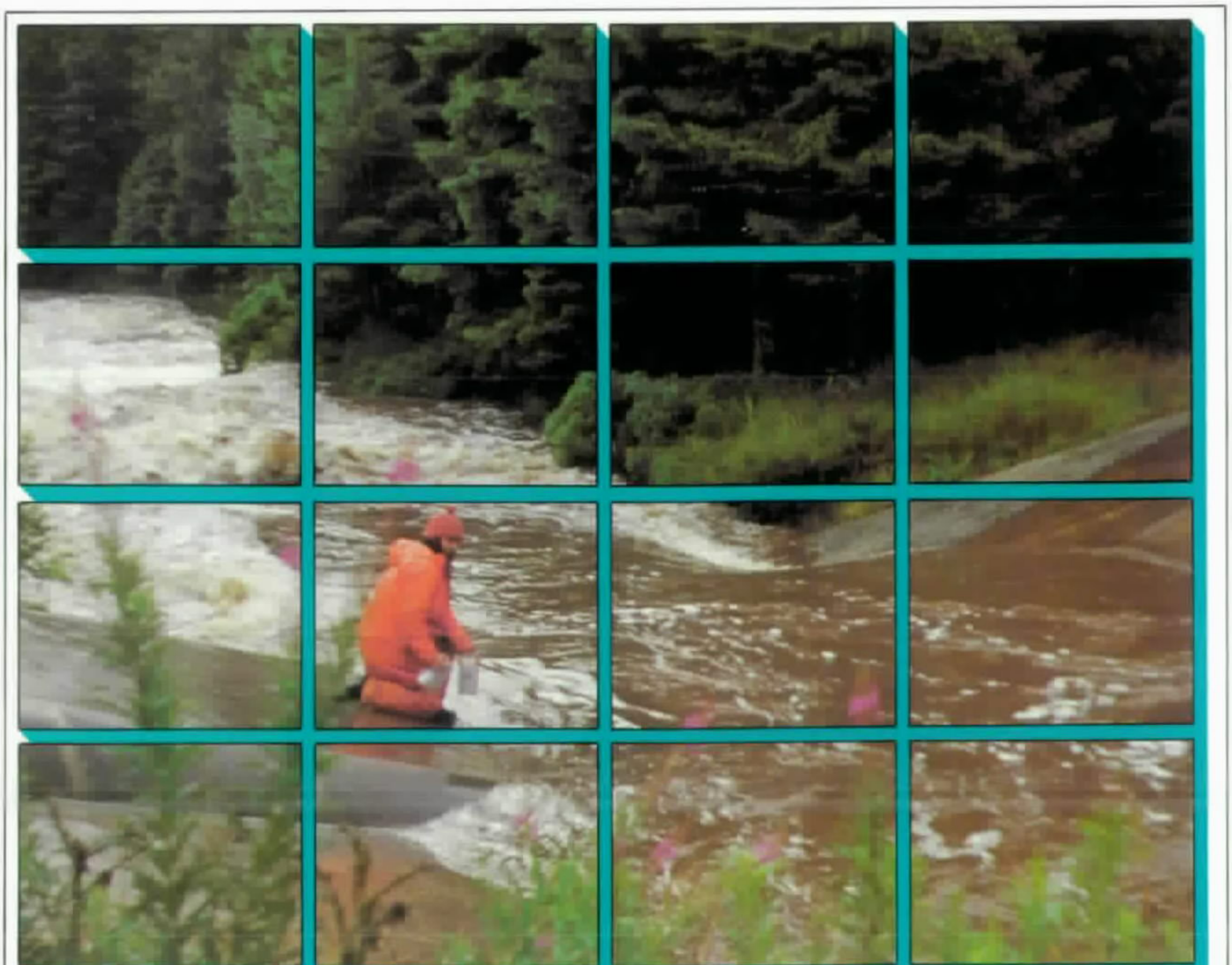
Institute of Hydrology

Report No. 117

Estimating contaminant loads in rivers: a review



Natural Environment Research Council

Estimating contaminant loads in rivers : a review

Ian G. Littlewood

Report No. 117

Institute of Hydrology
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB

ISBN 0 948540 38 9

IH Report No. 117

Published by the Institute of Hydrology

January 1992

A catalogue record for this book is available from the British Library

Executive summary

As the debate on probable and possible anthropogenic impacts on the natural environment becomes more scientific and international, the need has increased for reliable information concerning the magnitude and trends of mass loads carried by rivers through estuaries to the sea. River load information is also required to assist with the management of inland waters, environmental investigations and research into hydrochemical processes. This report reviews river mass load estimation procedures with particular reference to UK conditions.

According to the sampling strategy and calculation method employed, river load estimates can be highly variable in terms of their accuracy and precision. Low frequency sampling at regular intervals, combined with simplistic computational algorithms, can result in heavily biased and imprecise load estimates. Often, significant improvements can be obtained by exploiting 'continuous' flow records in the computational algorithm. At sites with a flashy hydrological response, and for determinands with a high coefficient of variation for concentration (e.g. suspended sediment), there may still be large errors if the sampling frequency is low (e.g. once or twice per month).

Levels of bias and precision in mass load estimates depend on many factors: the hydrological and hydrochemical dynamic responses at the site and for the determinand in question; calculation algorithm; sampling interval; length of estimation period; and measurement errors (systematic and random). No single combination of sampling frequency and calculation algorithm can be devised which will return load estimates of a prescribed quality for all sites and determinands. It may be necessary to consider a 'worst-case' determinand when setting the sampling strategy at a given site; often this will be suspended sediment.

Wherever a 'continuous' record of streamflow is available covering a period of sparse concentration data, a mathematical relationship between flow and concentration should be sought from which concentrations at non-sampled times can be estimated. Periods of relatively frequent sampling should be undertaken to quantify the flow - concentration relationship. The transfer function model, rather than the linear regression model, deserves much more attention in this respect. Other mathematical models, particularly those where the focus of attention is hydrograph separation into components with characteristic concentrations, could lead to improvements in river loads estimation. Wherever funds permit, flow-proportional or probability sampling using automatic sampling devices should be adopted. The report outlines briefly the problems which can arise in calculating river loads of determinands which are sometimes recorded as 'less than' values.

Given the wide range of techniques employed, and the need to make meaningful comparisons (e.g. between sites for a given year and over many years at a given site), more emphasis should be placed on qualifying river load estimates with uncertainties. Drawing from the literature, the report outlines some of the recent progress with theoretical aspects of selected estimation algorithms and discusses the results of some empirical investigations. Usually such case-studies have been concerned with the average annual performance of selected estimators applied to specific determinands (commonly nutrients or suspended sediment). A prototype computational framework known as SMILER (Simulation and Methods Investigation of Load Estimates for

Rivers) is described which allows heuristic assessment of the effect of prescribed combinations of a wider range of variables. As an example, the likely accuracy and precision associated with nitrate loads for the Stour at Langham (Essex) were calculated for a range of estimation periods employing concentration and flow data from the Harmonised Monitoring Scheme database and the Surface Water Archive respectively. Included in this exercise were estimates by the Paris Commission preferred method of Stour nitrate mass loads in 'high' (wet) and 'low' (dry) load years and a comparison of their accuracy and precision for regular sampling intervals ranging from 1 to 30 days.

The ability to derive and manipulate river loads is an important operational need which should be specified at the design stage of database and information systems.

Contents

1 Introduction

The need to use our water resources efficiently, and the growing awareness that some human activities may have undesirable effects on the environment, has led to increased monitoring of the physical and chemical quality of lakes and reservoirs, rivers, estuaries, coastal waters, continental-shelf seas and oceans. Associated with our need to know the state of the aquatic environment, and how it changes with time as a result of natural and anthropogenic influences, there is a requirement for good information about the spatial and temporal variation of constituent loads transported by rivers. River load data are employed increasingly by a wide range of individuals and organisations with interests in environmental protection, from researchers investigating particular hydrochemical processes through to government departments formulating national and international directives and agreements to reduce contaminant inputs to the seas. As the debate on probable and possible anthropogenic impacts on the natural environment becomes more scientific and international, the need has increased for reliable information concerning the magnitude and trends of mass loads carried by rivers through estuaries to the sea. River load information is also required to assist with the management of inland waters, environmental investigations and research into hydrochemical processes.

Recognising the need for better information on river loads, and acknowledging that some published river loads data may be inaccurate or problematical to compare (because of the various combinations of monitoring and calculation methods employed), the Department of the Environment (DOE) commissioned a study:

> To develop standard methods for estimating contaminant loads in rivers and to recommend methods to improve the accuracy of load estimates bearing in mind the associated costs.

The current report is very similar in content to the final report prepared for DOE (Littlewood, 1990); minor changes have been made in response to comments received, and the opportunity has been taken to make editorial improvements. The brief for the study was formulated and accepted in its highly generalised form to allow a non-site- and non-constituent-specific assessment of the problems involved in measuring river loads. This approach enables full use to be made of previous investigations into errors in river loads; usually these have been concerned with particular locations and constituents (determinands). By adopting a more general approach here, current load estimation practices can be assessed and recommendations made related to cost-effective monitoring of river loads for **types** of determinand and site, rather than specific cases. Other important factors which are considered include developments in computer systems and methods of data analysis, and also recent changes in the administrative arrangements in the United Kingdom for river quality control.

Accurate measurement of the load of substances transported by rivers (essentially the product of flow and concentration) presents demanding technical and operational difficulties. Routine load surveillance can be costly because both streamflow and concentration have to be recorded or sampled concurrently at a suitably high frequency, and often samples have to be analysed in the laboratory. Ion-specific electrodes and electronic data recording systems are becoming more widely deployed and can, in conjunction with 'continuous' flow measurement, provide good river load data, but only for those determinands amenable to 'continuous' detection by electrodes and provided that the flow and concentration monitoring sites are co-located.

Whereas the marginal costs associated with an estimate of instantaneous flow at a calibrated gauging station are fairly modest (though the cost of constructing and calibrating the station may be substantial), the marginal cost associated with a corresponding estimate of concentration can be high. Depending on the determinand in question, river water samples may have to be taken manually, then transported to a laboratory and submitted to a labour-intensive process of analysis. Automated laboratory procedures reduce the marginal costs of chemical analyses but the associated initial development and operational costs may be considerable. Although some groups of chemicals (e.g. certain metals) can be measured during the same laboratory procedure, the technical difficulties involved, and the typically large number of chemical species required to be monitored, mean that 'continuous' time series of concentrations are not often available to match the more commonly available 'continuous' time series of flow data.

In practice, truly continuous and exact measurement of flow and concentration is never possible and some level of error in derived river loads is inevitable. Even if streamflow and the concentration of a determinand for a large river (one which does not exhibit rapid changes in flow or concentration) are measured (or sampled) on successive days, and load is computed as the sum of the products of flow and concentration, the load calculated for any specified period will be an **estimate**. The quality of the estimate will depend on the errors in the flow and concentration measurements and on the length of the sampling interval relative to the variation of flow and concentration within the period in question.

Direct volumetric measurement is the most accurate method of measuring **flow** but, unfortunately, this is possible only for low flows on small streams. All operational methods of streamflow measurement (e.g. by hydraulic structure, calibrated natural section, electromagnetic or ultrasonic devices, dilution gauging) introduce error into the flow measurement process. However, with careful design (including selection of site and method) and operation, such indirect methods can usually provide flow data with small errors (Herschy, 1978).

There may be be uncertainty in **concentration** due to chemical changes in the sample during transport to the laboratory or due to errors associated with analysis in the laboratory. The established way of minimising these types of error is a combination of (a) good sampling practices, (b) continual review of the efficacy of methods of analysis and (c) analytical data quality control (AQC) systems agreed between the various measuring authorities. Only by some overall **data quality assurance** programme which includes (a), (b) and (c) can consistency be achieved in a record for a given site and between records at more than one site. An additional source of error, particularly just downstream of a natural confluence or an artificial discharge, is lack of representativeness of a sample. It is important to establish at each site that the concentrations in a sample are representative of average concentrations for the cross-section at which river flow is measured.

Due to the prohibitively high cost of measuring the concentration of certain determinands 'continuously', long unbroken time series of concentration data are rare. (For some interesting determinands the technology to detect the low concentrations typically found in river water has become available only relatively recently.) One method of computing load in such circumstances is to employ a mathematical relationship between flow and concentration to estimate 'continuous' concentration; the dynamic hydrological behaviour of the catchment above the monitoring site can be the dominant factor controlling variations in river chemistry, particularly in rural areas, though the effects of point discharges from farms and sewage treatment works may be superimposed on this behaviour. Load is then calculated (for example) as the sum of hourly streamflow and hourly concentration which, in turn, has been estimated from the relationship between flow and concentration. In such cases an additional error in load is introduced due to the uncertainty in the mathematical relationship, i.e. in the **model** relating concentration and flow.

Relationships between flow and concentration are rarely straightforward, even when it is clear from visual inspection of time series or a scatter plot that there is, indeed, a dependency of concentration on flow. At a given site some determinands may exhibit a general decrease in concentration as flow increases (a dilution effect) whilst other determinands exhibit the opposite behaviour (a purging effect). Superimposed on this broadly identifiable flow-concentration behaviour there may be **hysteresis** such that, for a given value of flow, concentration is systematically higher (or lower) when flow is decreasing, than when flow is increasing. Additionally, there may be evidence during high flows, or over successive periods of high flow, of **exhaustion** of the supply of material (a common behavioural trait of suspended sediment and, therefore, any material adsorbed on suspended sediment).

Factors other than natural river flow can influence the variation in stream concentration of particular determinands. Spillages and discharges from both agricultural and industrial plant (point source pollution) can affect concentration levels in streams and rivers episodically. In predominantly rural catchments the timing and amounts of agricultural applications of natural and manufactured chemicals areally (in addition to point-source discharges, as mentioned above) can affect the level and variation in the concentration of certain chemicals in streams and rivers. Also, a particular determinand may not behave similarly with respect to flow in all catchments. For example, in a mainly rural catchment subject to areal (non-point source) applications of fertilizer, nitrate in the river may exhibit a purging effect so that concentration increases with flow. In contrast, in an urban catchment where a high proportion of low flow comprises sewage effluent, the nitrate content may be diluted at high flows so that nitrate concentration decreases as flow increases. In some catchments, where there are important contributions from both rural and urban areas at different times of the year, nitrate concentration may increase with flow in winter but decrease with flow in summer. Indeed, heterogeneity of land-use, resulting in a 'mixed mechanism' catchment-scale response, is common in the UK.

Recent work in Wales (Littlewood, unpublished; Edwards *et al.*, 1990) and Scotland (Langan, 1987) has shown that the dynamic behaviour of stream chemistry in remote upland catchments can be sensitive to episodes of enhanced atmospheric deposition of natural sea-salts and anthropogenic products from the burning of fossil fuels. The effects of atmospheric deposition in such catchments can be modified by catchment characteristics (including land-use).

The physical and chemical processes (both natural and anthropogenic) which control how the concentration of a particular determinand varies with flow are many, and they interact in an exceedingly complex manner. At the catchment scale, the individual identities of distinct small-scale processes may become blurred to the extent that they cannot be clearly discerned solely from observations of streamflow and stream concentration. The detailed nature of the links and interactions between natural hydrochemical processes, anthropogenic processes and the dynamics of stream chemistry are not yet fully understood, and advances will be made only by continuation and extension of carefully planned and co-ordinated field process studies and mathematical modelling investigations. Such studies and investigations themselves require concurrent time series of flow and concentration from which good load estimates can be made.

A prerequisite to the development of standard methods for estimating river loads is a critical review of existing algorithms and procedures. A comprehensive investigation into errors in load estimates would need to consider all relevant aspects of (a) hydrometry, (b) sampling strategy, (c) chemical or sedimentological analysis in the laboratory (d) computer systems and database management and (e) estimation methods (including the scope for mathematical modelling). The **Analytical Quality Control** (AQC) aspects of water quality monitoring are being dealt with separately by water industry chemists and therefore are not considered in detail here. The main areas explored in this report are sampling strategies, databases and estimation algorithms, and their effects on errors in load estimates for given types of hydrological and hydrochemical dynamic behaviour.

2 Review of river load estimation methods

2.1 INTRODUCTION

The literature on river load estimation methods world-wide is substantial and growing rapidly. This chapter provides a brief review of the methods most commonly employed and is based principally on a selection of key papers and reports. Any emphasis here towards suspended sediment loads reflects the nature of much of the published work on load estimation methods. Although, for reasons discussed later, suspended sediment is perhaps a worst-case determinand in the context of load estimation, the methods which have been employed are generally transferable to other determinands. Except where pertinent, therefore, references to particular determinands have been omitted. The chapter commences with a brief introduction to the problems involved in calculating loads and to the associated statistical terminology.

If both flow and concentration data are available at a sufficiently high frequency (relative to the variation in flow and concentration during the period of estimation), then good load estimates can be calculated. When such data are spaced **regularly** in time the load may be calculated with little error as the sum of the products of flow and concentration, multiplied by the data time interval and a constant to account for the units used, as indicated in (2.1).

$$\text{Load} = K.\Delta t \ . \ \Sigma(C_i . Q_i) \qquad (2.1)$$

where K is a constant
Δt is the data time interval
C_i is the concentration of sample
Q_i is the flow at sample time

If the high-frequency data are spaced **irregularly** in time (allowing, for example, even better definition of flow and concentration during periods of high flows when variability may be greatest), load may be calculated with little error as the sum of the products of individual time intervals, concentration and flow, as indicated in (2.2).

$$\text{Load} = K.\Sigma(\Delta t_i . C_i . Q_i) \qquad (2.2)$$

where Δt_i is the short time interval over which C_i and Q_i are considered to apply

Loads calculated using high-frequency flow and concentration data and (2.1) or (2.2) can be expected to be very close to the true load (assuming the flow and concentration measurements are of high quality). In practice, however, because of the relatively high costs associated with determining concentrations (particularly for some chemical species), high-frequency flow **and** concentration data are rarely available for the same location. Additionally, there will be some error in both flow and concentration measurements.

The increase in environmental monitoring in recent years has been accompanied by many published papers and reports which propose, compare and assess various river load estimation methods. To draw out the main points of interest, and to discuss recent developments, a review of several key papers is given in the following Section. First, however, a brief introduction is given to some of the terms and concepts involved in error analysis, followed by definition of two broad categories of mass load estimation methods.

Accuracy and precision

The efficacy of any estimation method is usually assessed in terms of the **accuracy** (systematic error) and **precision** (random error) associated with the estimate. Consider the general case of taking 100 measurements of a fixed quantity; because of measurement errors the readings cover a range of values but there is one value (or one interval containing a small range of values) which occurs most frequently. The next measurement (the 101st) could lie anywhere in the range defined by the first 100 values (or, exceptionally, it may lie outside this range) but it is most likely to have the value which occurs most frequently in the set of 100 values. The **spread** of the 100 values about the value which occurs most often indicates the precision, or random error, associated with taking that value as a best estimate.

If the value which occurs most often in the first 100 is different to the true (but always unknown) value it would be by an amount known as the systematic error. Accuracy is inversely related to the magnitude of the systematic error; if the systematic error based on the 100 values is small then an estimate taken as the most likely from the 100 is said to be accurate. However, because the true value is never known exactly, accuracy can be difficult to quantify.

Ideally, we should like all 100 measurements to return the true value. Then there would be no imprecision or inaccuracy associated with our estimate. Indeed, in this situation we should need to take only one measurement. However, in all real measurement situations (as distinct from simple counting operations) there will always be some spread, or distribution, of readings about a most likely value, and therefore any future measurement (assuming the same physical situation) will be imprecise to some degree. Similarly, any estimate based on the most likely value taken from a set of replicated measurements will have associated with it some systematic error and therefore that measurement or estimation procedure will be inaccurate to some degree.

The terms accuracy and precision are summarised in Figure 2.1. The total error of an individual measurement comprises components of random and systematic error. Figure 2.2 shows four combinations of accuracy and precision: (a) accurate and precise (the desired quality of an estimate), (b) precise but inaccurate, (c) accurate but imprecise and (d) imprecise and inaccurate. In general terms, high accuracy can only be achieved by adopting good measurement practices; no amount of measurement replication can reduce systematic error. In contrast, measurement replication is beneficial where a fixed quantity is required to be known precisely. Assuming a Normal distribution of the observations, a standard measure of the random error associated with the mean, known as the standard error of the mean S_e, is given by (2.3).

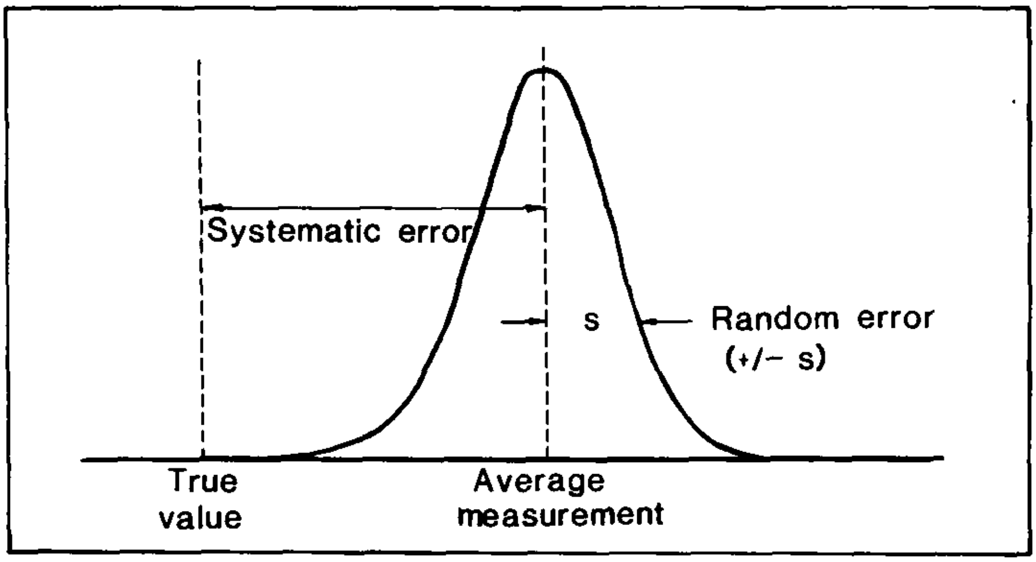


Figure 2.1 Definition sketch for random and systematic error

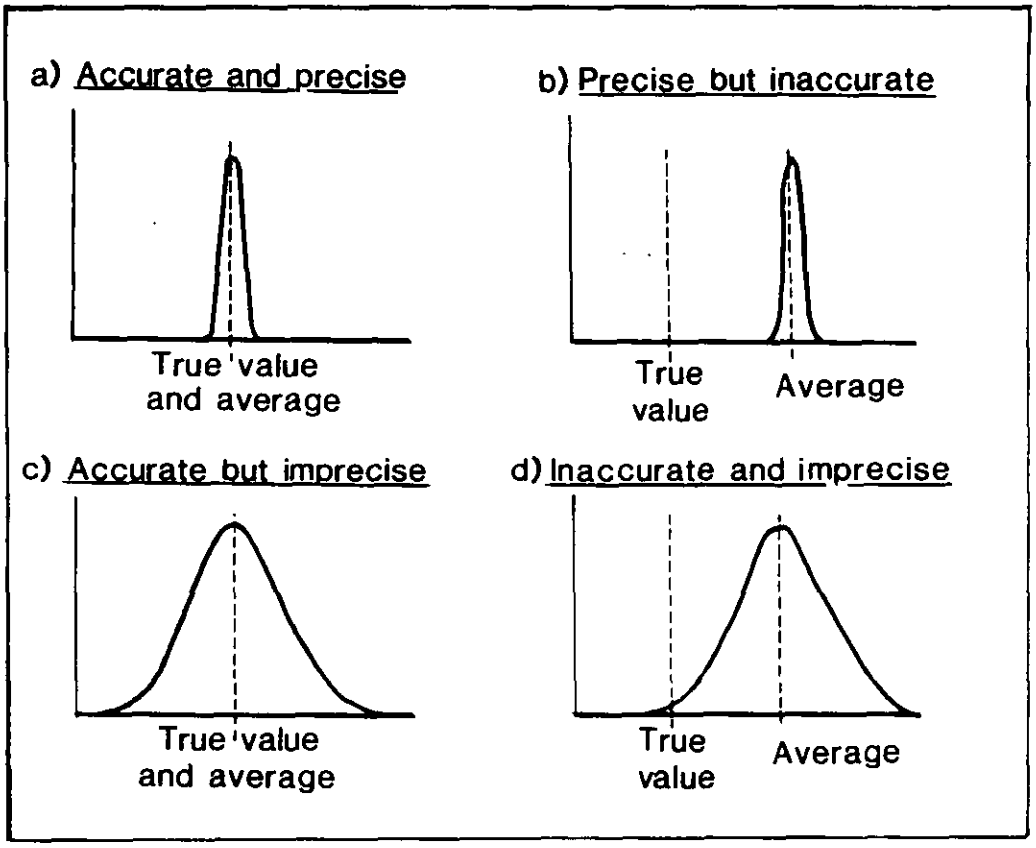


Figure 2.2 Definition sketch for four classes of estimate quality in terms of accuracy and precision

$$S_e = s / n^{0.5} \tag{2.3}$$

where s (the standard deviation of the observations x_i) is given by

$$\left[\sum_{1=i}^{n} (x_i - \bar{x})^2 / (n - 1) \right]^{½}$$

and n is the sample size

As sample size (n) increases the standard deviation remains much the same so S_e decreases; precision in an estimate of a fixed quantity can be improved by increasing the sample size. In the context of river load estimation, where flow and concentration are changing continuously, measurement replication is not possible. However, the principles regarding accuracy and precision outlined above can still be applied to river load estimates, as will become apparent later in the report.

Interpolation and extrapolation methods of load estimation

Whenever concentration data are sparse there are two broad categories of load estimation method. Methods from the first category, the **interpolation** category, are employed when both flow and concentration data are sparse (but available usually at the same times). Depending on whether the data are spaced regularly or irregularly in time, either equation (2.1) or (2.2) is used (or variants thereof, as discussed later). There is no information available with which to estimate how flow and concentration vary between samples (or if such information is available it is not used).

Methods from the second category, the **extrapolation** category, are employed when 'continuous' flow data are available for a period when concentration data are sparse. A relationship between flow and concentration is employed to estimate 'continuous' concentration data between samples and load is subsequently calculated using (2.1) or (2.2). Hence these methods are referred to as extrapolation methods. A pre-requisite of extrapolation methods is, therefore, a relationship between flow and concentration. Clearly, the better the relationship the better the final load estimate.

The Surface Water Archive contains some 26,000 station-years of daily streamflow data for United Kingdom rivers (more than 1000 stations are operational currently), whilst the Harmonised Monitoring Scheme database contains concentration data (rarely more frequent than weekly) for about 250 sites (most of which are co-located with, or in proximity to a Surface Water Archive site). It is evident, therefore, that in principle, and subject to there being useful relationships between flow and concentration, there is considerable scope for extrapolation methods of load estimation using existing United Kingdom records.

Unfortunately, periods of high-definition records (short-interval data) for both flow **and** concentration, from which to derive good relationships describing co-variability, are not common for the sites in the Surface Water Archive and the Harmonised Monitoring Scheme database. Often in the literature, recourse is made in such situations to crude models (e.g. linear regression equations), based on infrequent concentration data, and these exhibit a large scatter. Extrapolation methods do not necessarily lead to good load estimates. There is a pressing need to establish a monitoring and modelling programme to improve flow - concentration models for load estimation by the extrapolation method.

The next two sections review recent independent assessments of interpolation and extrapolation river load estimation methods.

2.2. INTERPOLATION METHODS

Several papers on the topic of errors in river load estimates have been published by Professor D. E. Walling and Dr B. W. Webb of Exeter University; their work has stimulated considerable interest both in the United Kingdom and internationally. The following algorithms, taken from Walling and Webb (1985), give several of the most commonly used interpolation methods of load estimation. There are alternative ways of expressing the algorithms but the forms given by Walling and Webb are used for convenience.

Method 1 $$\text{Load} = K\left[\sum_{i=1}^{n} C_i/n\right]\left[\sum_{i=1}^{n} Q_i/n\right]$$

Method 2 $$\text{Load} = K\sum_{i=1}^{n}\left[C_iQ_i/n\right]$$

Method 3 $$\text{Load} = K\sum_{i=1}^{n}\left[C_i\bar{Q}_p\right]$$

Method 4 $$\text{Load} = K\bar{Q}_r\left[\sum_{i=1}^{n} C_i/n\right]$$

Method 5 $$\text{Load} = \frac{K\sum_{i=1}^{n}\left[C_iQ_i\right]}{\sum_{i=1}^{n} Q_i}\cdot\bar{Q}_r$$

where K = a conversion factor to account for (a) the period of load estimation and (b) units

C_i = sample concentration

Q_i = flow at sample time

$\bar{Q}_r$ = mean flow for period of load estimate (derived from a 'continuous' flow record)

$\bar{Q}_p$ = mean flow over the period between samples (derived from a 'continuous' flow record)

n = number of samples

Walling and Webb (1985) point out that, by definition, load is essentially the product of mean flow and flow-weighted mean concentration and thereby Methods 1 to 5 can be assessed initially according to how well their components might be expected, from simple inspection of the formulae, to approximate to mean flow and flow-weighted mean concentration. Clearly, $\bar{Q}_r$ in **Method 5** should be a good estimate of the mean flow for the period of estimation though there will be some error in $\bar{Q}_r$ even if it is derived from a 'continuous' flow record. The efficacy of Method 5 depends largely, therefore, on the error in the flow-weighted mean concentration calculated (estimated) as the quotient of summed terms over a limited number (i) of samples.

Flow-weighted mean concentration in **Methods 1 and 4** is approximated as the arithmetic mean of sample concentrations. For determinands like suspended sediment, which tend to increase in concentration as flow increases, regular but infrequent sampling will be biased towards periods of relatively low determinand flux, so the arithmetic mean of a limited number (i) of concentrations will tend to underestimate the flow-weighted mean concentration. However, depending on the variation in flow, the mean of a limited number of sample concentrations might overestimate the flow-weighted mean concentration of a determinand which tends to decrease with flow. The degree of under- or overestimation of flow-weighted mean concentration will depend partly on the sampling interval with respect to the variability of concentration and flow during the period of estimation. Therefore, assuming the error in $\bar{Q}_r$ is small, Method 4 will tend to under- or overestimate load according to whether the flow-weighted mean concentration is under- or overestimated.

On inspection of the relevant formula it can be appreciated that there is an additional source of error in **Method 1** due to the approximation of mean flow as the mean of a limited number (i) of flows. A common feature of most streamflow regimes is that high flows occur for a lesser proportion of the time than low flows. Approximations based on a limited number (i) of flows might be expected, therefore, to underestimate mean flow. However, especially over short periods of estimation, the mean of a limited number of flows could overestimate the mean flow if, by chance, most of the individual flows at sample times are high.

Method 2 assumes that the sample concentrations, C_i, and the flows, Q_i, are representative of the concentrations and flows respectively during the time between samples. Method 2 is simply a discretization of the exact mathematical definition of load given by (2.4) and might be expected, therefore, to perform well when the data time interval is small with respect to the rates of change in flow and concentration. When the sampling frequency is low, however, Method 2 may underestimate load (particularly if concentration tends to increase with flow).

$$\text{Load} = \int_{t_1}^{t_2} c(t)\,.\,q(t)\,.\,dt \qquad (2.4)$$

where c(t) and q(t) are continuous functions describing the variation with time, t, of mean concentration and discharge respectively at the stream cross-section.

Method 3 makes the same assumption about the sample concentrations, C_i, as Method 2 but it employs mean flows derived from a 'continuous' flow record for the periods between samples. Because Method 3 uses more information than Method 2 it might be expected to give better load estimates.

A quantitative assessment of interpolation methods

To assess Methods 1 to 5 more rigorously, Walling and Webb (1985) use a two-year duration time series of hourly suspended sediment concentration, and corresponding hourly streamflow, for the Exe at Thorverton (a Harmonized Monitoring Scheme site), from which 'true' load is calculated using (2.1). The performances of estimation Methods 1 to 5 are then assessed against the 'true' load. The suspended sediment data were collected as part of a wider study by the authors and the hourly streamflow data were derived from records supplied by the South West Water Authority (now the South West Region of the National Rivers Authority).

For each of three regular sampling intervals (7, 14 and 28 days) fifty replicate data sets are derived from the two-year hourly record, starting each time from a different hour near the start of the two-year record. Each calculation Method is then applied to each data set to estimate load over the two-year period. For a given Method and sampling interval the difference between the mean of the 50 replicate load estimates and the 'true' load (calculated from the complete record of hourly data) is taken as a measure of accuracy. The dispersion of the 50 replicate estimates in each case is taken to be indicative of precision. Figure 2.3 (reproduced by kind permission of Professor Walling and Dr Webb) shows the results of this exercise.

The relatively small dispersions of the replicated estimates in Fig. 2.3 for **Methods 1 and 4** indicate that these are the most precise of the five Methods tested at each sampling interval. Figure 2.3 shows, however, that, as might be expected, precision (dispersion) worsens as sampling interval increases for both Method 1 and Method 4. Additionally, for both Methods 1 and 4, the difference between the mean of the replicate estimates and the 'true' load (i.e. the bias) increases as sampling interval increases. This increase in bias indicates the extent to which Methods 1 and 4 tend increasingly to underestimate load with increasing sampling interval, i.e. the arithmetic mean of sample concentrations becomes a poorer estimate of flow-weighted concentration as the number of samples decreases.

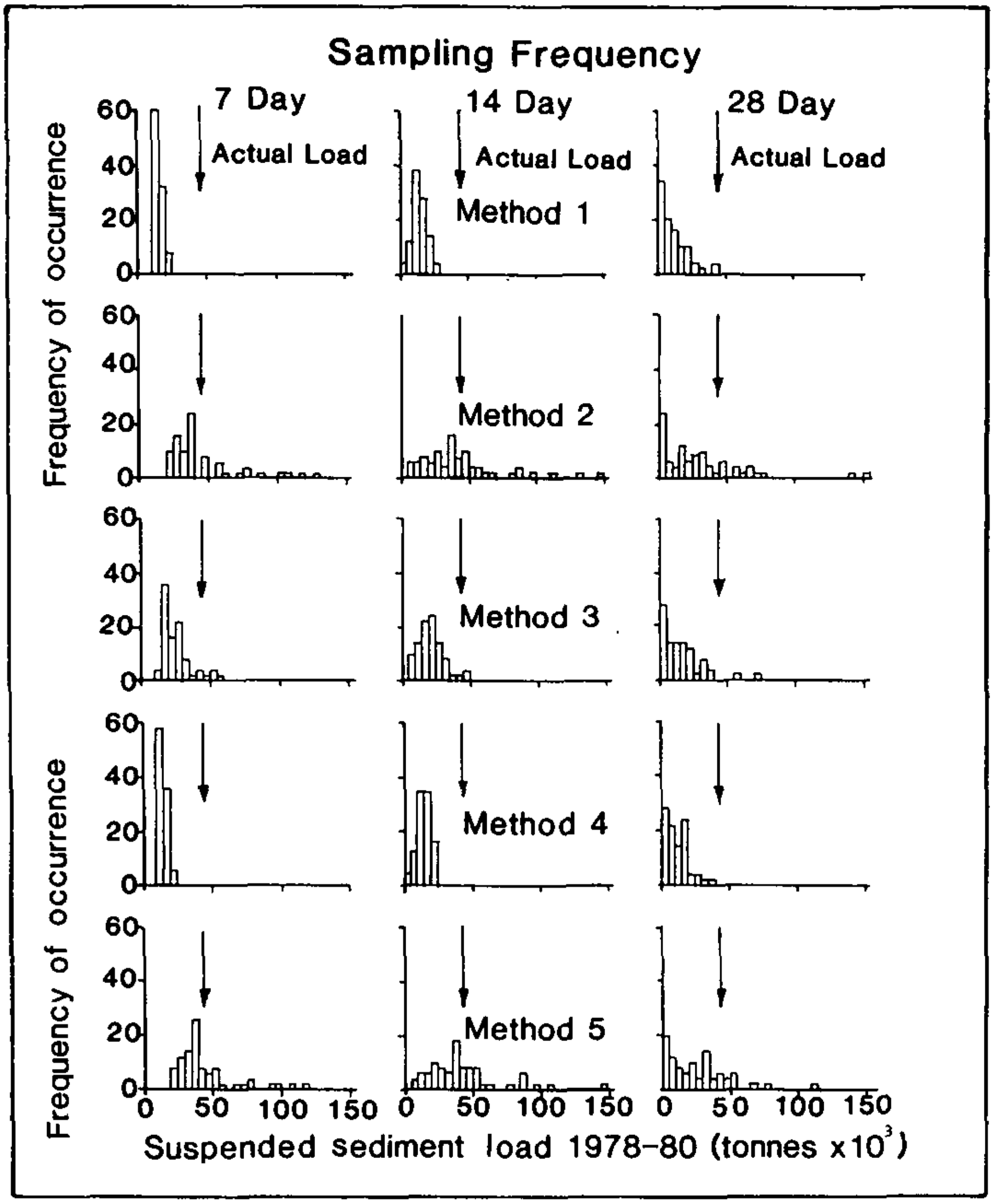


Figure 2.3 Distributions of replicated suspended sediment load estimates for the Exe at Thorverton, 1978-80 - Methods 1 to 5 (from Walling and Webb, 1985)

Since **Method** 1 uses an estimate of mean flow based on a limited number of samples, and error is inversely proportional to the number of samples, it is surprising, perhaps, that the accuracy associated with it at a particular sampling interval does not appear to be significantly worse than the accuracy associated with a corresponding Method 4 estimate (which uses the 'continuous' flow record to estimate mean flow for the period of load estimation). The bias introduced by Methods 1 and 4 for the example of a two-year load of suspended sediment, presented by Walling and Webb (1985), is considerable, ranging from about -75% at a sampling interval of 28 days to -62% for sampling intervals of 7 and 14 days.

It can be seen in Fig. 2.3 that the biases associated with estimation **Methods 2 and 5** are about the same and that they are relatively small. However, the dispersions of replicate estimates for Methods 2 and 5 are larger than for Methods 1 and 4. It appears, therefore, that none of the Methods 1, 2, 4 or 5 gives an estimate which is both relatively accurate **and** relatively precise; Methods 1 and 4 give relatively precise but inaccurate estimates whilst Methods 2 and 5 give relatively accurate but imprecise estimates.

Inspection of the histograms for **Method** 3 in Fig. 2.3 shows that this Method has the best combination of relative accuracy and precision of any of the five Methods assessed. It should be noted, however, that large errors (whether due to inaccuracy, imprecision, or both) are quite likely for any of the Methods investigated with weekly (or longer) sampling intervals to estimate suspended sediment load over a particular two-year period on the Exe at Thorverton. A single load estimate based on weekly samples could underestimate suspended sediment load by 65% (Method 4) or overestimate it by about 200% (Method 2); a single two-year estimate based on 28-day interval samples could be 5%, or more than 250%, of the true value.

The empirical assessments and comparisons made by Walling and Webb (1985) for Methods **1, 2 and 4** have been corroborated recently by a theoretical assessment of the statistical properties of the estimators for long periods of estimation (e.g. for annual loads). Given certain assumptions, including statistical independence between flow and concentration, Clarke (1990) presents an argument based on theory that Method 2 is an unbiased estimator and that Methods 1 and 4 are biased estimators (refer again to Fig. 2.3 for a summary of the empirical results of Walling and Webb). Clarke (1990) gives an equation which, in some circumstances, may be used for correcting the bias in published Method 2 load estimates. He argues also that the variance of estimates by Method 2 is of the order of $1/n$ whilst the variances of estimates by Methods 1 and 4 are smaller, of the order of $1/n^2$ (in broad agreement with the empirical and qualitative results shown in Fig. 2.3). . Equations in terms of the means, standard deviations and correlation coefficient of the logarithms of concentration and flow are given by Clarke (1990) for the variances of load estimates. From such equations it should be possible to say what sample size (approximately) would be required to achieve a given level of precision. Clarke (1990) points out that further theoretical work is being undertaken, to (a) assess the performance of different estimators under conditions of known serial correlation structure in the data and (b) to assess extrapolation methods of load estimation (see the next section in this report).

Method 5 is the preferred algorithm of two methods recommended by the Paris Commission for assessing river inputs of Red List and other substances to the North Sea. Method 2 is the second choice of the Paris Commission, for situations when there is insufficient streamflow information for Method 5 to be employed. At the request of the author of the current report, Professor Walling and Dr Webb estimated suspended sediment load for 1979 for the Exe at Thorverton using **Methods 2 and 5**, using a limited amount of data. Methods 2 and 5 were applied on the basis of 16 samples (typical in a year for the Thorverton Harmonised Monitoring Scheme site – and at many others), for comparison with the 'true' load for 1979 calculated from the hourly data (21984 tonnes). The results are shown as histograms in Figs. 2.4 and 2.5. The dispersion of the replicate load estimates is very large with respect to the 'true' value, and a modal frequency range from which to assess bias is not clearly discernible. It is clear, nonetheless, that both Paris Commission methods (Methods 2 and 5) tend severely to systematically underestimate annual suspended sediment load, probably by more than 50%.

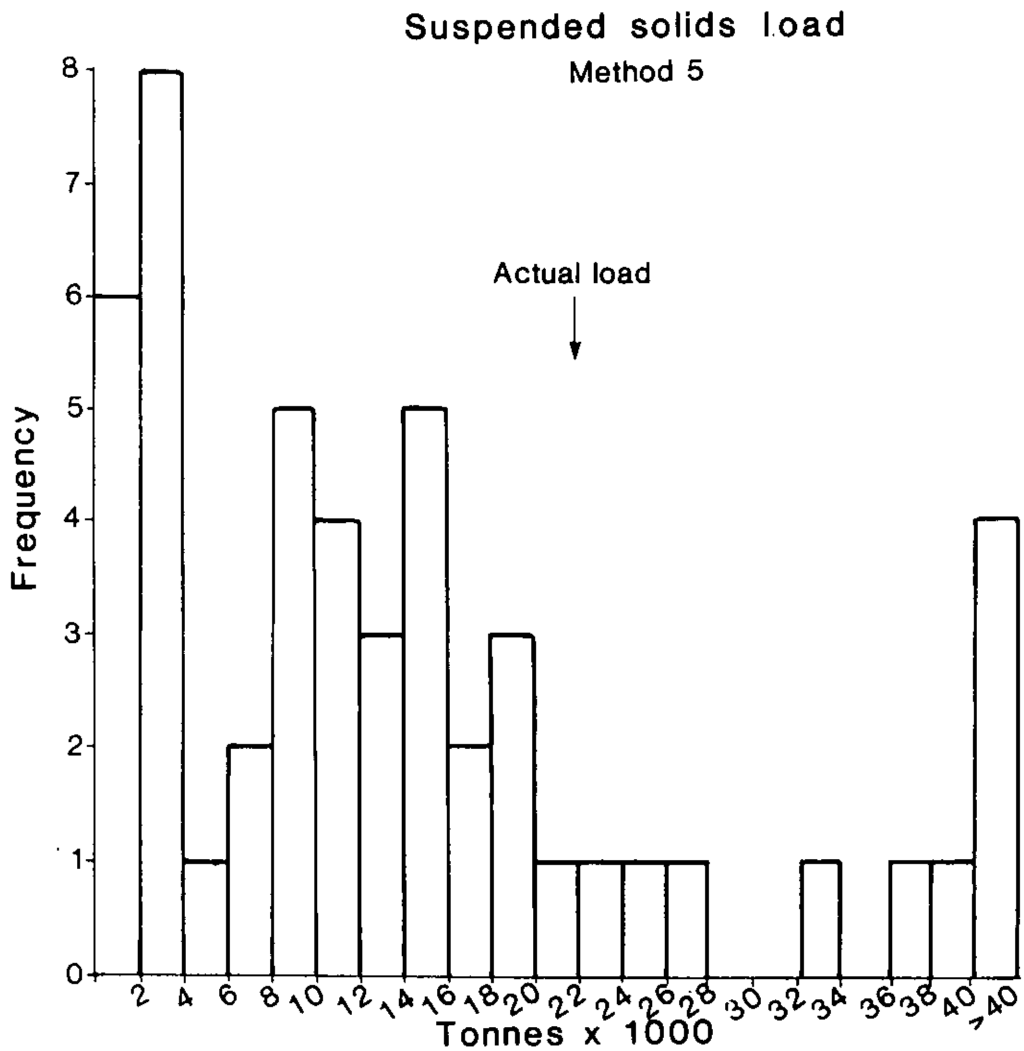


Figure 2.4 Distribution of replicated suspended sediment load estimates for the Exe at Thorverton, 1979 - Method 5

Although the analyses performed by Walling and Webb discussed above are for suspended sediment loads they are, in terms of their basic design, relevant to assessment of the errors in river loads for a wide range of determinands. Errors in suspended sediment loads may, however, give a pessimistic view of the expected quality of estimates for other determinands. Determinands which exhibit a more subdued concentration variation with flow (including negative

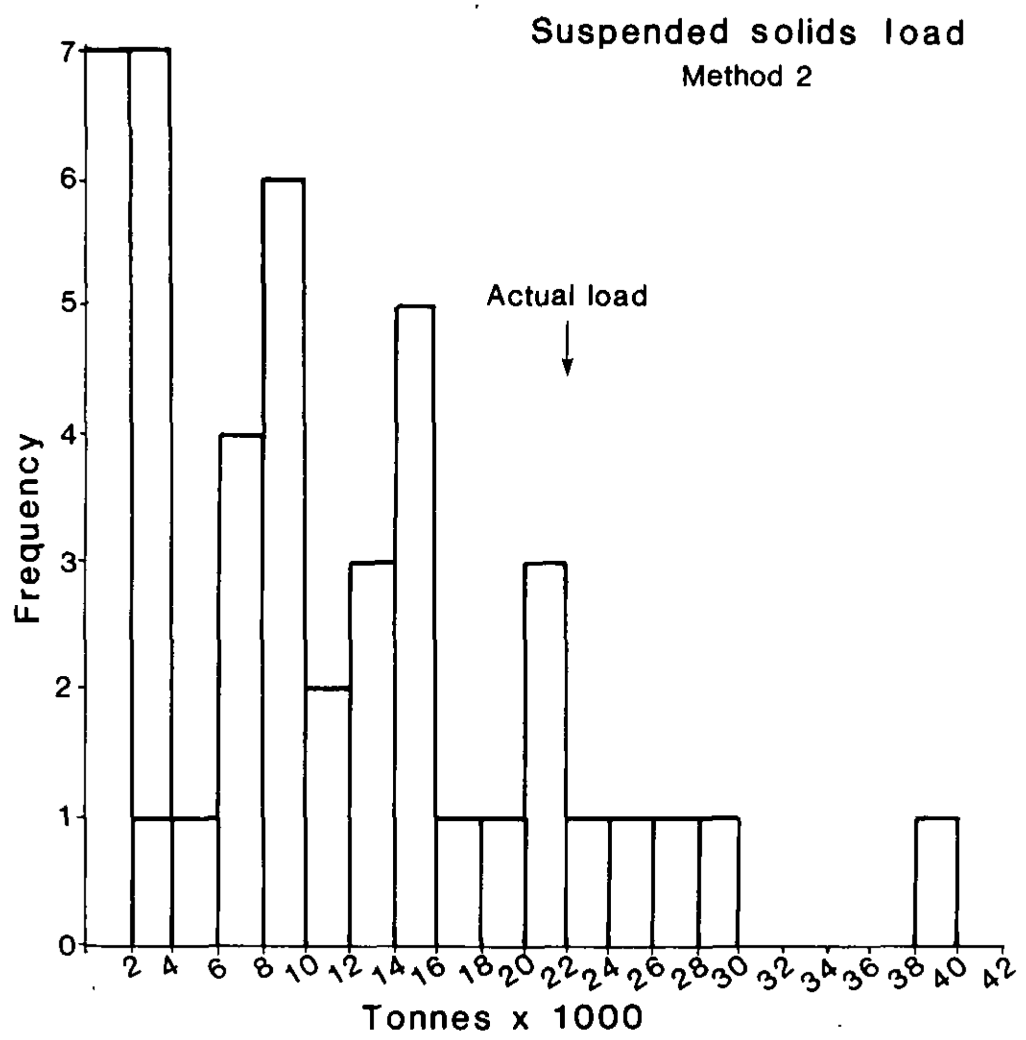


Figure 2.5 Distribution of replicated suspended sediment load estimates for the Exe at Thorverton, 1979 - Method 2

responses) should be estimated with better accuracy and precision by Methods 1 to 5, particularly at sites where there is also a subdued hydrological response. It is worth noting, however, that a significant proportion of the transport of many determinands (e.g. heavy metals, organochlorine residues) can occur during high flows as an adsorbed phase in association with the transport of particulate matter. Suspended sediment is, therefore, an important determinand.

Subject to the vagaries of concentration variation in time induced by episodic discharges to rivers from industry and centres of high population density, many determinands in solution (dissolved substances), and many rivers (especially at their tidal limits), will have relatively subdued responses in their variations of concentration and flow. However, the prerequisite 'continuous' data (over sufficiently long periods) with which to calculate 'true' loads for comparison purposes in such cases are not generally available. The suspended sediment datasets compiled by Walling and Webb for the Exe at Thorverton, and for a few other sites (discussed in the following section), are exceptional in this respect. It is difficult, though, to establish what the errors in load estimates might be for other determinands at other sites which exhibit different hydrological dynamic behaviours. Later in this report, the problem of a general lack of 'continuous' data from which to calculate the 'true' loads of various determinands is circumvented by adopting a technique using synthetic data. In this approach, combinations of typical hydrochemical and hydrological dynamic behaviour are **prescribed** and thus long time series of synthetic flow and concentration data can be generated from which 'true' load can be calculated.

2.3 EXTRAPOLATION METHODS

Extrapolation methods of load estimation employ a mathematical relationship between concentration, which operationally is normally measured infrequently, and some 'independent' variable (or variables), which operationally is (are) measured at high frequency. The most commonly employed single independent variable for these purposes is flow (Q) which is typically used in a simple power law rating curve of the form given by (2.5). At the expense of additional complexity, there is no reason why other available variables and other mathematical structures cannot be employed. However, the simplicity of (2.5) has encouraged its widespread use, sometimes with important procedural conditions which will be discussed.

$$C = K\,(Q + A)^{B} \qquad (2.5)$$

where C is concentration
K and A are constants
B is an exponent

Equations of the form given by (2.5) are usually derived by taking the logarithms of both concentration and flow (the latter adjusted by amount 'A' if necessary) and employing simple linear regression analysis. (A similar procedure could be adopted using several variables and multiple linear regression analysis.) In the context of suspended sediment, Walling (1977) highlights a number of inherent weaknesses of the simple rating curve (2.5) approach. The linear regression model with only one independent variable (flow) cannot take into account dynamic erosion processes in the catchment. One effect of such processes may be that suspended sediment concentrations at a given flow on the rising limb of a hydrograph are systematically, but variably, lower (or higher) than concentrations at the same flow on the falling limb – the response known as **hysteresis**. Another effect which cannot be accommodated by the simple linear regression approach is **exhaustion** of the supply of material during and between runoff events. This causes the concentration - flow relationship to be highly time variant, with concentrations generally becoming lower during events or over a succession of events as the source of material diminishes. Much of the scatter in concentration - flow relationships may be due to such factors, with which the simple linear regression model cannot cope because it is "essentially a univariate expression of a complex multivariate system" (Walling, 1977).

Recognising the deficiencies, identified above, of a single rating curve, many investigators have employed separate rating curves of the form given by (2.5) for (a) increasing and decreasing flows and/or (b) for the different seasons of the year. Such **stratification** of univariate rating curves by stage and time can reduce the scatter about the individual rating curves but it requires careful statistical analysis and substantial errors may remain. An alternative approach is to employ a multiple regression model where additional variables are incorporated to describe, for example, (a) whether flow is increasing or decreasing and (b) seasonality. On the basis that most of the load of many substances is transported during high flows, another variation of the simple rating curve method involves weighting it during calibration by including additional samples in the high-flow range. (This requires samples to be taken more frequently above a threshold flow.)

An additional source of error in the rating curve method, whereby calculated loads may be in error systematically by as much as 50%, has received considerable attention in the recent literature. In the following, a brief outline is given of the debate; careful scrutiny of the papers cited is recommended for a full appreciation of the complexity of the problem.

Ferguson (1986a, 1987) points out that unless a multiplicative correction factor is included in a simple rating curve of the form given by (2.5), a bias (in this case an underestimation) is the inevitable result of re-transforming to the 'arithmetic domain' after performing regression analysis on the logarithms of concentration and flow (adjusted by 'A' if required). Ferguson's analysis of the situation indicates that, for specified conditions, the bias thus introduced is proportional only to the degree of scatter about the logarithmic concentration - flow relationship (i.e. to the variance of the model residuals, S_e^2).

In many applications where an identical curve-fitting procedure is adopted (e.g. derivation of stage - discharge relations) the degree of scatter is usually small; the bias is then typically less than 1% or 2% and often is small enough to be ignored. However, for reasons already discussed, scatter in concentration - flow relationships is typically large, making the bias significant. Ferguson (1986a, 1987) suggests a parametric correction factor (a Normal distribution for the model residuals is assumed). For the case he presents, this reduces the underestimation of annual suspended sediment load from (-)43% (uncorrected) to (-)9% (corrected). Ferguson (1986a) concludes that the simple correction factor $\exp(2.65\ S_e^2)$ "removes most of the bias when log-log rating curves are approximately linear with additive normal scatter, and its use will improve the accuracy of estimates of river load". Ferguson's papers attracted much interest and comment.

Koch and Smillie (1986) apply the correction factor recommended by Ferguson to sediment data from two rivers in northwestern Colorado and find that corrected estimates are too high by 39%. Koch and Smillie also apply a non-parametric correction factor, $^1/n\Sigma\exp(e_i)$, based on the 'smearing estimate' presented by Duan (1983), but find an even greater overestimation. They report, however, that Thomas (1985) found the non-parametric correction to perform better than the parametric correction. On the basis of their results (and those of Thomas), Koch and Smillie (1986) doubt the applicability of a parametric bias correction factor for general use principally because the statistical distribution of residuals is not fixed. They conclude also that the non-parametric correction factor does not behave consistently. In his response, Ferguson (1986b) points out that correction for bias cannot eliminate random error, so the finding of individual overestimates of load after correcting by his parametric method does not invalidate the general conclusion that the un-corrected log-log rating curve method, on average, underestimates by the amount exp(2.65 S_e^2).

Walling and Webb (1988) also question the general effectiveness of Ferguson's parametric correction factor; they conclude that factors other than any bias introduced when re-transforming to (2.5) from the logarithms of the variables are more important in producing inaccurate load estimates. For three rivers in southwest England, Walling and Webb (1988) present annual and period-of-record 'true' loads of suspended sediment, against which they compare the means and dispersions of load estimates (for indications of bias and precision respectively) based on 50 different rating curves replicated by sampling from complete hourly datasets. They construct and replicate two types of rating curve for each river; the first type assumes regular weekly samples and the second assumes additional sampling above a threshold flow. Whilst Ferguson's results appear persuasive that biased load estimates obtained from simple rating curves can be effectively corrected, Walling and Webb's results indicate otherwise. The results for two rivers, where annual loads are computed by Walling and Webb, are discussed here.

As shown in Table 2.1, un-corrected rating curves based on regular weekly sampling lead to massive bias in annual loads of between -97% and -68%. And for rating curves weighted towards high flows, as in Table 2.2, the systematic errors are marginally less, between -95% and -51%. For the Dart at Bickleigh the systematic error in annual loads calculated using either type of rating varies over a remarkably narrow band of between -97% and -88%.

For comparison, Walling and Webb apply both the parametric and non-parametric correction factors but the effects vary from (a) only a slight improvement within the range of negative biases to (b) a positive bias of (+)38%. For the Dart at Bickleigh the parametric correction of estimates based on regular weekly sampling (Table 2.1) reduces the bias slightly, to between -93% and -83%, whilst for the Creedy at Cowley the similarly corrected estimates are still biased between -78% and -20%. The corresponding effect of the parametric correction factor on estimates based on flow-weighted rating curves can be seen in Table 2.2. For the Dart, bias is reduced to between -85% and -58%, whilst for the Creedy it is reduced to between -73% and +25%.

The non-parametric correction due to Duan (1983) appears to perform better than the parametric correction, though only slightly, and large systematic errors remain in the estimates of annual loads. For rating curves based on regular weekly sampling (Table 2.1), the bias in non-parametrically 'corrected' estimates is between -84% and -63% for the Dart and between -74% and -3% for the Creedy. For flow-weighted rating curves (Table 2.2), the bias remaining after non-parametric adjustment is between -75% and -34% for the Dart and between -70% and +38% for the Creedy.

Employing coefficients of variation (not given here) for the sets of 50 replicate estimates of load, Walling and Webb (1988) point out that the level of precision is also affected by the parametric and non-parametric adjustments, but not by consistent amounts between rivers. For example, the authors

Table 2.1 Bias in rating curve suspended sediment load estimates based on regular weekly sampling (original data from Walling and Webb, 1988)

PERIOD	ACTUAL LOAD tonnes	SIMPLE RATING %	PARAMETRIC CORRECTION %	NON-PARAMETRIC CORRECTION %
RIVER DART AT BICKLEIGH				
1975-1985	24499	-96	-90	-79
1975-1976	1072	-97	-92	-83
1976-1977	3779	-96	-91	-80
1977-1978	1872	-96	-90	-77
1978-1979	1476	-93	-83	-63
1979-1980	2475	-96	-90	-77
1980-1981	2684	-96	-91	-80
1981-1982	4451	-97	-92	-82
1982-1983	2672	-96	-91	-80
1983-1984	3046	-97	-93	-84
1984-1985	972	-95	-88	-73
RIVER CREEDY AT COWLEY				
1972-1980	82863	-80	-50	-40
1972-1973	7482	-86	-63	-56
1973-1974	20619	-85	-58	-49
1974-1975	10547	-84	-61	-53
1975-1976	1941	-92	-78	-74
1976-1977	16234	-80	-48	-37
1977-1978	10214	-68	-20	- 3
1978-1979	4717	-76	-40	-27
1979-1980	11109	-79	-47	-36

report that, after parametric adjustment, precision for the Creedy is improved slightly but that it is worsened considerably for the Dart. In some cases, therefore, it appears that, on correction for bias, a gain in accuracy is made at the expense of precision. Walling and Webb (1988) also report an increase in the inter-annual variability of 'corrected' loads and point out that this could cause additional problems in the interpretation of long time series of river loads.

Walling and Webb (1988) conclude that the bias introduced into suspended sediment load estimates due to the nature of the simple rating curve approach is not the major cause of systematic error in loads. The most important factor appears to be variability through time due to dynamic erosion processes. In

Table 2.2 Bias in rating curve suspended sediment load estimates based on regular weekly sampling plus flood period sampling (original data from Walling and Webb, 1988)

PERIOD	ACTUAL LOAD tonnes	SIMPLE RATING %	PARAMETRIC CORRECTION %	NON-PARAMETRIC CORRECTION %
RIVER DART AT BICKLEIGH				
1975-1985	24499	-93	-78	-66
1975-1976	1072	-95	-85	-75
1976-1977	3779	-93	-78	-64
1977-1978	1872	-93	-78	-65
1978-1979	1476	-88	-58	-34
1979-1980	2475	-93	-77	-64
1980-1981	2684	-94	-79	-67
1981-1982	4451	-94	-81	-69
1982-1983	2672	-94	-80	-68
1983-1984	3046	-95	-84	-74
1984-1985	972	-93	-75	-61
RIVER CREEDY AT COWLEY				
1972-1980	82863	-71	-25	-18
1972-1973	7482	-79	-46	-41
1973-1974	20619	-75	-35	-29
1974-1975	10547	-78	-44	-39
1975-1976	1941	-89	-73	-70
1976-1977	16234	-70	-24	-16
1977-1978	10214	-51	+25	+38
1978-1979	4717	-68	-17	- 9
1979-1980	11109	-68	-19	-11

the medium- to long-term, variations in erosion processes result in seasonal differences in the overall concentration - flow response (which might be accommodated to some extent by stratification of rating curves seasonally as already mentioned). In the short-term, the nature of dynamic erosion processes may be manifested as hysteresis, and as exhaustion during and between events. The problem of hysteresis is one that the simple linear regression model, upon which the rating curves are based, is unable to cope; an alternative model which can take hysteresis into account, at least to some extent, is discussed later in this report. The problem of exhaustion requires detailed mathematical modelling and has, therefore, been left to future developmental phases of work begun for the current study.

Walling and Webb (1988) also point out that although **some** improvement in load estimates can be obtained by employing rating curves weighted for high flows, the problems associated with hysteresis and exhaustion seem to preclude the use of any systematic adjustment to give major improvements in the reliability of rating curve estimates. The problem of bias in rating curves highlighted by Ferguson (1986a, 1987) is merely a contributing factor to the overall error in load estimates and the correction factors suggested by Ferguson and Duan may be only partially effective.

However, a more recent paper (Cohn *et al.*, 1989) gives a detailed theoretical analysis of the bias introduced by re-transformation to obtain simple rating curves of the form given by (2.5). This work shows that whereas the parametric correction factor proposed by Ferguson (1986a, 1987) performs satisfactorily in many cases it does not eliminate bias. Under certain conditions, Cohn *et al.*, (1989) show that the simple parametric correction suggested by Ferguson (1986a, 1987) can lead to systematic overestimation of loads, thus providing some theoretical insight into the empirical results of Walling and Webb (1988) presented here in summary in Table 2.2 (there is a positive bias in the parametrically 'corrected' load for the Creedy 1977-1978).

Cohn *et al.* (1989) present an alternative correction procedure which leads to a 'minimum variance unbiased estimator' (MVUE). The mathematics of the MVUE are more complex than the simple parametric correction factor described above and its detail is therefore omitted here. Provided that the log - log model for the concentration - flow relationship is valid (see below), the MVUE gives zero-biased estimates and the associated random error is usually nearly as good as, or better than, the basic rating curve method. (Simple parametric correction often increases the random error component of load estimates.)

The MVUE presented by Cohn *et al.* (1989) will undoubtedly receive further attention in relation to river load estimation methods but it seems clear already that there may be problems in its application to the relatively small and flashy rivers of the United Kingdom. For example, both Cohn *et al.* (1989) and Walling and Webb (1988) recognise fully the inappropriateness of the regression model in the logarithms of concentration and flow for describing the physical situation. Both groups of investigators refer to other work aimed at circumventing this problem, namely the smearing (or non-parametric) estimator (Duan, 1983), as discussed above, and also a probability-based sampling procedure (Thomas, 1985) which gives unbiased load estimates even when a log - log relationship is not strictly valid in physical terms.

Before describing the work of Thomas (1985) in Section 2.5, some further comment is warranted on the non-validity of the regression model in the context of hysteresis. A good candidate for an alternative model which can accommodate hysteresis, but which appears to have been largely neglected by many investigators, is introduced in the following section.

2.4 TRANSFER FUNCTION MODELS

The regression model assumes a unique (or one-to-one) relationship between concentration and flow but usually where 'continuous' records of the variables are available it is evident that this prescribed behaviour does not adequately characterise the physical situation. As discussed already, dynamic erosion processes cause hysteresis and exhaustion in suspended sediment concentration - flow behaviour; similar time-variant, non-linear responses are observed for other determinands (e.g. Edwards, 1973) and the reasons for such behaviour are many. Much of the scatter in a plot of concentration against flow (in the logarithms) may be due to hysteresis, and the residuals from a linear regression equation fitted to the observations will therefore be serially correlated – a condition which itself indicates that the regression model is inappropriate. However, the transfer function type of model (e.g. Box and Jenkins, 1970) can allow for non-uniqueness in a concentration - flow relationship.

There are several ways of expressing a transfer function model. The notation adopted here is consistent with that employed later in the report when describing a computer program for investigating errors in loads. The general form of a transfer function for relating (in this context) concentration and flow is given by (2.6).

$$C_t = \frac{B(z^{-1})}{A(z^{-1})} \cdot Q_{t-b} \tag{2.6}$$

where C_t is concentration at time t
Q_{t-b} is flow at time t-b
b is pure time delay
z^{-1} is the backward shift operator
i.e. $z^{-1}x_t = x_{t-1}$

and $B(z^{-1})$ and $A(z^{-1})$ are polynomials in z^{-1} given by (2.7) and (2.8) respectively

$$B(z^{-1}) = b_o + b_1 z^{-1} + \ldots + b_n z^{-n} \tag{2.7}$$

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_m z^{-m} \tag{2.8}$$

The definition of the general transfer function given above may seem rather formidable to readers unfamiliar with the concepts and notation involved but reference to a specific case can help clarify the techniques. The specific case when m is 1, n is zero and b is zero can be written as (2.9).

$$C_t = \frac{b_o}{1 + a_1 z^{-1}} \cdot Q_t \tag{2.9}$$

which may be re-written as (2.10)

$$C_t = b_o Q_t - a_1 C_{t-1} \tag{2.10}$$

Equation (2.10) states simply that the concentration at time t is given by b_0 times the flow at time t, plus a_1 times the concentration at time t-1 (note that a_1 in (2.9) is always negative, so (2.10) is effectively the addition of two positive terms); hence the serial dependency in concentration is accounted for – something with which the simple regression model is unable to cope.

Models of the form of (2.9)/(2.10) have been employed for investigating the dynamic behaviour of water quality (e.g. Whitehead, 1979; Littlewood, 1987) but, with limited exceptions (e.g. Gurnell and Fenn, 1984), their potential for characterising flow - concentration behaviour for use with extrapolation methods of load estimation appears to have gone largely unrecognised and unexploited.

To appreciate further the advantages of the transfer function model for characterising flow - concentration behaviour, consider the (observed) streamflow and (synthetic) concentration reponses shown in Fig. 2.6 where, clearly, there is a positive relationship between the two variables with hysteresis such that 'concentration' lags behind flow. (This is, in fact, the typical response of hydrogen ion concentration H^+ over several days and successive

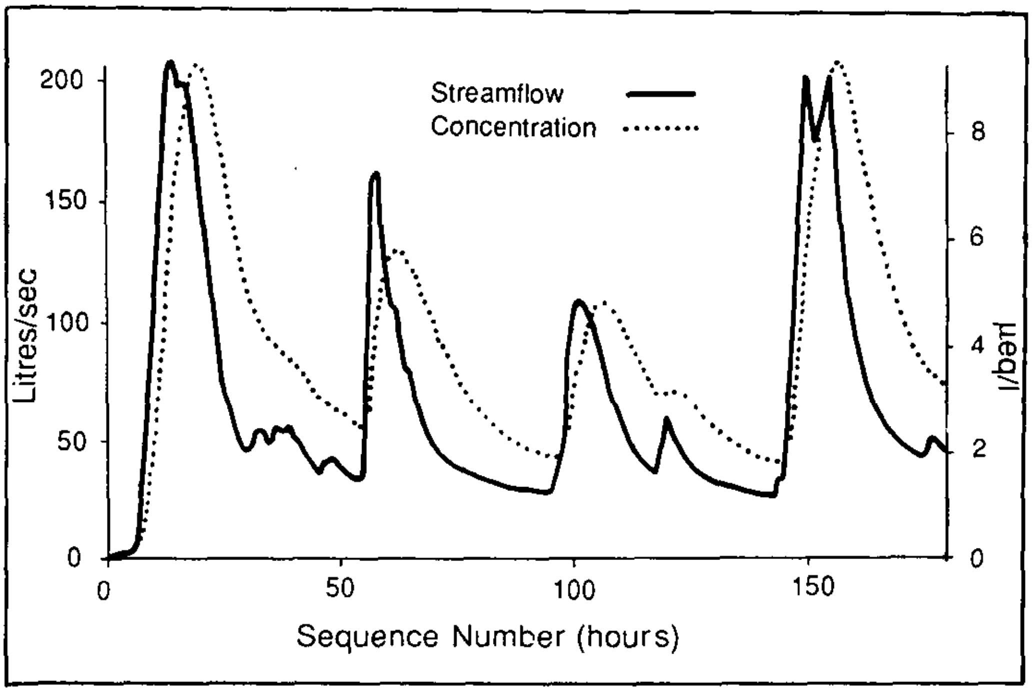


Figure 2.6 Demonstration streamflow and concentration time series

rainfall events for small streams in the Llyn Brianne catchment, Wales; other determinands will exhibit quite different responses to streamflow.) A linear regression model would simply appear as a straight line through the same data presented as a scatter plot in Fig. 2.7, and it would, therefore, be unable to represent the hysteresis.

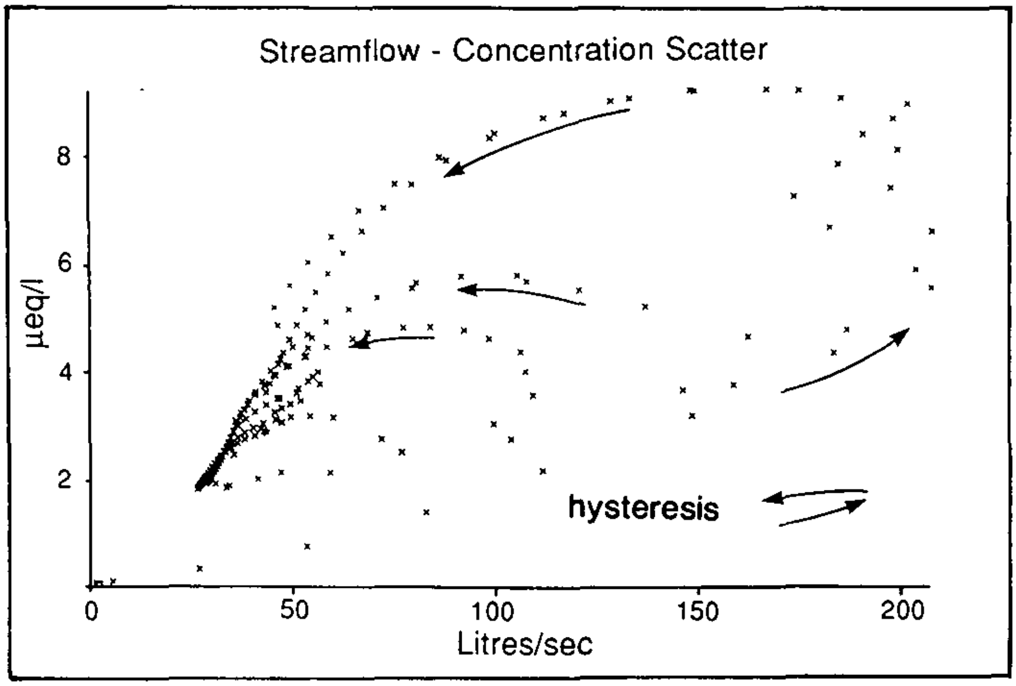


Figure 2.7 Demonstration streamflow and concentration scatter plot exhibiting hysteresis

In fact, the flow and 'concentration' data in Figs. 2.6 and 2.7 are related perfectly by the transfer function model given by (2.11).

$$H^+ = \frac{0.00893}{1 - 0.885z^{-1}} \cdot Q_t \tag{2.11}$$

where H^+ is hydrogen ion concentration in μeq. l^{-1} and flow Q is in $l.s^{-1}$

The transfer function model is therefore a useful way of summarising the dynamic behaviour of stream hydrochemistry (at least for H^+) and in principle, once a model has been calibrated from a short record of high frequency flow and concentration data it can be employed to estimate 'continuous' concentration from flow for periods of sparse or absent concentration data. Making the reasonable assumption that these estimates of concentration will be better than those which would be made from a simple linear regression model (because the dynamic behaviour will have been better characterised by the transfer function) we can expect improved river load estimates. The transfer function approach is, however, unlikely to hold all the answers.

Although the transfer function model has clear advantages over the regression model for flow - concentration relationships it should not be expected to give good results always. For example, a determinand like suspended sediment, which exhibits exhaustion, would require a model to accommodate that behaviour trait in addition to hysteresis (e.g. Moore, 1984); each case is required to be dealt with individually. Furthermore, even in situations where a transfer function model is a significant improvement on a regression model it may be subject to calibration 'drift' over time. It would be necessary, therefore, to periodically review and, if necessary, update such models.

To assist with testing various estimation algorithms over ranges of hydrochemistry dynamic types and sampling intervals, the transfer function model is employed extensively in a Simulation and Methods Investigation of Load Estimates for Rivers (SMILER) described in Chapter 4 of this report.

2.5 FLOW PROPORTIONAL AND PROBABILITY SAMPLING

Most routine sampling of river water quality is undertaken manually and is subject to the operational constraints imposed by daylight hours and the working day/week. Samples tend to be taken at fairly regular intervals (daily, weekly, monthly). Loads calculated using such data, especially by interpolation methods, can be heavily biased because the samples will tend to be taken at relatively low flows. For example, 80% of the suspended sediment load over a period of years may be transported in only 3% of the time (Walling and Webb, 1981), and similar figures can apply to annual and shorter periods of suspended sediment load estimation. The underestimation of supended sediment load can be severe since concentration can vary by orders of magnitude. The situation for other determinands may not be so extreme but Marsh (1980) estimated that 70% of the 1976 nitrate load for the Great Ouse at Bedford was transported during December of that year as a notable drought ended (the temporal distribution of runoff in 1976 was unusual).

Automatic bank-side apparatus permits samples to be taken at variable time intervals so that more information can be made available for load estimation during periods of high flow and high determinand flux. Careful design of such automatic sampling regimes can minimise or even (theoretically) eliminate bias in load estimates. There follows a brief introduction to two methodologies for automatic sampling to meet this objective, the first is a flow proportional sampling technique and the second is a probability sampling method.

Flow proportional sampling

Consider a perfect automatic system which allows continuous sampling of river water at a small but variable rate directly proportional to instantaneous discharge. The flow from the continuous sampler is directed into a 'bulk' sample. By definition, and in the absence of measurement errors, the product of the bulk sample concentration (assuming conservative behaviour of physical and chemical properties in the bulk sample) and the average flow over the period of sampling, will give the true load.

In practice, flow proportional sampling usually requires discrete samples to be taken each time a specified volume of water has flowed past the point of interest. These sub-samples can be mixed into a 'bulk' sample. The concentration of the bulk sample will be a good estimate of the true flow-weighted concentration provided the specified volume of flow for setting the variable sampling is sufficiently small.

Work undertaken for the Department of the Environment by the Water Research Centre (Harrison *et al.*, 1989) compared load estimation for the Thames at Kingston by manual 'routine' weekly grab samples with results from an automatic flow proportional sampling system linked to an ultrasonic gauging station (5 ml samples were taken at a frequency directly proportional to discharge - approximately every 5 to 10 minutes). An equation of the form given for Method 2 was employed to calculate the 'true' loads of Cd, Cr, Cu, Pb, Ni, Zn, SS, TP, SRP and NO_3-N, using the concentrations in the composite samples and the average flows over the periods of composite samples. An extremely valuable dataset was thus created.

Loads were estimated (Harrison *et al.*, 1989) from the grab samples employing Methods 2 and 5 (the Paris Commission recommended methods) for sampling intervals of 1 week, 2 weeks, 1 month and 2 months. Random error (precision) associated with the load estimates was approximated using appropriate formulae. As might be expected on the basis of previous work (summarised earlier in this report), this showed (a) that estimates by Method 2 have a larger random error component than estimates by Method 5 and (b) that for both methods precision (referred to by the authors as 'accuracy') decreases as sampling frequency decreases.

For almost all combinations of (a) determinand, (b) estimation method and (c) sampling interval, the approximate 95% confidence interval includes the 'true' load calculated from the flow proportional sample data. However, this observation on its own does not appear to be sufficient to support the view given by Harrison *et al.* (1989) that estimates based on grab sample data at Kingston exhibit "satisfactory consistency" with the 'true' loads derived from flow proportional sampling. The 'true' loads for chromium and total phosphorous are located near, or just outside, the corresponding 95% confidence intervals. Apart from this observation there appears to be no clear pattern which suggests which determinands are better estimated at a given sampling frequency in terms of either systematic or random error.

The Method 2 result for suspended sediment (monthly grab sampling) overestimates SS load by 37%; such a result may, of course, have arisen by chance but most previous work indicates that SS loads based on discrete data tend to be severely **under**estimated.

Amongst their conclusions, Harrison *et al.* (1989) point out that the error component in load estimates due to a small number of samples being available can be much larger than the error components due (a) analysis in the laboratory, (b) measurement of flow and (c) poor mixing at the measurement cross-section. The apparently good level of mixing at the relatively wide and slow Thames at Kingston is thought to be a fair test of typical conditions near the tidal limits of rivers in the United Kingdom. However, many other rivers exhibit a greater variation in flow, and in this respect the Thames at Kingston could be an undemanding test of the errors which arise from discrete sampling.

Probability sampling

Flow proportional sampling as outlined above can be expected to give good estimates of the load of substances which behave conservatively in a bulk sample. Continuous sampling is the ideal method but variable sampling (as performed for the Thames at Kingston) is the practical option. Provided that the sub-samples are taken at small (variable) time intervals, the estimate can be expected to have small bias and good precision. However, there appears to be no way to quantify either of these measures of estimate quality in particular cases; the work at Kingston adopted the flow-proportional load as the 'true' load against which to compare different estimation methods.

Probability sampling (e.g. Thomas, 1985) is a theoretical framework for deciding when to take individual samples (for separate analysis in the laboratory) such that (a) the final load estimate is unbiased and (b) its precision can be calculated. The larger the number of samples, the better the precision in the load estimate. Alternatively, given a fixed budget for sample analysis, probability sampling maximises the quality of a load estimated from a fixed (limited) number of samples. The samples could be taken manually at the required times but an automatic sampling system, as described by Thomas (1985) and summarised now, can be programmed to do this.

Simple random sampling (SRS) is the basic form of probability sampling, but this means that the probability of a sample being taken is constant in time (irrespective of variations in flow and therefore mass flux) and so it results in biased estimates. Thomas (1985) points out that probability sampling does not require that selection of sample times are equal in probability, only that the (variable) probabilities of sample selection times are known. Clearly, the aim is to relate the probability of a sample being taken to the magnitude of the variable being measured. Thus samples are required to be taken non-randomly in time. Thomas refers to a sampling technique "sampling with probability proportional to size" (PPS) which enables this approach to be applied in situations where the entire finite population of measured units is available for sampling. For example, if the objective is to estimate the total channel storage of sediments in a catchment, all the tributary channels are available for sampling. PPS is a methodology for selecting which tributaries to sample. Selection at List Time (SALT) is introduced by Thomas (1985) (see below) as a special case of PPS for use when only a sub-set of the population units are available - as in river load estimation.

The SALT technique involves an extrapolation method of estimating concentration (from flow) which, when multiplied by flow forms an "auxiliary variable", namely a crude estimate of the material flux at the mid-point of the ith interval. The extrapolation method can be based on a linear regression equation derived from initial survey data. Overall, some stratification may be beneficial, whereby SALT is applied above some threshold flow and SRS is applied below the threshold.

A preliminary estimate Y' is made of load over the period of interest. The value Y' is multiplied by a factor W to obtain Y^* which is (almost) certainly greater than the load which will be experienced in the period to be monitored. Random numbers are then selected from a uniform distribution between zero and Y* and arranged on a sampling interval axis as shown in the lower half of Fig. 2.8.

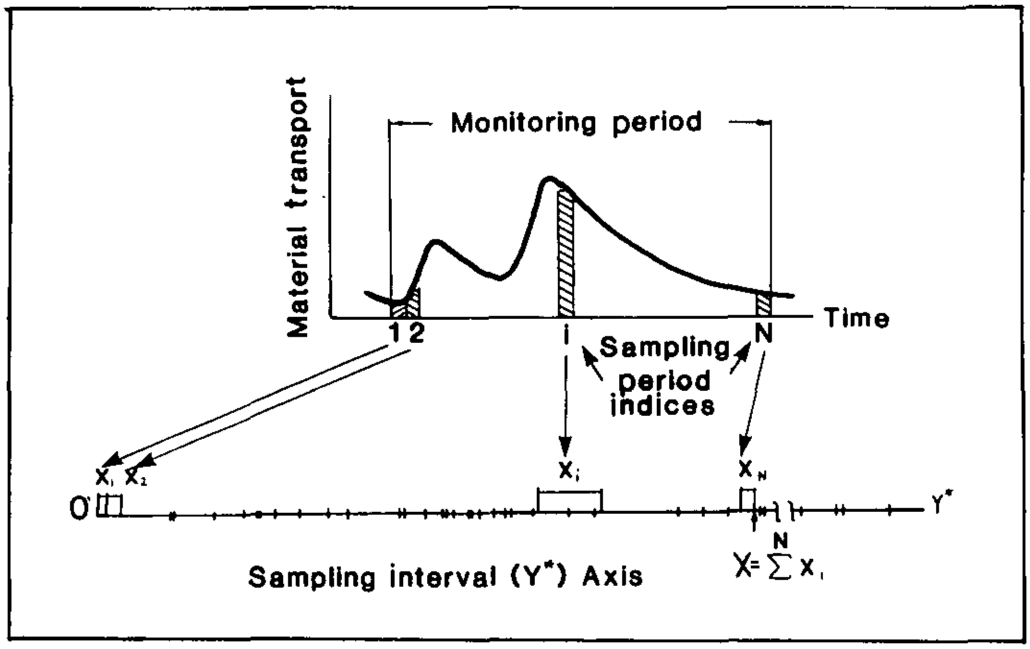


Figure 2.8 Selection at List Time (SALT) designation of sampling times (redrawn from Thomas, 1985)

The auxiliary variable is recorded in real time as shown in the upper half of Fig. 2.8. At each of the N equal intervals on the time axis of the auxiliary variable graph the decision whether, or not, to take a sample is made as follows (shown schematically in Fig. 2.8). The cumulative sum of the auxiliary variable at each of the N time intervals is located on the sampling interval axis, and if there is one (or more) random number(s) in the interval on the sampling interval axis corresponding to the *i*th time interval then a sample is taken. The quality of the flow - concentration relationship employed to derive the auxiliary variable does not affect the accuracy of SALT load estimates but it does affect their precision; the better the flow - concentration relation, the better the precision in the final load estimate. The precision of SALT load estimates depends also on the number n* of random numbers on the sampling interval axis in Fig. 2.8 and Thomas presents a procedure for determining n* for specified levels of precision or, conversely for determining the precision given a value of n*.

Employing a long synthetic time series, Thomas (1985) shows that whereas the flow-duration-curve / sediment-rating-curve method (e.g. Walling 1977) systematically underestimates annual suspended sediment load by significant amounts, a stratified application of SALT underestimates load by less than 1%. In other words, SALT gives very accurate load estimates. He also points out that whereas the precision of an individual estimate cannot be derived for the sediment rating curve method, SALT does permit precisions of individual estimates to be ascertained. The whole process can be controlled by a battery-powered computer.

For a full appreciation of SALT, and the benefits which accrue from it, direct reference to the work of R. B. Thomas is recommended (Thomas, 1983; 1985; 1986; 1988a; 1988b; 1989).

2.6 REAL-TIME UPDATED STRATIFIED SAMPLING

For estimating annual loads on fairly large rivers in North America, Burn (1990) proposes an alternative approach for deciding in real-time when a sample should be taken. Real-time Updated Stratified Sampling (RUSS) involves periodic adjustment during the current water-year of (a) threshold flows defining strata and (b) sampling frequencies within those strata, according to the cumulative flow as the year progresses. The motivation for this approach is that operational and resource constraints often mean that a fixed number of samples is allocated for any year (SALT, for example, returns a variable number of samples). The objective of RUSS is the efficient use of the allocated number of samples.

The detail of the RUSS strategy is rather complex but has been summarised by Burn (1990) as follows.

1 *Measure and classify the flow value as to stratum according to the current strata boundaries.*

2 *Compare the number of time periods since a sample has been taken in the stratum with the current sampling period for the stratum and, if required, take a sample.*

3 *Update the flow threshold value(s) if appropriate.*

4 *Update the sampling period based on the current flow threshold value(s) and the expected number of flow values yet to occur in the stratum.*

5 *Update the number of time periods since a sample has been taken in accordance with the outcome of step 2.*

6 *Repeat steps 1-5 for all time periods.*

Adjustment of the threshold flow between adjacent strata (more than two strata are permissible) is based on the non-exceedance probability of cumulative flow to day j, CF_j, in the current water-year and the correlation in historic data between the values of CF_j and corresponding threshold flows, QT^{α} (α% non-exceedence). In the current year an estimate of the non-exceedence probability for the threshold flow, F_j, is estimated recursively from:

$$F_j = F_{j-1} (1 - w_j) + w_j \pi_j$$

$$F_0 = 0.5$$

where w_j is a weighting factor defining the degree of updating, and π_j is the non-exceedence probability of cumulative flow, both derived from historic data. The weighting factor, w_j, can be the correlation between CF_j and QT^α (zero for no updating, unity for complete updating). The threshold update is the value of QT^α when F_j is substituted for the non-exceedence probability for QT^α according to the historic data. The longer the historic record the better the parameters in the updating strategy.

Initial allocation of the fractional number of samples (P_i) in each stratum (i) is by

$$P_i = \frac{N_i \; s_i}{\sum_i N_i \; s_i}$$

Where N_i is the total number of samples in a stratum and s_i is the standard deviation of daily load in a stratum. Periodic updating of the sampling period for each stratum (SP_i) is by

$$SP_i = \frac{NE_i}{NRS_i}$$

where, using the currently defined strata, NE_i is the expected number of flow values in the current year remaining in a stratum and NRS_i is the number of remaining samples in a stratum (i.e. the number of samples allocated initially to a stratum which have not been used).

Using synthetic daily 'total phosphorous' and flow data typical for a Canadian river (from which 'true' annual loads for 2000 years were calculated), Burn (1990) compared the average performance of RUSS with two other strategies: Fixed-Frequency Sampling (FFS); and Stratified Fixed-Frequency Sampling (SFFS). As their names imply, FFS comprises sampling at regular intervals irrespective of flow, and SFFS is FFS within constant-width strata. Annual loads were estimated on the basis of 30 samples per year (by RUSS, FFS and SFFS) from the synthetic time series. Method 5 (see Section 2.2) was used to calculate the annual loads.

The mean biases (accuracy) for the 2000 annual load estimates were -0.68%, -0.51% and -0.29% for FFS, SFFS and RUSS repectively. A measure of precision was 5.6%, 5.2% and 4.7% for FFS, SFFS and RUSS respectively. Thus the advantage, on average (i.e. for typical years), of using RUSS rather than FFS or SFFS was demonstrated. Although the benefit appears to be rather small for typical years, RUSS can be expected to out-perform SFFS (and therefore FFS) to a greater extent in unusually wet or dry years, but this remains to quantified. It remains to be seen to what extent the RUSS strategy would be beneficial for load estimation in UK rivers.

2.7 HYDROGRAPH SEPARATION

One of the factors which contributes to hysteresis in flow - concentration relationships is the variable degree of mixing during runoff events of waters from different sources (e.g. rainfall, groundwater, etc.) which may exhibit distinctly different chemical compositions. Whilst physically there is a continuum of sources which contribute to streamflow during runoff events, there is considerable circumstantial evidence to support the conceptual simplification that streamflow comprises a small number of dominant component flows. Practical descriptive hydrology usually propounds, therefore, that total streamflow comprises (a) baseflow or groundwater flow and (b) runoff. Each of these components may be sub-divided conceptually according to the needs of any particular investigation (e.g. runoff comprises interflow through the near-surface layers of a catchment and surface runoff, and so on).

We may surmise, therefore, that for determinands which mix 'conservatively' (i.e. do not react chemically in the catchment, including adsorption and desorption) there would be a better relationship (less scatter and hysteresis) between stream concentration and the **proportion** of baseflow at sample time than between stream concentration and **total** streamflow. However, a problem central to scientific hydrology is that it is not possible to observe during runoff events the baseflow (or any other) component of streamflow directly; measured streamflow is a mixture of flow components which physically cannot be unmixed. Many indirect techniques have been devised, however, for separating hydrographs into component flows. There is considerable activity and debate within the scientific community concerning hydrograph separation techniques, the outcome of which could have benefits for load estimation.

There are three broad categories of methods for separation of hydrographs into components flows;

(a) baseflow separation using only the information in the shape of the hydrograph and pre-conceived notions of how the baseflow component of streamflow varies over successive events,

(b) separation into 'age', or within-catchment-source, components based on mixing models and end-member chemistry and

(c) rainfall - runoff mathematical modelling.

Considerable overlap between these categories may be adopted in a strategy for any particular investigation but they will be considered individually here to help demonstrate the potential usefulness of hydrograph separation in the context of river load estimation.

Separation by hydrograph shape

One of the most widely used simple methods of hydrograph separation (by shape) was devised for assistance with 'low flows' water resource studies (Institute of Hydrology, 1980). Employing the 'geometry' of hydrographs, a simple algorithm identifies the low points on a hydrograph at which it is assumed that (almost) all the flow is from stored sources, and these points are joined together with straight lines to form an envelope enclosing what is loosely referred to as baseflow. The fraction of streamflow over a given period which is baseflow is referred to as the BaseFlow Index (BFI) and this has become a key parameter in studies of low flows at gauged and ungauged sites in the United Kingdom. The BFI statistic has been evaluated and published for about 700 gauged sites in the UK (Institute of Hydrology, 1988). An example of baseflow separation (BFI) is shown in Fig. 2.9.

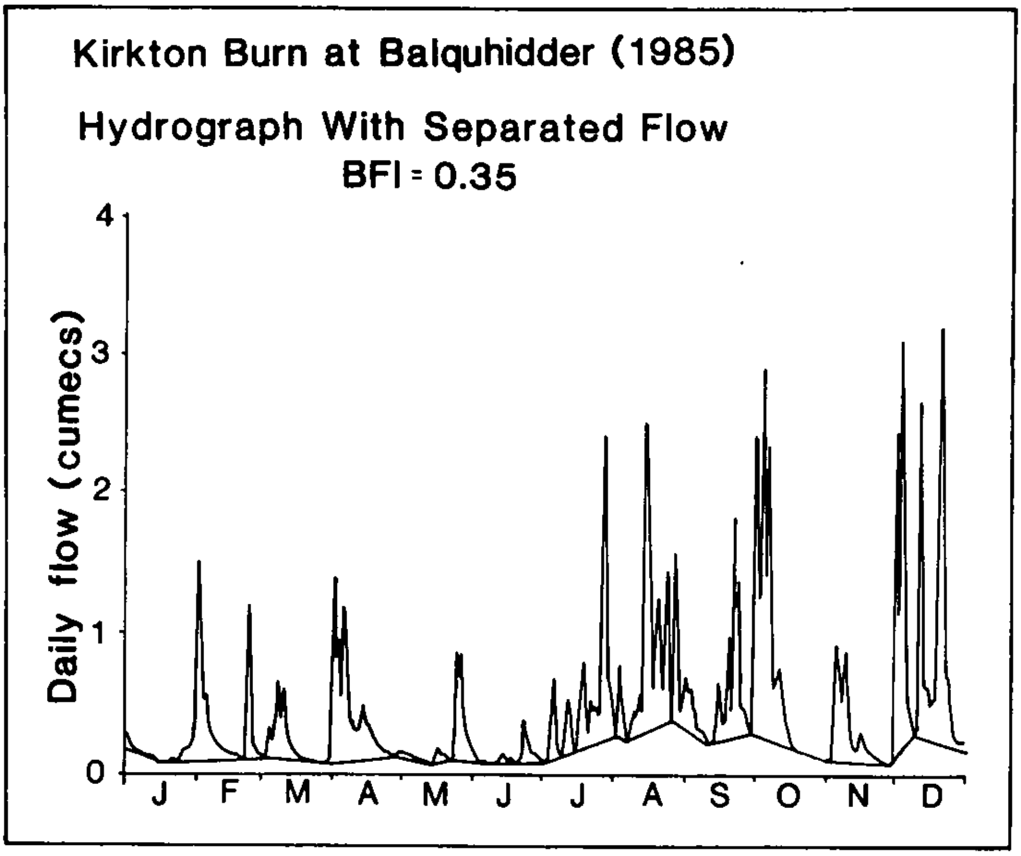


Figure 2.9 Baseflow separation by the Institute of Hydrology (1980) method

The BFI method of hydrograph separation has also found application for regional assessment of water quality. For example, a map of the BFI statistic has been employed to assist with identification of areas in the UK susceptible to surface water acidification (Edmunds and Kinniburgh, 1984). However, since the BFI hydrograph separation method is conceptually so simple (a characteristic strength perhaps for some engineering hydrology and water quality applications which require regionalisation), it would appear to have limited potential for assessing the dynamic proportions of flow components during runoff events. The BFI method may not, therefore, be suitable for designing improved load estimation procedures.

Chemical hydrograph separation

One of the first attempts to separate hydrographs dynamically using chemical signals was by Pinder and Jones (1969) and their method still forms the basis of many hydrological investigations where an understanding of streamflow generation processes is sought. The method, in its simplest form, assumes just two identifiable water sources of spatially and temporally constant (but different) concentrations of some 'conservative' tracer. The variable concentration of the tracer in streamwater is assumed to be the result only of dynamic mixing of flow components from the two sources considered.

Under these ideal conditions the proportion of total streamflow which is flow component 2 at any time is given by (2.12).

$$\frac{Q_2}{Q_s} = \frac{C_s - C_1}{C_2 - C_1} \qquad (2.12)$$

where $Q_s = Q_1 + Q_2$

and Q_s is streamflow
Q_1 is the component of Q_s from source 1
Q_2 is the component of Q_s from source 2
C_s is the concentration of Q_s
C_1 is the concentration of source 1
C_2 is the concentration of source 2.

In practice, the ideal conditions described above which are necessary for the validity of (2.12) do not occur naturally and they cannot be prescribed (e.g. concentration of the tracer in rainfall varies during the event). Some relaxation in the conditions is always necessary and it is not a simple matter to predict what will be the effect of the relaxation on the hydrograph separation.

Field investigations based on this method have given hydrograph separations quite different in character to those given by the BFI methodology (Fig. 2.9). For example, several tracer studies have indicated that peak streamflows comprise large components from a source of water which was in (or on) the catchment before the rainfall event started and not, as might be expected on the basis of the well-known 'contributing area' conceptual rainfall - runoff model, largely rain water which runs off quickly from wet areas (or areas which become wet during the rainfall event).

There are many variations on the basic method outlined above, and several investigations use natural isotopes – arguably the best approximation to a perfect tracer available. In principle, such hydrograph separations could be extremely useful for defining relationships between stream concentration and flow components but more research is required to check the validity of the method under a range of non-ideal conditions.

Rainfall - runoff mathematical models

The range of rainfall - runoff mathematical model types is large and there is a correspondingly large literature on this topic; it would not be appropriate to conduct a review of rainfall - runoff mathematical models in this report. It is sufficient to note here that, in principle, any mathematical model which either explicitly or implicitly routes rain water through elements of storage and via flow pathways to stream channels, and thence to the catchment outlet, can be employed for hydrograph separation.

An example of hydrograph separation using a mathematical model designed for water quality purposes (including load estimation by chemical mass balance) is provided by Birtles (1977, 1978). This model prescribes the broad details of physical within-catchment processes which occur during streamflow generation. It therefore routes rainfall through the catchment explicitly and, in its full form, requires a large amount of input data describing the spatial and temporal variation of rainfall, infiltration, evapotranspiration and aquifer properties. However, a reduced form of the model gives reasonable hydrograph separation into two baseflow components and 'runoff' (interflow plus overland flow) as shown in Fig. 2.10. The detail of the reduced model is still, however, fairly complex with respect to many other rainfall - runoff models.

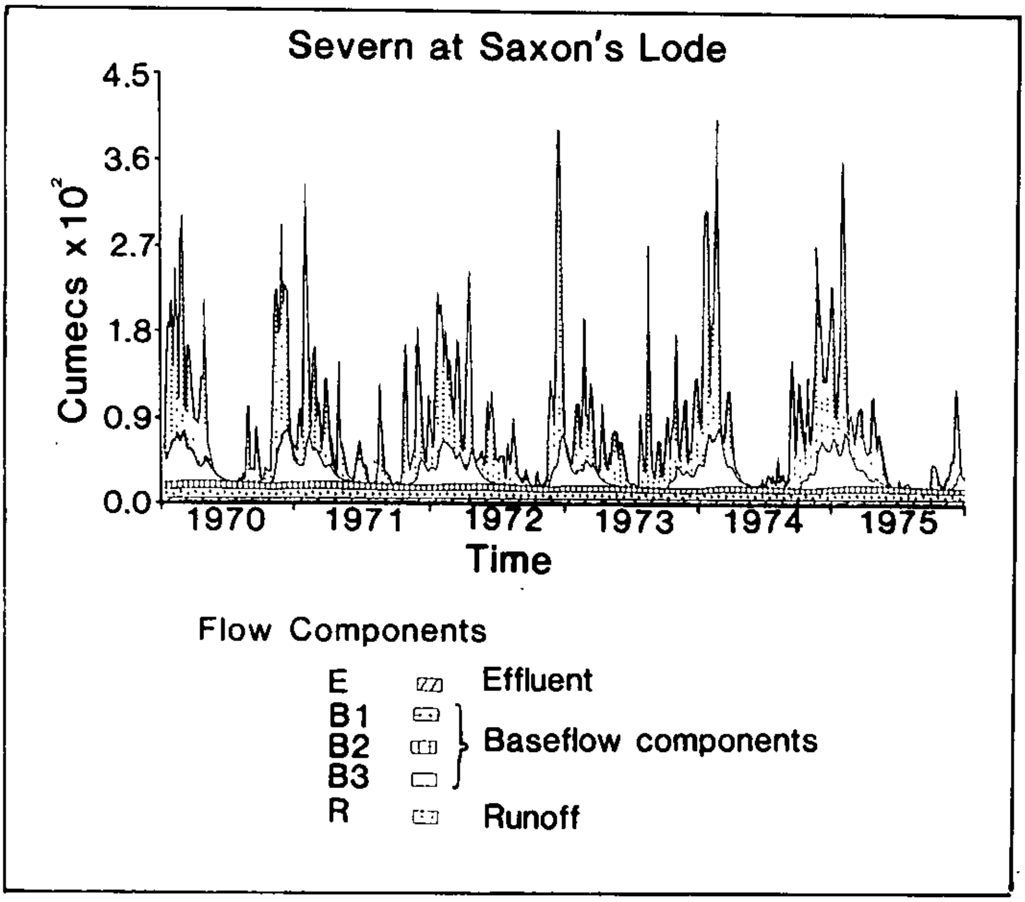


Figure 2.10 Hydrograph separation by mathematical modelling (Redrawn from Birtles, 1978)

An example of hydrograph separation which employs only rainfall and streamflow data (and, where necessary, whatever data are available on evaporation, or some surrogate for evaporation), is provided by the mathematical model known as IHACRES (Jakeman, Littlewood and Whitehead, 1990). This model is based conceptually not on within-catchment physical rainfall - runoff processes but on the theory of Unit Hydrographs (e.g. Chow, 1964) for (a) total streamflow and (b) quick and slow flow components of streamflow. IHACRES, therefore, accounts for rainfall - runoff processes implicitly only but, as the example in Fig. 2.11 shows, leads to reasonable and potentially extremely useful hydrograph separations. (Note from Fig. 2.11 that IHACRES gives a separation which broadly is similar to BFI hydrograph separations but with additional detail of a mixing ratio during runoff events). This model has been developed only recently and further research is required to establish for particular catchments the physical sources and chemical composition of the quick and slow flow components thus identified.

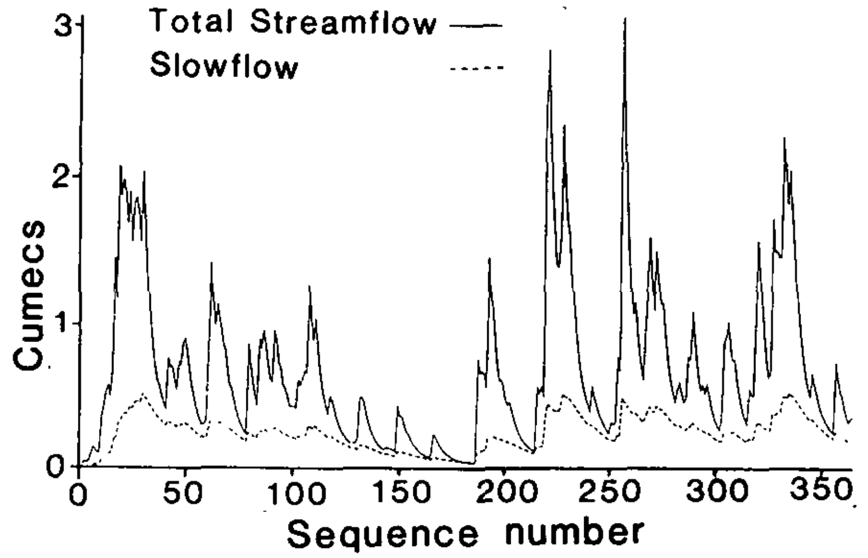


Figure 2.11 Hydrograph separation into quick and slow flow components by the IHACRES method

3 United Kingdom river load surveys

3.1 INTRODUCTION

Since 1970, River Pollution Surveys of England and Wales have been undertaken periodically by the Department of the Environment and the Welsh Office to give a 'snapshot' of the health of rivers. (Less formal national surveys were conducted at intervals during the period 1959 to 1970.) Largely on the basis of BOD, and knowledge of the existence of point discharges of toxic material or suspended matter which could affect the stream bed, river stretches are allocated to Classes (1 = unpolluted to 4 = grossly polluted). This classification provides assistance in identifying problem areas and monitoring trends. Changes in the lengths of river stretches in each class between Surveys indicate the extent to which the overall situation is improving or deteriorating. Such Surveys (similar ones are made for Scotland and Northern Ireland) are undoubtedly useful but it is recognised that the information they present is condensed and, to some extent, subjective (Simpson, 1978). Estimates of the mass loads of substances carried by rivers are not provided by River Pollution Surveys.

The Harmonised Monitoring Scheme was initiated in 1974 to provide river quality information of a complementary but less subjective nature than the information in River Pollution Surveys. Rivers with daily mean flows greater than about 2 cumecs are systematically sampled above the tidal limit or at the confluence with another river. About 250 sites are involved and samples are analysed for a wide range of determinands. One of the major purposes of the Scheme from the outset was to enable assessment of mass loads of materials carried by rivers into estuaries and the sea (Simpson, 1978), thereby fulfilling certain obligations to provide such information to international agencies (e.g. the Paris Commission) – see Section 3.3.

3.2 NITRATES

One of the earliest systematic attempts to survey river loads using the Harmonised Monitoring database was for nitrates (Marsh, 1980). For the four year period 1974 to 1977, between 60 and 200 pairs of flow and concentration data per station were available from the Harmonised Monitoring database. Mean annual nitrate loads for rivers in England and Wales, ranging from 2 to 40 kg N/ha, were estimated by the simple second-choice Paris Commission method defined earlier in the current report. Despite the acknowledged uncertainties involved, these estimates were judged to be of sufficient integrity to enable assessment of regional differences in nitrate loads, and on this basis many interesting observations were made (e.g. the relative importance of point and diffuse sources of nitrate). No attempt was made to assign numerical (e.g. percentage) uncertainties to the load estimates, though several points were discussed concerning the problems involved when using simple calculation methods which are likely to introduce large errors into estimates.

3.3 NORTH SEA INPUTS

Norton (1982) summarised available data to conclude that (a) rivers and atmospheric deposition together are the main input routes to the North Sea for the metals (Cu, Zn, Pb, Cr, Cd, Ni, Hg) and (b) rivers provide the largest input for nutrients, though direct coastal discharges also account for substantial amounts. Data and information for the summary were taken from existing sources (e.g. International Commission for the Exploration of the Sea) and no quantitative estimates of the associated uncertainties were given. Mention is made, however, of problems of comparability which arise due to (a) doubts in some cases about whether concentrations refer to filterable or total pollutant amounts and (b) the different calculation procedures adopted when 'less than' concentrations are encountered (some authorities use zero and some the limit of detection). This latter point is discussed again later.

The Water Research Centre has compiled budgets of toxic material inputs to the North Sea and other coastal waters of the United Kingdom (Hill *et al.*, 1984; O'Donnell and Mance, 1984a, b; Jolly, 1986). Estimation of annual river inputs for these budgets made use of the river flow information available in the Surface Water Archive and employed the first-choice Paris Commission method already described in the current report (flow-weighted mean concentration based on pairs of flow and concentration at sample times multiplied by a mean flow – based, in most cases, on a continuous record of daily mean flow from the Surface Water Archive). Curiously, though, the periods over which the flow-weighted concentration and the mean flow were computed appear to be different. For example, Firth of Forth estuary input estimates presented by Jolly (1986) are based on the product of flow-weighted concentration, derived from concentration and instantaneous flow data for the period February 1984 to August 1985, and a mean annual flow for 1984. The possibility that this practice introduces bias into the load estimates is, however, acknowledged by the author (Jolly, 1986). Nevertheless, the information on toxic inputs from rivers to the North Sea assembled by the Water Research Centre is probably one of the most comprehensive registers of river inputs so far.

Many toxic substances appear in rivers at very low concentrations and a large proportion of measurements may be recorded as 'less than' values. For the five major UK estuaries draining to the North Sea (Thames, Humber, Tees, Tyne and Firth of Forth), Jolly (1986) tabulates estimates of annual inputs of Cd, Cr, Cu, Pb, Hg, Ni, As, Lindane and DDT separately for rivers, sewage and industry. To investigate the effects of values recorded as 'less than', river load estimates are presented for three treatments of these values; replacement with (a) the 'limit of detection', (b) half the 'limit of detection' and (c) zero. As an example of the impact on computed loads, the annual river input of Pb via the Firth of Forth estuary is estimated at 18.3 tonnes or 7.1 tonnes depending on whether 'less than' values are replaced by the 'limit of detection' or zero respectively. Clearly, therefore, the frequency of 'less than' values can have a marked effect on the uncertainty associated with a load estimate but it is a problem with no easy solution. An additional factor of complexity is that the 'limits of detection' can themselves vary for a given determinand at a particular site (due to changes in analytical procedure).

A basic option exists for deriving river load estimates from the Harmonised

Monitoring database, involving simply the summation of the products of concentration and flow at sample time (instantaneous or daily mean - in that order of preference - depending on which type of flow data is present in the database). This is the second-choice Paris Commission method, or Method 2 as described earlier in the report, and it is being employed by Her Majesty's Inspectorate of Pollution for assessing annual inputs of nutrients (NH_3, NO_2, NO_3, TON) to the North Sea (B. G. Goldstone, personal communication). When a 'less than detection limit' value is encountered, a value of half that amount is used in the calculation. For those years when there are no measurements, the annual average load is used.

3.4 HYDROMETRIC AREA GROUP INPUTS AT TIDAL LIMITS

For the period 1975 to 1980, Rodda and Jones (1983) derived estimates of mean annual loads at 163 tidal limit sites grouped by Water Authority (England and Wales) and River Purification Board (Scotland) areas. In common with a previous survey of nitrate loads (Marsh, 1980), loads were estimated principally on the basis of data from the Harmonised Monitoring database and using the simple second-choice Paris Commission algorithm. Joint exploitation of the Harmonised Monitoring database and the Surface Water Archive, to allow use of the superior Method 5 (first-choice Paris Commission method) was precluded at the time by administrative and practical difficulties associated with bringing the necessary data together; the databases were, and remain, managed separately, though with modern database facilities the practical difficulties of jointly exploiting their contents is reduced. Mean annual loads were estimated for eight determinands; Cl, NO_3, SO_4, ortho-PO_4, Zn, Cd, Pb and Cu. No quantitative uncertainties are given for the estimates but many qualifying points are made. For example, the efficacy of regular fortnightly (or thereabouts) sampling is doubtful given that typically 80% of annual load might be transported in only 2% of the time (Rodda and Jones, 1983). Nevertheless, this survey provided much useful information on regional variations in loads and their causes.

3.5 OTHER

Following the third ministerial International Conference on the Protection of the North Sea, held in The Hague (1990), commitments were made, or reaffirmed, to reduce inputs of selected materials to the North Sea via water and air by about 50% by 1995 (based on 1985 levels). Clearly, a need exists therefore to establish baseline (1985) loads and to monitor inputs from year to year to determine whether the remedial measures being taken to improve the situation are effective. The Department of the Environment, the National Rivers Authority and the River Purification Boards are actively formulating and implementing North Sea Action Plans, and for marine inputs are basing their data on the Paris Commission methodology.

At the tenth meeting (June 1988) of the Paris Commission established by the

Convention for the Prevention of Marine Pollution from Land-based Sources, a comprehensive study of riverine inputs to the northeast Atlantic was planned. Measurements for this study are being taken from 1990 and a review by an *ad hoc* Working Group on Input Data is due in the autumn of 1991. Other existing and potential sources of river loads survey data and information are (a) reports of estuary management committees (e.g. The Water Quality of the Humber Estuary, 1987 (Howard and Urquhart, 1988)), (b) annual reports of the International Council for the Protection of the Sea (e.g. Anon., 1984; Bewers and Duinker, 1982), (c) the Joint Group of Experts on the Scientific Aspects of Marine Pollution (e.g. GESAMP, 1987) and (d) other reports prepared for the Department of the Environment (e.g. Grogan, 1984).

As part of the programme of North Sea research being undertaken by the Natural Environment Research Council, freshwater inputs from the major estuaries are required. However, given the practical problems of measuring net seaward flow in tidal areas, recourse has had to be made to adjusting measurements from further upstream - where flow is unidirectional. The Institute of Hydrology has recently supplied the Proudman Oceanographic Laboratory with estimates of daily freshwater inflows from the following estuaries for 1988 and 1989: Firth of Forth, Tyne, Tees, Humber, The Wash and Thames (T. J. Marsh, personal communication).

4 Simulation and Methods Investigation of Load Estimates for Rivers (SMILER)

4.1 INTRODUCTION

A basic problem in assessing load estimation methods, namely that only rarely are there sufficient streamflow and concentration data available from which to calculate 'true' loads, has been discussed earlier in this report. Some previous investigations have, therefore, determined the relationships between errors in loads and other factors (e.g. sampling frequency, estimation method) by employing synthetic data (e.g. Dolan *et al.*, 1981; Richards and Holloway, 1987; Young *et al.*, 1988). However, such investigations have been concerned invariably with specific load estimation periods (e.g. annual) and specific determinands at particular sites of local importance. Although, undoubtedly, the results of such exercises are extremely useful, it is problematical to transfer the information thus gained to other combinations of estimation period, site and determinand (exhibiting probably quite different hydrological and hydrochemical responses).

A requirement was perceived, therefore, for a scheme of synthesising long time series of flow and concentration data such that the dynamic hydrological and hydrochemical behaviour (of particular determinands) can be simulated for the full range of typical conditions likely to be encountered in the United Kingdom. Time series generated by such a scheme can then be sampled at any frequency (fixed or variable), and corresponding loads estimated by any method over various estimation periods can then be compared to 'true' values. In this way it should be possible to make generalisations about the relationships between errors in load estimates and a number of factors (sampling scheme, estimation method, hydrological response, hydrochemical response, estimation period). The program described below, Simulation and Methods Investigation of Load Estimates for Rivers (SMILER), is an initial attempt to provide such a scheme for data generation and includes facilities for comparing the performance of load estimators. Currently, SMILER allows assessment of fixed frequency sampling schemes only.

Concentration data can be generated by SMILER from streamflow data employing a (first order) transfer function model of the form introduced earlier in this report (see section 2.4). In some rare cases it may be possible, as a separate exercise, to calibrate a transfer function model for this purpose from a period of record when both flow and concentration are available at high frequency. Otherwise, the parameters of a transfer function model can be prescribed which, for a given streamflow data input, give an output which approximates to the response for the determinand of interest. In either case the objective is to obtain a time series for concentration which is reasonable with respect to the actual response; it is not necessary to reproduce exactly any complex behaviour due, for example, to exhaustion of the source of material, though the nearer to realistic behaviour the better will be the results. Factors which can be controlled by the current version of SMILER are (a) the coefficient of variation in concentration (b) whether concentration increases or decreases with flow and (c) the degree of hysteresis in the flow - concentration relationship.

Similarly, for applications where flow data are not available, streamflow data can be generated by a (second order) transfer function from an input of effective rainfall data (i.e. rainfall minus evapotranspiration 'losses'). In some cases the effective rainfall data and the parameters of a transfer function derived from rainfall - runoff modelling by IHACRES (Jakeman, Littlewood and Whitehead, 1990), referred to briefly in section 2.6, can be employed. In other cases it might be sufficient, for initial assessment purposes, to treat a time series of observed rainfall as effective rainfall and to simply prescribe transfer function parameters which produce a synthetic hydrograph of desired properties. Under certain circumstances it should be possible to employ the hydrograph separation facility of IHACRES (quick and slow flow) to investigate relationships between stream concentration and the variable mixing of component flows.

4.2 DEMONSTRATION OF *SMILER*

The series of Figures 4.1 to 4.3 illustrates the ability of SMILER to simulate a wide range of hydrological and hydrochemical behaviour types. Figure 4.1 shows the performance of an IHACRES (Jakeman, Littlewood and Whitehead, 1990) rainfall - streamflow model (daily data) calibrated for a 20 km^2 catchment in south-west England over a period of about 3.5 years. There are four parameters in the second-order effective rainfall - streamflow transfer function model. The goodness-of-fit is not of prime importance to the data generation exercise but it can be observed that the rainfall - streamflow model (6 parameters – 4 in the transfer function and 2 in the rainfall - effective rainfall part of the model) reproduces the measured streamflow tolerably well. Other hydrological response **types** can be simulated from the same effective rainfall shown in Fig. 4.1 by varying the parameters in the second-order transfer function.

Alternatively, SMILER can use long time series of observed flow data for a particular site retrieved from the Surface Water Archive, or any other data source.

Figure 4.2 shows the IHACRES hydrograph separation corresponding to the **modelled** flow in Fig 4.1. To estimate the flow components of an **observed** hydrograph, and to allow for differences between modelled and observed streamflow, the variable ratio of modelled component flow and modelled streamflow should be applied to observed streamflow.

Figure 4.3a reproduces the modelled flow data shown in Fig. 4.2 and shows additionally a time series of synthetic concentration generated from that flow data by a first-order transfer function (see equation 2.10). Four parameters control the scale and shape (including the degree of hysteresis) of the concentration response: the two parameters of a first-order transfer function; an initial concentration level (10 mg/l in Fig 4.3a); and a concentration/dilution (+ve/-ve flow - concentration relation) on-off switch. By varying these values a wide range of concentration behaviour types can be simulated. For example, the more pronounced +ve concentration response shown in Fig. 4.3b was

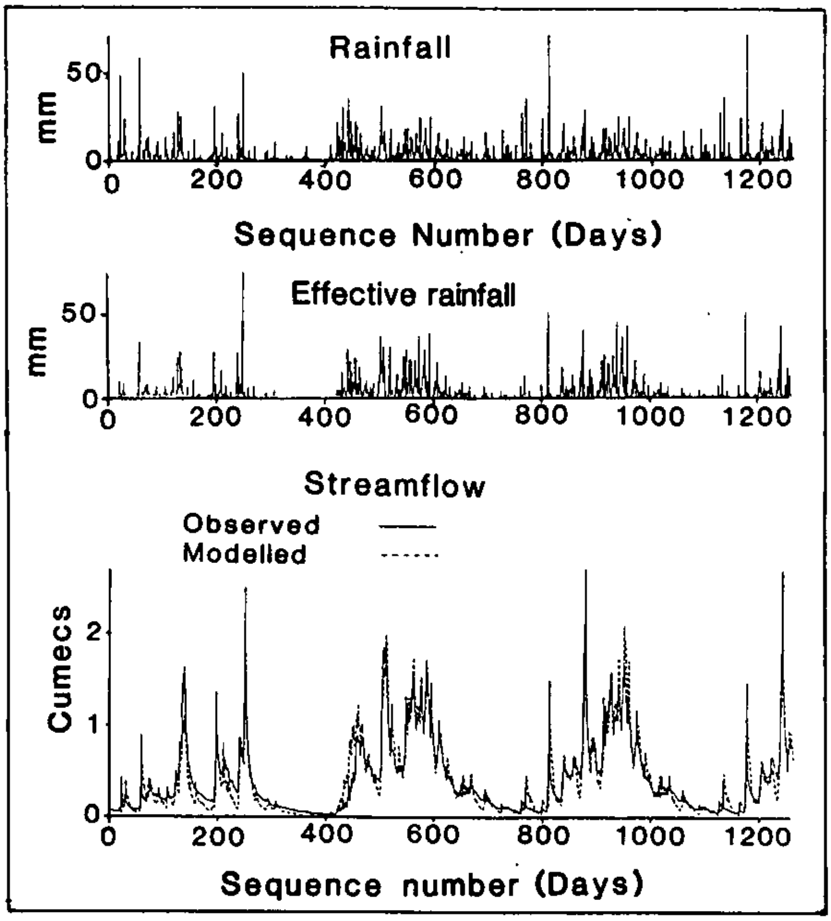


Figure 4.1 IHACRES rainfall - streamflow model-fit

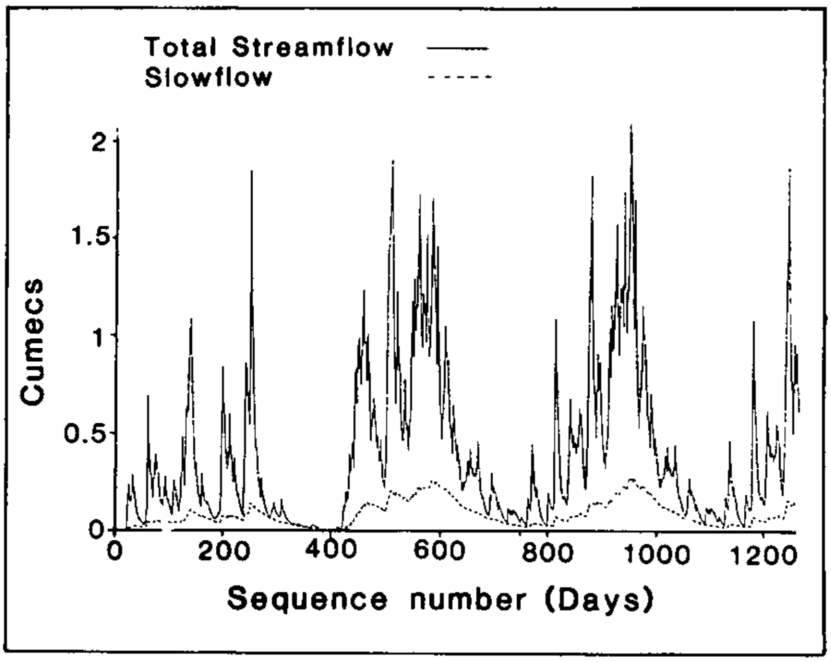


Figure 4.2 IHACRES hydrograph separation

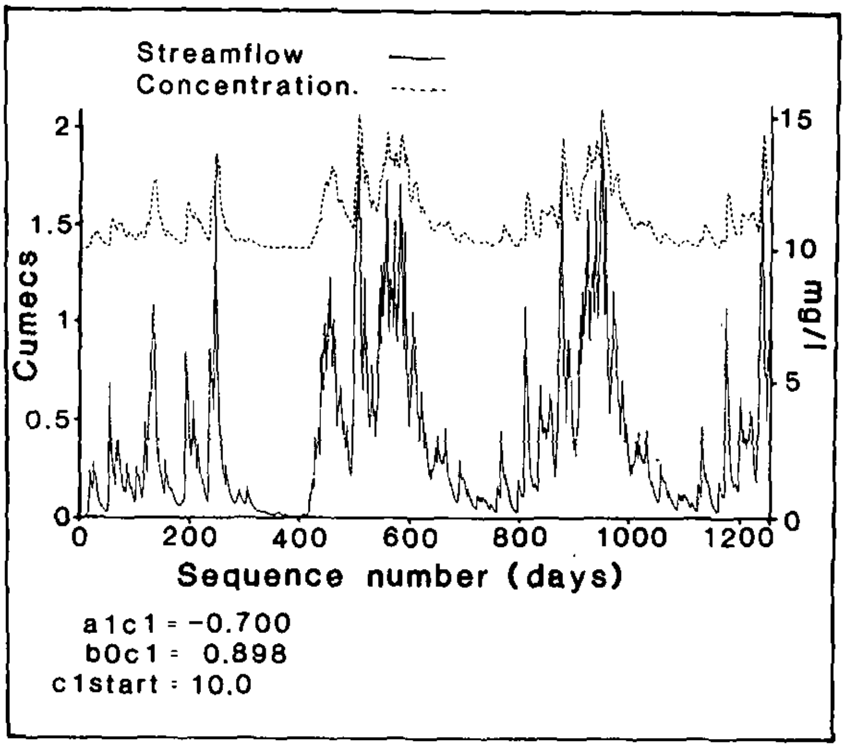


Figure 4.3a SMILER: Concentration data generated from streamflow data

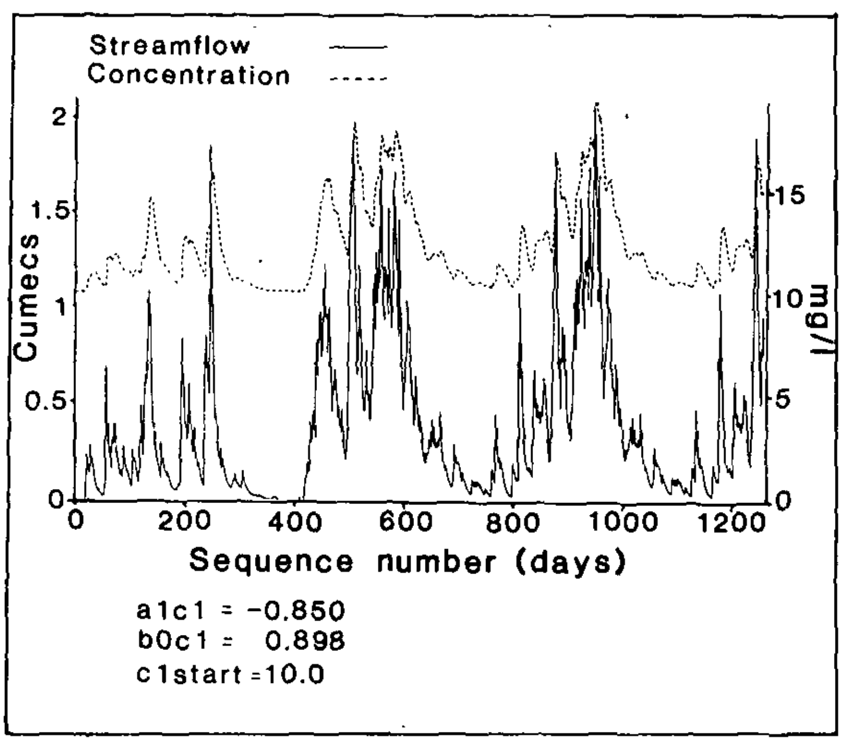


Figure 4.3b SMILER: Concentration data generated from streamflow data

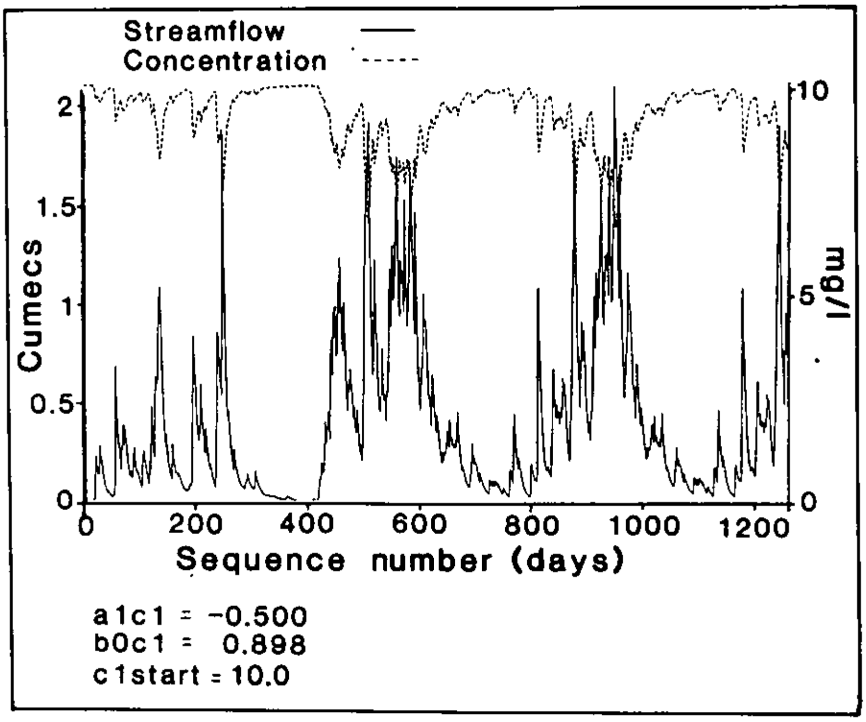


Figure 4.3c SMILER: Concentration data generated from streamflow data

obtained simply by changing the parameter in the numerator of the transfer function from -0.7 to -0.85. The more subdued -ve concentration response shown in Fig. 4.3c was obtained by reversing the on-off switch referred to above and setting the parameter in the numerator of the transfer function to -0.5.

Alternatively, in the rare event that matching time series of observed flow and concentration data for a particular determinand over a long period are available, and that these adequately define the continuous variations of concentration and flow, SMILER could receive an input of the observed concentration data. Most often, the user of SMILER will assume a dynamic response type (with respect to streamflow) for a determinand and site of interest, basing this on inspection of available flow and concentration data. The parameters in SMILER which control the scale and shape of concentration responses can then be adjusted by 'trial and error' until an approximation of the desired behaviour is obtained. Since it is not necessary to have time series which are excellent approximations of reality (they need only be reasonable) this procedure is deemed adequate for initial assessment and comparison purposes.

When flow and concentration time series of the desired dynamic characteristics and record length have been accessed from databases, or generated using SMILER, as described above, a 'true' load can be calculated. The accuracy of loads estimated from the same time series by a variety of methods and using different sampling frequencies can then be assessed against the 'true' load (taking the mean or median of load estimate replicates as the best estimate).

Precision can be expressed as some measure of the spread of the load estimate replicates about the mean or median. SMILER currently can assist with assessments of the efficacy of three interpolation methods of load estimation, namely **Method 2**, **Method 5** and the **Beale Ratio estimator**. Definitions of Methods 2 and 5 were given in Section 2.2. The Beale ratio estimator has been employed extensively in North America (e.g. Young *et al.*, 1988) and involves applying a multiplicative correction factor F, given by (4.1), to Method 5.

$$F = \left[\frac{1 + \frac{1}{N} \cdot \frac{S_{lq}}{\bar{l}\,\bar{q}}}{1 + \frac{1}{N} \cdot \frac{S_q^2}{\bar{q}^2}} \right] \tag{4.1}$$

where N is the number of samples for determination of concentration

$$S_{lq} = \left[\frac{1}{N-1}\right] \left[\sum_{i+1}^{N} \left[Q_i^2 C_i\right] - N \bar{q}\,\bar{l}\right]$$

$$S_q^2 = \left[\frac{1}{N-1}\right] \left[\sum_{i+1}^{N} \left[Q_i^2\right] - N \bar{q}^2\right]$$

C_i is the *i*th concentration

Q_i is the flow corresponding to C_i

$\bar{l}$ is the mean load calculated from the C_i and Q_i

$\bar{q}$ is the mean flow calculated from Q_i

SMILER enables the operator to specify calculation of load over all, or any part, of a synthetic record submitted to the program. Currently, two sets of units are permitted in SMILER; flow, concentration, load and time interval are assumed to be **either** (a) l/s, μeq/l, equivalents and hours respectively **or** (b) m^3/s, mg/l, tonnes and days respectively.

For the period of record selected and a user-specified range of sampling interval (incremented in user-specified steps), SMILER estimates the load by each of the three methods introduced above. In each case the estimated load is replicated 50 times by starting at the 1st, 2nd, . . . , 50th point of the specified period. This method of replication may not be rigorous statistically but it does provide a spread of estimate values which can be examined with respect to bias and precision. Unlike some other methods, therefore, SMILER permits assessment of the effects on errors of (a) length of record and (b) hydrochemical response **within a specified period of record.** In order that the 50 load replicates are based on nearly the same 'true' value of load and therefore that no apparent bias is introduced, periods of record for analysis by SMILER should be chosen such that the variation in flow and concentration over at least the first 50 points in the record is subdued with respect to the variation over the remainder of the record. In most cases the end of a fairly long period of flow recession meets this requirement.

Although SMILER can produce tabular output corresponding to the exercise described above, the most useful output is a plot showing how the median and user-specified percentiles of the distribution of replicated estimates for each estimation method change as the sampling interval increases. For example, Fig. 4.4 shows how, for the data shown in Fig. 4.3c (starting at sequence number 360), the bias (median) and precision (in terms of the 20 and 80 percentiles) vary for Method 2 and 5 loads as the sampling interval increases to 30 days. (The reader is asked to apply visual smoothing to the oscillations which arise because of the discrete, rather than continuous, nature of the source data.) The superiority of Method 5 over Method 2 is clear from Fig. 4.4. Whereas for Method 5 the bias is almost zero at all sampling intervals up to 30 days, and precision increases only to about 1% at 30 days, the bias for Method 2 is about +2% to +3% over the range of sampling intervals and the precision, even at small sampling intervals, is considerably greater than for Method 5.

Figure 4.5a corresponds to the data shown in Fig. 4.3b, i.e. concentration increasing with flow, and shows again the clear superiority of Method 5 over Method 2; the performance (bias and precision) of Method 5 is similar in both Fig. 4.4 and Fig. 4.5a but Method 2 precision is poorer in Fig. 4.5a, corresponding to a +ve flow - concentration relation, though Method 2 bias in both cases is about the same at just under +3%. The effect of applying the Beale Ratio estimator to the data shown in Fig. 4.3b is shown in Fig. 4.5b; in this case there appears little to be gained by incorporating this complex adjustment in the estimation procedure.

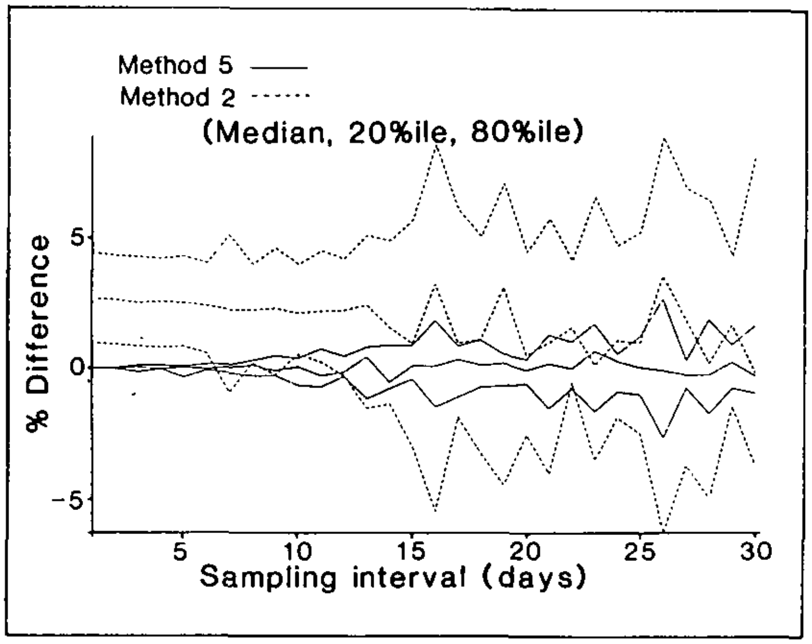


Figure 4.4 Methods 2 and 5 bias and precision versus sampling interval - concentration decreasing with streamflow - 2.5 years

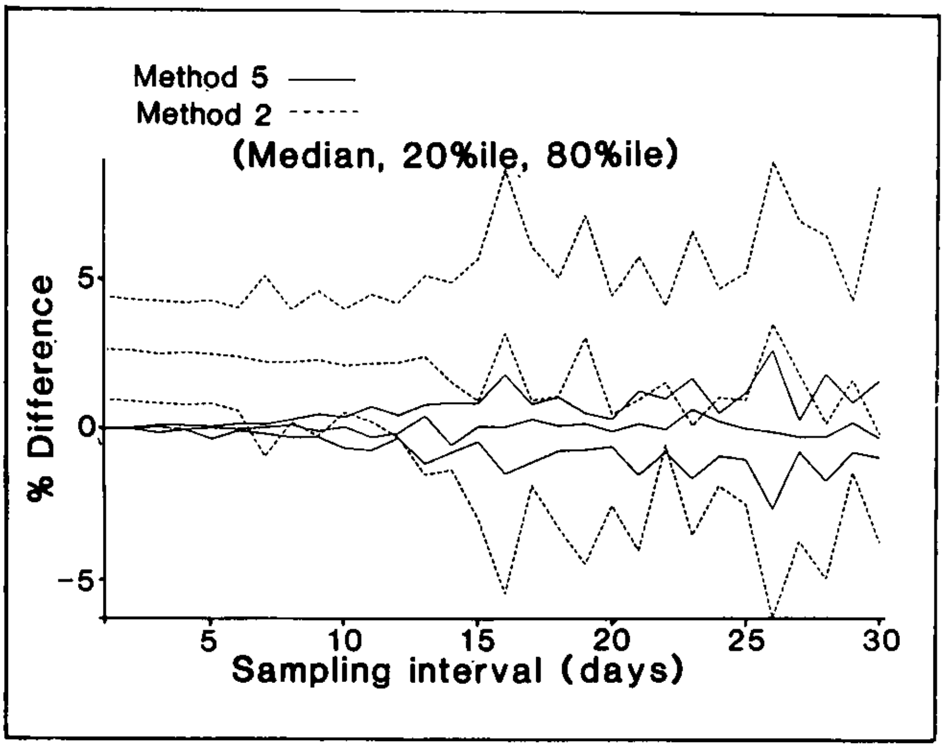


Figure 4.5a Methods 2 and 5 bias and precision versus sampling interval - concentration increasing with streamflow - 2.5 years

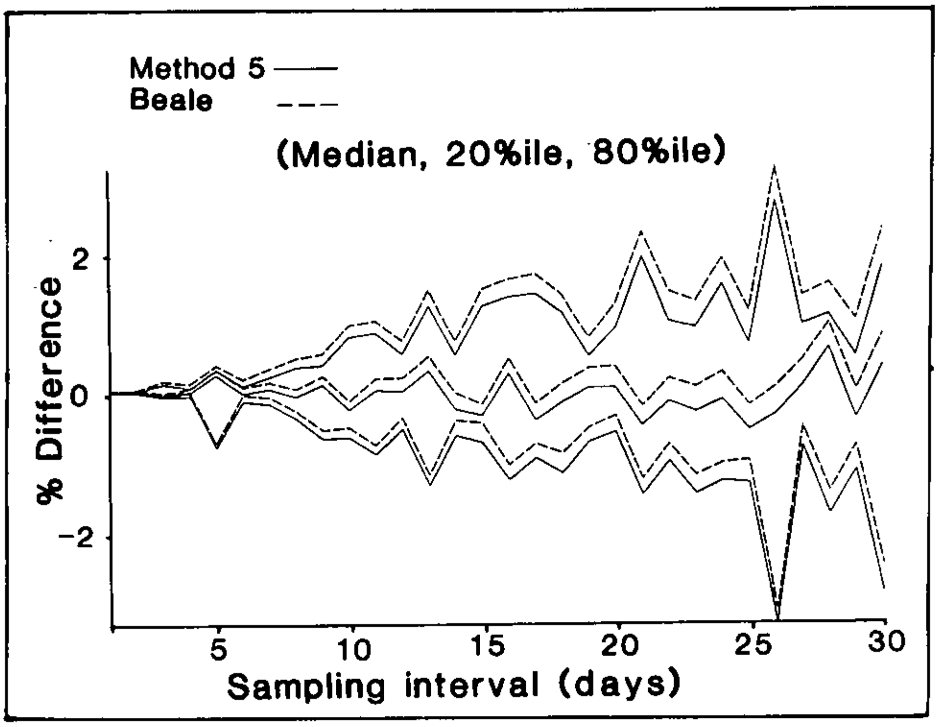


Figure 4.5b Method 5 and the Beale Ratio estimator bias and precision versus sampling interval - concentration increasing with streamflow - 2.5 years

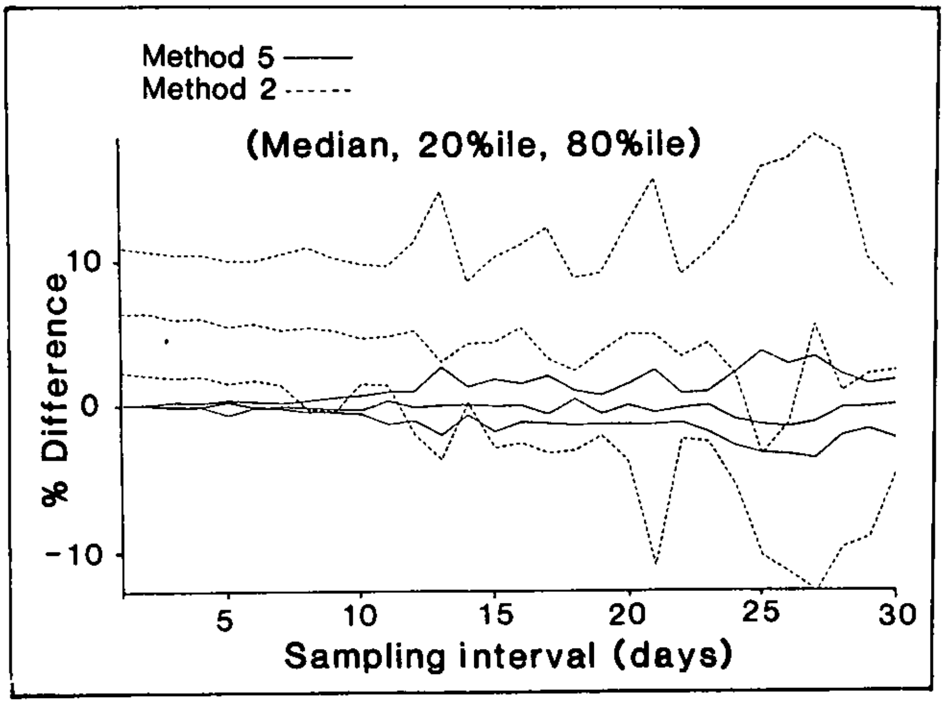


Figure 4.6 Methods 2 and 5 bias and precision versus sampling interval - concentration increasing with streamflow - 13 months

The examples given so far are for a load estimation period of about 2.5 years (sequence numbers 360 to 1261); for the flow and concentration shown in Fig. 4.3b the 'true' load over this period is 533 tonnes. Figure 4.6 shows SMILER profiles for Methods 2 and 5 for the load estimation period from sequence number 360 to 750 (about 13 months), during which the 'true' load for the data shown in Fig. 4.3b is 245 tonnes. Figure 4.6 shows quite clearly that the performance of Method 5 over the shorter period of estimation remains acceptable. However, the bias in Method 2 at high sampling frequencies has risen from about 3% (for the longer estimation period of about 2.5 years) to about +6%, decreasing gradually (for reasons not yet fully understood) to 3% or 4% at a sampling interval of 30 days. The precision of Method 2 for the shorter estimation period is worse than for Method 5 over the range of sampling intervals.

4.3 AN EXAMPLE - ERRORS IN NITRATE LOADS FOR THE STOUR AT LANGHAM

Consider now the specific problem of assessing errors in nitrate loads for a fairly large catchment in East Anglia. Figure 4.7 shows the flow and nitrate data taken from the Harmonised Monitoring Scheme database for the 13.5 years January 1974 to June 1987 for the 578 km^2 Stour at Langham. Clearly, there is a +ve flow - concentration relationship but the corresponding scatter plot shown in Fig. 4.8 indicates that a simple linear regression model would be a totally inadequate description of the dynamics involved. SMILER reads in the continuous 15-year (1974 to 1988) record of daily mean flows for the

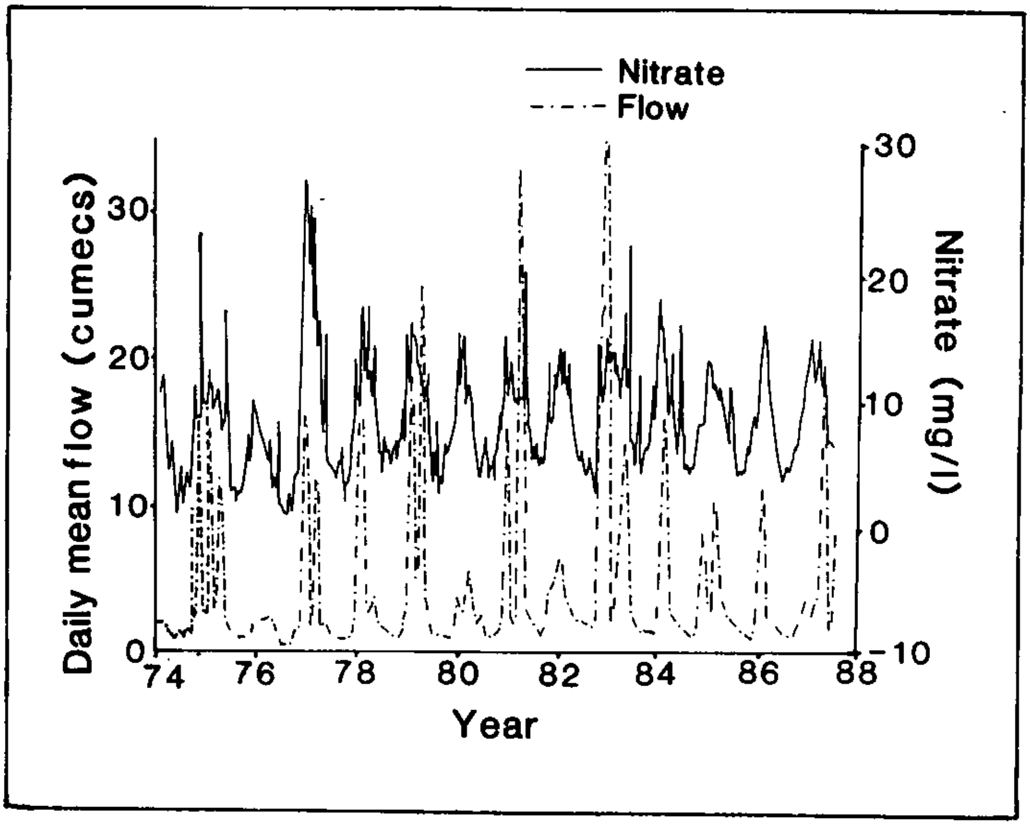


Figure 4.7 Harmonised Monitoring Scheme spot-sample nitrate concentration and corresponding daily mean flow for the Stour at Langham, January 1974 - June 1987

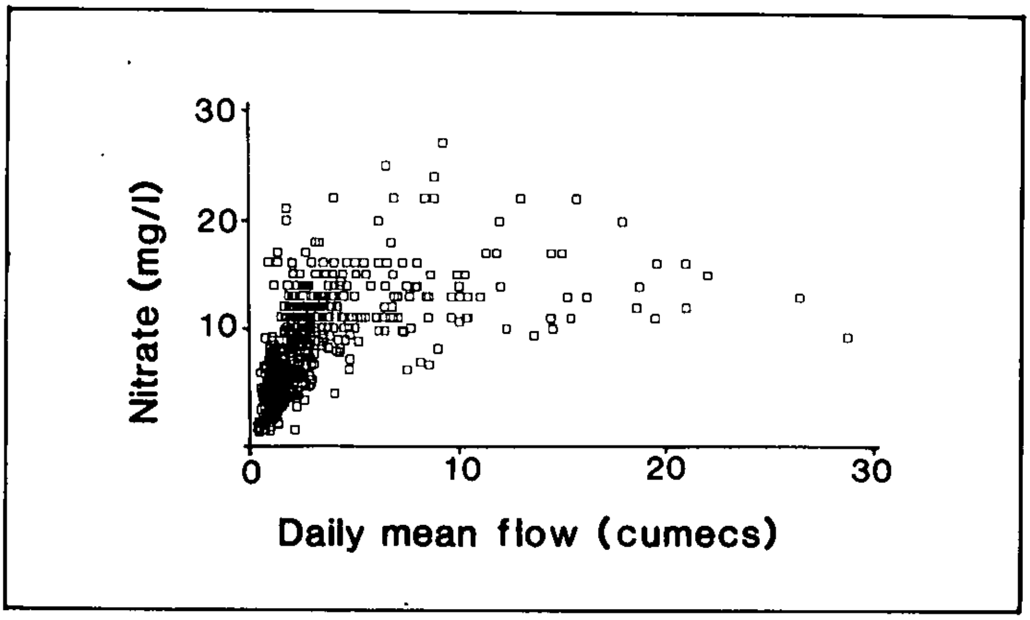


Figure 4.8 Scatter plot of Harmonised Monitoring Scheme spot sample nitrate concentration and daily mean flow for the Stour at Langham, January 1974 - June 1987

Stour at Langham from the Surface Water Archive and, assuming no other information is available, it then generates from those flow data a continuous time series of concentration data which conforms approximately, in terms of mean and range, to the nitrate data taken from the Harmonised Monitoring Scheme database. The degree of hysteresis introduced for this demonstration case is arbitrary but the scatter in Fig. 4.9 agrees fairly well with that in Fig. 4.8.

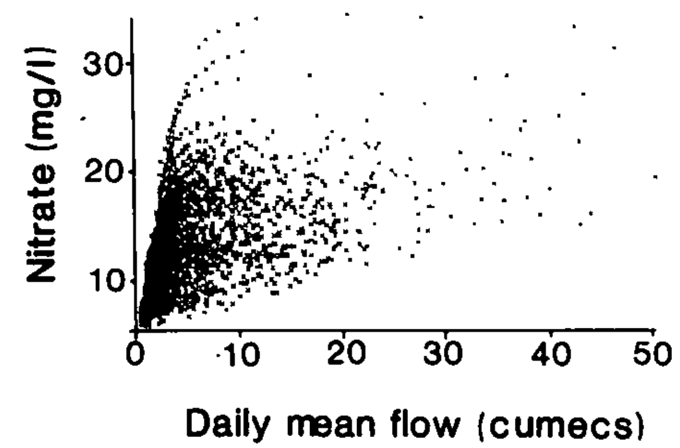


Figure 4.9 Scatter plot of 'nitrate' concentration generated by SMILER from 1974 - 1988 daily mean flow data for the Stour at Langham taken from the Surface Water Archive

No attempt has been made here to calibrate a transfer function model using the limited amount of nitrate data available from the Harmonised Monitoring database. These data are not spaced regularly in time and therefore would present difficulties in analysis to calibrate a model. However, the synthetic nitrate concentration data generated by SMILER are reasonable in that the annual 'true' loads are, in most cases, not totally dissimilar to the loads calculated by Methods 2 and 5, as shown in Table 4.1. Of the 11 years when all three estimates are available there are

6 where Method 2 < Method 5 < SMILER
3 where Method 5 < Method 2 < SMILER
and 2 where SMILER < Method 5 < Method 2.

Bearing in mind that individual Method 2 and Method 5 estimates could have large errors, and that these Methods are likely to underestimate river loads, the SMILER estimates are considered to be acceptable.

Figure 4.10 shows the 15-year record of observed daily mean flow (from the Surface Water Archive) and synthetic nitrate concentration (generated by SMILER). These time series may now be employed to compute a 'true' load

Table 4.1 Comparison of annual loads calculated by Methods 2 and 5 and by SMILER

		STOUR AT LANGHAM NITRATE LOAD (tonnes)		
YEAR	NUMBER OF SAMPLES	METHOD 2 ESTIMATE	METHOD 5 ESTIMATE	SMILER 'TRUE'
1974	41	640	767	1228
1975	48	855	918	1501
1976	*	*	*	413
1977	63	1305	1264	1112
1978	62	1066	1147	1185
1979	50	1790	1572	2091
1980	52	813	929	1067
1981	48	1386	1241	1289
1982	50	1465	1365	1761
1983	52	1505	1363	1542
1984	38	1212	1269	1451
1985	49	661	791	958
1986	*	*	*	924
1987	*	*	*	2556
1988	*	*	*	2303

* Sample concentrations do not adequately cover the calendar year or data not available.

for any sub-period against which combinations of sampling frequency and estimation method can be assessed. As given in Table 4.2, the 'true' load for the period between sequence numbers 150 and 5479 (14.6 years) in Fig. 4.10 is 21110 tonnes; 'true' average flux is 45.8 g/s; average flow is 3.3 cumecs; 'true' arithmetic average concentration is 10.5 mg/l. The other statistics in Table 4.2, describing the variability of, and association between, flow and concentration over the estimation period, will assist with further development of SMILER.

The variation in level of bias and precision of Method 5 load estimates for the 14.6 year period as sampling interval increases is shown in Fig. 4.11. Although Method 5 bias in this case, calculated as the median of replicates, becomes more variable as sampling interval increases it apparently introduces little bias. By visually smoothing the curves in Fig 4.11 it can be observed that precision worsens in an approximately linear manner from close to zero at a sampling interval of one day to about +/-3% at 30 days. Figure 4.12 shows that, for the same data and estimation period, Method 2 also introduces little bias but precision is about +/-10% at 30 days.

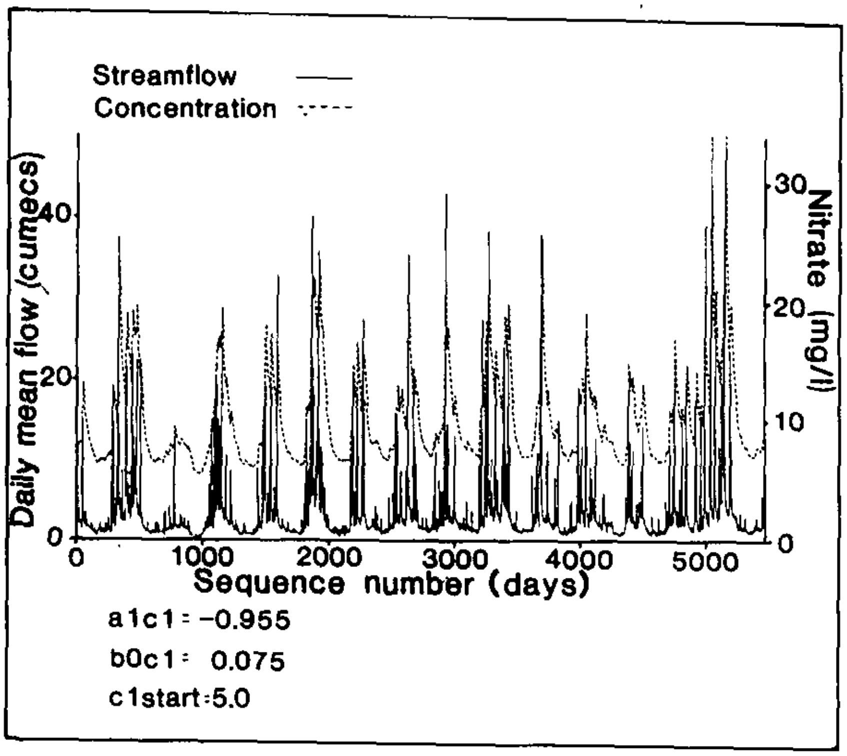


Figure 4.10 Time series plot of Surface Water Archive daily mean flow for the Stour at Langham 1974 - 1988 and corresponding 'nitrate' concentration generated by SMILER

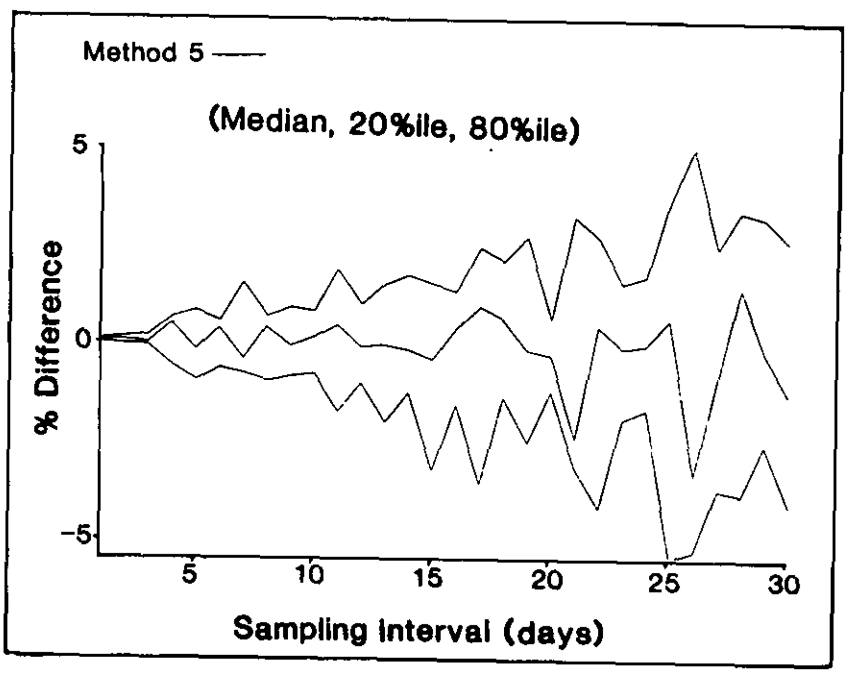


Figure 4.11 Method 5 bias and precision versus sampling interval for a 14.6 year 'nitrate' load for the Stour at Langham

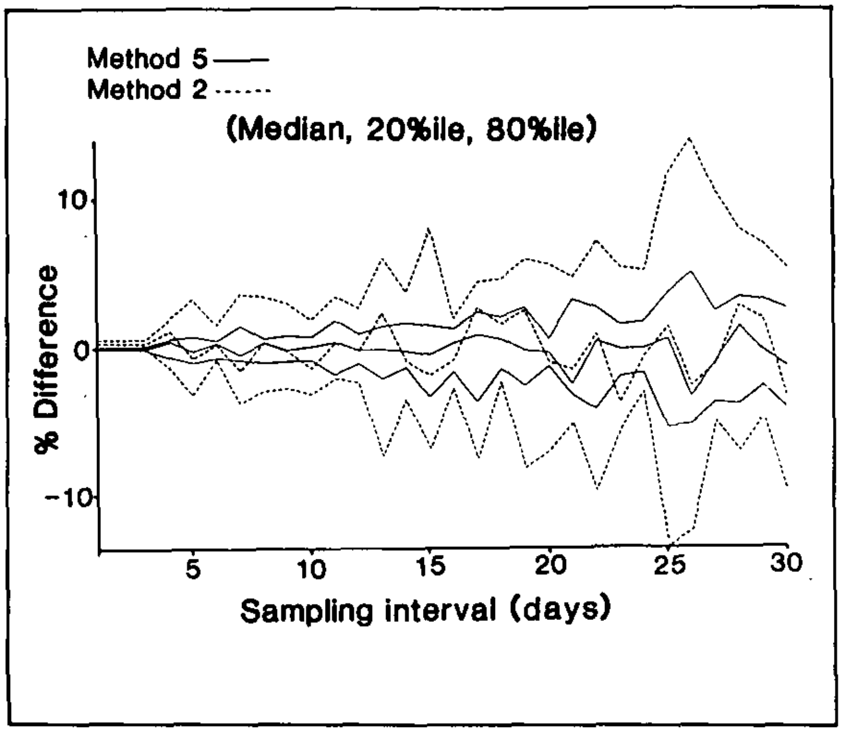


Figure 4.12 *Methods 2 and 5 bias and precision versus sampling interval for a 14.6 year 'nitrate' load for the Stour at Langham*

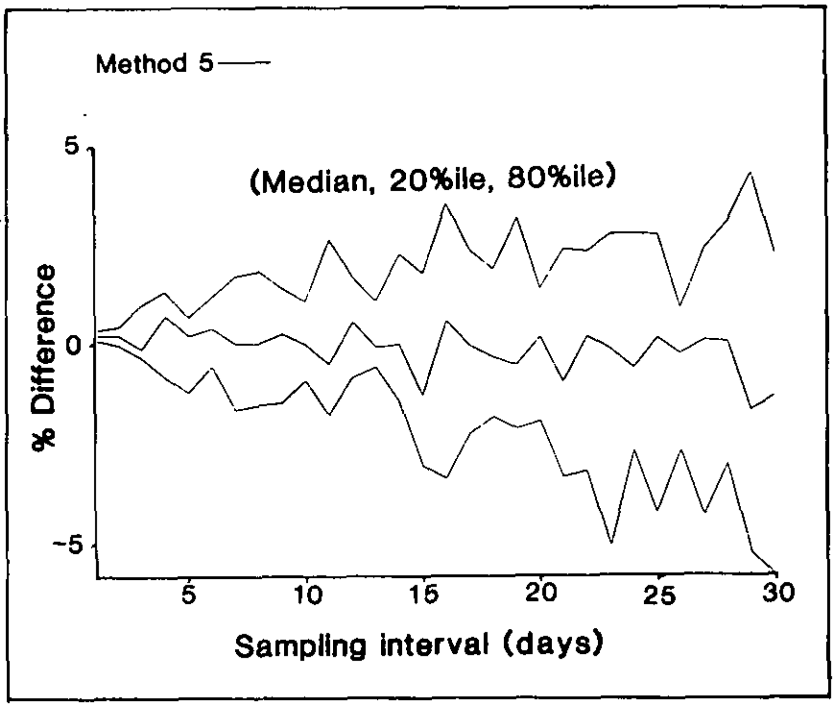


Figure 4.13 *Method 5 bias and precision versus sampling interval for a 6.4 year 'nitrate' load for the Stour at Langham*

Table 4.2 SMILER 'nitrate' load statistics for a 14.6 year period for the Stour at Langham

Stour (578 sq.km.)

Number of samples		=	5479 (days)	
First point in series		=	1	

Load has been estimated between
sequence numbers 150 and 5479 (days)

Flow - concentration model parameters

a1c1 = -0.955	c1start	=	5.000	mg/l
b0c1 = 0.075	idil	=	2	

"True" load	=	21110	Tonnes
"True" average flux	=	45.849	g/s
"True" average flow	=	3.309	cumecs
"True" arith. av. concl	=	10.507	mg/l
Coefficient of variation (flow)	=	1.368	
Coefficient of variation (concl)	=	0.399	
Ratio	=	0.292	
Correlation coeff. (flow, concl)	=	0.583	
Coefficient of variation (load)	=	1.950	
Variability Index, VI	=	1.239	
Coefficient of variation (CQ)	=	1.950	
Standard deviation, C	=	4.196	mg/l
Standard deviation, Q	=	4.528	cumecs
Standard deviation, CQ	=	89.420	g/s

Similar plots can be produced by SMILER for any specified period within the data shown in Fig. 4.10, enabling the user to investigate how errors in load estimates are influenced by the length of the estimation period and the variability of flow and concentration during that time. Figures 4.13 to 4.15 and Tables 4.3 to 4.5 correspond to Method 5 estimates for the periods between sequence numbers (Fig. 4.10) 150 and 2500 (6.4 years, 8139 tonnes), 150 and 1400 (2.8 years, 3925 tonnes) and 150 and 600 (1.2 years 2377 tonnes) respectively. It would appear from this series of plots that there is little deterioration in bias and precision in Method 5 estimates as the estimation period decreases from about six years to about one year. Bias remains small as estimation period decreases in this range, and the modest changes in precision could be due to differences amongst the periods in variability of flow and concentration (see next paragraph). Figure 4.16 reproduces the Method 5 curve in Fig. 4.15 on a different scale and shows also that Method 2 estimates for the 1.2 year period are about +/-20% at 30 days compared with about +/-10% at 30 days for the 14.6 year period (Fig. 4.12).

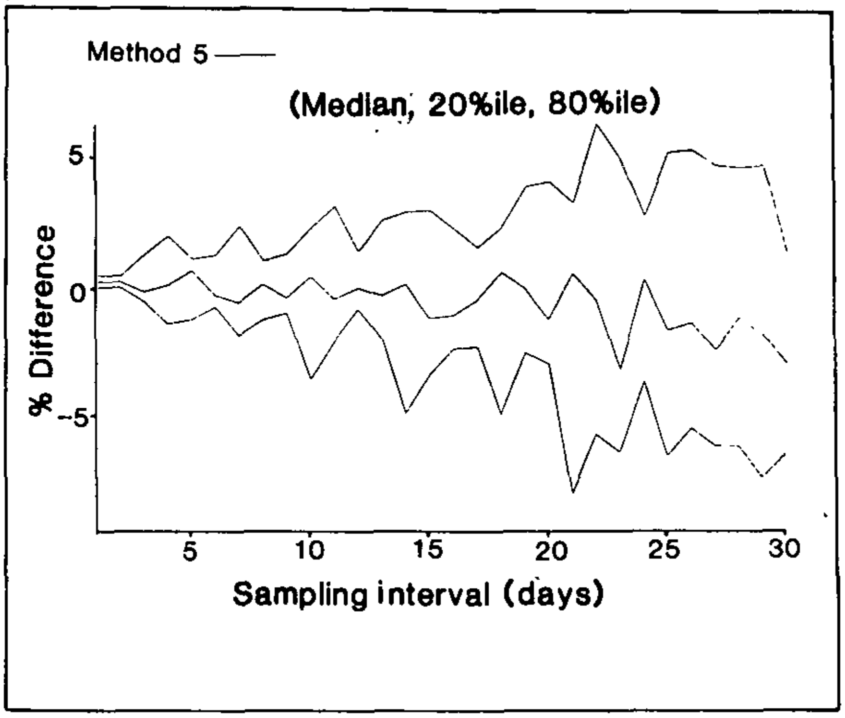


Figure 4.14 Method 5 bias and precision versus sampling interval for a 2.8 year 'nitrate' load for the Stour at Langham

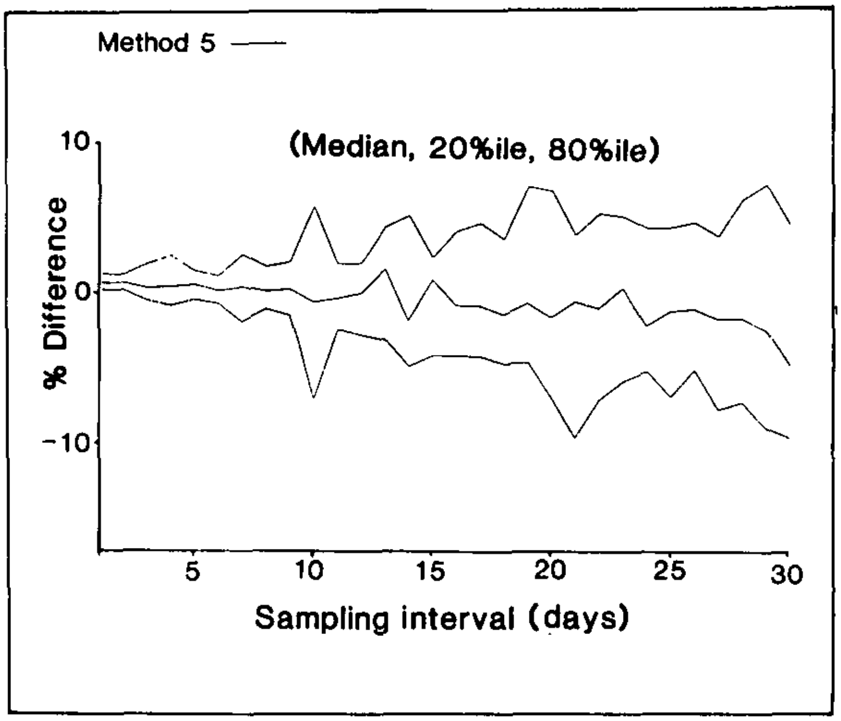


Figure 4.15 Method 5 bias and precision versus sampling interval for a 1.2 year 'nitrate' load for the Stour at Langham

Table 4.3 ***SMILER 'nitrate' load statistics for a 6.4 year period for the Stour at Langham***

Stour (578 sq.km.)

Number of samples	=	5479	(days)
First point in series	=	1	

Load has been estimated between sequence numbers 150 and 2500 (days)

Flow - concentration model parameters

alcl = -0.955	clstart	=	5.000	mg/l
b0cl = 0.075	idil	=	2	

"True" load	=	8139	Tonnes
"True" average flux	=	40.087	g/s
"True" average flow	=	2.945	cumecs
"True" arith. av. concl	=	9.898	mg/l

Coefficient of variation (flow)	=	1.409	
Coefficient of variation (concl)	=	0.424	
Ratio	=	0.301	
Correlation coeff. (flow, concl)	=	0.629	
Coefficient of variation (load)	=	1.873	
Variability Index, VI	=	1.242	
Coefficient of variation (CQ)	=	1.873	
Standard deviation, C	=	4.194	mg/l
Standard deviation, Q	=	4.149	cumecs
Standard deviation, CQ	=	75.102	g/s

Table 4.4 ***SMILER 'nitrate' load statistics for a 2.8 year period for the Stour at Langham***

Stour (578 sq.km.)

Number of samples	=	5479	(days)
First point in series	=	1	

Load has been estimated between sequence numbers 150 and 1400 (days)

Flow - concentration model parameters

alcl = -0.955	clstart	=	5.000	mg/l
b0cl = 0.075	idil	=	2	

"True" load	=	3925	Tonnes
"True" average flux	=	36.343	g/s
"True" average flow	=	2.713	cumecs
"True" arith. av. concl	=	9.532	mg/l

Coefficient of variation (flow)	=	1.453	
Coefficient of variation (concl)	=	0.431	
Ratio	=	0.296	
Correlation coeff. (flow, concl)	=	0.647	
Coefficient of variation (load)	=	1.964	
Variability Index, VI	=	1.255	
Coefficient of variation (CQ)	=	1.964	
Standard deviation, C	=	4.106	mg/l
Standard deviation, Q	=	3.943	cumecs
Standard deviation, CQ	=	71.360	g/s

Table 4.5 SMILER 'nitrate' load statistics for a 1.2 year period for the Stour at Langham

Stour (578 sq.km.)

Number of samples	=	5479 (days)
First point in series	=	1

Load has been estimated between sequence numbers 150 and 600 (days)

Flow - concentration model parameters

a1c1 = -0.955	c1start	=	5.000 mg/l
b0c1 = 0.075	idil	=	2

"True" load	=	2377 Tonnes
"True" average flux	=	61.138 g/s
"True" average flow	=	4.000 cumecs
"True" arith. av. concl	=	11.666 mg/l

Coefficient of variation (flow)	=	1.317
Coefficient of variation (concl)	=	0.411
Ratio	=	0.312
Correlation coeff. (flow, concl)	=	0.573
Coefficient of variation (load)	=	1.626
Variability Index, VI	=	1.213
Coefficient of variation (CQ)	=	1.626
Standard deviation, C	=	4.791 mg/l
Standard deviation, Q	=	5.270 cumecs
Standard deviation, CQ	=	99.403 g/s

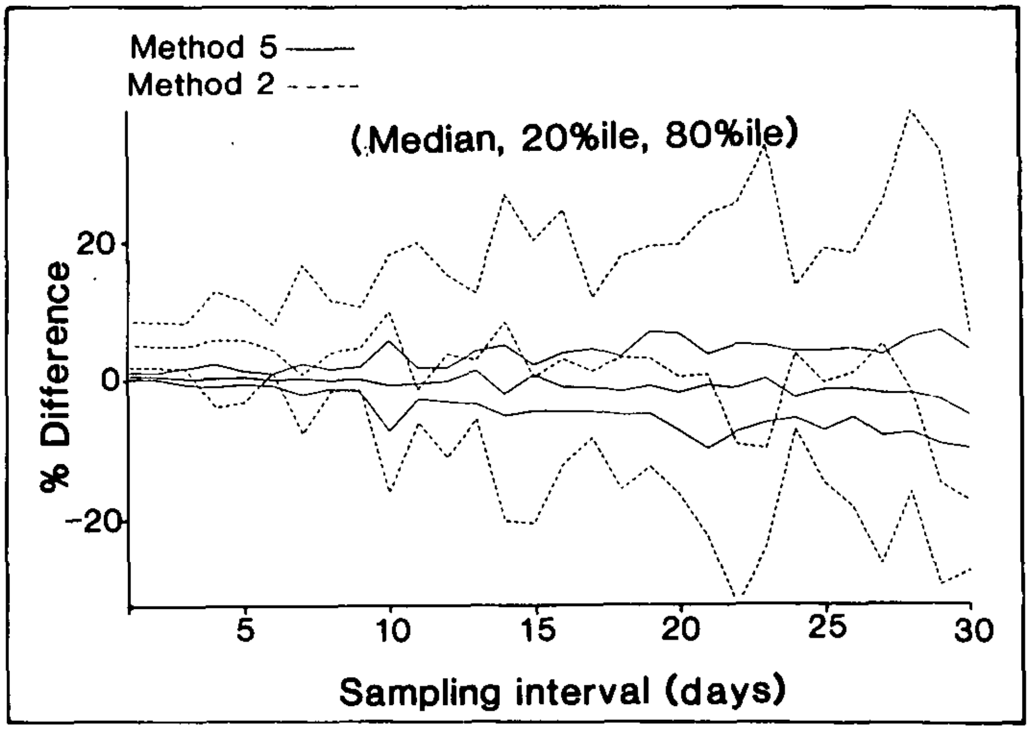


Figure 4.16 Methods 2 and 5 bias and precision versus sampling interval for a 1.2 year 'nitrate' load for the Stour at Langham

The influence on errors of the variability in flow and concentration during an estimation period can be assessed by applying SMILER to periods of equal length where the variability is quite different. Figure 4.17 shows Method 5 curves for the one-year period between sequence numbers (Fig. 4.10) 600 and 965 (low flows and relatively subdued concentration response - 298 tonnes) and Fig. 4.18 shows similar curves for the period between 1700 and 2065 (relatively high flows and more pronounced concentration response - 2051 tonnes). Tables 4.6 and 4.7 correspond to Figs 4.17 and 4.18 respectively. Fig. 4.17 shows that for the year with a low nitrate load the bias is less than 1% and the precision is better than 2% for a sampling interval of 30 days. For the year with a relatively high nitrate load, bias is again typically less than 1% but precision at a 30-day sampling interval is about +/-7%. (The reason for the downward trend in bias in Fig. 4.17 is not clear but it could be due to relatively high variability, in the first 50 samples, with respect to the remainder of the one-year record – see also the downward trends in Figs 4.14 and 4.15.)

The benefit of the additional information in the 'continuous' flow record, as utilised by Method 5 above, is clear. Precisions of corresponding Method 2 estimates (the SMILER plots are not given here) are inferior; about +/-10% for the 'low load' year and about +/-20% for the 'high load' year.

The surprisingly good precision in annual nitrate loads estimated from just 12 or 13 samples per year for the Stour at Langham (+/- 2% in a 'low load' year and +/- 7% in a 'high load' year) should be viewed with caution. In practice, there will be additional components of random error which arise during the measurement processes (not dealt with here). Furthermore, it should be remembered that the synthetic concentration data were generated using a model structure which, though arguably superior to a simple regression model (because it can incorporate at least some hysteresis), does not allow for time-variant flow - concentration responses which may exist in the real world (due to exhaustion or seasonal changes from a dilution to a purging mechanism). The parameters of the transfer function employed to generate the synthetic concentration data were necessarily selected in a somewhat arbitrary way.

The 'nitrate' concentration time series generated here by SMILER cannot be expected, therefore, to reproduce faithfully the unmeasured daily sequence of nitrate concentration for 1974 to 1988 for the Stour at Langham. Although Table 4.1 indicates a reasonable agreement overall between annual SMILER 'true' loads and loads estimated by Methods 2 and 5, there are significant differences in individual years. It is not possible to draw any firm conclusions from this observation (since we do not know any of the true annual loads) but it does warn against unconditional acceptance of the provisional SMILER results. If daily observed nitrate concentrations were available for the 15-year period (better – if hourly data were available), analysis might reveal that just 12 or 13 samples per year give annual loads with worse precisions than indicated by the SMILER results presented here. Further work is required to test and develop SMILER to ensure better representativeness of synthetic data.

However, the SMILER analyses related to the Stour at Langham presented here confirm at least two important points made earlier in the report. First, the precision in annual load estimates of a determinand like dissolved nitrate

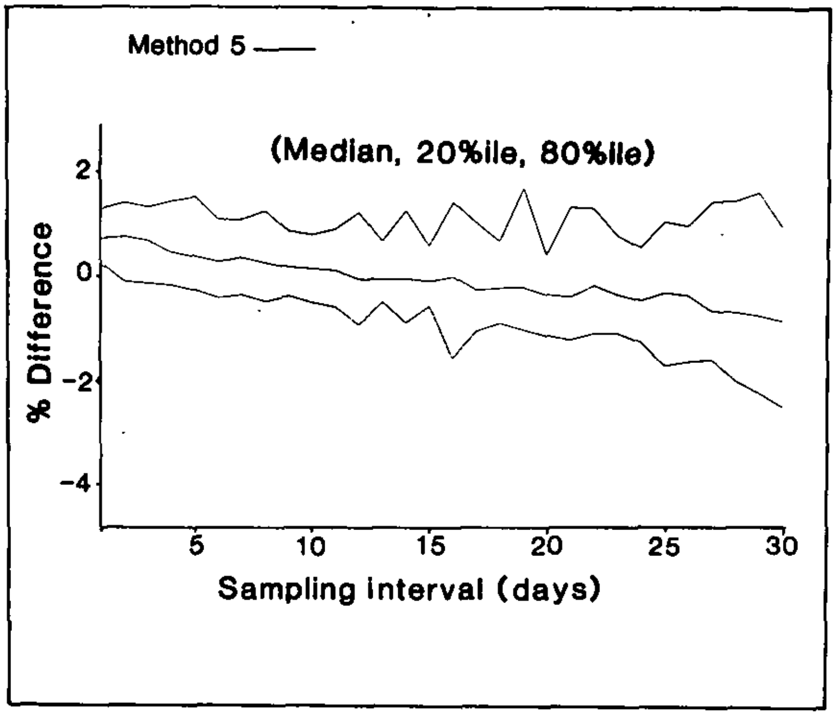


Figure 4.17 Method 5 bias and precision versus sampling interval for a one-year, low nitrate load period for the Stour at Langham

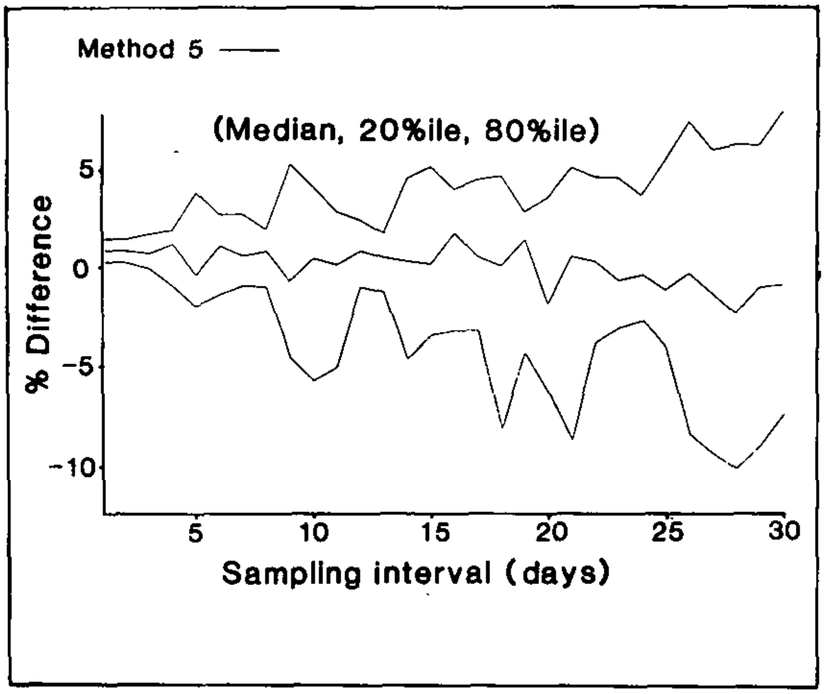


Figure 4.18 Method 5 bias and precision versus sampling interval for a one-year, high nitrate load period for the Stour at Langham

Table 4.6 SMILER 'nitrate' load statistics for a low-load year

Stour (578 sq.km.)

Number of samples	=	5479 (days)	
First point in series	=	1	

Load has been estimated between sequence numbers 600 and 965 (days)

Flow - concentration model parameters

a1c1 = -0.955	c1start	=	5.000	mg/l
b0c1 = 0.075	idil	=	2	

"True" load	=	298	Tonnes
"True" average flux	=	9.461	g/s
"True" average flow	=	1.246	cumecs
"True" arith. av. concl	=	7.141	mg/l
Coefficient of variation (flow)	=	0.815	
Coefficient of variation (concl)	=	0.127	
Ratio	=	0.156	
Correlation coeff. (flow, concl)	=	0.611	
Coefficient of variation (load)	=	0.938	
Variability Index, VI	=	0.845	
Coefficient of variation (CQ)	=	0.938	
Standard deviation, C	=	0.908	mg/l
Standard deviation, Q	=	1.015	cumecs
Standard deviation, CQ	=	8.870	g/s

is typically much better than precision in annual load estimates of suspended sediment – because of the relatively subdued degree of variation of nitrate concentration compared with that for suspended sediment. Second, the precision of annual load estimates varies considerably from year to year according to changes between years in the level of hydrochemical activity.

SMILER can be operated in a similar manner for almost any combination of hydrological and flow-concentration behaviour types characterised on the basis of available time series and other information. In this way SMILER can, subject to its current preliminary stage of development, assist with **either** (a) the design of a sampling strategy for a particular determinand at a given site to obtain specified levels of accuracy and precision for mass loads **or** (b) the converse of (a), namely assessing the likely levels of accuracy and precision in mass loads of a given determinand at a particular site and for a specified sampling strategy.

Table 4.7 SMILER 'nitrate' load statistics for a high-load year

Stour (578 sq.km.)

Number of samples	=	5479 (days)
First point in series	=	1

Load has been estimated between sequence numbers 1700 and 2065 (days)

Flow - concentration model parameters

a1c1 = -0.955	c1start	=	5.000 mg/l
b0c1 = 0.075	idil	=	2

"True" load	=	2051	Tonnes
"True" average flux	=	65.033	g/s
"True" average flow	=	3.966	cumecs
"True" arith. av. concl	=	11.608	mg/l
Coefficient of variation (flow)	=	1.316	
Coefficient of variation (concl)	=	0.478	
Ratio	=	0.363	
Correlation coeff. (flow, concl)	=	0.656	
Coefficient of variation (load)	=	1.643	
Variability Index, VI	=	1.181	
Coefficient of variation (CQ)	=	1.643	
Standard deviation, C	=	5.544	mg/l
Standard deviation, Q	=	5.219	cumecs
Standard deviation, CQ	=	106.849	g/s

5 United Kingdom river quantity and quality databases

5.1 INTRODUCTION

A clearly defined need has long been perceived for river flow databases for describing floods and droughts to assist with the planning, design and optimal operation of river channel flood control works and water supply schemes. At the national scale, the Surface Water Archive and other related databases maintained at the Institute of Hydrology service the need for information on water quantity. The Surface Water Archive receives data and information mainly from the regional Divisions of the National Rivers Authority covering England and Wales, the River Purification Boards covering Scotland and the Department of the Environment (Northern Ireland).

Environmental management today recognises more widely that the quality of our natural and engineered water resources is of equal importance to the quantity aspects. At the national scale, the repository of river water quality data is the Harmonised Monitoring Scheme database maintained by the Department of the Environment (Her Majesty's Inspectorate of Pollution). (It appears likely that in 1991, responsibility for the Harmonised Monitoring Scheme database will be transferred to the National Rivers Authority.) Data input to the Harmonised Monitoring database are principally from the National Rivers Authority and the River Purification Boards (river quality data from Northern Ireland are not input to the database currently).

Mainly for historical reasons, but also because of differences in the types of data, existing provisions for data retrieval, presentation and analysis vary between the national quantity and quality databases. In the context of river load estimation it is necessary to consider the best way of jointly exploiting such databases. The following sections give brief descriptions of the national databases, including details of their data retrieval and analysis facilities. The scope of Geographical Information Systems (GIS) generally, and the Water Information System in particular, for river loads data and information processing, is also discussed briefly.

5.2 THE SURFACE WATER ARCHIVE

The following points of interest are taken from a historical perspective of the Surface Water Archive given by Lees (1987).

In the 19th century regular river flow measurement was restricted to the Thames at Teddington and the Lee at Feildes Weir, but the need for a comprehensive survey of inland water quantity was already perceived. Early this century it was suggested that effluent standards should be adjusted according to the dilution available in the receiving watercourse, and some river flow measurements were made accordingly for planning purposes. The decision to undertake an Inland Water Survey was taken in 1934; systematic river flow measurement throughout the United Kingdom commenced, supported by regular publication of corresponding data and information in a Yearbook format.

The earliest Surface Water Yearbook, for the years 1935 and 1936, published daily mean river flow for 28 stations. Publication of Yearbooks has continued to the present day (albeit with some publication delays), evolving in style and content as administrative responsibilities for river flow measurement and maintainance of the national database have changed, and as information technology has developed. From 1982 onwards, management of the Surface Water Archive and publication of hydrological data covering the United Kingdom has been undertaken by the Natural Environment Research Council (at the Institute of Hydrology in collaboration with the British Geological Survey). Currently, the Archive contains data for about 1300 sites, about 1000 of which are extant. On a 'following year' basis the annual *Hydrological Data UK* publishes river flow data and information for about 200 sites. All the data in the Archive are stored in a computer system and *Hydrological Data UK* gives details of available data retrieval options (e.g. simple tabulations, hydrograph plots, flow-duration curve summaries). Effectively, and only for a modest handling charge, river flow data and information for almost any measuring station in the United Kingdom are readily available to a wide range of customers in pre-processed, quality controlled formats. Figure 5.1 shows the daily mean flows, and summary statistics, for the Stour at Langham taken from the 1989 Yearbook in the *Hydrological Data UK* series.

When environmental managers need to establish what river quality and quantity data are available it often requires separate enquiries within the same organisation, and sometimes enquiries to different organisations. In recognition of the increasing need for river quality and quantity data and information to be made available together, the series *Hydrological Data UK* now includes summaries of river quality data from a selection of Harmonised Monitoring sites. With agreement from Her Majesty's Inspectorate of Pollution, the 1986, 1987 and 1988 Yearbooks presented statistical summaries for the year, and for the period of previous record, for about a dozen determinands from 16 sites. In the 1989 Yearbook the number of sites has been increased to 32 (two of which are in Northern Ireland) and a greater range of determinands is featured. Figure 5.2 shows river quality statistical summaries for four sites (including the Stour at Langham) reproduced from the 1989 Yearbook. It is intended to extend this service by giving river loads – provided these can be estimated with acceptable accuracy and precision.

5.3 THE HARMONISED MONITORING SCHEME DATABASE

Systematic measurement of river quality, and the collection, collation and publication of river quality data on a national scale do not have the same depth of history as outlined above for quantity data.

The first modern River Pollution Survey was undertaken for 1970. Although improvements were introduced into the 1972 and 1973 Surveys, it was recognised that the nature of the river stretch classification, based largely on BOD levels and whether or not there were discharges to the stretch, rendered

036006 Stour at Langham — 1989

Measuring authority: NRA-A
First year: 1962

Grid reference: 62 (TM) 020 344
Level stn (m OD): 6.40

Catchment area (sq km) 578 0
Max alt. (m OD) 128

Daily mean gauged discharges (cubic metres per second)

DAY	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1	1 648	2 875	4 527	2.539	2 600	0 960	2.107	1 194	0 802	0 747	1.262	2 379
2	1 631	2 695	4 501	3 693	2 376	1.134	2 002	0 989	0 827	0 760	1 426	2 501
3	1 484	2.508	3 875	9 310	2 274	1 094	1 801	1 036	0 854	0 758	1 563	2 295
4	1,480	2 388	4 665	4 699	2 182	1,180	1,839	1 058	0,913	0,781	1 386	2 392
5	1 811	2,265	4 068	5 320	2 081	0 972	1 798	0 953	0 757	0 729	1 451	2 598
6	5 728	2,207	3 725	9 018	1 988	1 385	1 699	0 905	0 688	0 889	1 148	2 628
7	3 788	1,861	3 263	16 410	1 839	1 686	2 555	0 854	0 736	0 934	1 398	2 685
8	2,667	1 978	2 614	9 832	1 725	1,370	2,462	0 757	0,754	0,938	1 798	2 581
9	2,648	2 471	3 365	5 503	1 721	1 420	2 330	0 805	0 803	0 817	1 751	2 465
10	2,412	3 557	4 987	4 753	1 603	1 220	1,995	1 093	0 801	0 974	1 925	2 205
11	1,909	4 307	5 152	6 685	1 660	1 266	1,425	1 052	0 787	0,781	1,950	1.735
12	2 124	4 601	5 004	9 550	1.707	1 073	1 435	0 854	0 806	0 766	2 025	2 244
13	3,503	4 283	5 163	5 561	1 721	1 161	1,419	0 972	0 856	0 729	1 816	2 934
14	3 288	4 452	5 414	3 823	1 525	1 175	1,572	0 884	0 943	0 667	2 282	4 944
15	3 449	4 285	6 694	3 227	1 355	1 163	1 892	1 003	0 882	0 714	2 127	4 824
16	3 401	4 217	10 980	3 689	1 355	1 167	1 840	1 080	0 883	0 707	2 104	4,290
17	2 124	4 856	21 400	2 850	1 385	1 120	1,921	1 039	0 875	0,684	2 153	4 606
18	2,375	6 764	23 830	2 962	1 535	1 082	1 990	0 856	0 886	0 851	2 141	3 063
19	2,461	6,213	9 596	2 915	1 539	1 014	1 989	0 892	0 814	0 956	2 224	3 470
20	1,954	5 054	7 077	2 796	1 512	1 410	1 801	0 759	0,753	1,129	2,226	6 272
21	3 947	4 894	9 275	2 989	1 517	1 330	1.728	0,817	0 716	1,328	2,306	14 910
22	11 600	4 669	7 962	2 834	1 522	1 614	1 711	0,845	0 778	1 083	2.157	10 860
23	6 070	4.792	5 519	2 642	1 348	1 775	1.720	0 832	0 663	1,002	2 263	4.241
24	3,313	4 854	4 719	3 013	1 351	1.752	1.779	0,821	0 905	1 026	2,309	3 895
25	2.595	6.551	3 399	5 576	1 461	1 691	1 821	1 038	0 627	1 020	2 305	3 172
26	2,985	14 790	3 147	5 147	1 414	1 655	1 636	1,295	0 693	1 052	2,355	2 784
27	2,509	14 570	3 097	4 112	1 291	1 924	1 660	1 232	0 718	1 092	2 366	1 680
28	2 595	9,348	2 871	3 046	1 284	2 026	1,545	0 941	0 729	1 117	2 226	2 451
29	5,642		2 559	2 418	1 268	2 246	1,197	0 932	0 764	1 223	2 325	2 013
30	4 148		2 335	2 601	1 068	2 093	1 177	0 817	0 703	1 237	2 374	1 552
31	2,952		2 591		0 858		1 285	0 812		1 216		1 648
Average	3 234	4 939	6 044	4 984	1 615	1 405	1 778	0 949	0 790	0,926	1 971	3 623
Lowest	1 480	1 861	2 335	2 418	0 858	0 960	1 177	0,757	0 627	0 667	1 148	1 552
Highest	11 600	14 790	23 830	16 410	2 600	2 246	2 555	1 295	0 943	1 328	2 374	14 910
Peak flow	13 67	16 89	28 05	17 86	2 96	2 56	3 13	1 55	1.12	1,91	2 58	16 41
Day of peak	22	26	18	7	1	21	7	26	24	8	14	22
Monthly total (million cu m)	8 66	11 95	16 19	12 92	4 33	3 64	4 76	2 54	2 05	2 48	5.11	9 70
Runoff (mm)	15	21	28	22	7	6	8	4	4	4	9	17
Rainfall (mm)	37	39	54	66	7	49	45	30	11	42	27	103

Statistics of monthly data for previous record (Oct 1962 to Dec 1988)

		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Mean flows	Avg	5 708	4 983	4,774	3 643	2 416	1 653	1.101	1 178	1.176	1 996	2 898	4 061
	Low	1 398	0 883	1 597	1 217	0 758	0 454	0 191	0 210	0 395	0 510	0 578	0 692
	(year)	1965	1965	1976	1974	1974	1965	1976	1976	1964	1970	1964	1964
	High	16 080	12 980	9 775	9 334	7 253	5 999	2 957	6 236	4 945	13 170	11 340	10 550
	(year)	1988	1979	1981	1983	1983	1987	1987	1987	1968	1987	1974	1965
Runoff	Avg	26	21	22	16	11	7	5	5	5	9	13	19
	Low	6	4	7	5	4	2	1	1	2	2	3	3
	High	75	54	45	42	34	27	14	29	22	61	51	49
Rainfall	Avg	50	33	48	44	49	53	47	52	51	52	59	51
	Low	14	13	12	11	12	10	8	11	1	3	20	13
	High	125	63	93	99	100	132	93	105	118	128	155	107

Summary statistics

	For 1989		For record preceding 1989		1989 As % of pre-1989
Mean flow (m^3s^{-1})	2 674		2 957		90
Lowest yearly mean			1 428	1973	
Highest yearly mean			5 119	1987	
Lowest monthly mean	0 790	Sep	0 191	Jul 1976	
Highest monthly mean	6 044	Mar	16 080	Jan 1988	
Lowest daily mean	0 627	25 Sep	0 094	9 Jul 1976	
Highest daily mean	23 830	18 Mar	50 280	12 Oct 1987	
Peak	28 050	18 Mar	91 000	17 Sep 1968	
10% exceedance	5 075		6 450		79
50% exceedance	1 846		1 694		109
95% exceedance	0 757		0 523		145
Annual total (million cu m)	84 33		93 32		90
Annual runoff (mm)	146		161		90
Annual rainfall (mm)	510		589		87
[1941-70 rainfall average (mm)			598]		

Factors affecting flow regime

- Abstraction for public water supplies
- Flow reduced by industrial and/or agricultural abstractions
- Augmentation from surface water and/or groundwater

Station and catchment description
Twin-trapezoidal flume, throat tapping Spillway channel with weir constructed in 12/85 takes some flow above 1 45m. Bypassing also occurs over opposite bank above 1.85m More bypassing possible from 0.5km u/s during extreme events Naturalised flows to 9/76 Occasional high peaks due to gate action. Flow augmented by intermittent pumping from Ely/Ouse Transfer Scheme and occasional SAGS borehole pumping Mainly rural catchment Chalk outcrops in N, London Clay in S, all covered by semi-pervious Boulder Clay

Figure 5.1 Stour at Langham daily mean flows and hydrological summary from "Hydrological Data UK : 1989"

Bure at Horstead Mill — 1989

Harmonised monitoring station number : 05 722
Measuring authority : NRA-A NGR : 63 (TG) 267 198

Flow measurement station 034003 - Ingworth
C A (km²) : 164.7 NGR : 63 (TG) 192 296

Determinand	Units	1989 Samples	1989 Mean	1989 Max.	1989 Date	1989 Min	1989 Date	Period of record 1975 - 1988 Mean	Percentiles 5%	Percentiles 50%	Percentiles 95%	Quarterly averages J-M	Quarterly averages A-J	Quarterly averages J-S	Quarterly averages O-D
Temperature	°C	31	11.5	21.0	07/08	3.4	04/12	10.8	3.0	10.0	19.2	5.6	12.3	16.7	8.2
pH	pH units	34	8.0	8.5	21/08	7.3	18/09	7.8	7.3	7.8	8.2	7.7	7.8	7.9	7.7
Conductivity	µS/cm	25	774	835	13/11	717	24/07	728	610	730	828	742	702	725	744
Suspended solids	mg/l	25 (1)	3.6	13.5	26/01	0.0	24/07	8.3	1.2	5.2	23.9	25.8	6.4	4.2	6.2
BOD (inhibited)	mg/l O	34 (5)	1.5	3.4	22/05	<1.0	10/07	1.8	0.5	1.7	3.2	1.8	2.2	1.8	1.3
Ammoniacal nitrogen	mg/l N	33 (7)	0.116	1.000	20/11	<0.020		0.15	0.01	0.10	0.51	0.25	0.10	0.09	0.14
Nitrite	mg/l N	16	0.062	0.100	22/05	0.030	21/08	0.07	0.01	0.06	0.20	0.07	0.06	0.10	0.08
Nitrate	mg/l N	33	5.47	7.50	28/01	3.30	24/07	5.9	3.5	5.8	9.0	7.7	5.9	4.6	5.9
Chloride	mg/l Cl	33	60.3	127.0	11/12	25.0	21/02	57.9	47.5	56.0	74.7	60.3	55.6	56.1	60.1
Total alkalinity	mg/l $CaCO_3$	23	207.9	252.0	11/12	180.0	07/02	220.7	177.6	218.0	258.8	223.6	208.2	219.9	240.7
Silica	mg/l SiO_2	20	7.19	11.83	23/10	1.66	24/07	8.07	3.40	8.30	13.06	8.76	5.01	6.85	10.04
Sulphate	mg/l SO_4	23	84.4	111.0	20/11	65.9	21/02	80.0	54.9	79.0	112.0	83.9	82.9	71.8	85.4
Calcium	mg/l Ca	13	117.1	132.0	20/11	106.0	24/07	118.0	90.7	117.0	143.0	119.3	117.2	114.5	121.2
Magnesium	mg/l Mg	13	8.5	11.6	21/02	6.9	24/04	7.3	4.8	7.5	9.3	7.5	7.7	7.1	7.2
Potassium	mg/l K	13	4.4	8.4	25/09	3.2	24/07	4.0	2.4	4.0	5.6	4.2	3.7	3.9	4.5
Sodium	mg/l Na	13	27.6	29.2	24/07	25.7	21/02	30.0	20.0	27.7	48.0	30.3	29.5	29.6	29.9

Stour at Langham — 1989

Harmonised monitoring station number . 05 810
Measuring authority : NRA-A NGR : 62 (TM) 026 345

Flow measurement station 036006 - Langham
C.A (km²) : 578.0 NGR : 62 (TM) 020 344

Determinand	Units	1989 Samples	1989 Mean	1989 Max.	1989 Date	1989 Min	1989 Date	Period of record 1974 - 1988 Mean	Percentiles 5%	Percentiles 50%	Percentiles 95%	Quarterly averages J-M	Quarterly averages A J	Quarterly averages J-S	Quarterly averages O D
Temperature	°C	48	11.7	21.5	27/07	2.0	30/11	11.2	3.0	11.0	20.0	4.9	13.5	16.9	8.3
pH	pH units	49	8.3	9.0	03/08	7.8	07/12	8.2	7.8	8.2	8.9	8.1	8.4	8.3	8.1
Conductivity	µS/cm	30	871	1011	26/10	716	06/07	906	730	910	1100	927	879	883	946
Suspended solids	mg/l	30 (1)	11.6	122.0	21/12	<0.5	19/10	16.7	3.0	10.0	49.7	19.0	19.8	11.4	15.6
Dissolved oxygen	mg/l O	46	10.77	14.20	16/02	7.40	22/06	10.9	7.5	10.9	14.0	12.3	11.6	9.3	10.5
BOD (inhibited)	mg/l O	48	2.6	7.8	11/05	0.7	19/10	3.1	1.1	2.3	9.6	2.3	5.4	2.6	2.2
Ammoniacal nitrogen	mg/l N	49 (11)	0.088	0.600	06/04	<0.020	12/10	0.13	0.02	0.08	0.38	0.21	0.08	0.08	0.14
Nitrite	mg/l N	15 (2)	0.047	0.100	13/04	<0.004	03/08	0.08	0.02	0.07	0.16	0.08	0.09	0.04	0.09
Nitrate	mg/l N	49	6.33	15.10	28/12	0.50	07/12	8.3	2.1	7.6	16.0	12.5	7.9	4.4	8.8
Chloride	mg/l Cl	48	72.0	293.0	14/12	35.5	06/04	65.6	38.2	64.0	96.8	55.7	61.2	72.8	70.2
Total alkalinity	mg/l $CaCO_3$	21	253.0	285.0	15/05	194.0	21/12	243.3	191.5	250.0	280.0	243.8	242.4	249.4	250.3
Silica	mg/l SiO_2	19	10.57	46.10	15/05	3.70	12/06	7.47	0.20	7.95	13.00	7.51	3.40	8.39	10.27
Sulphate	mg/l SO_4	21	90.7	119.0	18/09	71.6	03/08	100.4	70.0	98.0	140.0	13.7	13.9	97.0	4.9
Calcium	mg/l Ca	11	128.8	145.0	30/03	108.0	17/07	134.2	93.0	139.0	167.5	148.3	134.9	119.7	140.3
Magnesium	mg/l Mg	11	8.3	11.8	18/09	6.2	02/02	9.9	5.0	8.4	22.0	7.8	8.9	9.7	8.9
Potassium	mg/l K	10	8.6	12.7	19/10	5.6	02/02	7.4	3.5	7.2	12.0	5.7	6.9	7.9	9.1
Sodium	mg/l Na	10	47.5	67.0	19/10	31.0	02/02	42.5	20.0	40.0	70.6	32.3	39.7	50.0	49.1

Thames at Teddington Weir — 1989

Harmonised monitoring station number : 06 010
Measuring authority . NRA-T NGR · 51 (TQ) 171 714

Flow measurement station . 039001 - Kingston
C A (km²) 9948.0 NGR . 51 (TQ) 177 698

Determinand	Units	1989 Samples	1989 Mean	1989 Max	1989 Date	1989 Min	1989 Date	Period of record 1974 - 1988 Mean	Percentiles 5%	Percentiles 50%	Percentiles 95%	Quarterly averages J M	Quarterly averages A J	Quarterly averages J S	Quarterly averages O D
Temperature	C	67	18.1	26.0	24/07	5.0	19/01	11.6	4.0	11.5	20.0	5.8	13.8	18.1	9.7
pH	pH units	70	7.7	9.0	18/05	7.1	05/10	8.1	7.6	8.0	8.8	8.0	8.3	8.0	7.8
Suspended solids	mg/l	56	9.3	62.5	05/10	1.5	20/07	22.1	5.0	14.3	76.9	29.2	22.2	13.4	24.4
Dissolved oxygen	mg/l O	69	7.97	11.75	19/01	5.30	12/09	10.2	7.1	10.2	13.2	11.3	10.8	8.6	10.0
BOD (inhibited)	mg/l O	69 (2)	2.4	7.1	16/05	<1.0	15/09	2.9	1.1	2.4	6.7	2.2	4.3	3.0	2.1
Ammoniacal nitrogen	mg/l N	70 (2)	0.422	2.300	15/08	<0.050	16/05	0.32	0.01	0.23	0.89	0.34	0.19	0.38	0.36
Nitrite	mg/l N	9	0.172	0.300	21/11	0.082	16/05	0.11	0.06	0.10	0.20	0.10	0.10	0.11	0.13
Nitrate	mg/l N	70	6.95	9.40	05/12	4.72	16/05	7.4	5.4	7.1	10.3	8.2	6.6	6.6	7.6
Chloride	mg/l Cl	70	58.1	69.0	12/10	39.0	27/04	41.2	30.0	40.0	57.0	39.7	38.4	45.0	42.4
Orthophosphate	mg/l P	62	3.330	4.600	12/10	0.910	09/03	1.30	0.39	1.08	2.72	0.78	1.06	1.99	1.41
Silica	mg/l SiO_2	32	12.91	21.50	05/09	0.30	16/05	10.49	2.90	11.50	14.60	11.03	6.77	11.16	13.23

Lee at Waterhall — 1989

Harmonised monitoring station number 06 101
Measuring authority NRA-T NGR · 52 (TL) 299 099

Flow measurement station · 038018 - Water Hall
C A (km²) . 150.0 NGR . 52 (TL) 299 099

Determinand	Units	1989 Samples	1989 Mean	1989 Max	1989 Date	1989 Min	1989 Date	Period of record 1975 - 1988 Mean	Percentiles 5%	Percentiles 50%	Percentiles 95%	Quarterly averages J M	Quarterly averages A J	Quarterly averages J S	Quarterly averages O-D
Temperature	C	24	11.9	21.5	20/08	4.0	28/02	12.0	4.5	12.0	19.9	6.9	13.6	16.7	9.3
pH	pH units	24	7.9	8.3	01/09	7.3	02/08	8.0	7.6	8.0	8.4	8.0	8.1	8.1	7.8
Conductivity	µS/cm	10	800	974	11/10	545	15/08	808	623	779	1066	885	775	784	814
Suspended solids	mg/l	11	21.0	78.0	25/04	3.0	10/12	15.1	3.0	11.0	42.7	15.9	12.3	13.6	14.9
BOD (inhibited)	mg/l O	24	2.9	5.2	15/08	1.4	04/01	2.7	1.3	2.4	4.6	2.7	3.1	2.2	2.4
Dissolved organic carbon	mg/l O	10	16.0	21.3	25/04	3.5	31/01	13.0	2.6	7.2	55.4	16.0	17.4	6.4	17.1
Ammoniacal nitrogen	mg/l N	24 (4)	0.349	3.900	07/11	<0.050	30/03	0.23	0.05	0.11	0.78	0.37	0.09	0.09	0.32
Nitrite	mg/l N	11	0.191	0.900	07/11	0.068	30/03	0.14	0.05	0.11	0.29	0.11	0.12	0.34	0.18
Nitrate	mg/l N	24	11.96	17.50	10/12	6.03	15/08	11.6	7.7	11.4	16.3	12.7	12.2	12.3	13.7
Chloride	mg/l Cl	24	88.3	132.0	24/05	54.0	15/08	69.4	44.4	67.0	105.0	86.5	65.1	75.6	74.9
Total alkalinity	mg/l $CaCO_3$	10	199.3	241.0	04/01	124.0	15/08	211.0	129.1	223.0	253.9	203.4	217.6	[illegible]	200.3
Orthophosphate	mg/l P	11	3.546	4.800	04/01	1.820	28/02	2.33	1.10	2.34	3.50	2.10	2.24	2.51	2.47
Sulphate	mg/l SO_4	9	77.4	104.0	11/10	52.0	15/08	79.6	54.3	77.5	112.7	77.8	79.6	77.6	81.5
Calcium	mg/l Ca	10	108.0	138.0	04/01	72.0	15/08	118.8	93.1	118.0	143.0	119.4	120.4	[illegible]	115.9
Magnesium	mg/l Mg	10	3.9	5.0	25/04	2.4	15/08	4.0	3.1	3.9	4.9	4.2	3.9	3.3	3.9
Potassium	mg/l K	10	10.5	15.5	11/10	7.5	28/02	8.4	5.6	8.0	14.2	7.7	7.1	8.9	9.8
Sodium	mg/l Na	10	74.2	108.2	07/11	42.1	15/08	63.9	35.0	[illegible]	[illegible]	63.9	63.5	[illegible]	58.7

Figure 5.2 Statistical summaries of river quality data from "Hydrological Data UK : 1989"

the Survey somewhat subjective. Nevertheless, River Pollution Surveys give a valuable 'snapshot' of the state of the nation's rivers.

The Harmonised Monitoring Scheme was initiated in 1974 to complement the River Pollution Survey information base by providing more quantitative information on the condition of rivers. From the outset, one of the specific objectives of the Harmonised Monitoring Scheme was

> "... to enable an assessment to be made, in connection with international obligations, of the materials carried down rivers into the sea; in due course, this will be supplemented by estimates of other polluting loads entering estuaries and the sea ..." (Simpson, 1978).

A further objective of the Harmonised Monitoring Scheme is

> "To enable long-term trends in river water quality to be identified." (Simpson, 1980).

Recognising the importance of such a database for environmental management generally, it was intended to publish a summary of information each year. However, publication of river quality data summaries from the Harmonised Monitoring Scheme database comparable to the Yearbook series for river flow data has yet to be achieved (but see the next paragraph). Mean annual concentrations for a selection of determinands and rivers have been published in the "Digest of Environmental Pollution Statistics" (e.g DOE, 1978; 1980) but this practice appears to have been discontinued. A recent "Digest" contains estimates of United Kingdom heavy metal inputs by rivers (and via other routes) to the North Sea (DOE, 1990). The source of this information is given as the Second International Conference on the Protection of the North Sea, 1987.

Understandably perhaps, a large part of the effort expended on the Harmonised Monitoring Scheme to date appears to have been in the area of Analytical Data Quality Control. It is essential, for comparisons between sites at a particular time and for data analysis of time series for a particular site, that concentrations are measured consistently over time by the different laboratories involved. The Water Research Centre co-ordinates the water industry effort in the area of Analytical Data Quality Control (Cheeseman and Wilson, 1978) and the Committee for Analytical Quality Control (Harmonised Monitoring) agrees and promulgates sampling and analytical procedures for particular determinands (e.g. Committee for Analytical Quality Control, 1984).

It appears that attention was given initally to the adequacy of the measurement network for assessing average concentrations, rather than river loads. A broad objective of the Scheme was that the average concentration for each determinand calculated from sample values should have a 95% probability of being within 20% of the true value (Simpson, 1980). Unfortunately, no similar broad objective appears to have been set for assessing the efficacy of the Harmonised Monitoring Scheme for estimating river loads.

Joint exploitation and assessment of the national river flow and quality databases for river loads estimation has been minimal and, where this has been attempted, the best methods may not have been used (see Chapter 3 for discussion of some of the ad hoc studies where Harmonised Monitoring and Surface Water Archive data have been employed for river loads estimation). The databases are managed quite independently of each other, though exchanges of data and information between the databases for a variety of purposes are becoming increasingly frequent. For example, data exchanges have been made for estimation, by Her Majesty's Inspectorate of Pollution, of nutrient loads to the North Sea (B.G. Goldstone, personal communication), and for this study.

5.4 THE WATER ARCHIVE AND CHANGING INFORMATION TECHNOLOGY

The Water Archive is a computer-based data and information storage system developed in the mid 1970s jointly by the Department of the Environment (Water Data Unit) and a consortium of Regional Water Authorities. It represents the first attempt in the United Kingdom to provide a single database system capable of handling any water-related data and information required by Government Departments, or by the Water Industry, for a wide spectrum of applications. In principle, therefore, the Water Archive can handle both river flow and concentration data and thereby ease the task of estimating river loads. However, whilst several Regions of the National Rivers Authority continue to use the Water Archive structure for holding river flow and quality data, it appears to have not had a great impact in the area of river loads information processing, either regionally or at the national level.

Information technology, in the form of Geographical Information Systems (GIS), is moving forward apace and the next section considers briefly if there might be implications for GIS in future joint exploitation of river flow and concentration data for river loads estimation and related data presentation.

5.5 GEOGRAPHICAL INFORMATION SYSTEMS AND *WIS*

Geographical Information Systems mark the next stage in the development of computer databases for environmental management. A full account of the nature of GIS, or of the exciting progress being made with this technology, cannot be given here. The purpose of this section is to outline some of the potential benefits of GIS in the context of manipulating river loads data and information.

A GIS can store a digitised 'image' of the land surface topography (including bathymetry) and register with this the position or nature of any other geographical feature such as rivers, coastlines, field boundaries, soil types, discharge points, roads, etc. Data input to a GIS from remote sensing devices is possible. The Water Information System (WIS) is a GIS being developed by the Institute of Hydrology with the support of International Computers Limited which will allow advantage to be taken of time series data within a GIS framework. Thus it will be possible to store details of the variation in time and space of such things as rainfall, streamflow, point-discharges of contaminants to rivers, river quality, crops, fertiliser applications, etc., and to display these types of information in a pictorial format. User-interaction with the pictorial output from WIS will facilitate data retrieval and analysis. The market lead of WIS is in its ability to bring to the screen the river network, maps of catchment characteristics and time series data of any type with equal facility at an interactive speed; normal GIS cannot handle the huge databases required for water industry purposes. The interface between WIS and mathematical models of environmental systems opens up new areas for research and use of data and information for environmental management purposes.

It is not clear yet just how this technology will impact on river loads data and information processing but it is clear that any GIS will not lessen the need for careful design and operation of river quality monitoring programmes and subsequent calculation of loads employing appropriate algorithms, though the latter could be incorporated into WIS.

Perhaps the greatest contribution WIS could make in the context of river loads is in 'environmental auditing'. Given the power of WIS to record and manipulate environmental data and information in four dimensions (the three dimensions of space, and time), and the potential for exploiting this power using mathematical models, it should be possible, theoretically, to keep an inventory of water and chemicals and predict their mass movements within catchments and across the tidal limits of estuaries. Direct discharges to estuaries could also be audited using WIS. The success of the whole idea is, of course, subject to the availability of data with which to create, test and operate the system.

6 Summary and Conclusions

For meaningful comparison of load estimates it is necessary to know the errors associated with specific estimates. In practice, a range of different load estimation procedures are adopted in response to the variable amount and quality of available flow and concentration data. Indeed, there is no single combination of measurement strategy and estimation algorithm which is suitable for the wide range of hydrological and hydrochemical response types exhibited by United Kingdom rivers. Different sampling frequencies and estimation algorithms lead to load estimates of varying accuracy and precision. However, qualification of load estimates in terms of numerical accuracy and precision is rare due to the difficulties of establishing statistical uncertainty from sparse data; because of the high marginal costs involved, volumes of concentration data seldom match those for river flows. Whenever a useful mathematical relationship (model) between flow and concentration can be established, the information in databases of 'continuous' river flow data can be exploited to estimate concentrations for periods between samples, thereby leading to improved river load estimates.

Methods

The literature review (Chapter 2) reveals that existing methods of river load estimation are many and extremely varied in detail. Some methods are simplistic because even when flow and concentration data are available at low frequency (e.g. monthly) they employ, for example, the arithmetic mean of the products of spot-sampled flows and concentrations. Except when the variability of flow and concentration are both small, such methods can result in load estimates which are biased or imprecise (or both) to an unacceptable level. The terms 'precision' and 'bias' are defined in the report. Refinements of simplistic methods attempt to reduce bias and improve precision by employing carefully designed sampling strategies where the frequency of samples spaced regularly in time is varied between 'strata' defined in terms of flow thresholds or seasons (or both).

In general, estimates based solely on infrequent spot-sample flow and concentration data are prone to large errors. Often, however, there is available a record of 'continuous' flow from which a 'continuous' concentration record may be estimated using a mathematical relationship (or model) between the two variables. And better relationships between flow and concentration lead to better load estimates. It is common practice to employ the linear regression model to relate (the logarithms of) flow and concentration for load estimation purposes but, because of the complexities of the physical and chemical processes which control flow - concentration dynamics, the uncertainty introduced by such models can be large and load estimates, therefore, can be rather imprecise. Furthermore, unless correction factors are applied the resultant estimates can be heavily biased.

The inability of the simple linear regression model to account for any part of the commonly observed hysteresis in flow - concentration relationships is one of its major weaknesses. The transfer function model is demonstrated to be able to account for at least **some** hysteresis; it is strongly recommended that the transfer function model be more widely considered for use in extrapolation load estimation procedures.

Most stream water sampling is undertaken manually and does not take into account the level and rate of change in flow which dominate the variation in flux of constituents. Manual sampling often leads to biased estimates. Where sampling can be controlled continually by a computer linked to a streamflow measuring device, it is possible, by taking samples at frequencies proportional to flow, or according to a scheme of probability sampling, to effectively eliminate bias. In the case of probability sampling, estimates of known precision can be obtained for specific periods of record. Potentially, such methods could lead to significant improvements in river load estimates, and greater understanding of flow-concentration dynamic behaviour generally.

Clearly, the costs of measuring river loads vary according to the required quality (accuracy and precision) in the estimate and are a function of the provisions made to obtain that quality. For a given determinand, there is no single simple manual sampling strategy which will give estimates of known accuracy and precision for the wide range of hydrological conditions encountered in the United Kingdom. Furthermore, for those methods which employ regular sampling irrespective of the hydrological conditions during the period of load estimation (e.g. a year), bias and precision can vary greatly between (annual) estimates according to differences between periods in the variation of flow and concentration. In general, load estimation is more prone to bias and imprecision in situations (i.e. the same period at different sites or the same site for different periods) where flow and concentration are highly variable.

Under the same hydrological conditions at a given site, loads estimated by a given method for determinands which exhibit a high coefficient of variation (e.g. suspended sediment) will tend to be the most imprecise.

Fundamental problems in assessing load estimation methods

A problem central to assessing methods of load estimation is the paucity of long records of concurrent high-frequency flow and concentration data from which 'true' load can be calculated for comparison purposes. Where such records do exist (e.g. for suspended sediment in certain southwest England rivers) it has been shown (Walling and Webb, personal communication) that an annual load (1979) computed (a) on the basis of 16 regularly-spaced samples (fairly typical for the Harmonised Monitoring network) and (b) by the preferred Paris Commission algorithm, can be biased by more than -50% (a systematic underestimation). An indication of the low level of precision in annual suspended sediment load estimates can be appreciated from a histogram showing the spread of values obtained from a scheme of load estimate replication (as discussed in the report).

In the context of load estimation, suspended sediment is a worst-case determinand because typically it increases in concentration with increasing flow by 'orders of magnitude' and therefore has a high coefficient of variation. It can be confidently predicted that mass loads of determinands in solution which exhibit lower coefficients of variation will, in general, be better estimated than suspended sediment loads. However, suspended sediment is a key variable because significant amounts of certain heavy metals and organochloride residues can be transported with the sediment as an adsorbed phase. Problems associated with measuring and assessing the errors in bed-load transport have not been considered in this report.

A systematic computational framework for assessing load estimation methods

The problem of a general lack of suitable data from which to compute 'true' loads of different determinands has been circumvented in this study by adopting a synthetic data approach. A prototype computer program for Simulation and Methods Investigation of Load Estimates for Rivers (SMILER) has been introduced. SMILER is flexible enough to make available for analysis long time series of data which typify almost any combination of hydrological and simple concentration responses likely to be encountered in the United Kingdom.

The rationale behind SMILER is that it uses any available 'continuous' data and other information on flow and concentration for the site of interest and generates synthetic concentration data (synthetic flow data also if necessary) which may be additionally required to establish a 'true' load. The synthetic responses are not required to be faithful representations of reality (though this is a goal to be kept in mind) - merely typical - for error assessment purposes.

Daily mean streamflow data for any available period of record can be input to SMILER from the Surface Water Archive maintained by the Institute of Hydrology. Information about the mean level and variability of concentration for a wide range of determinands since about 1974 (based on spot-sample data) can be accessed from the Harmonised Monitoring Scheme database maintained by the Department of the Environment. In principle, SMILER could receive flow and concentration data and information from other databases. Ideally, and especially for flashy rivers, the flow data should be at a finer time scale than daily (see later).

Demonstration time series for flow and concentration have been presented, and analysed by SMILER, to illustrate the complexity of the relationships between errors in load estimates and (a) types of hydrological and concentration response, (b) period of estimation and level of hydrochemical activity in that period and (c) estimation algorithm. The current version of SMILER can compare three methods of estimation using fixed frequency observations: the two methods recommended by the Paris Commission for monitoring North Sea inputs from rivers; and the Beale Ratio estimator employed for monitoring river inputs to the Great Lakes, North America. Several other methods are discussed in the report and selected additional methods will be incorporated into future versions of SMILER. Likewise, future versions of SMILER will be enhanced with improved statistical techniques.

An example — errors in loads of nitrate in solution

A particular application of SMILER, namely estimating the errors in nitrate loads carried in solution by the Stour at Langham, East Anglia, has been presented. In contrast to the massive errors which can arise in estimates of suspended sediment loads for certain rivers in southwest England, the SMILER estimated errors in Stour nitrate loads are modest. At this particular site it appears from the preliminary analysis presented here that annual load of nitrate in solution can be estimated by the preferred Paris Commission method (involving flow-weighted mean concentration), and about 12 samples taken regularly throughout the year, with a precision of about 2% in a 'low load' year and about 7% in a 'high load' year. Bias in both cases is probably less than +/-1%. **It must be stressed,** however, that these results assume zero measurement error and reflect, therefore, only the errors due to (a) 30 day sampling interval, (b) the representativeness of the synthetic concentration data and (c) a particular load estimation method. **The values of precision given above should not be used, in isolation, to assist in the design of sampling strategies**; the results are preliminary and further work is required for a range of determinand and site combinations. SMILER does incorporate facilities to simulate measurement errors in both streamflow and concentration but this is not discussed in the report. Clearly, the errors given above are minima. For any particular application of SMILER, information about streamflow and concentration measurement errors for the site and determinand in question should be taken into consideration.

Scope for hydrograph separation and mathematical hydrologic mixing models

Particular aspects of scientific hydrology which have considerable potential for application in the context of river loads estimation include hydrograph separation into component flows, and mathematical modelling generally. It can be envisaged that streamflow at any time comprises components which have different but distinctive chemical characteristics. For example, low streamflows occurring some considerable time after rainfall may be expected to be mainly from sources of stored water in the catchment which have a distinctive chemical signature. At, and near, peak flows, however, a significant proportion of streamflow is probably water added to the catchment as recent rainfall which has a quite different chemical signature. The higher the peak flow, the greater proportion of it is likely to be water from recent rainfall. Point-source contaminant inputs obviously complicate the situation and have to be considered separately.

Simplistically, therefore, we may assume that streamflow at any time, at least in relatively natural situations, is a mixture of flow components from sources with different but (relatively) fixed concentrations. It might be expected, therefore, that estimates of concentration based on knowledge of the proportions of flow components at any time would be better than estimates of concentration from total streamflow. Field studies and mathematical models can be employed to investigate the complex detail of the timing and proportions of component flow mixing. Examples have been given of different methods of separating hydrographs into component flows and, clearly, progress and developments in this area of research (and related mathematical modelling) should be closely monitored to assess the utility of the approach for river load estimation.

Databases for river load estimation

At the national level the Surface Water Archive and the Harmonised Monitoring Scheme database are the major sources of data and information for river flow and concentrations respectively. The limited amount of flow data in the Harmonised Monitoring database (either instantaneous or daily mean flow rates corresponding to sample time or date respectively) is sufficient only for crude estimation of loads carried by rivers. The Harmonised Monitoring concentration data, coupled with continuous records of daily mean flows from the Surface Water Archive, however, offers considerable scope for retrospective river load estimation, as demonstrated in the report with the example for nitrate loads based on synthetic concentration data approximating to those expected in a rural lowland catchment. Future arrangements for managing these databases should encourage the joint exploitation of their information contents.

It should be recognised, however, that use of Surface Water Archive daily mean flows and Harmonised Monitoring sample (instantaneous) concentrations for load estimation assumes that both values are representative for the sample time in question. This assumption will break down for sites and occasions where flow (and probably, therefore, concentration) can vary significantly within a day. For high-quality load estimates it will be necessary to consider data of higher frequency. The daily mean flows in the Surface Water Archive are returned by the measuring authorities (e.g. regional Divisions of the National Rivers Authority and the River Purification Boards) and are based typically on stage measurements at 15 minute intervals. Most measuring authorities have computer archives to accommodate short interval level and/or flow data. In principle, therefore, and subject to there being a stable identifiable relationship between the two variables over the period in question, concentration could be estimated from the 15 minute interval flow data for load estimation purposes. This will require relationships (or models) between flow and concentration which capture the essential dynamic behaviour of the constituent in question. Unfortunately, the relatively low frequency of observations in the Harmonised Monitoring database will render that source of information inadequate for such purposes. It will be necessary, therefore, to investigate the availability of high frequency concentration data from other databases maintained by the measuring authorities. Given that continual intensive measurement of concentration is prohibitively costly, except for those determinands which can be measured by ion-specific electrodes or some other automatic device, there should be intermittent periods of intensive monitoring at the site of interest to assess any dynamic flow - concentration behaviour and to calibrate (and periodically validate/update) suitable mathematical models.

The measuring authorities collect a far larger volume of river chemistry data than is returned to the Harmonised Monitoring database but these sources of data and information have not been inspected for this report. Their suitability for load estimation purposes should be assessed at an early date..

7 Recommendations and suggestions for further work

A full appreciation of several of the points referred to below cannot be gained without reading the relevant Chapter of the report.

Methods

Meaningful comparison of river load estimates is greatly enhanced by knowledge of the errors involved. Suppliers of such data and information should be strongly encouraged to provide a numerical level of precision (at an agreed confidence level) with individual river load estimates. Every effort should be made to minimise bias in river load estimates by adopting an appropriate sampling strategy and a suitable estimation algorithm.

In all situations where there is a 'continuous' flow record it should be employed – to give a better river load estimate than if just flows at sample times are used. Estimation methods based only on flow and concentrations at sample times ignore the valuable information in the 'continuous' flow record and therefore give inferior results.

A mathematical relationship (or model) between flow and concentration should always be sought and thereby a 'continuous' concentration record derived from a 'continuous' flow record. Wherever possible a programme of intensive sampling should be operated until a best relationship (or model) can be derived. It may be necessary to repeat relatively short periods of intensive sampling from time to time to check the stability of the relationship (or model).

In many cases a component of the scatter commonly observed in plots of flow against concentration could be due to hysteresis. The linear regression model cannot simulate hysteresis but a simple 'black box' model which can reproduce some degree of observed hysteretic behaviour is the transfer function model. The transfer function model, therefore, deserves more widespread use in extrapolation methods of river loads estimation. Models based on relevant physical and chemical processes may also be useful in extrapolation methods of load estimation, though in most cases they require more detailed input data than that required by the transfer function approach.

Given that a large proportion of the total mass load of many river constituents is transported during fairly short periods of high flow, it is evident that a combination of regular (but infrequent) sampling, and an interpolation estimation algorithm which does not use the information in a continuous record of flow, is likely to underestimate river loads. It has been shown by other investigators that for the Thames at Kingston an automatic system which takes samples at variable time intervals (depending on flow) reduces bias and improves precision in river load estimates. Similar test facilities should be deployed at other sites covering the range of hydrological regimes experienced in the United Kingdom. Particular attention should be paid to assessing for

United Kingdom rivers the efficacy of SALT (Thomas, 1986) – a sampling method which yields minimally biased loads of known precision for individual load estimates. The transfer function model, rather than the linear regression model, should be considered for generating the 'auxiliary variable' in SALT.

A systematic computational framework for assessing load estimation methods

A prototype computer program "Simulation and Methods Investigation of Load Estimates for Rivers" (SMILER) has been developed as part of the current study. Wherever hydrochemical budgets are required (including small research catchments) SMILER can be of assistance by providing information on the likely errors in river loads under particular hydrological conditions. It is recommended that SMILER be applied to strategic sites in the United Kingdom where errors in river load estimates are needed to assist with the assessment of the environmental quality control of the North Sea (and other bodies of water, as required).

Hydrograph separation and mathematical mixing models

Recent developments in hydrograph separation by chemical tracer methods and time series rainfall - streamflow modelling give new insights into the mixing dynamics of component flows which may have different but distinctive chemical compositions. Such developments, referred to in the report, could be extremely useful for extrapolation methods of river load estimation, and their continued development should, therefore, be supported.

Databases for river load estimation

Historically, the national databases for river quantity and quality have been managed independently and, although effective mergers of data and information from each are becoming more common, their separateness does not encourage their joint exploitation for river loads estimation. There is scope for estimation of historic river loads from river flow data in the Surface Water Archive and concentration data in the Harmonised Monitoring Scheme database. Similarly, there is scope using these databases for deriving 'best' estimates of river load inputs routinely, for example, to the North Sea, Irish Sea and English Channel. Careful thought should be given to data acquisition, archiving and data analysis procedures to facilitate improved estimates of mass flows.

Consideration is required also of the potential for developments in information technology (including GIS – Geographical Information Systems) in the context of river loads data processing and manipulation for environmental management purposes. Provision of river loads data is demanding in that it requires access to quantity **and** quality databases (or the use of one that handles both types of data). The ability to derive and manipulate river loads data is, therefore, an important operational need which should be specified for future database and information systems. Ideally, such systems should be able to transpose data to reaches where load estimates are needed, but where measurements have not been made.

The National Rivers Authority, the River Purification Boards, and other organisations, hold large volumes of data and information not in the Surface Water Archive or the Harmonised Monitoring Scheme database. There is a pressing need to identify such additional datasets and to explore ways of bringing them together for river load estimation.

Acknowledgements

The study was funded initially by the Department of the Environment (DOE). This report is published in accordance with the wishes of DOE to fulfil their agreement as a Member of the Foundation for Water Research (FWR) to disseminate research reports to other Members. Valuable discussions took place with Dr R.J. Otter, Mr K.P.D. Close, Mr D.E. Edwards and Mrs B.G. Goldstone of DOE, and Dr C. Reid (on secondment to DOE from the Plymouth Marine Laboratory of the Natural Environment Research Council). Any views expressed in the report are those of the author. Data from the Harmonised Monitoring Scheme were made available by Her Majesty's Inspectorate of Pollution.

At an early stage of the project, comments and data were gratefully received from officers of the National Rivers Authority; thanks are due to Dr C. Pattinson and Dr S.C. Bird of the Welsh Region, and Mr T. Reader and colleagues of Thames Region.

Professor D.E. Walling and Dr B.W. Webb of Exeter University made useful comments at an early stage of the project and, at the request of the author, readily agreed to undertake further limited analyses of one of their datasets.

Section 2.6 on Real-time Updated Stratified Sampling (RUSS) was a late addition to the report. Useful discussions were held with Dr D.H. Burn of the University of Manitoba, Canada (during his visit to the Institute of Hydrology in the second half of 1991) on RUSS and river loads estimation generally.

References

Anon. 1984. Estimation of river composition and riverine influxes of chemicals to the marine environment. Report of the Advisory Committee on Marine Pollution, ICES 132, 48-54.

Bewers, J.M. & Duinker, J.C. 1982. Methods of assessing gross riverine discharges of trace metals and organohalogens into the marine environment. In: Report of the International Council for the Exploration of the Seas, Cooperative Research Report No. 120, 60-72.

Birtles, A.B. 1977. River water quality models based on stream hydrograph components. Central Water Planning Unit, Technical Note No. 23.

Birtles, A.B. 1978. Identification and separation of major base flow components from a stream hydrograph. *Wat. Resour. Res.*, 14, 791-803.

Box, G.E.P. & Jenkins, G.M. 1970. *Time Series Analysis, Forecasting and Control.* Holden - Day, San Francisco.

Burn, D.H. 1990. Real-time sampling strategies for estimating nutrient loadings. *J. Wat. Resour. Planning & Management,* 116, 727-741.

Chow, V.T. (Ed.) 1964. *Handbook of Applied Hydrology.* McGraw-Hill, New York.

Clarke, R.T. 1990. Bias and variance of some estimators of suspended sediment load. *Hydrol. Sci. J.*, 35, 253-261.

Cohn, A.T., DeLong, L.L., Gilroy, E.J., Hirsch, R.M. & Wells, D.K. 1989. Estimating constituent loads. *Wat. Resour. Res.*, 25, 937-942.

Cheeseman, R.V. & Wilson, A.L. 1978. Manual on analytical quality-control for the water industry. Water Research Centre, Technical Report TR 66.

DOE 1978. Digest of Environmental Pollution Statistics: No. 1, 1978. HMSO, London.

DOE 1980. Digest of Environmental Pollution Statistics: No. 2, 1979. HMSO, London.

DOE 1990. Digest of Environmental Protection and Water Statistics: No.12. 1989. HMSO, London.

Dolan, D.M., Yui, A.K. & Geist, R.D. 1981. Evaluation of river load estimation methods for total phosphorous. *J. Great Lakes Research,* 7, 207-214.

Edmunds, W.M. & Kinniburgh 1986. The susceptibility of UK groundwaters to acidic deposition. *J. Geological Society,* 143, 707-720.

Edwards, A.M.C. 1973. Dissolved load and tentative solute budgets of some Norfolk catchments. *J. Hydrol.*, 18, 201-217.

Edwards, R.W., Gee, A.S. & Stoner, J.H. (Eds) 1990. *Acid Waters in Wales,* Kluwer Academic Press.

Ellis, J.C. 1989. Handbook on the design and interpretation of monitoring programmes. Water Research Centre plc.

Ferguson, R.I. 1986a. River loads underestimated by rating curves. *Wat. Resour. Res.*, 22, 74-76.

Ferguson, R.I. 1986b. Reply. *Wat. Resour. Res.*, 22, 2123-2124.

Ferguson, R.I. 1987. Accuracy and precision of methods for estimating river loads. *Earth Surf. Processes,* 12, 95-104.

GESAMP 1987. Report of Working Group 22 on land to sea transport of pollutants. GESAMP Reports and Studies No. 32.

Grogan, W.C. 1984. Input of contaminants to the North Sea from the United Kingdom. Institute of Offshore Engineering, Heriot-Watt University, Report for the Department of the Environment.

Gurnell, A.M. & Fenn, C.R. 1984. Box-Jenkins transfer function models applied to suspended sediment concentration-discharge relationships in a proglacial stream. *Arctic & Alpine Research,* 16, 93-106.

Harrison, R.M., Thorogood, G. & Lacey, R.F. 1989. Comparison of load estimation from grab samples and continuous flow-proportional sampling. Report to the Department of the Environment by the Water Research Centre, Medmenham.

Herschy, R.W. (Ed) 1978. *Hydrometry.* John Wiley, Chichester.

Hill, J.M. 1984. Nitrates in surface waters: observations from some rivers in the Lee drainage basin. Water Research Centre, TR 203.

Howard, K. & Urquart, C. 1988. The quality of the Humber estuary 1987. Humber Estuary Committee Report.

Institute of Hydrology 1988. *Hydrological Data UK : Hydrometric Register and Statistics 1981-5.* Natural Environment Research Council.

Jakeman, A.J., Littlewood, I.G. & Whitehead, P.G. 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to small catchments. *J. Hydrol.*, 117, 275-300.

Jolly, P.K. 1986. Estimates of the loads of certain List I and List II substances discharged to the North Sea. Report to the Department of the Environment by the Water Research Centre, Medmenham.

Jolly, P.K. & Ellis, J.C. 1989. Monitoring requirements for estimating pollutant loads in rivers. Interim Report to the Department of the Environment by the Water Research Centre, Medmenham.

Koch, R.W. & Smillie, G.M. 1986. Bias in hydrologic prediction using log-transformed regression models. *Wat. Resour. Bull.*, 22, 717-723.

Langan, S.J. 1987. Episodic acidification of streams at Loch Dee, S.W. Scotland. *Trans. Royal Society of Edinburgh: Earth Sciences,* 78, 393-397.

Lees, M.L. 1989. Inland water surveying in the United Kingdom - a short history. *Hydrological Data UK: 1985 Yearbook.* Institute of Hydrology.

Littlewood, I.G. 1987. Streamflow - pH dynamics in small moorland and conifer afforested catchments in the Upper Tywi valley, Wales. *Proc. British Hydrological Society Symposium,* Hull, 23.1-23.12.

Littlewood, I.G. 1990. Estimating loads transported by rivers. Final Report, DOE Contract PECD 7/7/317.

Marsh, T.J. 1980. Towards a nitrate balance for England, and Wales. *Water Services,* 84, No. 1016, 601-606.

Moore, R.J. 1984. A dynamic model of basin sediment yield. *Wat. Resour. Res.*, 20, 89-103.

Norton, R.L. 1982. Assessment of pollution loads to the North Sea. Water Research Centre, Technical Report No. 182.

O'Donnell, A.R. & Mance, G. 1984. Estimation of the loads of some List I and II substances to United Kingdom tidal waters -- a comparison with previous estimates. *Water Pollution Control,* 554-561.

Pinder, G.F. & Jones, J.F. 1969. Determination of the ground-water component of peak discharge from the chemistry of runoff. *Wat. Resour. Res.*, 5, 438-445.

Richards, R.P. & Holloway, J. 1987. Monte Carlo studies of sampling strategies for estimating tributary loads. *Wat. Resour. Res.*, 23, 1939- 1948.

Rodda, J.C. & Jones, G.N. 1983. Preliminary estimates of loads carried by rivers to estuaries and coastal waters around Great Britain derived from the Harmonized Monitoring Scheme. *J. Instn Wat. Eng. Scient.*, 37, 529-539.

Simpson, E.A. 1978. The harmonization of the monitoring of the quality of inland fresh water. *J. Instn Wat. Engrs Scient.*, 32, 57-66.

Simpson, E.A. 1980. The harmonization of the monitoring of the quality of rivers in the United Kingdom. *Hydrol. Sci. Bull.*, 25, 13-23.

Thomas, R.B. 1983. Errors in estimating suspended sediment. In: Li, Rhu-Ming & Lagasse, P.F. (Eds) *Proc. D.B. Simons Symposium on Erosion and Sedimentation,* Fort Collins, Colorado, 27-29 July 1983.

Thomas, R.B. 1985. Estimating total suspended sediment yield with probability sampling. *Wat. Resour. Res.*, 21, 1381-1388.

Thomas, R.B. 1986. Calibrating SALT: a sampling scheme to improve estimates of suspended sediment yield. In: Monitoring to Detect Changes in Water Quality Series, *IAHS Publication No. 157,* 79-88.

Thomas, R.B. 1988a. Monitoring baseline suspended sediment in forested basins: the effects of sampling on sediment rating curves. *Hydrol. Sci. J.*, 33, 499-514.

Thomas, R.B. 1988b. Measuring sediment yields of storms using SALT. In: Sediment Budgets, *IAHS Publ. No. 174,* 315-323.

Thomas, R.B. 1989. Piecewise SALT sampling for estimating suspended sediment yields. Report PSW-114, Pacific Southwest Forest and Range Experimental Station, US Department of Agriculture.

Walling, D.E. 1974. Suspended sediment and solute yields from a small catchment prior to urbanisation. In: K.J. Gregory & D.E. Walling (eds), Fluvial processes in instrumented watersheds, Inst. Br. Geogr. Spec. Publ., 6, 169-192.

Walling, D.E. & Webb, B.W. 1981. The reliability of suspended load data. In: Erosion and Sediment Transport Measurement, *IAHS Publ. No. 133,* 177-194.

Walling, D.E. & Webb, B.W. 1988. The reliability of rating curve estimates of suspended sediment yield: some further comments. In: Sediment Budgets, *IAHS Publ. No. 174,* 337-350.

Whitehead, P.G. 1979. Applications of recursive estimation techniques to time-variable hydrological systems. *J. Hydrol.*, 40, 1-16.

Young, T.C., DePinto, J.V. & Heidtke, T.M. 1988. Factors affecting the efficiency of some estimators of fluvial total phosphorous load. *Wat. Resour. Res.*, 24, 1535-1540.