Estimation of human impact
in the presence of natural fluctuations
EC Fifth Framework Program
SCA Project IST-1999-11313

Deliverable 17

Model reduction techniques for time series normalisation

December 2000

Document Reference:	IMPACT/JRC/WP6/D17
Version:	1
Security:	Pub.
Status:	Final
Authors:	M. Ratto, S. Tarantola, A. Saltelli Joint Research Centre of European Commission, Institute for Systems, Informatics and Safety, Ispra U. Callies GKSS Forschungszentrum Geestacht GmbH, Institute of Hydrophysics, Geestacht

TABLE OF CONTENTS

1 Introduction

1.1 Uncertainty and sensitivity analysis and model reduction

Computational models are often used to give a simplified mathematical representation of reality in many different fields of application. Model input is subjected to many sources of uncertainty including errors of measurement, inadequate sampling resolution, etc. Furthermore, the model itself can include conceptual uncertainty, i.e. uncertainty in model structures, assumptions and specifications. All this imposes a limit on our confidence in the response, or output, of the model. Good modelling practice requires the modeller to provide an evaluation of the confidence in the model predictions, possibly assessing the uncertainties associated with the outcome (response) of the model itself.

Uncertainty Analysis (UA) and sensitivity analysis (SA) are prerequisites for model building in any field where models are used. UA allows assessing the uncertainty associated with the model response as a result of uncertainties in the model input. SA is aimed at establishing how the variation in the model output can be apportioned to different sources of variation, in order to establish how the given model depends upon the information fed into it.

SA can be useful in model building for identifying, on one side, the relevant factors, and on the other those who do not drive significant variation on the output. In this way, SA can be used to reduce models: unimportant factors can be fixed to their nominal values and, if factors are clearly connected to particular processes included in the model, entire parts of the models can also be eliminated or simplified.

The role of UA-SA methods in framework of the IMPACT project is therefore clearly identified as the intermediate step in the merging of mechanistic models and statistical procedures, in which the original model is reduced by allowing unimportant factors to be fixed or eliminated.

1.2 Aims of the report

The present deliverable is focused on model reduction techniques in the presence of time dependent output. UA and SA are applied with the main objective of identifying quantitative criteria for model reduction for time series normalisation. The specific

problem of time-dependent model output required the identification of specific methodologies, with respect to the more 'classical' approach presented in Deliverable 16 (Ratto et al., 2000), regarding the uncertainty assessment. Moreover, the studies carried out aimed also at conditioning model reduction to the information available from observation data sets. Hence, the new methodology is based on the study of a 'summarised' model output: the object of the SA is not the time dependent-output, but a likelihood measure of the model outcome conditioned to observations (essentially a function of the mean square error of the model with respect to observations).

The methodological approach is carried through the following steps:

1) summarise the time-dependent model output through the definition of a likelihood measure, conditioning the model predictions to observations.
2) study the empirical distribution of the likelihood measure of model output due to propagation of the various input factors through the models themselves. This study is actually an uncertainty assessment, aimed at testing the overall robustness of both the underlying model and the available data.
3) investigate model performance as far as calibration is concerned. The variance of the model likelihood can be decomposed according to source by using global sensitivity analysis (GSA). This study is able to reveal to what extent a model factor affects the model capability of being a simulator of reality. This investigation aims at trying to distinguish “live” components of the model, which drive model response and are hence “relevant”, from “dead” ones, which make no contribution to the variation in the model predictions. As a consequence, SA is the basis for the elimination of unneeded complexity from the model and therefore for model reduction.
4) model reduction criteria. Criteria for model reduction are based on the results of the sensitivity analysis.

The methodological approach is applied in test cases of IMPACT. In particular, the Elbe River case study applying the WAMPUM model is described in the present deliverable.

1.3 List of publications

Participation to international conferences

1) M. Ratto, N. Giglioli, S. Tarantola, U. Callies, Å. Forsman, Sensitivity analysis and environmental data normalisation: eutrophication case studies, Abstract submitted for the *Third International Symposium on Sensitivity Analysis of Model Output SAMO* 2001, Madrid, June 2001.
2) M. Ratto, N. Giglioli, S. Tarantola, U. Callies, Å. Forsman, Abstract accepted for the *European Safety and Reliability International Conference ESREL 2001*, Torino, September 2001.

Papers for submission to the open literature are also in preparation:

1) Ratto M., Tarantola S., Saltelli A., Sensitivity analysis in model calibration: GSA-GLUE approach, (2000), submitted to *Computer Physics Communications*.

2 The method

2.1 Preliminary considerations

The model reduction problem discussed here considers time-dependent model outputs and the availability of observations. So, on one hand, the UASA approach is the reference tool for identifying model reduction criteria, on the other hand, conditioning of the technique to observations is an additional essential requirement with respect to the problem defined in Deliverable 16 (Ratto et al., 2000b). In the present report, a technique using UASA tools, allowing model reduction conditioned to observations is presented (Ratto et al., 2000a).

In the last decade, a method based on the concept of Bayesian Inference for uncertainty estimation, has been used in hydrology as the Generalised Likelihood Uncertainty Estimation Technique (GLUE) (Beven & Binley, 1992; Romanovicz et al., 1994). The GLUE technique is as an extension of the Generalised Sensitivity Analysis methodology, which has now come to be called Regional Sensitivity Analysis (RSA), by R.C Spear and G.M. Hornberger (Hornberger and Spear, 1980, Spear and Hornberger, 1980). GLUE has been developed from an acceptance of the possible equifinality of models, i.e. different sets of model factors/structures, later on lumped under the term 'input factors', may be equally likely as simulators of the real system. It works with multiple sets of factors, typically via Monte Carlo sampling, and applies likelihood measures to estimate the predictive uncertainty of the model. Model realisations are weighted and ranked on a likelihood scale via conditioning on observations and the weights are used to formulate a cumulative distribution of predictions. Applying the RSA terminology, model structures/parameter sets with almost zero likelihood can be classified as non-behavioural and rejected.

In the RSA-GLUE framework, the basic role of SA is also clear. In general, SA is aimed at establishing how the variation in the model output can be apportioned to different sources of variation, in order to establish how the given model depends upon the information fed into it. When one is mainly interested in the predictive uncertainty, a sensitivity analysis can help in better explaining the model structure and the main sources of model output uncertainty. Additionally, in a model identification/reduction problem, a

quantitative SA able to account for conditioning on observations, can provide useful information about the model internal structure and, above all, of the interaction structure between model factors resulting from the conditioning itself. The determination of an interaction structure between model factors, in fact, is a typical feature of the RSA-GLUE classification of model realisations, where (complex) structured parameter subsets having similar likelihood of being simulators of reality are identified. Spear at al. (1994) found that the application of conventional multivariate statistics like principal component analysis to analyse interaction is not very revealing and Spear (1997) showed also that complex parametric interactions do not become evident from looking at univariate marginal probability densities. So, the deep inspection of this interaction structure is a challenging aspect of RSA and the development of suitable methodologies is still an open problem.

The basic idea for the new methodology presented in Ratto et al. (2000a) consists of a combination of the GLUE technique with variance based sensitivity analysis (extended FAST, Importance measures, Sobol' indices). The use of GLUE allows both conditioning to observation and performing a sensitivity analysis on a scalar function (the likelihood measure) instead on a time-dependent output, so avoiding the calculation of time-dependent sensitivity indices. Moreover, GSA allows a quantitative decomposition of the likelihood variance with respect to the input factors.

In the GLUE approach, factors are never considered independently but as sets of values. The likelihood measure for each model realisation is associated with a particular set of factors, conditioned to the observed data. From the methodological point of view, when a detailed SA has to be performed on such kind of 'output', some peculiar aspects have to be taken into account:

- ✓ non monotonic input-output mapping;
- ✓ high level of interaction between input factors.

The former aspect is mostly due to the form of likelihood measures, inherently non-monotonic. The latter is connected to model equifinality or over-parameterisation, i.e. many different combinations of input factors give the same model performance when conditioning on observations. These aspects pose severe constraints about the SA methods to be applied. In particular, it can be expected that MC regression based

methodologies are poor for this kind of study and that global, model independent or model free techniques should be used.

In the RSA of Spear and Hornberger, a classification algorithm is applied to the model output, resulting in a classification of each model run as behavioural or non-behavioural. The parameter sets leading to the result are stored according to the behavioural outcome. Subsequently all parameter vectors are analysed to determine the degree to which the *a priori* distributions separate under the behavioural mapping. The separation, or the lack thereof, forms the basis of the generalised sensitivity analysis.

Another SA approach applied in the past has been to evaluate the marginal distribution of likelihood for each parameter by integrating across the parameter space (Romanowicz et al., 1994). A simplified approach is given by 'visual' SA based on scatter plots of the likelihood measure vs. single factors. When significant patterns are detected, a pronounced influence on model predictions can be concluded. Moreover, subsets of better model performance can be singled out for such influent factors.

Both the RSA and the scatter plots by Romanovicz et al. address well the problem characteristics, above all as far as the non-monotony is concerned. However, they are quantitatively poor (both) and are not very efficient in the case of strong interaction. A possible extension of RSA for the study of parameter interaction has been presented in Hornberger and Spear, 1981, based on the diagonalisation of the correlation matrix of the input factor sub-sample under the behaviour classification. More recently, Spear et al. (1994) provided a further extension of RSA, consisting of a tree-structured density estimation technique to characterise the complex interaction in the portion of the parameter space rising successful simulation. As a result, the parameter space can be partitioned into small, densely populated regions and relatively large, sparsely populated regions.

In order to improve the sensitivity analysis aspects of model calibration and uncertainty prediction variance-based Global Sensitivity Analysis methods (GSA) have been applied (Ratto et al., 2000a). Variance based methods are based on the decomposition of the model output variance into a sum of terms depending on single factors and on interaction terms of increasing order. They are quantitative methods and work without any restriction about monotony or additivity of the model. The only requirement is that all what we

desire to know about the model output is captured by its variance. Application of variance-based methods allows the determination not only of main effect of input factors (equivalent to RSA or scatter plots) but also of the total effect of each factor in combination with all the others. Such a quantification is very useful, since it allows a classification of factors according to, e.g.:

- ✓ factors with high main effect: such factors affect model output singularly, independently of interaction;
- ✓ factors with small main effect but high total effect: such factors influence the model output mainly through interaction;
- ✓ factors with small main and total effect: such factors have a negligible effect on the model output and can be fixed at a nominal value.

The first class of factors can be detected also with the other methodological approaches (RSA, scatter plots, regression analysis), while the second class could be qualitatively evaluated through the extension of the RSA in Hornberger and Spear (1981) [in case of second order interactions] and in Spear et al. (1994).

2.2 Short description of GLUE

The GLUE procedure is based upon making a large number of runs of a given model with different sets of factor values, chosen randomly form specified factor distributions. Different sets of initial, boundary conditions or model structures can also be considered. On a basis of comparing predicted and observed responses, each set of factor values is assigned a likelihood of being a simulator of the system. The definition of the likelihood measure is matter in the GLUE framework and the uncertainty prediction can strongly depend on that definition. In a Bayesian framework, this is connected to how errors in the observations and in the model structure are represented by a statistical model. However, such a differentiation is not particularly relevant as far as the application of GSA is concerned, which is compatible to any definition of the likelihood measure. Hence, for the sake of simplicity and without loss of generality, among the different possible likelihood measures (Beven & Binley, 1992; Romanowicz et al. 1994; Romanowicz et al. 1996; Romanowicz et al. 2000) the following is used here.

$$L(\underline{\theta}_i \mid \underline{Y}) = \left(\frac{1}{\sigma_i^{\,2}} \right)^N \tag{1}$$

where

$$\sigma_i^2 = \frac{1}{Nobs} \sum_{j=1}^{Nobs} \left(\hat{Y}_i(t_j) - \underline{Y}(t_j) \right)^2 \tag{2}$$

is the mean squared difference between predictions and observations for the *i*-th factor set.

Rescaling of the likelihood measures such that the sum of all the likelihood values equals 1 yields a distribution function for the factor sets. From this, the uncertainty estimation can be performed, by computing the model output cumulative distribution, together with prediction quantiles.

An interesting feature of this approach, is that correlation between factor values is reflected implicitly in the likelihood measure associated with the factor sets, so that no hypothesis about the correlation structure is necessary in defining the *a priori* distributions of the model factors. A covariance structure can be obtained *a posteriori* when each factors' combination is weighted via the likelihood measures. The GLUE methodology allows also combining or updating likelihood measures, by applying the Bayes theorem (Beven & Binley, 1992).

2.3 Short description of variance-based GSA

A thorough description of sensitivity analysis methods, including linear regression, correlation analysis, importance measures, variance-based and screening methods, can be found in Saltelli et al. (2000).

When using variance based techniques (see Archer et al., 1997 for a review), the SA is based on estimating the fractional contribution of each input factor to the variance of the model output. In order to calculate the sensitivity indices for each factor, the total variance *V* of the model output is decomposed as

$$V = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + \ldots + V_{12\ldots k} \tag{3}$$

where

$$V_i = V(E(Y|X_i = x_i^*)), \tag{4}$$

$$V_{ij} = V(E(Y|X_i = x_i^*, X_j = x_j^*)) - V(E(Y|X_i = x_i^*)) - V(E(Y|X_j = x_j^*)) \tag{5}$$

and so on. In the above formulas, Y denotes the output variable, X_i denotes an input factor, $E(Y|X_i = x_i^*)$ denotes the expectation of Y conditional on X_i having a fixed value x_i^*, and V stands for variance over all the possible values of X_i. The same variance decomposition underlies the theory of experimental design (Box, Hunter and Hunter, 1978). The decomposition is unique if the X_i are independent from each other. The sensitivity index S_i for the factor X_i is defined as V_i / V. The reason for that is intuitive: if the inner mean $E(Y|X_i = x_i^*)$ varies considerably with the selection of a particular value x_i^* for X_i, while all the effects of the X_j's, $j \neq i$ are being averaged, then surely factor X_i is an influential one. Estimation procedures for S_i are the Fourier Amplitude Sensitivity Test, FAST, (Cukier et al. 1973), the method of Sobol' (Sobol' 1993), and others (Iman and Hora 1990).

Higher order sensitivity indices, responsible for interaction effects among factors, are rarely estimated in computational experiments, as in a model with *k* factors the total number of indices (including the S_i's) that should be estimated is as high as $2^k - 1$. This problem is sometimes referred to as the curse of dimensionality. However interactions may have a strong impact on the output uncertainty especially when *k* is large and factors are varied on a wide scale, as often happens in numerical modelling.

A method, which is able of accounting for interactions and simultaneously coping with the curse of dimensionality, is the extended FAST (Saltelli et al. 1999). The extended FAST can yield estimates of the total sensitivity indices. S_{Ti} defined as the sum of all the indices (S_i and higher orders) where a given factor X_i is included. This concentrating in one single term all the effects involving X_i. For additive models, $S_i = S_{Ti}$ for all the

factors X_i. The estimation of the total sensitivity indices S_{Ti} makes the analysis affordable from a computational point of view, as we only need k total indices for decomposing quantitatively the output variance V. Furthermore, the extended FAST allows the simultaneous evaluation of the first and total effect indices. The estimation of the pair (S_i, S_{Ti}) is important to appreciate the difference between the impact of factor X_i alone on Y (S_i) and the overall impact of factor X_i through interactions with the others on Y (S_{Ti}). For a 3-factor model, the three total sensitivity indices are:

$$\begin{aligned} S_{T1} &= S_1 + S_{12} + S_{13} + S_{123} \\ S_{T2} &= S_2 + S_{12} + S_{23} + S_{123} \\ S_{T3} &= S_3 + S_{13} + S_{23} + S_{123} \end{aligned} \qquad (6)$$

where now each $S_{i_1,i_2,..,i_s}$ is simply $\frac{V_{i_1,i_2,..i_s}}{V}$. Clearly the $V_{i_1,i_2,..i_s}$ add up to V, and the corresponding $S_{i_1,i_2,..,i_s}$ add up to one; this is not true for the S_{Ti}'s, but a normalisation factor can be introduced.

Variance based methods such as Sobol' and the extended FAST display a number of attractive features for SA:

- Model independence: the sensitivity measure is model independent. It works for non-linear and non-additive models, unlike methods based on linear regression such as the standardised regression coefficients (Helton 1993).
- The measure captures the influence of the full range of variation of each factor.
- The measure captures interaction effects; this can be a crucial issue for a design problem, or for a risk analysis study.
- The methods can treat "sets" of factors as one single factor.

The last bullet means that the analysis can be performed by partitioning the k factors in a few subgroups and work on these rather than on the individual factors. It is sufficient, for a two-group example, to rewrite $Y = f(\mathbf{X})$ as $f(\mathbf{U}, \mathbf{Z})$ where now $\mathbf{X} = \mathbf{U} \cup \mathbf{Z}$, and apply to the two subsets $\mathbf{U}$ and $\mathbf{Z}$ the approach described above. The cost of obtaining S_{Tu}, S_{Tz} (and the corresponding first order terms) depends on the number of subgroups and not on the number of factors.

The added value of performing by an analysis on groups of factors is clear: in complex models uncertain factors might pertain to different logical levels, and it might be desirable to decompose the uncertainty according to these levels. For instance, aleatory and epistemic sources of uncertainty in the model could be appreciated separately[1].

2.4 Combined GSA-GLUE approach

The way of combining GSA and GLUE is straightforward. It is necessary that the sample generated for the GLUE analysis is designed also for the computation of variance-based sensitivity indices. So a Sobol' sample or a FAST sample should be used. In this way, by applying the same set of model runs, predictive uncertainty can be estimated, sensitivity indices computed and bootstrapping performed.

2.5 Model reduction technique

SA allows identifying, on one side, the relevant parameters, and on the other those who do not drive significant variation on the likelihood measure. As a consequence, SA is the basis for the elimination of unneeded complexity from the model and therefore for model reduction. Removing the unimportant factors/model structures, quantification of uncertainty can be performed with a reduced model. A big problem in this framework is usually given by the interaction structure of the model. The more complex is the structure, the more difficult is the decomposition of effects of input factors into elementary subsets. So, as already described in the preliminary considerations, the in-depth analysis of the interaction structure is a challenging task of our problem.

[1] Examples: aleatory (time to the next earthquake) and epistemic (earthquake frequent for a given severity and area). See Helton and Burmaster 1996.

3 Case study. WAMPUM model

3.1 Description of the WAMPUM model

A simple zero-dimensional model has been applied which describes the oxygen concentrations and major nutrient processes in the Elbe River at the Weir Geestacht (Schroeder, 1997). Period of study is 120 days (May 1 to August 28 1997). The model is run with a time step of 30 minutes.

A data set of daily data is available for model calibration (Petersen et al., 1999). Input data are meteorological time series and initial conditions for concentrations. Model outputs are time series of chlorophyll (CHL), phosphates (PO4), oxygen (O2). The Algal Biomass module of WAMPUM is analysed in the present study. Six input factors are considered:

- ✓ X_1 = depth: Water depth;
- ✓ X_2 = T_ref: Reference temperature;
- ✓ X_3 = k_light: Critical light intensity;
- ✓ X_4 = tendency_scale_factor: (for experimental purpose);
- ✓ X_5 = k_att_min: Non- algal light extinction coefficient;
- ✓ X_6 = k_att_shade: Algal self-shading coefficient.

Balance equation is expressed as:

$$\frac{\partial ALG}{\partial t} = X_4 \times (GROWTH - LOSS - RESPIRATION) \qquad (7)$$

1. *Algal growth term:*

$$\frac{\partial ALG}{\partial t} = X_4(GROWTH + ...) = X_4\left(G(X_1, X_3, X_5, X_6) \times fct(T - X_2, N, P) \times ALG + ...\right)$$

where

$$\lambda = X_5 + X_6 \times CHL_a \qquad CHL_a \propto ALG$$

function G describes growth limitation due to light conditions (I_O = global radiation):

$$G = \frac{1}{X_1}\int_0^{X_1} \frac{I_O e^{-\lambda z}}{\sqrt{X_3^2 + I_O^2 e^{-2\lambda z}}}\, dz$$

and *fct(T,N,P)* is the temperature dependence and growth limitation due to lack of nutrients (N,P).

2. *Algal Loss(sedimentation) term*

$$\frac{\partial ALG}{\partial t} = X_4(... + LOSS + ...) = X_4\left(... + C_1 \times fct(T - X_2) \times ALG + ...\right)$$

3. *Algal Respiration term*

$$\frac{\partial ALG}{\partial t} = X_4(... + RESPIRATION) = X_4\left(... + C_3 \times fct(T - X_2) \times ALG\right)$$

The reference model time series is shown in the continuous line of Figure 3.1 (parameter values of the reference curve are shown in the last column of Table 3.1). The measured data are the dots. A more detailed discussion of the observed data and their representation applying the WAMPUM model can be found in a companion report (Callies et al., 2000).

3.2 Description of the analysis

Factors

Six factors have been selected for the analysis, sampled from uniform distribution as shown in Table 3.1. The factors are statistically independent. Three samples have been generated: a FAST sample (dimension 3990), a Sobol' sample (256 for single factor, for a total of 3584 runs) and a pure Monte Carlo sample (dimension 4000).

Model output

The model considers three outputs: chlorophyll concentration, oxygen concentration and phosphate concentration. The model is run with a time step of 30 minutes for a total simulated time interval of 120 days (May 1 to August 28 1997). Two outputs are analysed: a time series (at 10 equally spaced times) and a likelihood measure. The likelihood measure is the same as defined in eq. (1) and is computed by considering available observed chlorophyll concentration time series.

3.3 GLUE analysis

Time dependent output

The confidence band for the prediction of the chlorophyll time series has been computed, based on the likelihood measures considered, with N=1, 4. This is shown in Figure 3.2.

With *N*=1, the low confidence bound remains quite distant from the lower bound of the observations, while the upper bound is very close to the observations. By considering *N*=4, the lower bound is translated closer to observed data, while the upper bound is below observations (i.e. inaccuracy for upper bound is increased) with a higher frequency than with *N*=1.

3.4 Sensitivity analysis

Sensitivity analysis has been done for both the time series and the likelihood measures.

Time series

FAST sensitivity indices for the time series of chlorophyll concentration are shown in Table 3.2-3 and Figure 3.3-4.

Parameter X_4 (`tend_sc`) is the most important parameter. It is the highest 1st order sensitivity index and also its total effect doubles the total effect of any other parameter. Among the remaining factors, X_1, X_3, X_5 (`depth, k_light_sm, k_att_min`) are the most influent, X_6 (`k_att_sh`) has a slight importance, while X_2 (`T_ref`) is not important. It is interesting to note that the trend of the sensitivity indices mimics the trend of the time series (Fig. 3.1-3.2).

Likelihood measures

The dotty plots are given in Figure 3.5-6. A significant pattern can be identified for tend_sc. It has to be expected that only tend_sc has a significant 1st order sensitivity coefficient and that interaction is mainly important. Following the results of the previous case study, this kind of portrait indicates that the model is over-parameterised and that the estimation problem is under-determined.

The sensitivity indices for the likelihood measure with *N=1* are shown in Figure 3.7 and Figure 3.8. As expected, the only significant main effect is detected for tend_sc, while a big interaction is obtained for all factors except T_ref (5 out of 6). This means that good runs (behavioural runs) are not driven by a particular factor, but by combinations of them. Moreover, T_ref is not important in all cases and can be surely fixed at a nominal value.

By increasing *N* (not shown here), the portrait is very similar, but the interaction structure is emphasised: the 1st order effect of `tend_sc` is smaller, while the difference between 1st order and total effect is increased.

Another interesting aspect is the difference between SA on raw data and on their likelihood. In particular, factors `depth, k_light_sm, k_att_min, k_att_sh`, have an almost zero effect for the likelihood measure, while for the physical output the effect is more significant. This is due to the fact that they induce a non-negligible increase in the variance of the algae concentration, but for most of such runs the likelihood is null. So a small variance in the likelihood can correspond to a large variance in the raw data. In this sense the GSA-GLUE analysis is much more informative about the model with respect to the analysis of raw data only.

3.5 Analysis of the correlation structure of the joint posterior distribution

This analysis aims at studying in more detail the properties of the posterior joint pdf of the input factors. This will show how any tool applied to represent in more detail the interaction structure confirms the basic features identified by the GSA. The basic idea is to study the covariance structure of the posterior distribution of the input factors obtained by applying the likelihood measure.

By normalising likelihood measures we obtain weights such as:

$$\sum_{i=1}^{n} w_i(\mathbf{x}) = 1, \tag{8}$$

where n is the sample size and $\mathbf{x}$ is the vector of input factors $\mathbf{x}=(x_1, x_2, \ldots, x_m)$.

The properties of the posterior marginal distributions can be evaluated as follows:

$$\hat{E}(x_j) = \sum_{i=1}^{n} x_{ij} w_i(\mathbf{x}) = \hat{\mu}_j \tag{9}$$

$$\hat{V}(x_j) = \sum_{i=1}^{n} x_{ij}^2 w_i(\mathbf{x}) - \hat{\mu}_j^2 = \hat{\sigma}_j^2 \tag{10}$$

and defining the new standardised factors:

$$\tilde{x}_j = \frac{x_j - \hat{\mu}_j}{\hat{\sigma}_j} \tag{11}$$

the correlation coefficients can be estimated as:

$$\hat{\rho}_{jk} = \sum_{i=1}^{n} \tilde{x}_{ij} \tilde{x}_{ik} w_i(\mathbf{x}) \in (-1,1) \qquad (12)$$

Such correlation coefficients would allow evaluating the pair-wise interaction structure, which is usually not observable from the total sensitivity indices. This analysis is similar to the approach presented by Hornberger and Spear (1981) but with a major difference: in RSA, correlation structure is analysed on the behavioural subset, while here all runs are used by applying different weights. This allows using the whole information.

Bootstrapping

Further inspection in the posterior joint pdf can be obtained by a bootstrapping procedure. Each parameter set can be sampled with a frequency proportional to the weight assigned by applying the likelihood measure (Russian roulette). As a result, a sample of the posterior joint pdf is obtained by using the same runs of the previous analyses.

The correlation coefficients of the posterior distribution have been evaluated via the bootstrapping procedure (sample size 4000), for N=1, 4. Values are shown in Table 3.5 and Table 3.6 respectively.

The correlation coefficients do not reflect any particular main interaction structure, other that T_ref is not correlated to the other variables. For N=4, the more selective classification procedure induces a higher correlation between X_1, X_3, X_5, X_6.

This kind of result implies some consequences about the analysis of the interaction structure of the WAMPUM model:

1. since coupled effects are not relevant, the interaction structure will be dominated by higher order effects;
2. hence, the interaction structure has a very high degree of complexity;
3. the application of conventional multivariate statistics like principal components analysis for analysing the posterior distribution will not be useful.

So, the analysis of the interaction structure of the present test case does not allow to 'summarise' into a simple scheme the interaction structure of 4 out of 6 factors (X_1, X_3, X_5, X_6). The only possible assessments are:

- ✓ unimportance of X_2 (T_ref): the role of GSA (total index) was essential, since first order effects (or scatter plots) alone could not allow this conclusion (who could distinguish between X_2 and [X_1, X_3, X_5, X_6] from scatter plots or 1st order sensitivity indices?);
- ✓ the relevance of X_4 (tendency scale factor), being the only factor directly affected by the conditioning on observations;
- ✓ model over-parameterisation and under-determination of the estimation/optimisation problem;
- ✓ very strong degree of interaction, of a degree higher than 2nd order.

A deeper comprehension of the interaction structure needs more powerful techniques. One possibility is the tree-structured density estimation technique by Spear. Another approach is to apply Bayesian networks (Pearl 1988). An example of this can be found in a companion report (Callies et al., 2000), where Bayesian networks are applied to describe the interaction structure of the six input factors of WAMPUM.

Another possibility lies always in framework of GSA and consists of computing the whole variance decomposition, in order to evaluate the importance of high order interaction terms.

Conclusions about the representation of the interaction structure

GSA allows a general, quantitative, model free identification of basic features of the interaction structure. On the other hand, it does not allow a complete representation of such a structure. Such a representation can be drawn applying other tools, which, on the other hand, require the introduction of more stringent assumptions about the interaction structure and have a less general applicability. In all cases, such representations confirm GSA results (in this case the interaction between the kinetic factors) and GSA, therefore, is a 'common denominator' to them.

In particular, the application of a GSA provides a quantitative evaluation about fundamental aspects of the calibration problem, such as:

- ✓ which factors are important for calibration, i.e. are somehow conditioned by observations;

- ✓ the degree of complexity of the interaction structure;
- ✓ which factors are involved in the interaction structure.

Such information has a general validity, since it is obtained without assumptions about the model structure and/or the error structure. So, even if it does not provide a complete representation of the interaction structure, GSA reveals some general and basic properties of such a structure, which are common to any more detailed representation and which are not affected by any "modeller's prejudice".

3.6 Complete GSA

In this section, the complete variance decomposition of the likelihood measure is considered, in order to verify the existence of privileged interaction structure of high order. This has been done computing the Sobol' indices of all orders, considering a sample size of about 32,000 runs. The elements in the variance decomposition, which are significant, are shown in Figure 3.7. The largest contribution is given by S_4 (i.e. the main effect of X_4). Moreover, interaction terms of $3^{rd}/4^{th}/5^{th}$ order between all factors except X_2 are present. No particular interaction term is prevailing, confirming the extreme difficulty, in the present test case, of identifying a clear structure for the set of factors [X_1, X_3, X_5, X_6].

Variable	Name	Minimum	Maximum	Reference
X_1	depth	1.2	4	2.5
X_2	T_ref	0	30	14
X_3	k_light	2	50	20
X_4	tendency_scale_factor	0.1	3	1
X_5	k_att_min	0.1	1.5	1
X_6	k_att_shade	0.002	0.01	0.005

Table 3.1. Factors of the analysis of the WAMPUM model.

Time (days)	10.4	20.8	31.2	41.6	52	62.4	72.8	83.2	93.6	104
X_1	0.29	0.09	0.11	0.06	0.08	0.09	0.06	0.11	0.13	0.10
X_2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X_3	0.08	0.14	0.03	0.02	0.09	0.03	0.02	0.05	0.15	0.03
X_4	0.17	0.29	0.31	0.42	0.44	0.39	0.49	0.31	0.30	0.43
X_5	0.09	0.09	0.06	0.03	0.06	0.05	0.02	0.06	0.10	0.04
X_6	0.02	0.02	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01
SUM	0.65	0.62	0.50	0.55	0.68	0.56	0.60	0.54	0.70	0.61

Table 3.2. Fast first order indexes for the Chlorophyll time series.

Time (days)	10.4	20.8	31.2	41.6	52	62.4	72.8	83.2	93.6	104
X_1	0.40	0.25	0.38	0.38	0.27	0.34	0.36	0.34	0.28	0.35
X_2	0.01	0.01	0.03	0.03	0.02	0.02	0.04	0.02	0.01	0.02
X_3	0.26	0.28	0.30	0.25	0.21	0.25	0.20	0.28	0.26	0.21
X_4	0.48	0.59	0.74	0.83	0.73	0.75	0.85	0.67	0.57	0.79
X_5	0.22	0.18	0.29	0.23	0.17	0.23	0.21	0.25	0.19	0.21
X_6	0.12	0.12	0.13	0.15	0.10	0.10	0.14	0.09	0.08	0.10

Table 3.3. Fast total order indexes for the Chlorophyll time series.

	1st order	Tot order
X_1	0.0147	0.623755
X_2	3.16E-05	0.020224
X_3	0.0321	0.501673
X_4	0.1562	0.725596
X_5	0.0227	0.634868
X_6	0.0026	0.290804

Table 3.4. FAST sensitivity indices for the likelihood measure with *N*=1.

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.0000	-0.0007	-0.1016	-0.0918	-0.1303	-0.0759
X_2	-0.0007	1.0000	-0.0163	-0.0261	0.0113	-0.0159
X_3	-0.1016	-0.0163	1.0000	-0.0642	-0.0902	0.0252
X_4	-0.0918	-0.0261	-0.0642	1.0000	0.0070	-0.0780
X_5	-0.1303	0.0113	-0.0902	0.0070	1.0000	-0.0432
X_6	-0.0759	-0.0159	0.0252	-0.0780	-0.0432	1.0000

Table 3.5: Correlation matrix of the posterior pdf for *N*=1 (estimated via bootstrapping).

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.0000	0.0626	-0.3422	-0.0343	-0.3894	-0.3220
X_2	0.0626	1.0000	-0.0960	-0.0264	-0.0104	0.0199
X_3	-0.3422	-0.0960	1.0000	-0.1756	-0.3513	0.0810
X_4	-0.0343	-0.0264	-0.1756	1.0000	-0.0860	-0.0070
X_5	-0.3894	-0.0104	-0.3513	-0.0860	1.0000	-0.2544
X_6	-0.3220	0.0199	0.0810	-0.0070	-0.2544	1.0000

Table 3.6. Correlation matrix of the posterior joint pdf for *N*=4 (estimated via bootstrapping).

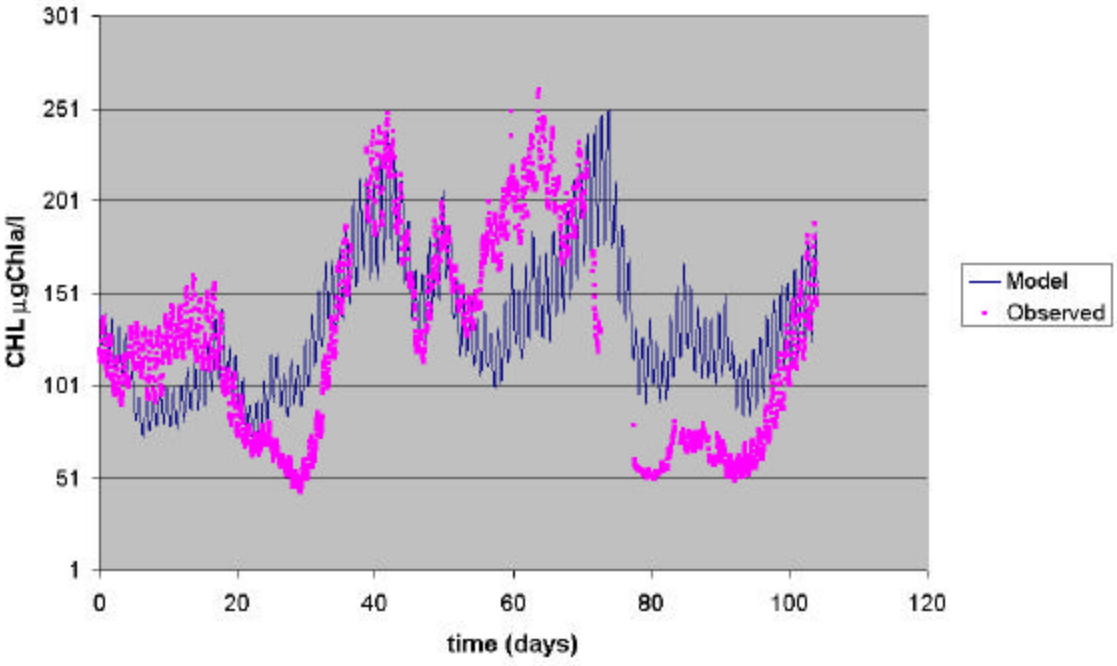

Figure 3.1: Model reference run and observations.

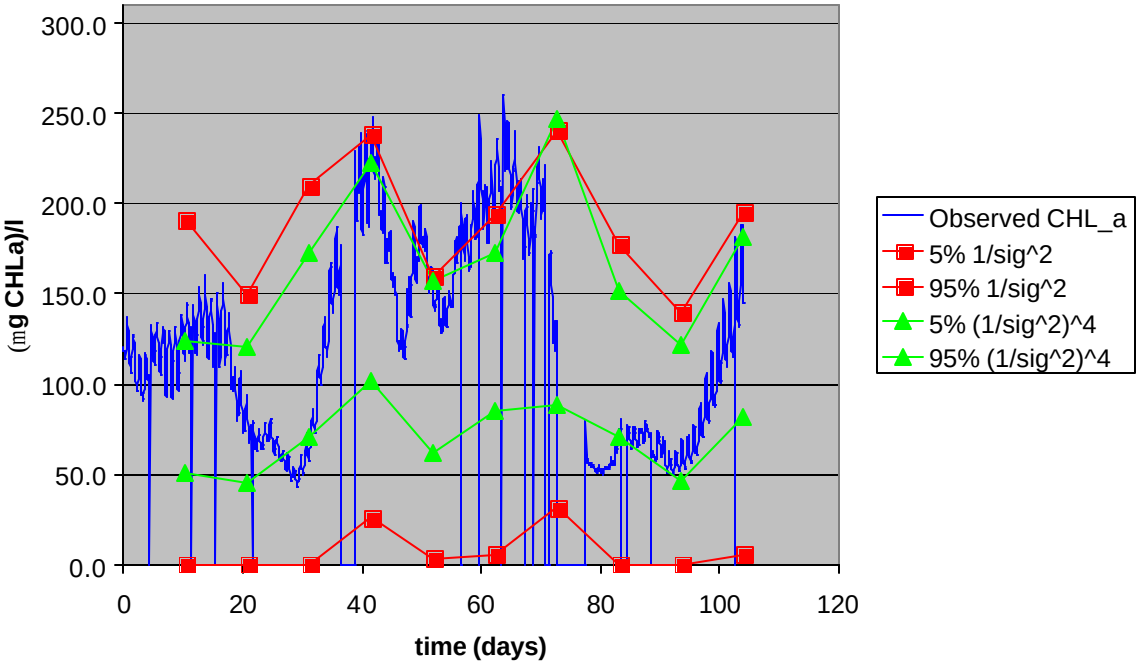

Figure 3.2. Confidence bands applying GLUE with *N*=1 and *N*=4.

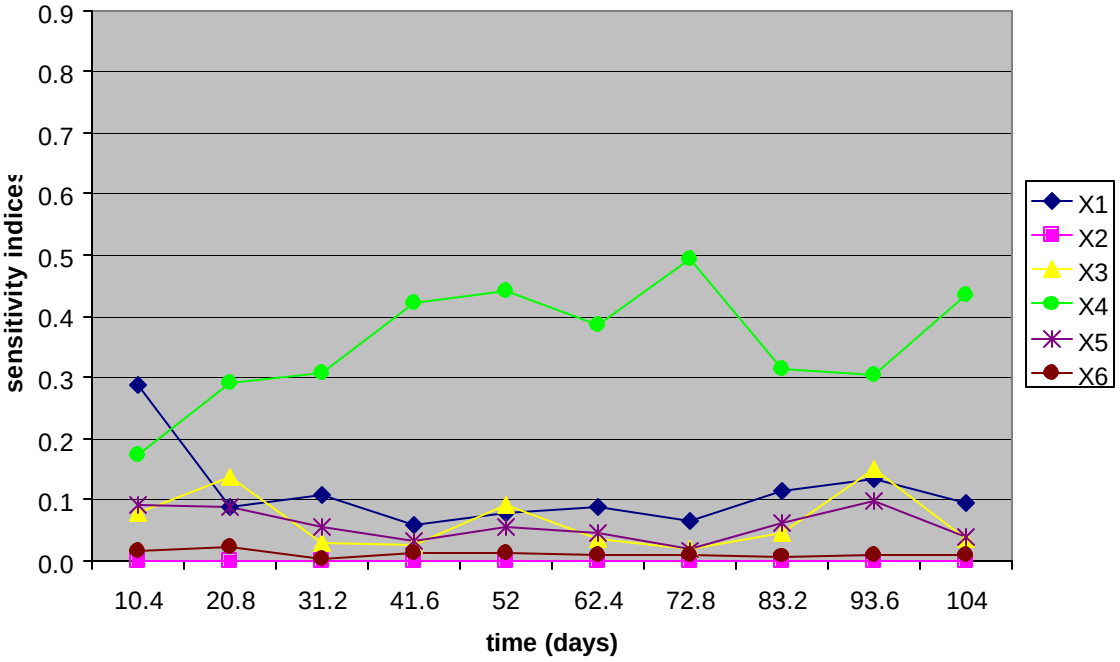

Figure 3.3: FAST first order sensitivity indices for chlorophyll concentration.

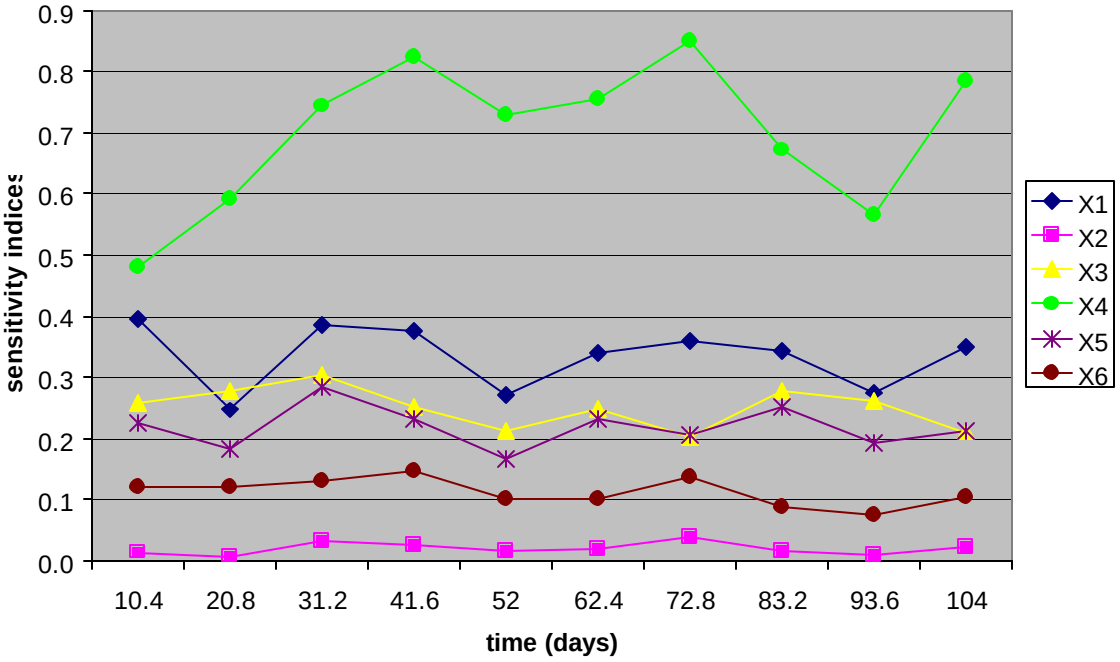

Figure 3.4: FAST total order sensitivity indices for chlorophyll concentration.

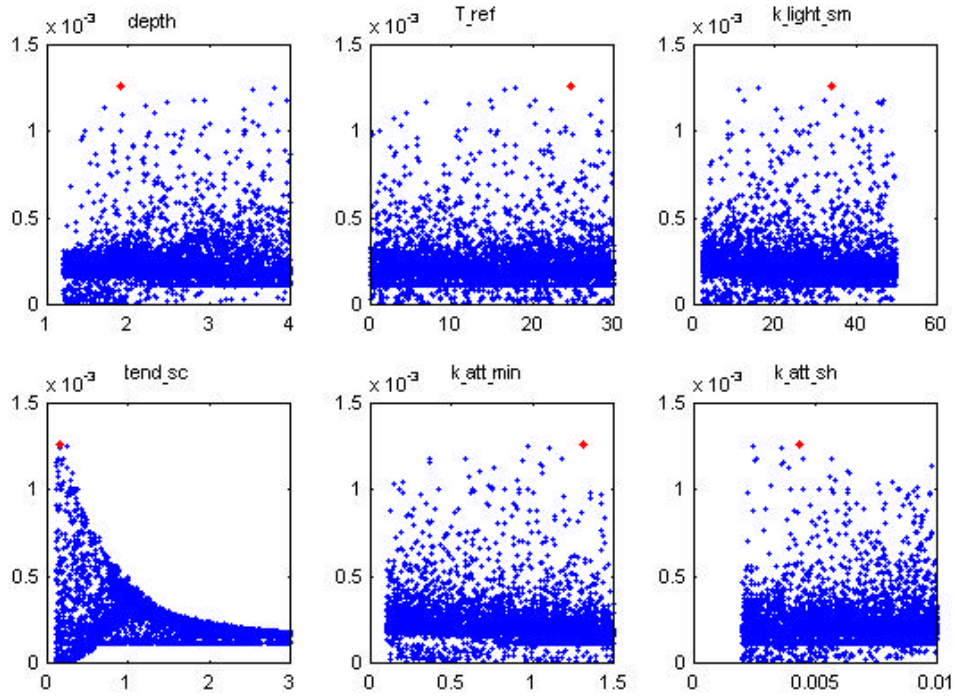

Figure 3.5. Scatter plots for the weights with *N*=1.

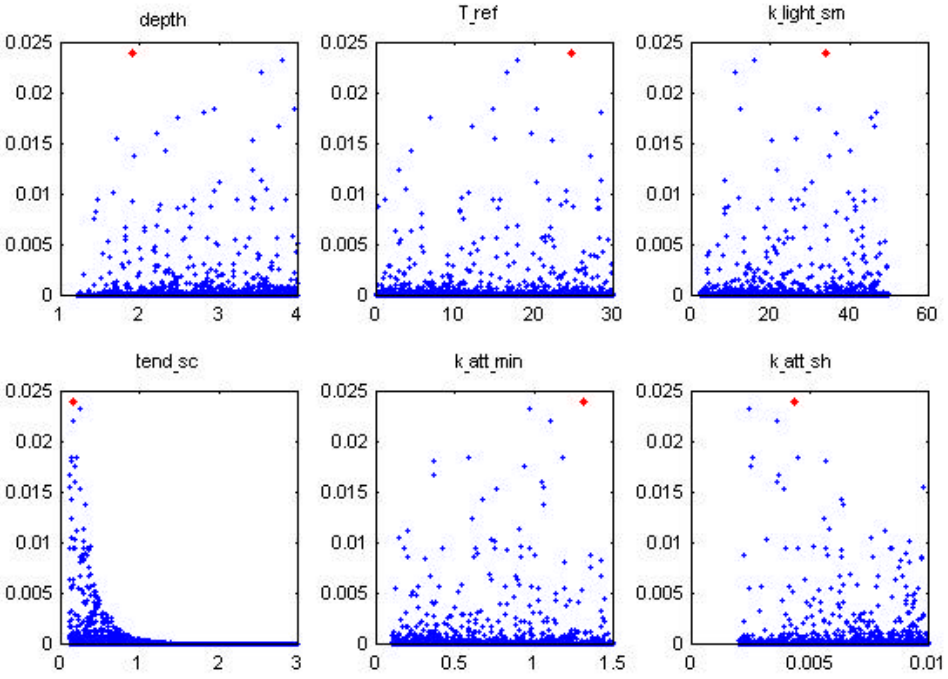

Figure 3.6. Scatter plots for the weights with *N*=4.

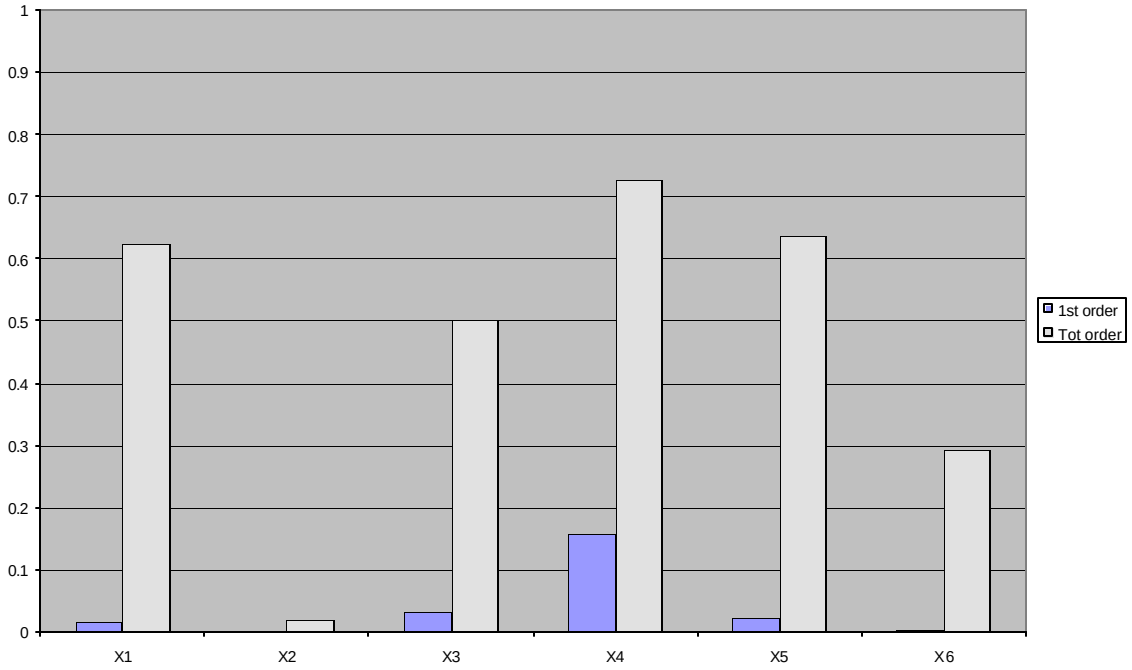

Figure 3.7. FAST sensitivity indices for the likelihood measure with *N*=1.

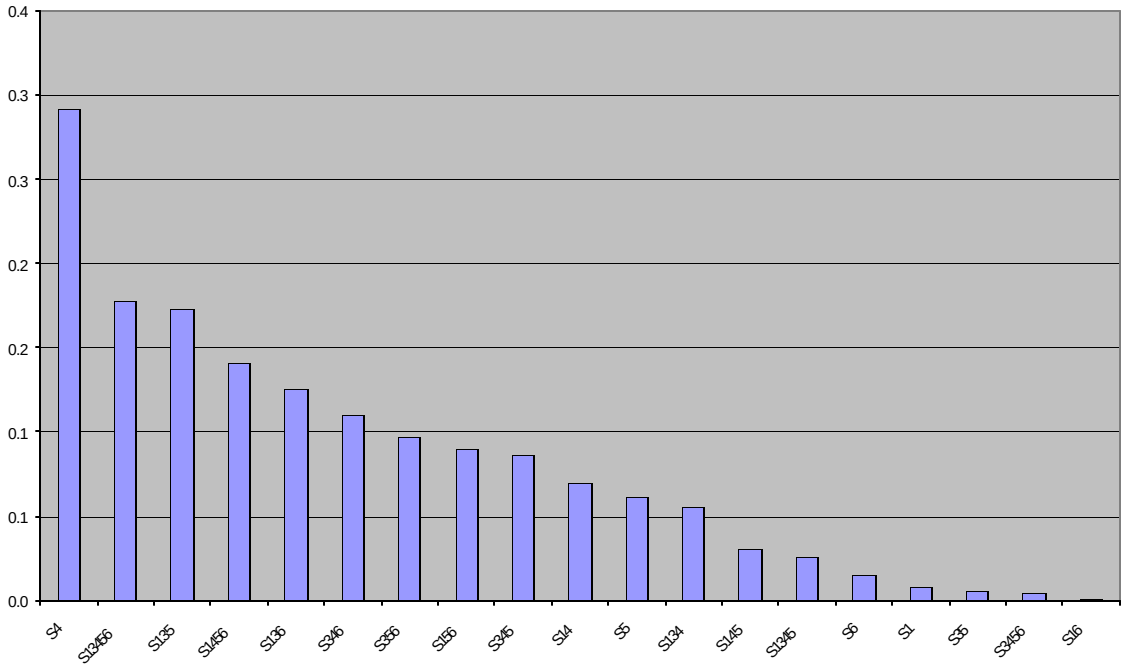

Figure 3.8. Elements of the complete variance decomposition of the likelihood measrue variance ranked by importance.

4 Concluding remarks

In the present report, a methodological approach is presented for the reduction of models of time-dependent output. The methodology consists of a combination of the GLUE and GSA methodologies. The use of GLUE allows both conditioning to observation and performing a sensitivity analysis on a scalar function (the likelihood measure) instead on a time-dependent output, so avoiding the calculation of time-dependent sensitivity indices. Moreover, GSA allows a quantitative decomposition of the likelihood variance with respect to the input factors, including high order terms. Factors providing negligible contributions to the likelihood variation can be clearly identified, allowing the modeller to exclude them from the calibration procedure and to fix them at a nominal value. On the other hand, factors having a significant impact on the likelihood measure (either as a main effect or as a total effect in interaction with all the other factors) have to be accounted in calibration, since they are able to drive behavioural runs of the model.

The model reduction technique is based on the results of the GSA. GSA allows partitioning the variation of the likelihood measure to the different input factors. So, factors providing negligible contributions to the likelihood variation are clearly identified and model reduction criteria are easily defined accordingly.

Model over-parameterisation usually implies that factors important for calibration hardly have an effect identifiable through elementary structures. On the other hand, a highly complex interaction structure is usually present. The analysis of the interaction structure is a challenging problem and a general method for assessing the posterior joint pdf is hardly to be identified. Some degree of arbitrariness in the construction of such tree-structures (Spear et al., 1994) or Bayesian networks (Pearl, 1988) cannot be avoided. Global SA can be very useful in this context, since it provides quantitative criteria for choosing 'leading' factors based on main and total effect. Such criteria do no necessarily provide a direct, complete representation of the interaction structure. However, the advantage of variance based GSA is that it makes few assumptions on the structure of the errors and of the input-output mapping. So, GSA results can be taken as a common denominator to all other tools applied to represent the interaction structure.

Applying the methodology to the Elbe River case study of IMPACT, enabled to assess the unimportant factors (T_ref) of the model. Moreover, the main features of the interaction structure of the remaining factors have been highlighted.

The methodological approaches identified will be applied in the further activities of the IMPACT project, where applicable.

5 References

Archer G., Saltelli A. and Sobol' I. M. (1997) Sensitivity measures, ANOVA like techniques and the use of bootstrap. Journal of Statistical Computation and Simulation 58, 99-120.

Beven K.J., Binley A., The Future of Distributed Models: Model Calibration and Uncertainty Prediction, Hydrological Processes, 6, 279-298, 1992.

Box, G. E. P., W. G. Hunter, J. S. Hunter (1978), Statistics for experimenters, John Wiley and Sons, New York.

Callies U., Blöcker G., Bülow C., Schroeder F., Ratto M., Normalisation involving mechanistic models: theoretical framework, IMPACT Deliverable 5, 2000.

Cukier, R. I., C. M. Fortuin, K. E. Schuler, A. G. Petschek, J. H. Schaibly (1973) Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory, *The Journal of Chemical Physics*, **59** (8), 3873-3878.

Helton J. C., and D. E. Burmaster, (1996), On the treatment of aleatory and epistemic uncertainty in performance assessment for complex systems, special issue: Reliability Engineering and System Safety, 54, 91-94.

Helton, J. C. (1993) Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal, *Reliability Engineering and System Safety*, **42**, 327-367.

Hornberger G.M., Spear R.C., Eutrophication in Peel Inlet, I. The problem-defining behavior and a mathematical model for the phosphorous scenario, Water Research, 14, 29-42, 1980.

Hornberger G.M., Spear R.C., An approach to the preliminary analysis of environmental systems, Journal of Environmental Management, 12, 7-18, 1981.

Iman, R. L., and Hora, S. C., (1990). A robust measure of uncertainty importance for use in fault tree system analysis. Risk Analysis, 10(3):401-406.

Jansen M. J. W., W. A. H. Rossing and R. A. Daamen, Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In

Predictability and Nonlinear Modelling in Natural Sciences and Economics, Gasman and van Straten (Eds.), 334-343. (1994).

Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of plausible Inference*, Morgan Kaufman Publishers, San Mateo, California, 1988.

Petersen W., Bloecker G., Mehlhorn N. and Schroeder F., Consequences of Altered Load of Pollutants on the Oxygen Budget of the Elbe River, *Vom Wasser*, 92, 37-50, 1999.

Ratto M., Tarantola S., Saltelli A., Sensitivity analysis in model calibration: GSA-GLUE approach, submitted to *Computer Physics Communications*, 2000a.

Ratto M., Tarantola S., Saltelli A., Model reduction techniques for quantification of uncertainty, IMPACT Deliverable 16, 2000b.

Romanowicz R., Beven K., Tawn J., Evaluation of predictive uncertainty in nonlinear hydrological models using a Bayesian approach. In *Statistics for the Environment 2*, Water Related Issues. Barnett V., Turkman F. (eds); pp. 297-315, Wiley: New York, 1994.

Romanowicz R., Beven K.J, Tawn J., Bayesian calibration of flood inundation models, in *Floodplain Processes*, M.G. Anderson, D. E. Walling, P.D. Bates, eds., John Wiley & Sons, 1996, pp. 333-360.

Romanowicz R., Higson H., Teasdale I., Bayesian uncertainty estimation methodology applied to air pollution modelling, Environmetrics, 11, 351-371, 2000.

Saltelli, A., K. Chan, M. Scott, Editors, (2000), Sensitivity analysis, John Wiley & Sons publishers, Probability and Statistics series.

Saltelli, A., S. Tarantola, K. P.-S. Chan, (1999), A quantitative model-independent method for global sensitivity analysis of model output, *Technometrics*, **41** (1), 39-56.

Schroeder, F. Water quality in the Elbe estuary: Significance of different processes for the oxygen deficit at Hamburg. *Environm. Model.&Assessm*, 2, 73-82, 1997.

Sobol', I. M., (1993), Sensitivity analysis for nonlinear mathematical models, *Mathematical Modelling & Computational Experiment*, **1**, 407-414.

Spear R.C. and Hornberger G.M., Eutrophication in Peel Inlet, II. Ideintification of Critical Uncertainties via Generalised Sensitivity Analysis, Water Research, 14, 43-49, 1980.

Spear R.C., Grieb T.M. and Shang N., Parameter Uncertainty and Interaction in Complex Environmental Models, Water Resources Research, 30, 3159-3169, 1994.

Spear R. C., Large simulation models: calibration, uniqueness and goodness of fit, *Environmental Modelling & Software*, 12, 219-228, 1997.