Distribution-Free Stochastic Model Updating with Copulas and Staircase Density Functions

Masaru Kitahara
Assistant Professor, Dept. of Civil Engineering, The University of Tokyo, Tokyo, Japan

Matteo Broggi
Deputy Head, Institute for Risk and Reliability, Leibniz University Hannover, Hannover, Germany

Takeshi Kitahara
Professor, Dept. of Civil Engineering, Kanto Gakuin University, Yokohama, Japan

Michael Beer
Professor, Institute for Risk and Reliability, Leibniz University Hannover, Hannover, Germany

ABSTRACT: In stochastic model updating, a probabilistic model is assumed for the model parameters and its hyperparameters (e.g., means and variances) are calibrated to minimize the stochastic discrepancy between model outputs and measurements. Thus, if the assumption about the probabilistic model is inappropriate, it may introduce a bias in the updating results. To avoid such inappropriate assumptions, we have recently developed a distribution-free approach that uses the staircase density function (SDF) to arbitrarily approximate a wide range of probability distributions. In this study, we aim to extend this approach to the calibration of dependent parameters. The dependence structure is represented by several types of copulas and the marginal distributions are modeled using the SDFs. An approximate Bayesian computation framework is then developed to calibrate the copula parameter as well as the SDF hyperparameters. Furthermore, in this framework, the most appropriate copula class is determined in the context of Bayesian model class selection.

1. INTRODUCTION

Stochastic model updating has attracted increasing attention as a fascinating technique to mitigate the discrepancy between model outputs and experimental measurements, taking into account uncertainties in both the modeling and measurement processes (Mares et al., 2006). To adequately deal with such uncertainties, the uncertainty quantification (UQ) metric plays an important role in stochastic model updating. Among the different types of UQ metrics, the Bhattacharyya distance has been shown to be a promising choice to quantitatively measure the stochastic discrepancy between model outputs and measurements (Bi et al., 2019).

On the other hand, in stochastic model updating, it is also important to assume an appropriate probabilistic model for the model parameters to be calibrated. In this context, the use of the staircase density function (SDF) has recently been investigated by the first author and his co-workers (Kitahara et al., 2022a; Kitahara et al., 2022b). The SDF is a discrete probability distribution that can arbitrarily approximate a wide range of bounded distributions given its four hyperparameters, namely the mean, variance, and third- and fourth-order central moments (Crespo et al., 2018). Kitahara et al. (2022a) have developed an approximate Bayesian computation (ABC) model updating framework, in which an approximate likelihood function is defined using the Bhattacharyya distance metric and the model parameters are then calibrated by inferring the posterior of the SDF hyperparameters.

While the above updating framework is based on the independent assumption among the model parameters, Kitahara et al. (2022b) have extended it to calibrate the dependent parameters.

Table 1: Summary of the copula functions employed in this study.

Copula	Copula function $C(\cdot;\theta)$	Range of θ	Prior distribution of θ
Gaussian	$\Phi_\theta\left(\Phi^{-1}(u_1),\Phi^{-1}(u_2)\right)$	$[-1\ 1]$	Truncated standard normal
Clayton	$\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-\frac{1}{\theta}}$	$(0\ \infty)$	Gamma ($a=2$, $b=1$)
Frank	$-\frac{1}{\theta}\log\left[1+\frac{\left(e^{-\theta u_1}-1\right)\left(e^{-\theta u_2}-1\right)}{e^{-\theta}-1}\right]$	$(-\infty\ \infty)$	Gaussian ($\mu=0$, $m_2=9$)
Gumbel	$\exp\left[-\left((-\log u_1)^\theta+(-\log u_2)^\theta\right)^{\frac{1}{\theta}}\right]$	$(1\ \infty)$	Gamma ($a=2$, $b=1$) for $\theta-1$

Note:Φ_θ is the bivariate standard normal CDF with the correlation coefficient θ; Φ is the standard normal CDF; a and b are the shape and scale parameters of Gamma distribution.

The dependence structure is characterized by a Gaussian copula function and its marginal distributions are represented as the SDFs. Thus, the correlation matrix in the Gaussian copula function is inferred as the posterior together with the SDF hyperparameters.

However, the Gaussian copula can only represent Gaussian dependence structures, and it may introduce a bias in the model calibration if the real dependence structure differs from this assumption. Several different types of bivariate copulas, such as Gaussian, Clayton, Frank, and Gumbel copulas, are commonly used to represent different types of dependence structures. In this study, we investigate the combination of these copulas with the marginal SDFs to calibrate the dependent parameters with different dependence structures. Furthermore, the most appropriate copula class is determined from the available measurement data in the context of Bayesian model class selection.

2. THEORIES AND METHODOLOGIES

2.1. Copula functions

Let the model parameters $\boldsymbol{x}=[x_1,x_2]$ be bivariate random variables. According to Sklar's theorem (Haff, 2013), the joint cumulative density function (CDF) of $\boldsymbol{x}$ is given by:

$$F(\boldsymbol{x})=C(F_1(x_1),F_2(x_2);\theta) \qquad (1)$$

where $u_i=F_i(x_i)$ $(i=1,2)$ is the marginal CDF of x_i; C is the copula function; θ is the copula parameter describing the dependence structure. The copula function is itself a joint CDF on $[0\ 1]^2$ with uniform marginal distributions on $[0\ 1]$.

There are many copulas in the literature, including the Gaussian, Clayton, Frank, and Gumbel copulas. Each copula is characterized by its own dependence structure. These four copula functions and the ranges of their parameter θ are summarized in Table 1.

2.2. Staircase density functions

Let x_i $(i=1,2)$ have the support set $\left[\underline{x}_i,\overline{x}_i\right]$ and the prescribed values for the hyperparameters $\boldsymbol{\theta}_{x_i}=[\mu_i,m_{2i},m_{3i},m_{4i}]$, where μ_i is the mean, m_{2i} is the variance, m_{3i} is the third-order central moment, and m_{4i} is the fourth-order central moment. Note that, in practice, m_{3i} and m_{4i} are normalized as skewness $\widetilde{m}_{3i}=m_{3i}/m_{2i}^{1.5}$ and the kurtosis $\widetilde{m}_{4i}=m_{4i}/m_{2i}^2$, respectively, in model updating. Any such variable needs to satisfy the feasible conditions $g\left(\boldsymbol{\theta}_{x_i}\right)\leq 0$ given in Crespo et al. (2018). The realizations of $\boldsymbol{\theta}_{x_i}$ satisfying these conditions constitute the feasible domain $\Theta_i=\left\{\boldsymbol{\theta}_{x_i}: g\left(\boldsymbol{\theta}_{x_i}\right)\leq 0\right\}$.

The CDF of x_i is then given as the SDF:

$$F_i(x_i)=\kappa\sum_{n=1}^{j} l_i^j\ \ \forall x_i\in(x_i^j\ x_i^{j+1}], \quad \forall j=1,\cdots,n_b \qquad (2)$$

where n_b is the number of bins; l_i^j is the density height at the jth bin; $x_i^j=\underline{x}_i+(j-1)\kappa$ with

$\kappa = (\overline{x}_i - \underline{x}_i)/n_b$. In this study, n_b is set to be $n_b = 25$. The density heights $\boldsymbol{l}_i$ are obtained by solving the optimization problem:

$$\min_{\boldsymbol{l}_i>0}\left\{J(\boldsymbol{l}_i): \mathbf{A}\left(\boldsymbol{\theta}_{x_i}, n_b\right)\boldsymbol{l}_i = \boldsymbol{b}\left(\boldsymbol{\theta}_{x_i}\right), \boldsymbol{\theta}_{x_i} \in \Theta_i\right\} \quad (3)$$

where J indicates the cost function; $\mathbf{A}\boldsymbol{l}_i = \boldsymbol{b}$ are moment matching constraints. In this study, the cost function is defined based on the principle of maximum entropy:

$$J(\boldsymbol{l}_i) = \kappa \log \boldsymbol{l}_i^T \boldsymbol{l}_i \quad (4)$$

2.3. Approximate Bayesian computation

Given the copula class $\mathcal{M}$ from its candidates listed in Table 1, ABC (Beaumont, 2019) is used to calibrate the model parameters $\boldsymbol{x}$ according to the well-known Bayes' theorem:

$$P(\boldsymbol{\vartheta}|\mathcal{M}, \mathbf{Y}_D) = \frac{\tilde{\mathcal{L}}(\mathbf{Y}_D|\boldsymbol{\vartheta}, \mathcal{M})P(\boldsymbol{\vartheta}|\mathcal{M})}{P(\mathbf{Y}_D|\mathcal{M})} \quad (5)$$

where $P(\boldsymbol{\vartheta}|\mathcal{M})$ is the prior distribution of $\boldsymbol{\vartheta} = \left[\theta\ \boldsymbol{\theta}_{x_1}\ \boldsymbol{\theta}_{x_2}\right]$ reflecting one's initial knowledge on $\boldsymbol{\vartheta}$; $P(\boldsymbol{\vartheta}|\mathcal{M}, \mathbf{Y}_D)$ is the posterior distribution of $\boldsymbol{\vartheta}$ reflecting the posterior state-of-knowledge on $\boldsymbol{\vartheta}$; $\tilde{\mathcal{L}}(\mathbf{Y}_D|\boldsymbol{\vartheta}, \mathcal{M})$ represents the so-called approximate likelihood function that serves as the connection between the available measurement data $\mathbf{Y}_D$ and $\boldsymbol{\vartheta}$; $P(\mathbf{Y}_D|\mathcal{M})$ is the evidence of $\mathcal{M}$ ensuring that the integral of the posterior distribution equals to one.

In this study, the inference parameters $\boldsymbol{\vartheta}$ are assumed to be independent of each other. The prior distribution of the copula parameter θ is defined for each copula function considering its range and is summarized in the last column of Table 1. In addition, the prior distribution of the hyperparameters $\boldsymbol{\theta}_{x_i}$ is assumed to be the uniform distribution over the feasible domain Θ_i.

Let $\mathbf{Y}_D = \left\{\boldsymbol{y}^{(i)}; i = 1, \cdots, N_D\right\}$ be N_D sets of the measurement quantities $\boldsymbol{y}$. The corresponding N_S sets of the model outputs can be obtained as $\mathbf{Y}_S = \left\{h\left(\boldsymbol{x}^{(i)}|\boldsymbol{\vartheta}, \mathcal{M}\right); i = 1, \cdots, N_S\right\}$, where h is the simulator. The Bhattacharyya distance can be then defined to measure the stochastic distance between $\mathbf{Y}_S$ and $\mathbf{Y}_D$:

$$d_B(\mathbf{Y}_S, \mathbf{Y}_D) = -\log\left\{\sum_{j=1}^{N_{bin}} \sqrt{P^j(\mathbf{Y}_S)P^j(\mathbf{Y}_D)}\right\} \quad (6)$$

where N_{bin} denotes the total number of bins and is set as $N_{bin} = 10^2$ in this study; $P^j(\cdot)$ denotes the probability mass function value at the jth bin. Therefore, the approximate likelihood function is expressed as:

$$\tilde{\mathcal{L}}(\mathbf{Y}_D|\boldsymbol{\vartheta}) \propto \exp\left\{\frac{d_B(\mathbf{Y}_S, \mathbf{Y}_D)^2}{\varepsilon^2}\right\} \quad (7)$$

where ε is the scaling parameter controlling the centralization degree of the posterior distribution and is set to be $\varepsilon = 0.02$ in this study.

Furthermore, the evidence $P(\mathbf{Y}_D|\mathcal{M})$ can be viewed as the plausibility measure of the copula class $\mathcal{M}$ given the measurement data $\mathbf{Y}_D$. Hence, in the context of Bayesian model class selection (Beck and Yuen, 2004), the most appropriate copula class that best represents $\mathbf{Y}_D$ can be determined as the one that provides the largest evidence value.

Finally, an analytical solution for the posterior distribution in Eq. (5) is generally not available and sampling methods are often used to estimate it. In this study, the transitional Markov chain Monte Carlo (TMCMC) algorithm (Ching and Chen, 2007) is used. TMCMC is a class of sequential Monte Carlo methods and samples from a series of intermediate distributions that will progressively converge to the true posterior distribution. In addition, the evidence is estimated as a byproduct. The reader is referred to Ching and Chen (2007) for details on TMCMC.

3. NUMERICAL EXAMPLES

3.1. Problem descriptions

The proposed stochastic updating framework is demonstrated using a three degree-of-freedom (DOF) spring-mass system shown in Figure 1. Three stiffness coefficients k_1, k_2, and k_3 are assumed as the model parameters to be calibrated and the prior knowledge on them is provided in Table 2. Here, k_2 follows a Gaussian distribution. Its mean and variance are not fully determined but fall within given intervals. Meanwhile, k_1 and k_3

Table 2: Prior knowledge on the model parameters.

Parameter	Uncertainty characteristics		Target values
	Distribution	Support set / hyperparameters	
k_2	Gaussian	$\mu_2 \in [3.0, 7.0]$, $m_{22} \in [0, 0.25]$	$\mu_2 = 5.0$, $m_2 = 0.01$
k_1, k_3	Unknown, dependent	$k_1 \in [2.5, 5.5]$, $k_3 \in [5.0, 7.0]$	Given in Tables 3 and 4
k_4–k_6, m_1–m_3	Deterministic	–	–

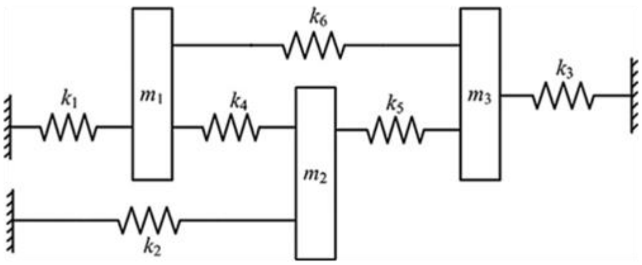

Figure 1: 3-DOF spring-mass system.

are dependent of each other within the given support sets, but their dependence structure and marginal distributions are unknown. Besides, the remaining parameters, three stiffness coefficients k_4 to k_6 and three masses m_1 to m_3, are set to be constants with determined values: $k_i = 5.0$ N/m ($i = 4, 5, 6$), $m_1 = 0.7$ kg, $m_2 = 0.5$ kg, and $m_3 = 0.3$ kg.

In total, four target joint distributions defined by different copulas are considered for k_1 and k_3. The target dependence structures and marginal distributions are summarized in Tables 3 and 4, respectively. While the marginal distributions are truncated, they cover more than 99.99 % of their original Gaussian distributions.

Table 3: Target dependence structures.

Copulas	Target value of θ
Gaussian	0.7
Clayton	2.0
Frank	-5.0
Gumbel	2.0

Table 4: Target marginal distributions.

Parameter	Distribution	Target value
k_1	Truncated Gaussian	$\mu_1 = 4.0$, $m_1 = 0.09$
k_3	Truncated Gaussian	$\mu_3 = 6.0$, $m_3 = 0.04$

The outputs of the system are three natural frequencies f_1, f_2, and f_3. The measurement data $\mathbf{Y}_D$ consists of these three natural frequencies and is generated by multiple model evaluations with multiple sets of the model parameters sampled from their target joint distributions. In this study, the number of data is set as $N_D = 3000$. Figure 2 shows the measurement data in the plane of f_1 and f_3 for each target copula class. It can be seen that different copula functions result in different data scatters. Therefore, the aim of model updating is to identify the most appropriate copula class that reproduces the uncertainty characteristics of the measurement data.

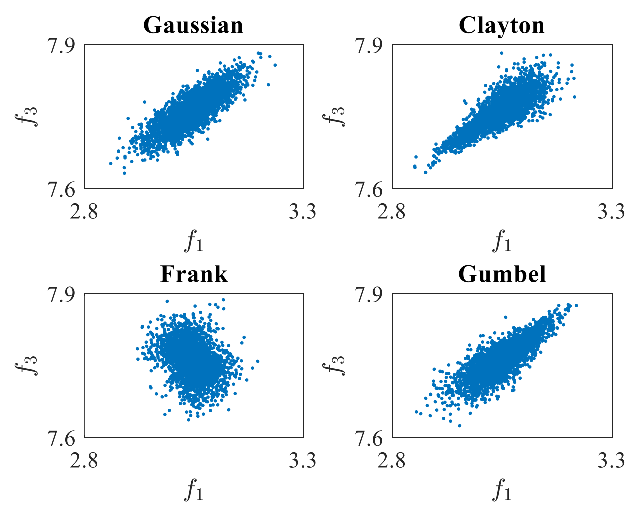

Figure 2: Measurement data with different target copula classes.

3.2. Model updating results

The TMCMC algorithm is employed to generate 1000 samples from the posterior distribution. In this section, the copula class $\mathcal{M}$ is chosen to be the same as the target copula class that is used to generate the measurement data $\mathbf{Y}_D$.

In the case where the Gaussian copula is used as the target dependence structure, a total of 16 iterations are performed to get converged. Figure 3 illustrates histograms of the posterior samples of 11 inferring parameters, i.e., μ_2, m_2, and $\boldsymbol{\vartheta}$. It can be observed that all the inferring parameters are significantly updated from the prior distribution and are converged around their target values.

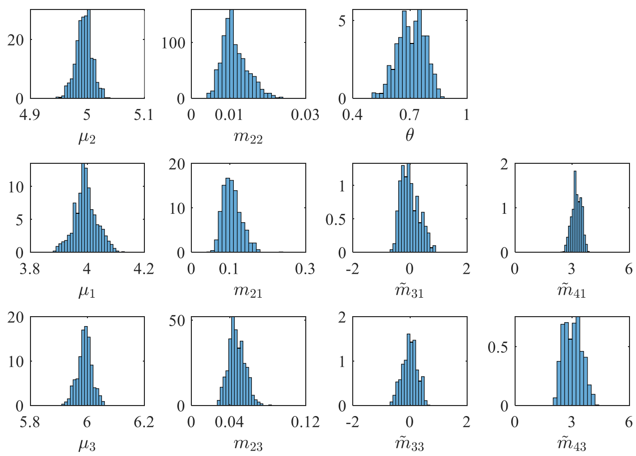

Figure 3: Posterior distribution of the inferring parameters in histograms.

The means of the posterior samples are then obtained as the posterior estimates of the inferring parameters. The same procedure is performed for all target copulas and the posterior estimates of the inferring parameters are summarized in Table 5. The percentage estimation errors are given in the parentheses after the posterior estimates.

Table 5: Posterior estimates of the inferring parameters.

(a) Gaussian copula

Inferring parameter	Target value	Posterior estimate
μ_2	5.0	4.99 (0.2)
m_{22}	0.01	0.0100 (0.0)
θ	0.7	0.707 (1.0)
μ_1	4.0	3.99 (0.3)
m_{21}	0.09	0.1063 (18.1)
$\tilde{m}_{31}$	0.0	0.002
$\tilde{m}_{41}$	3.0	3.23 (7.7)
μ_3	6.0	5.99 (0.2)
m_{23}	0.04	0.0475 (18.8)
$\tilde{m}_{33}$	0.0	-0.002
$\tilde{m}_{43}$	3.0	3.03 (1.0)

(b) Clayton copula

Inferring parameter	Target value	Posterior estimate
μ_2	5.0	5.01 (0.2)
σ_2	0.01	0.0104 (4.0)
θ	2.0	2.09 (4.5)
μ_1	4.0	3.99 (0.3)
m_{21}	0.09	0.1337 (48.6)
$\tilde{m}_{31}$	0.0	-0.133
$\tilde{m}_{41}$	3.0	5.69 (89.7)
μ_3	6.0	6.01 (0.2)
m_{23}	0.04	0.0563 (40.8)
$\tilde{m}_{33}$	0.0	0.113
$\tilde{m}_{43}$	3.0	4.07 (35.7)

(c) Frank copula

Inferring parameter	Target value	Posterior estimate
μ_2	5.0	4.99 (0.2)
σ_2	0.01	0.0096 (4.0)
θ	-5.0	-4.77 (4.6)
μ_1	4.0	4.02 (0.5)
m_{21}	0.09	0.1228 (36.4)
$\tilde{m}_{31}$	0.0	-0.206
$\tilde{m}_{41}$	3.0	3.52 (17.3)
μ_3	6.0	6.00 (0.0)
m_{23}	0.04	0.0467 (16.8)
$\tilde{m}_{33}$	0.0	-0.128
$\tilde{m}_{43}$	3.0	3.24 (8.0)

(d) Gumbel copula

Inferring parameter	Target value	Posterior estimate
μ_2	5.0	5.00 (0.0)
σ_2	0.01	0.0101 (1.0)
θ	2.0	2.13 (6.5)
μ_1	4.0	4.03 (0.8)
m_{21}	0.09	0.1319 (46.6)
$\tilde{m}_{31}$	0.0	-0.238
$\tilde{m}_{41}$	3.0	4.68 (56.0)
μ_3	6.0	6.01 (0.2)
m_{23}	0.04	0.0535 (33.8)
$\tilde{m}_{33}$	0.0	-0.221
$\tilde{m}_{43}$	3.0	3.29 (9.7)

Note: Percentage errors compared to the target values are shown in parentheses.

As can be seen, the posterior estimates for the mean μ_2 and the variance m_{22} agree well with the target values for all four cases, demonstrating that the Gaussian distribution of k_2 is well calibrated. More importantly, the posterior estimates for the copula parameter θ show good agreement with the target values for all copula classes, indicating that the dependence structure between k_1 and k_3 is precisely inferred from the available data. For the marginal distributions of k_1 and k_3, on the other hand, the means μ_1 and μ_3 are accurately estimated but relatively large errors are observed in the higher moments. The errors in the variances m_{21} and m_{23} may be caused by the very small target values, while the errors in the skewnesses $\tilde{m}_{31}$ and $\tilde{m}_{33}$ and the kurtoses $\tilde{m}_{41}$ and $\tilde{m}_{43}$ may be caused by the fact that they are relatively insensitive to the measurement data compared to the means and variances.

This can also be seen in Figure 4, where the calibrated marginal distributions of k_1 and k_3 are compared with the target distributions in the case where the Clayton copula is used as the target dependence structure. While the large errors are found in the higher moments, the SDFs assigning the posterior estimates to the hyperparameters show good agreement with the target Gaussian distributions.

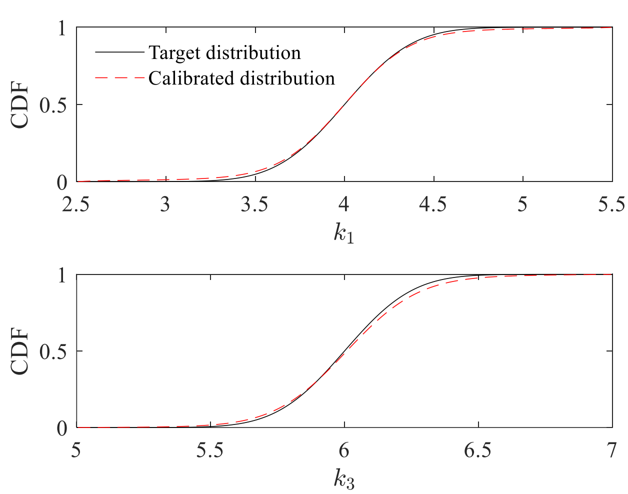

Figure 4: Calibrated marginal distributions of k_1 and k_3.

Figure 5 shows the relative position of the measurement data and the updated model outputs for each target copula class. The model outputs are generated by multiple model runs with a total of 1000 sets of the model parameters sampled from the calibrated joint distribution. As can be seen, the updated model outputs show favorable agreement with the target measurement data for all the cases, indicating that the proposed updating procedure can fully reproduce the uncertainty characteristics of the measurement data by using the appropriate copula class.

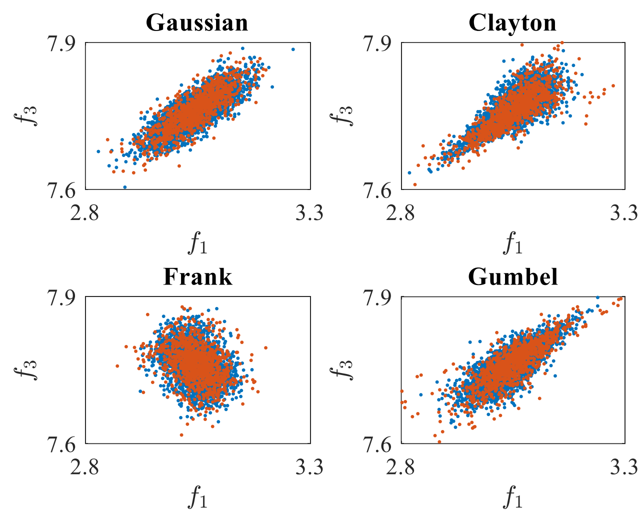

Figure 5: Relative position of the measurement data (in blue) and updated model outputs (in orange).

3.3. Model class selection results

While, in the last section, the copula class is chosen to be the same as the target copula class, the target copula class is commonly unknown a priori and thus the most appropriate copula class needs to be determined from the available data. In this section, all the four copula functions are taken into account as the candidate copula classes and the TMCMC algorithm is used to compute the evidence for each candidate. The results are given in Table 6 for each target copula class.

As can be observed from the tables, for all the cases, the candidate copula class that coincides with the target copula class results in the maximum evidence value. Hence, the copula class with the maximum evidence value is chosen as the most appropriate copula class that best represents the available measurement data in the context of Bayesian model class selection. From the results, it is indicated that the proposed approach enables to calibrate the joint probability distribution of the model parameters as appropriate by determining the best copula class and inferring its parameter together with the SDF hyperparameters.

Table 6: Evidence for each candidate copula class.
(a) Gaussian copula

Candidate copula class	Evidence
Gaussian	2.81×10^{-15}
Clayton	8.73×10^{-16}
Frank	6.09×10^{-16}
Gumbel	1.23×10^{-15}

(b) Clayton copula

Candidate copula class	Evidence
Gaussian	1.50×10^{-16}
Clayton	6.67×10^{-13}
Frank	5.89×10^{-18}
Gumbel	4.41×10^{-17}

(c) Frank copula

Candidate copula class	Evidence
Gaussian	2.46×10^{-17}
Clayton	2.33×10^{-30}
Frank	2.26×10^{-16}
Gumbel	1.78×10^{-28}

(d) Gumbel copula

Candidate copula class	Evidence
Gaussian	1.17×10^{-14}
Clayton	1.18×10^{-16}
Frank	2.26×10^{-15}
Gumbel	1.02×10^{-13}

4. CONCLUSIONS

System output may have dependencies, especially in the tail regions. Such dependencies strongly affect the system reliability and thus need to be properly dealt with. However, it is often assumed in model updating that the model parameters to be calibrated are independent of each other, which hinders us to quantify the dependence structure in the system output and makes the updating results unreliable.

This paper presents a novel stochastic model updating framework to calibrate the dependence structure between the model parameters. The joint distribution is modeled using different copulas to represent different dependence characteristics. Its marginal distributions are modeled using SDFs to avoid limiting hypotheses on distribution formats. Then, ABC is used to infer the copula parameter and the SDF hyperparameters from the available measurement data on the system of interest. In addition, Bayesian model class selection is used to determine the most appropriate copula that best represents the measurement data among different candidate copula classes. A simple 3-DOF system example is studied to demonstrate the capability of the proposed approach and the results indicate that it enables to reproduce different dependence characteristics as appropriate.

Open questions still remain. This study only focuses on the calibration of bivariate dependent model parameters. Among the copula classes used in this study, the Clayton, Frank, and Gumbel copulas have only a single parameter and cannot provide general dependence structures. While the Gaussian copula can provide general dependence structures through the correlation matrix, it can only represent Gaussian dependence structures. For the calibration of high-dimensional dependent model parameters, further research efforts are thus still needed in the future.

REFERENCES

Beaumont, M.A. (2019). “Approximate Bayesian Computation” *Annual Review of Statistics and Its Application*, 6, 379–403.

Bi, S., and Broggi, M., and Beer, M. (2019). “The role of the Bhattacharyya distance in stochastic model updating” *Mechanical Systems and Signal Processing*, 117, 437–452.

Ching, J., and Chen, Y.C. (2007). “Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging” *Journal of Engineering Mechanics*, 133(7), 816–832.

Crespo, L.G., and Kenny, S.P., and Giesy, D.P., and Stanford, B.K. (2018). “Random variables with moment-matching staircase density functions” *Applied Mathematical Modelling*, 64, 196–213.

Haff, L.H. (2013). "Parameter estimation for pair-copula constructions" *Bernoulli*, 19, 462-491.

Kitahara, M., and Bi, S., and Broggi, M., and Beer, M. (2022a). "Nonparametric Bayesian stochastic model updating with hybrid uncertainties" *Mechanical Systems and Signal Processing*, 163, 108195.

Kitahara, M., and Bi, S., and Broggi, M., and Beer, M. (2022b). "Distribution-free stochastic model updating of dynamic systems with parameter dependencies" *Structural Safety*, 97, 102227.

Mares, C., and Mottershead, J.E., and Friswell, M.I. (2006). "Stochastic model updating: Part 1-theory and simulated example" *Mechanical Systems and Signal Processing*, 20(7), 1674-1695.