Journal of Hydroinformatics | 03.3 | 2001

Water-quality modelling of the Upper Mersey river system using an object-oriented framework

Maria Spanou and Daoyi Chen

ABSTRACT

The aim of this paper is to present the recent advances in the development of an object-oriented software system for water-quality management, and discuss the results from its application to the study of the Upper Mersey river system in the United Kingdom. The software has been extended and includes tools for the construction of flow duration and low-flow frequency curves using different methods, the sensitivity analysis and parameter estimation of the water-quality model, and the stochastic simulation of the mass balance at the discharge points of point-source effluents. The application of object-orientation has facilitated the extension of the software, and supported the integration of different models in it. The results of the case study are in general agreement with published values. They also include low flow estimates at the ungauged river sites based on actual data for the artificial sources, and water-quality simulation results, which have not been presented earlier in the literature for the Upper Mersey system.

Key words | computer simulation, low flow analysis, object-oriented method, river system, Upper Mersey catchment, water-quality management

Maria Spanou
Daoyi Chen (corresponding author)
Civil Engineering Division,
University of Manchester,
Manchester M13 9PL,
UK
Tel: +44 161 275 4354;
Fax: +44 161 275 4361;
E-mail: *daoyi@water.eng.man.ac.uk*

INTRODUCTION

The application of hydroinformatics to the hydrological and water-quality modelling and management of river systems can provide efficient decision-support tools to the engineers and managers of water resources (Abbott 1999). Within the last decade object-oriented methods have been increasingly employed for the development of such tools, as they facilitate the integration of models, their modification and extension, and the interactive incorporation of changes related to the physical systems under study and their management schemes (Fedra & Jamieson 1996; Reitsma & Carron 1997).

A research effort to deliver the above benefits has been undertaken by the authors during the development of a comprehensive object-oriented framework for the performance of low-flow analysis, the simulation of water quality, and the control of point-source pollution in river basins (Spanou 2000). The resulting software, SMILE, provides a range of mathematical and statistical models for the above operations, includes a graphical user interface and a data management component, and communicates with the Microsoft Word™ and Excel™ systems for the tabular and graphical presentation of the results generated (Figure 1). The analysis and design of the system are based on the object-oriented paradigm, and have been implemented with the object-oriented programming environment Smalltalk Express™.

A significant part of the object model and the analytical tools that were included in the software at an earlier stage of development have been presented elsewhere (Spanou & Chen 1998, 2000). During the continuation of that work the software has been extended to include additional tools for low-flow studies, and to assist in the identification of the water-quality model and the stochastic calculation of the discharge consents of the point-source effluents.

The software has been validated though the study of the Upper Mersey river system in the UK. The Upper Mersey catchment is located in the North West of England

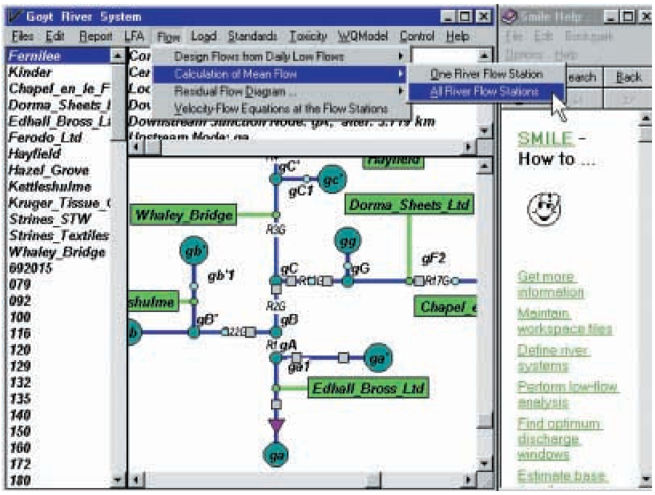

Figure 1 | Screenshot of the developed software tool.

that has a past of being one of the bases for the industrial revolution and has now suffered from the environmental pollution consequences. The river catchment includes areas of diverse land use such as urban development and agricultural land. The major sources of pollution are the continuous discharges of sewage treatment works (STW) and trade effluents, and the intermittent discharges from the combined sewer overflows (CSO) and the sewerage network. These discharges, the reservoirs, and the abstractions influence significantly the river flows. The improvement of water quality in the Upper Mersey river system is of current interest to the Environment Agency, the Mersey Basin Campaign, North West Water Ltd., and other pollution control authorities. Results for several aspects of the management study have also been presented in the literature during the last decade. These include (a) low-flow estimates at major river flow stations, (b) estimates of the natural mean and low flows at ungauged river sites, (c) the chemical grading of water quality for the year 1994, and (d) proposals for short-term and long-term *River Quality Objectives* (*RQO*s) which have been expressed in terms of the *River Ecosystem* (*RE*) classification scheme (Bullock *et al*. 1989; Gustard *et al*. 1992; National Rivers Authority North West Region 1996; Marsh 1997; Brown & Owens 1997). The *Mersey Measure* computer system was also developed for the assessment of water quality along the whole Mersey river (i.e., including the lower Mersey and the estuary), considering compliance with EC Directives and the calculation of water-quality indices (Department of the Environment 1996). North West Water Ltd. has further planned the improvement of several CSOs in the period 1995–2005, and the extension or rebuilding of a few STWs (Mersey Basin Campaign 1995).

The literature survey revealed that there were no estimates of the residual low flows at the ungauged river sites since the presentation of the residual flow diagrams for the four main rivers of the system from the North West Water Authority (1978). In addition, it suggests that there is a need for the simulation of river water quality and the detailed planning of point-source pollution control in the system (e.g., through the review of the corresponding discharge consents). The above parts of the management study, as well as those for which recent results are available, have been conducted using the developed software system.

For the performance of the study, data were provided by the Environment Agency North West Region, North West Water Ltd., and the National Water Archive Data Centre of the former Institute of Hydrology (IH), which is now the Centre for Ecology and Hydrology at Wallingford (CEH Wallingford) (Marsh 1997). Data were also retrieved from published reports, and were extracted from the *Water Information System* (*WIS*) of the IH (Tindall & Moore 1997). The collected data were subject to quality control, and were organised into an integrated data set. Indicative results of the case study that were obtained using the earlier version of the software have been presented in Spanou & Chen (1998, 2000). Improved results have been obtained thereafter for several aspects of the case study, using the added functionality of the software, as well as data that were recently provided and that were estimated in the past based on default values or other literature sources.

This paper summarises the mathematical analysis employed in the software and the object-oriented analysis that was followed for its development, describes in more detail the recently added methods and the extension of the object model, and presents thoroughly the software application to the case study.

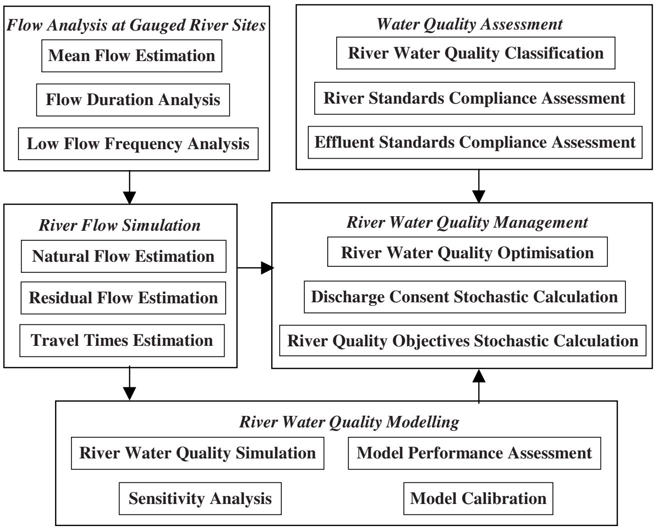

Figure 2 | Flowchart of the model components.

METHODS

The mathematical model employed consists of five components: (i) the flow analysis at gauged river sites, (ii) the estimation of flows at ungauged river sites, (iii) the modelling of river water quality, (iv) the assessment of water quality of the point-source effluents and receiving rivers, and (iv) the management of water quality in the river basin. Each component allows the performance of several tasks, as shown in the flowchart in Figure 2, and also integrates alternative methods for each task.

Flow analysis at gauged river sites

This component comprises methods for the calculation of the mean daily flow, the low-flow frequency analysis, and the flow duration analysis at a river flow station. The input to all methods is the record of daily flow data at the station.

The mean daily flow is calculated considering the whole period of record or a user-defined part of it.

Annual, seasonal or daily low-flow frequency curves of user-specified duration are derived applying parametric and non-parametric methods. The theoretical curves are prepared by fitting the three-parameter, asymptotic type III extreme-value probability distribution of smallest values $EX_{III,s}$ (ε,u,k) to the data (Spanou & Chen 2000). For the construction of the sample-based curves, the plotting positions p_i of the low-flow data are estimated using Eq. (1) with $c=0$ or $c=0.44$ (McMahon & Arenas 1982):

$$p_i = \frac{i-c}{n+1-2c} \tag{1}$$

where i is the rank number of a flow datum and n the total number of data.

Both the theoretical and sample-based curves are optionally plotted on linear graph paper. The latter are also plotted on probability paper of the Weibull distribution with parameters $\varepsilon=0$ and $k=4.0$; if the low flows follow the corresponding distribution, then these curves are approximately straight lines (Gustard *et al.* 1992; Johnson *et al.* 1994). The low flows with a specific return period (typically 2, 10 or 20 years) are finally estimated.

Annual or seasonal flow duration curves of user-specified duration are prepared applying non-parametric methods. They are derived through the estimation initially of the percentage frequency distribution, and subsequently of the percentage exceedence cumulative frequency distribution, using linear or logarithmic flow intervals within a user-defined range of flows. Alternatively, they are based on the estimation of the quantile values of the flow data. The resulting sample-based curves are plotted on linear or lognormal probability paper (Institute of Hydrology 1980). The flows exceeded for a specific percentage of time (e.g., 95%) are then estimated.

Flow estimation at ungauged river sites

The applied methods allow the estimation of mean or low flows at the ungauged river sites of natural and artificially influenced catchments, and the subsequent estimation of travel times in the river systems. The results are presented in detail in the output of the methods. Plots of the flow and travel time variation along a user-specified river are also provided.

More specifically, the river flows for natural catchments are derived applying a linear, drainage area–flow relationship. The proportionality constant (called the catchment contribution coefficient) is calculated from the data of a single downstream flow station or from a regional analysis among all stations; this value is then accepted or edited by the user (Spanou & Chen 2000).

For artificially influenced catchments, the river flows are estimated through the construction of *residual flow diagrams*. Their natural component is calculated as above, considering however the naturalised flows at the river flow stations. Their artificial component is derived using the licensed or actual flows of the surface-water abstractions, the STW and trade effluents, and the compensation reservoirs (Spanou & Chen 2000).

It is noted that the software system allows the user to select one or more data types to be used in the estimation of flows for each artificial source. The flows of the STW effluents, for example, can be set based on the design DWF, the discharge-consent flowrates, the actual DWF, and/or the actual flows to full treatment. When there are no data of a selected type for the time period under study, the user is informed about the methods that can be applied for their estimation, and about the data availability for other time periods. He can then specify interactively either the values of the parameters that are required for the calculations or the actual flow or volume estimates. For example, if the *actual* volume of water that is abstracted from a specific river location is not known for the year of study, it can be set equal to the volume of the previous or following year. Alternatively it can be calculated through the annual licensed abstraction volume and the user-defined value of the uptake factor (i.e., the ratio of the actual to the licensed volume). The software provides further suggestions and default values: for example, when an abstraction licence has been granted for cooling purposes and a trade-effluent discharge consent is not associated with it, the user is advised to apply a return factor (of 0.95) to quantify the volume of water returned to the river (Bullock *et al*. 1994).

The estimated flows at the river sites are used in velocity–flow relationships of power type to provide the cross-sectional velocities at the sites. The coefficients of the relationships are entered by the user or derived from a regression analysis of the corresponding measurements at a downstream river flow station. The travel time at a reach between two successive river sites is then estimated from the length of the reach and the velocity at the upstream site.

Modelling of river water quality

This component allows the simulation of water quality in a river system, the assessment of performance of the applied model, and the identification of the model through the sensitivity analysis and calibration.

The simulation of river water quality is performed for a specific year or a shorter time period, applying a deterministic, steady-state, node-reach model of dissolved oxygen deficit DOD, carbonaceous biochemical oxygen demand CBOD, total ammonia $(NH_3)_{tot}$, and un-ionised ammonia $(NH_3)_u$. The river system is described through a network of head and junction nodes, waste-discharge points, surface-water abstraction points, compensation reservoirs, reaches that connect the above sites, and stations of river flow or water quality. The accounted point sources of pollution are the STW and trade effluents. The diffuse sources of pollution are assumed to be added at the end of each reach and contribute both to the flow and water-quality variation along the river. The physical and (bio)chemical processes considered are the removal of the biodegradable organic material, the natural reaeration of the river, the benthic oxygen demand, the nitrification of ammonia, and the dissociation of the latter to ammonium and hydrogen ions. The structure of the model and the input requirements have been described in detail elsewhere (Spanou & Chen 2000). It is noted that the concentration of the diffuse sources of pollution and the coefficients used in the modelling of the biochemical processes can have either the same value for all reaches (as was assumed in the earlier version of the model) or can be reach-dependent. The output of the simulation includes reports with the modelling conditions and simulation results, as well as plots of the predicted and observed concentrations of all constituents along a user-selected river and/or at all sampling points in the river system.

The model performance is assessed subjectively (i.e., based on the visual comparison of the predictions $pred_i$ with the observed values obs_i on graphs for each variable), as well as through the calculation of the sum of squares of residuals (res_i) S, the chi-squared statistics χ^2, and the efficiency factor E. When more than one variable is considered, S is defined as the weighted sum of squares of residuals (Whitehead *et al*. 1997). When the variance of the observed concentration at each individual sampling point $\sigma^2_{obs_i}$ is not known, it can be set equal to 1 by the user or to another value, e.g., the variance of observations at all sampling points σ^2_{obs} (Press *et al*. 1986):

$$S=\sum_{i=1}^{N} res_i^2=\sum_{i=1}^{N}(obs_i-pred_i)^2 \tag{2}$$

$$\chi^2=\sum_{i=1}^{N}\left(\frac{res_i}{\sigma_{obs_i}}\right)^2=\sum_{i=1}^{N}\left(\frac{obs_i-pred_i}{\sigma_{obs_i}}\right)^2 \tag{3}$$

$$E=1-\frac{\sigma^2_{res}}{\sigma^2_{obs}}=1-\left(\frac{1}{N}\sum_{i=1}^{N}(obs_i-pred_i)^2\right)\left(\frac{1}{N}\sum_{i=1}^{N}(obs_i-\overline{obs})^2\right)^{-1} \tag{4}$$

Detailed results of the assessment are provided optionally by the software following each run of water-quality simulation.

The sensitivity analysis is performed considering a user-defined percentage of change (a%) of a single parameter p_i, and the response of one or more state variables of the model Var_j, calculated at the start and/or the end of user-specified reaches. The sensitivity values S_{Var_j,p_i} are calculated using the parameter perturbation method (Chapra 1997):

$$S_{Var_j,p_i}=\frac{\Delta Var_j/Var_j}{\Delta p_i/p_i}=\frac{0.5[Var_j(p_i+\Delta p_i)-Var_j(p_i-\Delta p_i)]}{0.01\ a\ Var_j(p_i)} \tag{5}$$

The analysis output includes detailed results for all state variables studied, and graphs with the variation of $S_{Var_j p_i}$ along the nodes of the selected reaches.

The model parameters are finally calibrated combining the subjective assessment of the model performance, and the application of the equal-interval search and Levenberg–Marquardt optimisation methods (Fletcher 1980; Press *et al*. 1986). The objective of the optimisation can be the minimisation of S or χ^2, or the maximisation of the E value. During a 'calibration run', the software user has to select the state variables of the model, the calibration parameters, and the data points (i.e., the water-quality sampling points) that will be considered in the formulation of the objective function. He can also specify the values of the calibration parameters for one or more river reaches or edit the values that were entered for them in previous 'runs'. The optimisation is then confined to the river reaches where the selected sampling points are, and to the upstream ones up to the head nodes of the river system, with the exception of those reaches for which optimum parameter values have been defined by the user. In that way the calibration can be performed for different river parts and eventually for the whole river system, having reach-dependent parameter values. Alternatively, it can be performed directly for the whole system considering uniform values for all reaches. In both cases, the user also specifies the values of the parameters that control the optimisation (e.g., the λ of the first iteration or the number of successive successful iterations, which are required by the Levenberg–Marquardt method). The output from the above procedure includes optionally detailed results for each iteration, and summary tables with the variation of the calibration parameters and the objective function throughout the optimisation.

Assessment of river and point-source effluent water quality

The monitored water quality at the sampling points of the river system and the point-source effluents is assessed considering relevant classification schemes and standards. More specifically, the river water quality is classified according to the *River Ecosystem* scheme. The compliance with *objectives*, which are expressed as RE classes or absolute in-stream standards, is also evaluated. The compliance of the point-source effluents with the percentile and/or absolute limits that are specified in their discharge consents is finally assessed (National River Authority 1994*a, b*). The above methods are described in more detail in Spanou & Chen (1998, 2000).

Management of river water quality

The available methods include the formulation of an optimisation scheme for the improvement of river water quality, the stochastic calculation of concentration limits for the point-source effluents, and the stochastic estimation of river water quality downstream of the effluent discharges.

The optimisation procedure is applied when absolute in-stream standards have to be met. It uses the simulation model of river water quality and a heuristic algorithm that combines the waste-load relocation and the upgrade of the effluent-treatment facilities (Spanou & Chen 2000).

When percentile standards of effluent or river water quality are considered, the mass balance equation at the discharge points of the effluents is used in Monte Carlo simulations (Warn & Brew 1980). The mass balance equation has the form:

$$Q_{r,us}C_{r,us} + Q_w C_w = Q_{r,ds}C_{r,ds} = (Q_{r,us} + Q_w)C_{r,ds} \tag{6}$$

where Q_w and C_w are the flow and concentration of the effluent, $Q_{r,us}$ and $C_{r,us}$ are the river flow and concentration upstream of the discharge, and $Q_{r,ds}$ and $C_{r,ds}$ are the corresponding features of the river downstream of the discharge.

Equation (6) is used to assess the impact of the discharges on the RQOs for the river systems, by estimating the $C_{r,ds}$ that results from the current or from a suggested consent C_w. Inversely, it is used to calculate the discharge-consent conditions which are required for the achievement of the current or proposed RQOs, by estimating the C_w that allows a target $C_{r,ds}$ to be met (National Rivers Authority 1995).

During the Monte Carlo simulations, the software user specifies whether the functionally independent variables will be treated as statistically independent or statistically correlated variables. In the first case he selects for each variable a univariate distribution, which can be the uniform, normal or lognormal one, while in the second case he assigns to the vector of all variables the multivariate normal or lognormal distribution (Johnson & Kotz 1972; Johnson *et al*. 1994). The user may further enter the parameters of the distributions or specify the methods for their estimation. The parameters of the univariate distributions can be derived from the mean and the variance, the mean and a percentile value, or the coefficient of variation and a percentile value of the corresponding random variable. Furthermore, the above statistics can be calculated using different types of data for the variable. For example, the mean and variance of the C_w can be entered by the user or estimated based on actual data, effluent limits, or typical values for the specific type of effluent treatment. Finally, the user enters the values of the parameters which are required by the random variate generators or control the Monte Carlo procedure, such as the seed number for the uniform variate generator or the number of realisations that will be performed (Fishman 1996). The analysis output includes optionally the values of the variates for each realisation, user-specified statistics for the derived distribution of the functionally dependent variable, and graphs with the histograms for all variables.

Object-oriented approach

The software system that allows the performance of the above tasks has been developed using the object-oriented methodology throughout its analysis, design, and implementation. For the first two stages the Coad/Yourdon/Nicola (C/Y/N) and several other methods were considered (Graham 1994; Hutt 1994; Coad *et al*. 1995, Pree 1995). The derived *object model* was implemented in the Smalltalk Express™ programming environment.

The present model includes a large number of objects that are grouped into distinct subjects. All objects know how to perform certain *functions*, using their *attributes*, and collaborating with other *associated* or *composite* objects. The Watershed subject is responsible for the representation of the river basin, and includes objects such as the *RiverNode* or *WastewaterTreatmentPlant*. Similar objects have been used in other models presented in the literature: however, there are differences in their names or structure (Behrens & Loucks 1993; Fedra & Jamieson 1996).

An additional subject integrates objects that describe the mathematical and statistical tools required in the calculations, such as the *ProbabilityDistribution* or the *MonteCarloMethod*.

Five more subjects describe the components of the employed mathematical model that was discussed earlier. The methods of flow and water-quality analysis, and the pollution control strategies are represented through conceptual entities (such as the *WaterQualityClassification Scheme* or *TreatmentPlantUpgrade objects*). The tasks of each component are performed by corresponding *controller* objects that coordinate the interaction of 'method' objects with other 'method', and/or 'river-basin' objects. It is noted that controller objects have been similarly used by Shane *et al*. (1996) during the development of PRYSM, a software package for the planning of optimal hydro-energy production from reservoirs in river basins.

Finally, five subjects have been defined to handle: (a) the graphical representation of the river basin, and the user interaction during the application of the software, (b) the search, storage and retrieval of data related to the river basin and the applied models, (c) the preparation of reports and plots with the analysis results, and (d), (e) the communication of the software with the Word™ and Excel™ applications.

The present analysis allows the integration of many mathematical or statistical models that perform different tasks. The application of object-orientation for the integration of models has been similarly followed during the development of *WaterWare*, a comprehensive information and decision-support system for integrated river-basin management (Fedra & Jamieson 1996). However, in that case the attributes of the river-basin objects are provided as input data or are updated by the output of models, which are external software systems and not objects themselves.

A detailed description of the analysis can be found elsewhere (Spanou 2000). The object model that corresponds to the earlier version of the software has been also summarised in Spanou & Chen (2000). A representative part of the recent model extension is discussed below. It concerns the calibration of the water-quality model, and is also shown in Figure 3 following the (C/Y/N) notation. The principal object is an *instance* of the *ModelCalibrationController* class, and is responsible for the objective estimation of values for the *calibrationParameters* of a model. To fulfil its responsibility it communicates with (a) a *simulationModelController* object (the *RiverWater QualitySimulation*) which performs simulations using specified values of the model parameters, (b) a *ModelPerformanceAssessment* object which assesses the goodness-of-fit of a model to the data based on a *criterion* value, and (c) an *optimisationMethod* object which searches for an optimum point of the *objectiveFunction* (i.e., of the *modelPerformanceCriterion*), changing the values of its *optimisationVariables* (i.e., of the *calibrationParameters*).

Object (a) has been presented in Spanou & Chen (2000). The *ModelPerformanceAssessment* interacts with a *simulationModelController*, a *dataManager* (the *River WaterQualityStationDataManager*) and a *ModelPerformanceCriterion*. The latter is an *abstract* class and specifies that each *concrete subclass* (the *SumOfSquaresOf Residuals*, the *ChiSquareCriterion*, or the *Efficiency Factor*) should know how to calculate the criterion value, as well as the values of the first and second derivatives of the criterion with respect to one and two parameters respectively. The *OptimisationMethod* is similarly an abstract class and defines that all its subclasses (at present the *LevenbergMarquardtMethod* and the *Ndimensional EqualIntervalSearch*) should know the *purposeOfOptimisation* (i.e., the minimisation or maximisation of the *objectiveFunction*), and how to *search for an optimum point* that will achieve this purpose. The *objectiveFunction* has been defined as a distinct object. It knows its *formulationBlock*, the *firstDerivativeBlock*, and the *secondDerivativeBlock*, and can implement these blocks (i.e., the pieces of code) in order to calculate respectively its *value*, and the elements of its *gradientVector* and *hessian Matrix* using the current values of the associated *optimisationVariables*.

From the above it can be realised that the operations required for the calibration of the model are distributed to many cohesive objects and this facilitates the reuse and extension of each object. In addition, it shows the significance of the well-defined *interfaces* of the objects. For example, the model of the *ModelCalibrationController* is specified through the attribute *simulationModelController*. So it is not restricted to be a *RiverWaterQualitySimulation* object but it can be any other object that describes a model (such as a *RiverFlowSimulation*). It is noted that the model calibration is also performed in AQUASIM, an object-oriented software system that provides enhanced flexibility in the formulation and assessment of alternative water-quality models. However, in that system it is handled by a single, overloaded object, and can be

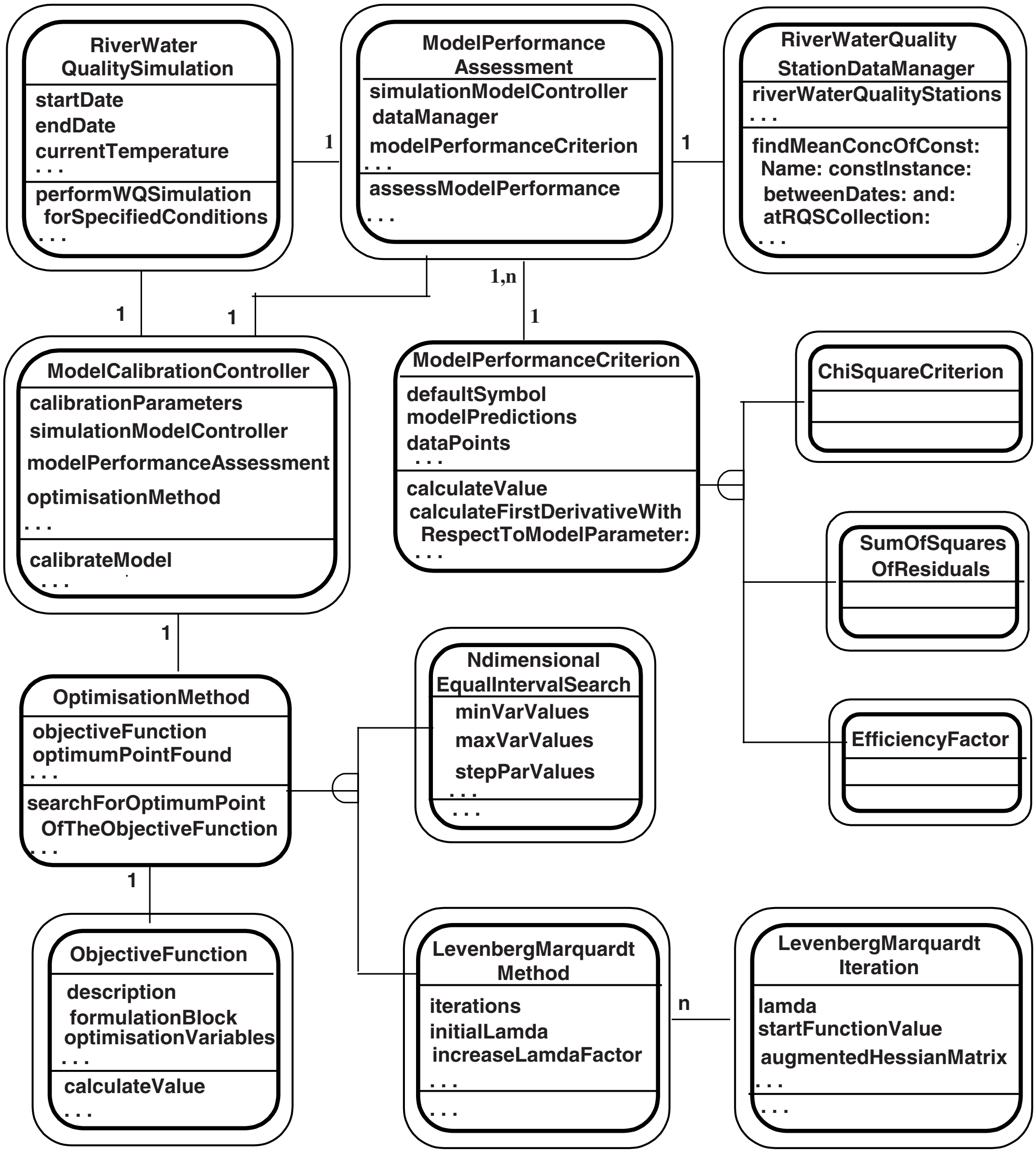

Figure 3 | Part of the object model for the parameter estimation of the water-quality model [C/Y/N notation in Coad *et al.* (1995)].

performed by applying only the weighted least squares technique (Reichert 1995).

RESULTS AND DISCUSSION

Benefits from the application of object-orientation

The application of the object-oriented method is advisable for the development of complex and user-friendly software systems. It also results in environments that are easier to evolve over time, and can communicate more efficiently with external systems, in comparison to traditional programming environments (Booch 1994). The above benefits have been delivered in the development of the present software system, and are discussed in the following.

Many objects that have been defined to describe the domain of water-quality management can be reused in object models for water-quantity or integrated river basin management. The *FlowDurationAnalysis* object, e.g., can be used for the flow estimation in low-flow as well as flood studies. The *MonteCarloMethod*, that at present is an associate of the *DischargeConsentStochasticCalculation* object, will be required for the stochastic simulation of water quality along the whole river system.

In addition, the structuring of objects in *class hierarchies* and the implementation of inheritance have delivered reusability of code. The *MinimalUniform VariateGeneratorWithShuffle*, for example, has inherited the *seedSequenceNumber*, *generatedDeviates*, and *probabilityDistribution* attributes from the *UniformVariate Generator*, *UnivariateGenerator*, and *RandomVariate Generator* objects that satisfy the specialised class–generalised class relationship. The application of *polymorphism* and *information hiding* has further supported the extension of the model. All *Univariate Distribution* subclasses, for example, can *calculate the theoretical mean from the distribution parameters*: however, each subclass applies different formulas when responding to this message. The use of *whole-part* structures also facilitates the model extension. The *RiverSystem*, for example, can be easily connected with additional *components* such as weirs.

The principle of *information hiding* and the communication of objects with *messages* have facilitated the incorporation of alternative models for the performance of each task of the management study. In addition, they have assisted the integration of models that perform different tasks of the study in a single environment.

The object-oriented approach to the design of the software has increased the efficiency in the use of data resources, offered flexibility in the form of presenting the model results, and facilitated the communication of the software with external systems. The monitored river water quality, for example, is required for the RE classification at the sampling points, the calibration of the simulation model, and the calculation of the discharge consents. In all cases, the data retrieval is performed by the same object, the *RiverWaterQualityStationDataManager*, and hence any errors in the applied procedure are minimised or at least localised. A *LevenbergMarquardtReport* can further prepare summary or detailed reports of the optimisation procedure, by communicating with *Levenberg MarquardtIteration* objects and requesting the values of their attributes (e.g., the *lamda*, the *augmentedHessian Matrix*, the *start* and *end value* of the *objective function*). The *ExcelSI* object can also handle the interaction between the software and different versions of the Excel™ application.

The Upper Mersey river system

The Upper Mersey river system that was studied with the developed software has a catchment area of 695 km^2 and includes the Etherow, Goyt, Tame, and upper-Mersey rivers, along with their tributaries. The length of the four main rivers is 148 km. The Etherow is a tributary of the Goyt, and the upper-Mersey is formed by the confluence of the Goyt and Tame rivers (National Rivers Authority North West Region 1996). In the present work the system of each one of the above rivers has been studied separately. Their node-reach representation and the point-pollutant sources they receive, as well as a map of the overall catchment, are shown in Figure 4.

Low-flow estimation at gauged river sites

Annual low-flow frequency curves and flow duration curves of duration D equal to 1, 7, 10, and 30 days have been derived for the main flow-monitoring stations of the catchment. These are the Compstall, Marple Bridge, and Ashton Weir stations, which are on the Etherow, Goyt, and upper-Mersey rivers respectively, and the Broomstairs Bridge and Portwood stations which are on the Tame river (Figure 4). Their daily-flow records were provided by the Environment Agency. When the number of days with missing (i.e., not measured) flows in a year was greater than the duration for a curve, then the data of the whole year were excluded from the calculations. The data that were finally used covered a period of 20 or more years (Tables 1 and 2), so they are expected to provide sufficiently accurate curves (Institute of Hydrology 1980).

All methods that are available in the software were applied for the analysis. Hence the sample-based, low-flow frequency curves were prepared using Eq. (1) with

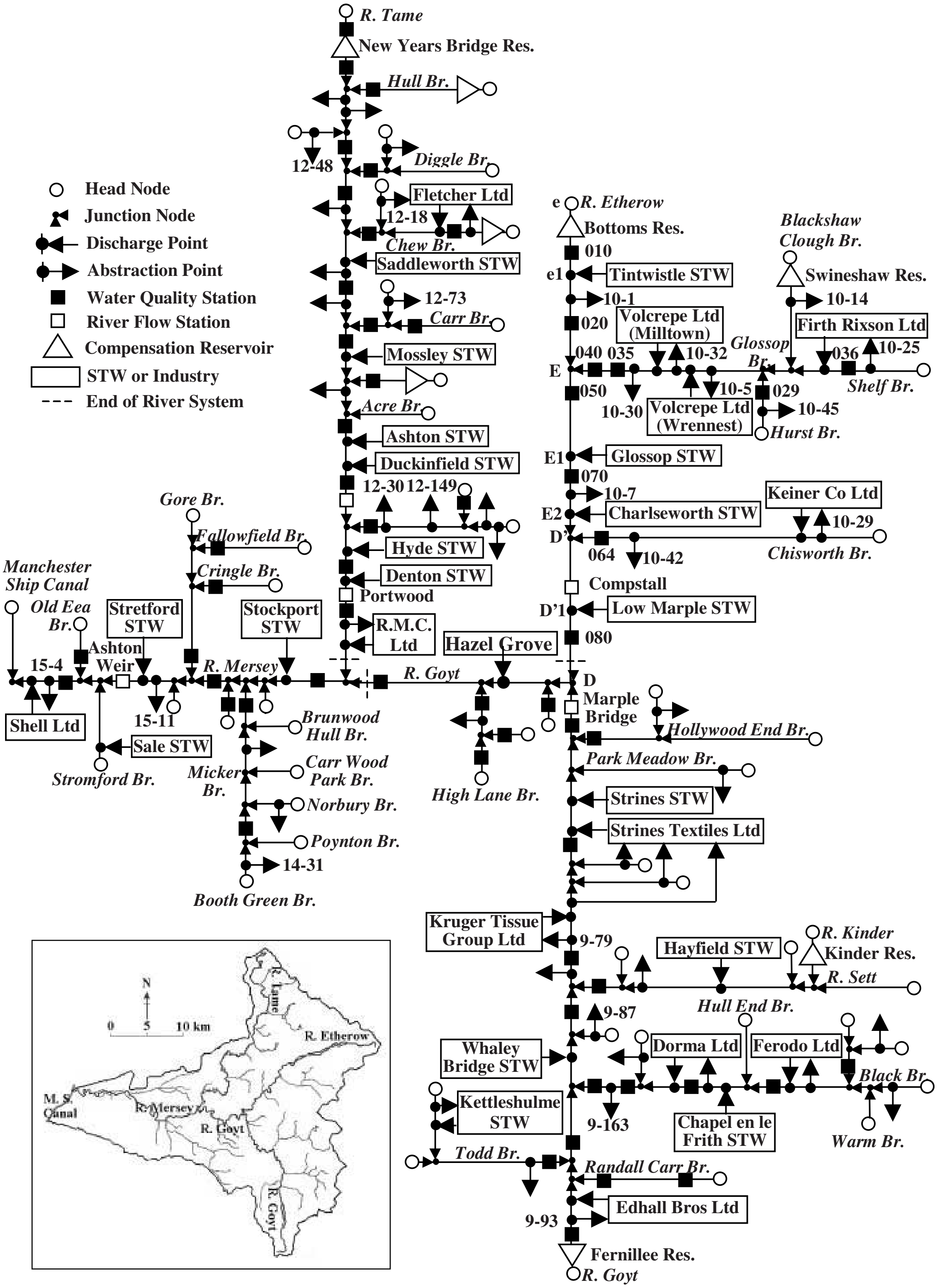

Figure 4 | Map of the Upper Mersey river catchment and its representation in the present study.

Table 1 | The annual, 7-day, low-flow frequency curves, and the flows of return period 2 years for the flow monitoring stations of the Upper Mersey catchment.

	$Q_{min,j}$ data[1]		Fitting of $EX_{III,s}$ (ε,u,k) distribution						Plot on W p.p.[4]	
Station	**N**	**Min flow (m^3/s)**	ε	**u**	**k**	**7Q2 (m^3/s)**	**Error[2] c=0.0**	**Error[2] c=0.44**	**7Q2 (m^3/s)**	**r^2 (–)**
Compstall	23	0.046	– 0.7061	0.8086	5.095	**0.703**	1.1722	1.0800	**0.656**	0.950
Marple Bridge	20	0.416	0.1770	0.8096	3.455	**0.746**	0.0472	0.0474	**0.720**	0.948
Broomstairs Bridge	21	0.010	– 1.2191	1.2116	5.095	**1.043**	6.2015	0.4041[3]	**1.063**	0.938
Portwood	20	0.868	0.5326	1.5544	2.995	**1.437**	0.0326	0.0249	**1.394**	0.976
Ashton Weir	20	2.026	1.6165	3.3969	2.220	**3.126**	0.0248	0.0205	**3.030**	0.956

[1]The $Q_{min,j}$ flow is the minimum 7-day averaged flow in the data of year *j*.
[2]The error is calculated as the sum {[sample predicted ln($Q_{min,j}$)]^2} for $Q_{min,j} \geq 0$ and *j=1* to N.
[3]A negative $Q_{min,j}$ flow has been predicted from the $EX_{III,s}$ (ε,u,k) distribution.
[4]The points of the curve have been plotted on Weibull probability paper with $\epsilon = 0$ and $k = 4$.

Table 2 | The annual, 1-day, flow duration curves, and the flows exceeded for 95% of the time for the flow monitoring stations of the Upper Mersey catchment.

			Linear regression[3]		
Station	**Time period with adequate data (years excluded)**	**Q95(1) (m^3/s) FDC reading[1] [IH result[2]]**	**Q95(1) (m^3/s)**	**Error[4] (%)**	**r^2 (–)**
Compstall	1972–1996 (1974, 1978)	0.676 [0.680]	0.404	– 40.2	0.926
Marple Bridge	1970–1996 (1971–1975, 1978–1979)	0.700 [0.718]	0.803	14.7	0.970
Broomstairs Bridge	1975–1996 (1994–1995)	1.097 [1.118]	0.643	– 41.4	0.954
Portwood	1976–1996 (1978)	1.365 [1.347]	1.426	4.5	0.971
Ashton Weir	1977–1996	3.090 [3.086]	2.984	– 3.4	0.974

[1]Readings of the *Q95*(1) values from the flow duration curves.
[2]Source: (Marsh, 1997).
[3]The linear regression model is of the form: log[Q(1)]=$a \times$NormalVariate+b.
[4]The error is calculated as [*Q95*(1)regression–*Q95*(1)curve]/*Q95*(1)curve×100%.

$c = 0$ and $c = 0.44$. Their points practically coincided for all return periods up to 6 years, and for some stations up to 8 years: however, significant deviations were encountered for longer return periods. The curves derived using $c = 0.44$ were plotted on Weibull probability paper with parameters $\varepsilon = 0$ and $k = 4.0$ to assess the assumption that the low flows follow the corresponding distribution. It is noted that the above procedure is also implemented by the CEH in the UK. Visual inspection of the obtained graphs showed that a linear pattern was also exhibited more clearly for the Portwood station (Figure 5). A linear regression analysis that was performed on the (Weibull variate,

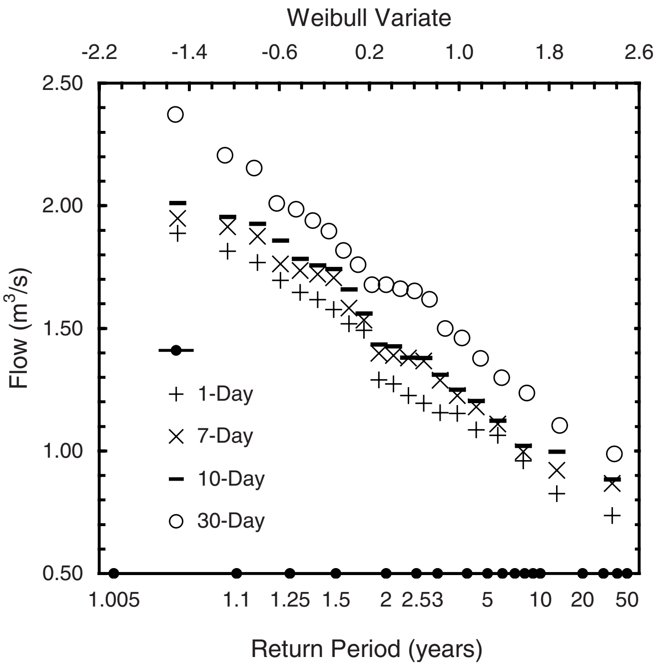

Figure 5 | Low flow frequency curves drawn on Weibull probability paper of ε=0 and k=4.0 for the Portwood flow station (no. 692423) of the Tame River, for durations of 1, 7, 10, and 30 days.

flow) points of the curves verified the above, as the highest values of the coefficient of determination r^2 were obtained for the same station, and were in the range of 0.976–0.994. The plots also showed that, for all stations and return periods, the low-flow estimates of longer duration had higher values.

The theoretical, $EX_{III,s}(\varepsilon,u,k)$, low-flow frequency curves were also derived for each station and duration. Smaller errors of fitting the distribution to the data were obtained for the Ashton Weir, Portwood, and Marple Bridge stations. Increased errors, and negative values of the lower flow limit ε which do not have physical meaning, were obtained for the other two stations for most durations studied.

The low flows with return periods of 2 and 20 years, *DQ2* and *DQ20*, were finally read from the sample-based curves, and also predicted from the theoretical ones. The *7Q2* flows are of particular interest as they are often used by the water industry in the UK for the design of wastewater treatment plants. Those derived from the sample-based curves ranged from 0.656–3.030 m^3/s (Table 1), and were higher than the 7*Q*20 flows that are used for design purposes in Canada and the United States and ranged from 0.267–2.128 m^3/s.

The overall analysis revealed the difficulty of identifying which distribution describes best the low flows at a given river site. Additional distributions (such as the log-Pearson Type III) and other measures of goodness of fit (such as the χ^2 criterion or the Kolmogorov–Smirnov procedure) can be considered (McMahon and Arenas 1982). However, these measures describe how well a distribution fits all the data, while most low flows of interest lie at the tails of the distribution.

The flow duration curves were prepared using the same method as the one applied by the CEH, i.e., considering logarithmic flow intervals in the range of 1–1000% of the mean flow over the years of record *MF* (Gustard *et al.* 1992). The curves were plotted on lognormal probability paper, and the flows exceeded for 95% of the time *Q*95(*D*) were estimated. The Q95(1) flows are of particular interest, as they are used by the Environment Agency in the UK for the setting of discharge consents and abstraction licences. These flows ranged between 0.676–3.090 m^3/s (Table 2) and were in general agreement with CEH results (Marsh 1997). They were also lower than the Q95 flows of higher duration. The same relationship applied to all flows exceeded for more than 20% of time. It is noted that the Q95(1) flows were higher than the estimated 7*Q*2 flows. Similarly the *Q*95(*D*) flows were higher than the *DQ*20 estimates. This is in agreement with theory, which states that the low flows obtained from low-flow frequency curves are more rare events than the corresponding ones obtained from the flow duration curves (Institute of Hydrology 1980).

If the flow data followed the lognormal distribution, the flow duration curves that were drawn on the corresponding probability paper should take the form of straight lines. Inspection of the plots revealed that most curves have a concave part at percentage exceedence probabilities between 60–80%, and/or convex parts closer to their ends, e.g., at probabilities smaller than 5% or greater than 99.5%. In addition, the curves tend to a sigmoid form for increased durations. To quantify their proximity or deviation from straight lines, a linear regression model was fitted to their (normal variate, logarithm of

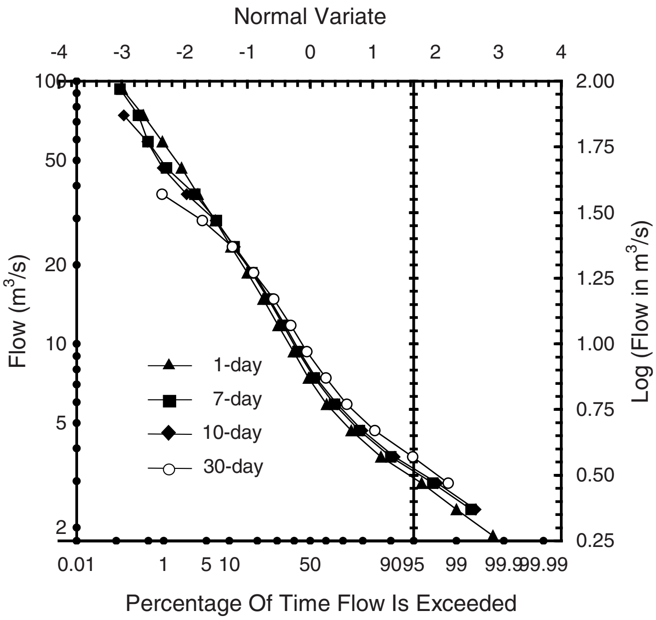

Figure 6 | Flow duration curves for the Ashton Weir flow station (no. 692726) of the upper-Mersey river, derived for durations of 1, 7, 10, and 30 days using logarithmic flow intervals.

flow) points. For most stations and durations studied, the values of r^2 were in the range of 0.970–0.985 showing a moderate fitting of the linear model. Smaller relative errors between the predicted and estimated $Q95(D)$ flows were obtained for the Ashton Weir station (Figure 6) and ranged from 2.5–7.4%.

The flow duration curves were also prepared based on the quantile values of the flow data or using a large number of linear flow intervals (e.g., 400). The obtained results were similar to the above, for flows exceeded from 0.1–99% of time.

Low flows at ungauged sites

The annual $Q95(1)$ flows and their natural and artificial components have been estimated for all ungauged nodes of the four river systems of the study. The flows at the final node of each system were boundary conditions for the downstream systems.

The artificially influenced flows were estimated based on (i) licensed and (ii) actual flow data. The licensed flows have been widely used in the UK for such calculations. The actual flows have attracted the recent interest of hydrologists, as they result in more realistic estimates of the waste assimilative capacity of the river and its potential for abstraction use. However, such flows are treated with caution, as there are errors in their measurement, and data are usually lacking for one or more influences (Bullock *et al.* 1994). Due to the above, the data availability and the assumptions that were made in the study of the Upper Mersey system will be presented in detail.

The flowrates of the STW effluents were set based on (i) the design and (ii) the actual DWF. The rates of the surface-water abstractions were calculated using (i) the licensed and (ii) the actual annual volumes of abstraction. In both cases the flowrates of the trade effluents and the reservoirs were calculated based on the annual discharge-consent volumes, and the compensation flows respectively, as these were the only relevant data available. It is noted that, although the annual actual abstraction volumes were made available by the Environment Agency for many abstractions for the years 1994–1997, the actual DWF were made available by North West Water Ltd. as mean values for the period 1994–5. Due to the above the artificial component of the $Q95(1)$ flows was estimated in case (ii) using the actual data for the year 1994, although ideally it should be derived using data of a drier year, e.g., of 1996 (Marsh & Sanderson 1997).

During the application of the first approach (i), a return factor of 0.95 was used for 7 abstractions that have been licensed for cooling purposes. The effluent volume of Firth Rixson Ltd. was not available, and was assumed to be equal to 105% of the annual licensed volume of the associated 10–25 abstraction.

In the second case (ii), the volumes of 20 of the 58 abstractions were unknown for the year 1994, and were estimated based on data of subsequent years. The resulting volumes were a small percentage (less than 4%) of the total volumes abstracted from the Etherow, Tame, and upper-Mersey river systems. This percentage was more significant (15%) only for the Goyt system, mainly due to the volume of abstraction 9-163 from the Black Brook that was estimated using data from 1996. In addition, uptake factors were used for 5 abstractions with no actual data. The factors were estimated based on the available actual and licensed data of all abstractions in the Upper Mersey system. Their values and the corresponding purposes of use of the abstracted water are: 0.75 (amenity purposes), 0.15 (spray irrigation), and 0.50 (manufacturing/boiler feeding, cooling, and domestic/agricultural purposes). These values are in general agreement with the uptake factors derived by the IH for different abstraction purposes and different regions in the UK (Bullock *et al.* 1994). The volumes estimated using the uptake factors were a small percentage (less than 2%) of the total volumes abstracted from the Goyt and Tame systems. The percentage was significant (95%) for the upper-Mersey system, due to the 15-4 abstraction. However, the latter is downstream of the Ashton Weir flow station, so it does not influence the naturalisation of flows for the upper-Mersey system.

The obtained results showed that the main artificial influences were the compensation reservoirs for the Etherow system, the STW effluents for the Tame and upper-Mersey systems, and the abstractions for the Goyt system. The lack of actual data for the reservoir-release rates and the trade-effluent flowrates was more important for the Etherow and Tame rivers respectively. When the artificially influenced flows were calculated using licensed data, they were 80–84% of the total flows estimated at the end of the Etherow and Tame rivers, and 58–69% of the total flows at the end of the Goyt and upper-Mersey rivers. When they were calculated using actual rather than licensed data, they were 5% and 1% smaller for the first two rivers respectively. However, they were 45% and 17% higher for the last two rivers, because the corresponding actual abstraction volumes were around 40% of the licensed abstraction volumes.

The approach that was followed can be refined by including the groundwater abstractions in the artificial influences under consideration (Bullock *et al.* 1994). However, in the Upper Mersey river system they correspond to the 10% of the total abstracted volume, so they are not expected to alter the results significantly.

The natural $Q95(1)$ flows were then calculated using the average catchment contribution coefficient of 0.0023 $m^3/(s\ km^2)$, and the catchment area of each point that was extracted using the WIS system. Improved estimates can be obtained by applying the regional equations that have been developed by the IH and relate the flow with the area, slope, rainfall, and geology of the catchment (Gustard *et al.* 1992).

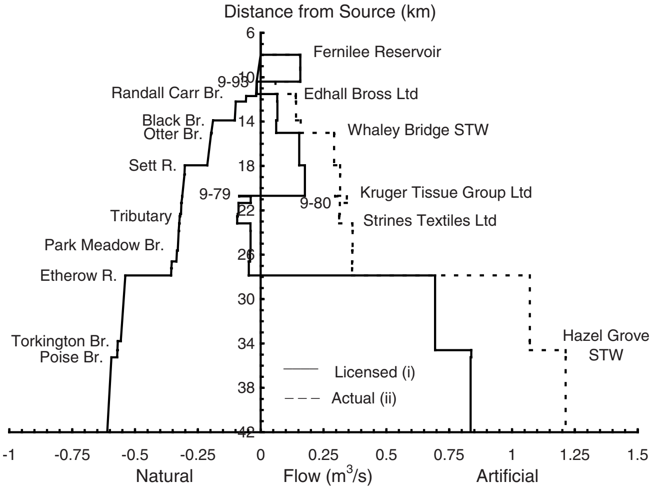

Figure 7 | Annual residual flow diagrams of the Goyt river based on the $Q95(1)$ low flows of the Marple Bridge (no. 692015) flow station and (i) the licensed or (ii) the actual flows of the artificial sources in 1994.

The residual flows were finally derived. The estimated flows had negative values for nine abstraction points, which are located on tributaries of the river systems and have a small drainage area (less than 9 km^2). For five of those points (the 10-25, 14-31, 12-73, 12-18, and 12-48), negative flows were estimated in both sets (i) and (ii) of the results. The ratio of their artificial to natural flows ranged from 2.59–3.38 based on set (i), and 1.22–1.98 based on set (ii). For the other four points (i.e., the 10-29, 9-87, 12-149, and 12-30) negative flows were obtained only in set (i) of the results. In addition, the ratio of artificial to natural flows ranged from 1.07–2.00, showing that the natural flows at those river sites have a smaller inadequacy to meet the abstraction requirements. Based on the above, the licenses of abstracting water from the nine points identified should be reviewed.

The residual flow diagrams along the four main rivers were finally constructed. Such a diagram is presented for indicative purposes for the Goyt River in Figure 7. The natural flow along the river is increased by several tributaries, the most important being the Black Brook, Sett River, and Etherow River. The artificial flow is increased

Table 3 | The v–Q and h–Q relationships derived for the flow monitoring stations using the data of 1994, and the corresponding v and h estimates for the river systems based on the mean simulated flows of the same year (Q in m^3/s, v in m/s, and h in m)

Station (river)	Q, v, and h data at the river flow station	v–Q and h–Q relationships	Q, v, and h estimates for the river system (main river)
Compstall (Etherow)	$0.8543 \le Q \le 2.4350$ $0.1627 \le v \le 0.4063$ $0.082 \le h \le 0.166$	$v = 0.1862Q^{0.751}$, $R^2 = 0.8561$ $h = 0.0959Q^{0.552}$, $R^2 = 0.9676$	$0.0100\ (0.5255) \le Q \le 2.8266$ $0.0059\ (0.1148) \le v \le 0.4063$ $0.008\ (0.067) \le h \le 0.170$
Marple Bridge (Goyt)	$0.5952 \le Q \le 4.1685$ $0.2400 \le v \le 0.5090$ $0.128 \le h \le 0.432$	$v = 0.3078Q^{0.335}$, $R^2 = 0.9262$ $h = 0.1748Q^{0.626}$, $R^2 = 0.9959$	$0.001\ (0.1579) \le Q \le 8.2597$[2] $0.0304\ (0.1659) \le v \le 0.6244$ $0.002\ (0.055) \le h \le 0.655$
Portwood (Tame)	$1.2266 \le Q \le 5.928$ (ul = 6[1]) $0.2984 \le v \le 0.7346$ $0.109 \le h \le 0.288$	$v = 0.2778Q^{0.557}$, $R^2 = 0.8931$ $h = 0.094Q^{0.609}$, $R^2 = 0.9805$	$0.0028\ (0.0263) \le Q \le 4.4951$ $0.0105\ (0.0366) \le v \le 0.6417$ $0.003\ (0.010) \le h \le 0.235$
Ashton Weir (Mersey)	$3.3807 \le Q \le 18.8619$ (ul = 20[1]) $0.1382 \le v \le 0.4384$ $0.321 \le h \le 0.733$	$v = 0.0536Q^{0.701}$, $R^2 = 0.9641$ $h = 0.1707Q^{0.497}$, $R^2 = 0.9434$	$0.0588\ (12.7547) \le Q \le 18.6685$ $0.0074\ (0.3193) \le v \le 0.4171$ $0.042\ (0.605) \le h \le 0.731$

[1]An upper limit (ul) was used for the collection of Q data in 1994. Any flows exceeding that limit, as well as the corresponding v and h measurements were excluded from the regression analysis.
[2]The flow of 8.2597 m^3/s is estimated for the end of the Goyt River, and includes the inflow from the Etherow River. The estimated flow at the Goyt River prior to the confluence with the Etherow is 4.4246 m^3/s, so it is close to the maximum measured flow of 4.1685 m^3/s.

significantly from the compensation of the Fernillee Reservoir, and the effluents of the Whaley Bridge and Hazel Grove STWs. In profile (i), the 9-93 and 9-79 abstractions of Edhall Bros Ltd. and Kruger Tissue Group Ltd. result in negative artificial flows. The flowrate of the first trade effluent is adequate to reverse that: however, the flowrate of the second effluent is much smaller than the corresponding abstraction flowrate (i.e., less than 18%). Hence, the artificial flow along the Goyt River remains negative up to the confluence with the Etherow River. The above problem does not occur in profile (ii), and in addition, the actual abstraction volumes of 9-79 in years 1994–1997 are around 7% of the annual licensed volume. It is noted that the corresponding license (9-79) has been granted for cooling, manufacturing, and boiler feeding purposes: however, a discharge consent has also been granted to the license holder, so a zero return factor was used for this abstraction in the calculations. Based on the above the license 9-79 should be reviewed, considering the reduction of the abstracted volume of water or the specification of a return factor.

Simulation of river flow and water quality

Simulations of flow and water quality along the four river systems have been undertaken for the year 1994. The artificial flows were calculated using actual data as in approach (ii) of the previous section: however, the STW-effluent rates were set based on the actual flows to full treatment rather than the DWF data. The natural flows were estimated using the catchment contribution coefficient of 0.0258 m^3/(s km^2) (i.e., the average value that was derived over the four river systems based on the mean gauged flows and the artificial-flow estimates for the simulation period).

Velocity-discharge and stage-discharge equations of the form $v = c_{1v}Q^{c_{2v}}$ and $h = c_{1h}Q^{c_{2h}}$ were developed for the flow stations of the four main rivers, applying a regression analysis to monthly data for the year 1994 that were provided by the Environment Agency. The derived equations were used to provide velocity and stage estimates for all ungauged sites of interest on the corresponding river systems (Table 3). The travel times along the river reaches were subsequently estimated. The cumulative travel times from the start up to the final node of the Etherow, Goyt, Tame, and upper-Mersey rivers were estimated as 14.2, 22.6, 64.5, and 21.7 hours respectively. The start nodes for the first three rivers represent the outflows from the Bottoms, Fernillee, and New Years Bridge reservoirs.

The above methods provided rough estimates of the mean river flows and the travel times for the year of simulation. If the simulation period was shorter or more accurate estimates were wanted, then an advanced flow-routing model should be applied, and detailed information of the channel geometry would be required.

The applied water-quality model does not describe the variation of temperature and pH along a river system. Hence these water-quality parameters were assumed to be the same for all river reaches, and were estimated by averaging the available data over all sampling points for the year 1994 (e.g., for the Etherow system they were found equal to 10.57°C and 7.273 respectively). The concentrations of CBOD and $(NH)_3)_{tot}$ in each point-source effluent were calculated as the mean values of the effluent data for the same year. When such data were not available the measurements in a previous or following year were used. The DO concentration of all effluents was set by the user to be equal to 5.5 mg/l. This was the mean value of two measurements that were available in 1994, at the effluents of the Duckinfield and Stockport STWs.

The concentration C_d of the diffuse-pollution load, and the $k_{BODrem,20}$, $k_{nitr,20}$, and $k_{reaer,20}$ parameters of the water-quality model (i.e., the specific rates of the CBOD removal, nitrification, and reaeration processes at 20°C) were estimated during the model calibration. The state variables of the model that were considered during the calibration were the river concentrations of CBOD, $(NH)_3)_{tot}$, and DOD. The $(NH_3)_u$ was not taken into account because it is derived from the $(NH_3)_{tot}$ so it is related indirectly to the above parameters. The rate of benthic oxygen demand r_{bod} was also ignored due to the high variability of its values in different rivers, and the lack of any information regarding the benthic activity in the Upper Mersey system. The temperature-correction factors θ_{BODrem}, θ_{reaer}, and θ_{nitr} were set equal to the values used in the QUAL2E model (Brown & Barnwell 1987). The values of the first two factors are also the same as those used in the QUASAR model (Whitehead *et al*. 1997).

The model was initially calibrated assuming uniform parameter values and including in the objective function all the state variables and parameters that were described above. Reach-dependent parameters and 'sequential' calibration were subsequently considered (Lewis *et al*. 1997). Hence, the $k_{BODrem,20}$ and the C_d of CBOD were estimated first based on the CBOD predictions. The $k_{nitr,20}$ and C_d of $(NH_3)_{tot}$ were estimated similarly based on the $(NH_3)_{tot}$ variation. The $k_{reaer,20}$ and C_d of DOD were calibrated last based on the DOD variation.

Detailed results of parameter estimation and water-quality simulation are presented in the following for the Etherow River system. The reach-dependent parameter values and the model performance assessment are summarised in Tables 4 and 5. The variation of the predicted and measured water quality along the main river is shown in Figure 8.

For the first reaches of the Etherow River (i.e., from the outflow of the Bottoms reservoir up to the 10-1 abstraction point), for all reaches of the Blackshaw Clough, Shelf, and Chisworth Brooks, and for the reaches of the Glossop Brook up to the 10-30 abstraction point, small C_d values were estimated (i.e., 2.2 mg/l for the CBOD, 0.08 mg/l for the $(NH_3)_{tot}$, and 0.15 mg/l for the DOD). A boundary condition of -0.01 mg/l DOD was further introduced for the Hurst Brook based on the data of its 029 sampling point. The estimates of the kinetic parameters for the above reaches are within the range reported in Brown & Barnwell (1987). The $k_{nitr,20}$ and $k_{reaer,20}$, for example, are 0.72 d^{-1} and 6.5 d^{-1} respectively. The $k_{BODrem,20}$ is 0.02 d^{-1} for the Etherow River and the Blackshaw Clough Brook that have water-supply reservoirs on their headwaters, while it is 0.46 d^{-1} for the Chisworth Brook that receives the effluent of Keiner Co. Ltd. This confirms that low $k_{BODrem,20}$ values are applied to rivers with potable water, and high values to rivers that receive municipal, mechanically treated waste (Jorgensen 1986). The model predictions also agree with the measurements at points 010 of Etherow River, 036 of Shelf Brook and 064 of Chisworth Brook. The $(NH_3)_{tot}$ and DOD at the last point are the highest in the river system, and result from the discharge of the Keiner Co. Ltd. effluent.

Table 4 | The rate coefficients and the diffuse-load concentration used in the simulation of water quality for the Etherow river system and the year 1994

River stretch			Reaction rate at 20°C $k_{process,20}$(d^{-1})			Diffuse-load concentration C_d(mg/l)		
Start node	End node	Length (km)	$k_{BODrem,20}$ [0.02–3.4][1] (θ=1.047)	$k_{nitr,20}$ [0.1–1.0][1] (θ=1.083)	$k_{reaer,20}$ [0.1–100][1] (θ=1.024)	DOD	CBOD	$(NH_3)_{tot}$
10-1	E	1.201	0.02	0.72	3.5	0.35	21.0	0.08
E	E1	1.876	2.50	0.72	14.0	0.15	2.2	4.50
E1	10-7	1.307	2.50	3.00	25.5	0.15	2.2	0.00
10-7	D	7.543	0.02	3.00	25.5	0.15	2.2	0.00
10-30	E	1.427	0.02	0.72	27.0	0.15	27.0	0.08
10-29	D′	2.366	0.46	0.72	6.5	0.15	2.2	0.08
All other		7.765	0.02	0.72	6.5	0.15	2.2	0.08

[1]The default range of values that is provided by the software for the $k_{process,20}$ parameters.

Table 5 | The performance of the water-quality simulation model for the Etherow river system for the year 1994, and statistics of the corresponding data obtained at the sampling points of the system

State variable *i*: concentration of DOD, CBOD $(NH_3)_{tot}$ (mg/l)								All state variables	
Variable	N_{obs}[1]	$\overline{obs}$	σ^2_{obs}	σ^2_{res}	E	S	χ^2	$\sum_i S_i$	$\sum_i \frac{(\sigma^2_{res})_i}{(\sigma^2_{obs})_i}$
DOD	9	0.317	0.1471	0.0006	0.9962	0.0050	0.0340		
CBOD	9	3.79	1.656	0.634	0.6169	5.7083	3.4470	5.7971	0.3938
$(NH_3)_{tot}$	9	0.566	1.3342	0.0093	0.9930	0.0839	0.0629		

[1]An *observation* for a state variable is the mean measured value of the state variable at a single sampling point of water quality in 1994.

Along the '10-30 E' reach of the Glossop Brook (i.e., between the point 10-30 and the junction with the Etherow River), and from point 035 to 040, the observed CBOD increases considerably while the DOD decreases. There is no point-source effluent to explain this variation, so high values of CBOD C_d (27 mg/l) and $k_{reaer,20}$ (27 d^{-1}) were applied. The resulting model predictions show the right trend of variation along the reach, and although they overestimate the CBOD at site 035, they match the observations at the reach end (site 040).

The CBOD also increases significantly from point 010 to 020 on Etherow River. The Tintwistle STW effluent has a small load, so a high CBOD C_d (21 mg/l) was used for the '10-1 E' reach. A high DOD C_d (0.35 mg/l) and a reduced $k_{reaer,20}$ (3.5 d^{-1}) were also applied to match the observed at DOD point 020.

The CBOD increases even more at site 050 after node E, although there is no additional waste discharge. It then decreases at site 070 to a value lower than that of site 020, although the river receives the high load of the Glossop STW effluent at node E1. To match the

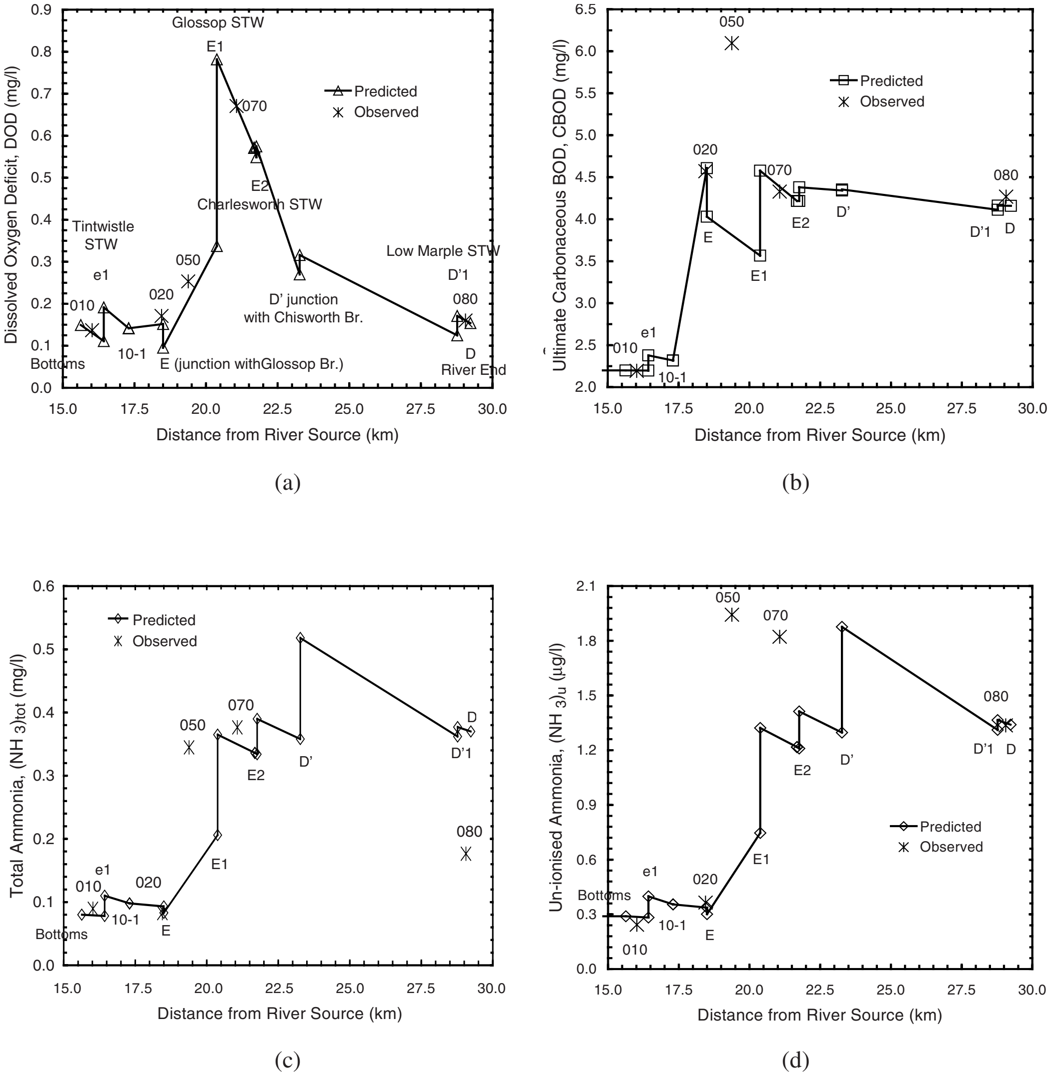

Figure 8 | Simulated and observed concentration of DOD, CBOD, $(NH_3)_{tot}$ and $(NH_3)_u$ along the Etherow River for the year 1994.

observed CBOD at point 070 a high $k_{BODrem,20}$ (2.5 d^{-1}) was applied to the reaches between nodes E and 10-7. If a high C_d was also used for the 'E E1' reach (to match the CBOD at point 050), then the required $k_{BODrem,20}$ would exceed the values reported in the literature. The measured $(NH_3)_{tot}$ increases similarly at site 050, and even more at site 070 after the Glossop STW discharge. The model cannot predict the value of $(NH_3)_{tot}$ at site 050: however, it approaches the value at site 070 using an increased $(NH_3)_{tot}$ C_d (4.5 mg/l) for the 'E E1' reach.

The CBOD decreases further at site 080 although a small load is discharged from the Charlesworth and Low Marple STWs and the Chisworth Brook. The $(NH_3)_{tot}$ decreases more at that site, although a significant load is discharged from the above sources. The model predicts well the CBOD using the C_d and $k_{BODrem,20}$ that were applied for the first reaches of the river. However, it overestimates the $(NH_3)_{tot}$ even when high $k_{nire,20}$ (3 d^{-1}) and zero C_d are used.

The DOD follows a similar pattern of increase between sites 020 and 070 and decrease at site 080. To predict this behaviour increased $k_{reaer,20}$ (14 to 25 d^{-1}) were used for the reaches between nodes E and D. Overall, the model predicts well the DOD at the nine sites along the Etherow River and its tributaries, the CBOD and $(NH_3)_{tot}$ at seven of those sites, and the $(NH_3)_u$ at six of them. The E value is small for the CBOD due to the residual at the 050 site: however, it is higher than 0.99 for the DOD and $(NH_3)_{tot}$.

The calibrated model was not validated mainly due to the absence of flow data for the point-source effluents in another year. Apart from these data, information is needed about the weirs of the river system to describe their impact on the DO variation. The applied water-quality model can be further developed to run in a stochastic framework, and include additional constituents and processes. It can also interact with more complex models of diffuse pollution. Advanced methods of sensitivity analysis and parameter estimation can be similarly applied (Young 1993).

Assessment of water quality

The monitored water quality along the four river systems was classified for the year 1994 using the RE scheme and the monthly measurements of DO, BOD[*ATU*], $(NH_3)_{tot}$, $(NH_3)_u$, and pH that were provided by the Environment Agency (Spanou & Chen 1998, 2000). For most sampling points the overall class was determined by the BOD and/or the $(NH_3)_{tot}$ class. The RE classification was also performed at the key monitoring sites at the downstream ends of the four main rivers for each year over the period 1986–1996 using weekly or fortnight data. The results show that, since 1991 or 1992, the water quality has been improved to *good* (class RE2) at the Etherow, *fair* (RE3) at the Goyt, and *fair* (RE4) at the Tame and upper-Mersey rivers.

The compliance of all STW and trade effluents with their discharge-consent limits was assessed for each year between 1990 and 1996. All STW effluents have to satisfy 95-percentile BOD[*ATU*] and SS limits. Some of them also have to satisfy absolute limits of the above constituents, or absolute along with percentile limits of $(NH_3)_{tot}$. All trade effluents have to comply with absolute BOD[*ATU*], SS, and pH limits, and some of them with temperature or $(NH_3)_{tot}$ limits as well. Seventeen STWs and the Shell Chemicals Ltd. complied with their standards for each year of the analysis. The Chapel en le Frith and Duckinfield STWs, and the other eight trade effluents violated one or more of their limits for one or more years of the study.

Management of river water quality

Monte Carlo simulations of the mass balance were performed for the discharge points of the STW and trade effluents, considering the BOD and $(NH_3)_{tot}$ constituents which, as was noted above, determine the RE class at most river sites.

In all simulations, the functionally independent variables were assumed to follow the multivariate lognormal distribution, the correlation coefficient between the upstream river flow $Q_{r,us}$ and the discharge flow Q_w was set equal to 0.6, and the distribution of the functionally dependent variable was considered to be defined after 500 shots. The above are in agreement with the assumptions adopted by the Environment Agency during routine discharge-consent calculations (National Rivers Authority 1995).

In addition, the $Q_{r,us}$ variable was described through the mean and the 5 percentile flow upstream of each discharge point. The 5 percentile flow was set equal to the residual $Q95(1)$ flow, and was calculated using the coefficient of 0.0023 $m^3/(s\ km^2)$ and licensed data of the artificial sources. The residual mean flow was calculated using the coefficient of 0.02301 $m^3/(s\ km^2)$, and assuming that the artificially influenced flow is the same for both mean and low-flow conditions.

The Q_w distribution was described by the mean and standard deviation. Due to the absence of detailed Q_w data, it was assumed that the mean flow is 25% higher than the design DWF, and the standard deviation is equal to 1/3 of the mean flow.

The remaining variables were described in different ways, depending on our knowledge about the effluent and river water quality. The procedure that was followed and the results obtained will be presented in more detail for the discharge points of 7 STW effluents with monitored upstream river water quality.

Initially, the $C_{r,us}$ and C_w for these points were described through the mean and standard deviation of the corresponding actual data for the years 1992–1994. Monte Carlo simulations were then performed to derive the $C_{r,ds}$ distribution. The coefficient of variation (c.v.) and the 90 percentile of the $C_{r,ds}$ deviates were also estimated. The $C_{r,ds}$ was subsequently described through the estimated c.v. and the *upper boundary* of the target RE class. Monte Carlo simulations were finally performed to provide the C_w limits that would allow the target $C_{r,ds}$ to be met.

The obtained results show that all effluent discharges deteriorate the water quality of the receiving rivers. In particular, the effluents from the Denton, Stockport, and Ashton under Lyne STWs result in a transition from a RE class of better water quality upstream of the discharge to a RE class of poorer water quality downstream of the discharge. For most discharge points, the $C_{r,ds}$ that were estimated using the actual effluent C_w are in agreement with the short or long-term *River Quality Objectives*, RQOs, which have been proposed by the Environment Agency in terms of the RE scheme. Similarly, the present discharge consents of most effluents allow these RQOs to be met. However, a review of the downstream RQOs and/or of the effluent limits should be considered for the Mossley and Glossop STWs.

Further research is required in order to check the validity of the adopted assumptions for the specific catchment, and perform the calculations with a stochastic catchment model.

CONCLUSIONS

An object-oriented framework for the management of river water quality has been extended to provide tools for (i) the flow duration and low-flow frequency analysis at gauged river sites using different parametric and non-parametric methods, (ii) the sensitivity analysis of a water-quality simulation model with reach-dependent kinetic parameters and diffuse load applying the parameter perturbation method, (iii) the estimation of the model parameters using the Levenberg–Marquardt and the equal-interval search optimisation methods, and (iv) the calculation of water-quality limits for the point-source effluents applying Monte Carlo simulations of the mass balance.

The framework has been employed in the study of the Upper Mersey river system in the UK. The low-flow estimates at the gauged sites are in agreement with the literature findings. Negative residual flows have been estimated for nine abstraction points, suggesting the review of the relevant licenses. The water-quality simulation model predicts well the measured values for the year 1994. Most STW effluents comply with their discharge-consent limits, while most trade effluents violate them. In many cases the present consent limits allow the Environment Agency RQOs to be met.

The object-oriented approach that has been followed for the development of the software has facilitated its extension, and supported the integration of several methods/models for each one of the tasks of the study. It also resulted in a user-friendly environment that applies efficient data management. Finally, it enhances the potential of reusing the object model and the code within a software system for integrated river-basin management. Future directions of work include the linking of the framework with external databases and GIS, the application of advanced methods of model identification, and the stochastic, dynamic simulation of river water quality.

ACKNOWLEDGEMENTS

Financial support from the School of Engineering, University of Manchester through the award of a Simon Scholarship is gratefully acknowledged. The authors would like to thank the Environment Agency North West Region and North West Water Ltd. for the kind provision of data related to the Upper Mersey catchment. They are also indebted to Mr R. V. Moore, Mrs C. I. Tindall, Mr D. Hill and Mr T. J. Marsh of the CEH Wallingford for the kind permission and assistance in using the *WIS* software and the useful discussions on low flows. Finally, they would like to thank the reviewers for their useful suggestions and comments.

REFERENCES

Abbott, M. B. 1999 Introducing hydroinformatics. *J. Hydroinform.* **1**(1), 3–20.

Behrens, J. S. & Loucks, D. P. 1993 Object-oriented river system simulation. In: *Proceedings of the 20th Anniversary Conference on Water Management in the '90s* (ed. K. Hon), pp. 260–263. American Society of Civil Engineers, Seattle, WA.

Booch, G. 1994 *Object-Oriented Analysis and Design with Applications*, 2nd edn. Benjamin/Cummings, California.

Brown, L. C. & Barnwell, T. O. Jr. 1987 *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual*. EPA/600/3-87/007, Environmental Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Athens, GA.

Brown, M. W. H. & Owens, M. D. 1997 Estimation of summer low flows at ungauged sites. In: *Proceedings of the Sixth National Hydrology Symposium, Manchester, UK*, pp. 9.1–9.8. British Hydrological Society, Institute of Hydrology, Wallingford.

Bullock, A. *et al*. 1994 Low flow estimation in artificially influenced catchments. *R&D Note 274*. National Rivers Authority, Bristol.

Bullock, A., Gustard, A. & Sekulin, A. E. 1989 *Low Flow Estimation in the North West Water Region*. North West Water Client Report, Institute of Hydrology.

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

Coad, P., North, D. & Mayfield, M. 1995 *Object Models: Strategies, Patterns, & Applications*. Prentice-Hall, Englewood Cliffs, NJ.

Department of the Environment 1996 *The Mersey Measure, A Guide to the Mersey Basin Campaign's River Water Quality Analysis and Presentation System*. HMSO, 307 (467).

Fedra, K. & Jamieson, D. G. 1996 An object-oriented approach to model integration: a river basin information system example. In: *Proceedings of the HydroGIS'96 Conference* (ed. K. Kovar & H. P. Nachtnebel), pp. 669–676. IAHS Publication, Wallingford, Oxfordshire, IAHS Press, No. 235.

Fishman, G. S. 1996 *Monte Carlo: Concepts, Algorithms, and Applications*. Springer-Verlag, New York.

Fletcher, R. 1980 *Practical Methods of Optimization, Vol. 1: Unconstrained Optimization*. John Wiley and Sons, Chichester.

Graham, I. 1994 *Object-Oriented Methods*, 2nd edn. Addison-Wesley, Cambridge, MA.

Gustard, A., Bullock, A. & Dixon, J. M. 1992 Low flow estimation in the United Kingdom. *IH Report No. 108*. Institute of Hydrology, Wallingford.

Hutt, A. T. F. 1994 *Object Analysis and Design*. John Wiley & Sons, New York.

Institute of Hydrology 1980 *Low Flow Studies*. Institute of Hydrology, Wallingford.

Johnson, N. L. & Kotz, S. 1972 *Distributions in Statistics: Continuous Multivariate Distributions*. John Wiley and Sons, New York.

Johnson, N. L., Kotz, S. & Balakrishnan, N. 1994 *Continuous Univariate Distributions Volume 1*, 2nd edn. John Wiley and Sons, New York.

Jorgensen, S. E. 1986 *Fundamentals of Ecological Modelling*. Elsevier, Amsterdam.

Lewis, D. R., Williams, R. J. & Whitehead, P. G. 1997 Quality simulation along rivers (QUASAR): an application to the Yorkshire Ouse. *Sci. Total Environ.* **194/195**, 399–418.

Marsh, T. J. 1997 River flow measuring station information sheets for the flow stations at the Mersey River system. *National River Flow Archive Data Retrieval Service*. Institute of Hydrology, personal communication.

Marsh, T. J. & Sanderson, F. J. 1997 A review of hydrological conditions throughout the period of the LOIS monitoring programme-considered within the context of the recent UK climatic variability. *Sci Total Environ.* **194/195**, 59–69.

McMahon, T. A. & Arenas, A. D. 1982 Methods of computation of low streamflow. *Studies and Reports in Hydrology, 36*. Unesco, Paris.

Mersey Basin Campaign 1995 *First Periodic Report*. Mersey Basin Campaign, Manchester.

National Rivers Authority 1994a *Discharge Consents and Compliance: The NRA's Approach to Control of Discharges to Water*. Water Quality Series No. 17, National Rivers Authority, Bristol.

National Rivers Authority 1994b *Water Quality Objectives: Procedures used by the National Rivers Authority for the Purpose of the Surface Waters (River Ecosystem) (Classification) Regulations 1994*. National Rivers Authority, Bristol.

National Rivers Authority 1995 *Guidelines for AMP (2) Periodic Review: Water Resources and Effluent Quality*. National Rivers Authority, Bristol.

National Rivers Authority North West Region 1996 *Upper Mersey Catchment Management Plan Consultation Report*. National Rivers Authority North West Region, Warrington.

North West Water Authority 1978 *First Report of Survey of Water Management and Use*. North West Water Authority, Warrington.

Pree, W. 1995 *Design Patterns for Object-Oriented Software Development*, 2nd edn. Addison-Wesley, Wokingham; ACM Press, New York.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. 1986 *Numerical Recipes: The Art of Scientific Community*. Cambridge University Press, Cambridge.

Reichert, P. 1995 Design techniques of a computer program for the identification of processes and the simulation of water quality in aquatic systems. *Environ. Software* **10(3)**, 199–210.

Reitsma, R. F. & Carron, J. C. 1997 Object-oriented simulation and evaluation of river basin operations. *J. Geo. Inf. Decision Analysis* **1**(1), 9–24.

Shane, R. M. *et al*. 1996 Project object. *Civil Engng.* **66**, 61–63.

Spanou, M. 2000 An object-oriented software system for river water-quality management. *PhD thesis*, University of Manchester, Manchester, UK.

Spanou, M. & Chen, D. 1998 An object-oriented framework for computational water-quality management of river systems. In: *Computational Methods in Water Resources XII* (ed. V. N. Burganos, *et al*.), vol. 2, pp. 503–510.

Spanou, M. & Chen, D. 2000 An object-oriented tool for the control of point-source pollution in river systems. *Environ. Modell. Software* **15**, 35–54.

Tindall, C. I. & Moore, R. V. 1997 The rivers database and the overall data management of the Land Ocean Interaction Study Programme. *Sci. Total Environ.* **194/195**, 129–135.

Warn, A. E. & Brew, J. S. 1980 Mass balance. *Water Res.* **14**, 1427–1434.

Whitehead, P. G., Williams, R. J. & Lewis, D. R. 1997 Quality simulation along river systems (QUASAR): model theory and development. *Sci. Total Environ.* **194/195**, 447–456.

Young, P. C. (ed.) 1993 *Concise Encyclopedia of Environmental Systems*. Pergamon Press, Oxford.