

## A copula-based sensitivity analysis method and its application to a North Sea sediment transport model

Tene, Matei; Stuparu, Dana E.; Kurowicka, Dorota; El Serafy, Ghada Y.

**DOI**

[10.1016/j.envsoft.2018.03.002](https://doi.org/10.1016/j.envsoft.2018.03.002)

**Publication date**

2018

**Document Version**

Accepted author manuscript

**Published in**

Environmental Modelling & Software

**Citation (APA)**

Tene, M., Stuparu, D. E., Kurowicka, D., & El Serafy, G. Y. (2018). A copula-based sensitivity analysis method and its application to a North Sea sediment transport model. *Environmental Modelling & Software*, 104, 1-12. <https://doi.org/10.1016/j.envsoft.2018.03.002>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

# A copula-based sensitivity analysis method and its application to a North Sea sediment transport model

Matei Ţene<sup>a,\*</sup>, Dana E. Stuparu<sup>b</sup>, Dorota Kurowicka<sup>c</sup>, Ghada Y. El Serafy<sup>b</sup>

<sup>a</sup>*Department of Geoscience and Engineering, Delft University of Technology  
P.O. Box 5048, 2600 GA Delft, The Netherlands.*

<sup>b</sup>*Deltares, P.O. Box 177, 2600 MH Delft, The Netherlands*

<sup>c</sup>*Department of Applied Mathematics, Delft University of Technology  
P.O. Box 5031, 2600 GA Delft, The Netherlands.*

---

## Abstract

This paper describes a novel sensitivity analysis method, able to handle dependency relationships between model parameters. The starting point is the popular Morris (1991) algorithm, which was initially devised under the assumption of parameter independence. This important limitation is tackled by allowing the user to incorporate dependency information through a copula. The set of model runs obtained using latin hypercube sampling, are then used for deriving appropriate sensitivity measures.

Delft3D-WAQ (Deltares, 2010) is a sediment transport model with strong correlations between input parameters. Despite this, the parameter ranking obtained with the newly proposed method is in accordance with the knowledge obtained from expert judgment. However, under the same conditions, the classic Morris method elicits its results from model runs which break the assumptions of the underlying physical processes. This leads to the conclusion that the proposed extension is superior to the classic Morris algorithm and can accommodate a wide range of use cases.

*Keywords:* sensitivity analysis, parameter dependencies, copula, latin hypercube sampling, sediment transport, North Sea

---

## 1. Introduction

Suspended particulate matter (SPM) is composed of fine-grained particles of both inorganic and organic origin, which are suspended in the water column. This material plays an important role in the ecology of coastal areas, as it influences the underwater light conditions (directly connected to the phytoplankton growth), the amount of nutrients in the water, the material transfers to the seabed and other environmental processes. As such, the SPM concentration plays a crucial role in the dynamics of aquatic ecosystems. At the same time, the increasing number of human activities along the shorelines (fishing, sand and gravel extraction, tourism, industry) often disturb the natural equilibrium of the sediment transport processes. To assess and monitor the possible impacts on the sediment transport patterns, models are used to estimate and forecast their movement, under the combined action of both natural factors and human interference.

The current study concerns the southern North Sea area, a marine system significantly affected by SPM, since

it receives the run-off from major rivers and coastal industries. This area has seen rising interest in the scientific community (Fettweis et al., 2006; Pietrzak et al., 2011), which has led to the continuous development of the Delft3D-WAQ (Deltares, 2010) sediment transport and water quality model (see El Serafy et al., 2011; Blaas et al., 2007). Delft3D-WAQ makes use of the hydrodynamic conditions (velocities, discharges, water levels, vertical eddy viscosity and vertical eddy diffusivity) and wave characteristics (important in the sediment re-suspension and settling) to simulate the complex interplay between the hydrodynamic, chemical and biological processes involved in the sediment transport system.

However, calibrating this model is made difficult by the large number of input parameters, some of which are strongly correlated, due to physical constraints. Also, the high running time for one simulation - approximately 11 hours in full resolution and 3 hours on a coarse grid - imposes additional restrictions on the calibration efforts. This gave rise to the question of whether the model parameters can be ranked, such that the calibration process can be focused on only the subset to which the output is most sensitive. The remaining parameters can be fixed to their maximum likelihood values (determined, for example, using an expert judgment exercise).

According to van Griensven et al. (2006), over-para-

---

\*Corresponding author

*Email addresses:* m.tene@tudelft.nl (Matei Ţene),  
dana.stuparu@deltares.nl (Dana E. Stuparu),  
d.kurowicka@tudelft.nl (Dorota Kurowicka),  
ghada.elserafy@deltares.nl (Ghada Y. El Serafy)

meterization is a widespread problem for environmental models. At the same time, Shin et al. (2013) point out that only few studies in the literature (see, e.g., Schmid et al., 2003; Francos et al., 2003; Shen et al., 2008; Plecha et al., 2010; Kurniawan et al., 2011) employ sensitivity analysis methods (Campolongo et al., 2000; Makler-Pick et al., 2011) to rank parameters and identify redundancies. Among these, the method developed by Morris (1991) is especially popular (Campolongo and Saltelli, 1997; Portilla et al., 2009; Arabi et al., 2008) due to its simplicity and computational efficiency. However, in its initial formulation, the method assumes independence between model parameters. This can be a limiting factor, since, in many cases, the physically-induced dependencies can not be overlooked. For example, in Campolongo and Gabric (1997) the authors had to eliminate certain parameters from their analysis, specifically because of this limitation. Also, in (Salacinska et al., 2010), the sensitivity of the simulated chlorophyll-a concentration to a subset of ecologically significant input factors has been carried out with the use of the Morris method and later enriched by the computation of the correlation ratios of the selected parameters on the model response at a few selected locations in the domain. The second step was crucial to obtain results in agreement with expert knowledge of the ecological processes in the North Sea.

This paper proposes an extension to Morris’ method which opens the possibility to control the sampling pattern of the necessary model runs for sensitivity analysis, based on prior information about dependencies between model parameters. More specifically, this work incorporates this information into the sampling strategy of the elementary effects in the Morris method. The dependencies can be specified in terms of parameter correlations or, more precisely, by providing their joint distribution. This leads to the construction of the corresponding copula Nelsen (2007) – a joint distribution on the unit hypercube with uniform marginals – which is, finally, used to determine the set of model simulations required to conduct the sensitive analysis study.

The application on the computationally expensive Delft 3D-WAQ sediment transport model confirms that the method is able to provide physically sound results regarding the parameter ranking, even in cases where the feasible number of simulations is limited. This confirms the relevance of the method in identifying the parameters having the strongest impact on the variability of the model predictions.

The content of the paper is structured as follows. First, the Delft3D-WAQ sediment transport model and the dependence relationships between the governing model parameters are introduced. Next, the classic Morris sensitivity analysis method is reviewed, which presents the opportunity to devise a geometrical reinterpretation of its elementary effect sampling strategy, separating it into three successive stages. This new insight leads to the formulation of mechanisms to constrain each stage of Morris method, by incorporating the prior information about pa-

rameter dependencies in the form of a joint distribution with corresponding copula. In Section 4, the basic theory concerning copulas is summarized and Section 5 presents the newly developed copula-based Morris method. Finally, both methods are applied to the Delft3D-WAQ model and the results are compared, leading to the conclusions.

## 2. The Delft3D-WAQ sediment transport model for the southern North Sea

With an extensive history of maritime commerce, the North Sea is one of the most intensively traversed sea areas. It is bordered by highly industrialized and densely populated countries, which are actively engaged in mineral extraction, diking, land reclamation and other activities. The main sources of sediments are the Dover straits, the Atlantic Ocean, river bed and coastal erosion (Kamel et al., 2014). The SPM concentration varies in both time and space, as a response of the seabed to the hydro-meteorological forces that result from the interaction between river inflows, waves, winds, currents and external factors.

For example, the breaking waves in the near-shore areas, together with various horizontal and vertical current patterns are constantly transporting beach sediments. Sometimes, this transport results in only a local rearrangement of sand. However, under certain conditions, extensive displacements of sediments along the shore take place, possibly moving hundreds of thousands of cubic meters of sand along the coast each year. During calm weather conditions, the SPM settles and mixes with the upper bed layers. Subsequently, strong near-bed currents, generated by tides or high surface waves, can trigger the resuspension of the SPM from the seabed into the water column.

The Delft3D-WAQ (Deltares, 2010) model is capable to describe the erosion, transport and deposition of SPM in the southern North Sea with a good degree of accuracy (El Serafy et al., 2011). In the model, SPM consists of three different fractions (Jiménez and Madsen, 2003): medium (IM<sub>1</sub>, diameter 40  $\mu m$ ), coarse (IM<sub>2</sub>, diameter 15  $\mu m$ ) and fine sediments (IM<sub>3</sub>, diameter 1  $\mu m$ ). These appellations are a Delft3D internal name and will be used to refer to three sediment types throughout the remainder of this paper. The model computes the convection-diffusion, settling and resuspension of the three silt fractions of SPM, given the transport velocities, mixing coefficients and bed shear stress adopted from the hydrodynamic and wave models. The spatial domain is covered by an orthogonal grid of  $134 \times 165$  cells, with a resolution that varies between  $2 \times 2 \text{ km}^2$  in the coastal zone and  $20 \times 20 \text{ km}^2$  further offshore, as illustrated in Fig. 1. Also, in order to capture the vertical structure of the flow, together with the stratification and mixing of SPM caused by the tidal influence in the domain, the water depth is modeled by 12 so-called *sigma layers*, with different thicknesses (increased resolution near the seabed). The water surface

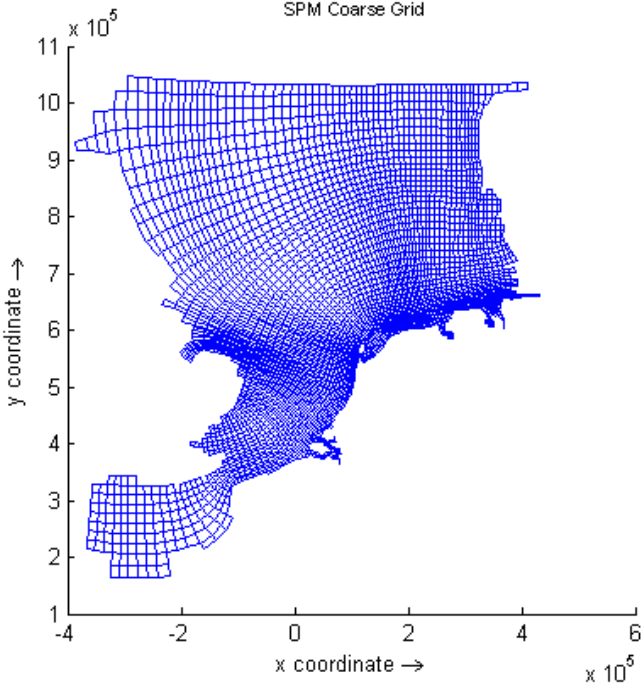


Figure 1: Delft3D-WAQ spatial discretization grid for the North Sea.

is represented by the first layer and represents 4% of the column depth.

Recently, Delft3D-WAQ has been extended with an improved parametrization of the resuspension and buffering of the silt fractions (related to both  $IM_2$  and  $IM_1$ ) from the seabed (van Kessel et al., 2011). This parametrization enables a more realistic description of the periodic and relatively limited resuspension during the tidal cycle and the massive resuspension from deeper bed layers observed during high wave events (El Serafy et al., 2011). Only the main features of this approach are described below; more details can be found in van Kessel et al. (2011).

The buffer model contains two bed layers, each interacting with the water column in a specific way (Fig. 2). The first layer, denoted as  $S_1$ , is a thin fluffy mud layer that is easily resuspended by tidal currents. On the other hand, the sandy buffer layer,  $S_2$ , can store fines for a longer time and releases SPM only during highly dynamic conditions, such as spring tides or storms. Both layers interact with the water column, but with different rates, depending on the different physical processes involved in either settling or resuspension mechanisms.

The deposition towards the layers  $S_1$  and  $S_2$  is influenced by the settling velocity  $V_{sed\ IM_i}$  [m/day] and the saturation factor  $Fr_{IM_i\ sed\ S_2}$ , which distributes the flux to the seabed. The main equations describing this process are:

$$D_{1\ IM_i} = (1 - Fr_{IM_i\ sed\ S_2}) V_{sed\ IM_i} C_{IM_i} \quad (1)$$

$$D_{2\ IM_i} = Fr_{IM_i\ sed\ S_2} V_{sed\ IM_i} C_{IM_i}, \quad (2)$$

where  $C_{IM_i}$  is the concentration of the inorganic fraction,

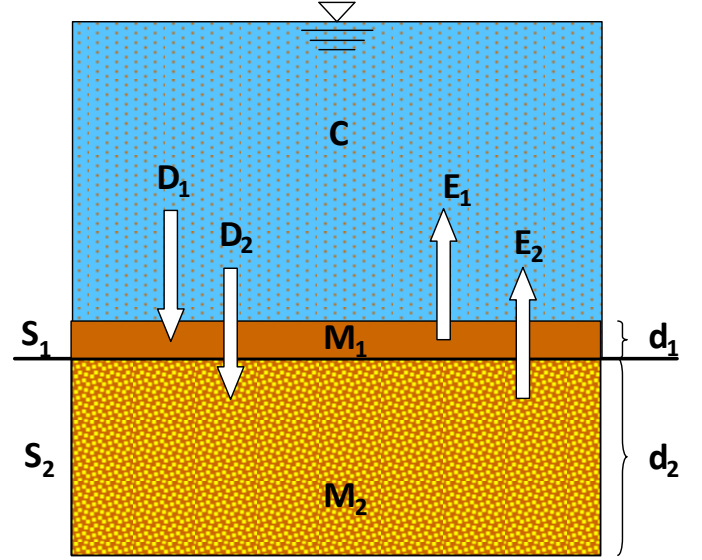


Figure 2: Schematic representation of the Delft3D-WAQ buffer model. Layer  $S_1$  is the thin fluffy mud layer of thickness  $d_1$ , while the layer  $S_2$  is the sandy sea bed infiltrated with fines of thickness  $d_2$ .  $D_j$  is the deposition flux towards layer  $S_j$ ,  $E_j$  is the erosion flux from layer  $S_j$ ,  $M_j$  is the mass sediment fraction in ( $j \in \{1, 2\}$ ) and  $C$  is the SPM concentration.

$IM_i$  ( $i \in \{1, 2, 3\}$ ).

Under certain conditions, resuspension events from the two layers occur. For the fluffy mud layer  $S_1$ , the resuspension of the SPM fractions is proportional to the respective critical shear stress levels  $\tau_{cr\ S_1\ IM_i}$  [Pa], as well as the resuspension rates  $V_{res\ IM_i}$  [1/day],

$$E_{1\ IM_i} = V_{res\ IM_i} M_{1\ IM_i} \left( \frac{\tau}{\tau_{cr\ S_1\ IM_i}} - 1 \right), \quad (3)$$

where  $\tau$  [Pa] is the bottom shear stress and  $M_{1\ IM_i}$  [g/m<sup>2</sup>] is the mass sediment fraction. For the buffer layer  $S_2$ , the fines can be mobilized only beyond critical mobilization conditions. Thus, the erosion process is mostly influenced by the critical shear stress  $\tau_{Shields}$  [Pa] and the overall pick up factor  $Fact_{res\ Pup}$  [kg/m<sup>2</sup>/s],

$$E_{1\ IM_2} = Fact_{res\ Pup} M_{2\ IM_i} \left( \frac{\tau}{\tau_{Shields}} - 1 \right)^{1.5}, \quad (4)$$

where  $M_{2\ IM_i}$  [g/m<sup>2</sup>] is the layer's mass sediment fraction and the exponent is due to the empirical pick-up function for a sandy seabed from van Rijn (1993).

These parameters and their relationships are further detailed in the next paragraph. The results presented in this paper were obtained using version 4.5208 (released on 12-08-2010) of the Delft3D-WAQ model.

### 2.1. Parameters and dependencies

The Delft3D-WAQ model captures the effect of complex mechanical, chemical and biological processes involved in the sediment transport system. Consequently, a total of

71 input parameters need to be specified in order to setup a simulation scenario. However, for the purposes of the current study, only the 14 parameters which govern the deposition/resuspension processes described earlier, will be considered for sensitivity analysis. They are listed in Table 1, along with their expert elicited value ranges see (van Kessel et al., 2011). The baseline values represent the model parametrization before the present study. The feasible ranges of the parameters were estimated based on measurements, such as the average settling velocity of mud, critical shear stress for erosion and erosion rate parameter under given conditions. For further information see (Winterwerp and van Kesteren, 2004; Gayer et al., 2006; Fettweis, 2008).

The values of the 14 parameters need to respect the physical laws and empirical relations governing the fluxes of sediment within and between the water column and the seabed. More specifically, the long term equilibrium between the buffer capacity (sediment in the  $S_2$  layer) and the water column needs to be preserved. Otherwise, the model would result in unrealistic outputs, for example localized accumulation or disappearance of sediments to/from the seabed. This necessity has resulted in a dependence structure between the model parameters, further described by the following relationships:

- an increase in parameter  $V_{\text{sed IM}_i}$ , needs to be accompanied by a decrease in parameter  $Fr_{\text{IM}_i \text{ sed } S_2}$  (or vice-versa), so that the settling into layer  $S_2$  is roughly preserved and the annual equilibrium is respected for each fraction  $i \in \{1, 2, 3\}$ ;
- the parameters  $\tau_{\text{cr } S_1 \text{ IM}_i}$  and  $V_{\text{res IM}_i}$  need to increase or decrease simultaneously, such that the year-average resuspension from layer  $S_1$  is roughly conserved for each fraction  $i \in \{1, 2, 3\}$ ;
- parameters  $\tau_{\text{Shields}}$  and  $Fact_{\text{res PUP}}$  need to increase or decrease simultaneously, so that that the year-average resuspension from layer  $S_2$  remains equal.

This leads to the specification of 7 pairs (as given in Table 2), each pair formed by two parameters which:

1. are completely rank-correlated
2. vary in the same or opposite directions (according to the rank-correlation)
3. vary simultaneously

Correlations between parameters belonging to different pairs are considered by experts as insignificant, hence independence is assumed. This amounts to a very sparse correlation matrix, with 7 non-zero correlations.

## 2.2. Model output and MERIS Remote Sensing SPM

The purpose of the sensitivity analysis is to identify the most important deposition/erosion parameters to be later used to calibrate the model against measured data.

For this purpose, this paragraph will introduce a suitable sensitivity objective function.

The model computes the total SPM concentration in each water surface grid cell on an hourly basis (calculated as the summation of the concentration of the three sediment fractions). In addition to this, SPM measurements retrieved from the optical remote sensing system ESA MERIS are available. This system supplies data from the visible, upper part of the water column, during the overpass of the Envisat satellite over the North Sea, occurring nominally once per day between 9:00 and 12:00 AM UTC. As SPM is a natural constituent of water, it affects the color of the sea. Therefore, the SPM concentrations in the water surface layer (several meters) can be derived from satellite snapshots, using the VU-IVM HYDROPT algorithm (Eleveld et al., 2008). However, some SPM pixels need to be rejected for technical or quality reasons (cloudiness, land, unreliable retrieval, etc.) and have, thus, been removed from the measurements data set (Eleveld et al., 2008). Fig. 3 illustrates an example of the MERIS data versus the model simulation results for the surface layer at the same time instance.

This allows the model error,  $\epsilon$ , to be defined as the spatial and temporal mean of the absolute differences between the model prediction and the MERIS data. If a measurement is not available in a given grid cell and time instance, that specific model output is discarded from the computation. Mathematically, this reads:

$$\epsilon = \frac{1}{N} \sum_{i=1}^N |\text{Model}_i - \text{MERIS}_i| \quad (5)$$

where  $N$  is the number of measurements (in both time and space). Note that the sampling by the MERIS sensor is irregular in time and space, mainly due to factors such as the low sun angle, cloudy weather and rough sea states. Therefore, the number of available measurements varies in time and space, with an average percentage coverage of the model domain of up to 60%. The estimated concentrations range from 0 to 100 mg/l. Using this function as sensitivity measure allows an assessment of the impact of each parameter on the ability of the model to forecast SPM concentrations.

The results of the sensitivity study will be detailed in the following paragraphs.

## 3. The classic Morris method

In this section, the concept of the classic Morris method is briefly presented, followed by a detailed discussion on the interpretation of the Morris sensitivity measures.

Given a model,  $M$ , with  $n$  model parameters,  $x = [x_1, \dots, x_n]$ , the goal of the Morris method is to rank the model parameters according to their average effect on a particular model output. The method explores all model parameters, with a so called *one-at-a-time* (OAT) design.

Table 1: Delft3D-WAQ deposition and erosion parameters, with ranges and baseline values.

Parameter	Minimum	Baseline	Maximum	Unit	Description
$V_{\text{sed IM}_1}$	5.04	10.8	43.2	$m/day$	Deposition velocities
$V_{\text{sed IM}_2}$	43.2	86.4	172.8		
$V_{\text{sed IM}_3}$	0.1	0.1	5.04		
$Fr_{\text{IM}_1 \text{ sed } S_2}$	0.05	0.15	0.4	–	$S_2$ deposition fractions
$Fr_{\text{IM}_2 \text{ sed } S_2}$	0.05	0.15	0.4		
$Fr_{\text{IM}_3 \text{ sed } S_2}$	0.05	0.15	0.4		
$V_{\text{res IM}_1}$	0.05	0.2	0.5	$1/day$	$S_1$ erosion velocities
$V_{\text{res IM}_2}$	0.2	1	1.2		
$V_{\text{res IM}_3}$	0.2	1	1.2		
$Fact_{\text{res Pup}}$	8e-9	3e-8	8e-8	$kg/m^2/s$	$S_2$ erosion rate
$\tau_{\text{cr } S_1 \text{ IM}_1}$	0.05	0.1	0.2	$Pa$	$S_1$ erosion crit. shear stresses
$\tau_{\text{cr } S_1 \text{ IM}_2}$	0.05	0.1	0.2		
$\tau_{\text{cr } S_1 \text{ IM}_3}$	0.05	0.1	0.2		
$\tau_{\text{Shields}}$	0.4	0.8	1.2	$Pa$	$S_2$ erosion crit. shear stress

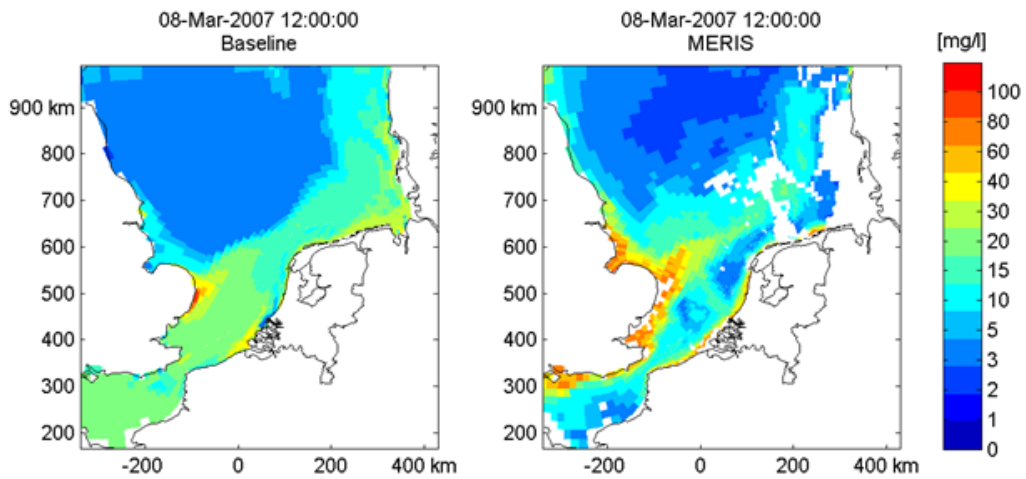


Figure 3: Comparison between the MERIS data and the corresponding Delft3D-WAQ simulation result.

Table 2: Completely rank-correlated pairs of parameters.

Pair	Rank correlation
$V_{\text{sed IM}_1} - Fr_{\text{IM}_1 \text{ sed S}_2}$	-1
$V_{\text{sed IM}_2} - Fr_{\text{IM}_2 \text{ sed S}_2}$	-1
$V_{\text{sed IM}_3} - Fr_{\text{IM}_3 \text{ sed S}_2}$	-1
$V_{\text{res IM}_1} - \tau_{\text{cr S}_1 \text{ IM}_1}$	1
$V_{\text{res IM}_2} - \tau_{\text{cr S}_1 \text{ IM}_2}$	1
$V_{\text{res IM}_3} - \tau_{\text{cr S}_1 \text{ IM}_3}$	1
$\tau_{\text{Shields}} - Fact_{\text{res Pup}}$	1

More precisely, the model parameters are varied in turn and the effect each variation has on the output is then measured. This is done using the so called *elementary effects*, which quantify the variation of the model output due to the variation in the model parameters.

This technique enables the identification of the model parameters  $x_j$  affecting the output in a way that is: (a) negligible, (b) linear and additive, (c) nonlinear or involved in interactions with other parameters (Campolongo et al., 2007). Note that in the case that the model has  $m > 1$  outputs,  $y_1, y_2, \dots, y_m$ , then, according to Shan and Wang (2009), the effects can either be measured separately for each  $y_k$  (the *split method*) or in terms of a scalar-valued function of the  $y_k$ , also referred to as *quantity of interest* (for example, an average or a norm).

After performing the sensitivity analysis, efforts can then be focused on calibration and fine-tuning of the parameters in category (c), while keeping the other parameters fixed to predefined values. Therefore, in its classic formulation, the Morris method is, essentially, a screening technique.

### 3.1. Elementary effect analysis

The Morris method (Morris, 1991) determines the statistics of the, so-called, *elementary effects*  $d_j$ , defined as

$$d_j = \frac{M(x_1, \dots, x_{j-1}, x_j + \delta, x_{j+1}, \dots, x_n) - M(x_1, \dots, x_n)}{\delta} \quad (6)$$

which serves as an approximation of the partial derivative of  $M$  with respect to  $x_j$ . In order to evaluate  $d_j$  independently of the parameter ranges, each  $x_j$  is first scaled to  $[0, 1]$ . This maps the parameter space to a unit hypercube,  $[0, 1]^n$ , which is subsequently discretized in  $p$  levels (an example is illustrated in Fig. 4). The Morris step,

$$\delta = \frac{s}{p-1} \quad s \in \{1, \dots, p-1\} \quad (7)$$

represents the magnitude of the variation and is chosen as a multiple of the grid cell size,  $\frac{1}{p-1}$ .

In order to measure the *average effect* of the parameter variation on the model output, elementary effects are calculated  $r$  times for each parameter at randomly chosen positions on the grid. This allows for the computation of

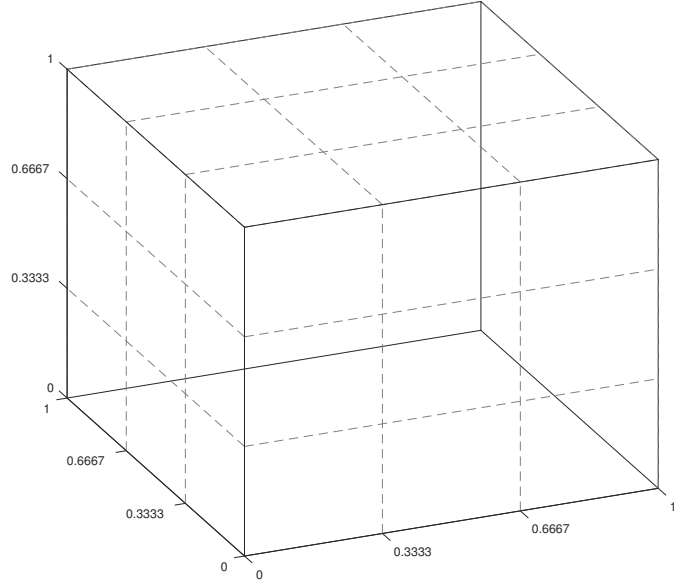


Figure 4: Unit hypercube representation of the parameter space for  $n = 3$  parameters and  $p = 4$  discretization levels.

two sensitivity measures, the elementary mean and standard deviation:

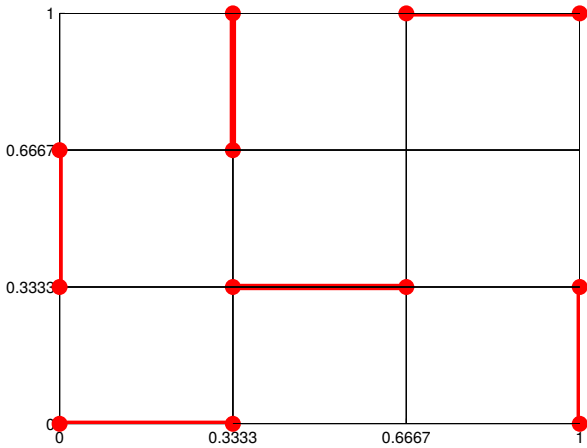
$$\mu_j = \frac{1}{r} \sum_{i=1}^r d_j^{(i)} \quad \sigma_j = \sqrt{\frac{1}{r-1} \sum_{i=1}^r (d_j^{(i)} - \mu_j)^2} \quad (8)$$

which provide insight into the relative sensitivity to  $x_j$ . Other sensitivity measures could be defined, for example, Portilla et al. (2009) use the value of  $\sqrt{\mu_j^2 + \sigma_j^2}$  to build a ranking of model parameters, while Campolongo et al. (2007) recommend using the absolute elementary mean,

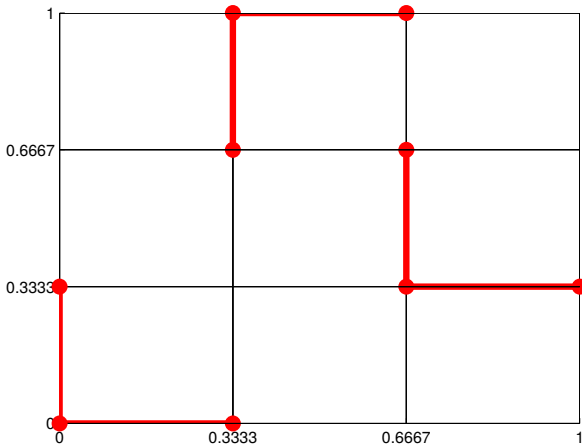
$$\mu_j^* = \frac{1}{r} \sum_{i=1}^r |d_j^{(i)}| \quad j = 1, \dots, n \quad (9)$$

instead of  $\mu_j$ , in order to better capture elementary effects of opposing sign (which cancel each other out in the calculation of  $\mu$ ).

The interpretation of  $\mu_j$ ,  $\mu_j^*$  and  $\sigma_j$  in assessing the overall influence of parameter  $x_j$  on the model output, is as follows. If  $\mu_j$  has a high amplitude, it implies not only that the parameter has a large effect on the output, but also that the sign of this effect does not vary significantly over model simulations. Meanwhile, in the case that  $\mu_j$  is relatively low and  $\mu_j^*$  is high,  $x_j$  has effects of opposing sign, varying with the point of evaluation. In addition, if  $\sigma_j$  is high, then the elementary effects relative to this parameter are significantly different from each other. This means that the value of  $x_j$ 's elementary effects are strongly dependent upon the choice of the point in the input space where it is evaluated, i.e., by the choice of the other parameters values. One may, therefore, conclude that this parameter has a high interaction with other parameters. On the other hand, a low value of  $\sigma_j$  indicates nearly con-



(a) randomly sampled elementary effects



(b) effects grouped in elementary paths

Figure 5: Efficient sampling in the Morris method ( $n = 2$ ,  $p = 4$ ,  $r = 3$ ,  $s = 1$ ): random sampling results in 12 model evaluations (left); this number can be reduced to 9 by forming elementary paths (right).

stant values of the elementary effects, therefore implying that the model is almost linearly dependent on  $x_j$ .

The analysis described above implies performing a total of  $2n \cdot r$  model evaluations. Morris (1991) proposed an efficient sampling scheme. It relies on elementary effects that share endpoints on the latin hypercube grid (Fig. 5b), effectively leading to  $(n + 1) \cdot r$  elementary paths. Such a path starts at a random position on the grid and sequentially travels one step of length  $\delta$  over each dimension. This effectively reduces the number of required model evaluations by a factor of 2.

The choices for  $p$ ,  $r$  and  $s$  have a significant impact on the outcome of the sensitivity analysis. If a high value of  $p$  is considered, which means that a high number of levels will be partitioned, one may think that the accuracy of the sampling has been increased. However, if this is not

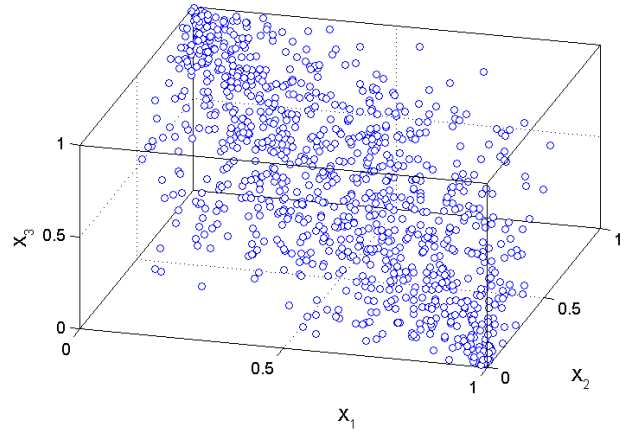


Figure 6: Scatter plot of 500 samples from a Gaussian copula with correlations  $\rho(x_1, x_2) = \rho(x_1, x_3) = -0.7$  and  $\rho(x_2, x_3) = 0.7$ .

related to a high value of  $r$ , many of the levels will remain unexplored. Also, the value of  $s$  depends on the choice of  $p$ . According to Morris (1991), a convenient choice is  $s = \frac{p}{2}$  (assuming  $p$  is even), while previous studies (Campolongo et al., 2007) have demonstrated that  $p = 4$  and  $r = 10$  produce valuable results in many cases.

A short introduction to copulas is presented in the next section.

#### 4. Copulas

A copula is a joint distribution, defined on an  $n$ -dimensional unit hypercube with uniform marginal distributions (Nelsen, 2007). It is a very popular way of representing the joint distribution, since it separates the influence of marginal distributions from the influence of parameter dependencies.

The joint cumulative distribution function  $F(x_1, \dots, x_n)$  of random variables  $X = (x_1, \dots, x_n)$  with marginal distributions denoted as  $F_i(x_i)$ ,  $i = 1, \dots, n$  can be represented with copula  $C$  as follows,

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)), \quad (10)$$

which is unique if  $(x_1, \dots, x_n)$  are continuous (Nelsen, 2007).

The most popular copulas used in practice are Gaussian, Student-t and copulas from the Archimedean family. In Figure 6, a scatter plot of samples from a three dimensional Gaussian copula with correlations  $\rho(x_1, x_2) = \rho(x_1, x_3) = -0.7$  and  $\rho(x_2, x_3) = 0.7$  is presented. The larger concentration of points close to the  $(1,0,0)$  and  $(0,1,1)$  is due to the negative correlation between the first parameter and the remaining two.

#### 5. Copula-based Morris method

The Morris method (Morris, 1991) was conceptually designed for models with independent parameters. However, most often, model parameters are related to each



other; disregarding this association results in an invalid description of the physical system. Sensitivity analysis based on independent random sampling, as is the one performed by the classic Morris method, is not applicable in these cases, since it breaks the underlying model assumptions, possibly leading to unrealistic behavior. This has motivated the need to develop a general method for sensitivity analysis. For this reason, this section introduces a novel copula-based approach, able to account for a wide range of dependencies between the model parameters.

As discussed before, the elementary paths are the building blocks of the Morris method. Without loss of generality, consider the case when the Morris step is equal to one cell, i.e.  $\delta = \frac{1}{p-1}$ . Then, as illustrated in Fig. 7, each path runs on the contour of a grid cell, starting in one of its corners and ending in the opposite one (since all coordinates are successively altered with  $\pm\delta$ ).

The copula-based method relies on the key observation that the sampling of a path can be done, equivalently, in the following three steps:

1. *Choosing the target grid block*
2. *Choosing the starting point* as one of the corners of the grid block
3. *Choosing the traversal order* of the contour segments, in order to reach the opposite corner

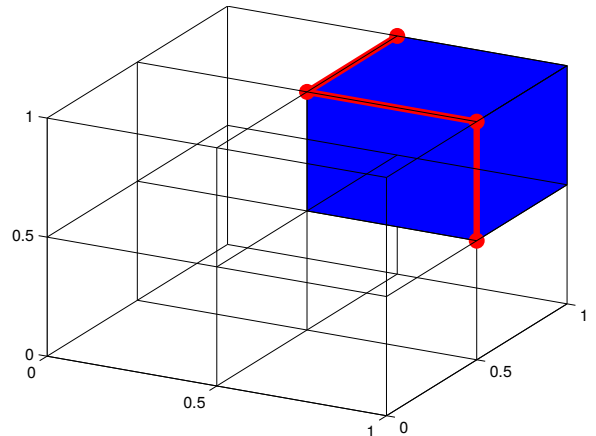
For example, the path in Figure 7 was obtained by first choosing the blue-shaded grid cell, then its lower-right corner as the starting point, A. In order to calculate the elementary effects, a path must be chosen such that all the parameters, three in this case, are varied, one at a time, with  $\delta$ . Note that there are  $3! = 6$  different ways of traversing this grid cell from A to B. In this case, parameter  $x_3$  is changed first, followed by  $x_1$  and finally,  $x_2$ . Thus, determining an order of traversal is equivalent to choosing a permutation of the set  $\{1, 2, 3\}$ .

Note that traversing a path in reverse (from B to A) does not produce new results, since it decomposes into the same elementary effects. Therefore, there are two different ways to sample the same path: choosing its start corner and corresponding permutation  $\pi$ , or choosing its end corner and the reverse of permutation  $\pi$ . Since this is true for all elementary paths, their probability of being selected remains uniformly distributed (in accordance to the classic formulation in Morris (1991)), without any alteration of the sampling strategy.

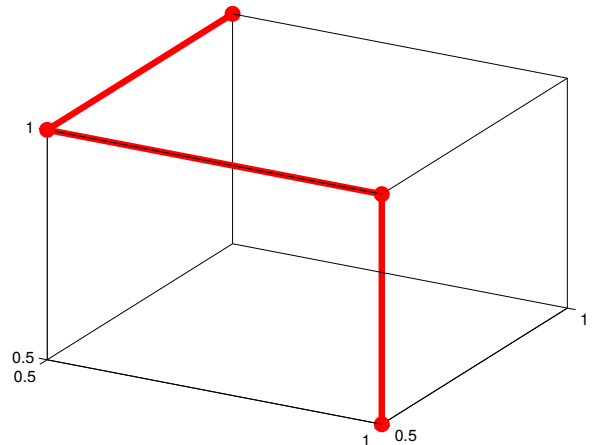
If the Morris step is higher than one grid cell (7), the only difference is that the path is drawn on the contour of a  $s \times s$  grid block (Figure 8). Note that, even though neighboring blocks intersect each other, they spawn different elementary paths and, hence, are *conceptually* disjunct.

This geometric interpretation allows us to compute the total number of possible paths on the unit hypercube as:

$$\begin{aligned} N_{\text{cells}} &= (p - s)^n, & N_{\text{corners}} &= 2^n, & N_{\text{orders}} &= n!, \\ N_{\text{paths}} &= N_{\text{cells}} \cdot N_{\text{corners}} \cdot N_{\text{orders}} / 2 \end{aligned} \tag{11}$$

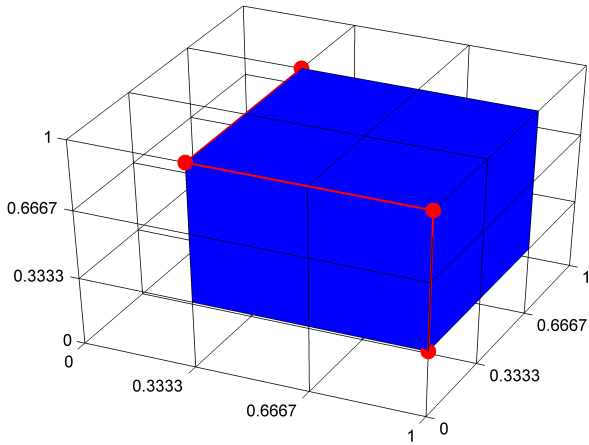


(a) the whole grid

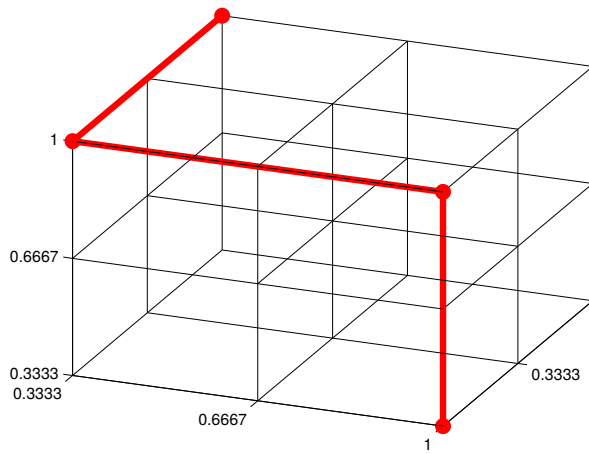


(b) zoom on the cell

Figure 7: Geometric reinterpretation of an elementary path ( $n = 3$ ,  $p = 3$ ,  $s = 1$ ).



(a) the whole grid



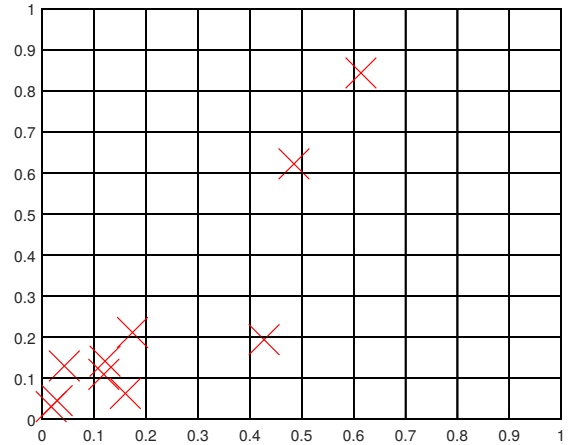
(b) zoom on the block

Figure 8: Elementary path for  $n = 3$ ,  $p = 4$ ,  $s = 2$ .

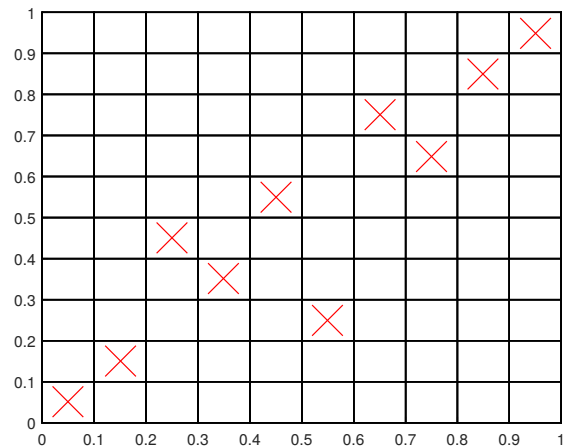
where  $n$  is the number of parameters,  $p$  is the number of discretization levels and  $s$  is the Morris step size. More importantly, sampling dependence constraints can now be introduced into each of the three steps enumerated above, by appropriately altering the sampling probabilities of the elementary effects.

### 5.1. Choosing the target grid block

The position of the grid block containing an elementary path gives the range of values within which the parameters are varied sequentially to compute elementary effects. Previous studies state that having the paths sufficiently spread within the unit hypercube is vital for the results of the analysis. For this purpose, Campolongo et al. (2007) introduce a penalty term based on Euclidean distances, while van Griensven et al. (2006) use Latin Hypercube Sampling (LHS, see McKay et al., 1979), instead of Monte-Carlo.



(a) Copula samples



(b) Latin hypercube reordered samples

Figure 9: Using LHS to ensure an even spread of the copula samples within the parameter space ( $n = 2$ ,  $p = 11$ ,  $s = 1$ ).

The goal of the new method is to constrain the sampling of the blocks in accordance to the available information about parameter dependencies. To this end, the first step is to specify a copula (Nelsen, 2007) which captures these dependencies. As presented in Section 4 one can extract a copula from the joint distribution between model parameters by transforming the margins to be uniform on  $(0,1)$ . This can be done simply by linear scaling if the parameters are uniformly distributed over their original ranges, otherwise marginal distributions need to be applied. Latin hypercube sampling is then performed on the copula, thus ensuring a good coverage of the parameter space.<sup>1</sup>

For example, in  $n = 2$  dimensions, there will be exactly one sample in each row and each column (compare Figure 9a with Figure 9b) as in the original latin hypercube sampling method (McKay et al., 1979), with the added effect of preserving dependencies between parameters, due to the copula. The algorithm used to achieve this is Latin Hypercube Sampling with Dependence (LHSD) and has been recently proposed by Packham and Schmidt (2008). Formally, considering a hypercube of  $l^n$  grid cells, LHSD operates by taking  $l$  samples from the copula,  $u^{(1)}, \dots, u^{(l)} \in \mathbb{R}^n$  (as in Figure 9a) and arranges them to get one sample in each row and column, while preserving their ranking. More precisely, the rank statistics of the  $i$ -th sample of parameter  $j$  are computed as

$$R_j[i] = \sum_{k=1}^l \mathbb{1}_{\{u^{(k)}[j] \leq u^{(i)}[j]\}} \quad i = 1, \dots, l \quad j = 1, \dots, n \quad (12)$$

where  $\mathbb{1}_S$  denotes the indicator function of set  $S$ .  $R_j[i]$  effectively represents the order of the sample in  $(u^{(1)}[j], \dots, u^{(l)}[j])$ .

Finally, the vector containing the coordinates of the origin of the target cell (i.e. its lower-left corner) is determined as:

$$\mathbf{cell}^{(i)}[j] = \frac{R_j[i] - 1}{l} \quad i = 1, \dots, l \quad j = 1, \dots, n \quad (13)$$

Note that, by the nature of LHSD, the number of samples needs to be a multiple the size of the hypercube. However, sensitivity studies may require an arbitrary number of samples. To maintain the flexibility of the Morris method, the LHSD algorithm can be repeated several times, until there is a sufficient number of samples (the excess can be discarded).

<sup>1</sup>This part of the algorithm could be done differently if only information about correlations between parameters was available. One could first obtain  $l$  Latin hypercube samples as specified in (McKay et al., 1979) and impose correlation constraints with the Iman and Conover (1982) method. However, since that method uses van der Warden scores (based on the normal distribution) that are linearly transformed with a lower triangular matrix obtained from the desired correlation matrix, it is approximately equivalent with the method presented in this paper, while employing a normal copula.

## 5.2. Choosing the starting point

For each sampled grid block,  $\mathbf{cell}^{(i)}$ , the starting corner of the path is randomly sampled such that the dependence information between parameters is preserved. The idea is simple – each corner in the grid is treated as a realization of an  $n$ -dimensional discrete distribution with  $p$  possible values, namely  $\{0, 1, 2, \dots, p-1\}$ , for each factor.

The advantage of using a copula is that the marginal distributions of the factors are removed through marginal transformation or linear scaling. The marginal probability of each factor  $x_i$  taking value  $j \in \{0, 1, 2, \dots, p-1\}$  is  $\text{Prob}(x_i = j) = \frac{1}{p}$ . As such, the finite difference formula presented in Nelsen (2007) can be used to compute the probability of each point specified on the grid,

$$\text{Prob}(X_1 = j_1, \dots, X_n = j_n) = \Delta_{a_1}^{b_1} \Delta_{a_2}^{b_2} \dots \Delta_{a_n}^{b_n} C, \quad (14)$$

where  $a_i = \frac{j_i}{p}$ ,  $b_i = \frac{j_i+1}{p}$  and

$$\Delta_{a_k}^{b_k} C = C(u_1, \dots, u_{k-1}, b_k, u_{k+1}, \dots, u_n) - C(u_1, \dots, u_{k-1}, a_k, u_{k+1}, \dots, u_n) \quad (15)$$

Hence when  $n = 3$

$$\begin{aligned} \text{Prob}(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \\ C(b_1, b_2, b_3) - C(a_1, b_2, b_3) - C(b_1, a_2, b_3) + C(a_1, a_2, b_3) \\ - C(b_1, b_2, a_3) + C(a_1, b_2, a_3) + C(b_1, a_2, a_3) - C(a_1, a_2, a_3). \end{aligned} \quad (16)$$

Consider the formula 16 when  $p = 2$ . In this case the hypercube is composed of only one cell with eight corners. Each factor can take only two possible values  $x_i = \{0, 1\}$ .  $p_{0,0,0} = \text{Prob}(X_1 = 0, X_2 = 0, X_3 = 0) = C(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ , with  $a_i = 0, b_i = \frac{1}{2}, i = 1, 2, 3$ , since any copula evaluated at point zero is 0. Using a normal copula with the correlation matrix presented in Section 4,  $p_{0,0,0} = 0.0633$ . Similarly,  $p_{0,1,0} = p_{0,0,1} = p_{1,1,0} = p_{1,0,1} = p_{1,1,1} = 0.0633$ . However, the probability of the point  $(0,0,1)$  (as well as  $(1,1,0)$ ) is much higher (see Figure 6) and it can be calculated as  $p_{1,0,0} = C(1, \frac{1}{2}, \frac{1}{2}) - C(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) = 0.3101$  with  $a_i = 0, b_i = \frac{1}{2}, i = 2, 3$  and  $a_1 = \frac{1}{2}, b_1 = 1$ . The starting point is sampled according to the calculated distribution, hence there is a much larger chance to choose point  $(1,0,0)$  or  $(0,1,1)$  over the other options.

This procedure can be used to compute distributions of corner points for each cell on the grid. However, for large  $p$  it can prove to be a computationally demanding task. It is only of interest to do this for the grid cells resulting from LHSD, however If this is still a large number, one can compute the distribution only once, as illustrated in the example above for  $p = 2$ , and assume that it applies to all grid cells. In the case of the Gaussian copula, this simplified procedure would be sufficient. For more complicated copulas, i.e. with asymmetries and tail dependencies, however, the assumption would not hold.

Finally, after the starting corner is determined, the corresponding elementary path will end in the opposite corner, since it is composed of one elementary effect for each

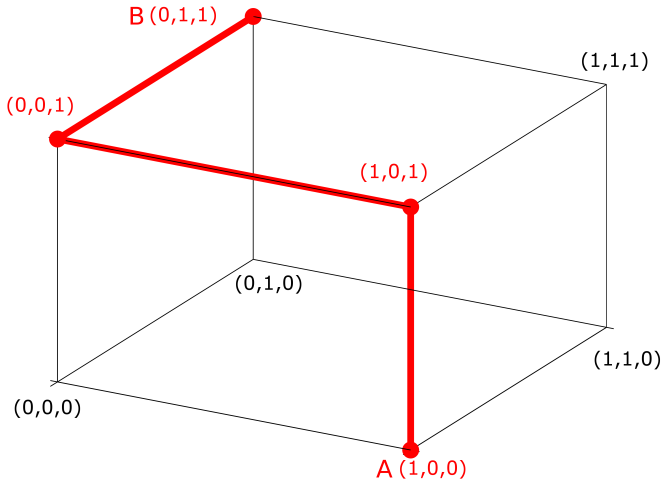


Figure 10: Binary representation of the nodes on the path, starting in corner A (i.e.  $\mathbf{b}_s = (1, 0, 0)$ ) and ending in B (i.e.  $\mathbf{b}_e = (0, 1, 1)$ ).

parameter. For convenience, each of the grid cell's corners are assigned a binary representation, starting with  $(0,0,0)$  in the origin (see Fig. 10). Then, the end point is determined by negating the representation of the sampled starting point. Note that there are  $2^n$  possible paths between these points – the choice between them is explained in the next paragraph.

### 5.3. Choosing the traversal order

The order of traversal is given by a randomly sampled permutation  $\pi^{(i)}$  which describes the way to get from the starting point to the opposite corner by changing one factor at the time. The path's vertices are determined by sequentially negating the components in the starting point's binary representation (similar to a Gray code sequence, obtained by swapping one bit at a time). For example, the path in Figure 10 was obtained using the permutation  $\{3, 1, 2\}$ , corresponding to the change along the  $\{z, x, y\}$  axes:

$$\begin{array}{l} \mathbf{b}_s \rightarrow \begin{pmatrix} x & y & z \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\ \mathbf{b}_e \rightarrow \end{array}$$

### 5.4. Method summary

To conclude, the copula-based Morris method follows the following outline:

#### Prerequisites

- A model that takes  $n$  parameters,  $M(x_1, \dots, x_n)$ , with their corresponding ranges.
- A copula  $C$  that best describes the dependence between the parameters. In the absence of any prior information the independence copula can be assumed,

whereas, if there are known correlations between the parameters,  $\rho_{i,j}$ , then a Gaussian copula is appropriate. For more complex dependency structures (e.g. tail dependence), one is free to use a copula from the Archimedean family (Clayton, Gumbell, etc.) or infer an empirical copula from a pre-existing set of model runs.

- The number of levels,  $p$ , and step size,  $s$ , for the Morris method.
- The number of desired paths,  $r$ , by taking into account that  $(n + 1) \times r$  model runs are necessary.

#### Algorithm

1. Define the grid as a  $p$ -level  $n$ -dimensional unit hypercube.
2. Sample  $r$  vectors,  $u_i$ , from the copula  $C$ .
3. Compute the rank statistics (12).
4. Compute the LHSD samples (13), which represent the grid blocks.
5. For each grid block, determine the start and corresponding end point, as explained in paragraph 5.2.
6. Determine the order of traversal of the path's segments by sampling a permutation,  $\pi^{(i)}$ , and determine the path, as explained in paragraph 5.3.
7. Evaluate the model at each point along the paths and compute the elementary effects (6).
8. Compute and interpret the sensitivity measures  $\mu_i$ ,  $\mu_i^*$  and  $\sigma_i$  using (8) and (9).

## 6. Sensitivity analysis results

The methodology proposed in section 5 has been applied to the Delft3D-WAQ model (described in Section 2). Recall that, in order to respect the interactions between the parameters, we separate them into 7 perfectly correlated pairs. This leads to the construction of a Gaussian copula with the rank-correlations given in Table 2, which enables the computation of cumulative elementary effects of each pair, rather than that of each individual parameter. Due to this specific choice of correlation structure, the sampling will favor grid cells lying on the hypercube's diagonal (anti-diagonal) for factors completely positively (negatively) correlated. Subsequently, the same dependencies are used to constrain the choice of starting points for the elementary paths, while the order of traversal is randomly sampled.

Since the Delft3D-WAQ model is computationally expensive (3 hours run time on a coarse grid and 11 hours for a fine grid), the number of simulations that can be performed for sensitivity analysis is limited. Therefore, the parameter space (unit hypercube) was divided into  $p = 4$  equidistant levels, on which  $r = 10$  elementary paths were sampled with a Morris step of  $s = 2$  cells. Therefore, a total number of 80 simulations were performed for the sensitivity study.

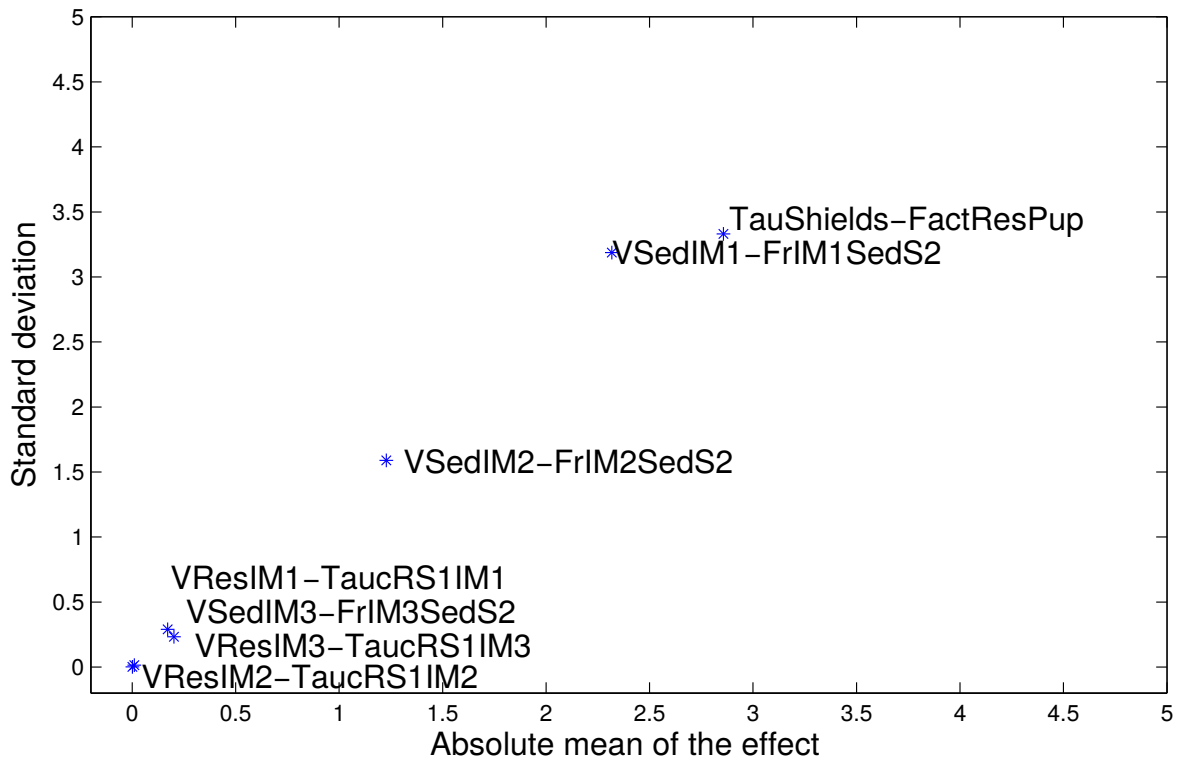


Figure 11: Copula-based Morris method sensitivity results.

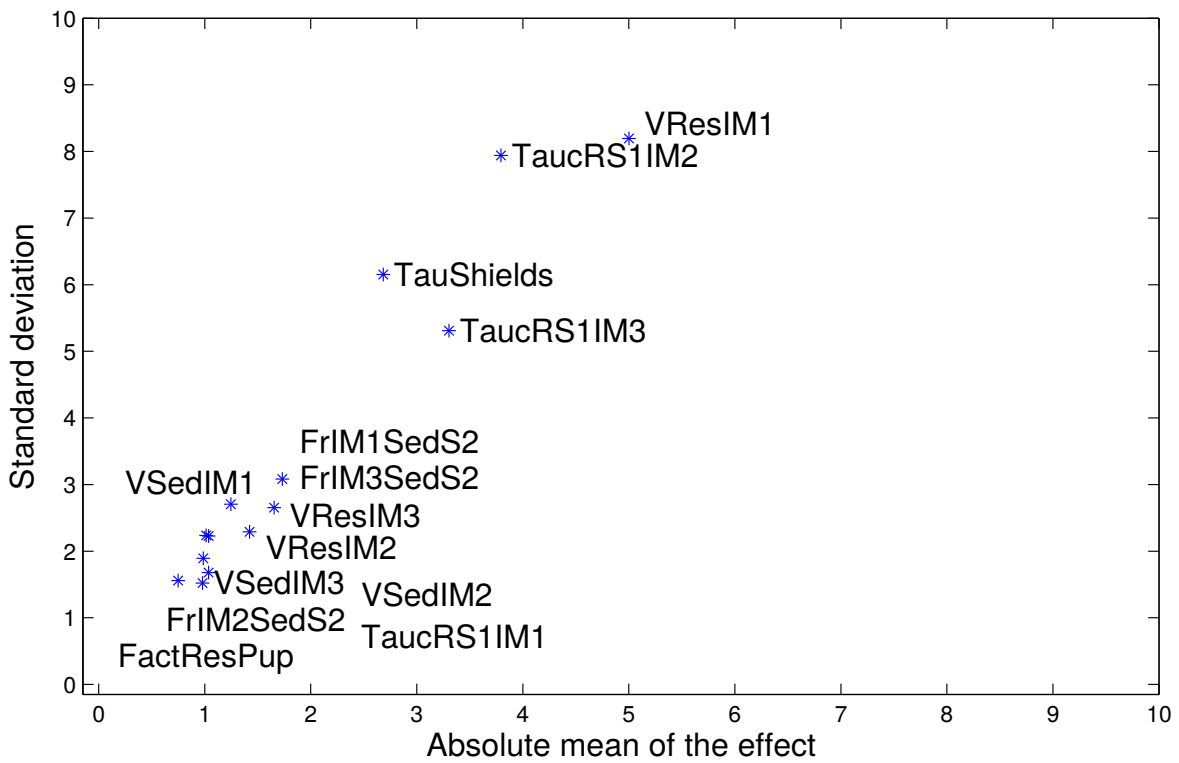


Figure 12: Classic Morris method sensitivity results.

For the sake of comparison, a separate set of simulations was performed, where the parameters were sampled using Morris' classic algorithm (thus, assuming complete independence). The comparative results are depicted in Fig. 11 and in Fig. 12 and detailed in Tables 3 and 4.

Table 3: Sensitivity measures for the correlated pairs, using the copula-based approach. The ranking is done ordered in decreasing order of  $\mu^*$ .

Pair	$\mu$	$\mu^*$	$\sigma$
$\tau_{\text{Shields}} - \text{Fact}_{\text{res Pup}}$	0.023	2.857	3.331
$V_{\text{sed IM}_1} - Fr_{\text{IM}_1 \text{ sed } S_2}$	-0.077	2.317	3.187
$V_{\text{sed IM}_2} - Fr_{\text{IM}_2 \text{ sed } S_2}$	0.496	1.228	1.589
$V_{\text{sed IM}_3} - Fr_{\text{IM}_3 \text{ sed } S_2}$	0.063	0.202	0.233
$V_{\text{res IM}_1} - \tau_{\text{cr } S_1 \text{ IM}_1}$	0.019	0.171	0.290
$V_{\text{res IM}_2} - \tau_{\text{cr } S_1 \text{ IM}_2}$	0.003	0.011	0.015
$V_{\text{res IM}_3} - \tau_{\text{cr } S_1 \text{ IM}_3}$	0.000	0.001	0.002

Table 4: Sensitivity measures for each parameter individually, using the classic Morris method, which does not support dependencies. The ranking is done ordered in decreasing order of  $\mu^*$ .

Parameter	$\mu$	$\mu^*$	$\sigma$
$V_{\text{res IM}_1}$	-0.160	5.002	8.193
$\tau_{\text{cr } S_1 \text{ IM}_2}$	-1.666	3.794	7.939
$\tau_{\text{cr } S_1 \text{ IM}_3}$	2.815	3.304	5.311
$\tau_{\text{Shields}}$	-2.037	2.684	6.153
$Fr_{\text{IM}_1 \text{ sed } S_2}$	-0.546	1.731	3.083
$V_{\text{res IM}_3}$	-1.221	1.655	2.652
$V_{\text{res IM}_2}$	-0.412	1.423	2.291
$Fr_{\text{IM}_3 \text{ sed } S_2}$	-0.450	1.247	2.706
$V_{\text{sed IM}_1}$	-0.732	1.037	2.228
$\tau_{\text{cr } S_1 \text{ IM}_1}$	0.660	1.037	1.678
$V_{\text{sed IM}_2}$	0.918	1.013	2.235
$Fr_{\text{IM}_2 \text{ sed } S_2}$	0.483	0.986	1.894
$V_{\text{sed IM}_3}$	0.642	0.979	1.523
$\text{Fact}_{\text{res Pup}}$	-0.251	0.749	1.559

The results of the copula-based Morris method match the expectations induced by the physics of the system and defined during the expert judgment exercise. The parameters to which the model is most sensitive to (presented in pairs) are, in this order:

1.  $\tau_{\text{Shields}} - \text{Fact}_{\text{res Pup}}$
2.  $V_{\text{sed IM}_1} - Fr_{\text{IM}_1 \text{ sed } S_2}$
3.  $V_{\text{sed IM}_2} - Fr_{\text{IM}_2 \text{ sed } S_2}$

As seen in Table 3, the values of  $\mu$  and  $\mu^*$  for these parameter pairs differ significantly, which suggests a high interaction with the other pairs. The pair  $\tau_{\text{Shields}} - \text{Fact}_{\text{res Pup}}$  is mainly responsible for the sand resuspension processes from the second bed layer releasing silt during high stress

events (e.g. high waves, spring tides) while the pairs  $V_{\text{sed IM}_i} - Fr_{\text{IM}_i \text{ sed } S_2}$ ,  $i \in \{1, 2\}$ , are involved in the deposition processes of the medium and coarse particles from the water column into the two bed layers. Note that the pairs  $V_{\text{res IM}_i} - \tau_{\text{cr } S_1 \text{ IM}_i}$ ,  $i \in \{1, 2, 3\}$ , which are involved in the resuspension process from the fluffy bed layer by weaker stress conditions (e.g. semi-diurnal tidal fluctuations), are of less impact on the model output variability. From this set, the resuspension for the medium size particles ( $\text{IM}_1$ ) has the highest impact.

On the other hand, the results of the classic Morris method rank the first-order resuspension rate for medium particles  $V_{\text{res IM}_1}$ , the critical resuspension stress from the layer  $S_1$  for the coarse and fine particles  $\tau_{\text{cr } S_1 \text{ IM}_i}$ ,  $i \in \{2, 3\}$ , and the critical shear stress  $\tau_{\text{Shields}}$  as the top four most influential parameters.  $\tau_{\text{Shields}}$  appears in both rankings as a sensitive parameter.

The comparison shows that, under the assumption of independence, the dominant process is the resuspension from layer  $S_1$  followed by the resuspension from the layer  $S_2$ , while under the copula-based approach, the dominant process is the resuspension from the second layer succeeded by deposition. In the model setup,  $S_1$  represents a thin fluffy layer consisting of rapidly eroding mud, while most sediment is stored in the sandy layer  $S_2$ . When the bed shear stress  $\tau_{\text{Shields}}$  exceeds a critical value (energetic conditions such as spring tides or storms) the sandy layer becomes mobile and the sediment is released in the water column. It is therefore expected that the total SPM concentration in the water column increases significantly. On the other side, during calm conditions, the presence of sediment in the water column is influenced by the deposition rates. As such, the results of the copula-based sensitivity analysis have a better correspondence with the expected system behavior.

## 7. Conclusions

Computer-based models for real-life processes often consist of systems of numerous nonlinear equations, with deterministic, as well as stochastic variables. Increases in the level of detail or accuracy within these models often imply an explosion in the number of degrees of freedom, sometimes to the point where a high number of simulations becomes unfeasible even on modern computing hardware.

This paper explored the prospect of performing sensitivity analysis on the Delft3D-WAQ sediment transport model, aiming to identify the parameters that have the strongest effects on the variability of the model predictions. The complexity and non-linearity of the model, along with the engagement of a great number of parameters, led to the application of the Morris method, due to its versatility and computational efficiency.

An extension to Morris' classical method was proposed, allowing it to incorporate prior information about the dependence structure between model parameters into the

sampling strategy. The extended method introduces copulas, which can accommodate a wide range of dependence constraints and are generally applicable. The sensitivity analysis results correspond well with the expected behavior and dynamics of sediment transport in shallow waters. More specifically, the analysis revealed that the critical shear stress and the factor responsible for re-suspension from the sandy layer  $S_2$  have the highest impact on the variance of the output. Consequently, and after expert assessment, the results of this study were used as a screening tool for subsequent model calibration, where the 5 significant pairs of parameters were subjected to a simulated annealing algorithm, in order to determine the optimal values which give the best fit between the model output and the remote sensing data.

The results of the sensitivity analysis applied for a set of dependent parameters demonstrate the potential use of the extended Morris method in determining the key driving factors of a complex model. The method may be representative for similar studies of complex models worldwide and has been implemented in a generic approach. To that end, the Matlab code used to obtain the results presented in this paper is openly available (Tene et al., 2015).

The dependence structure of parameters in the Delft3D-WAQ sediment transport model was very specific (complete positive / negative dependence between pairs of factors). This type of dependencies lead to exactly the same sensitivity behavior of these factors. However the method can accommodate other dependence structures, as well as complicated joint distributions of factors in the model. It would be of great importance to test the methodology presented in this paper for different types of models.

## 8. Acknowledgements

The authors would like to acknowledge Deltares for their openness in providing access to the Delft 3D-WAQ sediment transport model and their approval and technical support for the simulations necessary to obtain the results presented in this study.

## References

- Arabi, M., Frankenberger, J.R., Engel, B.A., Arnold, J.G., 2008. Representation of agricultural conservation practices with SWAT. *Hydrological Processes* 22, 3042–3055. doi:10.1002/hyp.6890.
- Blaas, M., El Serafy, G.Y., van Kessel, T., de Boer, G.J., Eleveld, M.A., van der Woerd, H.J., 2007. Data model integration of SPM transport in the Dutch coastal zone, in: *Proceedings of the joint 2007 EUMETSAT and 15th AMS Conference*, pp. 1–8.
- Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software* 22, 1509–1518. doi:10.1016/j.envsoft.2006.10.004.
- Campolongo, F., Gabric, A., 1997. The parametric sensitivity of dimethylsulfide flux in the southern ocean. *Journal of Statistical Computation and Simulation* 57, 337–352. doi:10.1080/00949659708811816.
- Campolongo, F., Saltelli, A., 1997. Sensitivity analysis of an environmental model: an application of different analysis methods. *Reliability Engineering & System Safety* 57, 49–69. doi:10.1016/S0951-8320(97)00021-5.
- Campolongo, F., Tarantola, S., Saltelli, A., 2000. Sensitivity analysis as an ingredient of modeling. *Statistical Science* 15, 377–395. doi:10.1214/ss/1009213004.
- Deltares, 2010. Delft3D-WAQ version 4.5208. URL: <https://oss.deltares.nl/web/delft3d/delwaq>.
- El Serafy, G.Y., Eleveld, M.A., Blaas, M., van Kessel, T., Gaytan Aguilar, S., van der Woerd, H.J., 2011. Improving the description of the suspended particulate matter concentrations in the southern North Sea through assimilating remotely sensed data. *Ocean Science Journal* 46, 179–204. doi:10.1007/s12601-011-0015-x.
- Eleveld, M.A., Pasterkamp, R., van der Woerd, H.J., Pietrzak, J.D., 2008. Remotely sensed seasonality in the spatial distribution of sea-surface suspended particulate matter in the southern North Sea. *Estuarine, Coastal and Shelf Science* 80, 103–113. doi:10.1016/j.ecss.2008.07.015.
- Fettweis, M., 2008. Uncertainty of excess density and settling velocity of mud flocs derived from in situ measurements. *Estuarine, Coastal and Shelf Science* 78, 426–436. doi:10.1016/j.ecss.2008.01.007.
- Fettweis, M., Francken, F., Pison, V., van den Eynde, D., 2006. Suspended particulate matter dynamics and aggregate sizes in a high turbidity area. *Marine Geology* 235, 63–74. doi:10.1016/j.margeo.2006.10.005.
- Francos, A., Elorza, F.J., Bouraoui, F., Bidoglio, G., Galbiati, L., 2003. Sensitivity analysis of distributed environmental simulation models: understanding the model behaviour in hydrological studies at the catchment scale. *Reliability Engineering & System Safety* 79, 205–218. doi:10.1016/S0951-8320(02)00231-4.
- Gayer, G., Dick, S., Pleskachevsky, A., Rosenthal, W., 2006. Numerical modeling of suspended matter transport in the North Sea. *Ocean Dynamics* 56, 62–77. doi:10.1007/s10236-006-0070-5.
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology* 324, 10–23. doi:10.1016/j.jhydro.2005.09.008.
- Iman, R.L., Conover, W.J., 1982. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics-Simulation and Computation* 11, 311–334. doi:10.1080/03610918208812265.
- Jiménez, J.A., Madsen, O.S., 2003. A simple formula to estimate settling velocity of natural sediments. *Journal of Waterway, Port, Coastal, and Ocean Engineering* 129, 70–78. doi:10.1061/(ASCE)0733-950X(2003)129:2(70).
- Kamel, A.M.Y., El Serafy, G.Y., Bhattacharya, B., van Kessel, T., Solomatine, D.P., 2014. Using remote sensing to enhance modelling of fine sediment dynamics in the Dutch coastal zone. *Journal of Hydroinformatics* 16, 458–476. doi:10.2166/hydro.2013.211.
- van Kessel, T., Winterwerp, H., van Prooijen, B., van Ledden, M., Borst, W., 2011. Modelling the seasonal dynamics of SPM with a simple algorithm for the buffering of fines in a sandy seabed. *Continental Shelf Research* 31, S124–S134. doi:10.1016/j.csr.2010.04.008.
- Kurniawan, A., Ooi, S.K., Hummel, S., Gerritsen, H., 2011. Sensitivity analysis of the tidal representation in Singapore Regional Waters in a data assimilation environment. *Ocean Dynamics* 61, 1121–1136. doi:10.1007/s10236-011-0415-6.
- Makler-Pick, V., Gal, G., Gorfine, M., Hipsey, M.R., Carmel, Y., 2011. Sensitivity analysis for complex ecological models – a new approach. *Environmental Modelling & Software* 26, 124–134. doi:10.1016/j.envsoft.2010.06.010.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245. doi:10.1080/00401706.1979.10489755.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, 161–174. doi:10.1080/

00401706.1991.10484804.

- Nelsen, R.B., 2007. An introduction to copulas. Springer Science & Business Media.
- Packham, N., Schmidt, W.M., 2008. Latin hypercube sampling with dependence and applications in finance. SSRN Electronic Journal doi:10.2139/ssrn.1269633.
- Pietrzak, J.D., de Boer, G.J., Eleveld, M.A., 2011. Mechanisms controlling the intra-annual mesoscale variability of SST and SPM in the southern North Sea. *Continental Shelf Research* 31, 594–610. doi:10.1016/j.csr.2010.12.014.
- Plecha, S., Silva, P.A., Vaz, N., Bertin, X., Oliveira, A., Fortunato, A.B., Dias, J.M., 2010. Sensitivity analysis of a morphodynamic modelling system applied to a coastal lagoon inlet. *Ocean Dynamics* 60, 275–284. doi:10.1007/s10236-010-0267-5.
- Portilla, E., Tett, P., Gillibrand, P.A., Inall, M., 2009. Description and sensitivity analysis for the LESV model: Water quality variables and the balance of organisms in a fjordic region of restricted exchange. *Ecological Modelling* 220, 2187–2205. doi:10.1016/j.ecolmodel.2009.05.004.
- van Rijn, L.C., 1993. Principles of sediment transport in rivers, estuaries and coastal seas. Aqua Publications.
- Salacinska, K., El Serafy, G., Los, F.J., Blauw, A., 2010. Sensitivity analysis of the two dimensional application of the Generic Ecological Model (GEM) to algal bloom prediction in the North Sea. *Ecological Modelling* 221, 178–190. doi:10.1016/j.ecolmodel.2009.10.001.
- Schmid, M., Lorke, A., Wüest, A., Halbwachs, M., Tanyileke, G., 2003. Development and sensitivity analysis of a model for assessing stratification and safety of Lake Nyos during artificial degassing. *Ocean Dynamics* 53, 288–301. doi:10.1007/s10236-003-0032-0.
- Shan, S., Wang, G.G., 2009. Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. *Structural and Multidisciplinary Optimization* 41, 219–241. doi:10.1007/s00158-009-0420-2.
- Shen, Z., Hong, Q., Yu, H., Liu, R., 2008. Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Science of The Total Environment* 405, 195–205. doi:10.1016/j.scitotenv.2008.06.009.
- Shin, M.J., Guillaume, J.H., Croke, B.F., Jakeman, A.J., 2013. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. *Journal of Hydrology* 503, 135–152. doi:10.1016/j.jhydrol.2013.08.047.
- Țene, M., Stuparu, D.E., Kurowicka, D., El Serafy, G.Y., 2015. Deltares OpenEarthTools. URL: [https://svn.oss.deltares.nl/repos/openearthtools/trunk/matlab/applications/probabilistic/sensitivity\\_analysis](https://svn.oss.deltares.nl/repos/openearthtools/trunk/matlab/applications/probabilistic/sensitivity_analysis).
- Winterwerp, J.C., van Kesteren, W.G.M., 2004. Introduction to the Physics of Cohesive Sediment Dynamics in the Marine Environment, Volume 56 (Developments in Sedimentology). Elsevier Science.