CHAPTER -6

UNCERTAINITY OF MATHEMATICAL MODELING

It is possible to predict an incorrect future with great accuracy or a correct future with great uncertainty" (Beck 1987).

CHAPTER-6

Uncertainty of Mathematical modeling

6.1 Introduction:

The case is given for attention to the evaluation of uncertainty in water quality modelling, in the contexts of new demands for assessment of risk to water quality status, and typical paucity of supporting data. A framework for the modelling of water quality is outlined and presented as a potentially valuable component of broader risk assessment methodologies, and potentially useful methods of numerical uncertainty analysis are reviewed and demonstrated. A selective library of dynamic models and numerical tools for model solving and uncertainty analysis are compiled into novel software for model uncertainty analysis and risk-based decision-support. This software is applied to a series of case studies in an exploration of the underlying numerical problems and their relevance to modelling and management objectives using relatively sparse data sets. Issues examined in some detail are the importance of reconciling numerical solution tolerances with overall model precision; relative effects of numerical approximations, data and model structural biases on optimal design of field experiments and on prediction reliability; and the value and limitations of extending established methods of uncertainty analysis to decision-support. These investigations lead to discussions about priorities for the water quality modelling research community, in the face of contemporary and emerging numerical, technological and management problems. The main conclusion is that the current generation of modelling software can make very limited contribution to risk-based decision support, due to general absence of formal uncertainty analysis capabilities. This restriction is becoming more important due to new, ambitious spatial and ecological management challenges. Further research into numerical issues is needed to provide tools that allow these new challenges to be met, as well as to resolve persistent deficiencies in modelling capability. A more pressing concern, however, is that practitioners and their clients begin to confront issues of uncertainty, and create a demand for software that facilitates risk-based planning.

In other side, any measurement is subject to imperfections; some of these are due to random effects, such as short-term fluctuations in temperature, humidity and air-pressure or variability in the performance of the measurement device. Repeated measurements will

show variation because of these random effects. Other imperfections are due to the practical limits to which correction can be made for systematic effects, such as the offset of a measuring instrument, drift in its characteristics between calibrations or the personal bias in reading an analogue scale.

These imperfections in measurement data are usually expressed by the word 'uncertainty'. This term does not inspire confidence. However, when used in a technical sense as in 'measurement uncertainty' or 'uncertainty in the result of a laboratory analysis of a water quality sample' it carries a specific meaning (Taylor, 1997). It is a parameter, associated with the result of the measurement that defines the range of the values that could reasonably be attributed to the measured quantity. When uncertainty is evaluated and reported in a specified way it indicates the level of confidence that the value actually lies within the range defined by the uncertainty interval.

6.2. Model Uncertainty

Physico-chemical water quality measurements are frequently applied for river quality management purposes, e.g. for the assessment of current and historical state of surface waters, for water quality risk assessment, for calibration and validation of river quality simulation models, etc. The measurements are, however, subject to errors, which might be considerable in some cases or for some variables. The errors cause uncertainties in the assessments and in the model calibration and validation.

A level of uncertainty applies to all models and application of any model should include testing and sensitivity analysis. Sensitivity analysis shows how variation of a single factor affects model outputs. Uncertainty affects data collection and all stages of the modelling process and tends to increase with both the number of processes that feed onto the model along the DPSIR chain, and with complexity within the relevant model domain. In predictive models, uncertainty arises from inherent variability in natural processes, model uncertainty and parameter uncertainty. While the importance of uncertainty analysis is well recognised (Reckhow 1994) it is usually not included in pollutant transport models. This is a serious omission because if variability of input variables are large, so too will be

output predictability. (Beck 1987),provide excellent discussion of this issue. If *within* ecosystem variability is large, many samples need to be analysed to provide a given, defined, level of certainty in a mean value. Combined spatial, temporal and analytical uncertainty may be particularly high for measurements of some of the most important chemical ecosystem drivers, e.g. total phosphorus. This has profound implications for the reliability of use of simple models that predict ecosystem response from, e.g. nutrient loadings. Model uncertainty is clearly of importance in the conceptualization of the process for which predictions are required.

For example, a one-dimensional hydrological model would be expected to have greater predictive power than an ecological food-web systems model for lakes.
Investigations into, e.g. nutrient response models, suggest that prediction errors in both empirical and mechanistic models are unlikely to be under +30% and can be more than + 100% (Beck 1987; Reckhow 1994). However it is possible that the impact on modelling of individual error terms may be overestimated compared with combined effect of pairs of related parameters (Reckhow & Chapra, 1998). However, it is also clear that error estimation is often neglected when it should not be.

Increasingly, however, techniques such as Monte Carlo simulation are applied to predict frequency distribution of variables, especially in sparse data sets (Shanahan *et al.*, 1998). Further discussion of uncertainty and techniques to address this are given in (Cox and Baybutt 1981; Inman and Helton 1988; Chapra 1997).

6.2.1 Sources of Uncertainty and their Propagation:

A definition of uncertainty analysis is 'the means of calculating and representing the certainty with which the model results represent reality'. The difference between a deterministic model result and reality will arise from,

a. model parameter error,

b. model structure error (where the model structure is the set of numerical equations which define the uncalibrated model),

c. numerical errors - truncation errors, rounding errors and typographical mistakes in the numerical implementation,

d. boundary condition uncertainties.

As reality can only be approximated by field data, data error analysis is a fundamental part of the uncertainty analysis. Data errors arise from,

e. sampling errors (*i.e.* the data not representing the required spatial and temporal averages),

f. measurement errors (e.g. due to methods of handling and laboratory analysis),

g. human reliability.

Realising that an error-free model would equate to the error-free observations, the relationship between the actual model result M and the actual observations O can be summarized by,

$$M - \varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4 = O - \varepsilon_5 - \varepsilon_6 - \varepsilon_7 \text{ ----------(6.1)}$$

where ε_1 to ε_4 represent the model error arising from the four sources in the order listed above, and ε_5 to ε_7 represent the data error arising from the sources listed above.

Representing the overall error on either side of Equation 6.1 is not generally a simple task of adding the error variances together, as might be implied by the equation. This is because the errors may be unknown, and/or not of a random nature (see below), and/or the model output may be interdependent on the various sources of error in a manner that precludes their simple addition.

It is the goal of the modeller to achieve, to within an arbitrary tolerance, an error-free model by removal of ε_1 to ε_4. However, the modeller is generally neither in control of model structure errors ε_2, nor numerical errors ε_3, nor boundary condition errors ε_4.

Commonly, only the values of the model parameters are under the direct control of the

modeller. The aim would then become one of compensating as far as possible for ε_2 to ε_4 by identification of optimum effective parameter values. Central to this Thesis is the argument that there is always some ambiguity in the optimum effective parameter values caused by the unknown natures of, and inseparability of, ε_2 to ε_7 , and that this ambiguity can be represented by parametric uncertainty. As such, the model parameters are used as error-handling variables, and are identified according to their ability to mathematically explain ε_2 to ε_7. In most environmental modelling problems, significant bias in one or more of these errors will inevitably lead to biased parameter estimates. While the ideal solution would be to eliminate bias, for example by compensatory adjustments to data or by model structure refinement, such measures are often not practical and never comprehensive. In recognition of this, the potential importance of biased model calibration will be illustrated in this chapter, and significant attention is given to methods of uncertainty analysis which aim to deliver some robustness to bias.

The difficult task of identifying parameter uncertainty is generally approached using methods of calibration which derive, from the pre-calibration (*a priori*) parameter distributions, calibrated (*a posteriori*) distributions. In hydrological modelling, due to lack of prior knowledge, the *a priori* distributions are often taken as uniform and independent (e.g. Hornberger and Spear 1980). On the other hand, the *a posteriori* distributions, constrained by the data, may be multi-modal and non-linearly interdependent (Sorooshian and Gupta 1995). Inter-dependency arises when the model result is simultaneously significantly affected by two or more parameters, such that the distribution of each parameter must be regarded as conditional on the value of all interdependent parameters. Therefore, it is necessary to refer to the *joint* parameter distribution which is defined by a continuous function of *all* the parameters, and to sampled parameter *sets* rather than individual parameter values.

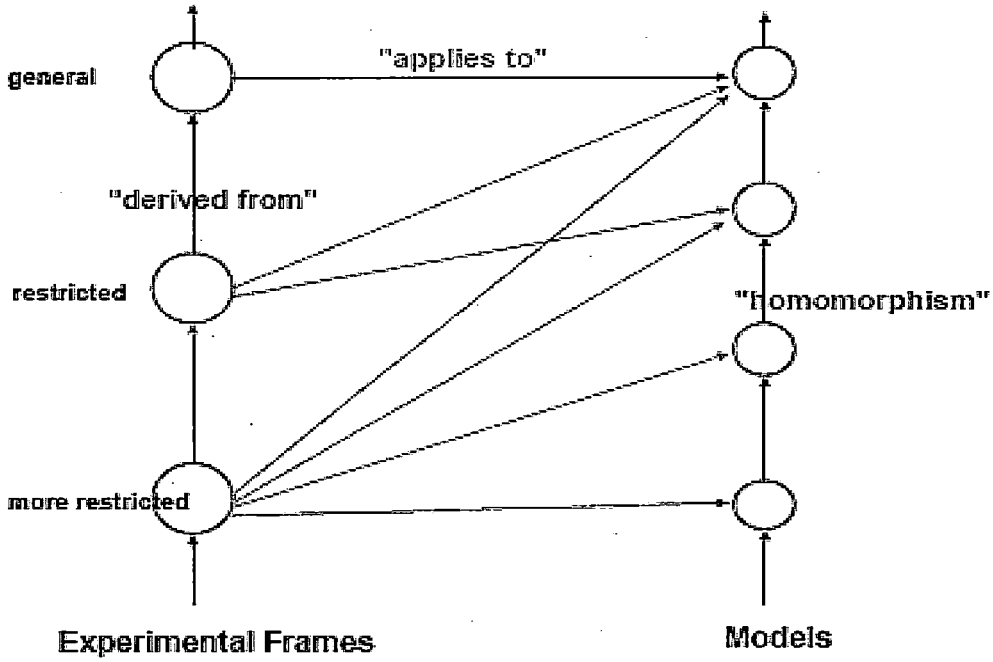

Fig 6.1: Experimental Frame – model relationship

6.2.2 Causes of uncertainty

Uncertainty in a water quality simulation model is inevitable due to the difficulty of identifying a single model (including grid-scale, process formulations and parameter values) which can accurately represent the water quality under all required model tasks (see the discussions of Beck 1987, Van Straten 1998, and Adams and Reckhow 2001). Although we have extensive knowledge about water quality processes from laboratory experiments, extrapolation of this knowledge to models of the real environment has consistently proven to be difficult. This is partly because the modelling scale is different to the laboratory scale, and the diversity of species and heterogeneity found in natural environments must (to some degree) be modelled approximately using lumped state variables. This means that formulations and parameter values identified at laboratory scale can only be used as a starting point for model design, rather than as a definitive end result. Nor is there yet any basis for regionalisation of water quality models. Therefore, models identified for one case study cannot be used with any confidence for another. Literature which describes established formulations and parameter values (Bowie et al. 1985, Thomann and Mueller 1987, Chapra 1997), is evidence of the wide range of models which are equally justified prior to observing a system's behaviour in detail, and that the uncertainty associated with modelling water quality on the basis of prior knowledge is extremely large. Given that it is

desirable to evaluate the performance of models with respect to observed water quality data, the accuracy, frequency and relevance of the available data dictates the attainable degree of certainty in the model. Unfortunately, water quality data can be expensive to collect and analyse, often requiring special handling and analysis in laboratories. This means that data to support model identification are generally sparse, often coming from sampling programmes which are fixed in frequency and location for regulation purposes, rather than designed to encapture the system's dynamic responses as required for successful model identification (Berthouex and Brown 1994). Also, water quality data are susceptible to noise and bias due to sampling, handling and measurement procedures (see Keith 1990). In addition, information about model boundary conditions such as sources of pollution, often suffers from the same short-comings, especially for distributed variables which are difficult to measure (pollution runoff, sediment quality, etc). In summary, lack of good quality data to support model identification is a major cause of model uncertainty. Closely related to the issue of data quality is model equifinality, whereby different models appear equally justified at the model design stage, but may give widely different realisations of the future. Equifinality is caused by interactions between model parameters, and by the near-equivalence of different model structures at the stage of model identification. This means that the same (or effectively the same within the context of the data errors) response can be achieved using different models. Clearly, the problem magnifies as both the number of interacting parameters increases, and as the precision of the data decreases. The use of parsimonious models, i.e. models which only include parameters which can be uniquely identified from the data, is one approach to avoiding equifinality. A parsimonious model implies that model components that are inactive during model identification are left out, and that strongly interacting components are combined into one (Young et al. 1996, Wagener et al. 2001). The inevitable omission of model components which are potentially relevant means that parsimonious models may seriously underestimate the uncertainty in model forecasts (Reichart and Omlin 1996). When the aim of the modelling is to investigate risks associated with proposed water quality interventions or other disturbances, it is essential that the uncertainty arising from previously unobserved behaviour is adequately allowed for, and so parsimonious models may be inappropriate. Thus, it may be said that identifying a single optimal model may not be a justifiable approach. The problem of equifinality and uncertainty in modelling environmental systems is inevitable and model predictions based on a single 'optimal' model will, in general, be rather arbitrary, and of very limited value. For this reason, a number of investigators have devoted their attention to rationalising the modelling problem, and redefining it as essentially stochastic whereby a population of feasible models (and by implication, a population of model predictions) are identified (Beven and Binley 1992).

6.2.3 Analysis of uncertainty

Water quality measurement uncertainty can be quantified either by evaluation of the results of several repeated measurements or by estimation based on data from records, previous measurements, knowledge of the equipment and experience with the measurement, or through comparison with a reference instrument.

The method, repeated measurements, has both advantages and disadvantages. Disadvantages are its limitations when using destructive tests. In the latter case, there must be the opportunity to repeat the test to another sample, often at significant extra cost and with the additional uncertainty due to sample variation, and at high economic cost. Advantages are that the method allows statistical analysis of the repeated measurement results. When these results are distributed around the average in the familiar bell-shaped curve, a normal distribution can be calibrated and the measurement uncertainty can be represented by the mean and standard deviation of this distribution. The mean reflects the bias (the systematic deviation) and the standard deviation the level of random error. Other measures of random error are the mean absolute error and the mean relative error. They may strongly depend on the magnitude of the water quality values, most often in a proportional way. In the latter case, the relative errors are independent on the water quality values, and the measurement uncertainty can be represented by the mean relative error as a single measure.

Identification of a population of feasible models can include both identification of alternative model structures (grid-scales and process formulations) and corresponding parameter distributions. Model structures should be of a complexity consistent with the difficulty and scale of the modelling task, and the supporting information and resources. They should be consistent with prior knowledge of how best to represent system processes at the scale and complexity in question. Given adequate supporting data, they can be assessed and amended using various identification techniques (e.g. Beck 1983). If one structure can be demonstrated as the most suitable for a particular modelling task (that is, for the particular system, and the particular information which the modeller aims to retrieve) then it would be reasonable to use this structure exclusively. On the other hand, if there are justified alternatives then ideally, from an analytical

point of view, the implications of these also should be considered (van der Perk 1997). This raises two issues. Firstly, it may be that no structures can be identified as 'suitable'. Then either an improved structure should be developed, or the stringency of the model assessment should be reviewed and the parameter uncertainty increased. Secondly, analysis of more than one structure may not be feasible given the available resources - such analysis will be costly, perhaps requiring purchase of additional software. Even using tools which offer some flexibility in the choice of water quality model structure, such as DESERT (Ivanov et al. 1996) or RWQM1 (Vanrolleghem et al. 2001) exploring candidate structures can significantly add to the burden on human and computer resources. In such a case (and this tends to *be* the case) all the significant model uncertainty must be represented, as far as possible, as parameter uncertainty within a single suitable structure. From a mathematical point of view, this has implications for the reliability of predictions (Draper 1995), but in a management context it is justifiable if it has relatively little bearing on the decisions being supported. In summary, investigating the sensitivity of decisions to different model structures is commendable, but may be neither viable due to resource constraints, nor worthwhile due to over-riding uncertainty in boundary conditions and parameter values.

Given a model structure, the identification of feasible sets of parameter values can be approached by conditioning (constraining) the prior population of parameter sets so that a specified modeling objective is better achieved. The modeling objective at this stage is generally to simulate observed data, and is generally expressed objectively as a function of the model residuals (the distances between the model result and the observed data). In traditional deterministic modeling, the response of this objective function (OF) to changes in the model parameters is used to estimate an optimum set of model parameters. This is achieved by manual perturbations of the parameters or, more suitably for complex models, by automatic algorithms. For uncertainty analysis, a joint distribution of parameters is identified rather than a single optimum, by recording the response of the OF across the parameter space. Depending partly on the algorithm which has been used, this joint distribution may be represented as a variance-covariance matrix, or as a discrete distribution (point estimates of probability mass over the parameter space), or as a population of feasible parameter sets.. Selecting an objective function to use for the conditioning of an environmental model is a difficult issue which involves a degree of speculation and subjectivity. This is because statistically-based identification of the parameter uncertainty requires knowledge of the combined error structure of the model, the data and the boundary conditions. However, especially when data are sparse or unreliable and the model structure is complex, there is little or no theoretical basis for estimation of the error structure . While parameter conditioning is often based on statistical likelihood functions (e.g. van Straten 1983), the result is dependent on the simplifying assumptions made about the error structure. As well as

being difficult to justify from prior information, such assumptions can lead to significant mis-representation of model uncertainty (Beven et al. 2001), in which case the model will fail to adequately explain the real system. In particular, the common assumption that the model and/or data are unbiased can lead to a serious underestimation of parameter and prediction uncertainties . As an alternative to statistical measures, the conditioning of the model can be based on subjectively derived rules, for example, "if the parameter set returns a model result that is highly consistent with my belief of true system behaviour then I will associate a relatively high weighting", or some objective expression of this, for example, "the relative probability of each parameter set will be equal to the proportion of the variance of the observed data explained by the model". Given that it is subjectively based, such an approach allows some freedom in achieving a satisfactory description of uncertainty, without the encumbrance of statistical rules and the long list of associated simplifying assumptions. Such conditioning of an environmental model, with the OF transformed to a probability without necessarily being related objectively to the error structure, was promoted by Beven and Binley (1992) in the context of their Generalised Likelihood Uncertainty Estimation. Once the uncertainty in the model is estimated, it can be propagated to give predictions. Methods of uncertainty propagation which are relevant to simulation modelling can be classified as variance propagation methods, point estimate methods, and Monte Carlo methods. Tung (1996) gives an overview of these methods, and a review and demonstration is included in the next chapter. The choice of method partly depends on the description of the parameter uncertainty, and partly on the computational resources, with the Monte Carlo methods generally (but not always) being more reliable and computationally demanding.

6.3 Surface water quality modelling applications

There is a variety of literature promoting understanding and application of uncertainty analysis in surface water quality modelling (e.g. Beck 1983, Beck 1987, Reckhow 1994, Adams and Reckhow 2001). However, the application of uncertainty analysis to surface water quality modelling seems to be relatively scarce, especially in practical decision support.

In the most widely used river water quality models, formal investigation of model uncertainty is very rare. Uncertainty identification in many contemporary models such as WASP5 (Ambrose et al. 1993), MIKE11 (Havnø et al. 1995), CE-QUAL-W2 (Cole and Wells 2000) and ISIS (Wallingford Software 2002) is difficult because they are relatively complex, and often linked to computationally intensive hydrodynamic, among other, modules. Although these models are well

founded in theory and well established in practice (see Ambrose et al. 1996), their usefulness is arguably limited by their high demand on resources, and the unknown uncertainty in their predictions. The large number of decision-support applications of these models which do not include analysis of uncertainty (amongst many others, Gunduz et al. 1998 and Warwick et al. 1999) is evidence of this. It is reasonable to assume that unpublished commercial applications of such models also under-represent the significance of uncertainty. The popular modelling tool QUAL2E-UNCAS (Brown and Barnwell 1987), which is a river modelling component of the US EPA's BASINS tool, has a built-in uncertainty analysis option. Reckhow (1994) recognises QUAL2E-UNCAS as an especially useful development, not only because it allows formal uncertainty analysis, but the associated documentation promotes uncertainty analysis amongst a large body of decision-makers. QUAL2E-UNCAS relies on estimation of prediction uncertainty through specification of feasible parameter and boundary condition ranges, and does not include a tool for conditioning the input uncertainties on observed data. Nor does the model allow covariance of inputs to be considered, meaning that uncertainty may be significantly over or under-estimated (Reckhow 1994, Brown 2002). Further to his commentary on QUAL2E-UNCAS, Reckhow (1994) notes that regulators in the USA tend to favour relatively simple water quality models, as complex models are too demanding on human resources, in addition to their high data demands. The UK Environment Agency have developed the relatively simple steady-state SIMCAT model to support regulation of river water quality (UK Environment Agency 2001a). SIMCAT is based on the recognition that model prediction uncertainties stem mainly from limitations in the calibration and pollution load data, rather than from the assumptions implicit to the model equations. SIMCAT was arguably a major step forward in the practice of river water quality modelling, in that parameter uncertainty can be identified from data sampling error by optimising the model parameters against different realisations of the data. As the model formulations used in SIMCAT are simple and easily solved, it is practical to use the computationally intensive sampling method. At the same time, the simplicity of the model structure makes the model less suitable for some tasks, such as extrapolation to changed boundary conditions, or simulation of dynamic events, when the effects of model structural error are more likely to be significant. The decision-support role of relatively simple models coupled with uncertainty analysis is evident from the continuing practices of both the UK and US environmental regulators. This contrast with the popularity of complex, resource-intensive models such as WASP5, MIKE11 and CE-QUAL. Accepting that both modelling approaches may have a role depending on the degree of detail sought and the resources available, there is arguably a benefit in providing tools that include a hierarchy of models. Supplementing this with uncertainty analysis facilities allows the limitations of both approaches to be evaluated for

specific modelling tasks. DESERT (Ivanov et al. 1996, also see Somlyody 1997) is a tool for catchments management optimization which provides a framework in which the user can design his own one-dimensional river water quality model. DESERT allows parameter conditioning, although the effect of parameter interactions cannot be included in application of the conditioned model. Based on dynamic programming, DESERT identifies all the sets of model inputs which conform to a series of constraints, which can include cost constraints for pollution control interventions, as well as in-river water quality criteria. In these respects, DESERT has the capacity for uncertainty analysis and flexibility of model design which will be needed for future water quality management problems, and is a valuable precedent for future developments.

6.4 Risks in Context

In the present context, risk may be usefully defined as "a combined measure of the degree of detriment to society or the aquatic ecosystem caused by a defined event (or combination of events), and the probability of that event occurring". Traditionally, in surface water quality management, the degree of detriment is simplified to a series of pass-fail criteria, each criterion representing a class of water quality (e.g. UK Environment Agency 1998). Risk can then be evaluated as probability of failure to achieve the target class. Modelling, then, has at least two potentially valuable roles – to extrapolate point measurements of water quality so that spatial and temporal criteria can be used in water quality classification rather than discrete, localised measurements of concentration; and to predict the response of risk to changing controls, to allow objective risk management.

This brief introduction to the role of modelling in risk-based water quality management raises a few issues. Firstly, it is important to differentiate between the frequency of failure that will actually occur due to system variability, and the modelled probability of failure, which includes (or *should* include) the influence of the uncertainty in the model and in the estimates of future boundary conditions. That is, there is a risk that any water quality intervention will fail to achieve its objectives due to the limitations of the modelling employed at planning stage. Consequently, where a modelling study implies a management option to be high-risk, this may be mainly due to the limited information and resources available for model and boundary condition identification, and a clear management priority would be to invest in more research. Also, there may be considerable risk associated with ill-defined objectives – that is, a water quality intervention may fail to be successful because at the time of planning the objectives were under-researched or

impossible to clearly define. For example, while it is reasonable to suggest that there will be lengthy debate over local and regional definitions of 'good ecological status' (Definition 22 in CEC 2000), the planning required to achieve such questionable status is already underway (e.g. UK Environment Agency 2001b). Finally on the point of associating risk, it is useful to distinguish between the risk stemming from anthropogenic system variabilities (for example diurnal variations in effluents) which are generally manageable, and risk stemming from 'natural' system variabilities (for example those due to meteorological influences) which are less manageable. In particular, if the risk of failure is predominantly due to unmanageable natural processes then reviewing the targets would be a logical way forward. With the capability of exploring reasons for risk, modelling has an essential role in, not only appraising pollution intervention options, but identifying sensible precursors to intervention.

6.5 A framework outline

Figure 6.2 outlines a general framework for risk-based modelling of water quality that will be further justified, developed, demonstrated and reviewed in the course of this chapter. Using such a framework it is intended that water quality managers have access to risk-based evaluation of surface water quality, and be able to respond to and develop this evaluation by,

· Identification of the principal factors affecting risk to water quality status.

· Evaluation of risk associated with alternative pollution control strategies, potentially with integration of external criteria, such as social and economic costs of water quality improvements.

· Consideration of alternative modeling criteria, in terms of identifying feasible water quality targets, and identifying acceptable compromises between no commensurate criteria (e.g. between water quality status and need for water abstractions).

· Consideration of different models for forecasting water quality response to pollution interventions (to reduce and evaluate risk associated with model structure uncertainty).

· Establishing priorities for collecting more data with which to improve model identification (reducing risk associated with data uncertainty),

Using modelling in this manner is consistent with more general risk assessment guidelines and frameworks used by environmental regulators. For example, UK environmental regulators (DETR et al. 2000) encourage proactive risk management using a tiered framework of quantitative risk assessments, whereby models, monitoring, and management options are reviewed as the analysis moves from risk screening to the advanced stages. This includes analysis of how the different sources of uncertainty contribute to the final risk estimate, and review of

costs and benefits. Such a tiered approach to risk assessment has been recommended for implementing the requirements of the Water Framework Directive (UK Environment Agency 2002). In applying this general risk assessment framework to management of water quality and aquatic ecology, there is clearly scope for iterative, model-based risk analyses, such as that promoted by Figure 6.2.

6.6 Technical considerations

In pursuit of a practical modelling tool that provides such a capacity for risk evaluation, the following tool features are considered essential;

1. Accessibility (ease of use), flexibility and extensibility (to cover a range of modelling tasks).
2. Efficiency of numerical techniques (to achieve the maximum benefit from Monte Carlo simulation).
3. Sensitivity analysis and risk evaluation capabilities.

Although the former three stipulations are common goals in the design and development of modelling tools in general, there are important implications in the water quality modelling context which deserve further discussion.

6.6.1 Need for accessibility, flexibility and extensibility

Accessibility of results is an important issue, as major management decisions usually must be supported using visually insightful reports, hence the benefit of an adequate interface for the graphical reporting of results. The value of advanced modeling techniques, for example Monte Carlo simulation, should not be diminished by perceptions that they are not transparent to decision-makers and stakeholders; effective interfaces may go a far way to avoid or resolve this concern. Furthermore, investigation of a variety of potential sources of risk, possibly including a large number of pollution sources and other system characteristics, requires careful attention to the thoroughness of model input specification. This draws attention to the value of an effective interface for model specification and data input.

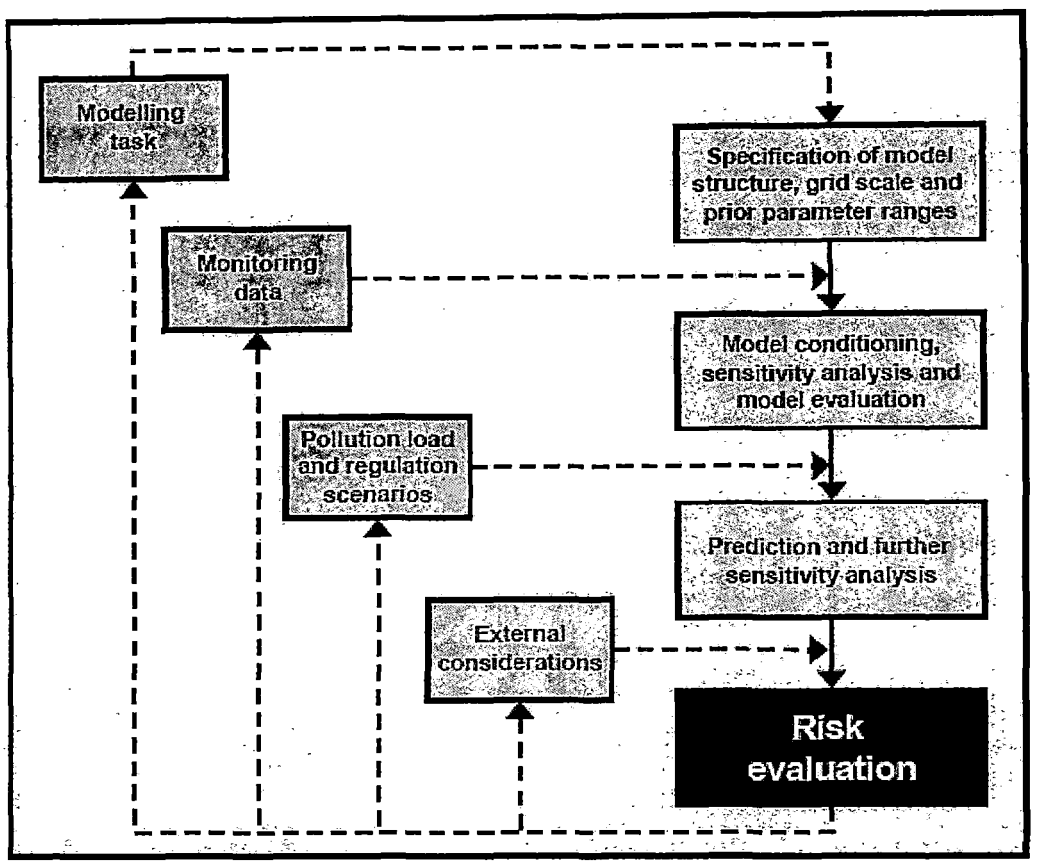

Figure 6.2 A framework for risk-based modelling of water quality

The requirement for flexibility is applicable to a number of aspects of a risk-based water quality modelling tool. Firstly, unavoidable subjectivity in estimating model uncertainty means that some choice of estimator should be provided, which is illustrated in studies by Freer et al. (1996) and Beven (1997). Application of multi-objective optimisation and sensitivity analysis (e.g. Bastidas et al. 1999) also requires flexibility in specification of model performance criteria. Central to the modelling procedure illustrated in Figure 1.1 is the capacity to explore different model structures, depending on the modelling task, data and computational resources available. If the model uncertainty is to be adequately represented by parameter uncertainty, the modeller should have the opportunity to identify a model structure which best allows this. In particular, the modelling grid scale (the spatial and temporal resolution of model) must be selected according to the water quality problem. Uncertainty introduced by spatial and temporal aggregations should be explored. Extensibility is essential so that new model structures and water quality determinands can be

incorporated, and so that the tool can be linked to new databases and other conjunctive software. In particular, as the directives driving water quality modelling promote integrated catchment management, and as the challenge of diffuse pollution management gathers pace, the increased use of Geographical Information Systems (GIS) as interfaces and platforms for water quality models is inevitable, and this might be borne in mind at the development stage, whatever the immediate modelling applications.

6.6.2 Need for Numerical Efficiency

Monte Carlo simulation provides us with the capability to retrieve a large amount of information about the sensitivity of model results to model inputs, which is extremely advantageous given the current limitations in the practice of water quality modelling. Although computational costs continue to diminish, the value of a Monte Carlo simulation will always depend on how well the continuum of possible model inputs/outputs is represented by a finite number of realisations. This would be especially relevant, for example, in catchment-scale distributed GIS-based modelling, due to the large amount of computation involved as well as the large number of spatially distributed model inputs which may be included in the analysis. There is therefore a need to either maximise the number of realisations achievable at a given computational cost, for example by implementing efficient numerical solvers and specifying numerical tolerances that are consistent with the overall reliability of the analysis, or to reduce the number of realisations required for an adequate representation by using variance reduction techniques (Cochran 1977). For example, a variance reduction technique which has been found useful in water quality modelling applications is Latin hypercube sampling (LHS; MacKay et al. 1979). LHS is, in the current context, designed to thoroughly sample the univariate distribution of each model input while leaving the sampling of interactions to chance. While some water quality modellers (e.g. Melching and Bauwens 2001) have successfully employed LHS to enormously improve the efficiency of sensitivity analysis, Press et al. (1988) note "if there is an important interaction between the design parameters, then Latin hypercube sampling gives no particular advantage (over simple random sampling)". Notwithstanding the merits of efficient sampling and solution schemes, more fundamental precursors to successful Monte Carlo analysis are, 1) appropriate limitation of model complexity, and 2) minimisation of the number of inputs to be sampled. Again, this draws attention to the need to match the model complexity to the specific modelling task, and the need to provide tools that offer some flexibility in model structure choice.

6.6.3 Need for Sensitivity Analysis and Risk Evaluation Capabilities:

Monte Carlo-based approaches to sensitivity analysis such as those implemented by Hornberger and Spear (1980), Beven and Binley (1992) and Kuczera and Parent (1998) have found wide application in environmental modelling, including a limited number of applications to surface water quality modeling. Incorporation of these methods into water quality modelling tools is an essential part of implementing the framework outlined in Figure 1.1. Firstly, they allow evaluation of the suitability of a model, in terms of reviewing the ability of the model and the associated parameter uncertainty to explain observed data. Thereafter, uncertainty in model forecasts can be estimated (e.g. Van Straten and Keesman 1991), avoiding the need for unqualified 'best estimate' forecasts. Monte Carlo methods not only have the potential to produce summary statistics of model sensitivities (e.g. Spear and Hornberger 1980, Wade et al. 2001), but can be used to evaluate risk to water quality status due to individual pollution sources and system properties, and can be extended to incorporate uncertainties in water quality criteria .Such evaluation has clear potential for risk-based decision making, particularly under conditions where data for identification of model and boundary conditions are limited. It also has the potential to be extended to simulating ecological risks, including spatial and temporal exposure as well as probability of occurrence.

Emphasis has been put on the value of Monte Carlo simulation because it is a relatively straightforward way of analysing how water quality objective functions respond over all feasible combinations of model inputs. This can be supplemented by alternative, computationally less demanding techniques of sensitivity analysis and uncertainty propagation. Using first order sensitivity analysis, the effect on a model response of perturbing each input variable around a specified value, while keeping the values of all other inputs fixed, is calculated. This has the advantage of allowing for the response components to be associated with individual inputs in a simple manner (e.g. Melching and Bowens 2001). However, the interactions between inputs are not explored and nonlinear responses are not estimated, so there is very restricted scope for exploring response surfaces, and effects (for example on risk) of low-probability values of model inputs are likely to be misrepresented. Also, the result will generally be dependent on the value around which the input is perturbed, as well as on the fixed values of all the other inputs, which may be quite arbitrary given the problem of model equifinality.

6.7. Quality of models

Models are, by their very nature, abstractions of reality used to simulate, rather than mimic, natural systems. They are seldom, if ever, truly correct (van Waveren, Groot et al. 1999) and application of models for management is often considered as much an art as it is a science. This does not imply a lack of rigour, but rather a recognition of inherent uncertainties and the need for the modeller to make intelligent choices in the development, use and reporting of models. This is both the strength and weakness of the modelling process. On one hand a model can reduce highly complex processes to simple output but, on the other, the strength of a model is determined by the relevance, and often extent, of the input data. Modelling can provide a powerful tool for management, but can be fairly meaningless if there is an ill defined objective, poor conceptualization of the causative relationships and their uncertainty, or if insufficient attention is paid to essential technical aspects of the modelling process. Failure to address, or at least be acutely aware of, these issues restricts sensible interpretation of results.

Mathematical models fall into a number of generic types that assemble and use data in different ways. All models have a *domain*, which provides the boundaries within which they were designed. Operating outside the defined domain is ill-advised.

Model development, however, often comprises hierarchical building blocks employing an array of methods and scales (e.g. bench scale, small field, field studies) in order to provide a conceptual model of the system. A good conceptual model is essential for the successful application of modelling which, itself, can help develop the concepts. A major distinction often made between models is the division into the, so called, empirical models or the process models. Empirical models provide relationships between variables without taking into account the dynamics of the processes modelled. They are based only on statistical or judgemental summary of data (Reckhow and Chapra 1998) and generally predict the magnitude of a response variable to a change in a driving variable. Time variant processes are not identified separately, but may be incorporated through amalgamation into long time-steps, such as annual means, to provide estimates of steady-state. In contrast, process (also known as dynamic or deterministic) models are very much concerned with changes of variables over time, commonly denoted as e.g. dN/dt, and generally depend on mathematical descriptions of the processes involved.

They are based strictly on scientific theory with processes described typically by scientifically reasonable equations. For example, they may include in the model such processes as change in partitioning of a chemical between states (e.g. dissolved or adsorbed onto a particle) in response to ambient conditions or rate of transport through a system by, e.g. advection-dispersion. Spatial resolution of models is also important, with broad distinction between *lumped* and *distributed* models which describes the extent to which models addresses areas, such as land-use categories, as homogenous units as opposed to treating, and modelling, spatial units separately.

Broadly, model complexity and data requirements, increase with spatial and temporal resolution. Existing models often link a number of domains and may include a mix of modelling approaches. It (Frisk et al. 1988) reviewed pros and cons of empirical and dynamic models for river eutrophication studies. Dynamic models can provide better causal relationships, which can clearly guide management, but often it may not be possible to collect the extensive data needed to develop or apply them.

The factors that guide the selection of an appropriate model or models relate to 1) applicability, 2) data requirements and 3) ease and cost of use, including necessary provisions for training. Complex mathematical models may not be transparent in their working and, therefore, difficult to explain to a non-technical audience. Models which require extensive and detailed data may not be feasible for operational purposes, unless they can reliably depend on one-off calibrations and validations. Simple models, on the other hand, may not be sufficiently intricate to be generally useful. For both complex and simple models, appropriate formulation of processes is important in order to provide a realistic conceptual model.

Applicability of models need to be guided by a realistic appreciation of what is required. Complex, data-hungry, models are not necessarily the best. Simple methods often suffice and are certainly appropriate where only rough or relative estimates are required and where overall estimates of annual net impact provides enough information to effect management (USEPA, 1999). Examples of simple methods would be nutrient export coefficient or simple regression models to predict nutrient loads. These can help particularly with catchment characterization under Article 5 of the (Water Framework Directive) WFD as part of the risk assessment process that water bodies may fail to meet

environmental objectives as outlined in Article 4 of the WFD. Both aspects can guide monitoring programmes required under Article 8 of the WFD.

Where a more detailed understanding is required, such as where a pressure may have important seasonal components but where there is, nevertheless, no capacity to apply very detailed modelling, the use of 'midrange' models (e.g. GWLF, AGNPS) might be employed (USEPA, 1999). In implementation of the WFD it could be envisaged that such an approach may be needed in some operational or investigative monitoring.

The most detailed models need only be applied to help management if it is clear that explicit analysis and understanding of underlying mechanisms are required. This could include a need to know high resolution temporal or spatial patterns of behaviour or impact of a pollutant. Under such circumstances, explicit process models (e.g. QUAL2E, HSPF, MODFLOW 3) can be powerful tools (USEPA, 1999). Mid-range and complex modelling are likely to be required mainly for investigative monitoring (Article 8) and for the implementation of Programmes of Measures (Article 11) where the cause of the failure to meet the Environmental Objectives are uncertain or contentious.

A handbook on Good Modelling Practice has been produced by a Dutch consortium (STOWA/RIZA, 1999; accessible free of charge from: http://www.info.wageningenur.nl/research%20projects/gmp.htm). The Handbook focuses on numerical and process orientated models, but contains principles of general applicability to all modelling. It provides a seven-step format and checklist for wise use of models. In addition to describing, in broad terms, the components required for each step, it also highlights essential considerations for the use of models and some common pitfalls.

General guidance relevant to application of models to support the implementation of the Water Framework Directive include:

- Define the objective that specifies the domain of the problem and the scenarios to be addressed;
- Determine if a mathematical model is needed to reach the objective;
- What reliability of model solution is required and does the expertise exist to apply the model;
- If a model is thought to be needed, provide a conceptual framework;

- Determine scope and boundary conditions, which guide data needs;
- Select a type of model;
- Verify that the conceptual model is effectively addressed by the computer programme chosen;
- Check the suitability or robustness of the model to extreme values of input data;
- Check the sensitivity of the model to changes in input values;
- Calibrate the model against empirical data sets;
- Validate the calibrated model with independent data; and
- Check if objective has been achieved (did the model answer the question that is was supposed to).

Failure to take account of any of these factors can result in a poor model output with limited descriptive or predictive value. For application to support the WFD, it may be tempting to use models which require data that, for whatever reasons, are not feasible to obtain or, conversely to use very simple models for which a very low degree of confidence can be placed in the results. Such models are not necessarily useless providing there is a realistic understanding of their strengths and weaknesses. But, inappropriate model use, or inexpert interpretation, can lead to results that lack credibility.

6.8 Conclusion:

The uncertainties in the water quality sampling data were estimated through comparison of the two or three sampling values for the same location and time moment (in few cases up to nine samples were taken for some specific variables). Both the absolute and the relative errors were calculated, and statistically analyzed. It was found that the errors are proportional. Mean relative errors therefore can be reported independent on the water quality concentration magnitudes.

The overall impression gained from the development of this chapter is that, at the moment, further research into improving river water quality models (and perhaps

environmental models in general) for decision-support is not warranted. Real improvements in the practical value of models lie in the willingness of modellers to seriously promote and confront the problem of uncertainty in research and tool development; and for decision makers to accept and use the outcomes, and integrate uncertainty in results into risk assessments. The complexity of commonly used models far surpasses the complexity of thought given to using the results properly. This view might be extended to say that further development of modeling methods (possibly including development of tools for uncertainty analysis) is not warranted by the current inclination of decision-makers to properly use them.

The reliability of the information produced by statistical analyses is only as reliable as the data from which it is generated. Greater commitment to consistency, frequency and accuracy is critical when accumulating the data. When dealing with frequency, it is important to note that observations that are accumulated too frequently, for instance weekly or bi weekly, tends to compromise data independence. Therefore, for purposes of temporal trend detection, monthly observations are considered sufficient (Cox *et al*, 2005, McBride & Loftis, 1994). It is also important to be consistent and accurate when collecting the data in order to curtail the limitations presented by missing, inaccurate and insufficient data.