Uncertainty and risk in water quality modelling and management

Neil R. McIntyre, Thorsten Wagener, Howard S. Wheater and Zeng Si Yu

ABSTRACT

The case is presented for increasing attention to the evaluation of uncertainty in water quality modelling practice, and for this evaluation to be extended to risk management applications. A framework for risk-based modelling of water quality is outlined and presented as a potentially valuable component of a broader risk assessment methodology. Technical considerations for the successful implementation of the modelling framework are discussed. The primary arguments presented are as follows. (1) For a large number of practical applications, deterministic use of complex water quality models is not supported by the available data and/or human resources, and is not warranted by the limited information contained in the results. Modelling tools should be flexible enough to be employed at levels of complexities which suit the modelling task, data and available resources. (2) Monte Carlo simulation has largely untapped potential for the evaluation of model performance, estimation of model uncertainty and identification of factors (including pollution sources, environmental influences and ill-defined objectives) contributing to the risk of failing water quality objectives. (3) For practical application of Monte Carlo methods, attention needs to be given to numerical efficiency, and for successful communication of results, effective interfaces are required. A risk-based modelling tool developed by the authors is introduced.

Key words | modelling, Monte Carlo, risk, water quality

Neil R. McIntyre (corresponding author)
Howard S. Wheater
Department of Civil and Environmental Engineering,
Imperial College of Science, Technology and Medicine,
London SW7 2BU,
UK
Tel: +44 207 594 6019
Fax: +44 207 823 9401
E-mails: *n.mcintyre@ic.ac.uk, h.wheater@ic.ac.uk*

Thorsten Wagener
SAHRA Hydrology and Water Resources,
Harshbarger Building,
University of Arizona,
Tucson, AZ 85721,
USA
Tel: +1 520 6268799;
Fax: +1 520 6211422;
E-mail: *thorsten@sahra.arizona.edu*

Zeng Si Yu
Department of Environmental Science and Engineering,
Tsinghua University, Beijing 100084,
People's Republic of China
Tel.: +86 106 2785610;
Fax: +86 106 2785687;
E-mail: *szeng@tsinghua.edu.cn*

INTRODUCTION

Motivation

In the European Community, the recently introduced Water Framework Directive (CEC 2000) requires that member states formulate River Basin Management Plans which identify objectives for achieving good water quality status on a catchment-wide basis. Similar standards apply throughout much of the world, for example catchment management in the United States has been guided by the Environmental Protection Agency's Water Quality Criteria and Standards Plan (USEPA 1998), in Australia by the National Water Quality Management Strategy (DAFF 2000) and in China by the Environmental Quality Standard for Surface Water (SEPA 1999). Simulation models are a central part of these basin management plans because they can apply the best available scientific knowledge, conditioned by historical evidence, to predict water quality responses to changing controls. For example, the development of the integrated catchment model BASINS is an explicit part of basin management plans in the United States (USEPA 1998, 1999).

The new high expectations for the aquatic environment, incorporated into the current wave of regulations, is prompting additional complexity with regard to modelling spatial variability, micro-pollutants and ecological indicators (Somlyody *et al.* 1998; Thomann 1998). Facilitated by improved computational resources, there is a trend for spatial discretisation to be increased, multi-media models to be developed (e.g. Havnø *et al.* 1995), and for traditional water quality determinands to be broken down into constituent species (Chapra 1999). As a consequence, the typical number of modelled components has risen exponentially over the past years, and this growth is expected to continue (Thomann 1998).

Despite the increasing expectations placed upon water quality models, contemporary deterministic models, when audited, frequently fail to predict the most local and basic biological indicators with a reasonable degree of precision (e.g. Jorgensen *et al.* 1986). Even when models are claimed to be 'reliable' following audits, a very significant margin of error is allowed (e.g. Hartigan *et al.* 1983). The application of modelling to the new era of high ecological standards presents severe challenges, especially given that our modelling experience is with relatively stressed ecological systems (Beck 1997; Shanahan *et al.* 1998), and that the economic implications of model errors may be relatively serious (Chapra 1997). While additional model complexity might be expected to improve the precision of model results, this has proven to be unfounded in a variety of studies (e.g. Gardner *et al.* 1980; Van der Perk 1997; also see Young *et al.* 1996). Furthermore, future driving forces such as climate (Parker 1993) and distributed pollution sources (Shepherd *et al.* 1999) are poorly defined and themselves cannot be modelled with much precision. Clearly, identification of suitable water quality policy must take account of the uncertainties associated with both the validity of the models and the driving forces. However, as increased model complexity hinders the formal evaluation of uncertainty, due to the large number of uncertain model components to be simultaneously analysed, there is a danger that our ability to evaluate uncertainty will decrease.

The challenges facing water quality modellers should be contemplated in the wider perspective of risk-based decision support. Firstly, a high degree of model uncertainty is not necessarily an undesirable outcome, and undoubtedly is preferable to no indication of reliability at all. Secondly, uncertainty in environmental models should be viewed as a source of risk, as is traditional in other fields of engineering, and should be used to establish and achieve an acceptable failure probability in terms of water quality status, rather than be used to decry the modelling approach (Beven 2000*a*). Given that risk is a concept that can be used to integrate external criteria, such as economics and safety, as well as integrating the model result over the relevant model responses, expressing results as risk is potentially attractive and seems inevitable. Thirdly, it is worth noting that, in the context of decision support, we are not justified in investing resources in modelling (including the identification of prediction uncertainty) unless this will be instrumental in the decisions that need to be made (Beven 1993). Therefore, we should keep sight of the modelling task and accept that (very) approximate solutions may be appropriate.

To allow intelligent use of complex simulation models, and to allow informed interpretation and application of model predictions, it is essential that a new generation of tools is developed and disseminated. These should be directed at evaluation of model uncertainty, as well as its minimisation, with respect to the modelling tasks. For results to be justified and interpreted properly, methods used for uncertainty analysis must be theoretically or intuitively well-founded and transparent to the modeller. For methods to be practical for day-to-day use, they should be relatively easy and fast to implement. These requirements are challenges which will be addressed in this paper, through review of the factors contributing to uncertainty, review of commonly employed methodologies used to address the problem and a proposed outline of a framework for risk-based water quality modelling. A tool for modelling of river and lake water quality where supporting resources are restricted, developed by the authors, is introduced. The proposed framework and associated tool are intended to be components of broader frameworks for risk assessment and management (e.g. DETR *et al.* 2000; Environment Agency 2002).

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 conditions.

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 the 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 (Thomann & Mueller 1987; Bowie *et al.* 1985; 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 & 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 shortcomings, especially for distributed variables which are difficult to measure (pollution runoff, sediment quality, etc.). In summary, lack of data to support model identification is a major cause of model uncertainty.

Closely related to the issue of data quality is model equifinality (Beven 1993), 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 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 (e.g., Auer *et al.* 1997), 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, 2002*c*). The inevitable omission of model components which are potentially relevant means that parsimonious models may seriously underestimate the uncertainty in model forecasts (Reichart & Omlin 1996; Beck 1999). 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.

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 (e.g. Hornberger & Spear 1980; Beven & Binley 1992; Reichart & Omlin 1996; Gupta *et al.* 1998).

Analysis of uncertainty

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; Qian 1997; Wagener *et al.* 2002*c*).

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 should also be considered (e.g. Gardner *et al.* 1980; Van der Perk 1997). This raises two issues. Firstly, it may be that no structures can be identified as 'suitable'. Then (as will be expanded upon later) 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 the 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 WaterRAT (McIntyre & Zeng 2002), 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 insignificant 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 modelling objective is better achieved. The modelling objective at this stage is generally to simulate observed data, and is expressed objectively as a function of the model residuals (the distances between the model result and the observed data). In traditional deterministic modelling, 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 estimators (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 misrepresentation 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 uncertainties (e.g. McIntyre *et al.* 2002*a*).

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 give it a high likelihood', or some objective expression of this, for example, 'the likelihood 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 likelihood without necessarily being related objectively to the error structure, was promoted by Beven & Binley (1992) in the context of their Generalised Likelihood Uncertainty Estimation, which is discussed in the next section.

Once the uncertainty in the model is estimated as outlined above, 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. McIntyre *et al.* (2002*a*) and Tung (1996) give an overview and examples of these methods. 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.

UNCERTAINTY ANALYSIS OF SURFACE WATER QUALITY MODELS: PREVIOUS RESEARCH

There is a variety of literature promoting understanding and application of uncertainty analysis in surface water quality modelling (e.g. Beck 1983, 1987; Reckhow 1994; Adams & Reckhow 2002; McIntyre *et al.* 2002*a*). An overview of the important developments in methodology and modelling tools is provided below.

Hornberger & Spear (1980) employed set theory to identify an array of feasible models of eutrophication in the Peel Inlet, Australia. The key step in this procedure is identifying a characteristic range of system responses from the observed data and subjective experience. Any model which gives a result falling within this range is considered to be a behavioural, feasible model, and all such models are considered equally likely. The particular value of this method is that it requires only simple and transparent interpretation of the data and no assumptions about the model structural error. Therefore, it is a particularly valuable approach to modelling systems for which observed data are sparse. The method allows the sensitivity of all uncertain parameters to be related to the relative distributions of behavioural and non-behavioural models, and hence it is commonly known as Regional Sensitivity Analysis (RSA). Van Straten & Keesman (1991) applied RSA to represent uncertainty in predictions of lake eutrophication, Whitehead & Hornberger (1984) to river algae simulation and Wade *et al.* (2001) apply the procedure to a river phosphorus assimilation model. DESERT (Ivanov *et al.* 1996) is a water quality modelling tool which employs RSA for estimation of parameter uncertainty, as discussed in more detail below.

Beven & Binley (1992) extended RSA in their Generalised Likelihood Uncertainty Estimation (GLUE) procedure, in which every feasible model is weighted with a likelihood. The likelihoods (for each model structure) are interpreted as point estimates of probability mass from the joint parameter distribution for that model. Then, subsequent conditioning on new data can be incorporated in a Bayesian manner. Similarly, alternative model structures with their own feasible parameter space can be combined using Bayes' method. A key feature of GLUE is that the modeller designs an objective function so that it can be used as a measure of likelihood. That is, the likelihood measure represents the modeller's degree of belief in the model validity given the available data, prior knowledge and modelling task, and the measure can be manipulated so that the data variance is satisfactorily explained by the identified parameter uncertainty (e.g. Freer *et al.* 1996). In this manner, the GLUE likelihood provides a framework whereby all sources of error, including model structural error, can be incorporated into parameter and prediction uncertainty. GLUE has been applied to hydrogeochemical modelling by Zak *et al.* (1997) and to solute transport modelling by Hankin & Beven (1998) and Camacho (2000). GLUE is available in generic uncertainty analysis tools of Beven (1998) and Wagener *et al.* (2002*b*), and further reviews are given by Melching (1995) and McIntyre *et al.* (2002*a*).

Another Monte Carlo-based technique for model conditioning is the use of Monte Carlo Markov Chains (MCMCs). Brooks (1998) gives a detailed description of this type of approach and Kuczera & Parent (1998) give a useful introduction and comparison with GLUE. MCMCs are related to GLUE in that a conditioned probability

distribution of parameters is identified using a likelihood measure of each of a series of sampled parameter sets. However, unlike the random sampling generally used for GLUE, an MCMC technique can favour samples from near the modes of the conditioned distribution, for example using the Metropolis algorithm (Metropolis *et al.* 1953). Because fewer parameter samples are taken from irrelevant regions of the parameter space, such an algorithm is expected to identify the modal regions more efficiently. Although parameter values with low posterior probability are sampled less frequently than by using GLUE, Kuczera & Parent (1998) found that the MCMC approach produced more reliable 90% confidence limits on model results. Other applications of a MCMC to hydrologic model conditioning include Mailhot *et al.* (1997) and Bates & Campbell (2001), although no evidence has been found of its application to conditioning of surface water quality models.

Gupta *et al.* (1998) advocate a multiple objective approach to parameter identification to increase the amount of information extracted from the available time series and provide insights into model structural error. Using their approach, different parameter sets may be optimised using different aspects of the model output, and may be associated with different components of the model. As the optimum parameter sets for each of the components are not mutually constrained, there is an added degree of freedom with which the model parameters can compensate for model structural error or data bias. The trade-off between alternative performance criteria can be defined by a surface of Pareto-optimal parameter sets (see Fonseca & Fleming 1995) and propagated to define a range of equally feasible model results (e.g. Yapo *et al.* 1998; Wagener *et al.* 2001). Another strength of multi-objective parameter identification is that sensitivities of a number of system components can be reported separately (e.g. Bastidas *et al.* 1999). A multi-objective approach to model conditioning is especially relevant in the context of water quality modelling, where it is common that multiple, inter-dependent determinands of water quality are simulated together. Wagener *et al.* (2002*a*) apply a novel form of multi-objective conditioning to a solute transport model and Meixner *et al.* (1999) apply a multi-objective sensitivity analysis to a hydro-geochemical model.

Often in water quality modelling, the prior parameter uncertainty is not conditioned by observations using one of the aforementioned methods, but is propagated to predictive results by Monte Carlo sampling of independent prior distributions of values, or through first-order approximation. Examples include Van der Perk *et al.* (1997), who apply Monte Carlo simulation to a steady-state river quality model, and Aalderink *et al.* (1996), who evaluate the effect of input uncertainties on a heavy metal model (interestingly, they conclude that distinguishing between the effects of different pollution control scenarios is impossible due to high uncertainty).

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) and CE-QUAL (Cole & Wells 2000), 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), Warwick *et al.* (1999) and Mishra *et al.* (2001)) 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 & 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 (Environment Agency 2001*a*). 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 in 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 contrasts 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 catchment management optimisation which provides a framework in which the user can design his own one-dimensional river water quality model. DESERT allows parameter conditioning using RSA, 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.

A FRAMEWORK FOR RISK-BASED SURFACE WATER QUALITY MODELLING

Following review of the driving forces behind water quality modelling, the inherent problems in this discipline and previously proposed directions for addressing these problems, an outline of a modelling framework is now proposed and some desirable facets of a potential modelling tool are discussed. Beforehand, it is worth reviewing the significance of the term 'risk' in the context of surface water quality modelling.

Risk 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. Environment Agency 1998). Risk can then be evaluated as the 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 the 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 regional definitions of 'good ecological status' (Definition 22 in CEC 2000), the planning required to achieve such a questionable status is already underway (e.g. Environment Agency 2001*b*). 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 were 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 also in identifying sensible precursors to intervention.

Given that we are entering an era of ecological objectives where degrees of achievement, as well as probabilities of failure, are extremely relevant to everyone (and everything) involved, there is arguably a need to develop the role of the 'degree of detriment' aspect of risk evaluation. This implies a need for dynamic modelling of ecological status, and the importance of continued research into quantifying ecological status and linking it to modelled frequency, concentration, assimilation and persistence of pollution (among other influences). Furthermore, there is going to be a need for significant trade-offs between protection of aquatic ecology and the resulting social and economic costs. The problem then becomes not only of identifying the risks of failing different criteria, but also of establishing acceptable compromises between the risks.

A framework outline

Figure 1 outlines a general framework for risk-based modelling of water quality. 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:

1. Identification of the principal factors affecting risk to water quality status.
2. Evaluation of risk associated with alternative pollution control strategies, including integration of external criteria, such as social and economic costs of water quality improvements.
3. Consideration of alternative modelling criteria in terms of identifying feasible water quality targets, and identifying acceptable compromises between non-commensurate criteria (e.g. between water quality status and need for water abstractions).
4. Consideration of different models for forecasting water quality response to pollution interventions (to reduce and evaluate the risk associated with model structure uncertainty).
5. 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

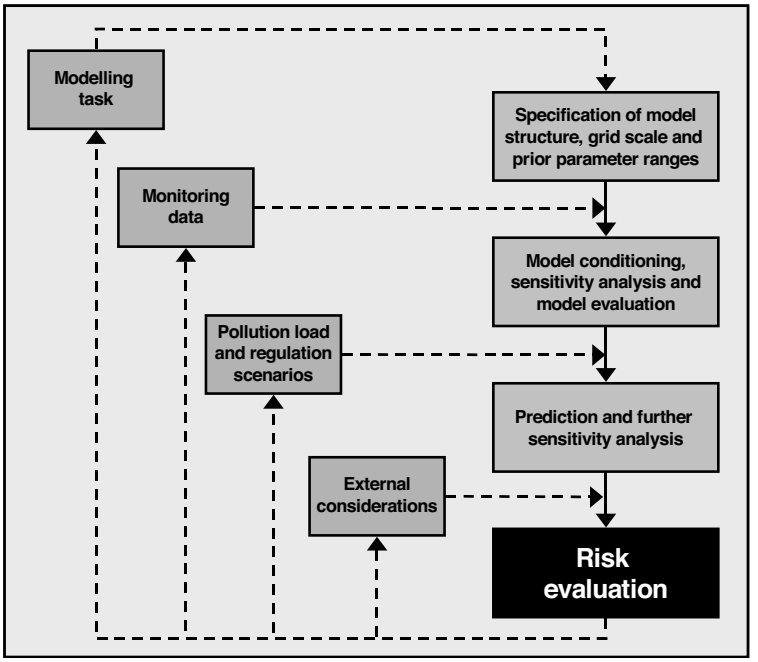

Figure 1 | A framework for risk-based modelling of water quality.

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 (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 1.

3.3 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.

The 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 obvious value of Monte Carlo-based approaches, such as RSA and MCMC, should not be diminished by perceptions that they are not transparent to decision-makers and stakeholders; effective interfaces may go a long 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 the model input specification. This draws attention to the value of an effective interface for model specification and data input.

With the general recognition that methods must suit individual problems, the requirement for flexibility is applicable to a number of aspects of a risk-based water quality modelling tool. Firstly, the unavoidable subjectivities in choosing measures of uncertainty, for example using GLUE likelihoods, means that some choice of OF must be provided which is clearly illustrated in studies by Freer *et al.* (1996) and Franks & Beven (1997). Application of multi-objective optimisation and sensitivity analysis (e.g. Bastidas *et al.* 1999) also requires flexibility in the specification of objectives. Central to the modelling procedure illustrated in Figure 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 the model) must be selected according to the water quality problem. Spatial aggregations are likely to introduce uncertainty that the modeller should be encouraged to explore, for example using grid-halving or advanced adaptive grid schemes (McIntyre *et al.* 2002*b*). 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, the increased use of Geographical Information Systems (GIS) as interfaces and platforms for water quality models is inevitable, and this should be borne in mind at the development stage, whatever the immediate modelling applications.

The 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 (as reviewed in Section 1 of this paper). 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 (McIntyre *et al.* 2002*b*), or to reduce the number of realisations required for an adequate representation by using variance reduction techniques. One variance reduction technique which has been found useful in water quality modelling applications is Latin hypercube sampling, LHS (e.g. Portielje *et al.* 2000). LHS is a highly factorised, stratified sampling scheme (MacKay *et al.* 1979) which, in the current context, is designed to thoroughly sample the prior marginal distribution of each model input while leaving the sampling of interactions to chance. While some water quality modellers (e.g. Melching & Bauwens 2001) have successfully employed LHS to enormously reduce the necessary number of random samples, 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.

The need for sensitivity analysis and risk evaluation capabilities

Monte Carlo-based approaches, such as RSA (Hornberger & Spear 1980), GLUE (Beven & Binley 1992) and MCMCs (see Kuczera & Parent 1998), have found wide application in environmental modelling, including a limited number of applications to surface water quality modelling, as cited

earlier. Incorporation of these methods into water quality modelling tools is an essential part of implementing the framework outlined in Figure 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 & 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 & Hornberger 1980; Wade *et al.* 2001) but can also 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 (e.g. McIntyre *et al.* 2002*c*). Such an evaluation has clear potential for risk-based decision-making, particularly under conditions where data for the 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.

The role of multi-objective optimisation has already been recognised in terms of identifying model structural error, and expressing it as parameter uncertainty, for example via the Pareto optimisation used by Gupta *et al.* (1998). The description of model uncertainty which may be obtained from such a multi-objective approach is different from that achieved using Bayesian-type approaches such as GLUE (Beven 2000*b*), and there is value in exploring the information given by the alternative definitions (e.g. McIntyre *et al.* 2001). Pareto optimisation also has a direct role in decision-making at a management level, for example to illustrate the trade-off between economic costs and water quality improvements (e.g. Hosoi *et al.* 1996).

Figure 2 shows an iterative approach to model structure and parameter uncertainty identification, within which any combination of the reviewed methods can be employed.

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 being simple, and allowing for the association of response components with individual inputs (e.g. Melching & Bauwens 2001). However, the interactions between inputs are not explored and non-linear 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.

To make the first-order analysis more robust to uncertainties, it can be extended to a factorial analysis (Henderson-Sellers & Henderson-Sellers 1993) which allows for two-factor interactions between model inputs. A two-factor factorial analysis may be regarded as the antithesis of Latin hypercube sampling—in the factorial analysis the prior marginal distributions of individual factors are represented by just two points so that two-factor interactions can be explored rigorously, while Latin hypercube sampling neglects exploration of interactions in favour of a much more thorough sampling of the marginal form. For model screening, two-factor factorial analysis may be preferred to Monte Carlo-based methods of analysis, as it can be used to concentrate on the sensitivities associated with extreme values (Kleijnen 1997) (although it still cannot identify non-linearities in the response surface). Due to this, and its attraction as a supplement to Monte Carlo methods, two-factor factorial analysis should be considered for inclusion in water quality modelling tools.

Another set of methods which are not designed to return extensive information about response surfaces, but which can be extremely useful for propagating input uncertainty to model forecasts, are point estimation methods. Of these, Rosenblueth's two-point estimation method (Rosenblueth 1981), and Harr's point estimation method (Harr 1989) have found most application in water resources (see Tung 1996). These methods are

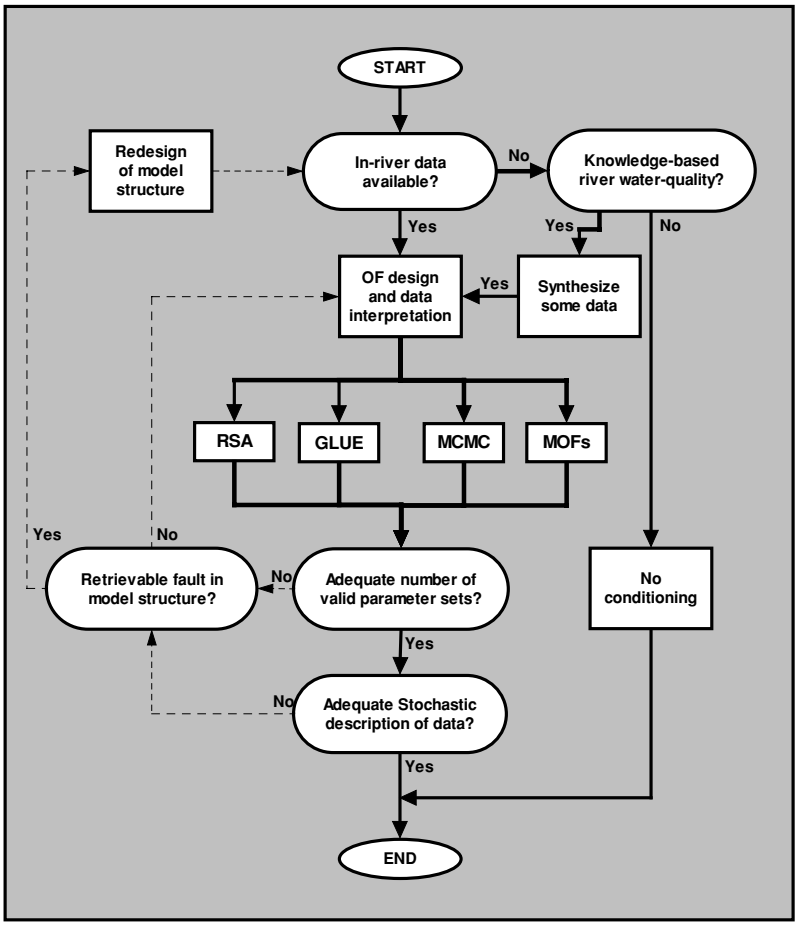

Figure 2 | An iterative approach to model identification. RSA=Regional Sensitivity Analysis (Hornberger & Spear 1980); GLUE=Generalised Likelihood Uncertainty Estimation (Beven & Binley 1992); MCMC=Monte Carlo Markov Chain using simulated annealing (Metropolis *et al*. 1953); MOFs=Pareto-based multi-objective optimisation (Gupta *et al*. 1998).

related to the GLUE methodology—in GLUE, a large number of randomly positioned point estimates of probability mass are assumed to represent the shape of the joint probability distribution of model inputs, and these are individually mapped onto point estimates from the joint distribution of model outputs; whereas in Rosenblueth's and Harr's methods it is hoped that a fewer number of carefully positioned and weighted point estimates will represent the lower moments and correlation structure of the input distribution and can be propagated through the model to represent the lower moments of the output distribution. Previous work (McIntyre *et al.* 2002*a*) suggests that these methods can provide useful estimates of the percentiles of model forecasts with substantially reduced computational cost.

A tool for risk-based management of water quality

As part of an investigation of the role of computational methods in the management of surface water quality in developing countries, where supporting data are unavoidably sparse, a modelling tool called WaterRAT (Water quality Risk Analysis Tool) has been developed (McIntyre & Zeng 2002). This tool is built around the methods and principles outlined above, and is designed to be employed in the manner illustrated by Figures 1 and 2. WaterRAT allows exploration of the uncertainties arising from all sources of prediction error—field data, model parameters, boundary and initial conditions, model structure, scale and numerical approximations. Model parameters, boundary and initial conditions can all be input as

distributions and can be conditioned to field data or other designed objectives using built-in algorithms (a genetic algorithm, Metropolis and GLUE). Four simultaneous objectives can be specified and Pareto-optimal trade-offs can be identified. Regional sensitivity analysis using Latin hypercube sampling is complemented by factorial sensitivity methods. WaterRAT allows the effects of output uncertainties to be evaluated in terms of the risk of failing water quality targets, and will plot risk of failure against any one input variable, supporting, for example, risk-based management of pollution control. Additionally, the water quality targets can themselves be assigned uncertainty, thus incorporating risk due to poorly defined objectives. Dynamic models are solved using an adaptive time-step procedure, with the temporal numerical tolerances pre-specified by the user.

The WaterRAT interface is a series of dialogue boxes and spreadsheets so that the need for the training of users is limited. WaterRAT contains a library of water quality simulation models, providing a hierarchy of models of varying complexity, and a capability for modelling organic pollution, phytoplankton, dissolved oxygen, various nutrients, a toxic substance, floating and suspended oil, and total suspended solids. This is supported by sediment models which include biochemical and physical sediment–water interactions. A thermodynamic model is available which models heat fluxes from the atmosphere and sediment, and includes a simulation of river ice cover. New models can easily be added to the library, allowing extension to new problems. The library of models is comprised of a set of Dynamic Link Libraries which minimise processing time and allow Monte Carlo simulation to be efficiently applied.

SUMMARY

Uncertainty is inherent to water quality model forecasts due to errors in model structures, boundary conditions, parameters and in the data used for model conditioning. Such uncertainty is not likely to diminish with advances in model complexity and computational power, due to the problem of model and boundary condition identification. Therefore, there will continue to be a degree of risk associated with using model predictions as a basis for water quality management decisions.

There have been significant developments in modelling uncertainty and risk in environmental systems in the last twenty years. In particular, modern computing resources allow Monte Carlo methods of uncertainty estimation and sensitivity analysis to be used more routinely to establish directions for environmental management and research. Despite the need, such methods have not been applied on a wide scale to water quality modelling. Instead, commercial model development has focused on the refinement of process representation, on the premise that the user can provide adequate supporting field data. Given that this premise is unrealistic in the majority of cases, this paper has argued that water quality modelling should be approached by the application of appropriate methods of uncertainty evaluation, and the development and dissemination of modelling tools that provide alternatives to more complex, resource-demanding models.

A computational framework, within which model uncertainty and associated risk can be extensively evaluated, has been outlined. Towards the implementation of such a framework, the need for a tool with a number of features and analytical methods has been argued. Monte Carlo-based approaches to regional sensitivity analysis and risk forecasting are fundamentally important. Notwithstanding the power of Monte Carlo simulation, alternative, less computationally onerous approaches, including factorial experiments and point estimation methods, have their own value for different modelling tasks. To allow uncertainty analysis to be employed for a range of practical applications, considerations in tool design have been summarised as accessibility, flexibility, extensibility and numerical efficiency. Such attributes allow the complex issue of model reliability to be addressed in modelling practice, and important aspects of risk to be communicated to decision-makers. A risk-based modelling tool, WaterRAT, which was developed to implement the outlined framework, has been introduced.

ACKNOWLEDGEMENTS

This research is undertaken as part of the EC 4th Framework programme project 'TOPLEM'. Particular thanks are due to Steven Chapra (Tufts University) and Andrew Billington (Northumbrian Water) for their contributions to this work.

REFERENCES

Aalderink, R. H., Zoeteman, A. & Jovin, R. 1996 Effect of input uncertainties upon scenario predictions for the river Vecht. *Wat. Sci. Technol.* **33**(2), 107–118.

Adams, B. & Reckhow, K. H. 2002 An examination of the scientific basis for mechanisms and parameters in water quality models. Unpublished. www2.ncsu.edu:8010/ncsu/CIL/WRRI.

Ambrose, R. B., Wool, T. A. & Martin, J. L. 1993 The water quality analysis simulation program WASP5 Version 5.10 Part A: Model Documentation. Environmental Research Laboratory, Office of Research and Development, US EPA, Athens, GA, USA.

Ambrose, R. B., Barnwell, T. O., McCutcheon, S. C. & William, J. R. 1996 Computer models for water quality analysis. In *Water Resources Handbook* (ed. L. W. Mays), pp 14.1–14.28. McGraw-Hill, New York.

Auer, M. T., Doerr, S. M., Effler, S. W. & Owens, E. M. 1997 A zero degree of freedom total phosphorus model: 1. Development for Onondaga Lake, New York. *Lake Reservoir Mngt.* **13**(2), 118–130.

Bastidas, L. A., Gupta, H. V., Sorooshian, S., Shuttleworth, W. J. & Yang, Z. L. 1999 Sensitivity analysis of a land surface scheme using multicriteria methods. *J. Geophys. Res.* **104**(D16), 19481–19490.

Bates, B. C. & Campbell, E. P. 2001 A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Wat. Res. Res.* **37**(4), 937–947.

Beck, M. B. 1983. Uncertainty, system identification, and the prediction of water quality. In *Uncertainty and Forecasting of Water Quality* (ed. M. B. Beck & G. Van Straten), pp 3–68. Springer-Verlag, Berlin.

Beck, M. B. 1987 Uncertainty in water quality models. *Wat. Res. Res.* **23**(8), 1393.

Beck, M. B. 1997 Applying systems analysis in managing the water environment: towards a new agenda. *Wat. Sci. Technol.* **36**(5), 1–17.

Beck, M. B. 1999 Coping with ever larger problems, models, and data bases. *Wat. Sci. Technol.* **39**(4), 1–11.

Berthouex, P. M. & Brown, L. C. 1994 *Statistics For Environmental Engineers*. CRC Press, Boca Raton, FL.

Beven, K. 1993 Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Wat. Res.* **16**(1), 41–51.

Beven, K. 1998 Generalised Likelihood Uncertainty Estimation (GLUE) User Manual. Lancaster University, Lancaster, UK.

Beven, K. 2000*a* On modelling uncertainty, risk and decision making. *Hydrol. Process.* **14**, 2605–2606.

Beven, K. 2000*b* *Rainfall-Runoff Modelling: The Primer*. Wiley, New York.

Beven, K., Freer, J., Hankin, B. & Schulz, K. 2001 The use of generalised likelihood measures in high order models of environmental systems. In *Non-linear and Non-stationary Signal Processing* (ed. W. S. Fitzgerald, R. C. Smith, A. T. Walden & P. C. Young), pp. 115–151. Cambridge University Press, Cambridge.

Beven, K. J. & Binley, A. M. 1992 The future of distributed models; model calibration and predictive uncertainty. *Hydrol. Process.* **6**, 279–298.

Bowie, G. L., Williams, B. M., Porcella, D. B., Campbell, C. L., Pagenkopf, J. R., Rupp, G. L., Johnson, K. M., Chen, P. W. H. & Gherini, S. A. 1985 Rates, constants and kinetics formulations in Surface Water Quality, 2nd edn. Report no. EPA/600/3-85/040, US EPA, Athens, GA, USA.

Brooks, S. 1998 Markov chain Monte Carlo method and its application. *J. R. Statist. Soc., Ser. D* **47**(1), 69–100.

Brown, L. C. 2002 Modeling uncertainty—QUAL2E-UNCAS A case study. Water Environment Federation. Unpublished. www.wef.org/pdffiles/TMDL/Brown.pdf.

Brown, L. C. & Barnwell, T. O. 1987 The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: documentation and user manual. Report no. EPA/600/3-87/007, US EPA, Athens, GA, USA.

Camacho, L. A. 2000 Development of a hierarchical modelling framework for solute transport under unsteady flow conditions in rivers. *PhD thesis*, Imperial College of Science, Technology and Medicine, London.

Chapra, S. C. 1997 *Surface Water-Quality Modeling*. McGraw-Hill, New York.

Chapra, S. C. 1999 Organic carbon and surface water quality modelling. *Prog. Environ. Sci.* **1**(1), 49–70.

CEC 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23rd October 2000 establishing a framework for Community action in the field of water policy. Council of the European Community, Brussels.

Cole, T. M. & Wells, S. A. 2000 CE-QUAL-W2: a two-dimensional, laterally-averaged hydrodynamic and water quality model, version 3.0. User manual. US Army Corps of Engineers. Instruction Report EL-00-1.

DAFF 2000 National Water Quality Management Strategy, 9, Rural Land Uses and Water Quality. Australian Department of Agriculture Fisheries and Forestry, Canberra, Australia.

DETR, Environment Agency, Institute for Environment and Health 2000 *Guidelines for Environmental Risk Assessment and Management*. HMSO, London.

Draper, D. 1995 Assessment and propagation of model uncertainty. *J. R. Statist. Soc., Ser. B* **57**(1), 45–97.

Environment Agency 1998 *The State of the Environment of England and Wales: Fresh Waters*. The Stationery Office, London.

Environment Agency 2001*a* *SIMCAT 7.6: A Guide and Reference for Users*. UK Environment Agency, Bristol, UK.

Environment Agency 2001*b* *Water Resources for the Future–A Summary of the Strategy for England and Wales*. UK Environment Agency, Bristol, UK.

Environment Agency 2002 *The Water Framework Directive: Guiding Principles on the Technical Requirements*, ch 7. UK Environment Agency, Bristol, UK.

Fonseca, C. M. & Fleming, P. J. 1995 An overview of evolutionary algorithms in multi-objective optimization. *Evol. Comput.* **3**(1), 1–16.

Franks, S. & Beven, K. J. 1997 Bayesian estimation of uncertainty in land-surface-atmosphere flux predictions, *J. Geophys. Res.* **102** D(20), 23991–23999.

Freer, J., Beven, K. & Ambroise, B. 1996 Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Wat. Res. Res.* **32**(7), 2161.

Gardner, R. H., O'Neill, R. V., Mankin, J. B. & Kumar, D. 1980 Comparative error analysis of six predator prey models. *Ecology* **6**(2), 323–332.

Gunduz, O., Soyupak, S. & Yurteri, C. 1998 Development of water quality management strategies for the proposed Isikli reservoir. *Wat. Sci. Technol.* **37**(2), 369–376.

Gupta, H. V., Sorooshian, S. & Yapo, P. O. 1998 Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Wat. Res. Res.* **34**, 751–763.

Hankin, B. & Beven, K. J. 1998 Modelling dispersion in complex open channel flows: 2. Fuzzy calibration. *Stochast. Hydrol. Hydraul.* **12**(6), 397–412.

Harr, M. E. 1989. Probabilistic estimates for multi-variate analyses. *Appl. Math. Model.* **13**(5), 313–318.

Hartigan, J. P., Friedman, J. A. & Southerland, E. 1983 Post-audit of lake model used for NPS management. *J. Environ. Engng. (ASCE)* **109**(6), 1354–1370

Havnø, K., Madsen, M. N. & Drge, J. 1995 MIKE 11—A Generalized River Modelling Package. In *Computer Models of Watershed Hydrology* (ed. V. P. Singh), pp. 733–782. Water Resources Publications.

Henderson-Sellers, B. & Henderson-Sellers, A. 1993 Factorial techniques for testing environmental model sensitivity. In *Modelling Changes in Environmental Systems* (ed. A. J. Jakeman, M. B. Beck & M. J. McAleer), pp. 59–76. Wiley, New York.

Hornberger, G. M. & Spear, R. C. 1980 Eutrophication in Peel Inlet, 1. Problem-defining behaviour and a mathematical model for the phosphorus scenario. *Wat. Res.* **14**, 29–42.

Hosoi, Y., Kido, Y., Nagira, H., Yoshida, H. & Bouda, Y. 1996 Analysis of water pollution and evaluation of purification measures in an urban river basin. *Wat. Sci. Technol.* **34**(12), 33–40.

Ivanov, P., Masliev, I., Kularathna, M., De Marchi, C. & Somlyódy, L. 1996 *DESERT User Manual*. International Institute for Applied Systems Analysis, Laxenburg, Austria/Institute for Water and Environmental Problems, Barnaul, Russia.

Jorgensen, S. E., Kamp-Nielsen, L., Christensen, T., Windolf-Nielsen, J. & Westergaard, B. 1986 Validation of a prognosis based upon a eutrophication model. *Ecol. Model.* **32**, 165–182.

Keith, L. H. 1990 Environmental sampling: a summary. *Environ. Sci. Technol.* **24**(5), 610–617.

Kleijnen, J. P. C. 1997 Sensitivity analysis and related analyses: a review of some statistical techniques. *J. Statist. Comput. Simul.* **57**(1–4), 111–142.

Kuczera, G. & Parent, E. 1998 Monte Carlo assessment of parameter uncertainty in conceptual catchment models. *J. Hydrol.* **211**(1), 69–85.

MacKay, M. D., Beckman, R. J. & Conover, W. J. 1979 A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**(2), 239–245.

Mailhot, A., Gaume, E. & Villeneuve, J. P. 1997 Uncertainty analysis of calibrated parameter values of an urban storm water quality model using Metropolis Monte Carlo algorithm. *Wat. Sci. Technol.* **36**(5), 141–148.

McIntyre, N., Lees, M. J. & Wheater, H. S. 2001 A review and demonstration of methods of uncertainty analysis in numerical environmental modelling. *Proceedings of the 8th Europia International Conference–Advances in Design Sciences and Technology, Delft, The Netherlands*, pp. 183–196. Europia Productions, The Netherlands.

McIntyre, N., Wheater, H. S. & Lees, M. J. 2002*a* Estimation and propagation of parametric uncertainty in environmental models. *J. Hydroinf.* **4**(3), 177–198.

McIntyre, N., Jackson, B., Wheater, H. S. & Chapra, S. 2002*b* Role of numerical efficiency in the viability of Monte Carlo simulation of the environment—a case study of river thermodynamics. *J. Environ. Engng. (ASCE)* to be published.

McIntyre, N. R., Wagener, T., Wheater, H. S. & Chapra, S. C. 2002c Risk-based modelling of surface water quality—a case study of the Charles River, Massachusetts. *J. Hydrol.* **274**, 225–247.

McIntyre, N. & Zeng, S. 2002 *WaterRAT User Manual (Version 1)*. A report to the European Commission. Department of Civil and Environmental Engineering, Imperial College, London, UK.

Meixner, T., Gupta, H. V., Bastidas, L. A. & Bales, R. C. 1999 Sensitivity analysis using mass flux and concentration. *Hydrol. Process.* **13**, 2233–2244.

Melching, C. 1995 Reliability estimation. In *Computer Models of Watershed Hydrology* (ed. V. P. Singh), pp. 69–118. Water Resource Publications, Colorado, USA.

Melching, C. S. & Bauwens, W. 2001 Uncertainty in coupled nonpoint source and stream water-quality models. *J. Wat. Res. Plann. Mngt. (ASCE)* **127**(6), 403–419.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, F. 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087.

Mishra, A., Anand, A., Singh, R. & Raghuwanshi, N. S. 2001 Hydraulic modeling of Kangsabati Main Canal for performance. *J. Irrig. Drain. Engng. (ASCE)* **127**(1), 27–34.

Parker, D. E. 1993 Climatic variation. In *Concise Encyclopaedia of Environmental Systems* (ed. P. C. Young), pp. 102–111. Pergamon, Oxford.

Portielje, R., Hvitved-Jacobsen, T. & Schaarup-Jensen, K. 2000 Risk analysis using stochastic reliability methods applied to two cases of deterministic water quality models. *Wat. Res.* **34**(1), 153–170.

Press, W. H., Teukolsky, S. A., Flannery, B. P. & Vetterling, W. T. 1988 *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge.

Qian, S. S. 1997 An illustration of model structure identification. *J. Am. Wat. Res. Assoc.* **33**(4), 811–824.

Reckhow, K. H. 1994 Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecol. Modell.* **72**(1–2), 1–20.

Reichart, P. & Omlin, M. 1996 On the usefulness of over-parameterised ecological models. *Ecol. Modell.* **95**, 289–299.

Rosenblueth, E. 1981 Two-point estimates in probabilities. *Appl. Math. Modell.* **5**, 329–335.

SEPA 1999 Environmental quality standard for surface water. Regulation GHZB1-1999. State Environmental Protection Administration, Beijing, People's Republic of China.

Shanahan, P., Henze, M., Koncsos, L., Rauch, W., Reichert, P., Somlyody, L. & Vanrolleghem, P. 1998 River water quality modelling: II. Problems of the art. *Wat. Sci. Technol.* **38**(11), 245–252.

Shepherd, B., Harper, D. & Millington, A. 1999 Modelling catchment-scale nutrient transport to watercourses in the UK. *Hydrobiologia* **396**, 227–237.

Somlyody, L. 1997 Use of optimization models in river basin water quality planning. *Wat. Sci. Technol.* **36**(5), 209–218.

Somlyody, L., Henze, M., Koncsos, L., Rauch, W., Reichert, P., Shanahan, P. & Vanrolleghem, P. 1998 River water quality modelling: III. Future of the art. *Wat. Sci. Technol.* **38**(11), 253–260.

Spear, R. C. & Hornberger, G. M. 1980 Eutrophication in peel inlet–2. Identification of critical uncertainties via generalized sensitivity analysis. *Wat. Res.* **14**(1), 43–49.

Thomann, R. V. 1998 The future ''golden age'' of predictive models for surface water quality and ecosystem management. *J. Environ. Engng. (ASCE)* **124**(2), 94–103.

Thomann, R. V. & Mueller, J. A. 1987 *Principles of Surface Water Quality Modeling and Control*. Addison-Wesley, New York.

Tung, Y. K. 1996 Uncertainty and reliability analysis. In *Water Resources Handbook* (ed. L. W. Mays), pp. 7.1–7.65. McGraw-Hill, New York.

USEPA 1998 Water quality criteria and standards plan–priorities for the future. USEPA 822-R-98-003, US EPA, Athens, GA, USA.

USEPA 1999 BASINS (Better Assessment Science Integrating Point and Nonpoint Sources) Version 2 User's Manual. Report no. EPA#: 823/B-98-006, US EPA, Athens, GA, USA.

Van der Perk, M. 1997 Effect of model structure on the accuracy and uncertainty of results from water quality models. *Hydrol. Process.* **11**(3), 227–239.

Van Straten, G. 1983 Maximum likelihood estimation for phytoplankton models. In *Uncertainty and Forecasting of Water Quality* (ed. M. B. Beck & G. Van Straten), pp. 157–171. Springer-Verlag, Berlin.

Van Straten, G. & Keesman, K. J. 1991 Uncertainty propagation and speculation on projective forecasts of environmental change: a lake eutrophication example. *J. Forecast.* **10**, 163–190.

Wade, A. J., Hornberger, G. M., Whitehead, P. G. & Flynn, N. 2001 On modeling the mechanisms that control in-stream phosphorus, macrophyte and epiphyte dynamics: an assessment of a new model using general sensitivity analysis. *Wat. Res. Res.* **37**(11), 2777–2792.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. & Sorooshian, S. 2001 A framework for the development and application of hydrological models. *Hydrol. Earth Syst. Sci.* **5**(1), 13–26.

Wagener, T., Camacho, L. A., Lees, M. J. & Wheater, H. S. 2002*a* Dynamic parameter identifiability analysis of a solute transport model. *J. Hydroinf.* **4**(3), 199–211.

Wagener, T., Lees, M. J. & Wheater, H. S. 2002*b* A framework for the development and application of parsimonious hydrological models. In *Mathematical Models of Large Watershed Hydrology*, vol. 1 (ed. Singh, & Frevert,), pp. 87–136. Water Resource Publications, Colorado, USA.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S. & Gupta, H. V. 2002c Reducing uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. *Hydrol. Process.* **17**(2), 455–476.

Warwick, J. J., Cockrum, D. & McKay, A. 1999 Modelling the impact of subsurface nutrient flux on water quality in the lower Truckee River, Nevada. *J. Am. Wat. Res. Assoc.* **35**(4), 837–851.

Whitehead, P. G. & Hornberger, G. M. 1984 Modelling algal behaviour in the River Thames. *Wat. Res.* **18**(8), 945–953.

Yapo, P. O., Gupta, H. V. & Sorooshian, S. 1998 Multi-objective global optimization for hydrological models. *J. Hydrol.* **204**(1–4), 83–97.

Young, P., Parkinson, S. & Lees, M. 1996 Simplicity out of complexity in environmental modelling: Occam's razor revisited. *J. Appl. Statist.* **23**(2), 165–210.

Zak, S. K., Beven, K. & Reynolds, B. 1997 Uncertainty in the estimation of critical loads: a practical methodology. *Wat. Air Soil Pollut.* **98**(3–4), 297–316.