ERROR AND UNCERTAINTY

This chapter identifies the issues associated with error and uncertainty. It reviews the difficulties with collecting field data as well as the potential problems with analysis and interpretation of outputs. Within this chapter, there is a section on assumptions, and the techniques used with establishing models. Furthermore, there is a discussion about uncertainty and uncertainty analysis and the ways in which confidence levels are used in statistics to outline levels of certainty in model outputs.

Errors

Field measurements in the marine environment are notoriously difficult because of its dynamic nature. As a result, errors in data collection are common but can be reduced. The types of error that can arise are:

- Instrument error, which reflects the limits of the measuring device in a controlled environment such as a laboratory;
- Movement of the measuring platform, such as the boat;
- Sampling limitations, such as frequency of measurement or spatial coverage;
- Background noise in the measured signal; and
- Operator error particularly when working in adverse weather conditions.

Nonetheless, there is generally reliance on field measurements, providing they have been collected with suitable quality control and are indicating physically realistic values. The results from field sampling are usually utilised within physical models to aid calibration and interpretation.

Model output and measurements often have very different spatial and temporal resolutions. This is because a measurement may be taken at a point in space but a model is averaging over an area or volume. Similarly, most instruments have a sampling frequency that produces an average value over some selected period. A model outputs values at discrete time intervals (although it can be argued that the time step used in a model is a form of temporal averaging). Consequently, differences between model and measured values combine the errors of both, as well as any differences due to spatial and temporal resolution. This issue is discussed more fully in a paper on data and models by Cunge (2003) and also in Data Requirements.

Assumptions

Assumptions are an inherent part of all assessments and modelling exercise, largely because knowledge of both the natural and the built environment is incomplete, and progress has however been based on idealised assumptions.

The same can be said for component models, such as the tide and wave models, which have been extensively researched and successfully applied to a wide range of situations all around the world. Field measurements and laboratory experiments have been used to check the predictions and develop an understanding of the range of applicability. This is by no means complete and the degree of knowledge varies with the complexity of the problem. For instance tide and wave models have been extensively tried and tested. Sediment transport models are more complex, combining more processes and seeking to represent more non-linear interactions. There is therefore a larger error likely in the predictions from such models. Particular attention needs to be given to specific site conditions, to ensure that the relevant processes are represented in the modelling process. None the less, these models have now been used extensively and there is a good body of knowledge concerning their use (HR Wallingford *et al*., 1996; 1994).

The area most recently researched is that of long-term prediction of morphological change, at both estuary and sub-component scales (e.g. individual mudflats). At its most simple, it is the application of sediment transport models for long enough to predict how the seabed changes. This is known as the *bottom-up* approach because it builds the results from the component processes. Unfortunately the errors in prediction, coupled with the potential for non-linear interactions that are chaotic in nature, place limits on the approach's ability to provide predictions over very long time periods (greater than a few years). One way of overcoming this is to develop a probabilistic description of the most likely outcomes.

An alternative approach has been to look for "target" states that the system or feature is trying to reach. This is known as the *top-down* approach because it seeks to establish a system view of change and equilibrium, rather than examining the internal detail. In due course, it may be possible to define the various possible states for the system and provide a probability that they will actually occur. This is equivalent to finding the system's possible positions on a fitness landscape, where the valleys represent stable positions for the system and perturbations are required to move the system over the peaks from one valley to another (Gell-Mann, 1995; Kauffman, 1993). Presently, there must be a reliance upon combining the various techniques that are currently available in order to synthesise a consensus of the likely outcome, or range of outcomes.

This is the procedure outlined in the section on Study approach and is the reason why such a wide range of different techniques and models may need to be applied to study the likely impacts of a development. To appreciate the assumptions that underpin the application of the various models, there needs to be understanding of the way the different models are formulated and how they have idealised, or simplified, the real world (e.g. the method of approximating the effects of turbulence in a tidal flow model). This is extensively covered in the literature (HR Wallingford *et al.*, 1996). Subsequently, this understanding informs the way in which particular models are selected for a given problem and how they are applied. For complex problems it is sometimes useful to document the key assumptions and to assess the consequences of making different assumptions.

The modelling approach is robust for some problems, and the results can be considered reliable. Whether these reflect what will actually happen, does however depend on whether a representative set of events has been modelled. For static structures, such as the construction of a quay wall or the deepening of a channel, this is quite straightforward. However, it may be much more difficult to be certain about the methods of construction that will be used. Thus, for example, the way in which dredging is carried out could critically affect sediment dispersion. One way round this is to identify acceptable levels of disturbance (e.g. levels of suspended sediment that can be tolerated) as a basis of defining thresholds. Modelling can be used to evaluate whether it is possible to maintain conditions below the critical thresholds, using accepted construction practices. The thresholds then provide a basis for monitoring activities and controlling impacts, in the knowledge that it should be possible for a contractor to meet them. Further discussion of this topic is found in the section on Assessing impacts in the context of impact assessment.

Uncertainty

Predictions of future change are hard to make because of the level of uncertainty in the environment. This means that when predictions are made about behaviour, there may follow a host of uncertainties. Individual models are limited by their inherent assumptions and the ability to prove the models is constrained by the difficulty and expense of obtaining good quality field measurements. This does not mean that the models are bad or of no use. Quite the contrary, they are significantly better than assertion or untested conjecture. The models are built around testable laws and hypotheses. By using them in the manner described they can help to build understanding of specific aspects of the problem, often with known levels of possible error. Therefore, it is possible to state the probability of a given outcome and the probability of being wrong. Combining the results from different models and methods of analysis allows the understanding to be expanded from one limited perspective to encompass a range of considerations. If this is done around the conceptual model of the system, this should allow confidence in the predictions or forecasts to be progressively developed.

There are many definitions of uncertainty. Perhaps the simplest and most complete is that: "Uncertainty is a general concept that reflects our lack of sureness about something or someone, ranging from just short of complete sureness to an almost complete lack of conviction about an outcome" (NRC, 2000).

Communication of uncertainty

In recognising uncertainty, there is an acknowledgement of the lack of knowledge of the behaviour of the physical world (knowledge uncertainty), its inherent variability (natural variability) and the complexity of the social/organisational values and objectives (decision uncertainty). Consideration of uncertainty within the decision process attempts to quantify a lack of sureness, and thereby provide the decision maker with additional information on which to base a decision. By investigating the sources of uncertainty, this type of analysis enables the decision-maker to identify the uncertainties that most influence the final outcome and focus resources efficiently. Understanding the sources and importance of uncertainty helps the decision maker to make more informed choices. However, uncertainties arise at every stage in the decision process (Sayers *et al*., 2003).

These uncertainties can be expressed in a number of different ways, both qualitative and quantitative (Sayers *et al*., 2003). These include:

- Deliberate vagueness, 'There is a high chance of breaching';
- Ranking without quantifying, 'Option A is safer than Option B';
- Stating possible outcomes without stating likelihoods, 'It is possible the embankment will breach';
- Probabilities of events or outcomes, 'There is a 10% chance of breaching';
- Range of variables and parameters, 'The design flow rate is 100 cumecs ±10 %'
- Confidence intervals, 'There is a 95% chance that the design flow rate lies between 90 and 110 cumecs';
- Probability distributions.

Two of the most widely used quantitative expressions of uncertainty are confidence intervals and probability distributions and these are discussed below.

Confidence intervals

A confidence interval specifies the probability that a variable falls within a range of values. For example, there is a 95% (this is the confidence level) chance that the design flow rate lies between the confidence limits of 90 and 110 cumecs. Confidence intervals can be formally calculated for some forms of uncertainty, e.g. statistical inference uncertainty. However, expert judgement can also be applied to specify confidence intervals. For example, an experienced wave modeller may judge the model output to provide results that are accurate to within ± 10%. The modeller may be able express this accuracy with a probability (90% for example) that reflects his strength of belief in the model results, based on the quality of the calibration procedure.

A confidence interval does not provide any information regarding how the probability of achieving different values within the range may vary. Using the wave modelling example above, although the interval has been specified as being symmetrical, the modeller may know from experience that the model is more likely to under predict than over predict. A symmetrical confidence interval does not contain this information and an asymmetrical description may be provided.

Probability distributions

A probability distribution describes the probability of obtaining different values of a variable or parameter and hence the associated uncertainty. Probability distributions can be discrete or continuous and a frequently used continuous probability distribution is the Normal, or Gaussian Distribution. Many natural phenomena conform well to the Normal Distribution, which makes it particularly useful.

Sources of uncertainty

Sources of uncertainty can be classified by natural variability and knowledge uncertainty. Natural variability refers to the randomness observed in nature and is also referred to as aleatory uncertainty (meaning to ‘gamble’), examples include:

- External uncertainty;
- Inherent uncertainty;
- Objective uncertainty;
- Random uncertainty;
- Stochastic uncertainty;
- Irreducible uncertainty;
- Fundamental uncertainty; and
- Real world uncertainty.

Knowledge uncertainty refers to the state of knowledge of a physical system and the ability to measure and model it, and is also referred to as epistemic uncertainty (meaning ‘knowledge’), examples include:

- Functional uncertainty;
- Internal uncertainty;
- Subjective uncertainty; and
- Incompleteness.

More information can be found in Sayers *et al*. (2003).

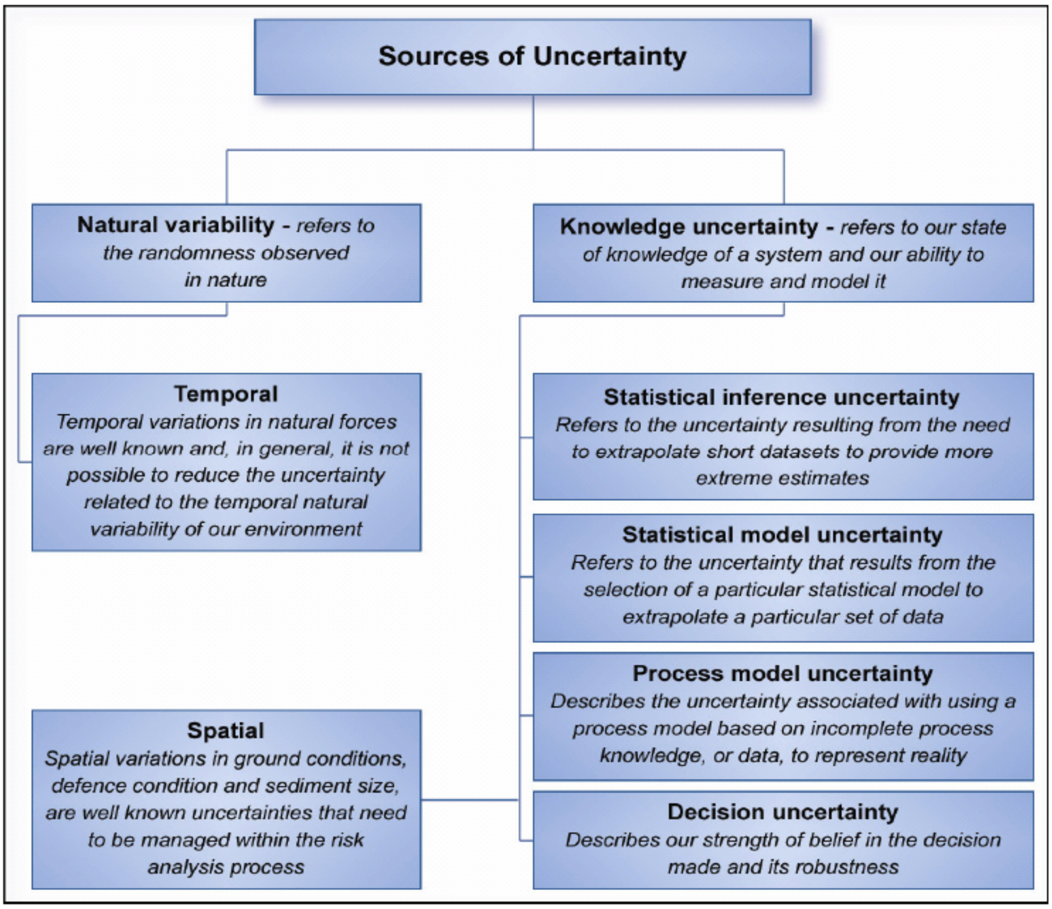

Figure 1. Generic sources of uncertainty (in the decision-making process)

Uncertainty analysis

A range of methods of uncertainty analysis are included here, however, for more information see the FloodRiskNet Uncertainty and Risk wiki at: http://www.floodrisknet.org.uk/.

Forward uncertainty propagation

Forward Uncertainty Propagation methods propagate uncertainty using prior assumptions about the different sources of uncertainty without the use of additional evaluation data. The assumptions that need to be made normally include prior distributions for parameters and other inputs. No model evaluation is necessary to apply forward uncertainty propagation although forward uncertainty propagation is often applied to an optimal model after a calibration exercise or some "best estimate" model. It is evident that for nonlinear models the results of a forward uncertainty propagation will depend on the model assumed, as well as the prior assumptions about the parameter and input uncertainties.

Examples of Forward Uncertainty Propagation include Error propagation equations:

- Monte Carlo propagation;
- Reliability methods; and
- Fuzzy and imprecise methods.

Model calibration and conditioning uncertainty on available data

There are two fundamentally different approaches to model calibration and conditioning given observational data. The first is based on treating model error as an additive term to the model prediction:

$$Y(x,t) = M(\theta, x, t) + \varepsilon(x, t) \qquad (1)$$

where Y is an observed value, the function M represents a model variable predicted using parameter set θ, ε is an error, x is space and t is time. Formal statistical assumptions about modelling errors are of this type and may involve additional parameters (such as bias and variance) in the error model $\varepsilon()$. Multiplicative errors can be treated in the same way by using log values of the observations and model predicted variables. The error model is normally evaluated using predictions based on some "optimal" parameter set. This is the basis of the regression and Bayesian methodologies. In some studies, the evaluation of uncertainties around an optimal model is undertaken after optimisation, rather than as an intrinsic part of the model calibration process.

The second approach rejects the concept of an optimal model in favour of the equifinality concept of allowing for multiple acceptable models. It is a rejectionist approach, in that only those models considered to give acceptable predictions in calibration will be retained for use in prediction. The errors associated with the predictions of a particular parameter set are treated implicitly. It is assumed that the structure of the errors found in calibration (in all their complexity) will be "similar" when that model is used in prediction. This is the basis of the Generalized Likelihood Uncertainty Estimation (GLUE) approach, in which informal model performance measures can be used to decide whether a model is retained (but can use formal error assumptions as a special case, treating the parameters of the error model as additional parameter dimensions). The distinction between this method and real-time data assimilation is fuzzy. Some methods from real-time data assimilation could be also under this heading.

Examples of this analysis include Nonlinear regression; Bayesian methods; and Generalized Likelihood Uncertainty Estimation (GLUE) Methods and Extended Generalized Likelihood Uncertainty Estimation (GLUE) (rejectionist) methods.

Real-Time Data Assimilation

This method has an important application in flood forecasting. The most important issue in real time flood forecasting is to allow for the fact that in any prediction event there will be departures between observations of water levels or discharges and predicted values. Any real time forecasting system should therefore allow for an updating or data assimilation strategy, with a view to minimising the uncertainties in the updated predictions. This will be possible in real applications as long as real time observations can be made available to the forecasting system, e.g. by telemetry. A number of data assimilation strategies are available, depending on the assumptions that the modeller is prepared to make.

Types of this analysis include the Kalman Filter; Extended Kalman Filter; Ensemble Kalman Filter; and Sequential Monte Carlo Methods.

Sensitivity Analysis

Sensitivity analysis involves identifying and investigating the sensitivity of the outcome or response variable to changes in input variables and parameters. The input variables/parameters are adjusted within what are thought to be plausible limits, and the impact on the response measured. Where a response is particularly sensitive to a variable/parameter, efforts can be directed to reducing the uncertainty on the 'key' variable/parameter. Uncertainty analysis that uses probability distributions to represent uncertainty and involves Monte-Carlo simulation techniques is a formal method of sensitivity analysis (Sayers *et al*., 2003).

Uncertainty analysis focuses on data, and how uncertainty in data propagates through computations. Sensitivity analysis sets out to establish the properties of computations (e.g. models) independently of particular data. The information generated by sensitivity analysis can be used to target uncertainty reduction activities, for example data collection or model improvement, where it will have maximum impact.

Many uncertainty analysis tools deal with parameter identifiability, ambiguity or uniqueness (e.g. Spear and Hornberger, 1980; Wagener *et al.*, 2003), and thus include some elements of sensitivity analysis. Plots showing model parameters vs. model results, such as the scatter plots of the GLUE methodology, provide a graphical indication of parameter sensitivity (Saltelli *et al.*, 2004).

Types of sensitivity analysis include: Local sensitivity analysis methods; and Global sensitivity analysis methods such as variance-based sensitivity analysis.

References

Cunge, J.A., 2003, Of data and models. Journal of Hydroinformatics 5(2):75-98.

Hall, J.W. and Anderson, M.G., 2002, Handling uncertainty in extreme or unrepeatable hydrological processes - the need for an alternative paradigm. Hydrological Processes, 16(9), 1867-1870.

Hall, J.W., Tarantola, S., Bates, P.D. and Horritt, M.S., 2005, Distributed sensitivity analysis of flood inundation model calibration. Journal Hydraulic Engineering-ASCE, 131(2), 117-126.

Oakley, J.E. and O'Hagan, A., 2004, Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of The Royal Statistical Society Series B-Statistical Methodology, 66, 751-769.

Pappenberger, F., Iorgulescu, I. and Beven, K.J. (2006). Sensitivity Analysis based on Regional Splits (SARS - RT). Environmental Modelling and Software, 21 (7), 976-990.

Ratto, M., Tarantola, S. and Saltelli, A., 2001, Sensitivity analysis in model calibration: GSA-GLUE approach. Computer Physics Communications, 136 (3), 212-224.

Saltelli, A., Tarantola, A., Campolongo, F. and Ratto, M., 2004, Sensitivity Analysis in Practice - A Guide to Assessing Scientific Models. John Wiley & Sons, Chichester, 232pp.

Sayers, P.B., Gouldby, B.P., Simm, J.D., Meadowcroft, I. and Hall, J., 2003, Risk, Performance and Uncertainty in Flood and Coastal Defence – A Review. R&D Technical Report FD2302/TR1, Defra, 141pp.

Spear, R.C. and Hornberger, G.M., 1980, Eutrophication in Peel Inlet, II, Identification of critical uncertainties via Generalised Sensitivity Analysis. Water Resource Research, 14, 43-49.

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S. and Gupta, H.V., 2003, Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. Hydrological Processes, 17 (2), 455-476.