Ocean Dynamics manuscript No.
(will be inserted by the editor)

Sensitivity analysis: extending the Morris method to handle parameter dependencies. Application on the Delft3D-WAQ sediment transport model for the southern North Sea

Matei Ţene · Dana E. Stuparu · Dorota Kurowicka · Ghada Y. El Serafy

Received: date / Accepted: date

Abstract This paper proposes a novel method for sensitivity analysis, able to handle dependency relations between model parameters. The foundation for this new method is the Morris algorithm, popular for its applicability and ease of implementation. However, in its classic formulation, this algorithm assumes independence between model parameters. We tackle this limitation by allowing the user to describe the dependence between parameters using a copula. The sampling strategy is further expanded using Latin hypercube sampling with dependence, leading to the identification of a set of model runs for deriving measures of sensitivity. This approach preserves the widely reported efficiency of the classical Morris method, while honoring the prescribed dependence structure between model parameters.

Our results show that, for the Delft3D-WAQ sediment transport model applied on the North Sea domain, the ranking of the model parameters is in accordance with the knowledge obtained from the prior expert judgment exercise. Under the same conditions, the ranking provided by the classic Morris method sees unexpected results, difficult to be explained by the underlying physical processes. We conclude that the extended Morris method is preferable over the classic Morris method when dependence relations between the model parameters are known. Due to its flexibility, the proposed method is applicable to a wide range of models and dependence scenarios.

Keywords sensitivity analysis · Morris method · dependence · correlation · copula · latin hypercube sampling · sediment transport · North Sea

M. Ţene
Delft University of Technology, Dept. of Geoscience and Engineering
Stevinweg 1, 2628 CN Delft, the Netherlands
Tel.: +31-(0)15-2784103
Fax: +31-(0)15-2787966
E-mail: M.Tene@tudelft.nl

D. E. Stuparu · G. Y. El Serafy
Deltares, Marine and Coastal Systems
Rotterdamseweg 185, PO Box 177, 2600 MH Delft, the Netherlands
Tel.: +31-(0)88-3358273
Fax: +31-(0)88-3358582
E-mail: Dana.Stuparu@deltares.nl
E-mail: Ghada.ElSerafy@deltares.nl

D. Kurowicka
Delft University of Technology, Dept. of Applied Mathematics
Mekelweg 4, 2628 CD Delft, the Netherlands
Tel.: +31-(0)15-2785756
Fax: +31-(0)15-2787295
E-mail: D.Kurowicka@tudelft.nl

1 Introduction

Suspended particulate matter (*SPM*) is composed of fine-grained inorganic particles and materials of organic origin that 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 natural sediment transport processes. To assess and monitor the possible impacts on the sediment transport patterns, models are used to estimate and forecast the sediment movement, under the combined action of both natural factors and human interference.

The current study concerns the southern North Sea area, a relatively heavily impacted marine system, which receives the run-off from major rivers and coastal industries. The sediment transport in this area has been subject to many studies (Fettweis et al 2006; Pietrzak et al 2011), among which

the continuous development of the Delft3D-WAQ sediment transport and water quality model (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 resuspension and settling) to successfully integrate the hydrodynamic, chemical and biological processes involved in the sediment transport system.

However, calibrating this model is made difficult by the large number of model parameters, some of which are strongly correlated, due to physical constraints. Also, the high running time for one simulation - approximately 3 hours per simulation on a coarse grid and 11 hours for a finer grid - imposes additional restrictions on the calibration efforts. This gave rise to the question of whether priorities can be defined among the model parameters or whether the number of parameters to be used for calibration could be reduced.

Fortunately, in previous studies (see, for example, Schmid et al 2003; Francos et al 2003; Shen et al 2008; Plecha et al 2010; Kurniawan et al 2011), sensitivity analysis has been successfully used to identify the set of parameters with the highest impact on the model output variability. After this assessment, the model calibration can focus solely on this significant set, while the other parameters can be fixed to their maximum likelihood values (for example, determined after an expert judgment exercise).

Among the various methods traditionally used for sensitivity analysis (see Saltelli et al 2000; Makler-Pick et al 2011, for a detailed overview), the method developed by Morris (1991), in particular, has seen widespread use (Campolongo and Saltelli 1997; Portilla et al 2009; Arabi et al 2008) due to its simplicity and computational efficiency. Unfortunately, in its initial formulation, the Morris 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, Campolongo and Gabric (1997) had to eliminate certain parameters from their analysis, specifically because of this limitation.

In this paper, we formulate an extension to Morris' method, which opens the possibility to control the sampling pattern of the parameters, based on prior information about dependency between (groups of) model parameters. The extension introduces a novel mechanism using the copula concept to constrain each stage in Morris' sampling strategy. The application on the computationally expensive Delft3D-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 strongest effects on the variability of the model predictions.

The content of the paper is structured as follows. First, we present the Delft3D-WAQ sediment transport model and the dependence relationships between the governing model parameters. Next, we give a brief review of the classic Morris sensitivity analysis method. We then propose a geometrical reinterpretation of Morris' sampling strategy, described in three successive stages. With this new insight at hand, we formulate generalized mechanisms to constrain each stage, by conditioning on prior information about parameter dependence. Finally, the results of the new sensitivity analysis method applied on the Delft3D-WAQ model are presented and compared with the results the classical Morris method.

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 conducting mineral extraction, diking, land reclamation and other activities. The main sources of sediments are the Dover straits, the Atlantic Ocean, the bed erosion sediment generated by river inflows, coastal erosion and other sources (Kamel et al 2013). 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 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, other times, there are extensive displacements of sediments along the shore, 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. On the other hand, strong near-bed currents generated by tides or high surface waves determine the resuspension of the *SPM* from the seabed onto the water column.

The Delft3D-WAQ 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* is represented as three different fractions (Jiménez and Madsen 2003): medium ($IM1$, diameter 40 μm), coarse ($IM2$, diameter 15 μm) and fine sediments ($IM3$, diameter 1 μm). The model computes the advection-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 134×165 cells, with a resolution that varies between 2×2 km^2 in the coastal zone and 20×20 km^2 further offshore, as illustrated in (Figure 1). Also, in order to

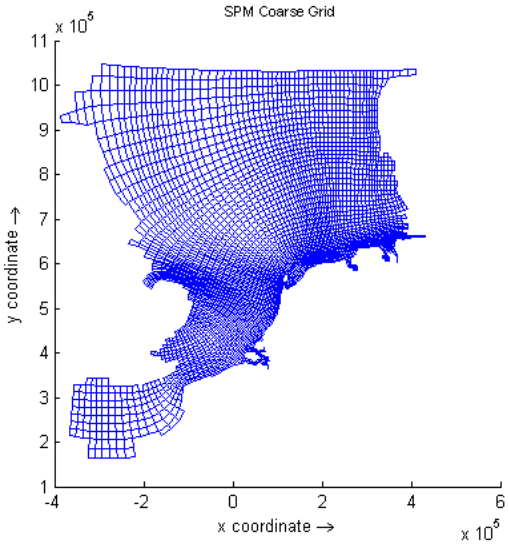


Fig. 1 Delft3D-WAQ spatial discretization grid for the North Sea

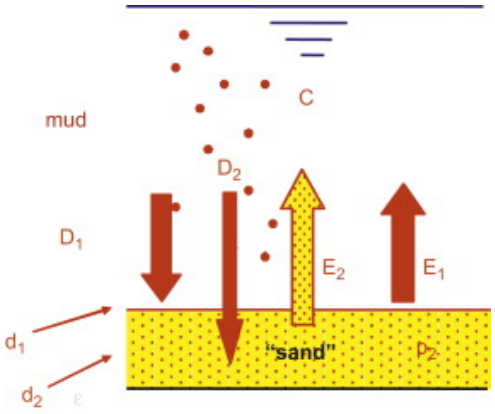


Fig. 2 Schematic representation of the buffer model. Layer S_1 is the "thin fluffy layer" of thickness d_1, while the layer S_2 is the "sandy sea bed infiltrated with fines" of thickness d_2. D_j is the deposition flux towards layer S_j, E_j is the erosion flux from layer S_j ($j \in \{1,2\}$) and C is the *SPM* concentration. Figure adopted from Van Kessel et al (2011)

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 sigma layers, with variable thickness (increased resolution near the seabed). The water surface is represented by the first sigma layer and it is calculated as 4% of the water depth.

Recently, Delft3D-WAQ has been extended with an improved parametrization of the resuspension and buffering of the silt fractions from the seabed (Van Kessel et al 2011). This parametrization enables a 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). We describe below the main features of this approach, 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. The first layer, denoted S_1, is a thin fluffy 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 ($VsedIMi$) and a saturation factor α ($FrIMiSedS2$) that distributes the flux to the seabed. The main equations describing this process are:

$$D_{1,IM_i} = (1-\alpha_{IM_i})V_{sed,IM_i}C_{IM_i} \quad (1)$$

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

where C_{IM_i} is the concentration of the anorganic fraction, IM_i, where i indexes the three fractions of *SPM*.

On the other hand, under certain conditions, resuspension events from the two layers occur. For the fluffy layer, the resuspension of the fractions is proportional to the critical resuspension stress layer ($TaucRS1IMi$) and to a resuspension rate from layer S1 ($VresIMi$). Then, a type of pick-up formulation is applied for the resuspension for the buffer layer. In this situation, the fines are detained from this layer only beyond critical mobilization conditions. Finally, erosion is mostly influenced by the critical shear stress ($TauShields$) and the overall pick up factor for resuspension pickup from the sandy layer ($FactResPup$). These parameters and their relationships are further detailed in the following paragraph.

2.1 Parameters and dependencies

In total, the model makes use of 71 model parameters which describe the sediment exchange fluxes between the seabed and the water column. Based on expert judgment and previous experience with the model, 14 parameters have been selected as possible candidates for calibration and for further sensitivity analysis (Van Kessel et al 2011). They are listed in Table 1, where the baseline values represent the model parametrization before the present study, given as maximum likelihood estimates by the experts. When asked to quantify their uncertainty with regard to these values, the experts also provided feasible ranges, which will be taken into account during the sensitivity analysis.

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 re-

Table 1 Model parameters, admissible ranges and baseline values

Parameter	Minimum	Baseline	Maximum
TauShields	0.4	0.8	1.2
VSedIM1	5.04	10.8	43.2
VSedIM2	43.2	86.4	172.8
VSedIM3	0.1	0.1	5.04
TaucRS1IM1	0.05	0.1	0.2
TaucRS1IM2	0.05	0.1	0.2
TaucRS1IM3	0.05	0.1	0.2
FactResPup	8e-9	3e-8	8e-8
FrIM1SedS2	0.05	0.15	0.4
FrIM2SedS2	0.05	0.15	0.4
FrIM3SedS2	0.05	0.15	0.4
VResIM1	0.05	0.2	0.5
VResIM2	0.2	1	1.2
VResIM3	0.2	1	1.2

sult in unrealistic outputs, for example localized accumulation of sediments in the seabed. This necessity has resulted in a dependence structure between the model parameters, further described by the following relationships:

- parameters *TauShields* and *FactResPup* need to increase or decrease simultaneously, so that that the year-average resuspension from layer S_2 remains equal;
- for each fraction $i = 1,2,3$, an increase in parameter *VSedIMi*, needs to be accompanied by a decrease in parameter *FRMiSedS*2 (or vice-versa), so that the settling into layer S_2 is roughly preserved and the annual equilibrium is respected;
- for each fraction $i = 1,2,3$, the parameters *TaucRS*1*IMi* and *VResIMi* need to increase or decrease simultaneously, such that the year-average resuspension from layer S_1 is roughly conserved.

This framework led 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

Following the expert advice, we assume independence between the pairs themselves.

Table 2 Completely correlated pairs of parameters

Parameters	Rank correlation
TauShields-FactResPup	1
VResIM1-TaucRS1IM1	1
VResIM2-TaucRS1IM2	1
VResIM3-TaucRS1IM3	1
VSedIM1-FrIM1SedS2	-1
VSedIM2-FrIM2SedS2	-1
VSedIM3-FrIM3SedS2	-1

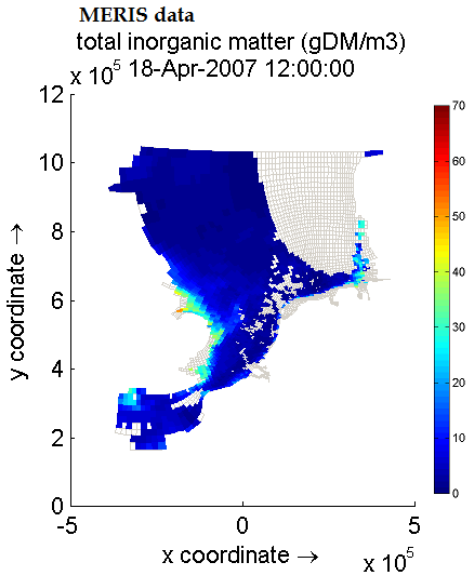


Fig. 3 Meris data

2.2 Model output and MERIS Remote Sensing *SPM*

The purpose of the sensitivity analysis is to identify the most important 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, we are provided with *SPM* measurements retrieved from the optical remote sensing system ESA MERIS. 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). Figures 3 and 4 illustrate an example of the MERIS data versus the model simulation results for the surface layer at the same time instance.

We define the model error, ε, 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,

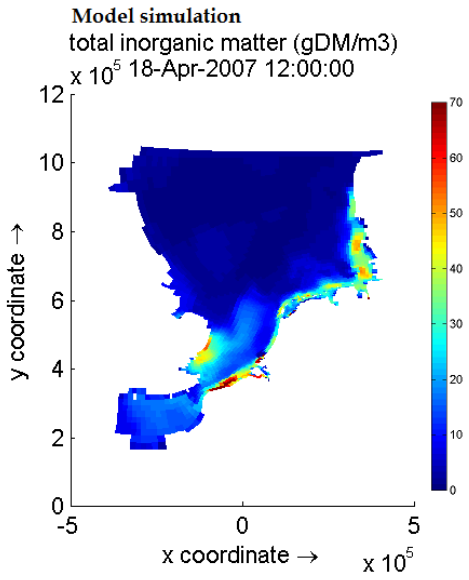


Fig. 4 Model simulation

this reads:

$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N} |Model_i - MERIS_i| \tag{3}$$

where N is the number of measurements (in both time and space). This function will be used as the sensitivity objective function. The use of this function allows for the investigation of the impact the variation in the model parameters has on the *SPM* concentration the model predicts.

Numerous other possible objective functions can be defined, depending both on the characteristics of the the modeled system and on the goals of the analysis. There is also the option of defining an objective function depending on multiple criteria, that would represent different aspects of the system behavior. However, the definition of the objective function should be carefully considered, as often an unclear definition can lead to results that are difficult to interpret.

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, $x = [x_1, \ldots, x_n]$, 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 simultaneously, with a so called *one-at-a-time (OAT)* design, which has the great advantage of reduced computational cost. 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 enable 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)). We note that in the case that the model has $m > 1$ outputs, $y_1, y_2, \ldots, y_m$, then, according to Shan and Wang (2010), 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 (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 (1991) method determines the statistics of the, so-called, *elementary effects* d_j, defined as

$$d_j = \frac{M(x_1, \ldots, x_{j-1}, x_j + \Delta, x_{j+1}, \ldots, x_n) - M(x_1, \ldots, x_n)}{\Delta} \tag{4}$$

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. 5). The Morris step,

$$\Delta = \frac{s}{p-1} \qquad s \in \{1, \ldots, p-1\} \tag{5}$$

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

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 two sensitivity measures, the elementary mean and standard deviation:

$$\mu_j = \frac{1}{r}\sum_{i=1}^{r} d_j{}^{(i)} \qquad \sigma_j = \sqrt{\frac{1}{r-1}\sum_{i=1}^{r}\left(d_j{}^{(i)} - \mu_j\right)^2} \tag{6}$$

which provide insight into the relative importance of 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

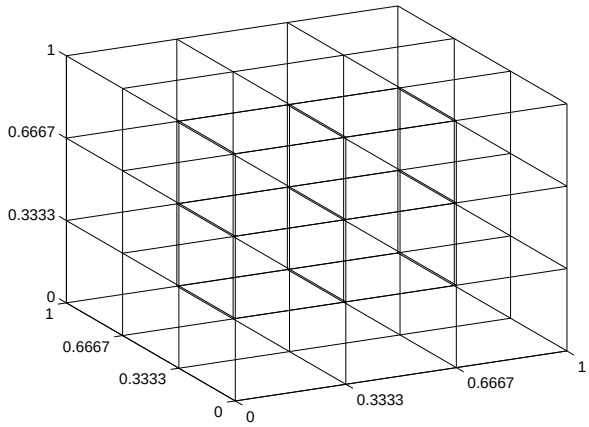


Fig. 5 Unit hypercube representation of the parameter space for $n = 3$ parameters and $p = 4$ discretization levels

ranking of model parameters, while Campolongo et al (2007) recommend using the absolute elementary mean,

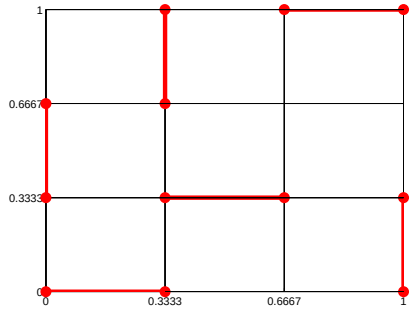
$$\mu_j^* = \frac{1}{r}\sum_{i=1}^{r}\left|d_j^{(i)}\right| \qquad j = 1,\ldots,n \tag{7}$$

instead of μ_j, in order to better capture elementary effects of opposing sign (canceling each other in the calculation of μ).

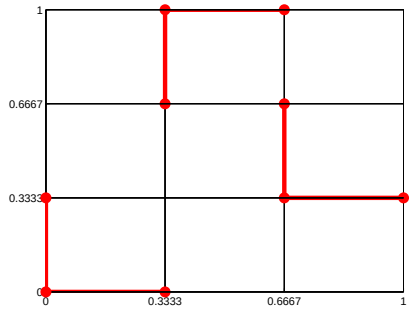
The meaningfulness of μ_j, μ_j^* and σ_j motivate our choice to use these measures of sensitivity to assess the overall influence of a parameter. If μ_j is high, it implies not only that the parameter has a large effect on the output, but also that the sign of this effect does not vary significantly over model simulations. Meanwhile, in the case that μ_j is relatively low and μ_j^* is high, it suggests that the examined parameter has effects of different signs depending on the point in space at which the effect is computed.

Assume that, for a parameter x_j we get a high value of σ_j. This can be explained by the fact that the elementary effects relative to this parameter are significantly different from each other, which means that the value of an elementary effect is strongly affected by the choice of the point in the input space where it is computed, i.e., by the choice of the other parameters values. We may therefore assume that this parameter has a high interaction with other parameters. On the other hand, a low value of σ_j indicates very similar values of the elementary effects, therefore implying that x_j is not affected by the values of the other model parameters, i.e. the model is almost linear with respect to x_j.

The analysis described above implies performing a total of $2n \cdot r$ model evaluations. Morris (1991) proposed a modification that enables an increase in efficiency of about a factor 2. The modification reuses model runs to compute different elementary effects, reducing the cost to $(n+1)\cdot r$ by grouping the effects into *elementary paths*. Such a path



(a) randomly sampled elementary effects (12 model evaluations)



(b) effects grouped in elementary paths (9 model evaluations)

Fig. 6 Efficient sampling in the Morris method ($n = 2,\ p = 4,\ r = 3,\ s = 1$)

starts at a random position on the grid and sequentially travels one step of length Δ over each dimension. As can be seen in Fig. 6b, this generates effects that share extremities, significantly reducing the number of required model evaluations.

We note that the choices for p, r and Δ 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 related to a high value of r, many of the levels will remain unexplored. Also, the value of Δ depends on the choice of p. According to Morris (1991), a convenient choice for is $\Delta = \frac{p}{2*(p-1)}$ (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.

4 Copula-based Morris method

The Morris method is conceptually designed for independent model parameters. However, most often, model parameters are related to each other; disregarding this association results in an invalid description of the physical system. Sen-

sitivity 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 model behavior (which is not of interest to the analyst). This has motivated the need to develop a general method for sensitivity analysis. For this reason, in this section we introduce 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 edges of 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 Δ. Note that there are 6 different ways of traversing this grid cell from A to B. In this case, we chose to first change parameter x_3, 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 π, or choosing its end corner and the reverse of permutation π. 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)).

If the Morris step is higher than one grid cell (5), 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* disjunctive.

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

$$N_{cells} = (p-s)^n, \qquad N_{corners} = 2^n, \qquad N_{orders} = n!, \\ N_{paths} = (p-s)^n \cdot 2^{n-1} \cdot n! \tag{8}$$

where n is the number of parameters, p is the number of discretization levels and s is the Morris step size. More importantly, we are now able to introduce sampling dependence constraints onto the three steps enumerated above.

4.1 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 (McKay et al 1979), instead of Monte-Carlo (see Figure 9a).

The goal of our method is to constrain the sampling of the blocks in accordance to the available information about parameter dependencies (point 1). To this aim, the first step is to specify a copula (Sklar 1959) which captures these dependencies. Then, in order to ensure a good coverage of the parameter space, the samples from the copula are distributed to form a latin hypercube. For example, in 2 dimensions, there will be exactly one sample in each row and each column (see Figure 9b). The algorithm used to achieve this is Latin Hypercube Sampling with Dependence (LHSD), as was recently proposed by Packham and Schmidt (2010). Formally, considering a hypercube of l^n grid cells, LHSD operates by taking l samples from the copula, $u^{(1)}, \ldots, u^{(l)} \in \mathbb{R}^n$, computing the rank statistics,

$$R_j[i] = \sum_{k=1}^{r} \mathbb{1}_{\{u^{(k)}[j] \le u^{(i)}[j]\}} \qquad i = 1,\ldots,l \qquad j = 1,\ldots,n \tag{9}$$

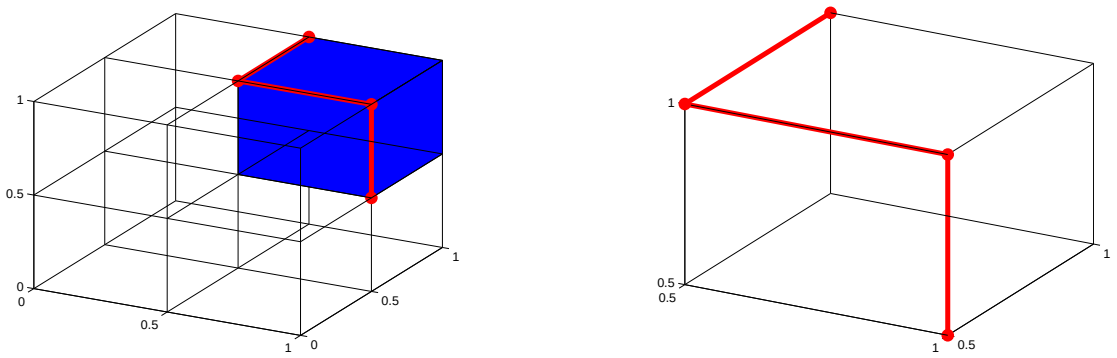
and, finally, the vector containing the coordinates of the lower-left corner of the sampled cells are determined as:

$$\mathbf{cell}^{(i)}[j] = \frac{R_j[i]-1}{r} \qquad i = 1,\ldots,l \qquad j = 1,\ldots,n \tag{10}$$

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).

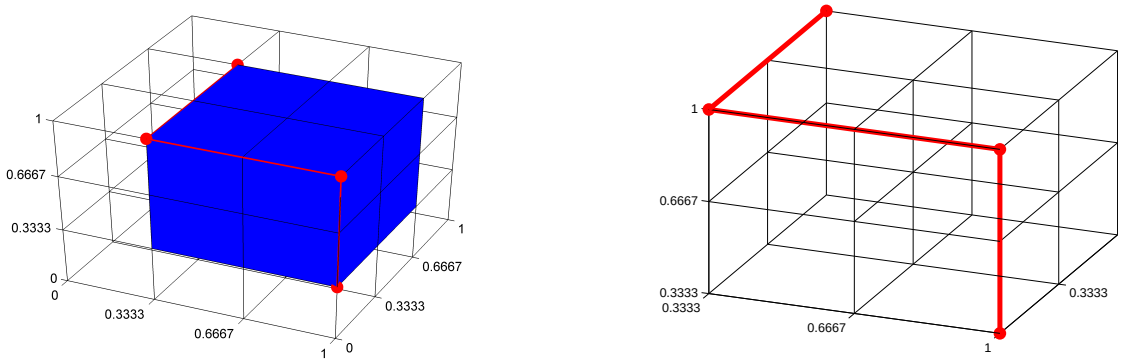
4.2 Choosing the starting point

For each sampled grid block, $\mathbf{cell}^{(i)}$, the starting corner of the path is randomly sampled. For this, we first assign an n component binary vector, $\mathbf{b}$, to each of the block's corners, as illustrated in Figure 10.



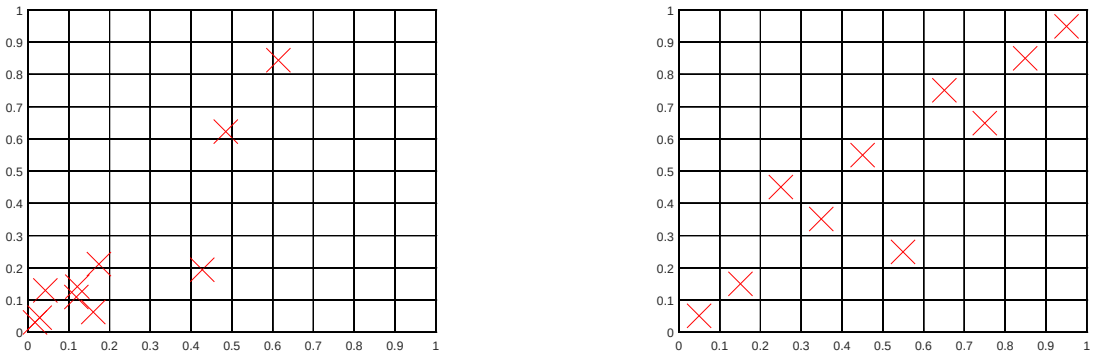
(a) the whole grid (b) zoom on the cell

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



(a) the whole grid (b) zoom on the block

Fig. 8 Elementary path for $n = 3,\ p = 4,\ s = 2$



(a) Copula samples (b) Latin hypercube reordered samples

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

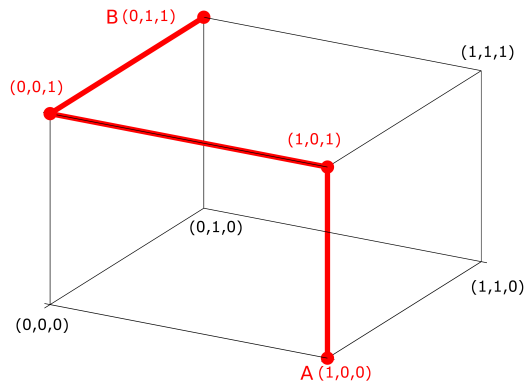


Fig. 10 Binary representation of the starting corner ($\mathbf{b}_s = (1,0,0)$)

With the strategy described so far, the grid block is chosen to respect the parameter correlations using the copula. For our application, as detailed in the model description (point 2), it is also required that the positively correlated parameters increase/decrease together, while the negatively correlated ones, vary in opposite directions. This variation is governed by the choice of starting point, $\mathbf{b}_s$, within the block. In order to accommodate this, in our extension to the Morris method, the sampling of $\mathbf{b}_s$ is constrained using a matrix of probabilities:

$$P[i,j] = \begin{cases} Prob(x_i \text{ increases} | x_j \text{ increases}) & \text{, if } i \neq j \\ Prob(x_i \text{ increases}) & \text{, if } i = j \end{cases} \quad (11)$$

A choice of $P[i,j] = P[i,i]$ signifies independence, while a higher or lower value dictates the nature of the influence. When $P[i,j] = 1$, the pair will always increase/decrease at the same time, while for $P[i,j] = 0$, they will always vary in opposite directions. P is then translated into appropriately correlated realizations of the vector $\mathbf{b}_s$ with the help of dichotomous multivariate Gaussian distributions, as described in (Macke et al 2009).

Another benefit of the binary representation is that, once $\mathbf{b}_s$ is sampled, we can identify the the end point of the path, $\mathbf{b}_e$, by simply negating the binary representation of the starting point.

4.3 Choosing the traversal order

Finally, the order of traversal is given by a randomly sampled permutation $\pi^{(i)}$. Having all of these ingredients, we can compute the path's vertices by sequentially negating the components in the starting point's binary representation. 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:

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

The last step is to integrate point 3 of the model's parameter interaction into this framework. To do this, our extension to the Morris method allows the definition of groups (in our case, pairs) of parameters. These groups also have an impact on the structure of the permutations. Specifically, if its members need to vary simultaneously, then the group will be represented as a single entity in π. For our model, this means that we need to sample permutations of length 7, each number representing the pair which needs to be varied to construct the path.

4.4 Method summary

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

Prerequisites

- A model that takes n parameters, $M(x_1,\ldots,x_n)$, 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.
- The probability matrix P, which will be used to determine the starting corner for each path. As a rule of thumb and if there is no prior evidence to suggest otherwise, P can be chosen in accordance to the already specified correlations, $\rho_{i,j}$, for example:

$$P[i,j] = \frac{1+\rho_{i,j}}{2} \qquad i > j \quad (12)$$

$$P[i,j] = \frac{P[j,i] * P[i,i]}{P[j,j]} \qquad i < j \quad (13)$$

$$P[i,i] = 0.5 \quad (14)$$

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 (9).
4. Compute the LHSD samples (10), which represent the grid blocks.
5. For each grid block, determine the start and corresponding end point, as explained in paragraph 4.2.
6. Determine the order of traversal of the path's segments by sampling a permutation, π_i, and determine the path, as explained in paragraph 4.3.
7. Evaluate the model at each point along the paths and compute the elementary effects (4).
8. Compute and interpret the sensitivity measures μ_i, μ_i^* and σ_i using (6) and (7).

5 Sensitivity analysis results

The methodology described in section 4 has been applied to the Delft3D-WAQ model (section 2). Recall that, in order to respect the interactions between the parameters, we separate them into 7 pairs of perfectly correlated parameters. The dependence relations between the parameters are defined using a Gaussian copula with the rank-correlations given in Table 2. This formulation allows us to compute the cumulative elementary effect of each pair, rather than that of each individual parameter.

At the same time, 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.

For the sake of comparison, we also performed separate sets of simulations 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 3a and 3b.

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 most important parameter pairs are, in this order:

1. $TauShields - FactResPup$
2. $VSedIM1 - FrIM1SedS2$
3. $VSedIM2 - FrIM2SedS2$

As seen in Table 3a, the values of μ and μ^* for these parameter pairs differ significantly, which suggests a high interaction with the other pairs. The pair ($TauShields$, $FactResPup$) 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 ($VsedIMi$, $FrIMiSedS2$), $i = 1,2$, are involved in the deposition processes of the medium and coarse particles from the water column into the two bed layers. We observe that the pairs ($VResIMi$, $TaucRS1IMi$), $i = 1,2,3$, 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 sides particles ($IM3$) has the highest impact.

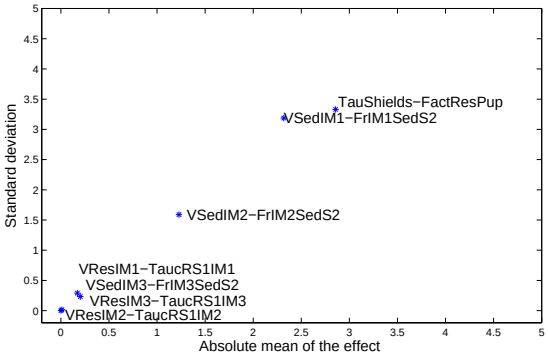


Fig. 11 Results - copula-based Morris method

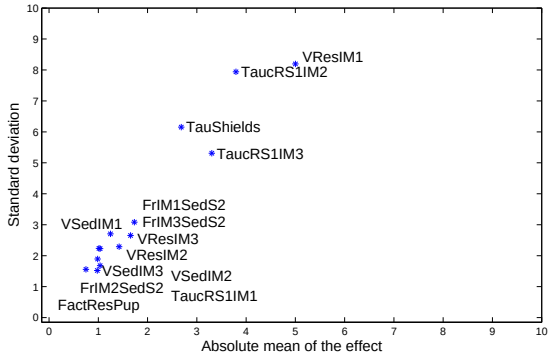


Fig. 12 Results - classical Morris method

On the other side, the results of the classic Morris method rank the first-order resuspension rate for medium particles $VResIM1$, the critical resuspension stress from the layer S_1 for the coarse particles $TaucRS1IM2$ and the critical shear stress $TauShields$ as the top three most influential parameters. $TauShields$ appears in both rankings as an important 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 fluff layer consisting of rapidly eroding mud, while most sediment is stored in the sandy layer S_2. When the bed shear stress $TauShields$ exceeds a critical value (energetic condi-

Pair	μ	μ^*	σ
TauShields-FactResPup	0.023	2.857	3.331
VSedIM1-FrIM1SedS2	-0.077	2.317	3.187
VSedIM2-FrIM2SedS2	0.496	1.228	1.589
VSedIM3-FrIM3SedS2	0.063	0.202	0.233
VResIM1-TaucRS1IM1	0.019	0.171	0.290
VResIM2-TaucRS1IM2	0.003	0.011	0.015
VResIM3-TaucRS1IM3	0.000	0.001	0.002

(a) Copula-based Morris method

Parameter	μ	μ^*	σ
VResIM1	-0.160	5.002	8.193
TaucRS1IM2	-1.666	3.794	7.939
TaucRS1IM3	2.815	3.304	5.311
TauShields	-2.037	2.684	6.153
FrIM1SedS2	-0.546	1.731	3.083
VResIM3	-1.221	1.655	2.652
VResIM2	-0.412	1.423	2.291
FrIM3SedS2	-0.450	1.247	2.706
VSedIM1	-0.732	1.037	2.228
TaucRS1IM1	0.660	1.037	1.678
VSedIM2	0.918	1.013	2.235
FrIM2SedS2	0.483	0.986	1.894
VSedIM3	0.642	0.979	1.523
FactResPup	-0.251	0.749	1.559

(b) Classic Morris method

Table 3 Sensitivity measures for the (a) correlated pairs, using the copula-based approach, and (b) each parameter individually, using the classic Morris method, which does not allow dependencies. The values are ordered in decreasing order of μ^*.

tions 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.

6 Conclusions

Computer-based models for real-life processes 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.

In this study, we 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.

We proposed an extension to Morris' classical method, 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 we obtained 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 resuspension 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. For this scope, the Matlab codes for the method can be download from: Adress repository.

Acknowledgements We 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 JR, Engel BA, Arnold JG (2008) Representation of agricultural conservation practices with SWAT. Hydrological Processes 22(16):3042–3055

Blaas M, Serafy E, GYH, van Kessel T, de Boer G, Eleveld M, van der Woerd H (2007) Data model integration of {SPM} transport in the dutch coastal zone. Proceedings of the Joint 2007 EUMETSAT/AMS Conference, Darmstadt Germany.

Campolongo F, Gabric A (1997) The parametric sensitivity of dimethylsulfide flux in the southern ocean. Journal of Statistical Computation and Simulation 57(1-4):337–352

Campolongo F, Saltelli A (1997) Sensitivity analysis of an environmental model: an application of different analysis methods. Reliability Engineering & System Safety 57(1):49–69

Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. Environmental modelling & software 22(10):1509–1518

El Serafy GY, Eleveld MA, Blaas M, van Kessel T, Aguilar SG, Van der Woerd HJ (2011) Improving the description of the suspended particulate matter concentrations in the southern North Sea through assimilating remotely sensed data. Ocean Science Journal 46(3):179–204

Eleveld MA, Pasterkamp R, van der Woerd HJ, Pietrzak JD (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(1):103–113

Fettweis M, Francken F, Pison V, den Eynde DV (2006) Suspended particulate matter dynamics and aggregate sizes in a high turbidity area. Marine Geology 235(14):63 – 74, DOI http://dx.doi.org/10.1016/j.margeo.2006.10.005, URL http://www.sciencedirect.com/science/article/pii/S0025322706002581, proceedings of the 6th International Congress on Tidal Sedimentology (Tidalites 2004) Proceedings of the 6th International Congress on Tidal Sedimentology (Tidalites 2004)

Francos A, Elorza FJ, 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(2):205–218

Jiménez JA, Madsen OS (2003) A simple formula to estimate settling velocity of natural sediments. Journal of waterway, port, coastal, and ocean engineering 129(2):70–78

Kamel A, El Serafy G, Bhattacharya B, van Kessel T, Solomatine D (2013) Using remote sensing to enhance modelling of fine sediment dynamics in the Dutch coastal zone

Kurniawan A, Ooi SK, Hummel S, Gerritsen H (2011) Sensitivity analysis of the tidal representation in Singapore Regional Waters in a data assimilation environment. Ocean Dynamics 61(8):1121–1136

Macke JH, Berens P, Ecker AS, Tolias AS, Bethge M (2009) Generating spike trains with specified correlation coefficients. Neural Computation 21(2):397–423

Makler-Pick V, Gal G, Gorfine M, Hipsey MR, Carmel Y (2011) Sensitivity analysis for complex ecological models – a new approach. Environmental Modelling & Software 26(2):124–134

McKay MD, Beckman RJ, Conover WJ (1979) Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2):239–245

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174

Packham NE, Schmidt WM (2010) Latin hypercube sampling with dependence and applications in finance. Journal of Computational Finance 13(3):81–111

Pietrzak JD, de Boer GJ, Eleveld MA (2011) Mechanisms controlling the intra-annual mesoscale variability of {SST} and {SPM} in the southern north sea. Continental Shelf Research 31(6):594 – 610, DOI http://dx.doi.org/10.1016/j.csr.2010.12.014, URL http://www.sciencedirect.com/science/article/pii/S0278434310003882

Plecha S, Silva PA, Vaz N, Bertin X, Oliveira A, Fortunato AB, Dias JM (2010) Sensitivity analysis of a morphodynamic modelling system applied to a coastal lagoon inlet. Ocean Dynamics 60(2):275–284

Portilla E, Tett P, Gillibrand P, 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(18):2187–2205

Saltelli A, Tarantola S, Campolongo F (2000) Sensitivity analysis as an ingredient of modeling. Statistical Science pp 377–395

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(3):288–301

Shan S, Wang GG (2010) Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. Structural and Multidisciplinary Optimization 41(2):219–241

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(1):195–205

Sklar M (1959) Fonctions de répartition à n dimensions et leurs marges. Université Paris 8

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(1):10–23

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(10):S124–S134