

UNCERTAINTY AND SENSITIVITY STUDIES OF MODELS OF ENVIRONMENTAL SYSTEMS

E. Marian Scott

Department of Statistics
University of Glasgow
Glasgow G12 8QW
UNITED KINGDOM

ABSTRACT

Traditionally, sensitivity and uncertainty analyses in environmental modelling contribute quantitative descriptions of the relative importances of individual parameters and processes, highlighting areas of significant contributors to the overall uncertainties of model predictions and giving markers for areas requiring substantial improvement (perhaps through directed experimental work). Both sensitivity and uncertainty analysis require extensive re-sampling from input data and simulation of model response and there is a large and growing literature concerning their use in environmental modelling. There is however growing interest in the sensitivity of predictions to model structure, and in evaluating the contribution of model structural uncertainty to the overall uncertainty within the general framework of sensitivity analysis (Draper, 1995, Beven, 1993, Beven and Binley, 1992).

Sensitivity and uncertainty analysis contribute to all stages of model development, testing and assessment and their impact on model reliability and validity will be described. After a general introduction to sensitivity and uncertainty analysis and discussion of their contributions, some applications of sensitivity analysis to several environmental modelling studies will be presented.

2 INTRODUCTION

An environmental model will very often be composed of a number of linked sub-models, representing physical processes understood to varying degrees. They may be dynamic and stochastic, and the model extent may also be spatial. Such models may be used for example, to predict levels of Cs-137 on pasture over several seasonal cycles following the Chernobyl power plant accident; they may be used to model the global carbon cycle in

climate studies and to model the dispersal of a radioactive pollutant discharged into the Irish Sea from Sellafield nuclear fuel reprocessing plant. The model, depending on its purpose and field of application will include many different processes, the sub-processes may operate at quite different time and space scales and for the modeller there will be choice of which processes to explicitly model and how to parameterise these processes. Some parameters will be known precisely, for others, there may be conflicting evidence so that the parameter can only be defined in terms of a range of values which may span several orders of magnitude. Some of the parameters may be time- and space-dependent.

Sensitivity analysis encourages the exploration of the interactions between the various modelled processes and helps throw light on the properties of complex computational models by in its simplest form perturbing one parameter at a time and studying its effect on the response. For large computer models, with perhaps hundreds of parameters, sensitivity analysis may involve perturbing all parameters simultaneously, but in such a way that the main effects of individual parameters and their interactions can be estimated. Uncertainty analysis contributes estimates of uncertainties on the final predictions by propagating through the model all quantified sources of uncertainty. Thus, sensitivity analysis identifies the key contributors to uncertainties, while uncertainty analysis quantifies the overall uncertainty, so that together they contribute to a reliability assessment for the model.

3 MODEL RELIABILITY EVALUATION

Modelling the environment may be done with a number of goals in mind: they include description of the system, prediction, impact assessment and to provide a basis for management decisions. It is crucial that the

model should be fit for the purpose to which it was designed.

There are a number of properties the modeller wishes the model and the model results to possess, but one of the most important is reliability. What properties might be considered to define model reliability? One fundamental requirement is that the model should reproduce the behaviour of the system under study, within the limits of uncertainty of the data and the model predictions.

As part of this requirement, therefore, an assessment of the uncertainty associated with the model predictions is required. This uncertainty should be acceptable for decision making purposes and it should also be possible to identify the factors contributing to this uncertainty.

3.1 Factors affecting reliability

The factors affecting reliability can be broadly classified corresponding to the main stages in model development: specification, conceptualisation, computation and parameterisation (IAEA, 1989).

In the specification of the problem, the modeller must consider the spatial and temporal resolution required in the predictions, the purpose of the modelling and the scales at which the predictions will be required. He must also consider the processes which operate.

He must formulate an appropriate conceptual model, defining the model structure, processes and the scale at which they operate and any interconnections.

The conceptual model must be translated into a computational model, whose code must be verified.

Finally, any unknown parameters must be defined. The parameters may have associated with them large uncertainties, arising from two sources: natural stochastic variation and lack of knowledge on the modeller's behalf (Hoffman & Hammonds, 1994).

Thus, at each stage of the modelling process uncertainties are introduced, not only in the definition of the parameters. Sensitivity analysis contributes at each stage and only after these stages have been completed, can the critical stage of model testing and validation be reached. The testing and validation must take into account the uncertainties introduced in the preceding stages (McKay, 1995).

3.2 The role of sensitivity and uncertainty analysis in model development

3.2.1 Parametric sensitivity analysis

Sensitivity analysis involves determination of the amount and kind of change produced in a given system

parameter by a change in another parameter. It is a tool in the investigation of factors which are important or influential in determining the system response. To perform a sensitivity analysis, it is necessary to first identify the parameters of interest, define for each a probability density function (p.d.f.) to reflect the belief that the parameter will take on various values within its possible range, account for dependencies amongst the parameters and propagate the uncertainties through the model to generate a p.d.f. of predicted values which can then be analysed.

Definition of the parameter p.d.f. can often prove difficult, occasionally there may be experimental information available which would allow empirical based estimation of the p.d.f., but more commonly there may be little information and the p.d.f. then fundamentally reflects the modeller's belief. The elicitation of belief based p.d.f.s in environmental contexts is a challenging task, but which if successfully achieved raises the possibility of directly incorporating expert opinion into the formal analysis through Bayesian methodology (Cooke, 1994a,b, EC Munvar report, 1995).

Given the p.d.f., there are a number of widely used methods of sampling, they include simple random sampling (SRS), (parameter values chosen at random from the p.d.f.) and Latin hypercube sampling (LHS), (the range of each parameter is partitioned into n intervals and one value is selected from each interval of each parameter). SRS is simple, reliable easy to analyse but inefficient. LHS is basically a stratified sampling procedure and is generally more efficient than SRS (Andres, 1987, Andres, 1997, Iman and Conover, 1980).

A large number of methods for the analysis of the results from the sensitivity analysis exist, and indeed there have been a number of comparisons of the methods (Iman and Helton, 1988, Saltelli and Homma, 1992, Hamby, 1995). Techniques used include response surface replacement, correlation and partial rank correlation coefficients, regression (on ranks) as well as a number of standard test statistics (Smirnov, Vramer-Von Mises and Mann-Whitney), (Saltelli et al, 1992). Whatever the method, its goal is to explore the model response to changing parameters.

However, parametric uncertainty reflects only one source (and perhaps not the dominant source) of uncertainty. The other key contributor is uncertainty about the model structure itself.

3.2.2 Model uncertainty

In the development of the model, the modeller must work with an imperfect and incomplete description of the physical system. He must select features and

processes to be parameterised, he must synthesise sometimes conflicting evidence and he must prioritise.

Model or structural uncertainty reflects that an acceptable model prediction may be achieved in many different ways and that the modeller must use judgement to decide on the structure of the model. In an attempt to explicitly consider this source of uncertainty in the predictions (particularly important when the model is being used as a management decision tool), a number of problems must first be resolved. It may be more difficult to quantify the uncertainty in the model structure when compared to a single parameter.

Models may differ in terms of the physical, chemical or biological processes included; they may have different state spaces; they may have been constructed using a different knowledge base; the complexity of the models may thus be different, yet all may be valid but still provide different responses. The issue of how uncertainty about model structure can be quantified and incorporated in a standard sensitivity analysis is important one which is yet to be fully addressed. The issue of structural uncertainty is one which can be addressed using a Bayesian formulation, within which, it is possible to elicit and make use of expert beliefs. Following Draper (1995):

let M denote the model, S the structural assumptions and Θ the model parameters. Suppose we have a discrete set of models, say m . Let x denote the data and y the quantity to predict, then

$$p(y | x, S) = \sum p(S_i | x) p(y | x, S_i) \quad (1)$$

and S_i , $i=1, \dots, m$ denotes the different structural alternatives within the models. The analysis presented in (1) allows, through the application of Bayes Theorem, the distribution of the predicted quantity to be defined in terms of prior beliefs $p(S_i | x)$ for the model structure given the data and the predictive distribution for y given the model structure S_i . However, the prior distributions $p(S_i | x)$ must be defined. Draper suggests creating cross-validation or bootstrap samples (Efron and Tibshirani, 1993) from the available data and conducting parallel modelling activities on each sample. However, in many environmental studies, there is little available data so that this may not be a viable alternative. A simple alternative to this approach which makes use of all the data, but rather re-samples from the modellers has often been used in international model inter-comparisons. Modellers are given access to the same information, and are tasked to independently develop models for the same, defined purposes. In this way, a more robust and hopefully reliable set of predictions can be achieved. In addition, the variability in the model

predictions provides a direct measure of uncertainty based on structural uncertainty.

4. EXAMPLES

A number of recent modelling studies have had as a stated objective an assessment of the uncertainty in predictions due to differences in model structures. In such work, often described as 'model benchmarking', a number of modellers have been provided with the same information concerning physical environment as well as clear modelling objectives. Model endpoints have been selected and the results compared both across the different models and also, when possible, with observational data. Results from three such studies are presented and discussed.

4.1 IAEA VAMP programme (IAEA, 1995)

The International Atomic Energy Agency programme had as one of its objectives:

'to provide a mechanism for validation of assessment models using Chernobyl related data'. As part of this work, a further objective was defined: 'to test predictive capabilities of models and identify the most important reasons for model misprediction'. The study was designed in the form of a test scenario, describing in some detail, conditions in a location (unknown to modellers) in central Europe. A wide range of model endpoints were identified including Cs-137 concentrations in a number of agricultural products at different points in time and also an estimate of dose to the local population. Modellers provided their best estimate as well as an estimate of the uncertainty on the result. Finally the model results could be compared both internally and against experimental data from the region (Hoffman & Thiessen, 1995, Hoffman et al, 1996).

The results showed variation amongst the predictions, and differences between the model predictions and experimental data. On further analysis, one of the main sources of variation in the results was found to be 'modeller interpretation' (IAEA, 1995). In addition to best estimates, the modellers also provided in some instances, uncertainty estimates, one of the main sources of difference in these estimates was the judgement on how representative the measured data.

4.2 BIOMOVs (1993)

BIOspheric Model Validation Study (BIOMOVs) was an international study for testing models for ecological transfer and bioaccumulation of radionuclides and other

trace substances. One of the primary objectives of BIOMOVs was 'to explain differences in model predictions due to structural deficiencies, invalid assumptions...' (BIOMOVs, 1993). Within the programme, one approach adopted involved the formulation of test scenarios based on suitable data and a comparison of model predictions against these independent data sets, while for the other approach, model predictions and associated estimates of uncertainty were compared. As in VAMP, when the model comparisons were made substantial differences were found, related in part to the expertise and judgement of the modeller. More recently, the BIOMOVs programme has developed a number of test cases to further investigate the role of expert judgement in determining uncertainty.

4.3 IAEA Arctic assessment programme (Scott et al, 1997)

The Arctic assessment programme was concerned with evaluating the present and potential impacts of the dumping of nuclear waste within shallow coastal waters in the Arctic (specifically the Kara and Barents Seas). The waste of most concern was reactor compartments from nuclear submarines, some of which still contained spent nuclear fuel. A modelling group was created with the task of providing reliable and realistic estimates of the potential radiological impact on biota and human populations.

The modelling group comprised 9 modellers, and in the first phase of the work they undertook an extensive benchmarking exercise, with a view to providing provisional estimates of the uncertainties in the predictions due to differences in the model structure. The endpoints of the benchmarking were spatially and temporally discriminating, and it was of interest to investigate the variation in the endpoints.

The models used in the work simulate the dispersion of radioactivity due to advection and diffusion within the water column and include interactions with suspended material and sediment. There were two main modelling approaches, one was based on compartmental models while the other used detailed 3-D hydrodynamic models.

In addition to a quantification of predictive uncertainty as a result of model structure, a directed sensitivity analysis was also carried out on sub-processes. Results again emphasised the importance of structural uncertainty in the overall predictive uncertainty.

5 SUMMARY AND CONCLUSIONS

Sensitivity and uncertainty analysis are important tools in the modelling toolkit. They should be performed routinely, as the modeller might assess the goodness of fit or check on outliers.

In model assessment and testing, sensitivity and uncertainty analysis provide the essential information for determining the acceptability of the uncertainties and the influential factors.

In model validation, the agreement of the model to observed or measured data can only be judged in the light of the predictive uncertainties.

In model design, sensitivity analysis encourages the identification of key processes to be further investigated and gives a metric by which to compare different model structures.

Taken together, sensitivity and uncertainty analysis contributes to a more reliable model, and one whose abilities are better understood.

REFERENCES

- Andres T H, 1987. Statistical sampling strategies. Procs NEA Workshop on Uncertainty analysis for Performance Assessments of Radioactive Waste Disposal Systems, Seattle 1987.
- Andres T H, 1997. Sampling methods and sensitivity analysis for large parameter sets. *Journal of Statistical Computation and Simulation*, to appear.
- Beven K, Binley A, 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279-298.
- Beven K, 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Research* 16, 41-51.
- BIOMOVs, 1993. Final report, BIOMOVs Technical report 15, Swedish Radiation Protection Institute, Stockholm.
- Cooke R, 1994a. Parameter fitting for uncertain models: modelling uncertainty in small models. *Reliability Engineering and System Safety* 44, 89-102.
- Cooke R, 1994b. Uncertainty in dispersion and deposition accident consequence modelling assessed with performance-based expert judgement. *Reliability Engineering and System Safety*, 45 35-46.
- Draper D, 1995. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society, Series B*
- EC, 1995. Review on development of methodologies for modelling with uncertainty and variability: Munvar project. European Commission report, EUR 16174 EN.

- Efron B, Tibshirani R, 1993. An introduction to the Bootstrap. Chapman & Hall, New York.
- Hamby D M, 1995. A comparison of sensitivity analysis techniques. *Health Physics*, 68(2), 195-204.
- Hoffman F O, Hammonds J S, 1994. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis* 14, 707-712.
- Hoffman F O, Thiessen K M, 1995. Use of Chernobyl data to test predictions and uncertainty estimates from exposure assessment models. Proceedings of 'Environmental impact or radioactive releases', IAEA-SM-339/20, 325-336.
- Hoffman F O, Thiessen K M, Watkins B, 1996. Opportunities for the testing of environmental transport models using data obtained following the Chernobyl accident. *Health Physics* 70, 5-7.
- Iman R L, Conover W J, 1980. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. *Communications in Statistical Theory and Methods*, A9(17), 1749-1842.
- Iman R L, Helton J C, 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* 8, 71-90.
- International Atomic Energy Agency, 1989. Evaluating the reliability of predictions made using environmental transfer models. *Safety Series*, 100.
- International Atomic Energy Agency, 1995. Validation of models using Chernobyl fallout data from the central Bohemia region of the Czech republic. IAEA-Tecdoc 795.
- McKay M D, 1995. Evaluating prediction uncertainty. Los Alamos National Lab Technical report, NUREG/CR-6311.
- Saltelli A, Homma T, 1992. Sensitivity analysis for model output. *Computational Statistics and Data analysis*, 13(1), 73-94.
- Saltelli A, Andres T H, Homma T, 1992. Sensitivity analysis of model output: an investigation of new techniques. *Computational Statistics and Data Analysis*, 14, 1-28.
- Scott E M, Gurbutt P, Harms I, Heling R, Nielsen S P, Preller R, Osvath I, Sazykina T, Wada A, Sjoebloom K L, 1997. Benchmarking of numerical models describing the dispersion of radionuclides in the Arctic Seas. *Science of the Total Environment*, to appear.

AUTHOR BIOGRAPHY

E. MARIAN SCOTT is a senior lecturer in the Department of Statistics, University of Glasgow. She is associate editor of the *Journal of Statistical Computation and Simulation*, and will be guest editor of a special issue on Sensitivity and Uncertainty analysis. Her research interests include environmental modelling, sensitivity analysis and model uncertainties.