A Copula-based Approach to Sensitivity to Correlations in Structural Reliability Problems

Nazih Benoumechiara
LSTA–UPMC & EDF Lab Chatou, France

Roman Sueur & Nicolas Bousquet & Bertrand Iooss
EDF Lab Chatou, France

Gérard Biau & Bertrand Michel & Philippe Saint-Pierre
LSTA–UPMC & Institut de Mathématique de Toulouse, France

To unsure the high reliability level of industrial structures, EDF conducts probabilistic studies [1]. They are based on a computational model, which aims at describing at best the physical behaviour of a structure under loading. A statistical model is build to describe the uncertainties of the parameters involved in the computational model. Unfortunately, little information is usually available on the stochastic dependence of variables. The statistical model is therefore partial and can be reduced to its margins only. Consequently, reliability studies in industrial practice are frequently carried out assuming independence of variables. A question that arises is how can we enhance the robustness of the studies without knowledge of the correlations? To answer this, our work aims to quantify the impact of potential dependencies on the structure reliability.

The methodology consider the input random vector $\mathbf{X} = (X_1, \dots, X_d) \in \mathcal{S}_{\mathbf{X}}$ and the output random variable $Y = g(\mathbf{X}) \in \mathcal{S}_Y$ of the model g. The quantity of interest of the output variable Y, used to quantify the risk faced by the structure, is denoted by $\mathscr{C}(Y)$. We use the notion of copulas to describe the dependence structure of $\mathbf{X}$, independently of its marginals. The joint Cumulative Distribution Function (CDF) of $\mathbf{X}$ is thus given as

$$F_{\mathbf{X}}(x_1, \dots, x_d) = C_{\boldsymbol{\rho}}\left(F_{X_1}(x_1), \dots, F_{X_d}(x_d)\right),$$

where $C_{\boldsymbol{\rho}} : [0,1]^d \to [0,1]$ is a copula with parameter $\boldsymbol{\rho} \in S_{\boldsymbol{\rho}}$ and F_{X_i} is the marginal's CDF of X_i. We also introduce the notation $\mathbf{X}^{\boldsymbol{\rho}}$ to describe a random vector $\mathbf{X}$ associated with a copula $C_{\boldsymbol{\rho}}$, and the related output variable $Y^{\boldsymbol{\rho}} = g(\mathbf{X}^{\boldsymbol{\rho}})$.

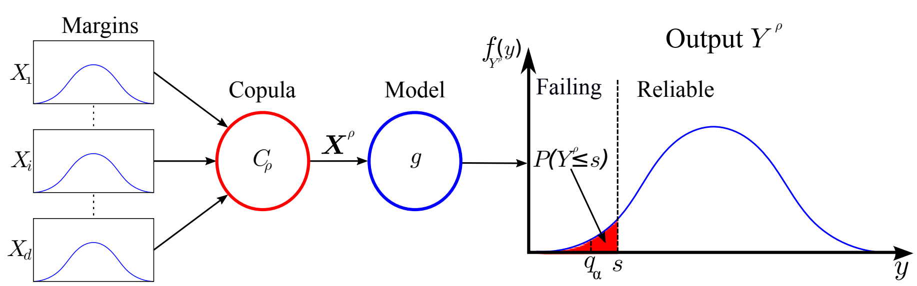

Figure 1: Uncertainty propagation of $\mathbf{X}$ with a copula $C_{\boldsymbol{\rho}}$ through the model g.

Some related studies focused on measuring the impact of a perturbation on a marginal X_i [2] or an incomplete joint density of $\mathbf{X}$ [3] on the model output Y. In this work, we propose a sensitivity index which quantify, for a chosen copula, the change on the quantity of interest between the worst case scenario and the independence case. Such an index is described by

$$\mathscr{I} = \frac{\mathscr{C}(Y^{\boldsymbol{\rho}^*})}{\mathscr{C}(Y)},$$

where $\mathscr{C}(Y)$ is the quantity of interest of Y at independence and $\boldsymbol{\rho}^*$ is the dependence configuration obtained by maximising the risk R, such as $\boldsymbol{\rho}^* = \operatorname{argmax}_{\boldsymbol{\rho} \in \mathcal{S}_{\boldsymbol{\rho}}} R(\boldsymbol{\rho})$. The index $\mathscr{I}$ would describes the general impact of the dependence structure on $\mathscr{C}$. Moreover, the index $\mathscr{I}_{ij}$ quantifies the impact of a one pair of variables dependence X_i-X_j on $\mathscr{C}$, while the other variables are independent. As for the general index $\mathscr{I}$, it is defined as

$$\mathscr{I}_{ij} = \frac{\mathscr{C}(Y^{\boldsymbol{\rho}^*_{ij}})}{\mathscr{C}(Y)},$$

where $\boldsymbol{\rho}^*_{ij} = \operatorname{argmax}_{\boldsymbol{\rho} \in \mathcal{S}_{\boldsymbol{\rho}}} R(\boldsymbol{\rho}_{ij})$ is the dependence parameter of the pair of variables X_i-X_j leading to the worst case scenario. There is, for the moment, no direct relation between $\mathscr{I}_{ij}$ and $\mathscr{I}$.

The estimation of such indices is almost entirely controlled by the estimation of the worst case dependence structure $\boldsymbol{\rho}^*$. This problem of *extremum-estimation* is consistent using a Monte-Carlo sampling. The figure 2 shows the related estimated one pair indices, for a Gaussian copula, applied to the Flood example [4] using the failure probability of Y as the quantity of interest. The closer an index $\mathscr{I}_{ij}$ is to 1 and the less impactful the dependence of the pair X_i-X_j is. And vice versa when the value is high.

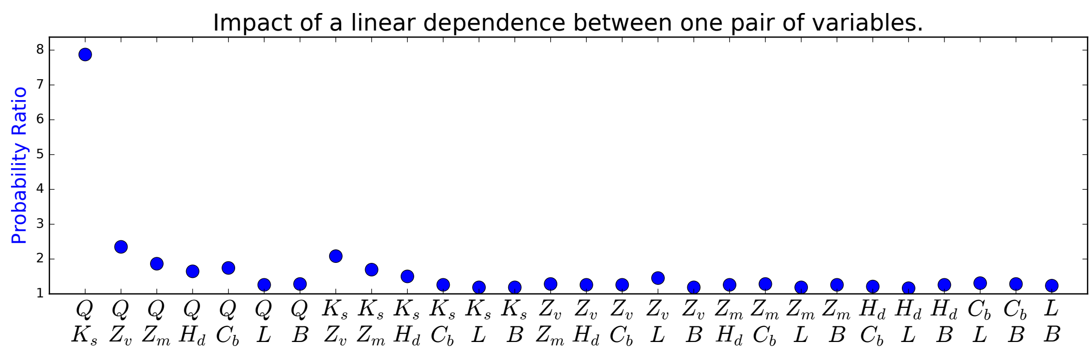

Figure 2: Monte-Carlo estimation of the indices $\mathscr{I}_{ij}$ for each pair X_i-X_j of the flood example.

Unfortunately, the Monte-Carlo sampling is costly and can hardly be performed for computationally expensive models. Thus, other estimation methods, such as Random Forests [5], could be considered to reduce the number of model evaluations. Moreover, such indices can be too pessimistic, because the worst case scenario can be very unlikely. Therefore, another perspective would be to consider every penalised dependence structures instead of the worst case configuration only.

References:

[1] de Rocquigny, E., Devictor, N., & Tarantola, S. (Eds.). (2008). *Uncertainty in industrial practice: a guide to quantitative uncertainty management*. John Wiley & Sons.

[2] Lemaître, P., Sergienko, E., Arnaud, A., Bousquet, N., Gamboa, F., Iooss & B. (2015). Density modification-based reliability sensitivity analysis. *Journal of Statistical Computation and Simulation*, 85(6), 1200-1223.

[3] Agrawal, S., Ding, Y., Saberi, A., & Ye, Y. (2012). Price of correlations in stochastic optimization. *Operations Research*, 60(1), 150-162.

[4] B. Iooss & P. Lemaître. A review on global sensitivity analysis methods. In: *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, C. Meloni and G. Dellino (Eds.), Springer, 2015.

[5] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.