Interim Review of Data Handling and Information Needs of Regulatory Agencies

WRc plc

R&D PRS 2273-M

NRA

National Rivers Authority

INTERIM REVIEW OF DATA HANDLING AND INFORMATION NEEDS OF REGULATORY AGENCIES

Report No: PRS 2273-M

September 1989

Author: J C Ellis

Contract Manager: J C Ellis

Contract No: 4743

Client's Reference No: 6.1.4a

RESTRICTION: This report is restricted to the National Rivers Authority and equivalent bodies in Scotland and Northern Ireland.

Any enquiries relating to this report should be referred to the author at the following address:

WRc Medmenham, Henley Road, Medmenham, PO Box 16, Marlow, Buckinghamshire SL7 2HD. Telephone: Henley (0491) 571531

SUMMARY

I OBJECTIVES

To establish the data handling and information needs of regulatory agencies.

II REASONS

The collection of monitoring data is not an end in itself. A clear specification of the information to be derived from routine monitoring data is therefore essential in ensuring that such data is used to best advantage. It is then important, having identified the information needs, to establish the data handling mechanisms - and, in particular, the computer software - needed to derive the desired information.

III CONCLUSIONS

The main point to emerge from the review is that, unsurprisingly, current practice regarding data handling and interpretation varies widely across the Regions. The report proposes a number of guidelines that would lead to a more unified national approach, and would also ensure that some of the information needs identified in the review were met more effectively than at present.

IV RECOMMENDATIONS

Broadly, the principal recommendations are as follows:

* that data transfer from mainframe to microcomputer is made easier;

* that a standard format is adopted for compliance reports and quality characterisation summaries;

* that confidence limits and other relevant measures of uncertainty are given greater prominence in reports;

* that a Test Data Facility is established within each Region to promote the more effective use and interpretation of monitoring data; and

* that a forum of quality officers from the Regions is established as a focal point for discussions on data handling matters.

V RESUME OF CONTENTS

The review is presented in three stages. First, we set out and discuss the information needs associated with various types of routine quality data. Six main categories were identified during our discussions with the Regions: compliance testing, quality characterisation, trend detection, planning, modelling, and responding to enquiries.

Next, we review the current position regarding data handling and interpretation software, particularly noting the extent to which the information needs established earlier can be met.

Then in the final section, we propose a number of guidelines for future data interpretation software, and also make some recommendations of a wider nature.

CONTENTS

CONTENTS continued

SECTION 1 - INTRODUCTION

1.1 BACKGROUND

WRc's current research programme for the NRA contains five projects relating to statistical aspects of routine quality monitoring. These are:

4739: Sampling Programme Design

4743: Data Handling and Information Needs
4745: Software for Data Interpretation
4744: Code of Practice for Processing Data

4761: Sampling and Statistics Service

Viewed as a whole, the five projects form a natural progression from WRc's recently published Sampling Handbook (Ellis 1989) - the emphasis now being very much on implementation. The first project - Sampling Programme Design - covers our involvement in a 'technical adviser' capacity with the various working groups set up earlier this year by NRA Central to establish general sampling guidelines for effluents, rivers, groundwaters and estuaries. At the other extreme, the Sampling and Statistics Service project is a general NRA helpline offering advice to anyone with a statistical enquiry or request.

The other projects focus in turn on three aspects identified in the Sampling Handbook as being of key importance to the success of a routine monitoring programme:

(i) The first - Data Handling and Information Needs - is concerned with the questions: 'What types of information are we hoping to obtain from our monitoring programme?' and 'What is the best way of extracting that information from the data?' That project forms the subject of the present report.

(ii) Once those information needs have been quantified, the next requirement is for suitable computer software. Some readers will already have seen WRc's AARDVARK package in action, running on their own micro and giving new insights into their own data. AARDVARK is a forerunner to other software products that we are in the process of developing in our Software for Data Interpretation project - the aim throughout being to help NRA quality officers extract the maximum information from their data. Some aspects of this project, too, are briefly mentioned later in the report.

(iii) Finally, our project on a Code of Practice for Processing Data is concerned with establishing standard protocols for the transfer and statistical interpretation of data. When data from the NRA Regions is collated to produce a national picture - or used to provide between-region comparisons - it is important that the assessments are made in a coherent and consistent manner. Otherwise, apparent variations could simply be artifacts due to the use of different statistical methods rather than indicative of real effects. Two examples of issues that would benefit from definitive guidelines are (i) the handling of 'less-than' values in statistical summaries, and (ii) the form of statistical distribution to assume when calculating 95%iles.

The Code of Practice can be thought of, therefore, as a mechanism for defining the practical details of the data interpretation systems that we identify in general terms here. A comprehensive discussion of the areas that might be covered by the Code of Practice is contained in the companion report PRS 2274-M.

1.2 STATUS OF REPORT

Following preliminary visits to several of the NRA Regions, a short note was circulated flagging up some of the topics to be discussed in subsequent more extensive meetings with the Regions. Seven such visits took place during July and August. One aim of this report, accordingly, is to summarise the main findings emerging from those meetings. In view, however, of the necessarily limited time available for consultation during the very busy period leading up to the NRA's September launch, we suggest that the report be regarded as an interim document for discussion. This will give the opportunity for us to receive more feedback - via further visits to and discussions with the NRA Regions, perhaps coupled with a joint meeting at Medmenham - following which we can produce a final version of the report.

A second reason for rescheduling the production of the final report is the degree of overlap between this WRc project and the activities of the groups established by NRA Central in March to plan the overall monitoring strategy. With hindsight some overlap was inevitable, given the need for WRc's research programme to be largely in place by early 1989. The main consequence of this is to lessen the degree of emphasis that we need place here on quantifying the information needs of routine monitoring (although it has still been useful to discuss this aspect from the perspective of quality officers out in the Regions). In particular, we will be saying nothing here about matters such as the choice of sampling locations or the specification of desired precision. The material that follows in Section 2 should therefore be regarded as complementary to the treatment of monitoring objectives given in the report to be produced shortly by the Chairman of the River and Effluent Sampling Groups.

There is also an element of overlap between our discussion in Section 4, on suggested guidelines for data handling software in the Regions, and the work of the NRA group that is currently addressing the whole issue of Information Technology (IT). We would hope, however, that the thoughts outlined here will be seen as broadly desirable goals whatever particular direction the NRA's future IT strategy might take.

1.3 SCOPE OF REPORT

The report is concerned with routine chemical quality monitoring data. One-off or ad hoc surveys, though of great importance, lie outside the present remit. Though the emphasis will vary, much of the discussion applies more or less equally to effluents, rivers, groundwaters and bathing waters.

1.4 OBJECTIVES

Following that preamble, we can now set out the objectives of the report. They are to:

(i) set out and discuss the information needs to be supplied by various types of monitoring data;

(ii) outline the types of statistical method that are appropriate for meeting those needs;

(iii) review the current position as regards statistical computing in the Regions; and

(iv) propose guidelines for the types of software ideally needed to apply the required statistical procedures.

SECTION 2 - ROUTINE MONITORING INFORMATION NEEDS

Data generated by routine quality monitoring programmes can be used for a variety of purposes. We have summarised our discussions with the Regions under the following six main headings:

(i) testing for compliance with standards;
(ii) determining present quality;
(iii) looking for trends;
(iv) planning;
(v) modelling;
(vi) responding to enquiries.

We discuss the information needs for each of these categories in turn in the following sections.

2.1 TESTING FOR COMPLIANCE WITH STANDARDS

Because there has long been a statutory requirement for the industry to provide annual compliance returns, the information needs in this area of routine monitoring are fairly clear-cut. Given (a) the relevant set of pass/fail criteria for a particular type of quality data, and (b) the numbers of excedences in relation to the total numbers of samples taken, the basic information need in compliance reporting is simply a factual statement of whether each sampling point has passed or failed.

In our discussions with the Regions, it was widely acknowledged that two additional types of information are needed beyond this basic minimum requirement. The first is a knowledge of how the level of compliance is changing through time, so as to provide advance warning of possible problems. The second need is for some measure of the confidence which can be placed in any individual pass/fail conclusion. For example, two sewage effluents that respectively failed their BOD consents (a) once in 50 samples, and (b) three times in 20 samples would both be judged to pass (on the basis of the Look-up Table); but we would clearly be more confident about the former judgement than the latter.

2.2 QUALITY CHARACTERISATION

Taken at face value, this is the simplest type of information requirement. All water authority archive systems can provide summary statistics such as means and 95%iles over any specified time period. (In some instances that is just about all they can provide!) As one user put it, 'Archive data provides a statement of quality'. So the information need is apparently well met. Or is it? Two basic questions can usefully be posed:

(i) Is the user given a quantitative measure of the uncertainty associated with each summary figure?

(ii) How does the user define the period over which to base the assessment?

On the first question, we quickly become involved in the innermost feedback loop of sampling programme design (see the discussion in Chapter 3 of the Sampling Handbook). In principle, the user should specify how precisely he needs to know a particular quantity over a certain period, and this dictates the required sampling frequency. (Indeed, a key task of the various NRA Sampling Groups which have been meeting this year has been to arrive at a sensible trade-off between desirable precision and required sampling effort.) But in practice, most users of quality data are confronted with a fait accompli: the samples have already been taken and the results are there on the archive. It is important, therefore, that the user is made aware of the degree of uncertainty that surrounds each reported summary statistic.

The other question concerns the choice of timespan. One year is probably the most commonly-used default period - the implication being that it is acceptable to aggregate data across any seasonal variations that might be present. For sampling points sampled less often than monthly or so, another common arrangement is to base the assessment on a two- or three-year period. But how should the choice be made? The answer, we suggest, hinges upon a key piece of information that the user should be seeking from the data, namely the time period over which quality has been running at its current level. The use of some form of trend-detection routine (see Section 2.3) may show that quality has been stable for the past five or six years; or it may show that there have been noteworthy changes even within the past year. This sort of information would enable the time period over which current quality was assessed to be tailored most effectively to the circumstances of each sampling point.

So to summarise, we believe the primary information needs for quality characterisation to be:

(i) a measure of the **uncertainty** surrounding each required estimate (not just the summary statistics themselves); and

(ii) a measure of the **timespan** over which quality has been running at its current level.

2.3 LOOKING FOR TRENDS

Given that the great majority of routine quality monitoring in the water industry has traditionally been (and is likely to continue to be) conducted at frequencies of fortnightly or less, probably the most useful objective to focus upon is medium- or long-term trend detection. In other words, everyone wants to know whether quality is getting better or worse.

But whilst there is general agreement that this is an important objective, it is quite another matter to attempt to quantify what information on trends is required from routine monitoring. What is a trend, anyway? Is it a linear progression of improvement or deterioration in quality? This may seem a plausible definition; but it is actually surprisingly difficult, in our experience, to find many genuine examples of linear trend. (Nitrate trends in certain groundwater sources are a notable exception to this rule.)

A simpler definition of trend, accordingly, and one that we have found usefully robust, is:

'A step change in the underlying mean quality between one (unspecified) time period and another'.

This definition has three main points in its favour. The first is that some changes in water quality do occur abruptly - as with the commissioning of a new treatment unit, or the closing down of a major industrial discharge. The second point is that, unfortunately, most routine monitoring programmes are themselves subject to step changes in sampling or analysis protocol - a change in the analytical method; a laboratory rationalisation; a switch in the sampling schedule - and these can induce step changes in perceived quality. The final point is that any more complex pattern of trend, if sufficiently pronounced in relation to the monitoring frequency and short-term sampling noise, can in any case be approximated by a series of steps (as we see, for example, when curves are drawn on a coarse resolution graphics screen).

Supposing that we accept this simple definition of trend, we are still faced with the problem of quantifying the size of step change that we would wish to detect, were it to occur, in a given time period. This is a perennially difficult question to which there is no easy answer. Our suggested solution - as we discuss at greater length in the Sampling Handbook - is to put that particular question on hold, and instead concentrate on answering the simpler question: 'What trends are we able to detect from past data?' To know the type of information that we have historically been capable of providing is an essential preliminary to deciding what level of information it is sensible and realistic to aim at in the future.

2.4 PLANNING

The use of routine quality data for planning purposes was mentioned by several Regions. This rather broad objective can perhaps usefully be subdivided into two more specific requirements:

(i) to determine the current position; or
(ii) to predict when a problem is likely to arise in the future (for example, in judging when a treatment works is likely to become overloaded).

Requirement (i) is a matter of characterising present quality, and so the discussion of Section 2.2 applies. In (ii), the interest lies in trend estimation and so the objective is essentially that discussed in Section 2.3 (although in this context something more sophisticated than a simple step model would clearly be needed if prediction is involved). In effect, therefore, the planning requirement is covered by these two more basic types of information need.

For major planning enquiries, it will often be the case that historically acceptable information requirements - and so the sampling frequency or sampling locations commensurate with those requirements - are insufficiently precise. It will then be necessary to 'move up a gear' and embark on a regime of more intensive sampling.

2.5 MODELLING

One objective of routine quality monitoring data mentioned by several Regions was the provision of data for model calibration and validation. Some diversity of opinions was noted, however, over the usefulness of routine monitoring data for modelling purposes - no doubt reflecting the fact that modelling can mean very different things to different people. For a simple black-box consent-setting model, for example, the information needs may well be accommodated by those already identified under the headings of quality characterisation or trend estimation. At the other extreme, a dynamic river quality model for, say, predicting and controlling the transient effects of intermittent discharges will have information demands that far outstrip routine needs and call for high-intensity special surveys.

2.6 RESPONDING TO ENQUIRIES

Finally, the Regions have a general need for routine quality data so as to be able to respond to ad hoc enquiries. As a quality officer in one Region put it, 'The archive is essential for the unforeseen questions'.

It is notoriously difficult, however, to progress from that very reasonable general statement to a more specific prescription of what is actually needed to meet this requirement. Thus, what is the benefit of providing a particular determinand at a particular location at a particular sampling frequency?... or the disbenefit of not having a particular set of data on the archive?

The classic method of finding out the usefulness of a particular report or service is to stop providing it routinely and see how long it takes before somebody shouts for it. But that approach is unhelpful here, as we cannot switch sampling points or determinand suites in and out of the routine programme just like that. A further problem is that the nature of people's enquiries tends to be conditioned by knowledge of what is available. For example, they might ask for 'minimum daily flow during August'; but they would not think of asking for 'minimum daily dissolved oxygen during August' - useful though such a statistic might be in assessing the health of the river.

The next difficulty follows on from the fact that, clearly, it is impossible to plan a routine monitoring programme to meet all possible enquiries. So what instead can often happen is that the present capability is used (maybe only by implication) to define the 'desirable' level of coverage. This then becomes a powerful agent for reinforcing the status quo.

But that begs the question of whether the present level is adequate anyway. There is a tendency for people to be too easily pleased when they get their hands on some data. ('This was the only data we had, but we used it because it was better than nothing.') The important thing, however - whether the aim is concerned with mean quality, percentile quality or anything else - should be the underlying truth, not the estimate obtained from the sample values. The smaller the available number of samples, the more imprecise is the estimate of that underlying truth, and so the greater is the risk that the person using that data will arrive at a misleading conclusion. So situations could arise in which a small amount of data was actually worse than no data at all!*

So to summarise, there is a strong tendency for the information needs relating to ad hoc enquiries to 'define themselves by default' - partly through historical precedent, and partly through people's blurring of the distinction between sample and population. This is unsatisfactory, given the general need for consistency between Regions. What we suggest, therefore, is some form of workshop discussion at which NRA quality officers could develop a consensus view as to what constitutes an acceptable minimum capacity for responding to ad hoc queries. WRc would be happy to host such a meeting if this were thought to be useful.

---

* The usefulness of routine data is further undermined, for some types of ad hoc query, by the 'sampling window' bias common to virtually all routine monitoring.

SECTION 3 - DATA HANDLING SOFTWARE IN THE NRA REGIONS

In the previous section we discussed the various types of information that are needed from routine quality monitoring data. To extract information from data, software is needed. In the present section, therefore, we move on to a brief review of the current position regarding data handling and interpretation software in the Regions. Then in Section 5 we will set out some suggested guidelines for future data interpretation software across the NRA.

3.1 QUALITY ARCHIVE RETRIEVALS SOFTWARE

The starting point of any examination of quality data is the retrieval routine. In one of the Regions we visited, a reorganisation of the archiving arrangements was imminent, and retrieval software was still being written. With that exception, every Region has a well-established retrieval system allowing the data to be sent either to the screen and/or printer, or to a batch file for subsequent access by various back-end programs. The procedures are more user-friendly in some Regions than others, but all are adequately accessible to the typical computer user after perhaps an hour's tuition.

3.2 OTHER MAINFRAME SOFTWARE

3.2.1 Testing for compliance with standards

All Regions have routines that apply the relevant pass/fail criterion to each type of quality data - the Look-up Table for sewage effluent determinands, for example; the 1-in-20 rule for bathing water quality; and so on. In several Regions, the routines are wrapped up into a single general program which can call up the appropriate pass/fail algorithm from a master file of compliance rules according to the type of data being assessed.

In reporting on sewage effluent compliance, the routine used by Thames Region produces a particularly attractive format. An example of the data summary produced for a treatment works is shown in Table 1. The most striking aspect of the output is the way that so much useful information is shoehorned into so small a space - a characteristic not always a feature of mainframe line-printer summaries.

In Section 2.1 we identified the need for additional information on (a) trends in compliance at a point through time, and (b) the level of confidence to be attached to any individual pass/fail judgement. Most Regions meet the first of these needs by the use of a 'moving average' approach, whereby the assessments are made on a rolling 12-monthly basis rather than at the minimum frequency of just once per annual reporting period.

We found a greater variety in the way the Regions tackled the second requirement. The approach used in several Regions is to calculate confidence limits around each 95%ile estimate of quality, and consult these to give some idea of which determinand/sampling point combinations are more borderline than others. This approach, though useful, does call for the exercising of an element of judgement as each case is individually scrutinised, and so scarcely lends itself to routine application over many river or effluent sampling points. In contrast, two other Regions have devised a mechanism that is applied automatically as a routine part of their compliance assessment programs. We briefly outline both schemes below.

(i) In Anglian Region, the 'Confidence of Failure' statistic (CoF) is calculated for each determinand in assessing 95% compliance with effluent consents or river class limits. The nearer to 100% the CoF is, the more confidently it can be asserted that there really has been a failure to meet the limit. (Thus, an exactly equivalent way of expressing a Look-up Table failure for a particular determinand is to say that the CoF is greater than 95%.) So by sorting and printing the sampling locations in decreasing order of this statistic, effluents or rivers are

Table 1: Example of the sevage effluent compliance report used by Thames Region

Group Manager	Works/Effluent Description	Site Code	Population Equivalent	Min. Samp Frequency
V HODGSON	CAMBERLEY	PLDE.0022	55500	52

Det	Consent Condtn.		Planned Samples 1988	Planned Samples 1989	Samples Taken	Results Obtained: Month by Month	JAS OND JFM AMJ	No Flow	Range of Values % 0–50	50–90	90–100	100–110	110–150	150–200	200+	Performance Indicators: Quality Achieved	Fail Risk	Form Fail
SS	20/	0	52	52	52	39 PASS 13 FAIL	444 412 332 444 001 042 203 010	0	26	9	4	1	4	4	4	49.3(246%)	100%	Yes
BOD	25/	0	52	52	52	44 PASS 8 FAIL	444 433 433 444 001 021 102 010	0	34	9	1	4	3	0	1	33.9(136%)	61%	Yes
AMM	1[illegible]/	0	52	52	52	47 PASS 5 FAIL	435 454 433 444 010 000 102 010	0	23	20	4	2	2	1	0	18.1(121%)	43%	No

separated not merely into passes and failures: the passes are themselves categorised on a quantitative scale ranging from 'dubious' (CoF = 94.9%) down to 'excellent' (CoF = 0%).

It might be thought that much the same effect could be achieved simply by ranking the cases in decreasing order of observed compliance. But this would fail to take account of how the number of samples qualifies that information. For example, in testing for compliance with a 95%ile standard, an observed compliance of 85% based on 20 samples has an associated CoF of 92.4%. But the same observed compliance figure derived from 40 samples (ie 6 excedences in 40 rather than 3 in 20) produces a CoF of 98.6% - much more convincing evidence of failure.

(ii) The effluent compliance program used by Thames Region generates a rather different statistic, though with much the same underlying purpose, known as the 'Risk of Failure'. On the assumption (for any particular determinand) that effluent quality is log-Normally distributed, an estimate is made of the proportion of time that the consent is exceeded. This figure is then fed into a binomial routine to calculate the corresponding chance of failing the Look-up Table - and this is the Risk of Failure.

An example will clarify the purpose of the calculation. In the rolling assessment shown in Table 1, five ammonia values exceeded the consent out of 52 samples. As we can see, the Risk of Failure turned out to be 43%. This tells us that if the current year's data really is typical, there would be a 43% chance of the treatment works failing its Look-up Table ammonia consent in any future 12-month assessment.

(It is interesting to note that an equivalent non-parametric calculation can be made, rather more simply, as follows. Our best estimate of the true excedence rate is simply 5/52 = 9.6%. If we insert this figure into the same binomial routine as before, we get a value of 38% for the Risk of Failure - little different from the parametrically derived figure.)

3.5 IMPORTING SOFTWARE

In the discussion document circulated prior to our meetings with the Regions, one of the questions was this:

> Suppose you were given a FORTRAN subroutine that, say, performs a test for trends in river quality upon any given five-year run of data. How easy would it be for you to 'graft' this subroutine onto existing software so as to try out the test on 50 or more of your own archived data sets?

In our discussions we learnt that three of the Regions would find such an exercise very difficult, having nobody with the necessary FORTRAN expertise. Wessex Region would have problems, too, as its SAS-based system does not support FORTRAN. All the other Regions with whom we discussed the idea were confident that they could cope successfully.

Even in the Regions where there would at present be problems, there was general enthusiasm about the potential benefits to be gained by circulating statistical routines nationally in this way.

SECTION 4 - ROUTINE DATA HANDLING NEEDS

As the previous section indicated, the current position regarding statistical software in the NRA varies markedly between Regions - understandably so, given the historical background. It is also the case that, in certain respects, typical present capabilities fail to satisfy the corresponding information needs identified in Section 2. We conclude the report, therefore, by setting out some suggested guidelines for future data interpretation software across the NRA.

4.1 BATCH VERSUS INTERACTIVE

It is helpful at the outset to draw the distinction between:

(i) repetitive routines required to be applied sequentially to data from many sampling points; and

(ii) more detailed, ad hoc or open-ended investigations of data relating to individual sampling points.

Applications of the former type are most effectively carried out as batch runs on the mainframe system holding the data archive. In contrast, the latter type of application is ideally suited to the micro environment. The user should feel able to select either option depending solely on the type of study to be undertaken. For this to be possible, adequate data transfer systems must be readily available for:

(i) extracting data for specified sets of determinands and sampling points from the archive, and writing this to a batch file in a format suitable for subsequent input to various back-end analysis programs; and

(ii) downloading specified data files from the mainframe to individual users' micros.

4.2 COMPLIANCE ASSESSMENT

To provide a better understanding of the uncertainty attaching to each pass/fail judgement, all Regions would benefit from the routine calculation of a supporting statistic such as the 'Risk of Failure' measure used by Thames, or Anglian's 'Confidence of Failure' statistic. Indeed, in view of the importance of the regulatory function to the NRA, there would be advantages in all Regions using the same reporting format. Both these issues would be useful topics for discussion should the idea of a workshop, mentioned earlier, be pursued.

Routine compliance assessments will normally of course be run on the mainframe. For exploring the statistical principles underlying compliance issues, however, no 'real' data is needed and so the micro is the ideal choice. For example, what is the reasoning underlying the

Confidence of Failure and Risk of Failure measures discussed earlier? It is to help illuminate such questions that we are developing the PC-based ZEBRA package. ZEBRA can be used to produce binomial distribution plots; power curves for any specified 'up to r out of n' compliance rule; confidence intervals around any observed % compliance figure; Look-up Table calculations; and a variety of other charts and diagrams. A comprehensive account of the functions provided by ZEBRA can be found in the Sampling Handbook, and the package is planned to be available early in 1990.

On the question of time trends in compliance, we endorse the practice followed in most Regions of maintaining a rolling 12-month assessment. Further insights into compliance trends would be gained indirectly by the use of the procedures advocated in Section 4.4.

4.3 QUALITY CHARACTERISATION

There is a strong case for the NRA adopting a common format for quality characterisation summaries across the Regions. This would not, of course, prevent Regions from using other reporting formats for internal purposes - although the apologetic way in which several Regions referred to the lacklustre appearance of their standard summaries suggests that a general redesign of output formats might not be a bad idea in any case.

The details of what the standard summary should contain would need to be discussed with the Regions. But in principle our recommendations are for (a) greater emphasis to be placed on **non-parametric** methods for percentile estimation, and (b) **confidence intervals** to be quoted along with each summary statistic.

4.4 TREND DETECTION

The most notable shortcoming in the industry's current software capability is in trend detection. The lack of effective software in this area has in fact long been recognised, and WRc has for this reason been running an exercise for a number of years on the development of programs for detecting and quantifying trends in routine quality data. One major output from this programme has been our recently released AARDVARK package. AARDVARK's trend detection routine is built around the 'cusum' technique. This allows the user to browse through any selected time series, searching for and testing the statistical significance of step changes in quality. AARDVARK can also be used to test for linear time trends and to highlight seasonal patterns.

Another of our forthcoming software products is concerned with quantifying trends in quality between sampling points along a river. A micro version of this program is currently being developed from the mainframe version, and we aim to have this available to the NRA early in 1990.

For larger-scale batch applications, personal examination of individual determinands site by site is impracticable. Rather, the need is for some type of automatic scrutiny that can screen out the relatively few sampling points that might benefit from a more careful scrutiny. The potential value of such a mechanism was confirmed in our discussions with the Regions, and the reaction from the NRA River and Effluent Groups reinforced this view. We will shortly be constructing a prototype cusum-based exception-reporting algorithm, and hope to put this to the test on real-life quality data during the next stage of the Code of Practice study (see the companion report PRS 2274-M).

4.5 TEST DATA FACILITY

In the discussion of Section 3.5 we noted that, although not all Regions were able to cope at present with importing and running software from outside sources, there was general enthusiasm as to the value of such exercises. We share that enthusiasm, and believe that the development of a 'Test Data Facility' within each Region would be of major benefit to the NRA in promoting the more effective use and interpretation of monitoring data across its Regions.

3.2.2 Quality characterisation

With the temporary exception mentioned in Section 3.1, all Regions have software for producing statistical summaries - available, in most cases, as an integral part of the archive system. As we remarked in Section 2, there are marked differences between Regions in the type of information provided. In particular, a variety of methods are used - both parametric and non-parametric - for estimating percentiles, whilst confidence intervals are provided by some Regions but not others.

3.2.3 Trend detection

Most Regions offer users the ability to plot water quality data as a time series, but very little non-specialist software seems to be available on how to analyse or interpret such plots in situations when the answer is not obvious merely by inspection. Three Regions do, however, have in-house routines for carrying out trend analyses by the cusum method.

3.2.4 Statistical packages

In most of the Regions, one or more of the main statistical packages are available on the mainframe, such as GENSTAT, SAS, SPSS, BMDP or Minitab. (Indeed, at Wessex Region the entire computer system is built around SAS.) The computer language FORTRAN also is available in all but a couple of Regions.

Statistical tools of this sort are essential to the specialist. However, a clear message emerging from our discussions with the Regions was that they are of limited help to the typical user, who generally has insufficient statistical experience to know which technique to turn to in a given situation. Several Regions have tackled this problem by identifying various common statistical tasks (comparing two sample means, for example), and then writing appropriate routines or 'macros' which serve as a more accessible interface between the package and the user. In North West Region, for example, a number of GENSTAT and BMDP macros have been successfully developed in this way.

3.3 TRANSFERRING DATA FROM MAINFRAME TO MICROCOMPUTER

With the rapid growth of microcomputers over the last few years, there is an increasing need to be able to transfer or 'download' data files from the mainframe to a user's desktop micro. In perhaps half of the Regions, downloading is a fairly straightforward operation provided the user is equipped with the necessary communications link and software (Tangogate is used by several Regions). Elsewhere matters are not so simple. One Region remarked: 'It is awkward, cumbersome and unfriendly to transfer extracts to the micro'. And in the Region with evidently the worst problems, the comment was: 'Downloading is very, very difficult. It's almost as though the Computer Department makes it so in order to maximise its income!'

3.4 MICROCOMPUTER SOFTWARE

In a majority of Regions there is a positive policy of encouraging the use of micros - particularly for detailed investigations of particular sites. As well as the obvious benefits of colour and graphics (features not widely available to mainframe users in most Regions), other reasons mentioned were: greater user-friendliness, better interactive facilities, speed, ease of access, and 'the huge expense of the mainframe'. A variety of commercial products are already in use across the Regions - reflecting the wide range of excellent and inexpensive micro software on the market - including SuperCalc, the various Lotus products, and more dedicated statistical packages such as Statgraphics, Minitab and Instat.

As yet, few Regions have developed in-house programs for other than local use. One exception is Anglian Region, whose pioneering suite of routines relating to compliance, consent-setting and sampling error developed in the early 1980s gave an early pointer to the enormous potential of the micro.

SECTION 5 - CONCLUSIONS AND RECOMMENDATIONS

The main point to emerge from this interim review of data handling and information needs is that, unsurprisingly, current practice regarding data handling and interpretation varies widely across the Regions. The report proposes a number of guidelines that would lead to a more unified national approach, and would also ensure that some of the information needs identified in the review were met more effectively than at present.

The main recommendations are these:

* that each Region establishes data transfer facilities making it a straightforward matter for the data user to (a) build up a batch file of archive extracts from a sequence of sampling points, and (b) download data files from the mainframe to his micro;

* that the NRA adopts a standard reporting format - both for compliance reports and for quality characterisation summaries;

* that pass/fail decisions in compliance reports are accompanied by some auxiliary measure of statistical confidence, such as the Confidence of Failure statistic used by Anglian Region;

* that the standard summary report places primary emphasis on non-parametric estimates of percentiles, and quotes confidence intervals along with each summary statistic;

* that the NRA encourages the development within each Region of a Test Data Facility to promote the more effective use and interpretation of monitoring data across its Regions;

* that the Test Data Facility is used at an early stage to explore the potential of an exception-reporting trend detection procedure to be supplied by WRc; and

* that a forum of quality officers from the Regions is established to serve as a focal point for discussion - and meets in the first instance to debate the various questions flagged up in this report.

REFERENCES

ELLIS J C (1989) Handbook on the Design and Interpretation of Monitoring Programmes. NS29, Water Research Centre.

ELLIS J C (1989) Code of Practice for Data Handling - a Discussion Document. PRS 2274-M, Water Research Centre.