

## Article

# A Vine Copula-Based Global Sensitivity Analysis Method for Structures with Multidimensional Dependent Variables

Zhiwei Bai <sup>1</sup>, Hongkui Wei <sup>2</sup>, Yingying Xiao <sup>2</sup>, Shufang Song <sup>1,\*</sup> and Sergei Kucherenko <sup>3</sup>

<sup>1</sup> School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China; bzw1020@mail.nwpu.edu.cn

<sup>2</sup> State Key Laboratory of Intelligent Manufacturing System Technology, Beijing Institute of Electronic System Engineering, Beijing 100854, China; weihongkui89@hotmail.com (H.W.); xiaoyingying504@126.com (Y.X.)

<sup>3</sup> Centre for Process Systems Engineering, Imperial College London, London SW7 2AZ, UK; s.kucherenko@imperial.ac.uk

\* Correspondence: shufangsong@nwpu.edu.cn; Tel.: +86-135-7295-9529

**Abstract:** For multidimensional dependent cases with incomplete probability information of random variables, global sensitivity analysis (GSA) theory is not yet mature. The joint probability density function (PDF) of multidimensional variables is usually unknown, meaning that the samples of multivariate variables cannot be easily obtained. Vine copula can decompose the joint PDF of multidimensional variables into the continuous product of marginal PDF and several bivariate copula functions. Based on Vine copula, multidimensional dependent problems can be transformed into two-dimensional dependent problems. A novel Vine copula-based approach for analyzing variance-based sensitivity measures is proposed, which can estimate the main and total sensitivity indices of dependent input variables. Five considered test cases and engineering examples show that the proposed methods are accurate and applicable.

**Keywords:** Vine copula; global sensitivity analysis; multivariate correlation analysis; Nataf model; probability density function



check for updates

**Citation:** Bai, Z.; Wei, H.; Xiao, Y.; Song, S.; Kucherenko, S. A Vine Copula-Based Global Sensitivity Analysis Method for Structures with Multidimensional Dependent Variables. *Mathematics* **2021**, *9*, 2489. <https://doi.org/10.3390/math9192489>

Academic Editor: Zdeněk Kala

Received: 9 September 2021

Accepted: 29 September 2021

Published: 4 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Variance-based global sensitivity analysis (GSA) is very commonly used in the area of structural safety [1–3]. It can calculate the contribution of each input variable or group variables to the variance of the output response, thus allowing for selection of the important variables that have great influence on the output, while the randomness of unimportant variables can be ignored. Several methods exist to perform variance-based GSA of model outputs when the inputs are independent [2–6], but it is still a challenge for the dependent cases.

For dependent input variables, there are some techniques proposed to solve Sobol' indices. Xu and Gertner [7] divided the contribution of inputs into two components: the independent and dependent contributions, and proposed a regression-based method for estimating different contributions of the inputs. Similarly, Li et al. [8] introduced a unified framework including covariance in the decomposition of the model output variance, which can distinguish the structural and correlative contribution of a nominal input. Kala [9] performed GSA based on entropy, and proposed a novel method from differential entropy to alternative measures. Kucherenko et al. [10] applied a Gaussian copula method to variance-based global sensitivity analysis, and introduced a more general approach. On this basis, Song et al. [11] proposed using an adaptive copula method to solve Sobol' indices, which can describe different types of correlations. The above methods are mainly aimed at the case of bivariate correlation. Yet, considering that many input variables are multidimensional and interrelated in engineering problems, it is necessary to propose an effective method with which to measure multidimensional correlation.

Over the past few years, the Nataf model proposed by Kiureghian et al. [12] has been widely used in structural safety engineering with multidimensional correlated variables. Lebrun and Dutfoy [13] proved that the Nataf model is essentially equivalent to multivariate Gaussian copula. Besides, Eryilmaz [14] studied multivariate Archimedean copula and applied multivariate Clayton and Gumbel copulas to the reliability analysis of weighted-k-out-of-n systems with dependent components. Daul et al. [15] discussed the grouped t-copula with an application to credit risk. Cossette et al. [16] reviewed the Farlie-Gumbel-Morgenstern (FGM) copula and expanded it to multidimensional situations. The above copula methods require that the correlations between different variables conform to the same copula function. However, many of the multidimensional variables in practical engineering have different correlations. For example, in the variable-amplitude load fatigue problem, the fatigue life under different stress levels has different correlations [17]. In flood frequency analysis, the flood peak value, total amount and duration are complexly correlated [18]. For such problems, the above multivariate copula models are subject to errors [19], and the correlations should be measured by different copula functions.

In this paper, we introduce Vine copula [20–22] to the variance-based global sensitivity analysis of complex multidimensional correlation problems. Vine copula can decompose the joint PDF of multidimensional variables into some bivariate copula functions of original variables and conditional variables [20], so as to measure different correlations among the multivariate variables with different copula functions. After the decomposition, the Akaike Information Criterion (AIC) method [23] or one of other methods [24,25] is required to select the optimal bivariate copula function, which can separately measure the correlation among the input variables. For engineering problems with multidimensional correlated variables, the Vine copula method can use optimal copula to separately measure the correlation of different variables. Thus, the complex dependence structure of variables can be fitted to the maximum extent, and the most accurate possible solution of  $S_i$  and  $S_i^{tot}$  can be obtained. Due to the excellent properties of Vine copula, it has been applied in structural safety and reliability. Torre et al. [26] formalized the framework needed to build Vine copula models of multivariate inputs and combine them with virtually any uncertainty quantification method. Nazih et al. [27] studied the specific problem of structural reliability as an application of Vine copula. Sarazin et al. [28] used Vine copula to perform reliability-oriented sensitivity analysis in the presence of data-driven epistemic uncertainty. Inspired by the above works, we applied Vine copula to perform variance-based sensitivity analysis.

This paper is organized as follows: Section 2 presents a detailed introduction of Vine copula, including its basic theory and classification. The process of solving variance-based global sensitivity analysis based on Vine copula is introduced in Section 3. Section 4 analyzes some numerical and engineering tests. Conclusions are summarized in Section 5.

## 2. Vine Copula

Vine copula is an efficient mathematical tool to decompose the multidimensional joint PDF into the continuous product of several two-dimensional copula functions and marginal PDF [20]. Consider a set of  $n$ -dimensional correlated variables  $X = (X_1, X_2, \dots, X_n)$ , the joint PDF  $f(x_1, x_2, \dots, x_n)$  of that can be decomposed as follows:

$$f_X(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2|X_1}(x_2)f_{X_3|X_1, X_2}(x_3) \cdots f_{X_n|X_1, X_2, \dots, X_{n-1}}(x_n) \quad (1)$$

where  $f_{X_i}(x_i)$ ,  $i = 1, 2, \dots, n$  is the marginal PDF of the input variable  $X_i$  and  $f_{X_j|X_1, X_2, \dots, X_{j-1}}(x_j)$ ,  $j = 2, 3, \dots, n$  is the conditional PDF of  $X_j$  under  $X_1, X_2, \dots, X_{j-1}$  constraint. For the multidimensional independent cases,  $f_{X_j|X_1, X_2, \dots, X_{j-1}}(x_j) = f_{X_j}(x_j)$ , i.e., the joint PDF of the variables is equal to the product of marginal PDFs.

For the case of two-dimensional variables, the joint PDF is  $f_X(x_1, x_2) = f_{X_1}(x_1)f_{X_2|X_1}(x_2)$  and the conditional PDF and be expressed as:

$$\begin{aligned}
 f_{X_2|X_1}(x_2) &= \frac{f_X(x_1, x_2)}{f_{X_1}(x_1)} \\
 &= \frac{c_{12}(F_{X_1}(x_1), F_{X_2}(x_2))f_{X_1}(x_1)f_{X_2}(x_2)}{f_{X_1}(x_1)} \\
 &= c_{12}(F_{X_1}(x_1), F_{X_2}(x_2))f_{X_2}(x_2)
 \end{aligned}
 \tag{2}$$

where  $c_{12}(\cdot)$  is copula density function of input variables  $X_1$  and  $X_2$ , while  $F_{X_i}(x_i)$  is the marginal CDF of  $X_i$ .

For the case of high-dimensional variables, the decomposition of Equation (1) is not unique. Bedford and Cooke [20,21] proposed Vine copula to describe the different decomposition methods and introduced a regular Vine (R-Vine) structure, in which C-Vine and D-Vine are two common structures. C-Vine can be well used to describe correlations among variables when a particular variable is known to be a key variable that governs interactions in the dataset (e.g.,  $X_1$  is the key variable of C-Vine in Figure 1), while D-Vine is suitable for describing correlated variables of the same status. An  $n$ -dimensional Vine copula model consists of  $n - 1$  trees  $T_j$  ( $j = 1, 2, \dots, n - 1$ ), which has  $n - j + 1$  nodes and  $n - j$  sides. Each node represents an unconditional or conditional variable, and each side represents a two-dimensional copula density function. The nodes in tree  $T_j$  are made up of the sides in tree  $T_{j-1}$  and related only to the sides that have a common node in tree  $T_{j-1}$ . According to the logical structure of C-Vine and D-Vine models, the joint PDF of  $n$ -dimensional variables can, respectively, be expressed as Equations (3) and (4):

$$f_X(x_1, x_2, \dots, x_n) = \prod_{k=1}^n f_{X_k}(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j, j+i|1, \dots, j-1} \left( F_{X_j|X_1, \dots, X_{j-1}}(x_j), F_{X_{j+i}|X_1, \dots, X_{j-1}}(x_{j+i}) \right)
 \tag{3}$$

$$f_X(x_1, x_2, \dots, x_n) = \prod_{k=1}^n f_{X_k}(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i, i+j|i+1, \dots, i+j-1} \left( F_{X_i|X_{i+1}, \dots, X_{i+j-1}}(x_i), F_{X_{i+j}|X_{i+1}, \dots, X_{i+j-1}}(x_{i+j}) \right)
 \tag{4}$$

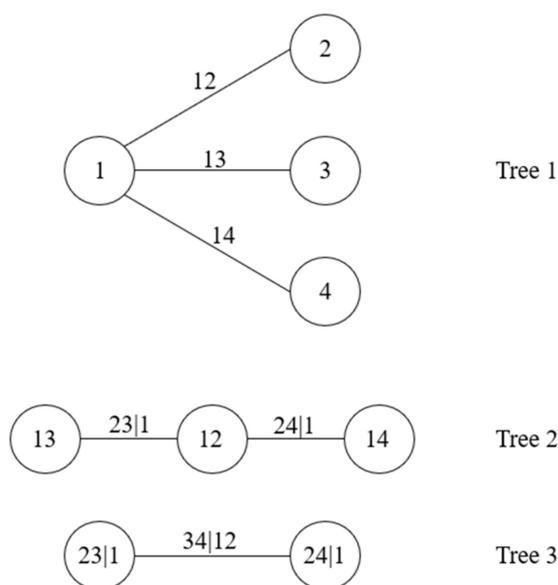


Figure 1. C-Vine model of four-dimensional variables.

Figures 1 and 2 show the structural diagrams of C-Vine and D-Vine with four dimensional variables; the corresponding decomposition of the joint PDF in Figures 1 and 2 can be expressed separately as follows. It can be seen that the notation of D-Vine resembles

independence graphs more than that of C-Vine. Therefore, we mainly adopted C-Vine in order to emphasize the influence of dependence on sensitivity indices.

$$\begin{aligned}
 f_X(x_1, x_2, x_3, x_4) &= f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3)f_{X_4}(x_4) \\
 &\cdot c_{12}(F_{X_1}(x_1), F_{X_2}(x_2)) \cdot c_{13}(F_{X_1}(x_1), F_{X_3}(x_3)) \cdot c_{14}(F_{X_1}(x_1), F_{X_4}(x_4)) \\
 &\cdot c_{23|1}(F_{X_2|X_1}(x_2), F_{X_3|X_1}(x_3)) \cdot c_{24|1}(F_{X_2|X_1}(x_2), F_{X_4|X_1}(x_4)) \\
 &\cdot c_{34|12}(F_{X_3|X_1X_2}(x_3), F_{X_4|X_1X_2}(x_4))
 \end{aligned}
 \tag{5}$$

$$\begin{aligned}
 f_X(x_1, x_2, x_3, x_4) &= f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3)f_{X_4}(x_4) \\
 &\cdot c_{12}(F_{X_1}(x_1), F_{X_2}(x_2)) \cdot c_{23}(F_{X_2}(x_2), F_{X_3}(x_3)) \cdot c_{34}(F_{X_3}(x_3), F_{X_4}(x_4)) \\
 &\cdot c_{13|2}(F_{X_1|X_2}(x_1), F_{X_3|X_2}(x_3)) \cdot c_{24|3}(F_{X_2|X_3}(x_2), F_{X_4|X_3}(x_4)) \\
 &c_{14|23}(F_{X_1|X_2X_3}(x_1), F_{X_4|X_2X_3}(x_4))
 \end{aligned}
 \tag{6}$$

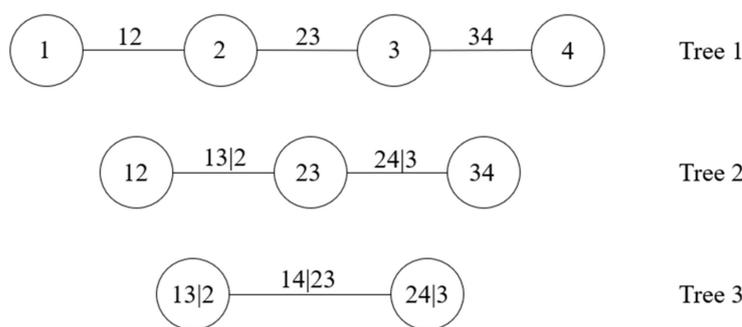


Figure 2. D-Vine model of four-dimensional variables.

The above decomposition process involves conditional distribution  $F_{X_j|X_1, \dots, X_{j-1}}(x_j)$  and  $F_{X_i|X_{i+1}, \dots, X_{i+j-1}}(x_i)$ . For the convenience of expression, we introduce the function  $h(\cdot)$  to express the binary conditional distribution, while  $h^{-1}$  denotes the inverse function of  $h(\cdot)$ .

$$h(u_i, u_j) = F_{X_i|X_j}(x_i) = \frac{\partial C(u_i, u_j | \alpha)}{\partial u_j}
 \tag{7}$$

where  $u_i = F_{X_i}(x_i)$  and  $u_j = F_{X_j}(x_j)$  are the marginal cumulative distribution functions (CDF) of  $X_i$  and  $X_j$ ,  $C(u_i, u_j)$  denotes copula function and  $\alpha$  denotes the parameter of copula function, which can be estimated by the AIC method [23] or other methods [24,25]. Table 1 lists the  $h(\cdot)$  and  $h^{-1}(\cdot)$  functions of some common copula functions [29].

According to the decomposition process in Equation (3) or (4), the multidimensional joint PDF can be transformed into the product of marginal PDF and several binary copula functions. Thus, we can use the conventional binary copula functions for variance-based global sensitivity analysis with multidimensional correlated input variables.

Table 1. The functions  $h$  and  $h^{-1}$  of some frequently-used copula functions.

Copula	$C(u_1, u_2   \alpha)$	$h$	$h^{-1}$
Gaussian	$\Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2)   \alpha)$	$\Phi\left(\frac{\Phi^{-1}(u_1) - \alpha \Phi^{-1}(u_2)}{\sqrt{1 - \alpha^2}}\right)$	$\Phi\left(\Phi^{-1}(u_1)\sqrt{1 - \alpha^2} + \alpha \Phi^{-1}(u_2)\right)$
Clayton	$(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha}$	$u_2^{-\alpha-1}(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1-1/\alpha}$	$\left((u_1 u_2^{\alpha+1})^{-\frac{\alpha}{\alpha+1}} - u_2^{-\alpha} + 1\right)^{-1/\alpha}$
Gumbel	$e^{-((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{1/\alpha}}$	$C(u_1, u_2   \alpha) \frac{1}{u_2} (-\log u_2)^{\alpha-1} ((-\log u_1)^\alpha + (-\log u_2)^\alpha)^{1/\alpha-1}$	-
Frank	$-\frac{1}{\alpha} \left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right)$	$\frac{e^{-\alpha u_2}}{\frac{e^{-\alpha} - 1}{e^{-\alpha u_1} - 1} + e^{-\alpha u_2} - 1}$	$-\log\left(1 - \frac{1 - e^{-\alpha}}{(u_1^{-1} - 1)e^{-\alpha u_2} + 1}\right) / \alpha$

### 3. Variance-Based GSA Based on Vine Copula

Considering a model function  $Y = g(\mathbf{X})$  defined in  $R^n$  with an input vector  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ , Sobol [1] and Saltelli [2] put forward the variance-based global sensitivity indices based on ANalysis Of VAriance (ANOVA) decomposition:

$$S_i = \frac{D_i}{D} = \frac{\text{Var}(E(Y|X_i))}{\text{Var}(Y)} \tag{8}$$

$$S_i^{tot} = \frac{D_i^{tot}}{D} = \frac{E(\text{Var}(Y|\mathbf{X}_{\sim i}))}{\text{Var}(Y)} \tag{9}$$

where  $D$  denotes the variance of model output  $Y$ , while  $D_i$  and  $D_i^{tot}$  are the conditional variances of  $Y$  under  $X_i$  constraint. The full expressions for  $S_i$  and  $S_i^{tot}$  are given by the following formulas:

$$S_i = \frac{1}{D} \left[ \int_R f_{X_i}(x_i) dx_i \left[ \int_{R^{n-1}} g(\mathbf{X}) f_{\mathbf{X}_{\sim i}|X_i}(\mathbf{x}_{\sim i}) d\mathbf{x}_{\sim i} \right]^2 - g_0^2 \right] \tag{10}$$

$$S_i^{tot} = \frac{1}{D} \left[ \int_{R^{n-1}} f_{\mathbf{X}_{\sim i}|X_i}(\mathbf{x}_{\sim i}) d\mathbf{x}_{\sim i} \left[ \int_R g^2(\mathbf{X}) f_{X_i}(x_i) dx_i - g_0^2 \right] \right] \tag{11}$$

where  $g_0 = \int_{R^n} g(\mathbf{X}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$  denotes the expectation of model output  $Y$ , and where the variance is given by  $D = \int_{R^n} g^2(\mathbf{X}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} - g_0^2$ . Kucherenko et al. [10] derived the formula of variance-based global sensitivity indices under correlated variables based on the conditional variance theory:

$$S_i = \frac{1}{D} \left[ \int_{R^n} g(\mathbf{X}) f(\mathbf{x}) d\mathbf{x} \left[ \int_{R^{n-1}} g(\overline{X}_1, \dots, X_i, \dots, \overline{X}_n) f_{\overline{X}_{\sim i}|X_i}(\overline{\mathbf{x}}_{\sim i}) d\overline{\mathbf{x}}_{\sim i} - \int_{R^n} g(\mathbf{X}') f(\mathbf{x}') d\mathbf{x}' \right] \right] \tag{12}$$

$$S_i^{tot} = \frac{1}{2D} \int_{R^n} \left[ g(\mathbf{X}) - g(X_1, \dots, \overline{X}_i, \dots, X_n) \right]^2 f(\mathbf{x}) f_{\overline{X}_i|X_i}(\overline{x}_i) d\mathbf{x} d\overline{x}_i \tag{13}$$

where  $\mathbf{X}$  and  $\mathbf{X}'$  are unconditional sample matrices generated according to the joint PDF of input variables, while  $(\overline{X}_1, \dots, \overline{X}_{i-1}, \overline{X}_{i+1}, \dots, \overline{X}_n)$  and  $\overline{X}_i$  are conditional samples generated by conditional PDFs  $f_{\overline{X}_{\sim i}|X_i}(\overline{\mathbf{x}}_{\sim i})$  and  $f_{\overline{X}_i|X_i}(\overline{x}_i)$ , respectively. Tarantola and Mara [30] indicated that  $S_i$  measures the amount of variance of  $X_i$  due to  $X_i$  and its dependence with  $\mathbf{X}_{\sim i}$  but does not include the interactions of them, while  $S_i^{tot}$  accounts for the contributions of  $X_i$  and its dependence with  $\mathbf{X}_{\sim i}$  by ignoring the correlations among variables. Therefore, the dependence among variables may lead to  $S_i > S_i^{tot}$  (dependent effects larger than interacted effects). By comparing the values of the two indices, we can get the independent contributions, dependent contributions and interacted contributions of the variables, so as to perform factor ranking/fixing, variance cutting, etc.

According to Equations (12) and (13), Song et al. [11] introduced two copula-based methods for solving  $S_i$  and  $S_i^{tot}$  under the case of two-dimensional correlated variables. The first method uses the copula function to approximate the real joint distribution, while the other approach uses the copula function to decompose the joint PDF into the product of the marginal PDF and copula PDF. Song et al. [11] also concluded that the second method is less computationally efficient than the first method, so we mainly discuss the first copula-based method in this paper. For the sake of description, we denote the two methods based on Vine copula as VC1 and VC2. In this section, we extend VC1 to the case of multidimensional correlation, where there is a need to decompose the joint PDF of the

input variables into the product of several two-dimensional copula functions and marginal PDF using the Vine copula function (details of VC2 are presented in Appendix A).

The basic idea of VC1 is to use the copula function to approximate the real joint distribution, and generate unconditional and conditional samples according to the selected optimal copula function. This method can rewrite (12) and (13) as follows:

$$S_i = \frac{\frac{1}{N} \sum_{j=1}^N g(\mathbf{X}^{(j)}) \left( g(\overline{X}_1^{(j)}, \dots, X_i^{(j)} \dots \overline{X}_n^{(j)}) - g(\mathbf{X}^{(j)}) \right)}{D} \tag{14}$$

$$S_i^{tot} = \frac{\frac{1}{N} \sum_{j=1}^N \left( g(\mathbf{X}^{(j)}) - g(X_1^{(j)}, \dots, \overline{X}_i^{(j)} \dots X_n^{(j)}) \right)^2}{2D} \tag{15}$$

where  $N$  is the number of samples. After the decomposition of the joint PDF with the Vine copula function, each optimal two-dimensional copula function and its parameters needs to be inferred by using the known samples. According to the selected optimal copula function, we generate samples according to the following ideas and steps [20,31].

Assuming that  $(x_1, x_2, \dots, x_n)$  is a group of samples of input variables  $\mathbf{X}$ , according to Rosenblatt model [32], the marginal and conditional distribution of  $(x_1, x_2, \dots, x_n)$  can be expressed as follows:

$$\begin{aligned} r_1 &= F_{X_1}(x_1) \\ r_2 &= F_{X_2|X_1}(x_2) \\ &\dots \\ r_n &= F_{X_n|X_1, \dots, X_{n-1}}(x_n) \end{aligned} \tag{16}$$

where  $(r_1, r_2, \dots, r_n)$  are independent and uniformly distributed between 0 and 1. Correspondingly, we can get the inverse transform of Equation (16).

$$\begin{aligned} x_1 &= F_{X_1}^{-1}(r_1) \\ x_2 &= F_{X_2|X_1}^{-1}(r_2) \\ &\dots \\ x_n &= F_{X_n|X_1, \dots, X_{n-1}}^{-1}(r_n) \end{aligned} \tag{17}$$

Assuming that  $(u_1, u_2, \dots, u_n)$  are the marginal CDF values of the input variables, i.e.,  $u_1 = F_{X_1}(x_1)$ ,  $u_2 = F_{X_2}(x_2)$ ,  $\dots$ ,  $u_n = F_{X_n}(x_n)$ , the sampling procedure is as follows:

1. Generate two independent vectors,  $(r_1, r_2, \dots, r_n)$  and  $(r'_1, r'_2, \dots, r'_n)$ , uniformly distributed between 0 and 1;
2. Let  $u_1 = r_1$ , and obtain  $x_1 = F_{X_1}^{-1}(u_1)$ ;
3. Let  $r_2 = F_{X_2|X_1}(x_2) = \frac{\partial C_{12}(F_{X_1}(x_1), F_{X_2}(x_2))}{\partial F_{X_1}(x_1)} = h(u_2, u_1)$ , and obtain  $u_2 = h^{-1}(r_2, u_1)$ ,  $x_2 = F_{X_2}^{-1}(u_2)$ ;
4. Let  $r_3 = F_{X_3|X_1, X_2}(x_3) = \frac{\partial C_{13|2}(F_{X_3|X_2}(x_3), F_{X_1|X_2}(x_1))}{\partial F_{X_1|X_2}(x_1)} = h(h(u_3, u_2), h(u_1, u_2))$ , and obtain  $u_3 = h^{-1}(h^{-1}(r_3, h(u_1, u_2)), u_2)$ ,  $x_3 = F_{X_3}^{-1}(u_3)$ ;
5. Take the iteration repeatedly, and get a group of unconditional samples  $(x_1, x_2, \dots, x_n)$ .

Considering the above sampling method involves a conditional copula function, the generation of conditional samples is similar to the above procedure. There is just a need to construct independent uniform variables as  $(r'_1, r'_2, \dots, r'_i, \dots, r'_n)$  and  $(r_1, r_2, \dots, r'_i, \dots, r_n)$ , and thus the specific steps are omitted. The above procedure can be briefly summarized as in the following figure. After getting the unconditional and conditional samples,  $S_i$  and  $S_i^{tot}$  can be computed according to Equations (14) and (15). The solution process is concluded as Figure 3.

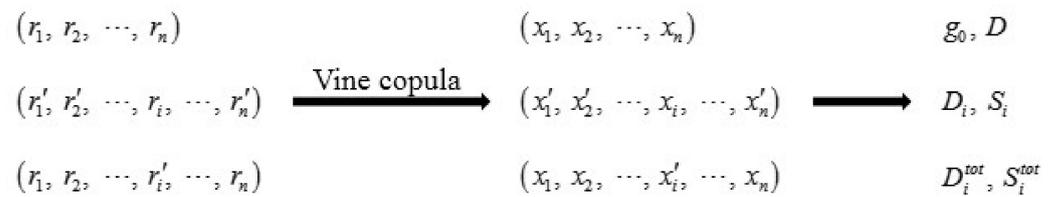


Figure 3. The solution process of VC1.

4. Test Cases

In this section, the developed approach is illustrated by five different models with multidimensional correlated input variables.

The first test with lognormal distributed variables’ numerical solutions is computed and compared with the reference results. We also compare the solutions obtained by the two copula-based methods and discuss the effect of different Vine copulas on the results. Test 2 compares the results obtained by the developed approach and traditional Nataf model by analyzing a four-dimensional dependent model. The third case’s variables are uniformly distributed and the influence of Kendall  $\tau$  on the results is analyzed. The last two tests are engineering cases with complex model functions and correlations, in which Test 5 combines the proposed method and the finite element models.

4.1. Test 1. Example with Complete Probability Information

Consider a model function

$$g(\mathbf{X}) = X_1X_2 - X_3^2 \tag{18}$$

where the input variables are normally distributed variables with  $\mu = (1, 1.4, 1.6)$ ,  $\sigma = (0.3, 0.2, 0.1)$ . The linear correlation coefficients are  $\rho_{12} = 0.3$ ,  $\rho_{13} = 0.5$ ,  $\rho_{23} = 0.8$ , so

that the covariance matrix is  $Cov = \begin{bmatrix} 0.09 & 0.018 & 0.015 \\ 0.018 & 0.04 & 0.016 \\ 0.015 & 0.016 & 0.01 \end{bmatrix}$ . According to the theory of

multivariate normal distribution [10], a sufficient number of samples  $X_N$  can be generated to accurately estimate the mean and variance of the output response.

Further, we assume the input variables to have lognormal distribution with parameters  $\mu$  and  $\sigma$ . By exponent transformation of the above normally distributed samples  $X_N$ , we can obtain the lognormal distributed samples  $X_L$  with unknown correlations, which are used for copula parameter estimation, and the optimal copula functions can be selected by the AIC method. The results are shown in Table 2, and the AIC values of the optimal copula are marked in bold.

Table 2. The results of parameter estimation and AIC method.

Copula	$\alpha_{12}$	$\alpha_{13}$	$\alpha_{23}$	$\alpha_{12 3}$	$\alpha_{13 2}$	$\alpha_{23 1}$
Gaussian	0.2999	0.5000	0.8000	-0.1925	0.4543	0.7868
Clayton	0.3298	0.6592	1.7598	0.0001	0.5717	1.6681
Gumbel	1.1995	1.4269	2.2462	1.0001	1.3649	2.1767
Frank	1.7998	3.2955	7.5508	0.0025	2.9129	7.2215
Copula	AIC <sub>12</sub> ( $\times 10^4$ )	AIC <sub>13</sub> ( $\times 10^4$ )	AIC <sub>23</sub> ( $\times 10^5$ )	AIC <sub>12 3</sub> ( $\times 10^3$ )	AIC <sub>13 2</sub> ( $\times 10^4$ )	AIC <sub>23 1</sub> ( $\times 10^5$ )
Gaussian	<b>-2.4711</b>	-7.5399	-2.6779	-9.8929	-6.0581	-2.5297
Clayton	<b>-1.7792</b>	-5.5229	-2.0079	0.0099	-4.4245	-1.8934
Gumbel	<b>-2.0010</b>	-6.5179	-2.4497	0.0174	-5.1750	-2.3091
Frank	<b>-2.2542</b>	-6.8754	-2.4359	0.0419	-5.5241	-2.3014

Table 2 shows that the optimal bivariate copula functions among all variables are Gaussian copula, so we use bivariate Gaussian copula to construct Vine copula. The results obtained by the two methods based on the Vine copula function are listed in Table 3, while the reference solutions obtained by exponential transformation and the approximate solutions based on the Nataf model are also listed in Table 3.

**Table 3.** Values of  $g_0$ ,  $D$ ,  $S_i$  and  $S_i^{tot}$ .

	$g_0$	$D$	$S_1$	$S_2$	$S_3$	$S_1^{tot}$	$S_2^{tot}$	$S_3^{tot}$
Reference	-13.0510	13.1197	0.3177	0.0271	0.1286	0.8123	0.1778	0.5647
Nataf ( $N = 2^{16}$ )	-13.0511	13.1298	0.3173	0.0277	0.1286	0.8114	0.1779	0.5640
VC1 ( $N = 2^{16}$ )	-13.0511	13.1304	0.3173	0.0279	0.1279	0.8122	0.1784	0.5641
VC2 ( $N = 2^{23}$ )	-13.0477	13.1498	0.3170	0.0134	0.0956	0.8058	0.1757	0.5612

Table 3 shows that the two Vine copula methods obtain almost the same results as reference solutions, which illustrates the feasibility and accuracy of the developed methods. For this case, Vine copula consists only of Gaussian copula functions, which means Vine copula is essentially the same as the Nataf model (multivariate Gaussian copula) for this model. Therefore, the results obtained by Vine copula methods and Nataf methods are very close.

Furthermore, we notice that the results obtained by VC2 need more sample points than VC1 but have less accuracy. The convergence curve of the results obtained by the two methods with the number of samples is shown in Figure 4. We can see that the convergence of VC1 is better than that of VC2, which is consistent with the previous conclusion of Song et al. [11]. This may be because VC2 needs to calculate the copula density function when calculating Sobol’ indices, which may lead to numerical errors and cause abrupt changes in the results. Therefore, we only discuss VC1 in the following tests.

The above method uses C-Vine to construct the joint PDF of input variables ( $X_1$  as the key variable), i.e., the two-dimensional variables required for sampling are  $(X_1, X_2)$ ,  $(X_1, X_3)$  and  $(X_2|X_1, X_3|X_1)$ . We change the key variable and use a different C-Vine structure to describe the joint distribution of variables. The values of  $g_0$  and  $D$  obtained by C-Vine copulas with different key variables are listed in Table 4. It can be seen that almost the same results can be obtained by different C-Vine copulas, which means C-Vine copulas based on different key variables have little effect on the mean and variance of the output response for this model.

**Table 4.** Values of  $g_0$  and  $D$  obtained by different C-Vine copulas ( $N = 2^{16}$ ).

Key Variable	Vine Structure	$g_0$	$D$
$X_1$	$(X_1, X_2), (X_1, X_3), (X_2 X_1, X_3 X_1)$	-13.0511	13.1304
$X_2$	$(X_1, X_2), (X_2, X_3), (X_1 X_2, X_3 X_2)$	-13.0509	13.1611
$X_3$	$(X_1, X_3), (X_2, X_3), (X_1 X_3, X_2 X_3)$	-13.0512	13.1345

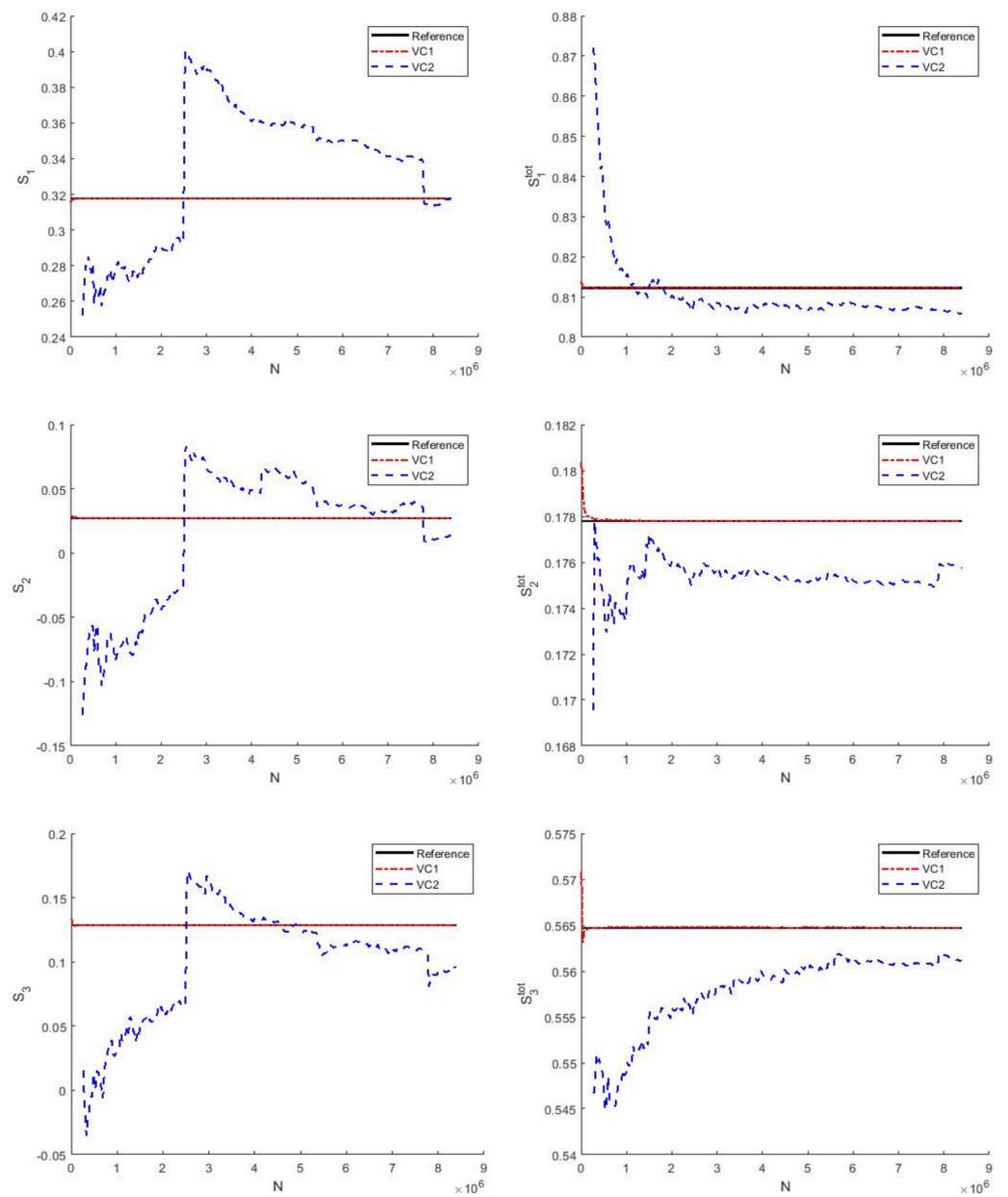


Figure 4. Convergence plots of the methods for  $S_i$  and  $S_i^{tot}$ .

4.2. Test 2. Portfolio Model

Consider the Portfolio model [10]:

$$g(\mathbf{X}) = X_1X_2 - X_3^2 \tag{19}$$

The input variables are normally distributed with  $\mu = (0, 0, 2.5, 4)$  and  $\sigma = (4, 2, 2, 3)$ . There are several known samples ( $N = 500$ ) in Figure 5, which shows all the variables are correlated.

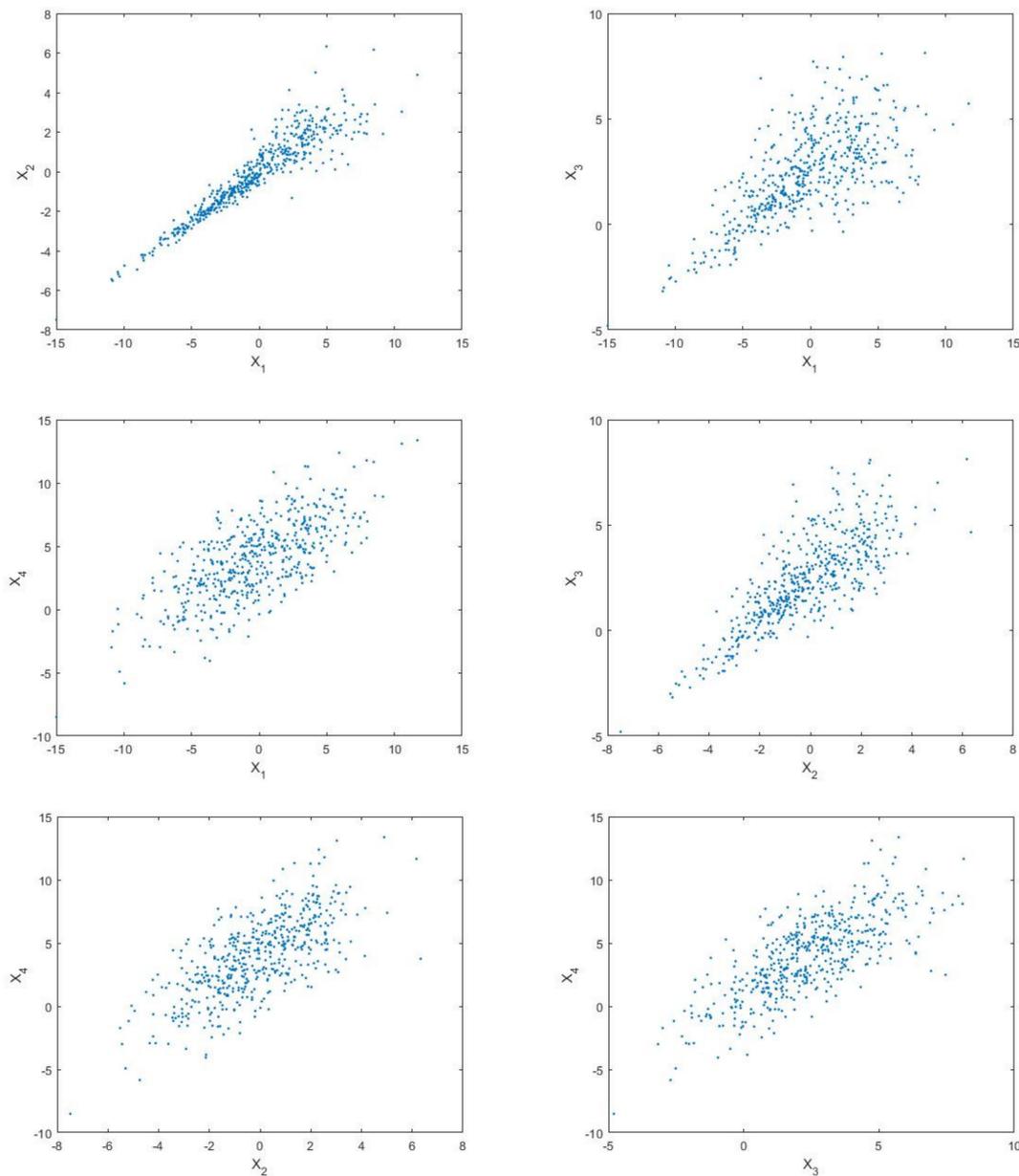


Figure 5. Sample distribution of input variables ( $N = 500$ ).

According to Equation (5), the four-dimensional input variables are decomposed as six two-dimensional variables:  $(X_1, X_2)$ ,  $(X_1, X_3)$ ,  $(X_1, X_4)$ ,  $(X_2|X_1, X_3|X_1)$ ,  $(X_2|X_1, X_4|X_1)$  and  $(X_3|X_1X_2, X_4|X_1X_2)$ . The AIC values of each alternative bivariate copula function are listed in Table 5, which were calculated by using the above 500 samples. We mark the AIC values of the optimal copula in bold and display the corresponding parameters in Table 5.

Table 5. Values of parameter estimation and AIC method ( $N = 500$ ).

	Gaussian	Clayton	Gumbel	Frank	Copula $\alpha$
$X_1, X_2$	-1033.3	-1364.7	-864.4	-855.9	8.0377
$X_1, X_3$	-324.6	-456.3	-229.5	-300.6	1.9490
$X_1, X_4$	-347.5	-264.8	-310.7	-294.0	0.7100
$X_2, X_3 X_1$	-224.7	-173.1	-193.0	-212.4	0.6038
$X_2, X_4 X_1$	-22.1	-15.0	-14.0	-24.0	1.3584
$X_3, X_4 X_1X_2$	-160.2	-117.7	-134.2	-162.2	3.6438

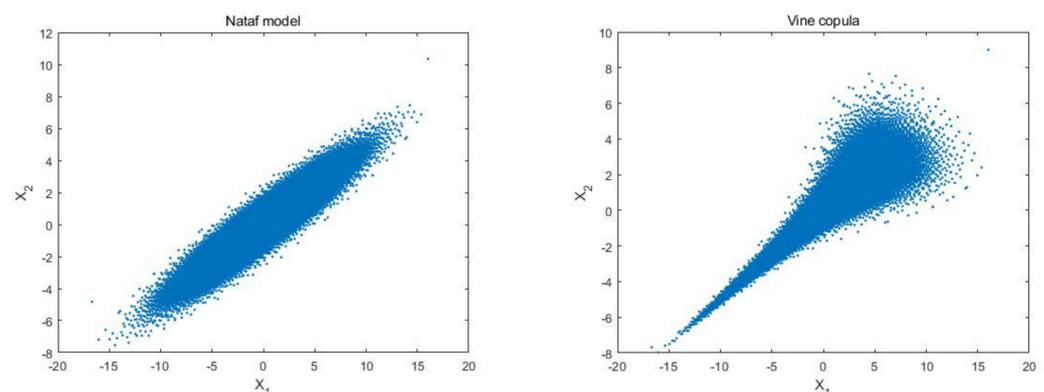
It can be seen that the optimal copula functions of two-dimensional variables are Clayton copula, Gaussian copula and Frank copula. Thus, there will be unknown errors if using the traditional Nataf method to measure the correlation. Table 6 gives the values obtained by the Nataf method and VC1.

**Table 6.** Values of  $g_0, D, S_i$  and  $S_i^{tot}$  ( $N = 2^{15}$ ).

	$g_0$	$S_1$	$S_2$	$S_3$	$S_4$
Nataf	9.7296	0.8338	0.8687	0.5843	0.5188
VC1	9.6056	0.7619	0.7702	0.3610	0.5877
	$D$	$S_1^{tot}$	$S_2^{tot}$	$S_3^{tot}$	$S_4^{tot}$
Nataf	581.1910	0.0294	0.0142	0.0285	0.0221
VC1	452.2249	0.0584	0.0311	0.0280	0.0279

The results obtained by the traditional Nataf method and the proposed Vine copula method are different, as well as the importance orders of input variables. According to the values of  $S_i$ , the Nataf method gives the importance order of input variables as  $X_2 > X_1 > X_3 > X_4$ , while  $X_2 > X_1 > X_4 > X_3$  is obtained if using the Vine copula method. Considering that an accurate importance order of variables is vital in practical problems, if we only use the conventional Nataf method for this problem, inaccurate results may be obtained, thus affecting the structural safety of a design.

Figure 6 shows the distribution of the samples  $(X_1, X_2)$  of the two methods. It can be seen that the samples generated by the Nataf method have symmetry but no tail correlation, while the Vine copula samples show a significantly lower tail correlation, which perfectly matches the distribution of variables  $X_1$  and  $X_2$  in Figure 5. This model illustrates that the tail correlation of variables has great impacts on Sobol’ indices. Since the traditional Nataf method cannot capture the tail correlation between variables, the estimated results of the Nataf method will have unknown errors if there is obvious tail correlation between the variables, which is also the reason for the large gap between the results of the Nataf method and the Vine copula method for this model.



**Figure 6.** Sample distribution generated by two methods  $((X_1, X_2), N = 2^{15})$ .

Furthermore, it can be seen that  $S_i^{tot} < S_i$  for this model, which means the correlations among input variables have greater effects than interactions on the variance of the output. Therefore, we should focus on the main indices  $S_i$  in practical engineering when analyzing correlations, especially for non-normally dependent and nonlinear problems.

4.3. Test 3. Ishigami Function

Consider the Ishigami function [10]:

$$g(\mathbf{X}) = X_1 X_2 - X_3^2 \tag{20}$$

with input variables being uniformly distributed:  $-\pi \leq X_i \leq \pi, i = 1, 2, 3$ . This function is widely used as a benchmark for sensitivity analysis since it has strong non-linearity and non-monotonicity.

We also analyzed this model in [11] and considered that only the variables  $X_1$  and  $X_3$  are dependent. In this paper, we assume that the three input variables are all correlated, and the distribution of 500 samples of input variables is shown in Figure 7. It can be seen that there is upper tail correlation between  $X_1$  and  $X_2$ , and lower tail correlation between  $X_1$  and  $X_3$ .

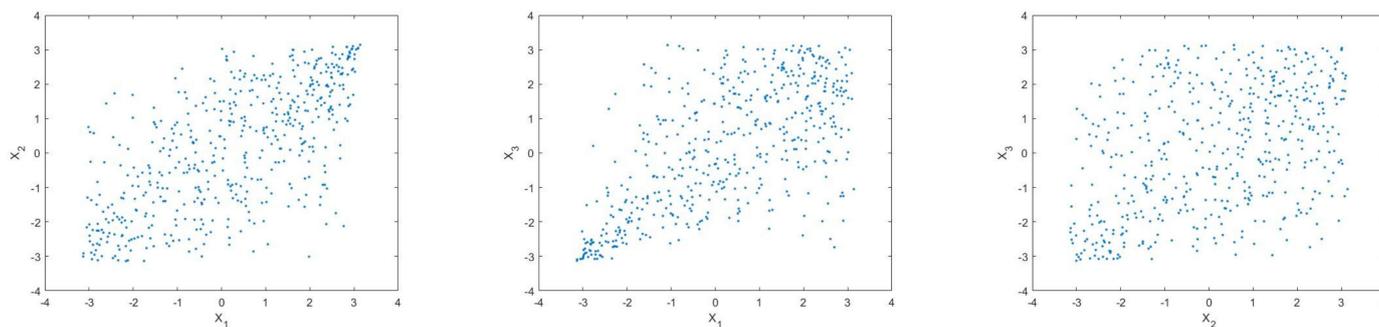


Figure 7. Sample distribution of input variables ( $N = 500$ ).

By using the above samples, the optimal copula functions and their parameters among the two-dimensional variables are selected by the AIC method, and the results are listed in Table 7 (where the AIC values of the optimal copula functions are in bold).

Table 7. Values of parameter estimation and AIC method ( $N = 500$ ).

	Gaussian	Clayton	Gumbel	Frank	Copula $\alpha$
$X_1, X_2$	-307.42	-194.53	<b>-318.29</b>	-294.48	1.9063
$X_1, X_3$	-255.70	<b>-367.33</b>	-185.25	-273.38	1.8789
$X_2, X_3$	-91.34	<b>-95.63</b>	-61.17	-90.73	0.6513
$X_2, X_3 X_1$	2.00	<b>1.77</b>	2.00	1.99	0.0192
$X_1, X_3 X_2$	-154.66	<b>-204.43</b>	-104.82	-155.01	1.1335
$X_1, X_2 X_3$	-169.16	-88.75	<b>-189.89</b>	-162.15	1.5972

The optimal copula functions between two-dimensional variables are mainly Clayton copula and Gumbel copula, the parameters of which are listed in Table 7. Table 8 gives the values obtained by the Nataf method and the proposed method. It is shown that the importance orders of variables obtained by the two methods are the same ( $X_2 > X_1 > X_3$ ), but the specific values of  $S_i$  and  $S_i^{tot}$  are different. This is because the Vine copula method proposed in this paper takes into account the tail correlation between variables, which cannot be measured by the Nataf method.

Table 8. Values of  $g_0, D, S_i$  and  $S_i^{tot}$  ( $N = 2^{14}$ ).

	$g_0$	$D$	$S_1$	$S_2$	$S_3$	$S_1^{tot}$	$S_2^{tot}$	$S_3^{tot}$
Nataf	3.5000	12.4260	0.3125	0.5714	0.2289	0.2468	0.4760	0.2129
VC1	3.5189	12.3998	0.3450	0.6018	0.2468	0.2400	0.4655	0.1887

We also analyze the influence of degree of correlation of input variables on the importance analysis results. It is assumed that the Kendall  $\tau$  between the decomposed two-dimensional variables is the same, and the optimal copula functions in the previous step are used. The proposed Vine copula method and Nataf method are used for analysis, and the independent situation of the three variables is considered at the same time. The curve of the calculated results changing with Kendall  $\tau$  is shown in Figure 8.

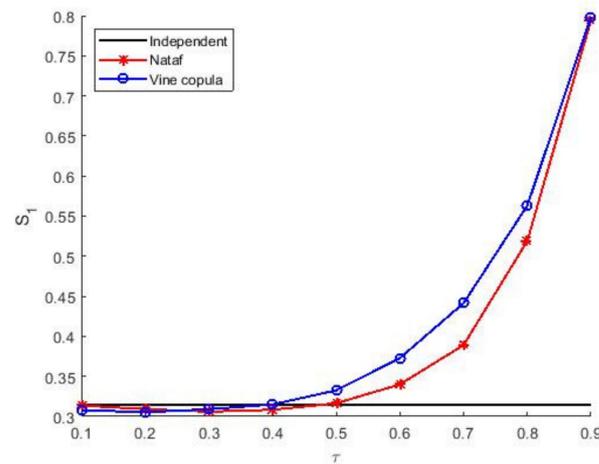


Figure 8. Values of  $S_1$  obtained by two methods versus  $\tau$ .

It can be seen that the results obtained by the two methods show the same trend with Kendall  $\tau$  and are significantly different from the results in the independent case. This shows that the correlation between variables has a great impact on the Sobol' indices. If the variables are simply assumed as independent variables for processing, this may cause a large error in the importance analysis and affect the safety of designs in engineering. Besides, the results obtained by the two methods are relatively close in the case of small and large Kendall  $\tau$ , while the results are significantly different when  $0.5 < \tau < 0.8$ . This indicates that the Vine copula method with better correlation measurement should be selected for analysis within the correlation range  $0.5 < \tau < 0.8$ .

#### 4.4. Test 4. Fatigue and Creep of Materials

Fatigue and the creep of materials are two common failure modes at high temperatures. Based on experimental data using the linear damage accumulation rule, Mao et al. [33] proposed a probabilistic model for reliability analysis of the creep and fatigue of materials as follows:

$$g(\mathbf{X}) = X_1 X_2 - X_3^2 \tag{21}$$

where  $D_c$  and  $D_f$  denote the creep damage and fatigue damage, respectively, and  $\theta_1$  and  $\theta_2$  are the parameters obtained from the experimental results. Two instances of damage can be expressed as  $D_c = \frac{n_c}{N_c}$  and  $D_f = \frac{n_f}{N_f}$ , where  $N_c$  and  $N_f$  correspond to the creep life and fatigue life, respectively, while  $n_c$  and  $n_f$  are the number of the creep and fatigue load cycles.

Assuming that the creep and fatigue lifetimes of a material follow log-normal distribution,  $\theta_1$  and  $\theta_2$  are normally distributed. The assumed values of the distribution parameters are shown in Table 9.

Table 9. Values and distribution parameters of input variables.

Input Variables	Mean $\mu$	Coefficient of Variation (COV)	Distribution
$N_c$	5490	0.20	Log-normal
$N_f$	17,100	0.20	Log-normal
$n_c$	549	0.20	Log-normal
$n_f$	6000	0.20	Log-normal
$\theta_1$	0.42	0.20	Normal
$\theta_2$	6.0	0.20	Normal

The input variables  $N_c$ ,  $N_f$ ,  $n_c$  and  $n_f$  are dependent according to the theory of material mechanics. There are several experimental data, the distributions of which are

shown in Figure 9. The sample values of  $D_c$  and  $D_f$  can be calculated by using these data. Figure 10 shows the curve of creep-fatigue failure.

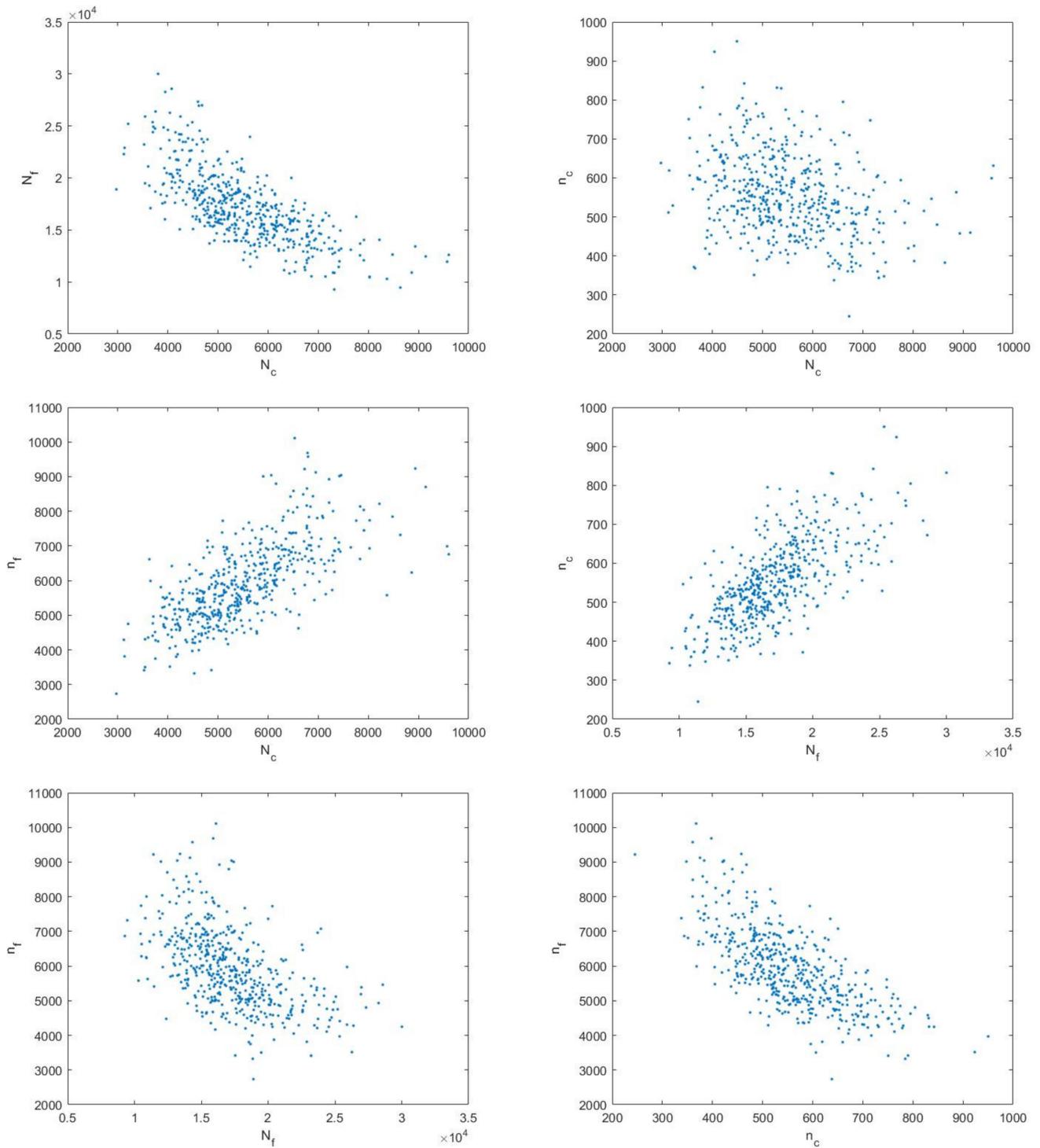


Figure 9. Sample distribution of input variables ( $N = 500$ ).

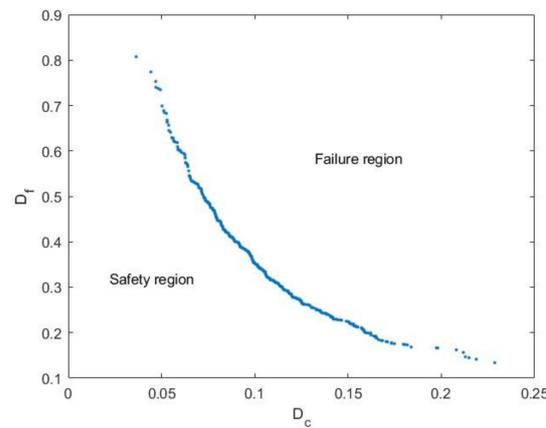


Figure 10. Creep-fatigue failure function.

Figure 10 shows that the relationship between creep damage and fatigue damage is nonlinear, i.e., the ultimate damage to the material is a nonlinear function of fatigue damage and creep damage, which means creep damage and fatigue damage affect each other. The contribution of each input variable to the limit state equation is analyzed. The samples in Figure 9 are statistically inferred to determine the optimal copula function and correlation parameters between the two-dimensional variables by the AIC method. The results are listed in Table 10.

Table 10. Values of parameter estimation and AIC method ( $N = 500$ ).

	Gaussian	Clayton	Gumbel	Frank	Copula $\alpha$
$N_c, N_f$	−386.7	2.0	2.0	−368.1	−0.7364
$N_c, n_c$	−48.8	2.0	2.0	−52.3	−2.1534
$N_c, n_f$	−331.7	−239.8	−292.5	−334.8	6.0953
$N_f, n_c   N_c$	−354.6	−249.8	−336.4	−351.0	0.7132
$N_f, n_f   N_c$	−0.1	2.0	2.0	−1.8	−0.5304
$n_c, n_f   N_c N_f$	−1376.1	2.0	2.0	−1412.1	−27.3777

Figure 11 gives the results of variance-based global sensitivity indices with the Nataf method and Vine copula method. It can be seen that the results obtained by the two methods are relatively similar, and the same importance order of variables is obtained. According to the values of  $S_i$  in Figure 11a, the fatigue life  $N_f$  and the fatigue load cycle  $n_f$  have greater influences on the limit state function than  $N_c$  and  $n_c$ , which is consistent with the AIC values in Table 10. This means the effects of input variables can be predicted preliminarily after the correlation estimation by the AIC method. For the parameters in the experiment, the influence of  $\theta_1$  is greater than  $\theta_2$ . We can conclude that the fatigue life and fatigue load cycle of materials most affect the creep-fatigue failure of the structure.

In addition, the importance measurement results obtained by the two methods are different since the Vine copula method uses the optimal copula (Gaussian copula and Frank copula, respectively) to measure the correlation between variables, while the traditional Nataf method only uses a single Gaussian copula to compute.

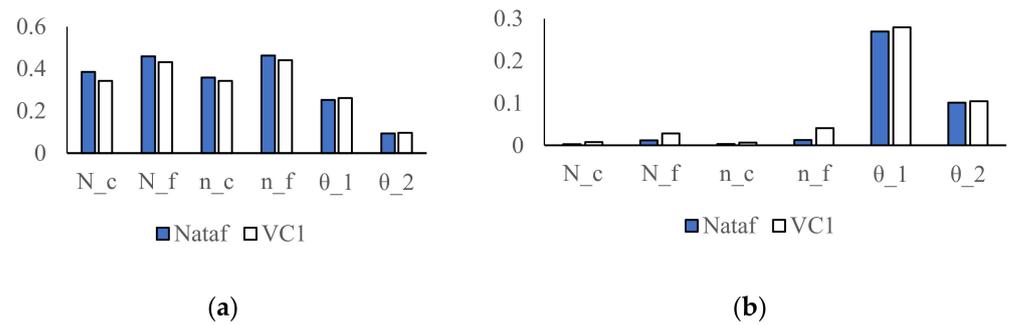


Figure 11. The values of Sobol' indices; (a)  $S_i$ , (b)  $S_i^{tot}$ .

4.5. Test 5. Three-Bay Five-Story Linear Elastic Frame Structure

Consider a three-bay five-story linear elastic frame structure [34,35], as shown in Figure 12, which is characterized by 21 random variables: 3 applied loads, 2 Young's moduli, 8 moments of inertia and 8 cross-sectional areas. The properties of the frame structures and the distribution types and parameters of random variables are given in Tables 11 and 12, respectively. Statistical dependence among the basic variables is considered. All loads are correlated with linear correlation coefficient  $\rho = 0.5$ , while the cross-sectional area and the moment of inertia of each element type are correlated with  $\rho = 0.95$ . All cross-sectional properties are correlated as

$$g(\mathbf{X}) = X_1X_2 - X_3^2 \tag{22}$$

We consider the horizontal displacement at node 1 as output function:

$$g(\mathbf{X}) = X_1X_2 - X_3^2 \tag{23}$$

where  $u_X$  denotes the actual horizontal displacement as an implicit function of all basic variables.

According to engineering qualitative analysis, input variables  $P_1, P_2, P_3, E_4, A_{18}, A_{19}$  and  $A_{20}$  contribute more than 95% to the output variance. Thus, we mainly analyze the importance orders of these seven input variables. In the following, the frame structure is modeled and analyzed by ANSYS with element type BEAM188. Several samples ( $N = 2^{15}$ ) are generated using the Nataf model with the linear correlation coefficients. According to these samples, Gaussian copula is selected as the optimal bivariate copula of the Vine structure through the AIC method, and the corresponding parameters are computed as well (given in Appendix B). The values of Sobol' indices given by the Nataf method and Vine copula method are listed in Table 13.

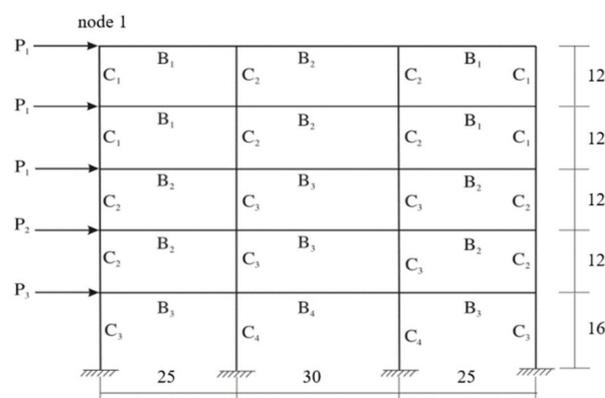


Figure 12. Three-bay five-story linear elastic frame structure.

**Table 11.** Properties of the frame structures.

Element	Young's Modulus	Moment of Inertia	Cross Section Area
B <sub>1</sub>	E <sub>4</sub>	I <sub>10</sub>	A <sub>18</sub>
B <sub>2</sub>	E <sub>4</sub>	I <sub>11</sub>	A <sub>19</sub>
B <sub>3</sub>	E <sub>4</sub>	I <sub>12</sub>	A <sub>20</sub>
B <sub>4</sub>	E <sub>4</sub>	I <sub>13</sub>	A <sub>21</sub>
C <sub>1</sub>	E <sub>5</sub>	I <sub>6</sub>	A <sub>14</sub>
C <sub>2</sub>	E <sub>5</sub>	I <sub>7</sub>	A <sub>15</sub>
C <sub>3</sub>	E <sub>5</sub>	I <sub>8</sub>	A <sub>16</sub>
C <sub>4</sub>	E <sub>5</sub>	I <sub>9</sub>	A <sub>17</sub>

**Table 12.** Statistical properties of random variables.

Variables	Distribution	Mean	SD	Variables	Distribution	Mean	SD
P <sub>1</sub>	Rayleigh	30	9	I <sub>12</sub>	Normal	2.69	0.52
P <sub>2</sub>	Rayleigh	20	8	I <sub>13</sub>	Normal	3.00	0.60
P <sub>3</sub>	Rayleigh	16	6.40	A <sub>14</sub>	Normal	3.36	0.48
E <sub>4</sub>	Normal	454,000	40,000	A <sub>15</sub>	Normal	4.00	0.64
E <sub>5</sub>	Normal	497,000	40,000	A <sub>16</sub>	Normal	5.44	0.80
I <sub>6</sub>	Normal	0.94	0.10	A <sub>17</sub>	Normal	6.00	0.96
I <sub>7</sub>	Normal	1.33	0.12	A <sub>18</sub>	Normal	2.72	0.80
I <sub>8</sub>	Normal	2.47	0.24	A <sub>19</sub>	Normal	3.13	0.88
I <sub>9</sub>	Normal	3.00	0.28	A <sub>20</sub>	Normal	4.01	1.04
I <sub>10</sub>	Normal	1.25	0.24	A <sub>21</sub>	Normal	4.50	1.16
I <sub>11</sub>	Normal	1.63	0.32				

SD: Standard deviation.

**Table 13.** Values of  $g_0$ ,  $D$ ,  $S_i$  and  $S_i^{tot}$  ( $N = 2^{15}$ ).

	$g_0$	$S_{P_1}$	$S_{P_2}$	$S_{P_3}$	$S_{E_4}$	$S_{A_{18}}$	$S_{A_{19}}$	$S_{A_{20}}$
Nataf	0.1149	0.4395	0.1473	0.1247	0.0112	0.0306	0.0751	0.2789
VC1	0.1148	0.4344	0.1523	0.1322	0.0116	0.0416	0.0819	0.2796
	$D$	$S_{P_1}^{tot}$	$S_{P_2}^{tot}$	$S_{P_3}^{tot}$	$S_{E_4}^{tot}$	$S_{A_{18}}^{tot}$	$S_{A_{19}}^{tot}$	$S_{A_{20}}^{tot}$
Nataf	0.0071	0.3132	0.0053	0.0005	0.0006	0.0005	0.0120	0.0617
VC1	0.0071	0.3137	0.0052	0.0005	0.0015	0.0004	0.0115	0.0638

It can be seen that two methods can get the same importance ranking of input variables. According to the main indices, the load  $P_1$  has the largest contribution to the uncertainty of the output of the frame structure, and the importance of  $P_2$  is larger than  $P_3$ , which is consistent with the conclusion of the engineering qualitative analysis. In all interface properties, the cross-section area  $A_{20}$  of element B<sub>3</sub> contributes the most to the uncertainty of the output. Besides, the total effects of input variables are generally smaller than the main effects, which indicates that the correlations among variables have greater effects on the output than the interacted effects. Therefore, the correlations among variables must be considered in importance analysis in engineering.

Comparing the results of the two methods, we can see that the values of the mean and variance of the output as well as Sobol' indices are almost the same. This is because the optimal bivariate copula functions of Vine structure are all Gaussian copula functions, which means the Vine copula method is essentially equivalent to the Nataf method under these circumstances.

### 5. Conclusions

Two novel Vine copula approaches (VC1 and VC2) for estimating variance-based global sensitivity indices for models with multidimensional correlated input variables are presented. VC1 generates dependent unconditional and conditional samples according to copula instead of the real joint CDF, while VC2 solves the indices respectively by

decomposing the joint PDF into independent and dependent parts. The approaches provide a new method for variance-based GSA of multidimensional cases with different correlations. By introducing Vine copula, the complex multidimensional related cases are transformed into two-dimensional related problems, and the optimal copula functions are used to measure different correlations separately. In the general case of arbitrary distributions with unknown multidimensional correlations, the proposed methods have a wider application than the traditional correlation analysis methods, such as the multi-Gaussian copula method (Nataf model) or multi-Archimedean copula method. VC1 is mainly used in five different test functions for testing and comparison. The results show the feasibility and accuracy of the proposed methods compared with the Nataf model. Through the calculation of the AIC method, the number of decomposed items of the joint PDF can be reduced preliminarily, which can help to improve the computational efficiency of reliability and sensitivity analyses based on the joint PDF for the further work.

**Author Contributions:** Conceptualization, S.S.; methodology, S.S. and Z.B.; software, Z.B.; formal analysis, Z.B.; investigation, H.W.; validation, H.W. and Y.X.; writing—original draft preparation, Z.B.; writing—review and editing, S.S. and S.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the anonymous reviewers for their constructive suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

The basic idea of VC2 is based on the properties of copula [29] that copula can divide the joint PDF  $f(\mathbf{x})$  into two parts: the independent part and the dependent part.

$$f(\mathbf{x}) = \frac{\partial^d C(u_1, u_2, \dots, u_d)}{\partial u_1 \partial u_2 \dots \partial u_d} \prod_{i=1}^d f(x_i) = c(u_1, u_2, \dots, u_d) \prod_{i=1}^d f(x_i) \tag{A1}$$

where  $c(u_1, u_2, \dots, u_d)$  is copula density function and  $f(x_i)$  is the marginal PDF of  $x_i$ . The independent part can be measured by the product of marginal PDF  $\prod_{i=1}^d f(x_i)$ , while the dependent part can be described by copula density function  $c(u_1, u_2, \dots, u_d)$ . Therefore, we generate independent samples according to marginal PDF, and calculate the independent and dependent parts respectively. Equations (12) and (13) can be formulated as

$$S_i = \frac{\frac{1}{N} \sum_{j=1}^N g(\mathbf{X}^{(j)}) c(\mathbf{u}^{(j)}) \left( g(\bar{X}_1^{(j)}, \dots, X_i^{(j)} \dots \bar{X}_n^{(j)}) c(\bar{u}_1^{(j)}, \dots, u_i^{(j)} \dots \bar{u}_n^{(j)}) - g(\mathbf{X}^{(j)}) c(\mathbf{u}^{(j)}) \right)}{D} \tag{A2}$$

$$S_i^{tot} = \frac{\frac{1}{N} \sum_{j=1}^N \left( g(\mathbf{X}^{(j)}) - g(X_1^{(j)}, \dots, \bar{X}_i^{(j)} \dots X_n^{(j)}) \right)^2 c(\mathbf{u}^{(j)}) \frac{c(u_1^{(j)}, \dots, \bar{u}_i^{(j)} \dots u_n^{(j)})}{c(u_1^{(j)}, \dots, u_{i-1}^{(j)} u_{i+1}^{(j)} \dots u_n^{(j)})}}{2D} \tag{A3}$$

Unlike Equations (14) and (15), the variable samples in Equations (A2) and (A3) are independent samples generated according to marginal PDF. We can see that Equations (A2) and (A3) consist in weighting the different terms of the sum in the MC estimator formula with the copula density function values at the corresponding data point. The basic idea

of this method is analogous to the one on which importance sampling (IS) estimators [36] are based to reduce the variance of MC estimators (the weighting by the copula density function can be seen as a likelihood ratio). The detailed procedure is as follows:

- i. generate independent samples  $(X_1, X_2, \dots, X_n)$  and  $(X'_1, X'_2, \dots, X'_n)$  according to the marginal PDFs;
- ii. construct conditional samples  $(X'_1, \dots, X'_i, \dots, X'_n)$  and  $(X_1, \dots, X'_i, \dots, X_n)$ ;
- iii. compute the values of marginal CDF using the above samples and get  $(u_1, u_2, \dots, u_n)$ ,  $(u'_1, u'_2, \dots, u'_n)$ ,  $(u'_1, \dots, u'_i, \dots, u'_n)$  and  $(u_1, \dots, u'_i, \dots, u_n)$ ;
- iv. compute the values of output function and copula density function, and get  $S_i$  and  $S_i^{tot}$  according to Equations (19) and (20).

### Appendix B

**Table A1.** Parameters of the Optimal Bivariate Copula Functions, test 5.

Variables	Copula $\alpha$
$P_1 P_2, P_1 P_3$	0.5000
$P_2 P_3   P_1$	0.3332
$E_4 E_5$	0.9000
$I_6 I_{7-13}, I_6 I_{15-21}$	0.1300
$I_6 A_{14}$	0.9500
$I_7 I_{8-13}   I_6, I_7 A_{16-21}   I_6$	0.1148
$I_7 A_{14}   I_6$	0.0207
$I_7 A_{15}   I_6$	0.9491
$I_8 I_{9-13}   I_6 I_7, I_8 A_{17-21}   I_6 I_7$	0.1030
$I_8 A_{14-15}   I_6 I_7$	0.0184
$I_8 A_{16}   I_6 I_7$	0.9484
$I_9 I_{10-13}   I_6 I_7 I_8, I_9 A_{18-21}   I_6 I_7 I_8$	0.0933
$I_9 A_{14-16}   I_6 I_7 I_8$	0.0166
$I_9 A_{17}   I_6 I_7 I_8$	0.9478
$I_{10} I_{11-13}   I_6 I_7 I_8 I_9, I_{10} A_{19-21}   I_6 I_7 I_8 I_9$	0.0853
$I_{10} A_{14-17}   I_6 I_7 I_8 I_9$	0.0152
$I_{10} A_{18}   I_6 I_7 I_8 I_9$	0.9474
$I_{11} I_{12-13}   I_6 I_7 I_8 I_9 I_{10}, I_{11} A_{20-21}   I_6 I_7 I_8 I_9 I_{10}$	0.0785
$I_{11} A_{14-18}   I_6 I_7 I_8 I_9 I_{10}$	0.0139
$I_{11} A_{19}   I_6 I_7 I_8 I_9 I_{10}$	0.9470
$I_{12} I_{13}   I_6 I_7 I_8 I_9 I_{10} I_{11}, I_{12} A_{21}   I_6 I_7 I_8 I_9 I_{10} I_{11}$	0.0728
$I_{12} A_{14-19}   I_6 I_7 I_8 I_9 I_{10} I_{11}$	0.0129
$I_{12} A_{20}   I_6 I_7 I_8 I_9 I_{10} I_{11}$	0.9467
$I_{13} A_{14-20}   I_6 I_7 I_8 I_9 I_{10} I_{11} I_{12}$	0.0120
$I_{13} A_{21}   I_6 I_7 I_8 I_9 I_{10} I_{11} I_{12}$	0.9464
Other conditional variables	0.0020

### References

1. Sobol, I.M.; Kucherenko, S. Global sensitivity indices for nonlinear mathematical models. Review. *Wilmott* **2005**, *1*, 56–61. [\[CrossRef\]](#)
2. Saltelli, A.; Annoni, P.; Azzini, I.; Campolongo, F.; Ratto, M.; Tarantola, S. Variance based sensitivity analysis of model output: Design and estimator for the total sensitivity index. *Comput. Phys. Commun.* **2010**, *181*, 259–270. [\[CrossRef\]](#)
3. Kala, Z. Sensitivity analysis in probabilistic structural design: A comparison of selected techniques. *Sustainability* **2020**, *12*, 4788. [\[CrossRef\]](#)
4. Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280. [\[CrossRef\]](#)
5. Li, G.; Rosenthal, C.; Rabitz, H. High Dimensional Model Representations. *J. Phys. Chem. A* **2001**, *105*, 7765–7777. [\[CrossRef\]](#)
6. Kala, Z. Global sensitivity analysis of reliability of structural bridge system. *Eng. Struct.* **2019**, *194*, 36–45. [\[CrossRef\]](#)
7. Xu, C.; Gertner, G.Z. Uncertainty and sensitivity analysis for models with correlated parameters. *Reliab. Eng. Syst. Saf.* **2008**, *93*, 1563–1573. [\[CrossRef\]](#)
8. Li, G.; Rabitz, H.; Yelvington, P.E.; Oluwole, O.O.; Bacon, F.; Kolb, C.E.; Schoendorf, J. Global Sensitivity Analysis for Systems with Independent and/or Correlated Inputs. *J. Phys. Chem. A* **2010**, *114*, 6022–6032. [\[CrossRef\]](#)

9. Kala, Z. Global sensitivity analysis based on entropy: From differential entropy to alternative measures. *Entropy* **2021**, *23*, 778. [[CrossRef](#)]
10. Kucherenko, S.; Tarantola, S.; Annoni, P. Estimation of global sensitivity indices for models with dependent variables. *Comput. Phys. Commun.* **2012**, *183*, 937–946. [[CrossRef](#)]
11. Song, S.; Bai, Z.; Wei, H.; Xiao, Y.; Kucherenko, S. Variance-based importance measure analysis based on copula under incomplete probability information. *Probabilistic Eng. Mech.* **2021**, submitted for publication.
12. Der Kiureghian, A.; Liu, P. Structural Reliability under Incomplete Probability Information. *J. Eng. Mech.* **1986**, *112*, 85–104. [[CrossRef](#)]
13. Lebrun, R.; Dutfoy, A. An innovating analysis of the Nataf transformation from the viewpoint of copula. *Probabilistic Eng. Mech.* **2009**, *24*, 312–320. [[CrossRef](#)]
14. Eryilmaz, S. Multivariate copula based dynamic reliability modeling with application to weighted-k-out-of-n systems of dependent components. *Struct. Saf.* **2014**, *51*, 23–28. [[CrossRef](#)]
15. Daul, S.; De Giorgi, E.G.; Lindskog, F.; McNeil, A. The Grouped t-Copula with an Application to Credit Risk. *Soc. Sci. Electron. Publ.* **2009**, *16*, 11. [[CrossRef](#)]
16. Cossette, H.; Côté, M.P.; Marceau, E.; Moutanabbir, K. Multivariate distribution defined with Farlie-Gumbel-Morgenstern copula and mixed Erlang marginals: Aggregation and capital allocation. *Insur. Math. Econ.* **2013**, *52*, 560–572. [[CrossRef](#)]
17. Gao, K.; Liu, G. Novel nonlinear time-varying fatigue reliability analysis based on the probability density evolution method. *Int. J. Fatigue* **2021**, *149*, 106257. [[CrossRef](#)]
18. Fischer, S.; Schumann, A.H. Multivariate flood frequency analysis in large river basins considering tributary impacts and flood types. *Water Resour. Res.* **2021**, *57*, e2020WR029029. [[CrossRef](#)]
19. Jiang, C.; Zhang, W.; Han, X.; Ni, B.Y.; Song, L.J. A Vine-Copula-Based Reliability Analysis Method for Structures with Multidimensional Correlation. *J. Mech. Des.* **2015**, *137*, 061405. [[CrossRef](#)]
20. Bedford, T.; Cooke, R.M. Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann. Math. Artif. Intell.* **2001**, *32*, 245–268. [[CrossRef](#)]
21. Bedford, T.; Cooke, R.M. Vines: A new graphical model for dependent random variables. *Ann. Stat.* **2002**, *30*, 1031–1068. [[CrossRef](#)]
22. Kurowicka, D.; Joe, H. *Dependence Modelling: Vine Copula Handbook*; World Scientific: Singapore, 2011.
23. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
24. Genest, C.; Verret, F. Locally most powerful rank tests of independence for copulas model. *J. Nonparametric Stat.* **2005**, *17*, 521–535. [[CrossRef](#)]
25. Huard, D.; Evin, G.; Favre, A. Bayesian copula selection. *Comput. Stat. Data Anal.* **2006**, *51*, 809–822. [[CrossRef](#)]
26. Torre, E.; Marelli, S.; Embrechts, P.; Sudret, B. A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas. *Probabilistic Eng. Mech.* **2019**, *55*, 1–16. [[CrossRef](#)]
27. Benoumechiara, N.; Bousquet, N.; Michel, B.; Saint-Pierre, P. Detecting and modeling critical dependence structures between random inputs of computer models. *Depend. Modeling* **2020**, *8*, 263–297. [[CrossRef](#)]
28. Sarazin, G.; Morio, J.; Lagnoux, A.; Balesdent, M.; Brevault, L. Reliability-oriented sensitivity analysis in presence of data-driven epistemic uncertainty. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107733. [[CrossRef](#)]
29. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer: New York, NY, USA, 2006.
30. Tarantola, S.; Mara, T.A. Variance-based sensitivity indices of computer models with dependent inputs: The Fourier Amplitude Sensitivity Test. *Int. J. Uncertain. Quantif.* **2017**, *7*, 511–523. [[CrossRef](#)]
31. Aas, K.; Czado, C.; Frigessi, A.; Bakken, H. Pair-copula constructions of multiple dependence. *Insur. Math. Econ.* **2009**, *44*, 182–198. [[CrossRef](#)]
32. Rosenblatt, M. Remarks on a Multivariate Transformation. *Ann. Math. Stat.* **1952**, *23*, 470–472. [[CrossRef](#)]
33. Mao, H.; Mahadevan, S. Reliability analysis of creep–fatigue failure. *Int. J. Fatigue* **2000**, *22*, 789–797. [[CrossRef](#)]
34. Guan, X.L.; Melchers, R.E. Effect of response surface parameter variation on structural reliability estimates. *Struct. Saf.* **2001**, *23*, 429–444. [[CrossRef](#)]
35. Kaymaz, I. Application of kriging method to structural reliability problems. *Struct. Saf.* **2005**, *27*, 133–151. [[CrossRef](#)]
36. Au, S.K.; Beck, J.L. Importance sampling in high dimensions. *Struct. Saf.* **2002**, *25*, 139–163. [[CrossRef](#)]