the parameters was completed through their coefficient of variation and the determination of the correlation between parameters was mainly a process to transform the test measurements in design parameters.

Ching & Phoon (2014) presented an example of multivariate distribution, applied to some clay parameters, that, as the correlation coefficient, may be applicable to site-specific data and used as a prior model that may be updated via, for example, Bayesian updating. As an example of this, the work of Zhang et al. (2018) is a worthy illustration. With the use of the multivariate distribution, the entire probability distribution of a design parameter may be updated covering all data, which represents an obvious advantage compared with the popular pairwise regression, where updates of the design parameter result from a single value of another parameter.

The copula theory (Nelsen, 2006) has found widespread applications in the last years and there are also recent examples of its application to geotechnical problems, as is the case of the pioneering works of Li et al. (2012) and Tang et al. (2013). Tang et al. (2013) studied the application of several types of copulas to the cohesion and friction angle data from four different sites. Zhang et al. (2014) clearly stated that previous probability models used in geotechnical engineering, such as multivariate normal distribution, is indeed based on the Gaussian copula, which can only consider the linear dependence relationship between random variables and may not always be optimal. Therefore, it is important to consider other copula functions for constructing probability models in geotechnical reliability analysis. The copula theory provides thus an advanced tool to model geotechnical problems more realistically (Tang et al., 2013; Li et al., 2015; Zhang & Lam, 2016). Particularly when using the Mohr-Coulomb failure criteria for soils, described by the two parameters, cohesion, *c'*, and friction angle, ϕ'. It is widely accepted that there exists a negative correlation between them, which results from the linearization of the failure envelope. Tang et al. (2013) presented a list of correlation coefficients between these two parameters found by several authors, but also stated that the Gaussian copula is commonly adopted without rigorous validation. There are also recent tentative to adjust non Gaussian dependence, though not abundant (Wang & Li, 2017).

Residual soils are cemented materials but have low cohesion values. Having in mind that the cohesion is always positive, this can create an asymmetry in the distribution, and thus asymmetric copulas might arise as an interesting solution to cope with real data. Additionally, the fact real data is used to test several copula constitutes an enormous advantage to evaluate the advantages of using asymmetric copula.

3. Copula theory and dependence measures

As mentioned in the previous section, copula models provide an alternative way to model the multivariate soil data. The concept of copula theory has already been used for modeling a wide range of engineering data, for example, in reliability studies (see, Noh et al., 2009; Wang et al., 2017), as well as offshore engineering (Zhang et al., 2015; 2018). Several former works have provided a thorough survey: for the theoretical background see Nelsen (2006), and Joe (2014); for the practical applications see Genest and Favre (2007), Salvadori and De Michele (2007), and Hong et al. (2015).

3.1 Definition and basic properties

The theoretical definition of a copula can be specified by the marginal distributions as introduced in Sklar's theorem (Sklar, 1959):

Sklar's Theorem: Let F be an n-dimensional distribution function with marginal distributions $F_1, \ldots, F_n$. A copula C is therefore defined as an n-dimensional distribution function such that for all $x \in R^n$

$$F\left(x_1,\ldots,x_n\right)=C\left(F_1\left(x_1\right),\ldots,F_n\left(x_n\right)\right) \quad (1)$$

If $F_1, \ldots, F_n$ are all continuous, then C should be unique. Conversely, if C is a copula and $F_1, \ldots, F_n$ are all continuous marginal distribution functions, then the distribution function F must be a multivariate distribution function with marginal distributions $F_1, \ldots, F_n$.

Compared to the other joint distribution models, the copula approach has the freedom of selecting any marginal distributions for the variables which makes this approach much more flexible in characterizing individual variable's behaviors. Many existing copula functions have been formulated in the literature, see e.g. (Hutchinson and Lai 1990; Trivedi & Zimmer, 2007). Each specific copula could characterize a certain kind of dependence in the multivariate data.

3.2 Dependence measures

In order to emphasize the significance of the copula approach in modeling geotechnical data, the dependence concepts are interpreted with details herein. It is said the key characteristic of a copula model is its dependence structure. Traditionally, the Pearson's correlation coefficient ρ is used as the most common and convenient way for measuring the data dependence. Because of its ease of handling, it is widely adopted in many statistical approaches. However, the weakness of ρ is also obvious and many researchers tend to criticize it. For instance, it is realized the linear correlation coefficient is invariant with respect to linear transformations of the variables. But it is not invariant to strictly increasing nonlinear transformations. The property of linear dependency may not be preserved through such transformations. Therefore, based on these concerns, other concepts of dependencies have been developed in the literature such as Kendall's τ_k and Spearman's ρ_s. Kendall's τ_k is a measure of the possible excess of concordance/discordance in the sample, and Spearman's ρ_s measures the "distance" between the chosen copula and the one modeling independent variables (see Salvadori et al. ,2007). These two measures are also known as the most well-established concordant measures of rankings among the variables. The concepts of Kendall's τ_k and Spearman's ρ_s are well integrated in a copula model. For example, for any bivariate copula, these two coefficients can be directly linked to the copula function as

$$\tau_k\left(u_1,u_2\right)=4\int_0^1\int_0^1 C\left(u_1,u_2\right)dC \qquad (2)$$

$$\rho_s\left(u_1,u_2\right)=12\int_0^1\int_0^1 C\left(u_1,u_2\right)dC-3 \qquad (3)$$

where $u_i=F_i(X_i)$. This linkage provides a feature in copula model that can describe various kinds of dependencies, including association concepts such as concordance, linear correlation and other related measures.

However, the traditional copulas have many weaknesses (e.g. Archimedean copulas) when they are applied to model soil parameters. A key drawback is that most well established copulas can only model symmetric dependent variables whereas the soil data usually display non-symmetric dependencies. For example, the feasible domain of soil parameters restricted by the physical phenomenon is a major reason for asymmetric dependencies. For instance, a large value of soil cohesion strength is unlikely to be accompanied by a large value of friction angle because of the physical limit. Negative values for cohesion are not physically possible. In other words, the realization of some variable combinations should not exist in the real nature. This effect can be illustrated by means of an example scatter plot as shown in Fig. **1**. As demonstrated in the figure, it is impossible to have observations in the right-lower region (marked with a cross), while observations can be available in the left-upper region (marked with a tick). In other words, implicit physical phenomena could exert limit of occurrence for some data combinations. Thus, the feasible domain reduces and becomes asymmetric. More typical examples can be illustrated by Fig. 2 which show the scatter plot of soil data from the database provided by TC304 webpage. The dependences among the chosen soil parameters undrained shear strength s_u, preconsolidation stress σ'_p and vertical effective stress σ'_v are not perfect linear. In fact, they are inherently dependent on the liquid limit and overconsolidation ratio which makes their dependences quite complex. From these scatter plots, it can be observed that no data is distributed in the upper-lower domain (as marked by the red star symbol). This generally means the considered bivariate dataset has a restricted domain which can only allow data to be distributed asymmetrically. Therefore, considering this physical feature in the multivariate soil data modeling, especially copula approach, is not straightforward and still needs further development.

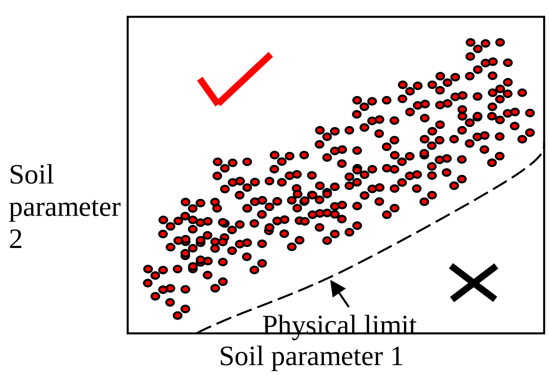


Figure 1 Asymmetric domain of soil data caused by physical phenomenon.

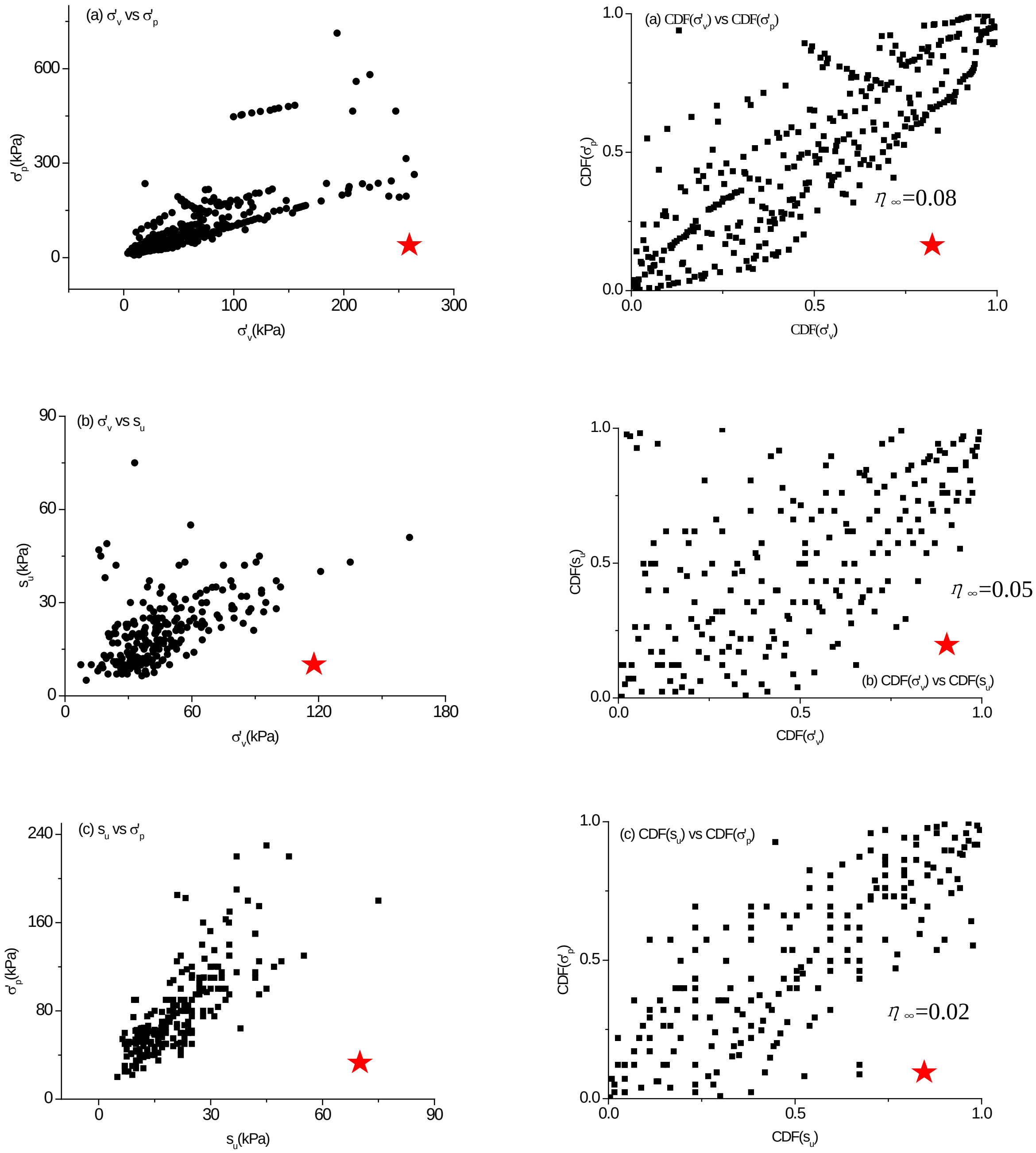


Figure 2 Examples of soil data having asymmetric domain (data retrieved from Ching and Phoon, (2012), Ching et al. (2014), D'Ignazio et al. (2016) and Zhang et al. (2019))

However, these effects can be frequently observed in most collected soil datasets. The ignorance of such asymmetric dependencies in the multivariate modeling might create some unreliable estimates for the design. More advanced statistical techniques are therefore required on the improvement of traditional copula model to further enhance this approach.

4. Asymmetric copulas

In order to have a more accurate modeling of asymmetrically dependent variables, several groups of asymmetric copulas as well as the basic concepts in measuring the asymmetry of a copula model are introduced herein.

4.1 Measure of asymmetry and tail dependency

The fundamental definition of symmetry in a copula model can be defined as following. For a given copula $C(u_1, \dots, u_n)$, if

$$C\left(u_1,\dots,u_i,\dots,u_j,\dots,u_n\right) = C(u_1,\dots,u_j,\dots,u_i,\dots,u_n) \text{ is true for any pair } u_i, u_j \in \mathbf{I},$$

then we can say u_i and u_j are exchangeable within the copula $C(u_1, \dots, u_n)$ and this copula is said to be symmetric (Genest and Nešlehová, 2013). Therefore, if this copula function cannot satisfy the above condition, it is believed to be asymmetric. Following this idea, a measure of asymmetry in a copula model can be formulated by the following equation (Klement and Mesiar, 2006)

$$\eta_p\left(C\right) = \left\{\int_0^1\int_0^1 \left|C\left(u_1,u_2\right) - C\left(u_2,u_1\right)\right|^p du_1 du_2\right\}^{1/p} \tag{4}$$

where p is a factor which can be set at any value greater than or equal to 1, $p \geq 1$. In other words, the function calculates the distance between C and its transpose C^T, like the norm. Usually, it is more convenient to set the value of p to infinity for calculating the measure of asymmetry. This gives a simplified formula as

$$\eta_\infty\left(C\right) = \sup_{\left(u_1,u_2\right)\in[0,1]^2} \left|C\left(u_1,u_2\right) - C\left(u_2,u_1\right)\right| \tag{5}$$

Therefore, if the value of this measure is too large, the copula is considered to be asymmetric. Meanwhile, when it is applied to bivariate data, the measure of asymmetry as calculated by Eq. (5) has the same meaning of a measure of exchangeability for the data.

Another indicator that can be used to detect the asymmetric characteristics is the tail dependencies. Based on the concept of tail dependence, four coefficients are defined to describe the tail dependences, namely, *lower-lower*, *lower-upper*, *upper-lower*, *upper-upper* tail dependence coefficients. For example, for a bivariate copula $C(u_1, u_2)$, the tail dependence coefficients can be calculated by (Nelsen 2006)

$$\lambda_{1|2}^{l,l}\left(C\right) = \lim_{u\to 0+} P\left(x_1 \leq F_1^{-1}\left(u\right) \mid x_2 \leq F_2^{-1}\left(u\right)\right) = \lim_{u\to 0+} \frac{C\left(u,u\right)}{u} \tag{6}$$

$$\lambda_{1|2}^{l,u}\left(C\right) = \lim_{u\to 0+} P\left(x_1 \geq F_1^{-1}\left(1-u\right) \mid x_2 \leq F_2^{-1}\left(u\right)\right) = 1 - \lim_{u\to 0+} \frac{C\left(u,1-u\right)}{u} \tag{7}$$

$$\lambda_{1|2}^{u,l}\left(C\right) = \lim_{u\to 0+} P\left(x_1 \leq F_1^{-1}\left(u\right) \mid x_2 \geq F_2^{-1}\left(1-u\right)\right) = 1 - \lim_{u\to 0+} \frac{C\left(1-u,u\right)}{u} \tag{8}$$

$$\lambda_{1|2}^{u,u}\left(C\right) = \lim_{u\to 0+} P\left(x_1 \geq F_1^{-1}\left(1-u\right) \mid x_2 \geq F_2^{-1}\left(1-u\right)\right) = 2 - \lim_{u\to 0+} \frac{1-C\left(1-u,1-u\right)}{u} \tag{9}$$

where $F_1^{-1}(.)$ and $F_2^{-1}(.)$ are the inverse marginal distribution functions for x_1 and x_2. Therefore, these equations provide measures of the tail dependence for the two variables in four different extremes. The tail coefficients have a value range between 0 and 1, where a value of 0 indicates asymptotical independence.

Tail dependencies can provide useful information about the dependences of extreme values from the intrinsic information. It gives a measure for relating one margin exceeding a certain quantile threshold while the other has already exceeded that quantile threshold. The *lower-upper* and *upper-lower* tail coefficients are especially useful for assessing the asymmetry of a copula. If these coefficients are observed to be different, the copula is generally an asymmetric one.

4.2 Asymmetric copulas constructed by products

There are various ways of constructing asymmetric copulas. Many recent works have been done in this direction (Grimaldi and Serinaldi, 2006; Mesiar and Najjari, 2014; Mazo et al., 2015). Plenty of techniques able to capture the asymmetric dependencies in the multivariate data are utilized in the copula function establishment (Patton, 2006). Nevertheless, not all the asymmetric copulas are really useful in practice. Some asymmetric copulas may need very sophisticated extra functions to characterize the asymmetric dependencies which are quite cumbersome for the calculation. A typical example could be the Archimax copula which requires complex statistical derivations for obtaining the Pickhands dependence function for its construction (Charpentier et al. 2014). Therefore, from the engineering point of view, we choose to review the most popular and practical alternatives among these asymmetric copulas in this study. Meanwhile, this work tends to focus on the asymmetric copula families that can be built based on the traditional symmetric copulas, e.g. Archimedean copulas. Therefore, the asymmetric copulas with a very complicated mathematical formulation would not be discussed in the present study.

One of the most popular ways of constructing asymmetric copulas is by means of a product of copulas (Liebscher, 2008). The general form for constructing this type of asymmetric copula is given as following

$$C_{product}\left(u_1,...,u_n\right)=\prod_{i=1}^{m}C_i\left(f_{i1}\left(u_1\right),...,f_{in}\left(u_n\right)\right), \quad (10)$$

where $C_1,\ldots,C_m$ are all copulas for the *n*-dimensional variables, f_{ij}: $[0,1] \rightarrow [0,1]$ for i=1,…,m, j=1,…,n are the individual functions for describing the individual variable's behavior which should be strictly increasing or identically equal to 1. To guarantee Eq. (10) is also a copula, the individual functions f_{ij} must satisfy the following additional properties:

1. $f_{ij}(1)=1$ and $f_{ij}(0)=0$,
2. f_{ij} is continuous on]0,1],
3. If there are at least two functions f_{i_1j}, f_{i_2j} with $1 \leq i_1, i_2 \leq m$ which are not identical and equal to 1, then $f_{ij}(x) > x$ holds for $x \in (0,1)$, i=1,…,m.

From the above formulation, it is easy to see the constructed copula could be asymmetric if the individual functions are different for the variables. Each individual functions f_{ij} characterizes a specific property of the variables in the asymmetric dependence modeling. The idea of this construction is also known as an extension of Khoudraji's device (1995). For instance, by adopting type I individual function in constructing the asymmetric copula (see Table 1) and setting m, n=2, Eq. (10) becomes exactly the Khoudraji copula. On the other hand, various groups of parametric copulas can be selected for the *n*-dimensional copulas $C_1,\ldots,C_m$, e.g. Archimedean copulas. As for the individual functions f_{ij}, many candidate functions which are suitable for the copula construction have been proposed by Liebscher (2008) - see Table 1. Moreover, it is also possible to choose the number and type of individual copulas.

Table 1 Examples of individual functions

Individual function	Parameters	Value range
I. $f_{ij}\left(u\right)=u^{\theta_{ij}}$	$\sum_{i=1}^{m}\theta_{ij}=1$	$\theta_{ij}\in\left[0,1\right]$
II. $f_{ij}\left(u\right)=u^{\theta_{ij}}e^{(u-1)\alpha_{ij}}$	$\sum_{i=1}^{m}\theta_{ij}=1$, $\sum_{i=1}^{m}\alpha_{ij}=0$	$\theta_{ij}\in\left(0,1\right)$, $\alpha_{ij}\in\left(-\infty,1\right)$, $\theta_{ij}+\alpha_{ij}\geq 0$
III. * $f_{1j}\left(u\right)=\exp\left(\theta_j-\sqrt{\left\lvert\ln u\right\rvert+\theta_j^2}\right)$, $f_{2j}\left(u\right)=u\exp(-\theta_j+\sqrt{\left\lvert\ln u\right\rvert+\theta_j^2})$	θ_j *for* $j\in\{1,...,n\}$	$\theta_j\geq\frac{1}{2}$

*Note: type III individual functions can only be used for the asymmetric copula having two individual copulas (e.g. m=2).

4.3 Asymmetric copulas constructed by linear convex combinations

Another way of constructing an asymmetric copula could be done through the linear convex combinations of copulas. However, it should be noted the direct linear convex combination of copulas is not able to create asymmetric copulas. The main reason is most fundamental copulas are symmetric. Such linear convex combination of these copulas could not change their dependence characteristics and would also only produce symmetric copulas. One way to change the symmetric dependence characteristics is to modify the fundamental copulas to account for asymmetric properties (Wu, 2014). A change on the new kind of copula is proposed as:

$$\breve{C}_h\left(u_1,\ldots,u_n\right)=C\left(u_1,\ldots,u_{h-1},1,u_{h+1},\ldots,u_n\right)-C\left(u_1,\ldots,u_{h-1},1-u_h,u_{h+1},\ldots,u_n\right) \quad (11)$$

where $C(.)$ is the original n-dimensional base copula. It is easy to see that any variable u_h in the copula model is not exchangeable with other variables. Such developed model is also called flipped copula as mentioned in the literature (Nelsen 2003). Therefore, the flipped copula can be used to fit data exhibiting unequal tail dependencies. By combining all the possible flipped copulas, one may use the following copula to model asymmetric properties in multiple variables:

$$C_{addition}\left(u_1,\ldots,u_n\right)=\sum_{h=0}^{n} p_h \breve{C}_h\left(u_1,\ldots,u_n\right) \quad (12)$$

where p_h is a weighting factor which needs to satisfy the conditions $0 \le p_h \le 1$ and $\sum_{h=0}^{n} p_h = 1$. And when h=0, the flipped copula downgraded to the original one, e.g. $\breve{C}_0(u_1,\ldots,u_n) = C(u_1,\ldots,u_n)$. Same as the copula in Section 4.2, various types of copula families can be utilized as the base copula $C(u_1,\ldots,u_n)$. When it is applied for the bivariate data, Eq. (12) can be expressed as following

$$\breve{C}_1\left(u_1,u_2\right)=u_2-C\left(1-u_1,u_2\right), \quad (13)$$

$$\breve{C}_2\left(u_1,u_2\right)=u_1-C\left(u_1,1-u_2\right), \quad (14)$$

where we can also call Eq. (13) and Eq. (14) the horizontal-flipped and vertical-flipped copulas (Salvadori et al. 2007). A typical bivariate asymmetric copula in this case can be given as

$$C_{addition}\left(u_1,u_2\right)=p_0 C\left(u_1,u_2\right)+p_1\breve{C}_1\left(u_1,u_2\right)+p_2\breve{C}_2\left(u_1,u_2\right) \quad (15)$$

where $p_0, p_1, p_2 \ge 0$ and $p_0 + p_1 + p_2 = 1$. The asymmetric properties of the bivariate data can be simply modeled by adjusting the values of weight factors assigned to each base copula in this formula. That is, the flipped copula $\breve{C}_1(u_1,u_2)$ or $\breve{C}_2(u_1,u_2)$ are used to model the asymmetry in each of the variables. This is also the main difference between the current construction method and Liebscher's method. The current method constructs asymmetric copulas by modeling the asymmetric property for variables each at a time. However, on the other hand, Liebscher's method constructs the asymmetric copulas for variables all at a time.

4.4 Skewed copula

Despite the algebraic construction methods, another convenient way of constructing asymmetric copulas is by means of the skewed copula. The idea of this approach is from the skewed multivariate Guassian distribution which allows non-zero skewness. The general concept is to transform a multivariate Gaussian distribution to an asymmetric one by introducing a parameter (Kollo et al., 2013). The most famous and commonly adopted one is the *skewed Gaussian copula*.

The skewed Gaussian copula originates from the the Gaussian copula. By definition, an n-dimensional Gaussian copula is expressed by

$$C_{Gaussian}\left(u_1,\ldots,u_n\right)=\Phi_n\left(\Phi^{-1}\left(u_1\right),\ldots,\Phi^{-1}\left(u_n\right);\boldsymbol{\Sigma}\right) \quad (16)$$

Where $\Phi_n(.)$ represents the n-dimensional normal distribution function, $\Phi^{-1}(.)$ denotes the inverse of the standard normal distribution function, and $\sum$ stands for the covariance matrix. In the skewed Gaussian copula, the basic formula is modified to account for asymmetries by adding the shape parameter. A general n-dimensional skewed Gaussian copula can be written as

$$C_{skew-Gaussian}\left(u_1,\ldots,u_n;\mu,\Sigma,\beta\right)=F_{n,skew}\left(F_{1,skew}^{-1}\left(u_1;\mu_1,1,\beta_1\right),\ldots,F_{1,skew}^{-1}\left(u_n;\mu_n,1,\beta_n\right);\mu,\Sigma,\boldsymbol{\beta}\right) \quad (17)$$

where $F_{n,skew}(.)$ is the n-dimensional skew normal distribution with mean parameter μ**,** $F_{1,skew}^{-1}(.)$ is the inverse of the univariare skew normal distribution $SN(\mu_i, 1, \beta_i)$, $\boldsymbol{\beta}$ are the shape parameters and $\sum$ is the covariance matrix. Therefore, the density function of a multivariate skewed Gaussian copula for n-dimensional random variables can be given by

$$f_n\left(u_1,\ldots,u_n;\mu,\Sigma,\beta\right)=2\phi_n\left(u_1,\ldots,u_n;\mu,\Sigma\right)\Phi_n\left(\boldsymbol{\beta}^T u_1,\ldots,u_n;\mu,\Sigma\right) \quad (18)$$

where $\phi_n(.)$ and $\Phi_n(.)$ are the probability density function and cumulative distribution function for n-dimensional Gaussian distribution (Azzalini and Valle, 1996). In this constructed asymmetric copula, the asymmetric property results from the shape parameters. For example, when $\boldsymbol{\beta}$=0, the skewed Gaussian copula downgrades to the standard Gaussian copula with no skewness. If $\boldsymbol{\beta}$ increases, the skewness of the skewed Gaussian copula increases.

Moreover, it should be pointed out the skewed Gaussian copula is in fact a special case of the constructed copulas as given in Section 4.2. Compared to the copula constructed by Eq. (10), the skewed Gaussian copula is a special one with only one individual copula (m=1). This base copula (C_i) are all skewed Gaussian distributions. Nevertheless, it is still worth to see the performance of skewed copulas compared to the other approaches. There are no previous works done on its application in the modeling of real collected soil data. The following will provide a case study to demonstrate the key advantages of using the asymmetric copulas in modeling soil data.

5. Case Study – Site Soil Data

The soil data used in this paper results from tests performed in a residual soil from Porto granite. Pinheiro Branco (2011) and Pinheiro Branco et al. (2014) conducted an extensive characterization of a localized area of residual soil, collecting more than 40 samples in an area of approximately 1 m^2. Detail of the area where the samples were collected is shown in Fig. 3.

Figure **3** Detail of the area where the samples were collected

All the samples were carefully collected in situ, by cutting the residual soil around the sampler (0.1×0.1×0.03 m^3), isolated and transported to the geotechnical laboratory. For all the specimen, the dry unit weight (γ_d), the water content (w), the void ratio (e) and subsequently the saturated unit weight (γ_{sat}) were all measured (Pinheiro Branco et al., 2014). For three representative samples the granulometric curve of the material (Fig. 4), as well as the unit weight of the soil particles (γ_s), were also determined.

All the samples were subjected to direct shear tests, with different normal stresses: 25 kPa, 50 kPa, 75 kPa, and 100 kPa. The normal stresses were intentionally low, in order to avoid particle breaking or sample disturbance during the installation of the initial stress. The *in situ* vertical stress where the samples were located was approximately 120 kPa. In such conditions all the tests were performed with normal stresses lower than the *in situ* vertical stress. The consolidation time was established as 1 hour. After several minutes there were no additional vertical settlements which allowed to conclude that there was no further consolidation. The shear rate of the tests was 0.03 mm/min. This reduced shearing rate guarantees no excess water pressures appear during shear, corresponding to drained conditions.

The 40 samples were divided into 10 samples for each stress level. During each shear test, the peak shear stress τ_p, the residual shear stress τ_r, and the dilation angle ψ were measured. The residual strength was simply defined by the constant volume friction angle, ϕ'_{cv}. The peak strength was defined by a unique friction angle, ϕ'_s, although its value is dependent on the normal stress of the test. Table 2 presents the complete list of variables measured or calculated for the 40 samples, during the direct shear tests.

The parameters presented in Table 2 correspond to each individual sample. In geotechnical practice, the peak strength is usually defined as the Mohr-Coulomb failure criteria, namely the cohesion, c'_p, and the peak friction angle, ϕ'_p. To determine these parameters, soil data samples have to be grouped and utilized to estimate them from Mohr-Coulomb circle. With this purpose, the 40 samples were combined in groups of 3, resulting the 40 values of the Mohr-Coulomb parameters presented in Table 3.

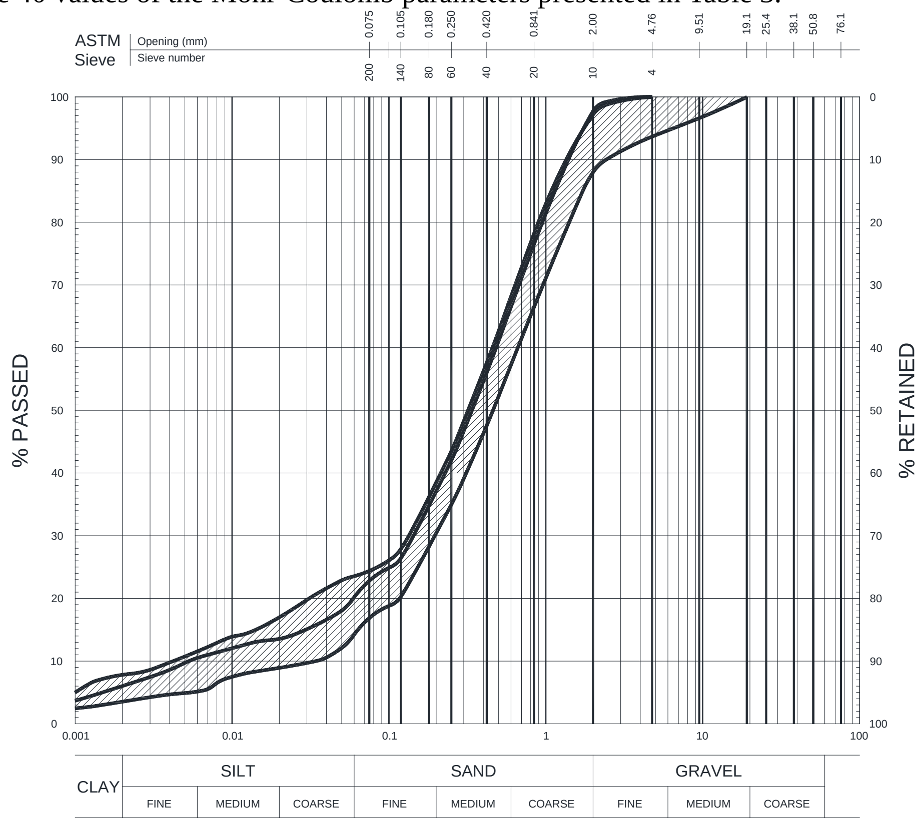


Figure **4** Granulometric curves of the tested soil samples

Table 2 Collected soil property data from the site

σ'(kPa)	**$\phi'_{cv}(^0)$**	**$\phi'_s(^0)$**	**e**	**γ (kN/m^3)**	**γ_d (kN/m$^{3)}$**	**$\psi(^0)$**
25.0	39.35	53.45	0.578	19.19	16.37	13.55
25.0	41.96	49.00	0.574	19.41	16.41	8.72
25.0	36.42	50.46	0.573	19.44	16.42	14.68
25.0	34.78	47.53	0.558	19.52	16.58	16.20
25.0	41.19	41.80	0.640	18.29	15.75	3.03
25.0	41.11	47.18	0.568	19.03	16.47	13.85
25.0	35.87	50.15	0.453	20.27	17.78	14.28
25.0	39.83	47.32	0.551	19.10	16.66	14.93
25.0	40.46	40.46	0.694	17.23	15.25	7.88
25.0	38.70	53.61	0.525	19.19	16.94	19.20
50.0	39.30	46.61	0.717	17.72	15.05	6.80
50.0	38.36	47.83	0.574	19.29	16.42	10.95
50.0	38.10	50.98	0.577	19.27	16.39	11.01
50.0	39.14	52.34	0.530	19.20	16.88	14.15
50.0	40.37	50.67	0.589	18.87	16.26	8.82
50.0	38.61	47.24	0.543	19.23	16.74	7.77
50.0	37.78	48.51	0.489	19.87	17.35	13.12

50.0	37.43	44.88	0.530	19.07	16.88	12.33
50.0	35.96	41.02	0.649	17.78	15.66	4.43
50.0	38.20	41.96	0.589	18.35	16.26	8.13
75.0	37.20	45.51	0.571	19.40	16.45	9.32
75.0	42.57	51.28	0.575	19.40	16.41	10.66
75.0	37.33	47.89	0.557	19.57	16.59	14.54
75.0	38.49	45.89	0.567	19.43	16.48	10.22
75.0	38.74	38.74	0.581	19.06	16.34	0.53
75.0	38.40	40.77	0.609	18.50	16.05	3.76
75.0	38.12	47.08	0.499	19.58	17.23	6.76
75.0	37.86	45.22	0.625	18.02	15.90	7.23
75.0	40.03	48.67	0.517	19.30	17.02	11.24
75.0	37.45	39.74	0.663	17.46	15.53	2.79
100.0	36.33	44.57	0.611	18.96	16.03	8.23
100.0	37.59	40.42	0.576	19.20	16.40	4.41
100.0	33.22	40.16	0.581	19.02	16.33	8.74
100.0	38.30	43.45	0.599	19.11	16.15	3.14
100.0	35.44	39.97	0.588	18.78	16.27	4.57
100.0	37.95	46.16	0.549	19.22	16.68	8.89
100.0	33.46	40.02	0.563	18.86	16.53	6.32
100.0	39.00	46.03	0.562	18.70	16.54	5.25
100.0	36.67	38.73	0.599	18.14	16.16	2.86
100.0	36.53	45.70	0.479	19.81	17.46	6.34

Table 3 Estimated friction angle and cohesion

c'_p (kPa)	$\tan(\phi_p')$	c'_p (kPa)	$\tan(\phi_p')$	c'_p (kPa)	$\tan(\phi_p')$	c'_p (kPa)	$\tan(\phi_p')$
11.68	0.85	14.61	0.76	10.89	0.75	1.22	1.01
10.91	0.87	55.00	0.32	1.96	1.02	30.38	0.68
12.04	0.86	6.44	1.00	0.00	1.19	0.00	1.19
36.69	0.58	12.34	0.85	34.14	0.60	2.98	0.77
0.00	1.19	5.56	0.91	5.00	1.01	5.23	0.85
13.79	0.73	53.75	0.37	16.23	0.69	33.74	0.53
13.84	0.82	10.03	0.75	14.04	0.76	18.65	0.56
47.85	0.45	10.95	0.76	48.22	0.38	5.16	0.94
5.62	1.05	1.40	0.81	2.36	0.96	8.60	0.86
18.93	0.68	51.12	0.25	0.17	1.02	22.79	0.70

6. Data Analysis

The total sample size of 40 soil data is selected for the analysis in this study. All of these data are obtained from the same site and therefore are believed to have the same statistical characteristics. To understand the statistical properties of the collected data, a general statistical summary of c'_p, $\tan(\phi'_p)$, $\tan(\phi'_s)$, $\tan(\phi'_{cv})$, e, γ, γ_d and ψ is provided in Table 4. It can be seen the variations in c'_p is much higher compared to other soil parameters. The mean and variations of the friction angle are generally small, particularly for $\tan(\phi'_{cv})$. However, the differences between $\tan(\phi'_p)$, $\tan(\phi'_s)$ and $\tan(\phi'_{cv})$ are very obvious. The statistical values of the unit weight and dry unit weight are quite close. Individual characteristics of the soil parameters c'_p, $\tan(\phi'_p)$, $\tan(\phi'_s)$, $\tan(\phi'_{cv})$, e, γ, γ_d and ψ have to be investigated separately.

Table **4** Statistical summary of the collected soil data

	Number of data	Mean	Standard deviation	Minimum	Maximum
c'_p(kPa)	40	16.35	16.38	0	54.99
$\tan(\phi'_p)$	40	0.78	0.23	0.25	1.19
$\tan(\phi'_s)$	40	1.03	0.15	0.80	1.35
$\tan(\phi'_{cv})$	40	0.78	0.05	0.65	0.91
e	40	0.57	0.05	0.45	0.71
γ (kN/m^3)	40	18.97	0.66	17.23	20.27
γ_d(kN/m^3)	40	16.42	0.54	15.04	17.78
ψ (°)	40	8.99	4.38	0.53	19.2

As an initial step in the copula statistical analysis, the marginal distribution functions are determined for all the soil parameters. For example, in order to make a fair comparison, we choose a group of parametric statistical models to fit the collected data. For this list, we include Weibull, Normal, Lognormal, Logistic, Extreme value, Exponential and Gamma models. To compare all the candidate models, the standard Akaike Information Criterion (AIC) is utilized herein as a reference. The calculation of AIC is generally given by

$$\mathrm{AIC} = -2l\left(p\right) + 2p \tag{19}$$

where p is the number of parameters used in each statistical model, and $l(p)$ is the maximized log-likelihood for that model. Generally speaking, the concept of AIC takes into account both the simplicity of the model and the goodness-of-fit. A smaller AIC value implies a better model.

Table 5 summarizes the calculated AIC values for each of the parametric models. From the results, the best models are Gamma for c'_p, Extreme Value for $\tan(\phi'_p)$, Lognormal for $\tan(\phi'_s)$, Normal for $\tan(\phi'_{cv})$, Lognormal for e, Weibull for γ, Normal for γ_d and Weibull for ψ. Based on the selected models, the statistical model parameters are estimated by the maximum likelihood method. The results of these parameter estimates, including the statistical errors are presented in Table 6. As indicated by the model parameters, e, γ_d and ψ are quite symmetric in the distribution density function, γ and c'_p have quite high skewness. The good thing is, in the copula model, all these parameters will be converted to their CDF values based on marginal distributions. Therefore, after the transformation, the individual parameters will all be uniformly distributed variables between 0 and 1. Thus, the individual behavior could be removed at this initial step before the copula modeling. The following would be mainly focusing on the dependence characterizations.

Table **5** Calculated AIC statistics for the marginal distribution model fitting (Chi square test p-value with significance level of 5% are provided in the bracket)

	Weibull	Normal	Lognormal	Logistic	Extreme value	Exponential	Gamma
c'_p(kPa)	299.8 (0.172)	340.2 (0.009)	329.2 (0.127)	339.1 (0.133)	356.3 (0.141)	303.5 (0.130)	295.2* (0.199)
$\tan(\phi'_p)$	-0.8646 (0.679)	0.0341 (0.556)	10.21 (0.060)	0.6868 (0.759)	-0.8672* (0.277)	63.01 (0.005)	5.404 (0.244)
$\tan(\phi'_s)$	-30.04 (0.3199)	-32.51 (0.298)	-33.18* (0.264)	-30.02 (0.301)	-28 (0.341)	86.92 (0.007)	-33.18 (0.276)
$\tan(\phi'_{cv})$	-135.4 (0.103)	-140.3* (0.371)	-139.7 (0.349)	-139.4 (0.513)	-133.1 (0.076)	63.08 (0.001)	-139.9 (0.371)
e	-151.2 (0.003)	-151.9 (0.036)	-152.5* (0.173)	-150.1 (0.155)	-149.8 (0.001)	39.64 (0.001)	-151.1 (0.043)
γ(kN/m^3)	51.36* (0.090)	52.12 (0.089)	55.62 (0.056)	54.92 (0.029)	53.74 (0.062)	319.4 (0.001)	54.4 (0.074)
γ_d(kN/m^3)	49.48 (0.003)	36.38* (0.082)	38.86 (0.080)	38.76 (0.278)	50.84 (0.001)	307.92 (0.001)	37.02 (0.091)
ψ (°)	188.7* (0.822)	189.9 (0.825)	201.5 (0.001)	190.5 (0.793)	189.1 (0.297)	259.6 (0.001)	196.2 (0.294)

*The lowest AIC indicates the best model.

Table **6** Estimated model parameters for the best marginal distribution model for each soil parameter (standard errors are provided in the bracket)

	c'_p(kPa)	tan(ϕ'_p)	tan(ϕ'_s)	tan(ϕ'_{cv})	e	γ (kN/m^3)	γ_d (kN/m^3)	ψ (°)
Parameter Estimates	a=0.5477 (0.0094) b=29.8625 (8.4622)	k=-0.4184 (0.181) σ=0.2442 (0.0031) μ=0.7228 (0.0425)	μ=0.0256 (0.0007) σ=0.1500 (0.0004)	μ=0.7854 (0.0006) σ=0.0576 (0.0046)	μ=-0.5563 (0.0083) σ=0.0619 (0.0061)	A=19.1829 (0.0666) B=48.1434 (5.7492)	μ=16.4277 (0.0581) σ=0.3673 (0.0419)	A=9.8905 (0.3834) B=4.2735 (0.5455)

In order to have a full understanding of the relationships among all the soil parameters, the dependence measure concepts including Kendall's tau, Spearman's rho and correlation coefficient are calculated for each of the dataset and recorded in Table 7. As can be seen from the table, the dependences between several pairs of data are quite strong, namely, (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d). For the other pairs of data, the dependence is not very strong. From a statistical point of view, if the dependence is very weak, a multivariate modeling is not very meaningful. Thus, the following study will be limited to the datasets (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d) for the asymmetric copula modeling.

Table **7** Summary of the dependences among collected soil data

Pearson Correlation

	c'_p(kPa)	tan(ϕ'_p)	tan(ϕ'_s)	tan(ϕ'_{cv})	e	γ(kN/m^3)	γ_d (kN/m^3)	ψ (°)
c'_p(kPa)	-	-0.91353	-	-	-	-	-	-
tan(ϕ'_p)	-0.91353	-	-	-	-	-	-	-
tan(ϕ'_s)	-	-	-	0.36488	-0.44835	0.54173	0.44988	0.78078
tan(ϕ'_{cv})	-	-	0.36488	-	0.15556	-0.09402	-0.15627	0.05931
e	-	-	-0.44835	0.15556	-	-0.87407	-0.99857	-0.58744
γ(kN/m^3)	-	-	0.54173	-0.09402	-0.87407	-	0.8677	0.61339
γ_d(kN/m^3)	-	-	0.44988	-0.15627	-0.99857	0.8677	-	0.57945
ψ (°)	-	-	0.78078	0.05931	-0.58744	0.61339	0.57945	-

Spearman's ρ_s

	c'_p(kPa)	tan(ϕ'_p)	tan(ϕ'_s)	tan(ϕ'_{cv})	e	γ(kN/m^3)	γ_d (kN/m^3)	ψ (°)
c'_p(kPa)	-	-0.9116	-	-	-	-	-	-
tan(ϕ'_p)	-0.9116	-	-	-	-	-	-	-
tan(ϕ'_s)	-	-	-	0.37317	-0.50544	0.59981	0.50544	0.78837
tan(ϕ'_{cv})	-	-	0.37317	-	0.08818	-0.11445	-0.08818	0.06323
e	-	-	-0.50544	0.08818	-	-0.86224	-0.99872	-0.58819
γ(kN/m^3)	-	-	0.59981	-0.11445	-0.86224	-	0.85366	0.6081
γ_d(kN/m^3)	-	-	0.50544	-0.08818	-0.99872	0.85366	-	0.58037
ψ (°)	-	-	0.78837	0.06323	-0.58819	0.6081	0.58037	-

Kendall's τ

	c'_p(kPa)	tan(ϕ'_p)	tan(ϕ'_s)	tan(ϕ'_{cv})	e	γ(kN/m^3)	γ_d (kN/m^3)	ψ (°)
c'_p(kPa)	-	-0.77864	-	-	-	-	-	-
tan(ϕ'_p)	-0.77864	-	-	-	-	-	-	-
tan(ϕ'_s)	-	-	-	0.26667	-0.31282	0.39744	0.31282	0.57692
tan(ϕ'_{cv})	-	-	0.26667	-	0.05641	-0.08974	-0.05641	0.03333
e	-	-	-0.31282	0.05641	-	-0.67318	-0.97378	-0.41115
γ(kN/m^3)	-	-	0.39744	-0.08974	-0.67318	-	0.66382	0.42598
γ_d(kN/m^3)	-	-	0.31282	-0.05641	-0.97378	0.66382	-	0.40519
ψ (°)	-	-	0.57692	0.03333	-0.41115	0.42598	0.40519	-

A general feeling of the data scatterness can be seen in Fig. 5. The figure indicates the datasets (tan(ϕ'_s), tan(ϕ'_{cv})) and (γ, γ_d) are having positive dependence while (c'_p, tan(ϕ'_p)) and (e, ψ) are having negative dependence. This agrees well with the results in Table 7. From the plot we can see that the dependences of these four paired datasets are not perfectly linear. Especially, the paired dataset (γ, γ_d) has some particular concentrations in its domain (around the mean). The dependence of dataset (c'_p, tan(ϕ'_p)) is also observed to be quite high when c'_p is close to zero which was resulting from the physical limitation imposing positive values for the cohesion. To better understand the dependences among the soil parameters, the datasets are transformed into the copula domain for the analysis. Figure 6 presents the scatter plot of these transformed soil data in the copula domain. As expected, the transformed paired soil data in the copula domain are not perfectly symmetric. From the density plot it can be observed that the probability density of (tan(ϕ'_s), tan(ϕ'_{cv})) centralizes at several parts in the copula domain which is quite asymmetric. The probability density of (c'_p, tan(ϕ'_p)) also shows a much higher concentration at the minimums compared to the maximums. This also causes asymmetric dependences in the copula domain.

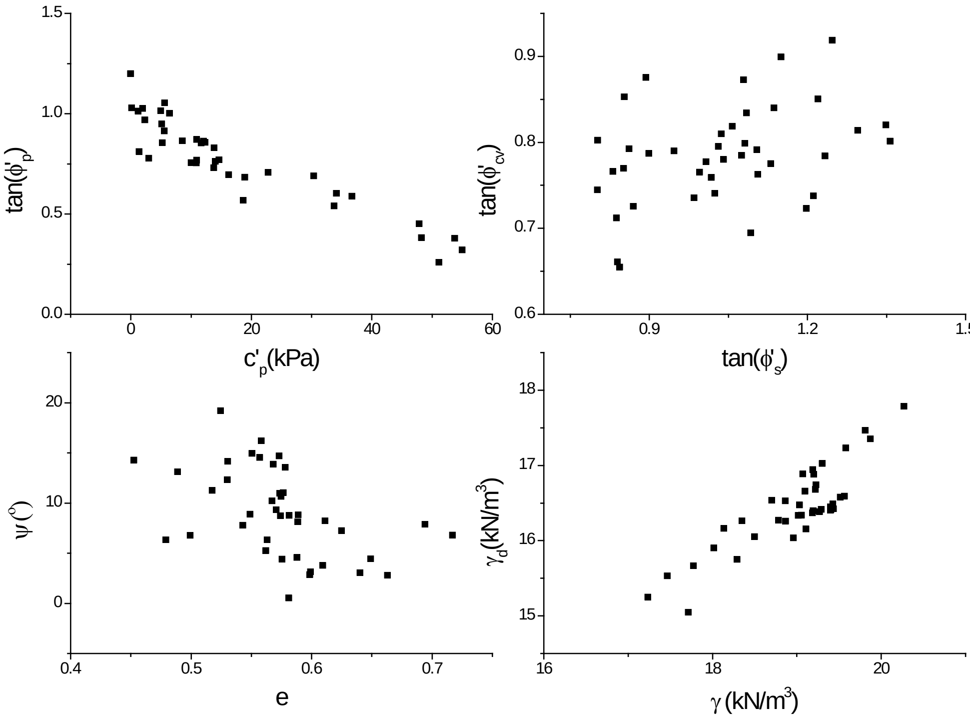


Figure 5 Scatter plot of (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d)

To further investigate the asymmetric dependence, the measure of the asymmetry as introduced in Section 4.1, is computed for the paired data and presented in Table 8. Here, the value of p is set to be infinity in the calculation of the measure of asymmetry as given by Eq. (5). The results show that the dataset (tan(ϕ'_s), tan(ϕ'_{cv})) has a larger asymmetric dependence compared to the others. This may be explained by the fact the secant friction angle ϕ'_s is dependent of the normal stress, as previously referred, while ϕ'_{cv} does have this dependency.

Another way of depicting this asymmetric dependence can be done by checking the tail dependence coefficients. By utilizing the concepts of tail dependence, the *upper-lower* and *lower-upper* tail dependence coefficients are calculated for the paired data based on Eqs. (7)-(8). The results are plotted in Fig. 7. It is seen that the *upper-lower* ($\lambda^{u,l}$) and the *lower-upper* tail ($\lambda^{l,u}$) dependence coefficients have some differences for all the considered datasets when the quantile values are close to zero (e.g. $u \rightarrow 0$). Generally, if any differences between the *upper-lower* ($\lambda^{u,l}$) and the *lower-upper* ($\lambda^{l,u}$) tail dependence coefficients are observed, the bivariate data is believed to be asymmetrically dependent. Therefore, it is necessary to utilize asymmetric copulas to model the data in this case.

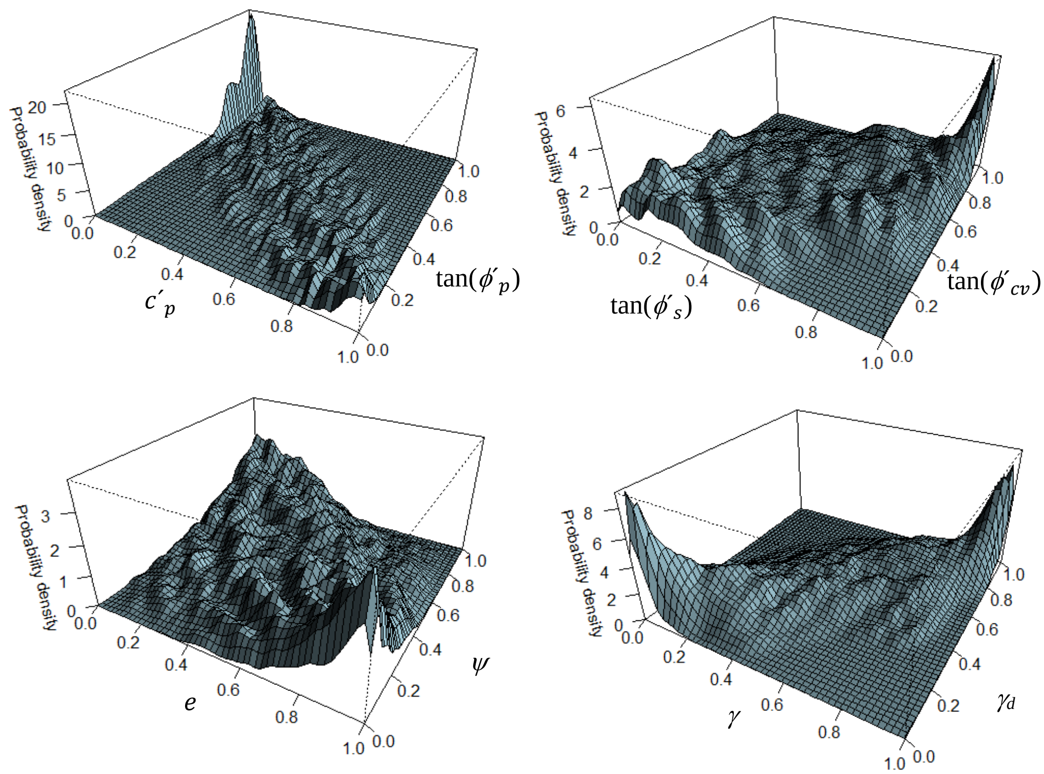
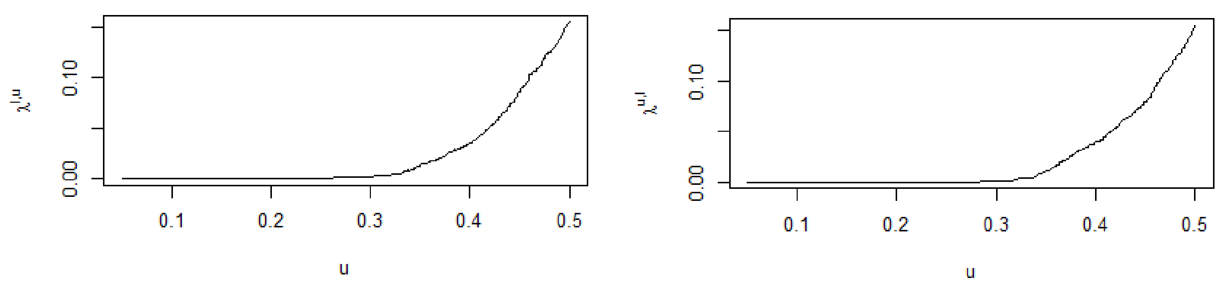


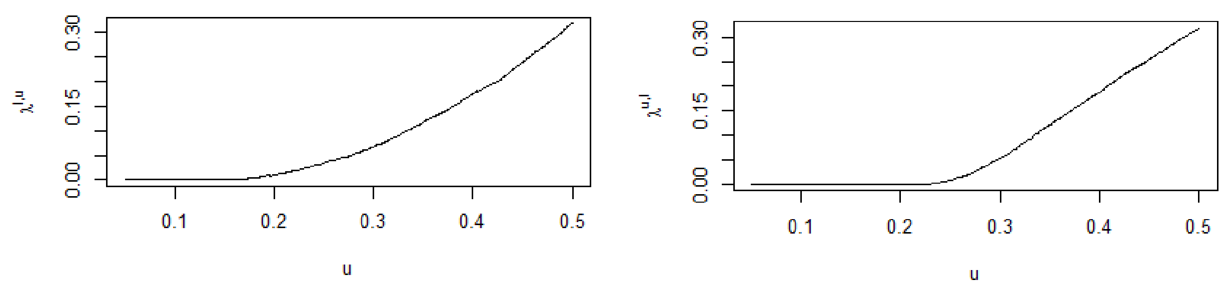
Figure 6 Empirical probability density of (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d) in the copula domain

Table 8 Measure of asymmetry in the bivariate data (-c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (-e, ψ) and (γ, γ_d)

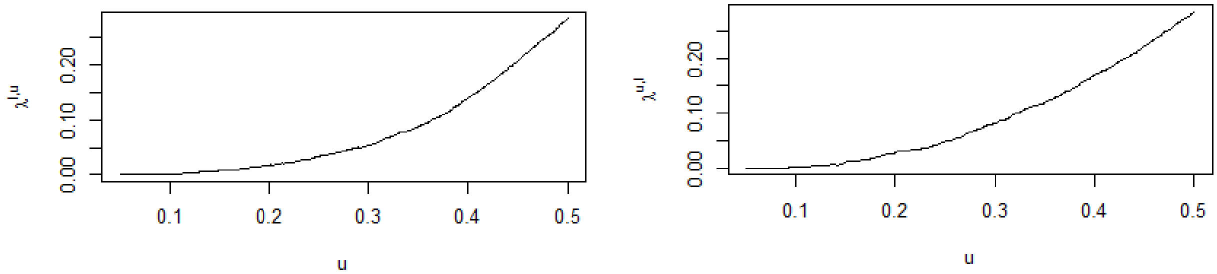
	(-c'_p, tan(ϕ'_p))	(tan(ϕ'_s), tan(ϕ'_{cv}))	(-e, ψ)	(γ, γ_d)
Measure of asymmetry η_∞	0.011	0.033	0.009	0.012



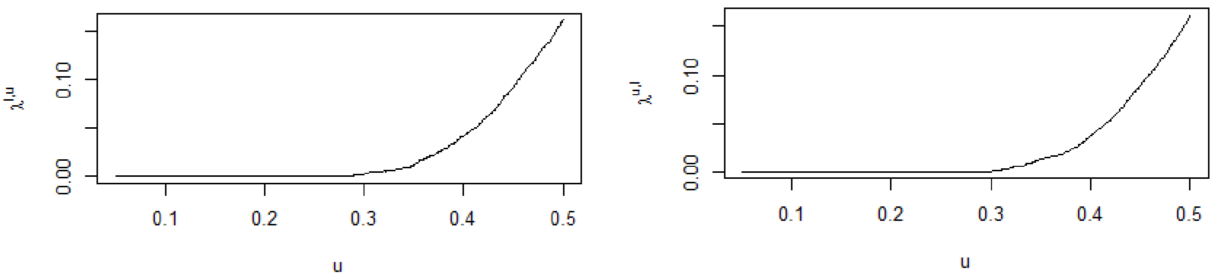
(-c'_p, tan(ϕ'_p))



(tan(ϕ'_s), tan(ϕ'_{cv}))



(-e, ψ)



(γ, γ_d)

Figure 7 Estimated empirical tail dependences for (-c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (-e, ψ) and (γ, γ_d)

Several asymmetric copulas, as introduced in Section 4, are utilized here to model the soil data. To compare with the symmetric copula, the commonly adopted symmetric Archimedean copulas are also considered. Moreover, the combination rule allows much more possible expansions for the asymmetric copula. Thus, in order to make the problem simpler, this study will only utilize the Archimedean copulas as the base copulas for the construction of asymmetric copulas. We choose the most commonly applied Archimedean copulas that can characterize different tail dependences in this study, namely, Gumbel, Clayton and Frank copulas. Following the construction rules, the asymmetric copulas are established based on these selected copulas. More specifically, the following categories of copulas are been investigated:

1. *Symmetric copulas*: The original symmetric Archimedean copulas are considered herein. They are one parameter copulas, Gumbel, Clayton and Frank copulas.
2. *Asymmetric copulas constructed by products*: We adopt the Khoudraji's device for the construction of asymmetric copulas. Following Eq. (10), we combine two base copulas from the selected Archimedean copulas. This gives three combinations namely, Gumbel-Clayton, Gumbel-Frank and Clayton-Frank. For the individual functions, the Type I function in Table 1 is selected for the asymmetric copula construction.
3. *Asymmetric copulas constructed by linear convex combinations*: This group of asymmetric copulas is constructed by the rules introduced in Section 4.3. The selected base copulas for constructing this asymmetric copula are Gumbel, Clayton and Frank copulas.
4. *Skewed Gaussian copula*: The last asymmetric copula has its exact formulation as given in Section 4.4. No base copulas are needed in this category.

Meanwhile, it is noted the Gumbel, Clayton and Frank copulas are usually used to feature positive dependences. For bivariate data having negative dependences, the use of these copulas will have problems in the parameter estimation. Therefore, for the ease of modeling, a slight change is made to the negative dependent paired datasets (c'_p, tan(ϕ'_p)) and (e, ψ) in the copula modeling. Instead of directly modeling the original data, the copula models are utilized to model the (-c'_p, tan(ϕ'_p)) and (-e, ψ). As copula model is established based on variables' cumulative distribution function values, such change of magnitude will not affect the quality of a copula model. However, the marginal distribution models for the individual variables will remain unchanged.

The results for the AIC statistics for all the considered models fitting to ($-c'_p$, $\tan(\phi'_p)$), ($\tan(\phi'_s)$, $\tan(\phi'_{cv})$), ($-e$, ψ) and (γ, γ_d) are reported in Table 9. The model parameters are estimated by the method of minimization of Cramer-von Mises statistic, which is explained in Appendix A. The best models among all the candidate models are marked in the tables. The results show that the best copula models for ($-c'_p$, $\tan(\phi'_p)$), ($\tan(\phi'_s)$, $\tan(\phi'_{cv})$), ($-e$, ψ) and (γ, γ_d) are Gumbel-Clayton Type I, Gumbel-Frank Type I, Frank and Gumbel-Clayton Type I copulas. Generally, the asymmetric copulas show an AIC value lower than the one parameter Archimedean copulas except for the dataset ($-e$, ψ). For example, the dataset ($-e$, ψ) is very symmetric in the copula domain as indicated previously in Table 8. Thus, the use of asymmetric copulas does not show clear advantages in this case. For the other three datasets, the asymmetric copulas all showed a lower AIC value. The quality of asymmetric copulas highly relies on the utilized base copulas. For instance, in modeling the data ($\tan(\phi'_s)$, $\tan(\phi'_{cv})$), the Gumbel and Frank copulas show better performance compared to Clayton copula when they are used as base copulas (e.g. the AIC value in either Clayton-Gumbel Type I or Clayton-Frank Type I is larger than Frank-Gumbel Type I). This indicates the dependence characteristic in Clayton copula may not be very suitable for the data ($\tan(\phi'_s)$, $\tan(\phi'_{cv})$). Despite the selection of base copulas, the construction rules are also a dominant factor for the quality of asymmetric copulas. The AIC values show that the overall performance of asymmetric copulas constructed by Khoudraji's device is quite prominent. However, the asymmetric copulas constructed by linear convex combinations are not very desirable as AIC values are quite large. This indicates the way of constructing the asymmetric copulas by linear convex combinations is not adequate for modeling the soil data dependences in this case. Compared to these combined asymmetric copulas, skewed Gaussian copula gives moderate performance. However, the key feature of using skewed Gaussian copula is that no base copulas are needed. It does not need to consider the selections of base copulas which might not be appropriate for the data.

Table 9 Comparison of copula parameter estimates and AIC statistics to the data of ($-c'_p$, $\tan(\phi'_p)$), ($\tan(\phi'_s)$, $\tan(\phi'_{cv})$), ($-e$, ψ) and (γ, γ_d)

Copula type		AIC			
		($-c'_p$, $\tan(\phi'_p)$)	($\tan(\phi'_s)$, $\tan(\phi'_{cv})$)	($-e$, ψ)	(γ, γ_d)
1. One parameter copula	Gumbel	-63.62	6.574	1.56	-28.9
	Clayton	-57.62	9.05	0.442	-33.28
	Frank	-56.34	5.848	-2.504*	-23.5
2. Asymmetric copulas constructed by products	Gumbel-Clayton Type I	-64.9*	6.502	-1.946	-35.5*
	Gumbel-Frank Type I	-64.4	5.764*	-0.392	-24.92
	Frank-Clayton Type I	-62.96	10.312	0.358	-31.04
3. Asymmetric copulas constructed by linear convex combinations	Gumbel-LCC	-26.8	13.198	11.336	-4.626
	Clayton-LCC	-27.08	14.796	22.182	0.256
	Frank-LCC	-22.94	11.428	1.822	-8.506
4. Skewed copula	Skewed Gaussian	-44.02	11.936	8.61	-19.542

*Minimum AIC value indicates the best model in each type.

To further check the quality of fitted asymmetric copulas, a comparison is made between the empirical data and the simulated data from the established models. Based on the best copula models in Tables 9-12, the simulated data for (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d) are plotted in Fig. 8. The simulations are performed based on the method introduced in Appendix A. It can be seen the simulated data and the original data can fit each other very well in the scatter plot. The concentrations of the simulated data generally overlap the concentrations of original data in all the plots. Even the nonlinear dependences between the variables are also well handled by the asymmetric copula, see (tan(ϕ'_s), tan(ϕ'_{cv})). A more clearer view of the fitting quality can be seen from the contour plots of the probability densities of the empirical data and the simulated data. Generally, the contour line could be used as an indicator of the quality in predicting extreme values in the bivariate data. The selection of the most accurate multivariate model has to be made based on the tail fitting capabilities. Figure 9 shows four levels of the probability density function values for both the original data and the simulated data. As expected, the quality of the model fitting to the empirical data is decreasing with the drop of contour level values. Nevertheless, the similarities of the contour lines are still quite high in all the cases. For example, as for (e, ψ), the contour lines from both original data and the simulated data can be very well fitted even for level value equals to 0.01. The rest adopted asymmetric copulas also show prominent performance in the contour fitting. These have further validated that the asymmetric copulas are very applicable to soil data modeling, and also demonstrating advantages for geotechnical reliability analysis.

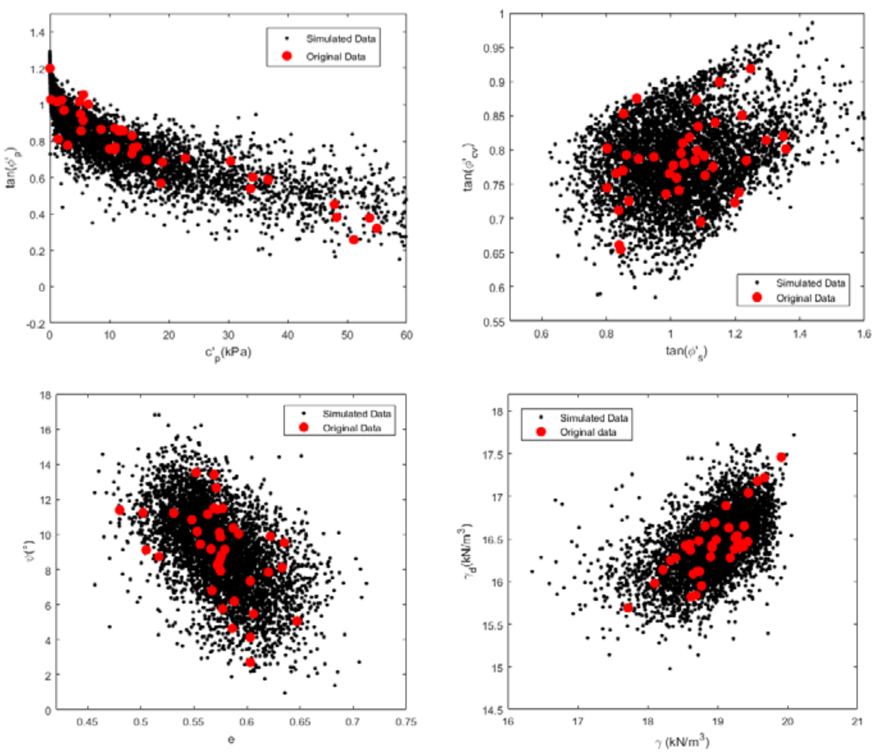


Figure 8 Comparison of scatterplots between original data and simulated data for (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d)

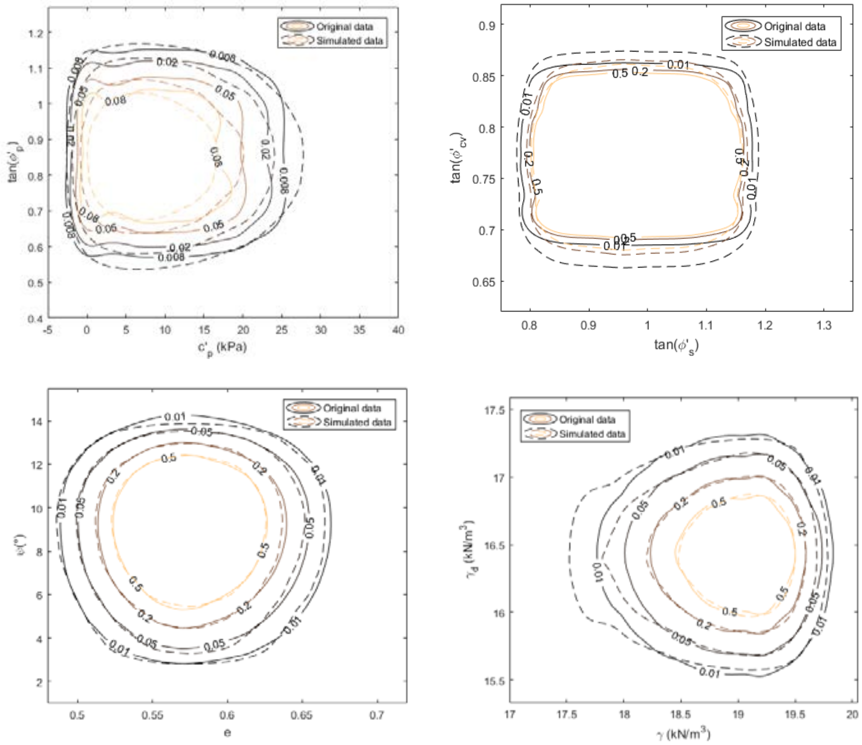


Figure 9 Comparison of contour plot between original data and best fitted copula models for (c'_p, tan(ϕ'_p)), (tan(ϕ'_s), tan(ϕ'_{cv})), (e, ψ) and (γ, γ_d) (black line indicates the empirical data; dash line indicates the fitted model).

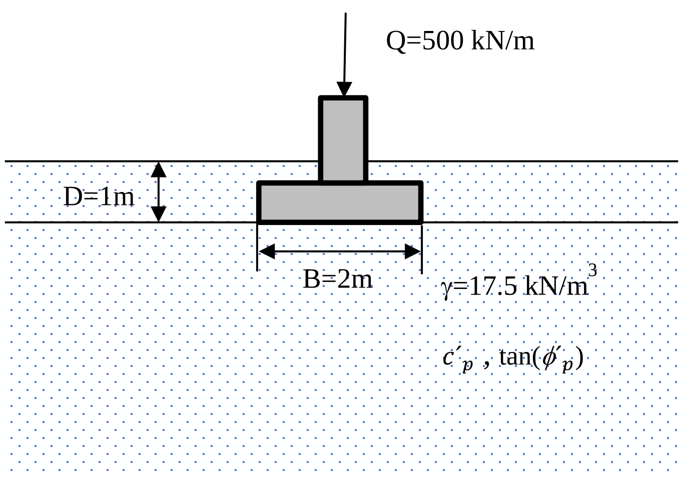


Figure 10 Strip foundation reliability analysis

In order to demonstrate the significance of using the asymmetric copulas, a reliability analysis is performed for a typical geotechnical problem by using the constructed copulas. In this example, a common strip foundation on the granite residual soil is been analyzed, see Fig. 10. The foundation is located 1 m below the ground surface, D=1 m and the width of the foundation is 2 m, B=2 m. The filled soil has a unit weight of 17.5 kN/m^3 whereas the soil cohesion c'_p and friction angle tan(ϕ'_p) are assumed to be characterized by the copulas constructed in Table 9. In this example, the load exerted on the foundation is set at Q=500 kN/m. The design formula for calculating the bearing capacity of the foundation is defined as

$$q_{ult} = c'_p \cdot N_c + q' \cdot N_q + \frac{1}{2} \times \gamma \cdot B \cdot N_\gamma \quad (20)$$

where the capacity factors N_c, N_q and N_γ are depending on the friction angle of the ground soil and estimated by

$$N_q = e^{\pi \cdot tan\phi'_p} \cdot tan^2\left(45° + \frac{\phi_p{}'}{2}\right) \quad (21)$$

$$N_c = \left(N_q - 1\right) \cdot cotg\phi'_p \quad (22)$$

$$N_\gamma = e^{\frac{1}{6}*(\pi+3\pi^2 tan\phi'_p)} \times \left(tan\phi'_p\right)^{2\pi/5} \quad (23)$$

The effective stress at the base of the foundation ′ , in the present case, is calculated by

$$q' = \mathrm{D} \times \gamma \quad (24)$$

where *D* is the depth of the footing and γ is the unit weight of the residual soil. Thus, the ultimate vertical load strength of the foundation is determined by

$$Q_{ult} = q_{ult} \times B \quad (25)$$

Therefore, the overall performance function can be formulated by the following equation.

$$G = Q_{ult} - Q \quad (26)$$

In reliability calculations, Monte Carlo simulations with 10^6 samples are performed to calculate the failure probability of Eq. (26). The associated copulas as listed in Table 9 are utilized in the reliability analysis separately. In order to show the significance of using the asymmetric copulas, a comparison is made on the failure probability between using the symmetric copulas and asymmetric copula. The computed results is shown in Table 10. It can be seen the failure probabilities differs quite a lot among the constructed copulas. The highest failure probability is $2.59{\cdot}10^{-3}$ in Gumbel copula and the lowest failure probability is $3.05{\cdot}10^{-5}$ in Clayton copula. The asymmetric copula produces a failure probability of $1.44{\cdot}10^{-4}$ which is a moderate value among all the copulas. However, it is noted the asymmetric copula could produce a failure probability that is very different from the symmetric copulas. The results of failure probability is very sensitive to the adopted copula. Therefore, even the goodness-of-fit statistics (e.g. AIC) is very close, it could not simply imply a similar value in the failure probability. Either symmetric or asymmetric dependences could have great influences in the safety assessment.

Table 10 Comparison of the failure probability using different copulas (B=2 m)

	Gaussian	Gumbel	Clayton	Frank	Gumbel-Clayton Type I
Failure probability	$9.05{\cdot}10^{-4}$	$2.59{\cdot}10^{-3}$	$3.05{\cdot}10^{-5}$	$2.49{\cdot}10^{-3}$	$1.44{\cdot}10^{-4}$

In the final part of this study, a short reference is made to discuss the possibility of extending the current bivariate asymmetric copulas to multivariate ones. This extension can be achieved with the aid of "*pair copula construction* (PCC)" techniques. There is an extensive literature on PCC techniques and their properties, for example, see Joe (2014), Bedford and Cooke (2001) and Aas et al. (2009). The key idea is to derive a general principle for decomposing a multivariate distribution into bivariate copulas and the distribution margins. The most common way is to utilize the conditional distributions in relating the multivariate distribution to bivariate distributions. However, the accuracy of a multivariate model highly relies on the choices of copulas in each step. A "clever choice" would make the multivariate model much more adequate. For more advanced techniques in PCC, one can refer to some technical books in discussing the construction of vine copulas, for example, see (Matthias and Mai, 2017).

It should be pointed out the results obtained from the present study can only be interpreted for the collected soil data. The soil parameter may exhibit different dependences in other situations when geological/geotechnical conditions change. Moreover, it also should be realized the sample size in this study is quite small. Such small sample size dataset sometimes may not be enough to represent the soil parameters. Thus, the conclusions may be distorted in other situations. For more references discussing the influence of data scarceness uncertainties to the multivariate modeling, one can read Ching et al. (2010), Beer et al. (2013), Ching and Phoon (2014a,b) and (Ching and Phoon 2015). In fact, in this study, the asymmetric copulas are only proved to be more accurate in depicting the data when they are asymmetrically dependent. In this context, if the geotechnical data is not expected to be asymmetrically dependent, then the application of asymmetric copulas is not very necessary. Although this analysis is only valid for the selected dataset, the results can be used to explain significant features of using asymmetric copulas for modeling soil data in general. Meanwhile, we should also note the asymmetric copulas are more flexible compared to the traditional copula models. Various types of base copulas and individual functions can be chosen and implemented for the construction of asymmetric copulas. This flexibility provides the asymmetric copula a great feature in its application to the data analysis. The findings of this study can help geotechnical engineers or researchers to have a better understanding of the soil data. The guidelines presented in this paper can support the design and analysis of geotechnical problems when considering soil dependences.

7. Conclusions

In this paper, the soil data have been analyzed by means of the asymmetric copulas in a multivariate setting. The fundamental formulation and theoretical basics of asymmetric copulas have been reviewed in details regarding the modeling of soil parameters. These include the concepts of measuring the asymmetric dependences and tail dependences. Several ways of constructing an asymmetric copula were introduced. These introduced asymmetric copulas were then compared with several Archimedean copulas on the modeling of soil parameters collected from a site located in Portugal. The soil parameters are divided into four groups of bivariate dataset. The copula models were constructed for each of the data group and compared based on the goodness-of-fit statistics. The results showed that the asymmetric copula can provide more appropriate characterizations of the asymmetric dependences and tail dependences in the soil data. It was found that the asymmetric copula can also provide accurate predictions of extreme values from the empirical data. However, if the soil data does not possess an obvious asymmetric dependence, the use of asymmetric copula would not be very necessary. The study also demonstrated that the asymmetric copulas can be quite powerful in capturing the extreme contours. Therefore, it is expected that asymmetric copula can contribute to improve the reliability analysis or risk assessment of geotechnical problems due to soil data modeling. Future work seems necessary to investigate the ways of selecting base copulas and individual functions in the construction of asymmetric copulas. Also, applications of the obtained asymmetric copula to real geotechnical problems, as well as different site data, may prove to have relevant interest regarding Geotechnical Reliability Based Design.

Acknowledgments

This study is supported by Tsinghua University Initiative Scientific Research Program and grants from the Alexander von Humboldt Foundation. The first author, Yi Zhang, is sponsored by "Humboldt Research Fellowship for Postdoctoral Researchers" Program. Such financial aids are gratefully acknowledged.
Meanwhile, The authors would like to thank the members of the TC304 Committee on Engineering Practice of Risk Assessment & Management of the International Society of Soil Mechanics and Geotechnical Engineering for developing the database 304dB used in this study and making it available for scientific inquiry. We also wish to thank Jaksa, Stuedlein and Grashuis for contributing this database to the TC304 compendium of databases.

Appendix A Parameter estimation and simulation of asymmetric copulas

In this section, a brief introduction of the fundamentals of parameter estimation and simulation of asymmetric copulas is provided. For more detailed fundamental basics and theoretical proofs, one can read from Nelsen (2006). This section will provide some discussions only on a simplified bivariate problem. The same concept can be easily expanded to high dimensional models.

Many parameter estimation methods have been developed by the former researchers. The most well known method is the maximum likelihood method. The concept of maximum likelihood method is to maximize the likelihood value of a distribution function when it is fitted to the empirical data. The idea of this method is quite straight forward and it has been widely used to estimate the parameters for copulas having only one parameter. However, when multiple parameters exist in the copula, the maximum likelihood method becomes quite difficult as the maximization tend to be quite tedious. The computation can become quite cumbersome for most computers.

An easy way to estimate the parameters of copulas having multiple parameters can be done through the distance based estimation method. For this concern, the Cramer-von Mises statistic $\boldsymbol{S}$ can be employed here to seek the most appropriate model parameters $\boldsymbol{\Theta} = \{\theta_1, \dots, \theta_n\}$ of the copula. In Cramer-von Mises statistics, $\boldsymbol{S}$ generally calculates the distances between the empirical copula distribution function and the theoretical copula distribution function. The minimization of this statistic will produce the most desirable estimates for the copula parameters. For instance, in estimating the parameters for a bivariate copula, the Cramer-von Mises statistic based estimation method can be formulated as

$$\Theta = \underset{\theta_1,\dots,\theta_n}{\arg\min}\, S = \underset{\theta_1,\dots,\theta_n}{\arg\min} \sum_{i=1}^{N} \left\{ C_{empirical}\left(u_1^i, u_2^i\right) - C_{\Theta}\left(u_1^i, u_2^i\right) \right\}^2 \tag{A.1}$$

where N is the number of data, $C_{empirical}$ is the empirical copula function, C_{Θ} represents the fitted parametric copula and Θ stands for the set of copula parameters that need to be estimated. Thus, the concept is to minimize the distances of cumulative distribution functions by evaluating the statistic for each of the observed data point (u_1^i, u_2^i).

The simulation method for asymmetric copulas can follow the traditional algorithm used for symmetric copulas. For instance, the most commonly applied simulation approach is the conditional distribution approach which is developed based on the Rosenblatt transform (Devroye 1986). Similar concepts for simulating random vectors from asymmetric copulas are also developed by other researchers (Matthias and Mai, 2017). The key weakness of the conditional distribution based simulation approach is that it requires a root finding procedure. If the conditional distribution can be easily derived from the copula function, this simulation technique can be well applied. Unfortunately, due to the complex formulation of an asymmetric copula, the derived conditional distribution is quite complicated. As such, the conditional distribution based simulation is too cumbersome. There are many other ways of simulating an asymmetric copula. Here we will introduce a simple way to simulate data from an asymmetric copula constructed by products. For example, suppose that we need to generate a set of n-dimensional multivariate data from an asymmetric copula constructed by products by two base copulas (e.g. m=2) and type I individual function (see Section 4.2):

$$C_{product}\left(u_1,\dots,u_n\right) = C_1\left(u_1^{\theta_1},\dots,u_n^{\theta_n}\right) C_2\left(u_1^{1-\theta_1},\dots,u_n^{1-\theta_n}\right). \tag{A.2}$$

Simulating these uniform variates from this copula can be accomplished through the following steps:

1. Generate n uniform variates $(v_1, \dots, v_i, \dots, v_k)$ from the first base copula C_1(.);
2. Generate n uniform variates $(t_1, \dots, t_i, \dots, t_k)$ from the second base copula C_2(.);
3. Then the random data $(u_1, \dots, u_n)$ from the asymmetric copula can be obtained by the following

$$u_i = \max\left\{ v_i^{1/\theta_i}, t_i^{1/(1-\theta_i)} \right\} \quad \text{for } i=1,\dots n. \tag{A.3}$$

One can easily see that the other copulas having different number and types of base copulas can also be included in this simulation technique. With the same concept, it is straightforward to formulate the simulation algorithm for other asymmetric copulas constructed by products with different types of individual functions.

Lastly, in order to facilitate the practical use of asymmetric copula for geotechnical engineers, it is worth to mention some statistical software which already contains certain simulation techniques for the asymmetric copula modeling. For example, the package named "copula" in *R* (Yan 2007; Hofert et al., 2014) can easily perform the simulation of Khoudraji copula. For instance, for simulating a bivariate asymmetric Gumbel-Frank Type I copula (e.g. like the one in Table 9), the following code can be directly used in R:

```
C1 <- khoudrajiCopula (copula1=gumbelCopula(param=5), copula2=frankCopula(param=5),
shapes=c(0.7,0.4))
X1 <- rCopula(copula=C1,n=5000)
plot(X1)
contour(C1, dCopula, nlevels = 20)
C2 <- khoudrajiCopula (copula1=gumbelCopula(param=5), copula2=frankCopula(param=5),
```

```
shapes=c(0.5,0.5))
X2 <- rCopula(copula=C2,n=5000)
plot(X2)
contour(C2, dCopula, nlevels = 20)
C3 <- khoudrajiCopula (copula1=gumbelCopula(param=5), copula2=frankCopula(param=5),
shapes=c(0.4,0.7))
X3 <- rCopula(copula=C3,n=5000)
plot(X3)
contour(C3, dCopula, nlevels = 20)
```

A general view of the simulated data can be seen in scatter plot given in Fig. A.1.

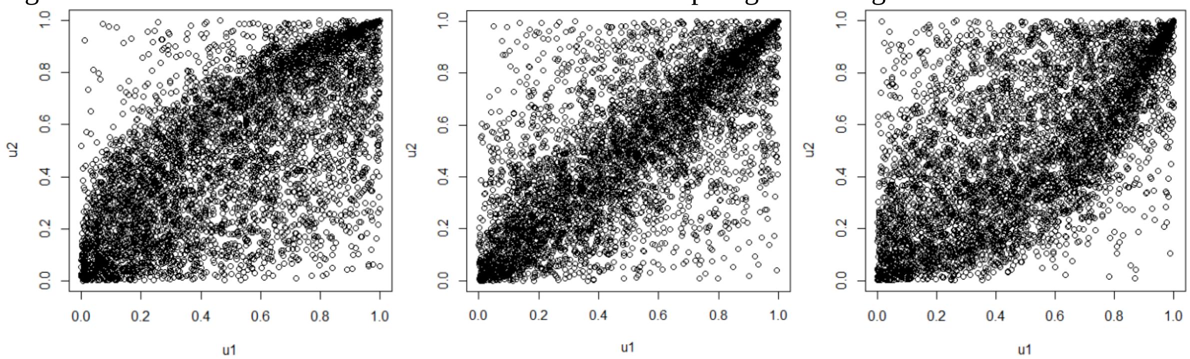


Figure A.1 Scatter plot of 5000 samples from bivariate asymmetric Gumbel-Frank Type I copula with parameters (θ_1, θ_2)=(0.7,0.4) (left), (θ_1, θ_2)=(0.5,0.5) (middle) and (θ_1, θ_2)=(0.4,0.7) (right)

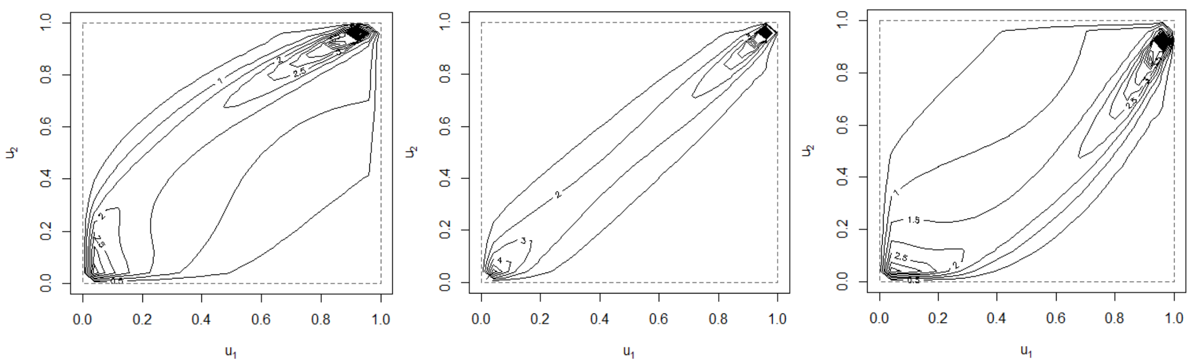


Figure A.2 Contour plot of 5000 samples from bivariate asymmetric Gumbel-Frank Type I copula with parameters (θ_1, θ_2)=(0.7,0.4) (left), (θ_1, θ_2)=(0.5,0.5) (middle) and (θ_1, θ_2)=(0.4,0.7) (right)

The fitting of asymmetric copulas by using the Khoudraji's device can also be done by using the "copula" package. The following code can be used to estimate the parameters for a bivariate asymmetric Gumbel-Frank Type I copula:

```
fitCopula(khoudrajiCopula(copula1 = gumbelCopula(),copula2 = claytonCopula()), start = c(4,4,
0.5, 0.5), data = X1, optim.method = "Nelder-Mead")
```

The associated likelihood values can be simply determined by typing the following code in R:

```
loglikCopula(c(5,   5, 0.5,   0.5), u = X1, copula = C1)
```

However, the speed of the parameter estimation calculation highly relies on the starting values. An appropriate starting value could reduce the calculation time tremendously. Meanwhile, as mentioned in Section 6, it should be emphasized the selection of the base copulas is very important in constructing the asymmetric copula. Wrong use of the base copulas may lead to undesirable results in the modeling.

References

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. Insurance: Mathematics and economics, 44(2), 182-198.

Angeli, M. G., Pasuto, A., & Silvano, S. (2000). A critical review of landslide monitoring experiences. Engineering Geology, 55(3), 133-147.

Azzalini, A., & Valle, A. D. (1996). The multivariate skew-normal distribution. Biometrika, 83(4), 715-726.

Baecher, GB, and JT Christian. (2003). Reliability and Statistics in Geotechnical Engineering: John Wiley & Sons.

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial intelligence, 32(1-4), 245-268.

Beer, M., Zhang, Y., Tong Quek, S; Phoon, KK (2013). Reliability analysis with scarce information: Comparing alternative approaches in a geotechnical engineering context, Structural Safety, Volume 41, 2013, Pages 1-10.

Charpentier, A., Fougères, A. L., Genest, C., & Nešlehová, J. G. (2014). Multivariate archimax copulas. Journal of Multivariate Analysis, 126, 118-136.

Ching J. Y. and Phoon, K. K. (2015). Reducing the transformation uncertainty for the mobilized undrained shear strength of clays. Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 141(2), 04014103.

Ching, J. & Phoon, K. K. (2014). Transformations and correlations among some clay parameters – the global database. Canadian Geotechnical Journal, 51(6), 663-685.

Ching, J. and Phoon, K. K. (2012). Modeling parameters of structured clays as a multivariate normal distribution. Canadian Geotechnical Journal, 49(5), 522-545.

Ching, J., Phoon, K. K. & Chen, Y. C. (2010). Reducing shear strength uncertainties in clays by multivariate correlations. Canadian Geotechnical Journal, 47(1), 16-33.

Ching, J., Phoon, K. K., and Chen, C. H. (2014). Modeling CPTU parameters of clays as a multivariate normal distribution. Canadian Geotechnical Journal, 51(1), 77-91.

D'Ignazio, M., Phoon, K. K., Tan, S. A. and Länsivaara, T. T. (2016). Correlations for undrained shear strength of Finnish soft clays. Canadian Geotechnical Journal, 53(10), 1628-1645.

Devroye, L. (1986). Sample-based non-uniform random variate generation. In Proceedings of the 18th conference on Winter simulation (pp. 260-265). ACM.

Duncan, J. M. (2000). Factors of Safety and Reliability in Geotechnical Engineering. Review of. Journal of Geotechnical and Geoenvironmental Engineering 126 (4):307-16. doi: 10.1061/(ASCE)1090-0241(2000)126:4(307).

Forrest, W. S., and Trevor L. L. Orr. (2010). Reliability of shallow foundations designed to Eurocode 7. Review of. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 4 (4):186-207. doi: 10.1080/17499511003646484.

Genest, C., & Favre, A. C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. Journal of hydrologic engineering, 12(4), 347-368.

Genest, C., & Nešlehová, J. G. (2013). Assessing and modeling asymmetry in bivariate continuous data. In Copulae in Mathematical and Quantitative Finance (pp. 91-114). Springer, Berlin, Heidelberg.

Genest, C., & Rivest, L. P. (1993). Statistical inference procedures for bivariate Archimedean copulas. Journal of the American statistical Association, 88(423), 1034-1043.

Grimaldi, S., & Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis. Advances in Water Resources, 29(8), 1155-1167.

Harris, C., Smith, J. S., Davies, M. C., & Rea, B. (2008). An investigation of periglacial slope stability in relation to soil properties based on physical modelling in the geotechnical centrifuge. Geomorphology, 93(3), 437-459.

Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2014). copula: Multivariate dependence with copulas. R package version 0.999-9, URL http://CRAN. R-project. org/package= copula.

Hong, Y., Wang, J. P., Li, D. Q., Cao, Z. J., Ng, C. W. W., & Cui, P. (2015). Statistical and probabilistic analyses of impact pressure and discharge of debris flow from 139 events during 1961 and 2000 at Jiangjia Ravine, China. Engineering Geology, 187, 122-134.

Hutchinson, T. P. T. P., & Lai, C. D. (1990). Continuous bivariate distributions emphasising applications (No. 04; QA278, H8.).

Joe, H. (2014). Dependence modeling with copulas. CRC Press.

Kazianka, H., & Pilz, J. (2010). Copula-based geostatistical modeling of continuous and discrete data including covariates. Stochastic Environmental Research and Risk Assessment, 24(5), 661-673.

Khoudraji, A. (1995). Contributions a l'etude des copules et a la modelisation de valeurs extremes bivariees. Ph.D. Thesis, University Laval Quebec, Canada.

Klement, E. P., & Mesiar, R. (2006). How non-symmetric can a copula be?. Commentationes Mathematicae Universitatis Carolinae, 47(1), 141-148.
Kollo, T., Selart, A., & Visk, H. (2013). From multivariate skewed distributions to copulas. In Combinatorial Matrix Theory and Generalized Inverses of Matrices (pp. 63-72). Springer India.
L'Heureux, J. S., & Long, M. (2017). Relationship between Shear-Wave Velocity and Geotechnical Parameters for Norwegian Clays. Journal of Geotechnical and Geoenvironmental Engineering, 143(6), 04017013.
Li, D. Q., Zhang, L., Tang, X. S., Zhou, W., Li, J. H., Zhou, C. B., & Phoon, K. K. (2015). Bivariate distribution of shear strength parameters using copulas and its impact on geotechnical system reliability. Computers and Geotechnics, 68, 184-195.
Li, DianQing, XiaoSong Tang, ChuangBing Zhou, and Kok-Kwang Phoon. (2012). Uncertainty analysis of correlated non-normal geotechnical parameters using Gaussian copula. Review of. Science China Technological Sciences 55 (11):3081-9. doi: 10.1007/s11431-012-4937-z.
Liebscher, E. (2008). Construction of asymmetric multivariate copulas. Journal of Multivariate analysis, 99(10), 2234-2250.
Matthias, S. and Mai, J. F. (2017). Simulating copulas: stochastic models, sampling algorithms, and applications. World Scientific.
Mazo, G., Girard, S., & Forbes, F. (2015). A class of multivariate copulas based on products of bivariate copulas. Journal of Multivariate Analysis, 140, 363-376.
Mesiar, R., & Najjari, V. (2014). New families of symmetric/asymmetric copulas. Fuzzy Sets and Systems, 252, 99-110.
Nelsen, R. B. (2003). Properties and applications of copulas: A brief survey. In Proceedings of the First Brazilian Conference on Statistical Modeling in Insurance and Finance,(Dhaene, J., Kolev, N., Morettin, PA (Eds.)), University Press USP: Sao Paulo (pp. 10-28).
Nelsen, R. B. (2006). An Introduction to Copulas. Springer New York.
Noh, Y., Choi, K. K., & Du, L. (2009). Reliability-based design optimization of problems with correlated input variables using a Gaussian Copula. Structural and multidisciplinary optimization, 38(1), 1-16.
Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. International economic review, 47(2), 527-556.
Phoon, K. K., and Ching, J. (Eds.). (2014). Risk and reliability in geotechnical engineering. CRC Press.
Phoon, K. K., J. V. Retief, J. Ching, M. Dithinde, T. Schweckendiek, Y. Wang, and L. M. Zhang. (2016). Some observations on ISO2394:2015 Annex D (Reliability of Geotechnical Structures). Review of. Structural Safety 62 (Supplement C):24-33. doi: https://doi.org/10.1016/j.strusafe.2016.05.003.
Phoon, Kok-Kwang, and Kulhawy, F. (1999). Characterization of geotechnical variability. Review of. Canadian Geotechnical Journal 36 (4):612-24. doi: 10.1139/t99-038.
Pinheiro Branco, L. (2011). Reliablity concepts applied to Granite Residual Soils. (in portugueses). MSc thesis, university of Porto.
Pinheiro Branco, L.; Topa Gomes, A: Silva Cardoso, A; and Santos Pereira, C. (2014). "Natural Variability of Shear Strength in a Granite Residual Soil from Porto." Geotechnical and Geological Engineering 32 (4):911-22. doi: 10.1007/s10706-014-9768-1.
Robertson, P. K. (2009). Interpretation of cone penetration tests—a unified approach. Canadian geotechnical journal, 46(11), 1337-1355.
Salvadori, G., & De Michele, C. (2007). On the use of copulas in hydrology: theory and practice. Journal of Hydrologic Engineering, 12(4), 369-380.
Salvadori, G., De Michele, C., Kottegoda, N. T., & Rosso, R. (2007). Extremes in nature: an approach using copulas (Vol. 56). Springer Science & Business Media.
Santoso, A. M., Phoon, K. K., & Tan, T. S. (2013). Estimating strength of stabilized dredged fill using multivariate normal model. Journal of Geotechnical and Geoenvironmental Engineering, 139(11), 1944-1953.
Sideri, D., Modis, K., & Rozos, D. (2014). Multivariate geostatistical modelling of geotechnical characteristics of the alluvial deposits in West Thessaly, Greece. Bulletin of Engineering Geology and the Environment, 73(3), 709-722.
Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris.
Tang, C., & Phoon, K. K. (2018). Statistics of Model Factors and Consideration in Reliability-Based Design of Axially Loaded Helical Piles. Journal of Geotechnical and Geoenvironmental Engineering, 144(8), 04018050.
Tang, X. S., Li, D. Q., Zhou, C. B., & Phoon, K. K. (2015). Copula-based approaches for evaluating slope reliability under incomplete probability information. Structural Safety, 52, 90-99.

Tang, Xiao-Song, Dian-Qing Li, Guan Rong, Kok-Kwang Phoon, and Chuang-Bing Zhou. (2013). Impact of copula selection on geotechnical reliability under incomplete probability information. Review of. Computers and Geotechnics 49 (Supplement C):264-78. doi: https://doi.org/10.1016/j.compgeo.2012.12.002.

Topa Gomes, A. (2009). Elliptical Shafts by the Sequential Excavation Method. The case of Metro do Porto, (in Portuguese), PhD Thesis, University of Porto.

Trivedi, P. K., & Zimmer, D. M. (2007). Copula modeling: an introduction for practitioners. Foundations and Trends® in Econometrics, 1(1), 1-111.

Vanapalli, S. K., Fredlund, D. G., Pufahl, D. E., & Clifton, A. W. (1996). Model for the prediction of shear strength with respect to soil suction. Canadian Geotechnical Journal, 33(3), 379-392.

Viana da Fonseca, A. (1996). Geomechanics of Residual Soils from Porto Granite. Design Criteria for foundations (in Portuguese), PhD Thesis, University of Porto.

Viana da Fonseca, A., and R. Coutinho. (2008). Characterization of residual soils - Keynote paper. In 3rd International Conference on Site Characterization, 195-248. Taiwan: Taylor&Francis/Balkema.

Wang, C., Zhang, H., & Li, Q. (2017a). Reliability assessment of aging structures subjected to gradual and shock deteriorations. Reliability Engineering & System Safety, 161, 78-86.

Wang, F. and H. Li (2017b). "Stochastic response surface method for reliability problems involving correlated multivariates with non-Gaussian dependence structure: Analysis under incomplete probability information." Computers and Geotechnics 89(Supplement C): 22-32.

Wu, S. (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. European Journal of Operational Research, 238(2), 476-485.

Wu, X. Z. (2013). Trivariate analysis of soil ranking-correlated characteristics and its application to probabilistic stability assessments in geotechnical engineering problems. Soils and Foundations, 53(4), 540-556.

Yan, J. (2007). Enjoy the joy of copulas: with a package copula. Journal of Statistical Software, 21(4), 1-21.

Yan, W. M., Yuen, K. V., & Yoon, G. L. (2009). Bayesian probabilistic approach for the correlations of compression index for marine clays. Journal of Geotechnical and Geoenvironmental Engineering, 135(12), 1932-1940.

Zhang, J., H. W. Huang, C. H. Juang, and W. W. Su. (2014). Geotechnical reliability analysis with limited data: Consideration of model selection uncertainty. Review of. Engineering Geology 181 (Supplement C):27-37. doi: https://doi.org/10.1016/j.enggeo.2014.08.002.

Zhang, L., D.-Q. Li, X.-S. Tang, Z.-J. Cao and K.-K. Phoon (2018). "Bayesian model comparison and characterization of bivariate distribution for shear strength parameters of soil." Computers and Geotechnics 95(Supplement C): 110-118.

Zhang, Y., & Lam, J. S. L. (2016). A copula approach in the point estimate method for reliability engineering. Quality and Reliability Engineering International, 32(4), 1501-1508.

Zhang, Y., Beer, M., & Quek, S. T. (2015). Long-term performance assessment and design of offshore structures. Computers & Structures, 154, 101-115.

Zhang, Y., Kim, C.W., Beer, M., Dai, H., Guedes Soares, C., (2018) Modeling multivariate ocean data using asymmetric copulas. Coastal Engineering, 135, 91-111.

Zhang, Y., Gomes, A., Beer, M., Neumann, I., Nackenhorst, U., (2019) Reliability analysis with consideration of asymmetrically dependent variables: discussion and application to geotechnical examples. Reliability Engineering & System Safety, 185, 261-277.

Zhu, H., L. M. Zhang, T. Xiao and X. Y. Li (2017). Generation of multivariate cross-correlated geotechnical random fields. Computers and Geotechnics 86(Supplement C): 95-107.