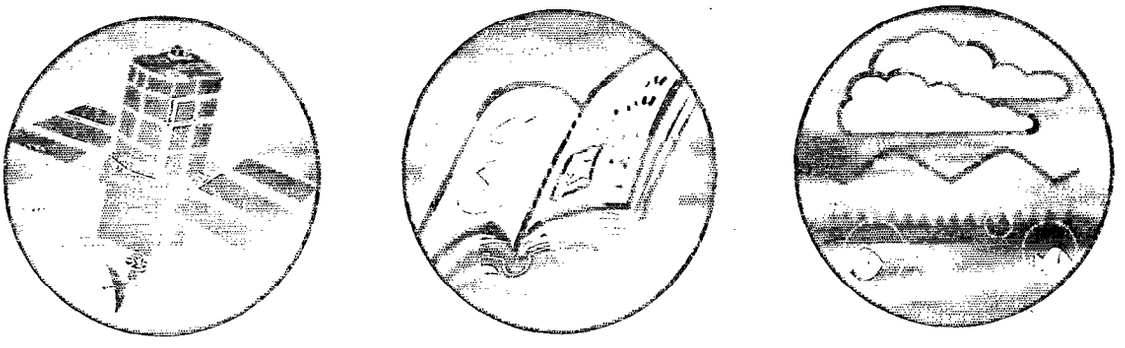
Data Transmission Requirements for Self-monitoring



Research and Development
Technical Report
E60

ENVIRONMENT AGENCY

All pulps used in production of this paper is sourced from sustainable managed forests and are elemental chlorine free and wood free

Data Transmission Requirements for Self-monitoring

Technical Report E60

D J Waters, P Marsh, J Humm and T Marsland

Research Contractor:
QuantiSci Ltd and Aspinwall and Company Ltd

Further copies of this report are available from:
Environment Agency R&D Dissemination Centre, c/o
WRc, Frankland Road, Swindon, Wilts SN5 8YF

tel: 01793-865000 fax: 01793-514562 e-mail: publications@wrcplc.co.uk

Publishing Organisation:
Environment Agency
Rio House
Waterside Drive
Aztec West
Almondsbury
Bristol BS32 4UD

Tel: 01454 624400 Fax: 01454 624409

ISBN:HO-07/98-B-BBYH

Dissemination status
Internal: Released to Regions
External: Released to the Public Domain

Statement of use
This report recommends options for the transmission and management of operator self-monitoring data. The information within this document is for use by Environment Agency staff with data management responsibilities.

Research contractor
This document was produced under R&D Project E1-006 by:

QuantiSci Ltd
Chiltern House
45 Station Road
Henley-on-Thames
RG9 1AT

Tel: 01491 410474 Fax: 01491 576916

Environment Agency Project Leader
The Environment Agency's Project Leader for R&D Project E1-006 was:
Dr Richard Saull, Environment Agency, NCEDS, Twerton

Amendments
Any corrections or proposed amendments to this manual should be made through the regional Agency representative on the Water Resources National Abstraction Licensing Group.

Summary

An essential component of the Environment Agency's overall environmental strategy is an effective environmental monitoring and assessment programme. In certain circumstances the environmental monitoring is performed by operators and serves as an essential data source for checking compliance with EC directives, International commitments, and national standards and targets. In addition, certain data are used for determining financial charges. This latter function is increasing in significance with time.

The primary objective of this project was to make a number of recommendations concerning the transmission of operator self-monitoring data to the Agency. The information was gathered from:

- background information;
- strategic meetings and telephone calls;
- a 'Business Needs' and a 'Data Collection' questionnaire; and
- contacts with external organisations by telephone and questionnaire.

The Agency has been in a state of flux and identifying key personnel within each functional area, Region and Area office, was not a trivial task. Ultimately this process took longer than originally envisaged, although the final contact list was deemed satisfactory.

The preliminary phone calls were an essential information source and were critical in the formulation of the structure of the questionnaires. A number of iterations were discussed with key personnel before the questionnaires were disseminated more widely within the Agency. The level of response to 130 questionnaires, sent to targeted individuals within the Agency, was below our expectations and was probably affected by the workloads of the targeted personnel. Considerable effort was expended in trying to elicit information from respondents, as it was important that any recommendations would reflect, as far as possible, the views of existing staff and current practices. It is possible that for future elicitation exercises the use of targeted interviews rather than large numbers of questionnaires may be more effective.

Approximately 25 questionnaires were used in the final analysis with the quality of response varying across the topic areas. Some aspects such as data checks and quality assurance, and equipment used and data volume were rather vague, but others such as transmission method, software used and frequency of transfer were much more clearly specified and understood. The industry response provided useful corroborating evidence of the self-monitoring situation.

In general there was considerable variation in the approach between functional areas, and in some cases variation between Regions within the same function. Hence, short and medium term guidance and protocols are presented that work towards a better common solution which takes into consideration wider Agency initiatives, such as convergence and public access policy.

The main conclusions from the study are:

- there are a range of practices, even within functions, that need to be standardised;

- the Data Strategy Steering Group (DSSG) policies need to implemented;
- wider initiatives need to be better publicised, the 'best interim solution' clearly presented and potential impacts assessed as early as possible;
- clearly specified quality assurance procedures for sampling and analysis to improve data integrity are required;
- data management responsibilities should be better defined, i.e. assignment of dedicated personnel to standardise and improve the handling of data;
- a requirements gathering and specification exercise is needed for WIMS (in a wider Agency role);
- master and working files should be clearly identified to avoid data loss/corruption and ensure that audited data is used for all secondary purposes; and
- the need for better overall Agency understanding of the system components of the 'Best Interim solution' that affect data storage and management, e.g. clearer guidance as to the implications of the convergence and public access policies;

Several technical recommendations can be made and include:

- standard guidelines for data formats should be established;
- quality assurance procedures for sampling and analysis should be specified in order to improve data integrity;
- data transfer in electronic format should be requested where practical;
- in the short term data transfer via floppy disk and Zip cartridges is appropriate;
- Direct Network Transfer (DNT) could be integrated with the above where operators are already connected (but this needs to be managed by the Agency); and
- production of data flow sheets for the improved understanding of the needs of secondary and tertiary uses and users of data.

Finally the data transmission requirements of the UWWTD are considered in more detail and are presented in Annex E.

Contents

1 Introduction

The Environment Agency has a duty to undertake Research and Development to further the effective protection and management of the environment. Legislation enforced by the Agency often requires authorised dischargers to monitor their own discharges (to all media) and report the results to the Agency. The self-monitoring data are required for a number of primary (e.g. compliance, authorisation, raising charges, licence checking, public register etc.) and secondary purposes (e.g. Chemical Release Inventory (CRI), Environment Snapshot, adhoc statistical analyses, presentation to sponsoring bodies, Internet etc.).

The main objective of this project is to review the type and volume of self monitored data (automated and manual collection) and present a number of recommendations for effective data management (including data checking and verification, version control, access and security) and transmission from these remote sites to the Agency. It is important that we recommend a number of practical scenarios that are forward looking, but appreciative of existing practices in each of the functional areas.

The Agency has a Data Strategy Steering Group (DSSG) which is responsible for implementing Agency Data Policy. Their primary objective is to provide guidance to ensure environmental data are properly managed and preserved so that they can promote use and re-use by internal and external users, and enable commercial and educational exploitation. One of the aims of the DSSG is to have each of the Agency's major datasets under the custodianship of a designated data co-ordinator. At present however, there is a range of practices from central custody and stewardship to individual dataset management [Environment Agency, 1996b] and policy development in this area would obviously have an impact on the management of self-monitoring data.

There is a considerable literature base concerning information networks and data transfer protocols. QuantiSci and Aspinwall have various sources of generic and Agency background knowledge, and have utilised these sources of information in conjunction with strategic phonecalls and interviews, and questionnaires, to formulate an appropriate short and medium term action plan, that is compatible with wider initiatives.

The Agency requires self-monitoring data mainly to assess compliance with standards and limits and for charging purposes. The main activities that involve Agency staff liaison with the operator are listed below and are discussed in greater detail in Section 2.

- pre-application investigations;
- pollution prevention and incident inspections;
- monitoring discharges (audit monitoring) for consents and enforcement;
- inspection of licensed waste facilities;
- review of site inspection reports to assist in the response to planning consultations, e.g. landfills; and
- general site inspections.

Following the data gathering exercise the report covers three main areas:

- information derived from within the Agency - Chapters 2 - 5;
- assessment of external opinion (operators and interested parties); and

- short and medium term recommendations for self-monitoring data management and transmission.

Chapter 2 provides an assessment of the business needs, covering:

- the present and future regulatory requirements and procedures;
- the processes that require self-monitoring;
- the categories of data that are transferred;
- the secondary uses of self-monitoring data (e.g. CRI, annual statistical reports, other functional databases);
- a review of the receipt, storage and transfer mechanisms - i.e. internal data flows;
- and related Agency initiatives that may impact on procedures (e.g. convergence, public access policy etc.).

Chapter 3 offers a review of self monitoring techniques and recording instrumentation including:

- discharge type and monitoring requirements;
- the types of parameters transmitted;
- recording instrumentation and sampling procedures;
- quality control and data management; and
- the form of the data transferred (e.g. original or pre-processed).

Chapter 4 is an assessment of data transmission scenarios including:

- a review of data transmission technologies;
- and comparison of existing Agency practices with potential transmission options.

Chapter 5 -covers data management issues including:

- an assessment of the robustness and security of the data transfer;
- database storage, archive and verification checks;
- an assessment of the self-monitor performance; and
- version control and access rights.

Chapter 6 - provides an external perspective including:

- an industry view; and
- the thoughts of other interested parties.

Chapters 7 and 8 report the recommendations and conclusions of the study.

2 Assessment of Business Needs

There are numerous activities that discharge potentially polluting substances to the environment and their impact must be prevented, reduced or minimised, wherever possible. For the Agency, self-monitoring is a cost effective method of assessing the potential significance of discharges and their impact on receiving medium. It is important that the Agency's policy and practice in pursuance of its statutory duties keep track of technology and ensure that the appropriate results are transferred. Industry also has a duty to monitor discharges and provide the Agency with the relevant data to assess compliance and in some cases for charging purposes. In order to maximise the benefit of these data to the Agency, secondary and tertiary uses and users of the data should be recognised, and where possible, their requirements incorporated into the final recommendations.

In order to efficiently establish the basis of the business needs a number of strategic meetings and telephone conversations were held with key Agency staff. Due to restructuring within the Agency this was not a trivial task and the final key people are listed below:

Richard Saull - National Centre for Environmental Data and Surveillance (NCEDS) (Twerton)

Terry Long - NCEDS (Twerton)

Alistair Gordon - National Data Policy (Bristol)

Rob Gemmil - National Centre for Compliance Assessment (NCCA - Westbury-on-Trym)

Stuart Newstead - NCCA manager (Lancaster)

Stephan Carlyle - Scientific and Technical Information Service (SATIS - Bristol)

Chris Chubb - Pollution Prevention and Control (PPC - Bristol)

John West Water Services Association (WSA)

Chris Moore - Central Information Systems (CIS - Bridgwater)

Don Munns - Integrated Pollution Control (IPC - Bedford)

Richard Coward - CIS (Bristol)

Gillian Hill - Technical Information (Warrington)

Freda McDonald - Environmental Quality Technical Support (Bangor)

Aileen Kirmond - Water Resources (Bristol)

Cathy Greenhall - Data Management Team (Warrington)

Keith Harsham - IPC (Wallingford)

Vic Whiteley - IPC (Bedford)

Dave Wardle - Data Services (York)

Ruth Wolstenholme Scottish Environmental Protection Agency (SEPA)

These contacts also provided further names that received copies of the questionnaire. A complete list of the primary contacts can be found in Annex A. Some of these primary contacts also circulated the questionnaires more widely in their sections, so that the total number of questionnaires distributed was at least 150. A list of eventual respondents can be found in Annex C. The profile of their involvement in self-monitoring is illustrated below and shows a good spread of activities.

Involvement	Number of Persons
• specification of policy	4
• specification of need for self-monitored data	8
• reception of self-monitored data	7
• checking data for compliance with self-monitoring requirement	8
• checking data against prescribed limits	9
• summarising for internal reporting	6
• data management or archiving	7
• management of officers carrying out one or more of the above	6
• data distribution (including external distribution)	9

Table 2.1: Respondent 'Involvement' Profile

Although it would have been useful to have a full organisational structure when attempting to derive data flow tables for secondary and tertiary uses and users of data, unfortunately this was not available. However an Area Office list, a regional boundaries map, and a list of National Centres were provided.

Before discussing the regulatory instruments in the next section it is useful to briefly consider how the requirements are currently specified to the operator and how the data are currently transmitted.

There are various vehicles by which the requirements of these instruments are specified to the operator and include:

- conditional prohibition notices;
- letters of agreement;
- discharge consents;
- waste management licences;
- IPC authorisations;
- approval/agreement of working plans; and
- water abstraction licences.

Numerous documents could be offered by respondents that were considered good examples of how the requirements were specified.

The technical adequacy of the self-monitoring procedure was established in a number of ways including:

- specified in the documents; i.e. consents, working plans;
- negotiation;
- discussed and agreed at meetings; and
- acceptance by the field officer.

However, most respondents could not supply copies of documents that illustrated how these technical details are approved. Monitoring Certification for Continuous Emission Monitors (MCERTS) is being established by the National Centre for Compliance Assessment (NCCA) and its purpose is to approve instrumentation inline with Agency performance standards. This could also provide useful guidelines for operators.

Having reviewed the legislation and summarised how the requirements are communicated to the operator the final part of the loop is the transfer back to the Agency. The subdivision of this response is shown in Table 2.

mainly in hard-copy form	7
mainly in electronic form	1
neither predominantly hard-copy or electronic	3

Table 2.2: Mode of operator data transfer responses

N.B. The evidence following the phonecalls, interviews and operator questionnaire responses pushed the balance of existing practice further towards paper transfer, even though all operators stored data electronically.

2.1 Regulatory Requirements and Procedures

The Agency's principal aim, as set out in the Environment Act 1995, is to protect or enhance the environment, in order to achieve the objective of sustainable development. This Act has introduced new environmental management principles and the associated statutory guidance requires review in respect of the kind of environmental monitoring and assessment programmes that are needed to support the achievement of this objective. Although monitoring is carried out by both the Agency and discharging operators, this project is only concerned with the latter.

The Agency has inherited numerous environmental monitoring and surveillance programmes across its range of functional areas and these include a broad range of statutory requirements, formal commitments and environmental management needs. The main statutes and key regulations that give rise to self-monitored data are summarised in Annex B.

From a higher level perspective the following categories of commitments are relevant:

- **Global** - contributions to international conventions such as long-range transport of pollutants;
- **International** - North Sea Conferences and Oslo and Paris conventions;
- **European** - European Comission (EC) Directives and European Environment Agency (EEA) initiatives;
- **National** - national surveillance programmes, e.g. Harmonised Monitoring Scheme;
- **England and Wales** - e.g. GQA surveys of river quality;
- **Regional and Local** - e.g. regional surveys and local environmental management issues, e.g. LEAPs.

Environmental monitoring is required in support of EC Directives, International Commitments and National Standards and Targets and is most frequently enacted in UK legislation in the form of the Regulations that support UK Statutes. European policy and legislation has had an increasing impact on UK environmental practice in recent years and this has led to many new or amended Regulations. For example, the Agency has direct responsibility for 26 pieces of existing legislation that require specific reporting to the European Commission. These include requirements for:

- the regulation of emissions from different processes;
- the management and disposal of certain types of wastes; and
- the achievement of specific environmental quality standards and targets.

A significant number of these EC Directives are directed at water or wastewater quality. These relate primarily to:

- pollution control of individual substances (e.g. Dangerous Substances Directive, Nitrate from Agricultural Sources Directive);
- protection of different uses of the environment (e.g. Surface Water Abstraction Directive, Bathing Water Directive); and
- setting minimum requirements for processes (e.g. Urban Waste Water Treatment Directive).

Some directives are very specific and include details of determinands to be measured, sampling methods and sampling frequencies; others are left to the jurisdiction of the controlling Member State. Operators supply data primarily to assess compliance with individual environmental quality standards and targets, or occasionally to support decisions on designations of sites and areas such as "Sensitive Areas" and "Nitrate Vulnerable Zones", under the UWWTD and Nitrates Directive respectively. The detail of these directives is not the concern of this report, but their importance in determining future monitoring requirements should be recognised and allowed for in the the design of future data transmission systems. Purely as an illustration, the self-monitroing implications for the UWWTD are briefly discussed in Section 7.

International commitments that require monitoring are generally associated with achieving contaminant concentration targets, that is, controlling discharges to controlled waters or coastal waters. However, there are numerous international conventions that regulate

contaminant releases to the atmosphere, but so far these are generally assessed by measurements of emissions rather than sampling of the receiving medium.

At the national level, the UK has maintained a stance over the use of environmental quality standards and objectives as the basis for pollution control, but generally most legislative requirements have arisen from EC Directives. However, one set of standards that may eventually be derived nationally is the water classification scheme for river ecosystems under the Water Resources Act - 1991. The Government has also issued guidance on standards but few are yet set in statute, for example, DoE Circular 7/89 - List II dangerous substances. Nationally agreed, but non-statutory standards also exist for air quality, that is the Government's Expert Panel on Air Quality Standards (EPAQS).

One criticism that has been made [Environment Agency, 1997] is that process technology has outpaced the traditional mandatory transfer of data. Hence, existing legislation does not cover all the substances released. In some circumstances monitoring is carried out for substances that are banned, have been phased out or are no longer in use. It is recommended that any operator self-monitoring data that are measured and transmitted are reviewed for discrepancies and current relevance, though the minimum needs to comply with the legislation must not be overlooked.

From a functional perspective the existing legislation is summarised in Annex B. This list is a complete list of legislation taken from 'Enforcement Practise General Guidance' Ref :OP/OP/009 V1 05/96 and many of these clearly affect self-monitoring data.

Other instruments of guidance mentioned in the questionnaires include:

- Waste Management Papers 4, 26 and 27; and
- Chief Inspectors Process Guidance notes.

2.1.1 The Future

As discussed in Section 2.1 most legislation that has an impact on self-monitoring originates from EC Directives that are translated into UK Statutes and Regulations. There are some directives and guidance that are currently being considered and the additional requirements that these may place on operators in a monitoring context should be taken into account in any subsequent work. Specific examples include:

- the Urban Waste Water Treatment Directive (UWWTD);
- the Landfill Directive;
- the Groundwater Regulations (due to be implemented in January 1999);
- the Contaminated Land Regulations (to be implemented in 1999?);
- the Water Framework Directive;
- the Integrated Pollution Prevention and Control Directive
- internal Agency guidance on the implementation of Regulation 15 of the Waste Management Regulations; and
- Waste Management Paper 26D.

Some of the above which are imminent, such as the Groundwater Regulations and wider implementation of Regulation 15 of the Waste Management Regulations, both of which are associated with the Groundwater Directive, will potentially have a wide ranging effect on the volume and scope of data that operators will be required to collect and transmit to the Agency. Whilst most of the data will consist of water quality analyses, a significant amount of text to describe and support these data will also be required.

2.2 Primary Interface with the Operator

Primary business need.

From the information gathered during this project, and from background knowledge, the primary business needs of the Agency related to self monitoring data at the interface with the operators are:

- specification of monitoring requirements;
- monitoring compliance;
- making monitoring data available to those permitted to access it; and
- other needs (e.g. 'scientific investigations'). These are considered here as secondary uses of the data, although within some functions they may form part of compliance monitoring or the users may be the same as those with the primary needs as grouped here.

There are several steps in monitoring compliance:

- ensuring that the monitoring scheme is compliant:
    - the correct sampling methods are used;
    - monitoring is carried out at the correct frequency;
    - the correct determinands are measured;
- checking that the values reported are within the prescribed limits.
- checking that the field data are accurate:
    - equipment is calibrated;
    - equipment is operated by properly trained staff;
    - free from transcription errors;
- checking that the data is real - not subject to dishonesty; and
- checking that an appropriate analytical method has been used and has been reported correctly.

During the later part of this study the importance of the distinction between those needs that require the values present in self-monitoring data as opposed to those that are an audit of the monitoring process will be discussed.

Making monitoring data available to those permitted to access it can be broken down into:

- access by internal staff (includes 'restricted' data);
- access as part of Public 'Service' (controlled information);
- commercial use (controlled data); and
- sufficient access to the EC for statutory requirements.

2.3 Processes that require self-monitoring

The most important relationships between the Agency's business needs and the transmission requirements for self-monitored data are best seen from a categorisation of the characteristics and relationships between processes, rather than a list of individual processes. For some aspects of need (e.g. the specification of equipment) two processes in different industry sectors or regulatory regimes may have much in common, whilst in others (e.g. the requirement to interface with the Agency) they may fall into different groupings.

Thus several overlapping groupings can be recognised.

Categorisation by Regulation. Regulatory regimes may bring together processes with different technical monitoring requirements which are treated together for interface with the Agency. The technical requirements of individual processes that are controlled under one set of regulations may have more in common with, or be identical to, those falling under different regulations. For example, they may all be 'aqueous liquids that come out of pipes'.

Categorisation by Industry Sector. Many industries carry out a large range of different processes, often covered by different regulations. Thus contacts with industry associations, or the focus of policy within the Agency may draw together concerns over processes which would be grouped differently by regulation or technical requirements. From the perspective of operators, a centralised monitoring function may have requirements arising from more than one set of regulations and several interfaces with different Agency functions.

Categorisation by Operator size. Operator size often has a major influence on the technical resources available for monitoring. For example, in a large organisation monitoring may be the responsibility of a specialist group with a dedicated manager (e.g. 'Environmental Manager'). In smaller operators these roles are more likely to be only part of the job function of the staff involved, or may be contracted to a third party monitoring organisation. These factors have a large influence on the data transmission techniques that an operator can, at present, be expected to perform reliably

Categorisation by Media Sampled. The media that is being sampled by monitoring has the dominant effect on the equipment used and may be summarised as:

Gases
- at vents and stacks;
- the atmosphere;
- in the soil;

Aqueous liquids	-	at outfall or discharge point;
	-	surface water;
	-	ground water;
	-	soil water;
Solids	-	deposited material;
	-	soils;
	-	particulates in gases or liquids.

Categorisation by Mode of Discharge or Release. How the potentially contaminating or polluting material reaches, or might reach, the environment affects both monitoring equipment and procedures. Two main criteria can be recognised:

- Whether the impact on the environment is;
    - part of the normal operating process and must be kept within prescribed limits,
    - or whether it occurs only as a result of malfunction or accident (e.g. fugitive) and must be detected.
- Whether the discharge is;
    - of a controlled nature through a pipe or stack or as a deposit,
    - or is of a diffuse nature which must be limited (e.g. evaporation during handling).

Monitoring of controlled discharges aims to ensure that what is permitted stays within prescribed limits and that unauthorised discharges do not occur. Monitoring of ambient conditions seeks to ensure that anticipated diffuse emissions or discharges have not caused specified limits to be exceeded and to detect unauthorised releases to the environment.

Recognition of these groupings has been prompted by comments in many sections of the questionnaires returned (both business needs and data), supplemented by our background knowledge. It appears that different groupings are important to different Agency functions, and that for secondary use, self-monitored data are often transmitted within the Agency so 'packaged'. For the recommendation of data transmission methods the 'media sampled' and 'mode of discharge' categorisations are of most importance, but the other categorisations are likely to be of importance in introducing any change in methods.

The questionnaire results and our background knowledge do not permit a complete cross-referenced tabulation of these different categorisations. Such a tabulation is not needed in order to assess data transmission needs. However, it is suggested that such an analysis may be useful when planning the implementation of any revised transmission methods.

2.4 Categories of Data

Monitoring data can also be categorised in several main ways. Interfaces with the operator and data flow within the Agency are determined mainly by Agency function and, at a lower level, by the governing regulations. Regional differences in procedures and data flow exist at

present, but these will progressively diminish as part of the convergence/harmonisation process.

For the purpose of development of transmission methods a more general categorisation emphasising mode of discharge (emission, release etc) and the receiving medium is more relevant.

Data representing concentrations of substances, mainly chemical compounds include:

- gas, vapours, and mists entering the air through vent, stacks and escapes;
- aqueous discharges from outfalls;
- particulates in gas and aqueous discharges;
- soil gas;
- surface waters;
- the atmosphere;
- groundwater and soil water;
- slurries; and
- solid materials deposited (mainly wastes).

Measurements of mass flow include:

- discharge rates from stacks and outfalls;
- water abstraction rates;
- surface water flows; and
- quantities of material deposited (Solid waste and slurries).

These data are often accompanied by other information such as:

- meteorological data;
- groundwater level data;
- gauge board readings, reservoir levels;
- operational data (e.g. waste management facility returns);
- quality data (e.g. calibration data and procedural records); and
- a variety of descriptive text.

For the purpose of data management and interpretation, categorisations include:

- is it raw data, or pre-processed (e.g. statistical summaries or collation);
- transmission method (e.g. on paper, on computer disk, by e-mail, by modem link);

- archiving method (e.g. paper files, computer database); and
- access (e.g. public register, commercial data, internal restricted data use for raising charges).

This variety of categorisation allows a set of descriptive attributes to be attached to each dataset for use in developing data transmission and management methods.

2.5 Secondary Uses and Needs of the Data ?

As discussed in Section 2.1 and Section 2.2, there are numerous primary reasons for receiving operator data. However, having fulfilled its primary objective there are other uses and users of the data and these secondary and tertiary purposes naturally enhance the data value. Secondary and tertiary uses include:

- placement of raw and filtered data on the public register;
- transfer of annual release data (from IPC) to the Chemical Release Inventory;
- additional databases for central data management;
- transfer to users for the production of annual performance statistics, e.g. Environment Snapshot;
- resource evaluation and status;
- adhoc analyses for broader environmental monitoring initiatives and trend detection;
- compilation for reports to Agency sponsors (i.e. DETR, MAFF and the Welsh Office).

2.6 Internal Receipt, Transfer and Storage

The Agency reported a basic data flow diagram its 'Data & Information Policy Handbook' [Agency, 1996b] and this is replicated in Figure 2.1.

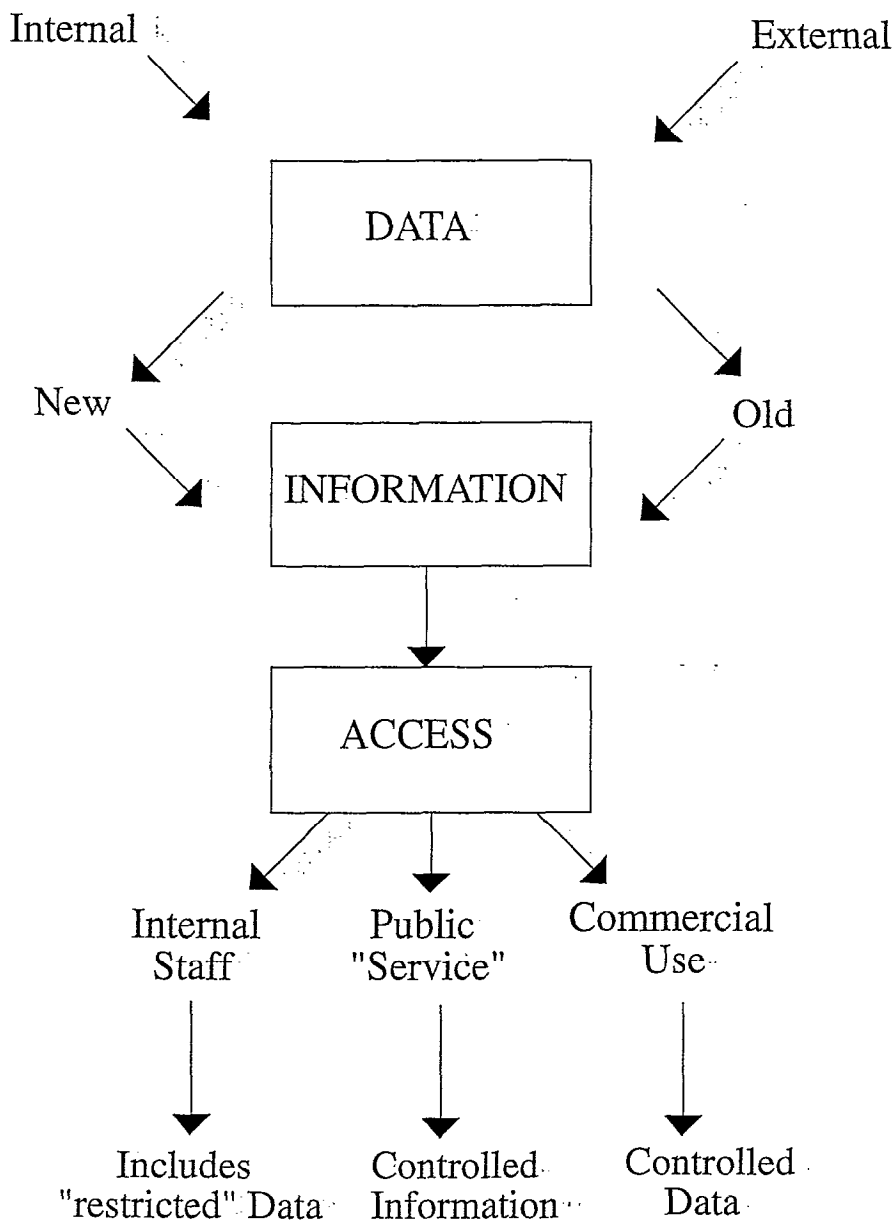


Figure 2.1: Data Flow in the Agency (Source: Data and Information Policy Handbook, [Agency, 1996b])

A more detailed example of a data flow diagram was received from the Waste function and is illustrated in Figure 2.2

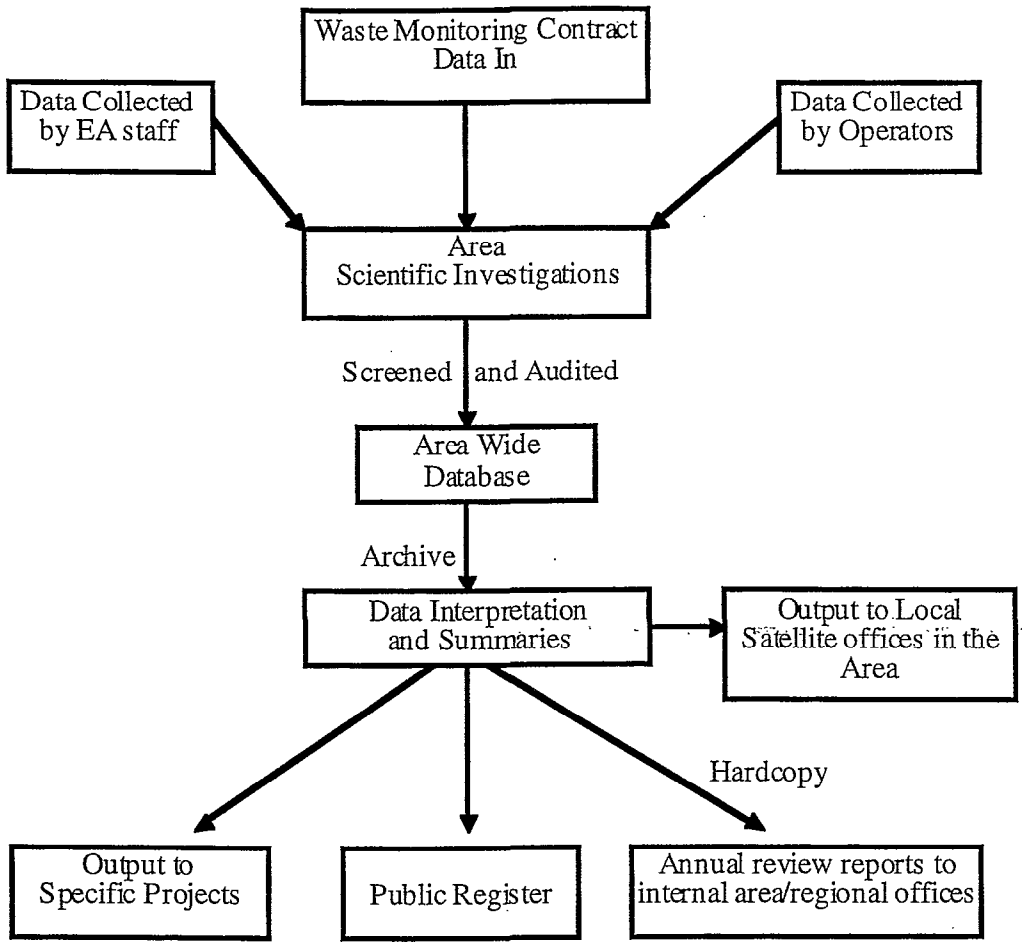


Figure 2.2: Example Data Flow Diagram for Waste Function

Clearly it would be of benefit if each function could agree on a generic data flow diagram. This would improve the understanding of the secondary and tertiary uses and users of data and perhaps enable their requirements to be better appreciated and incorporated.

2.7 Related Agency Initiatives

It is important that we do not look at the issue of self-monitoring transmission in isolation. It is relevant to consider wider Agency initiatives that may have a direct or indirect influence on the future requirements and management of self-monitoring data, both internally and externally. Many respondents indicated they were not aware of any Information Systems strategies that would influence the way self-monitored data is managed, the remainder suggested:

- convergence;
- Water Information Management System (WIMS); and

- public access policy.

2.7.1 Convergence

Convergence is the Agency initiative towards standardisation of information systems throughout its functions to improve the consistency in the customer interface and protect their position from the year 2000 threat. Unfortunately the adoption of an integrated information system across all functions and regions, that is year 2000 compliant, is beyond the timescale and available budget. Hence a 'best interim solution' is being pursued that involves cutting the number of applications used and performing a technology transfer of the most suitable systems, which will also undergo some improvement. Securing a pan-Agency consistency by standardising large parts of the system will clearly impact some of the systems that are currently used for managing self-monitoring data. We cannot make any recommendations concerning convergence until the following detailed issues are resolved:

- which systems are used for managing self-monitoring data;
- how widely they are adopted;
- what systems will remain as part of the convergence strategy;
- what new systems will be introduced; and
- what systems will be transferred across regions and functions as best practice.

However, following a series of workshops held across all regions and functions [Agency, 1997c] the requirements of the new information system included:

- the year 2000 time-bomb be defused;
- the technical infrastructure should be strengthened;
- geographical information systems should be incorporated;
- an integrated planning authorisations system be adopted;
- more accessible public registers;
- improvements to the authorisations systems - particularly for waste;
- support for people working away from the office; and
- improved internal communications and sharing of data.

Recommendations from CIS will be important and early awareness of the strategy will enable an early migration towards the preferred solution. It is unclear at present what the exact impact will be on the data management responsibilities and processes. These need to be published as soon as possible if potential changes to self-monitoring processes and practices are to efficiently incorporated.

2.7.2 Water Information Management System (WIMS)

Currently WIMS is a software package for the recording and storing of data relating to water quality archive details, consent information, incident details, water flow details and abstraction information. This archive system was originally developed to replace Wessex Water's IBM mainframe-based water quality archive and has three main functions:

- to store and allow selective retrieval of the results of chemical analysis of samples taken by the Agency;
- to produce reports on: sample results by sampling point, date range and option determinands including basic statistical analysis; non-compliance with consents; and sampled points to compare progress against targets; and
- to produce and manage operators letters.

Agency sample results are transmitted to the system electronically via an overnight process whilst external laboratory data is entered manually.

The system is currently used in the South Western and Southern Regions. Numerous conversations and meetings raised the potential of WIMS as a much wider database system for archive and retrieval. Post-processing capability is limited, but FTP protocols allow fixed format ASCII files to be exported from the VAX to PC's for further processing [Green, 1993].

It would useful to perform a detailed user evaluation study of the existing system before recommending it. For example, its primary purpose was to provide a water quality archive at the regional level, is it fully scalable in terms of data volumes and types ? What are the data access and post-processing requirements of the different functions ? A requirements-gathering and specification phase would be a very valuable and timely exercise.

2.7.3 Public Access

The Agency has two main statutory obligations for public access to environmental information, namely the public register service and the 1992 Environmental Information Regulations (the UK implementation of the EC Directive 90/313/EEC - Freedom of Access to Information on the Environment). A number of early initiatives were set up to ensure the Agency met its statutory obligations and these included the appointment of a Public Access Co-ordinator and a pre-Agency Public Access/Public Registers Working Group, which has now been superseded by the Statutory Public Access Working Group (SPAWG).

Following hard work from operational staff and SPAWG, a National Library and Information Service (NLIS) is being set up based on a network of information centres. This service includes repositories to all the Agency's past and current R&D reports, training material, publications, leaflets, pamphlets and CD ROMs.

There are a number of formal mechanisms whereby monitoring data are made available to the public, including:

- public registers;
- Environmental Information Regulations; and
- Open Government Code of Practice.

In addition, data are made available following individual requests and occasionally in response to planning consultations, legal proceedings and Public Inquiries.

The current public registers include:

Water Abstraction and Impounding Licences (S.189 WRA 1991);

Water Quality and Pollution Control (S.190 WRA 1990);

Agency Works Discharges (S.191 WRA 1991);

Maps of Freshwater Limits (S.193 WRA 1991);

Main River Maps (S.193 WRA 1991);

Maps of Waterworks (S.195 WRA 1991);

Integrated Pollution Control (S.20 EPA 1990);

Radioactive Substances (S.39 RSA 1993);

Industrial Air Pollution (SI No.318/1989 & E&SIA 1988);

Genetically Modified Organisms (S.122 EPA 1990);

Waste Management Licences (S.65 EPA 1990);

Carriers and Brokers of Controlled Waste (SI No.1056/1994);

Special Waste Notifications (SI No.1056/1994).

The current problem is that there are few staff dedicated to the issue of public access, hence the result is that managing the registers or dealing with requests, which is ever increasing, does not get sufficient attention. Secondly, there are a large number of different corporate, regional and area information systems which provide elements of the public register and only two dedicated IT-based public register systems. Moreover, the Agency must be able to demonstrate that the registers are readily available to the public, which has implications for the siting of access facilities. Hence, future effort needs to be directed towards common standards for new IS/IT technologies (see convergence above) consistency of data quality and where these systems will be managed and accessed.

In 1996 further support was also offered in the form of Customer Service Centres (CSCs). These were officially trialed in Welsh and South West Regions. These were a useful one-stop-shop but highlighted the need to ensure they are appropriately staffed. These have proved successful and now customer contact teams are available at the Area Customer Service Centres. (N.B. Regions, HQ and National Centres also have some teams to deal with public information requests). Two initiatives are being explored in this area at present and include:

- co-location of public registers with CSCs; and
- computerised access to electronic data.

Although the resource implications require detailed examination both issues would affect the handling of operator self-monitoring data.

3 Review of Self Monitoring Techniques and Recording Instrumentation

This section reviews the monitoring data generated in terms of type of discharge, parameters measured, monitoring strategies and the types of instrumentation employed. The review was carried out with the aim of identifying generic features of monitoring. These are used later to examine the options for data structures, formats, transmission and archiving.

The survey using the Data Questionnaire, provided overviews and examples of monitoring data from a wide range of Agency functions and Regions, and a large number of opinions and suggestions. Some respondents did not distinguish completely between Agency monitoring and monitoring by operators. This does not disadvantage this study since the results indicate that the Agency's own methods are in many cases the most convenient descriptions of those that are specified in new authorisations. For example, the Agency's 'Regulatory Environmental Monitoring Contract Specifications' were provided by two respondents as good examples of how self-monitoring would be specified.

The survey confirmed the huge range of processes monitored, parameters measured and equipment used by operators during self-monitoring. Tabulation of the processes, parameters, equipment and methods would be lengthy and difficult. The most useful approximation is to assume that the Agency regulates almost every type of industrial, commercial and agricultural process in England and Wales, and requires self-monitoring by a large proportion of operators, that almost every imaginable parameter is measured, and that operators use almost all of the equipment available to measure those parameters during monitoring. One respondent ended the most comprehensive list of parameters with 'for others see the periodic table'.

The results of the survey were used to identify generic features of self-monitoring; these are described in this section in the main groupings that are found to occur. As indicated in Section 2, the groupings often cross industrial sectors, regulatory regimes and Agency functions. Views expressed by respondents to the survey are included in this section. Together with our analysis of the survey results and background knowledge, they form the basis of many of the recommendations of this report.

3.1 Media, Discharge Type and Monitoring Requirements

The media sampled and how discharges, emissions or other releases to the environment occur are the main factors controlling the monitoring procedures and equipment used, hence the nature of the data to be transmitted to the Agency.

3.1.1 Stacks and vents to the atmosphere.

These fall almost entirely within the IPC (PIR) function, and are subject of thousands of pages of detailed guidance notes. They are mainly part of continuous processes, usually with some form of control instrumentation. They are thus suitable for the installation of in-situ monitoring equipment which can, and often is, linked to centralised recording equipment. For many processes air monitoring is also required, often in combination with local meteorological data. As with most other processes, sampling for laboratory analysis is required for measurement of some determinands and checks on in-situ equipment, and spot sampling is also undertaken.

Monitoring equipment is often used for process control (at the simplest level, maintaining combustion efficiency) and because of this they are the main processes where operators may do more monitoring than is required by the Agency. The Agency may thus receive statistical summaries of determinand values over the required periods, the source data for which may come from continuous monitoring.

3.1.2 Discharges to surface waters.

These predominantly fall within the water quality and waste management functions. As with stacks they are amenable to fixed monitoring installations, but geographic location often makes installation of centralised recording facilities less convenient. For these discharges it is less common for process control to require more frequent monitoring than required by the Agency and it appears that sampling and laboratory analysis makes up a higher proportion of the data. However, from the survey, it appears that transmission of data and alerts by remote telemetry seem more common than for most processes under IPC, probably because of geographic location and lower staffing levels.

3.1.3 Surface waters.

The survey suggests that the majority of surface water data are received from the water industry. The more detailed accounts given were based largely on the Agency's own monitoring, but the methods and techniques used by operators can be assumed to be similar. Laboratory analysis of water samples appears to be the main source of data, sometimes derived from automatic samplers. Flow, level and abstraction data, although not related to discharge of potential pollutants, are part of the operator monitored data. In some regions technical links remain between the regulators and abstractors, with shared telemetry and environmental data.

3.1.4 Groundwater.

Groundwater quality and abstraction rate data are routinely provided by operators in the water industry. Although not discharge-related they are the source of considerable volumes of self-monitored data, of concern to different functions within the Agency. Chemical data are predominantly from laboratory analysis of samples, often in the operators own laboratories.

A separate source of self-monitored data is from monitoring boreholes related to specific processes. Examples include chemical plant with a long operating history, which often represent a threat to groundwater quality, and landfill sites where groundwater monitoring is a test of effective leachate management. Data (other than water level data) are mainly from samples taken for laboratory analysis, with many determinands specific to individual situations.

3.1.5 Air

The survey provided little information on air monitoring by operators. However, it appears to be largely related to fugitive emissions and releases under IPC. Both fixed monitoring installations on masts and spot sampling using portable equipment, are carried out. Meteorological data, especially wind speed and direction, normally form part of the data required.

3.1.6 Soil Gas

Routine monitoring of soil gas is required of landfill-site operators as a test of landfill gas management procedures. Data provided to the Agency are mainly from fixed standpipes using mobile equipment, although laboratory analyses are also received. On a less routine

basis the Agency may receive data from temporary probes using mobile equipment, testing for landfill gas or a range of volatile organic compounds.

3.1.7 Solids and sludges

Self-monitored data from solid material appears to be largely related to the regulation of waste management and radioactive substances, although little information was available from the survey. For waste management, much of the data are descriptive (e.g. the source, nature and volume of material spread on land or deposited at landfill). Analytical data invariably are the result of laboratory analysis.

3.2 Types of Parameters

A list of parameters reported by self-monitoring would be large and open-ended. However, some loose groupings and common features important to data transmission can be recognised.

3.2.1 Units of measure

Clear communication of units of measurement is an essential element of data transmission. They are part of the data, and it is equally as important that they are not corrupted or lost, as it is for the determinand values. In some examples given in the survey, they are specified as part of the licence or consent, in others they are not and may change with time. In receiving and assessing data, conversion of units can be a significant part of the effort. One respondent observed that because of the wide range of monitoring instruments in use there was often confusion over what the units were used where.

In our own experience, for many important chemical determinands there is frequently confusion between results quoted in milligrams and micrograms per litre, in elemental or compound form (e.g. N or NO_3) and for Cation Exchange Capacity (used in landfill assessments) results can be quoted in mg/100grams or mg/kg.

Measurement conditions are a related concern. Gas data may or may not have been converted to Standard Temperature and Pressure (STP), and for other parameters they may be similar variations in the way results are presented. Need for some form of conversion may also arise from the way in which analytical data are presented. For example, some determinands may be reported as equivalents of some other substance. Landfill gas is reported by some instruments as a percentage of the 'lower explosive limit'.

3.2.2 Concentrations of substances.

If concerns over units of measure and conversions between different representations are set aside, concentration data represent a relatively simple and predominant component of most data sets. As several respondents to the survey pointed out, the list of determinands could include almost every chemical compound. Depending on the process, a data-set could contain two or three determinands, or many tens. There is currently discussion within the Agency over the extent to which a standard list of 79 List 1 substances should be covered in screening analyses for landfill leachate.

For transmission and receipt considerations an important factor is the groupings within which determinations are carried out. In the majority of situations it is necessary to use more than one instrument or analytical technique, such that the data relating to one sampling event may be transmitted in different data sets.

3.2.3 Detection limits.

Representations for recording values below detection limits, the limits themselves, and missing data vary. For most analytical processes there some value below which determinands are reported as 'not detected'. From our experience it is sometimes difficult to establish from manufacturers or laboratories what the 'detection limit' means. For example, how does the precision of a value at the limit of detection compare to one fractionaly below it. It must be recognised that detection limits vary with the analytical technique employed for a single determinand and may even vary between analytical runs.

For some sources (e.g. flow meters) readings of zero are often recorded and reported as such. For others values, below detection limits are variously reported as 'less than the detection limit', 'half of the detection limit' or zero. In some cases the value reported by the instrument is recorded even when below the detection limit; in some situations where statistical analysis is required this may be a correct approach.

How values below the detection limit are treated must be taken into account when applying statistical techniques, or reviewing data provided as statistical summaries. However, in most cases equipment used and acceptable levels of determinands are specified such that treatment of low readings does not compromise regulatory control. That is, the critical values are sufficiently higher than detection limits for this not to be a concern.

3.2.4 Missing data

Missing data also increase the effort associated with statistical analysis. A greater concern is that some data management software has difficulty handling null values, or non-numeric data used to indicate them. This also applies to non-numeric characters used to indicate values below detection limits. The most common result of failure is that a dataset goes forward with such fields replaced by zero.

Some respondents pointed out that if data were missing an important part of regulatory control was to find out why.

3.2.5 Electrical conductivity, salinity, pH, radioactivity

A number of chemical characteristics are measured indirectly, by some measure which indicates their environmental impact. These parameters are handled in much the same way as concentration data, except that care must be taken when making and using statistical summaries. For example, whether the maximum, minimum and mean of a set of pH values is meaningful depends on how the environmental impact may occur, and how prescribed limits are set.

3.2.6 Rates of flow, weights and volumes.

Many data from outfalls, stacks and abstractions include flow data. In some instances these have been calculated indirectly, for example from pump capacities and running hours or even depth of water over a weir. The data are converted before transmission to the Agency but this may raise concerns over calibration.

Waste management operators are often required to provide monthly returns of waste inputs and remaining capacity (for landfills).

3.2.7 Level data

Reporting of the level of groundwater in boreholes requires that surface and datum levels also form part of the data set. The specification normally requires that the levels be reduced

to Ordnance Datum before transmission. A similar procedure is often required for gauge board readings on controlled watercourses or reservoirs.

3.2.8 Quality control data

Increasingly, there is a requirement to add quality control information to the raw data noted above. In the case of chemical analyses, the method of analysis or accreditation status of the technique may be needed. For data measured in the field and groundwater level data, the type/number of the equipment used to take the readings is an essential piece of quality control information, as changes in the equipment may lead to fault lines in the data.

A similar situation may arise with flow data, where a change in the method of measurement or measurement equipment may lead to differences between data points in the overall dataset.

Therefore, it is essential that adequate provision is made for the addition of quality control data, usually in the form of qualifying text, to the raw or processed data undergoing transmission.

3.3 Recording Instrumentation and Sampling Procedures

Recording instrumentation and sampling procedures, in part specified by the Agency, determine the 'packages' of data produced by the monitoring. Data management procedures used by the operator for collation or pre-processing of the data, discussed below, generally result in a smaller number of larger data sets being transmitted to the Agency.

The survey shows that in many cases almost 'raw' data (e.g. certificates of analysis from laboratories) are passed to the Agency. Even where data are collated by the operator, the diversity of monitoring specified, for example, for a landfill site or petrochemical plant, results in a large number of distinct sets of data being received by the Agency.

Sampling or monitoring procedures resulting in different types of data set include:

- continuous monitoring with time averages or statistics at intervals;
- fixed-interval monitoring by fixed equipment, individual readings or summaries transmitted;
- fixed-interval monitoring by mobile equipment;
- fixed-interval sampling, automatically or by operator visits;
- multi-point sampling or monitoring by operator making 'tour' with mobile equipment.

Fixed-intervals may vary from one minute to one year, although the range 15 minutes to 3 months excludes all but the most extreme cases. One respondent gave the range as 'one sample per month for a cement works, daily for a petrochemical plant'. Where movement of the operator between sampling points is required the time intervals between visits, monitoring or sampling events may vary. Some nominally contemporaneous data (e.g. quarterly monitoring of points on a site) may be spread over several hours.

Each monitoring event at a monitoring point may result in several distinct data sets. For example a visit to a borehole may result in descriptive data, manual measurements (of water level), data from one or more portable instruments, and results from one or more laboratory

procedures on groundwater and/or gas samples. Monitoring of many outfalls or stacks results in simpler datasets, for example, simultaneous readings from several in-situ sensors, with lower frequency analyses of samples collected. Some programmes simply require monthly analysis of single samples.

The survey showed that, except for monitoring using in-situ equipment, most of the data recording is manual. The Agency typically receives tabulations of results, with formats that differ between operators, compiled from written records and presented on paper; these are often accompanied by certificates of analysis from laboratories. One respondent, in giving the estimated data volume as '400 A4 pages per month', provides a succinct summary of typical current data transmission methods.

For some processes, particularly those producing discharges from stacks and outfalls, electronic logging of data is more common. This provides larger quantities of data and succinct summaries. Because of the relative ease of handling these data it produced disproportionately few comments in the survey.

The use of data loggers contained within mobile equipment, attached to multi-channel probes, or providing electronic alternatives to notebooks is becoming more common. Many Agency respondents to the questionnaires suggested that more extensive use by operators would result in data sets that were easier to assimilate into Agency systems. Operators with a large monitoring requirement or centralised monitoring function were also enthusiastic about their introduction.

The character of the datasets produced by different recording equipment, sampling and monitoring systems is discussed in more detail in the next section.

4 Assessment and Recommendation of Data Transmission Scenarios

The output from this task will be guidelines for data transmission from remote collection points to an Agency database. There are numerous options, and it may be that the most effective solution will be site-specific. For example, the recommended solution will be dependent on the type, volume and format of data to be transferred, as well as the architecture available (e.g. present lines of communication including type of link, alternatives available, databases). From the industry's perspective it will be important to utilise the available infrastructure where possible, but this must be balanced with the need for a robust and standardised approach for consistency and efficiency for database storage. Potential lines of communication vary from hardcopy transfer (post and fax), disk, e-mail, real-time, modem (ISDN) and internet.

4.1 Data Transmission Technologies.

Following an examination of current practice, we recommend an approach based on legislative requirements, and the type and volume of data transfer. The system recommendation must be 'fit for purpose' i.e. based on the risk of error or non-compliance. By comparing each process to a generic class of "gathering/transmission" types, a more uniform approach can be adopted. This will generate a more robust system thus reducing the effort for the Agency and operator alike.

While there are many possible methods of transmitting data between collection sites and a data repository, we can generally characterise any transmission method as either:

- traditional (that is hardcopy);
- via a portable magnetic medium; and
- via direct network connection.

The traditional methods would involve hand transcription of data, or downloaded printout and transmission by post or by FAX. Examples of portable media include floppy disks, zip disks, jaz cartridges, Digital Audio Tapes (DAT), SyQuest cartridges and CD's. Examples of direct network connections include, modem connection, ISDN connection and use of the Internet.

Section 4.4 discusses the potential longevity of the technologies and attempts to assess future developments and the likely on-line support that will be available. It also attempts to understand the practicalities of implementing alternative solutions within operators IT frameworks (including the Agency's).

In the following discussion we shall review some of the advantages and disadvantages of each of the approaches mentioned above (with respect to the operator and the Agency).

4.1.1 Traditional Methods

Traditional methods are likely to be applicable only to those situations where small data streams are transferred infrequently. Data transferred by these methods are unlikely to be automatically validated and rely on human interaction for checking and filtering of anomalous results. This method also requires the most effort at the Agency end to cross-check and

digitally store the results. It is also the prevailing means of transmission and for many small datasets will remain so for a considerable period in the future.

4.1.2 Portable Magnetic Media

A general note

It should be clear from the outset that the use of any portable medium for data transfer can only be a partial solution at best, since the actual transfer must be achieved by moving the medium itself from site to site. Despite this, it is still possible that such technologies (in conjunction with a courier system or perhaps the Royal Mail) would give the most appropriate solution for certain types of data transfer, e.g. infrequent transfers, annual report to the CRI. Although this area has been slow to develop, new technologies are entering the marketplace and the medium term future may look completely different.

Floppy disks

The most common magnetic media with which personal computer users are most familiar is the floppy disk. Although almost all personal computers are capable of reading and writing to 3.5 inch floppy disks, their limited capacity (approximately 1.4 Megabyte) and slow access times means that the use of the floppy disk is likely to limited in the future. However, where data files are small, the convenience, reasonable reliability and ubiquity of floppy disks mean that they still can be serious contenders for data transfer.

Zip cartridges

Zip drives were introduced as an alternative to floppy disk drives, allowing storage of up to 100 megabytes of data on a single 3.5 inch disk. These disks have much quicker access times than floppy disks as well as greater capacity, but are more expensive. A Zip drive for a personal computer would be expected to cost approximately £100, with Zip disks costing about £10 each. Some personal computer manufacturers are beginning to provide zip drives in addition to floppy disk drives in their hardware.

Jaz cartridges

Jaz cartridges are useful where very large data files need to be transferred (up to 1 GigaByte). The larger capacity of the Jaz cartridges comes with a correspondingly higher price, around £300 for a Jaz drive and around £90 per cartridge. The performance of these drives is good, but the high price of the cartridges means that these should probably only be considered where it is essential to send very large files.

Digital Audio Tape

Digital Audio Tapes allow extremely large files to be transferred, with specifications varying from 2 Gigabytes up to 24 Gigabytes. The tapes are relatively cost effective at about £10 each, but the DAT drive itself tends to be expensive (up to £1000). Perhaps the biggest disadvantage of the DAT solution is the relatively slow read/write times, due to the serial manner in which data is stored on the tapes.

SyQuest Cartridges

SyQuest cartridges have been with us for some time and can be considered dated technology. However, SyQuest has launched a low-cost device called, Sparq which costs £150 and a set of three cartridges at £50-£60 can store 4GB. They feel this product will be a serious contender to the Iomega Zip drive and the LS120 drive (this drive also supports 3.5in floppy disks).

Compact Disks (CD's)

CD's have good storage capacity and are a robust medium. Their biggest disadvantage to date is that they can only be written to once and users obviously require a CD writer. Although not currently widely available a read/write version has been released, which will obviously overcome this problem. The current CD's do however, provide a stable medium to transfer and subsequently store large volumes of data. The fact that they cannot be written over can be a benefit.

4.1.3 Direct Network Transmission

A general note

All approaches to direct network transmission of data are very similar in their implementation, requiring some form of communications hardware and appropriate software to interact with the hardware. With an appropriate choice of software, it should be possible to use almost any hardware option with little or no difference to the end users. Assuming that an appropriate software mechanism can be decided upon, the critical choice is the hardware and the way in which the hardware communicates. As a result of this we concentrate here on the different hardware solutions which should be considered and outline their advantages and disadvantages.

Dialup modem connection

Modem connections are probably the cheapest form of direct network transmission. They make use of existing telephone lines and can transfer information at up to a maximum of 7,000 characters per second (56,000 bits per second). Modem technology is very common and many personal computers now come supplied with a modem as part of a standard package. For a single computer transferring moderate amounts of data, a single modem connection could be quite effective.

It is important to acknowledge that many users however, are not able to achieve 56K as this relies heavily on access to a digital telephone exchange and a high quality local loop. Improved performance can be achieved by keeping the data signal digital for as long as possible, hence some users may even struggle to achieve a 33.6K speed. The good news is that the US Robotics, X2 product and Rockwell, K56flex product (which also has a GSM mobile phone connection) are being supported by Internet service providers, and prices for both type of modem are falling. The installation process is simple with Windows 95 able to detect the drive and request the appropriate driver.

In the US further improvements have been announced, i.e. speeds of 128K but the limitation is the requirement for two telephone lines. This two-line modem is likely to cost $200 in the US. The higher uptake of ISDN in Europe may limit the exploitation of this system.

ISDN connection

Unlike the analogue technology used by a modem, an ISDN link uses fibre optics and the digital capabilities of the telephone network to gain significantly higher performance and flexibility. A single ISDN line allows transfer rates of up to 64,000 bits per second, but multiple ISDN lines can be seamlessly combined to provided transfer rates many times that of a single line. The ISDN line will require installation by a specialised provider, unlike a modem, but should be as transparent to use after installation. Thus ISDN connections are faster and more flexible than modem connections, but with a correspondingly higher price. N.B. with a standard 64k line a 1MB file would take approximately 2 minutes (i.e. 8 million bits at 64,000 bits/sec) at peak line performance.

Internet connection

Both modem and ISDN connections require charges for the calls made via the telephone system, while data is transferred. Where offices are widely separated, these calls will be charged at national rates, and if significant quantities of data are to be transferred these costs could be an important factor. One way to reduce these costs is to make use of the Internet to transfer the data. The use of a local Internet service provider allows all the calls to be made at a local rate to the service provider. The data transfer over the Internet itself is then free, although the Internet Service provider will charge a small monthly fee.

The major issue for the use of the Internet is likely to be the security of the data which is being transferred. This issue can be addressed relatively easily by using one of the many simple encryption mechanisms available for secure data transfer. Ideally, such encryption and decryption should take place automatically, without the need for significant intervention by the users of the system.

4.1.4 Preliminary Appraisal

Although we cannot ignore the current IT situation within the Agency and operators we must recommend practical solutions that look to the future. Given the rate of change of technology it is unwise to make any long term recommendations. This section briefly appraises the potential adoption of the above discussed methods in the short and medium term.

Short Term

The future has to be electronic transfer of data, even for small volumes. Hence, in the next 12-18 months paper transfer should be abolished, wherever posible. The medium to transfer this data, particularly for volumes <1Mb should be the floppy disk. For much larger transfers, Zip cartridges and CD's could be used. CD's are also a robust storage /archive medium. Digital Audio Tape and Jaz cartridges are only acceptable for very large transfers.

Medium Term

As described above, it is recommended that, wherever possible all data transfer should be in electronic form. Floppy disks, Zip cartridges and CD's are recommended as the main transfer medium. Direct Network Transmission is more practical for frequent transfers (e.g. daily). The training and set-up is greater for these methods but where they currently exist they could be used. It would require proper routing and set-up at the Agency to ensure a transparent transfer to a server.

4.2 Comparison of Existing Practices with Potential Options

Most of the raw and summary data received by the Agency is in paper form. There is some electronic transfer, via diskette but this is not the norm. Hence, there is little to compare with potential transmission options. In some ways this is a positive position because it means that existing practices do not have to be changed or modified. Also we are able to fit in with any other strategies, for example DSSG strategy, that may be introduced in the near future.

5 Data Management

Data management is a particularly important issue for a data-rich organisation. The integrity and consistency of the data is also vitally important from a public perspective. The DSSG have presented a number of objectives in their 'Data and Information Policy Handbook' and best practice should be implemented wherever possible.

5.1 Quality Control and Quality Data

Quality data include identification of equipment and operators, equipment calibrations records, training records of operators, details of laboratories used, precision of analytical methods, and procedural records forming part of a quality assurance system. Detailed monitoring procedures, often on a process by process basis, are given in documents such as guidance notes and waste management papers. Monitoring specifications typically include requirements to record information to allow audit of the monitoring process. It is vitally important that these data are always available for secondary users.

Monitoring guidelines set out in IPC authorisations specify that preference should be given, where possible, to analytical methods published by the British Standards Institution and related committees such as the Standing Committee of Analysts and so priority is given to these methods. Other standards highlighted are ISO, US EPA quoted Standards and ASTM Methods, and Standards from other European Standards Organisations such as DIN and AFNOR.

The results of the survey suggest that primary users see a clear distinction between quality data and the determinand values. All would prefer determinand values in digital form, but many point out that quality data, particularly that relating to monitoring procedures, is best assessed from the original documents. It was suggested by some that quality data not be transmitted to the Agency, but be available at operators sites in a form suitable for audit ("preferably unannounced"). These primary users appear to believe that secondary and tertiary users, and databases, are mainly concerned with determinand values, with compliance checking and verification having been carried out at the receiving office..

In contrast, we have received clear advise from central Agency functions that quality data *are* an important part of the datasets to be transmitted within the Agency. The presence of such data is important in allowing maximum use to be made of databases internally and externally.

Comments received during the survey suggest that this dichotomy of views is partially a result of resource limitations in the receiving offices. Several respondents commented on problems in resourcing data entry and verification of determinand values. The quality data, rarely used by them in analysis of the data represent considerable extra effort. They are also more heterogeneous and so less readily accommodated in the software most commonly used.

Some of these concerns can probably be addressed by identifying how and where different elements of the quality data are used. Some data required to allow primary users to confirm regulatory compliance may not be required by secondary users providing that a 'pointer' into an audit trail is included in the database. Other data, for example confidence limits on determinand values and indications of the data source (at varying degrees of resolution), will more commonly be needed the maximise subsequent use.

Quality data are complex in structure and storage and transmission can make up a large proportion of the cost of data transmission and management. Hence, the Agency should review:

- which quality data can be retained by operators for audit during site visits;
- which quality data must be transmitted for storage within the Agency; and
- which quality data must always accompany determinand data in Agency databases.

5.2 Summary of Data Management Techniques

Data management techniques by operators range from manual systems for logging of samples, and paper storage of analytical results, to integrated environmental data management systems. The former are typical of small operators from whom the Agency receives manual (or word-processor) tabulations and laboratory results. The latter are more common in large organisations which involve several regulated processes, with specialist monitoring groups, who routinely transfer data electronically.

The survey shows a similar diversity of methods within the Agency. In general those offices receiving small amounts of data on paper from a large number of operators (for example, in the waste regulation function) are limited by resources to paper-based management systems, i.e. data entry and validation are extremely expensive. Offices receiving large amounts of data in a standard format from a smaller number of operators (e.g. the water industry), or those with a less diverse monitoring requirement, tend to have a higher proportion of data held electronically.

However, there are variations across the regions and functions, and most offices have more sophisticated management systems for their own monitoring data than that received from operators. Where individual monitoring events produce diverse data, for example, those drawn from different instruments, the spreadsheet is used for manual collation; the results are copied from the original output into the appropriate fields.

Almost all portable data loggers, either stand-alone or integrated into monitoring equipment, transfer data via serial connection to a desktop PC. Data are normally received by a logger-specific utility which produces a computer file. This typically this consists of a header carrying parameters relevant to the whole set of readings, followed by sequential records of the same structure, usually either fixed format or delimited ASCII. Users typically load this file into the spreadsheet. More sophisticated utilities incorporate relational database functions and insert the data into a more fully structured database.

It appears that for many operators a spreadsheet fulfils their recording and reporting needs. Several examples of operator-supplied data provided by respondents to the survey consist of hard copies from spreadsheets. Typically these have sequential rows of similar format, separated by header rows and rows with data relating to a set of rows (e.g. date, operator, instrument and calibration data). Receipt of such data in electronic form, in spreadsheet format or as ASCII files was suggested by several Agency respondents to the survey. We are aware that several laboratories make results available in a similar form.

Despite use of spreadsheets by both operators and the Agency, there is a general awareness in both sectors that databases with more formal structures are a preferred option. All databases in use appear to be relational databases and as such fulfil the needs of both

operators and regulators to 'slice' data in different ways. For example, they may need time series data for a single point, data for points over a particular area for one time, or data for a particular instrument or operator.

From general knowledge we are familiar with the structure of several such database applications. They are generally called upon to manage data from a wide range of sources. A common feature is that, as the range of data to be stored increases, the data model becomes increasingly complex, incorporating more data tables and hence requiring increased development and maintenance effort. In other systems the data model becomes more abstract, with less obvious 'mappings' between the data as normally received or used and the internal structure. These are inevitable consequences of the large range in character of environmental monitoring data.

A number of proprietary database applications designed for monitoring data are available. The survey found one (Monitor Pro) to be in use in at least two Agency offices in different Regions. As the range of data to be stored increases, such systems show the same disadvantages for other relational databases (as discussed above). They do have the advantage that in aiming for a diverse marketplace, many of the problems presented by the wide range of input datasets have been addressed, with varying degrees of success. Import functions for data from a range of logging devices are usually available.

Within the various datasets and database applications of which we are aware a 'lowest common denominator' data structure is often present.

Spatial data.

Spatial data typically centre around either a *location*, including surface x, y and z co-ordinates amongst its attributes, or a *level* having height (or depth) offset from the *location. Locations* are grouped within the database into aggregates typically called *sites*, which carry attributes common to all *locations* (e.g. operator and co-ordinate datums). Conventions often apply as to what areal extent a nominally dimensionless location may have, and how far apart different sensors may be. All point data should be as accurate as possible for later visualisation in a GIS.

Temporal data

Temporal attributes often consist of combined *date and time* value. However, many datasets from loggers store date as 'header' information with clock time against individual records. Some datasets contain only date, with perhaps time as 'a.m.' or 'p.m.'. In the analysis and display of data some degree of aggregation of data collected at different time intervals is often required. Data collected at different times, for example, by an operator touring a set of monitoring locations are often regarded as contemporaneous during interpretation.

Parameter data

A common feature in systems with an 'open ended' parameter list is a table with attributes of parameters identified elsewhere only by an identifier. This may include the name, chemical symbol, units of measurement, trigger levels, analytical method, detection limits, null value substitutes etc. A common practice is for determinations of the same substance by different methods to be treated as different parameters, to be brought together during interpretation.

Quality data

Quality data such as operator and equipment identifiers usually have their own table in the database. Other quality data may be included as parameters or consigned to a general purpose 'note field'.

Determinand values

Determinand values are often received and required in a form such as (location, date, time, parameter1 value, parameter2 value...) against some defined parameter list. Database designers adopt a number of strategies to cope with the diversity of data sets. Some use multiple tables for different data sets, so that for example, daily data from fixed sensors is not held with monthly data from samples taken for laboratory analysis. Others adopt a more extreme strategy of storing the data in a form such as (location, date, time, parameter identifier, parameter numeric value, flags for such things as 'not determined', and pointers to quality data.). This results in great flexibility during data storage, but considerable effort when bringing data together for interpretation.

This review of data structures displays many of the issues to be addressed when transmitting monitoring data. In general terms, data within one structure in the operator's storage system must be re-structured and described for transmission, then re-mapped into the Agency's internal data structure. Some form of re-mapping from the structure generated by the original monitoring equipment or laboratory will already have occurred. The extent to which this is carried out manually, through electronic data import and export functions, or through a combination of the two (e.g. manual cutting and pasting between spreadsheets) has a large influence on the sources of error in transfer.

5.3 Storage and Verification Checks

There are three components to this task which include:

- the integrity of the data that is monitored and measured;
- the robustness of the transfer method; and
- the validation checks at the Agency end (identification of transcription errors, outliers etc.).

According to the 'Data and Information Policy Handbook', all data are subject to validation and quality assurance before being used or archived. This system does not appear to be consistently used with respect to self-monitoring data. There does not appear to be a consistent scheme for selection and appointment of a data manager and who is ultimately responsible for the quality of the data. The response to the question, "how is it confirmed that the received data meet the required specification ?", or even "include the correct parameters ?" included the following variations:

- data is manually checked by Agency;
- data is checked by electronic manipulation;
- the document is examined;
- the data is assumed to be OK;

- IPC field officer inspects returns;
- Manual check by field officers;
- all data validated by waste staff before entered into database; and
- checked against prescribed limits.

Appreciating that we are dealing with a number of different functions, which have inherited different protocols and guidelines for data management, there does not appear to be a standard procedure for data verification and checking as prescribed in the 'Data and Information Policy Handbook'. This corroborates the findings of Agency (1996a) where there was considerable variation in the responsibility and management of Public Register duties.

Clearly a more structured procedure is required for the verification and validation of external data once it arrives in the Agency, which is beyond simply relying upon the operator. Some procedures and standards were mentioned but they are not widespread. There may be an opportunity here to accrue best practice and perhaps deliver guidance to operators to ensure good procedures are adopted which naturally minimise errors.

It may be that it is not acceptable to prescribe responsibilities for data management across all functions, perhaps due to resource limitations. However, each function should attempt to clarify who is responsible and standardise as far as possible. This responsibility should be with staff that understand the data that are received and not just left to IT staff that are responsible for data entry and database management. There may already be internal processes for data management that could be naturally extended to self-monitoring data, e.g. WIMS. We do not want to make recommendations here that may interfere with the DSSG strategy. The NCCA have also offered its assistance in the process of checking returns.

5.4 Assessment of the robustness and security of data transfers

The current approach is fairly robust being reliant on the postal system. If and when data is lost further copies can be dispatched from the operator. Electronic transfer is open to corruption but this is rare and should not cause any operational difficulties. When confidential data are being transferred some form of encryption could be used but this seems rather excessive. Although floppy disks are a good transport mechanism they are not robust for long term storage.

5.5 Assessment of the data storage / archive

Having assessed the robustness and security of the data transfer, some self-monitoring data is entered into a database and it is important to understand how the data are maintained and accessed. The Agency has a statutory duty to archive and make accessible environmental information, and compliance data must reside on the Public Register. Most of the data are received in paper form and are therefore stored in paper files. It appears that some of these data are entered into a database but this is not consistent across the regions and functions. Some basic checks exist but there appears to be no standard method for checking the original robustness and transcription errors. Adhoc methods, such as random checks and spotting typographical errors, visualisation in spreadsheets, and examination of outliers were sometimes used. More robust techniques are required when the public have access to such

information and the procedures that are used for Agency data should be adopted where possible.

One potential problem of single data stores, particularly if access is not strictly controlled, is that the working files also become the public registers (particularly for the waste function). More robust systems with access control would be preferred to duplication/replication of datasets, particularly from a resource perspective. Instances have been recorded where extracts have been removed from the registers and not replaced. This can be embarrassing (and can lead to the Agency not fulfilling its statutory duties) when these are the only copies and the operator has to be approached to forward the data again. Clearly a more effective solution is required and electronic data may improve access and prevent loss of, or even corruption of data. The systems eventually used will obviously be dictated by convergence, but there are some general procedures that would improve the management of these data and include:

- specification of access control and security;
- defined responsibility for data entry and checking;
- appropriate guidelines for data checking, verification and validation;
- appropriate designation of master copies and working files (particularly important for subsequent internal transfer and use);
- a better appreciation of secondary uses and necessary caveats for use etc.

6 External Perspective

One important component of the project was to elicit an industry perspective. The external consultees were informed that the Agency planned to review their current approach and wanted the 'outside' view before proposing any recommendations. These organisations were thus given the opportunity to comment on existing practices and suggest what changes would support their position. Anonymity was offered, in case individual organisations felt that they may be prejudiced in the light of their responses. However, no organisation asked for this.

The Water Services Association (WSA) was also contacted on behalf of the water industry and asked to give their overview of this issue. Points of contact and relevant initiatives are discussed further in Section 6.2.

SEPA and DoE Northern Ireland were also contacted and sent a questionnaire. They were informed of the project and asked if they would like to contribute their own perspective. It appears that no similar projects are being funded by these bodies. They recognised it as an area that required further effort and they seemed keen to express their views.

6.1 Industry

A dozen industrial operators were contacted concerning the transmission of monitoring data into the Agency. These included a range of operator size and industry type. Finally however, only five organisations, namely Shell UK, Anglian Water, Oaktree Environmental, 3C Waste and AgrEvo UK Ltd found time to return the completed 'Data Questionnaire' and the results are summarised below.

The activities that these organisations are involved with in connection with self-monitoring include:

- drinking water supply;
- sewage treatment;
- trade effluent control;
- water leachate and gas monitoring;
- IPC part A processes.

A range of fixed and portable equipment is used. Different equipment is used depending on the purpose. A range of calibration frequencies are used, some daily, some annually, others left to the equipment providers, but generally they conform to national accreditation standards. Laboratories mentioned were mainly NAMAS accredited.

6.1.1 The Data

A range of parameters were measured and listed in the questionnaire but they are obviously related to the function and the specific conditions agreed in the licence/authorisation. When asked about data quantities transmitted the response varied from:

- approximating the number of pieces of data;

- naming specific data sets and frequencies of transfer;
- as little as required.

Not all send individual measurements, but everyone supplies summaries of the data. Not all measured data are sent to the Agency, but they obviously record other parameters for their own internal use. However, one operator indicated that additional data are sent. The reporting of missing data varies from operator to operator, some followed ISO9002 procedures (i.e. all missing data reported as less than), others added explanatory text to support missing values others said 'no guidelines available'.

6.1.2 Data Processing and storage

All operators store data in paper and electronic format with an increasing trend towards computer files. The Software used to hold data included:

- Microsoft Excel;
- Microsoft Access; and
- QuatroPro.

All data are sent in paper form, but all said there was a trend towards electronic transfer.

No guidelines were in place for the format of the data that are transferred, hence internal formats were used.

6.1.3 Future Data Transmission Methods

The general conclusion was a decision to move to electronic data transfer. One operator however, thought the existing system was appropriate given the amount of monitoring data transferred, but suggested e-mail if the volume increased. Other operators also suggested email. One operator felt electronic transfer would be more appropriate if their databases were compatible with the Agency's. The Internet was mentioned by two operators.

One operator was concerned with the idea of sending raw data that may not have been properly validated and felt more secure with sending data summaries.

Some new technical developments have been made in this area, particularly advances with portable and fixed sampling equipment, but no specific reasons were given of how this would affect the data transfer.

6.1.4 Recommended Changes

A number of personal comments were made on the pros and cons of the existing system and suggested recommendations included:

- more stringent quality assurance procedures;
- more precise guidelines and trigger levels;
- site specific rather than nationally applied guidelines;
- better on-line determination of measurements avoiding laboratories;
- reduce volume of data transfer;

- standardised proformas for monitoring returns; and
- electronic transfer of data.

In general, the operators already adopt there own quality assurance and data validation tests and use databases to store the monitoring information. They send paper files but in general are willing to submit returns electronically. They recognise the need for greater standardisation of formats and more site-specific guidelines.

6.2 Other Interested Parties

Detailed discussions were held with SEPA, DoE NI and WSA. Both SEPA and DoE NI were informed of the project goals and objectives and were offered the opportunity to also complete both questionnaires. They indicated that no similar projects were being independently funded, but recognised the importance of the subject. Regular contact was maintained after the questionnaires were sent but to date none have been completed and returned.

WSA were contacted and provided some useful background information. Brian Sparks and Ted Theirs were the appropriate contact points. They had established a telecommunications advisory group of which Keith Edwards at Anglian Water was the Chairman. Keith, in his capacity as Control and Communications Network manager, agreed to complete a questionnaire.

7 Conclusions and Recommendations

The study allows us to draw conclusions about the current and future data transmission methods and requirements, and to make a number of recommendations. Much time was allocated to the data elicitation exercise as it was deemed vitally important to assess current views and practices, rather than rely on contractor generic knowledge and perceptions. A list of key people was drawn up, and interviews and telephone calls used to formulate and develop the questionnaires. Subsequently 130 questionnaires were sent out to people in the Agency. The level of response was below our expectations and was probably affected by the workloads of the targeted personnel.

In hindsight the respondent performance may have been more effective if further interviews were arranged, rather than rely on proactive questionnaire completion. However, supervised questionnaire completion may lead to biased responses. Data questionnaires were completed by a number of external organisations and this proved a useful corroboration exercise.

Following the initial kick off meeting, the specification was amended to concentrate more on the business needs, as it was recognised that the actual transfer method is unlikely to be the critical factor affecting the management of operator data. However, electronic transfer may be the first significant step towards more efficient integration of the data within the Agency.

First and foremost a requirements gathering and specification exercise is required for handling data. This may already be underway under the auspices of the DSSG. This specification can then interact with the best interim solution to enhance the way forward.

A number of transmission technologies have been reviewed, some general technical developments may be useful, but others will be superseded by the time any recommendations are acted upon. The main recommendations are for:

- more stringent quality assurance procedures in monitoring procedure and data management;
- more precise guidelines and trigger levels;
- site-specific rather than nationally applied guidelines;
- increased on-line (field) determination of measurements avoiding laboratories;
- reduced volume of data transfer;
- standardised proformas for monitoring returns;
- electronic transfer of data; and
- that the implications of these recommendations should be properly discussed with the Agency's legal department.

7.1 Need for Business Objectives

There are a number of wider initiatives that need to be considered and the impacts of 'best interim solutions' assessed as early as possible. With respect to this and other activities it is important for the Agency to state clearly its objectives now and in the future. There is no

point in recommending protocols and guidelines that are not practical for both the Agency and operators (large and small alike).

From a business objective perspective the following recommendations are made:

- wider Agency initiatives should be published as soon as possible and broader implications assessed;
- data flow diagrams should be produced to support all secondary and tertiary uses and users of data;
- clear version control and robust access rights need to be established due to the wide range of interested parties, and to avoid costly duplication and management of datasets;
- data requests in all functions should be reviewed to ensure they are still relevant and not placing outdated requests on the operator;
- the implications of new regulations and legislation need further investigation and incorporation into the framework as soon as possible;
- to reduce the time taken to enter self-monitoring data the transfer should be in electronic form;
- a number of fixed data formats should be specified (if only to enhance the readability of the data);
- standard verification and validation tests should be specified for certain datasets (e.g., qualitative data / explanatory text should also be sent to primary data users to support the audit process);
- clearly defined quality assurance standards should be specified which include some operator responsibility for data integrity;
- there are numerous instruments used to inform operators of self-monitoring requirements and these seem satisfactory;
- clear internal data management responsibilities should be assigned that focus on data quality issues;
- functional level recommendations are appropriate, a data driven approach should be used.
- where possible standard systems should be adopted for data management (e.g. WIMS) to improve inter-agency data sharing.

7.2 Data Management

Having received the data it is the responsibility of the Agency to quality assure the data, provide adequate storage, enable multi-access and retrieval and provide clear data management responsibilities. The DSSG policy will clearly define these issues and improve the handling of self-monitoring data. Specific issues are discussed below.

7.2.1 Archiving

Regulations dictate that environmental data has to be stored and archived and be made available to public, education and commercial audiences. Convergence is recommending a best interim solution that will dictate the software systems available for data storage. A number of recommendations can be proposed:

- common systems should be used across functions and regions to support secondary uses and users of data (i.e. commonality of the interface for members of the public and Agency staff; same formats for subsequent users etc.)
- systems should have the flexibility to hold vast range of parameters at different frequencies and be able to hold explanatory character strings;
- they should enable efficient entry, archiving and access; and
- a detailed user evaluation study of WIMS would be a useful follow up of suggestions of its adoption as the basis of a standard Agency system.

7.2.2 Checking

Questionnaires have revealed wide ranging views and practices concerning the validation of data. Although the 'Data and Information Policy Handbook' talked of standard procedures, these were not recognised by Agency staff and were certainly not specified in any licences / authorisations. A more structured approach is required for the verification validation of external data. The following suggestions are put forward:

- clear specification of quality assurance concerning data collection and analysis;
- ensure data integrity is not the sole responsibility of the operator;
- establish standard data check procedures;
- establish standard reporting mechanisms including change control.

7.2.3 Format

There is a great range of parameters measured and transmitted to the Agency. The data include qualitative explanatory text which is particularly important the further the data is used away from the source. The only formats used by operators at present follow their own internal procedures. Standard data reporting formats would reduce the data management effort of the Agency. It is recommended is that:

- where possible, standard formats (proformas or computer file formats) should be provided to operators.

7.2.4 Version Control

Version control is important when managing key data. The study has found that within many Agency offices there are failures both to isolate working versions from wider access and to recognise master copies of data. The first is important because working files may be non-static and uncontrolled access may interfere with the integrity of the dataset. The second is important when data checking procedures are used and cleansing naturally removes or upgrades data entries. It is important that secondary users of data are using the final version

(and are so aware) and that there is no means by which conflicts can arise. It is recommended that:

- there should be appropriate designation of master copies and working files for all paper and electronic data sets.

7.2.5 Access Control and Security

Access control and security is a vitally important part of the data management process and is particularly relevant to a data-rich organisation. It has been established from phone conversations that working files in the waste function are regularly used as the public registers. It is unclear as to the access rights, but this is not good practice. First and foremost the public needs access to stable data and secondly they must never be in a position where they can violate the integrity of the database. Similarly, access needs to be controlled internally and the DSSG protocols need mentioning. Recommendations include:

- access control and security procedures for paper files should be specified and enforced;
- data management responsibilities need to be assigned and appropriate internal and external access rights awarded;
- powerful database systems that have strictly controlled access rights should be adopted (this will avoid replication of data and thus reduce maintenance effort).

7.3 Transmission Options

The current self-monitoring data transmission mechanism is nearly all paper. Although any recommendation for change will impact on everyone, at least we have a clean sheet and the most practical short and medium term solution can be selected that is in keeping with the 'best interim solution' and potential enhancements. Recommendations in the short and medium term are presented in the light of the volume and frequency of data transfer extracted from the questionnaires and include:

- Short term
    - gradual change to electronic data transfer;
    - use of floppy disks for small volumes and Zip cartridges for larger volumes;
    - DAT, Jaz and CD only reserved for very large transfers;
- Medium Term
    - continued use of floppy disks and Zip cartridges;
    - where DNT (including e-mail) exists enable transparent transfer with proper Agency set-up and routing (N.B. this technology is better suited to higher frequency transfers, i.e. daily).

If the Internet continues to grow in the future and security controls are improved it may be appropriate to enable connected users to transfer data via this mechanism. Ultimately the

Agency should try to reduce the effort taken to receive, archive and manage operator self-monitoring data. Migration to the above in the short and medium term will achieve this if the system is set-up correctly.

7.4 Legal Implications of Potential Changes

We must be aware that there are legal implications associated with any possible changes to data transmission methods. It is possible that whilst a move to electronic transfer of data is both desirable and acceptable to most operators, some maybe unwilling, and it may be illegal to insist on such transfer. It is recommended that the Agency's legal department is consulted with respect to this in the first instance.

References

Cremer and Warner (1993) Continuous Monitoring Instrumentation for Emissions to Air from Large Combustion Plant. DoE Report DOE/HMIP/RR/93/041, June 1993.

Cremer and Warner (1994) Calibration and Verification of Continuous Emission Monitoring Systems for Large Combustion Plant. DoE Report DoE/HMIP/RR/94/033, September 1994.

Environment Agency (1996a) Status Report: Public Access to Environmental Information. Internal Report, Environment Agency, August 1996

Environment Agency (1996b) Data and Information Policy Handbook. Internal Document, Environment Agency 1996

Environment Agency (1997a) Storyboard to Assists in the Population of The Agency's Catalogue of Data Sources, Version 6.0. Environment Agency, April 1997.

Environment Agency (1997b) Viewpoints on the Environment: Developing a National Environmental Monitoring and Assessment Framework. Date ?

Environment Agency (1997c) The Convergence programme: Why are we changing our information systems - and how will it affect us. Environment Agency internal note.

Green C (1993) NRA Wessex Region WIMS Water Quality Archive System, NRA User Manual Version 2.0, 01 March 1993.

HMIP (1993a) Sampling Facility Requirements for the Monitoring of Particulates in Gaseous Releases to Atmosphere. Technical Guidance Note (Monitoring) M1, (EPA 1990), January 1993, HMSO, London.

HMIP (1993b) Monitoring Emissions of Pollutants at Source. Technical Guidance Note (Monitoring) M2, (EPA 1990), November 1993, HMSO, London.

HMIP (1995a) Standards for IPC Monitoring: Part 1 - Standards organisations and the Measurement Infrastructure. Technical Guidance Note (Monitoring) M3, (EPA 1990), 1995, HMSO, London.

HMIP (1995b) Standards for IPC Monitoring: Part 2 - Standards in support of IPC monitoring. Technical Guidance Note (Monitoring) M4, (EPA 1990), August 1995, HMSO, London.

HMIP (1995c) Routine measurement of gamma ray air kerma rate in the environment. Technical Guidance Note (Monitoring) M5, (RSA 1993), September 1995, HMSO, London.

Annex A: Environment Agency Contacts and Respondents

Contact Name	Fn/Role	Contacted	Status
Chris Chubb	(PPC)	Yes	Interview/Questionnaire
Peter Lloyd	Control Systems	No	
Jeff Dolby	Area WQ Mgr	No	
Sue Stocks	WQ	Yes	Further Contacts
Shiela Sowerby	Planning/Modelling	Yes	Further Contacts
John Tyson	Control Systems	No	
Richard Freestone	Modelling	Yes	Questionnaire
Richard Streeter	Water Resources	No	
Martin Bigg	IPC	No	
James Hunt	Waste	No	
Dave Wardle	Data Services Mgr.	Yes	Questionnaire
Stuart Wright	CRI Mgr	No	
Clive Williams	RAS	No	
Richard Coward	Head CIS	Yes	Questionnaire
Alistair Gordon	Nat. Data Policy	Yes	Interview/Questionnaire
Paul Arrigoni	IS/IT Policy	Yes	Interview/Questionnaire
Ruth Wolstenolme	(SEPA)	Yes	Questionnaire
Roy Ramsay	(Env. Her. Serv.)	No	
John Dalton	R&D	Yes	Further Contacts
John West	WSA	Yes	Further Contacts
Ted Theirs	WSA	No	
Brian Sparks	WSA	No	
Vic Whiteley	IPC Manager	Yes	Questionnaire
Keith Harsham	IPC Manager	Yes	Further Contacts
Stuart Newstead	Reg. monitoring	No	
Mark Mardell	"	No	
Tim Reader	Auto. Monit. Sys.	No	
Tony Warne	WQ	Yes	Further Contacts
Paul Culverhouse	WQ	No	Questionnaire
John Braughton	?	Yes	Questionnaire
Chris Moore	WQ	Yes	Interview/Questionnaire
Paul Hughes	RAS	Yes	Interview/Questionnaire
Gillian Hill	PS - Tech.Info.	Yes	Questionnaire
Rob Gemmil	NCCA	Yes	Interview/Questionnaire
Andrew Dixon	?	Yes	Questionnaire
Paul Barraclough	CIS	Yes	Questionnaire
Kathy Greenhall	?	No	Questionnaire
Simon Powell	?	Yes	Interview/Questionnaire
George Marshall	IPC	Yes	Questionnaire
Barbara Evans	CIS	No	Questionnaire

David Treader	CIS	No	Interview/Questionnaire
David Glen	CIS	No	Interview/Questionnaire
Aileen Kirmond	Water Resources	Yes	Interview/Questionaire
Freda McDonald	?	Yes	Interview/Questionnaire
Charlotte Henderson	IPC ?	No	Questionnaire
Paul Moorhouse	IPC	Yes	Questionnaire
Neal Smith	Waste	No	Questionnaire
Neal Smith (2)	WQ	Yes	Questionnaire

Annex B: The Environment Agency - Enforced Acts

This list is taken from 'Enforcement Practise General Guidance' Ref :OP/OP/009 V1 05/96.

WATER MANAGEMENT

WATER RESOURCES

Water Resources Act 1991

Water Industry Act 1991

FLOOD DEFENCE

Water Resources Act 1991

Land Drainage Act 1991

Land Drainage Act 1976

[Flood Defence Bylaws]

FISHERIES

Diseases of Fish Act 1937

Sea Fisheries Regulation Act 1996

Salmon & Freshwater Fisheries Act 1975

Wildlife & Countryside Act 1981

Diseases of Fish Act 1986

Salmon Act 1986

Water Resources Act 1991

NAVIGATION

Water Act 1989

Water Resources Act 1991

Land Drainage Act 1976

Sea Fish Industry Act 1951

Pilotage Act 1987

Harbour Docks & Piers Clauses Act 1847

Anglian Water Act 1977

Upper Medway Navigation & Conservancy Act 11 & 14

Southern Water Authority Act 1982

Thames Conservancy Acts; 1932, 1950,1959,1966,1972

RECREATION

Water Resources Act 1991

CONSERVATION

Water Resources Act 1991

POLLUTION REGULATION

DISCHARGES TO WATER

Water Act 1989

Water Resources Act 1991

Water Industry Act 1991

D.2 How is it confirmed that received data meet the required specification (for example include the correct parameters, in the correct format at the required frequency and are free from blunders in copying or transmission)?

D.3 Who is responsible for this data management?

D.4 Are copies (i.e. duplicates) of the data placed in a paper filing system (e.g. an additional working copy)? If so, please expand.

Yes	No

D.5 Are the data loaded into a computer system for storage or analysis?
If so is it:

Yes	No

- a 'PC-compatible' computer?
- another type of desktop computer or workstation?
- a multi-user computer accessed through a terminal or PC?
- is the computer connected to a network in your office (LAN)?
- is the computer connected via a network to other offices (WAN)?

Annex C: List of Respondents

Business Needs Questionnaire

Name	Job Title	Function	Location
P Pearce	Team Leader Env. Prot.	Env. Protection	Thames West
C Evans	Monitoring Officer	Operational Monitoring	NW South Area
M Williamson	Regional Poll. Officer	IPC	NW
J Daniels	Team Leader WQ Consents	Env. Planning	Anglian- Eastern
N Ingrey	Team Leader Scientific Invest.	Env. Protection	Thames NE Area
Dr G Fulcher	Team Leader Tactical Planning	Env. Protection	Midlands
S Newstead	Head of National Centre	National Centre	Lancaster
J Broughton	WQ Scientist	?	NE
A Brewster	Scientific Support Officer	Env. Planning	Anglian Northern
N Smith	Team Leader Monitoring/Tech	Env. Protection	NE Ridings
R Stevens	Collaborative completion		

Data Collection Questionnaire

Name	Job Title	Function	Location
P Pearce	Team Leader Env. Prot.	Env. Protection	Thames West
S Coble	Team Leader Scientific Invest	Env. Protection	Thames West
N Ingrey	Team Leader Scientific Invest.	Env. Protection	Thames NE Area
Dr G Fulcher	Team Leader Tactical Planning	Env. Protection	Midlands
C Evans	Monitoring Officer	Operational Monitoring	NW South Area
J Daniels	Team Leader WQ Consents	Env. Planning	Anglian- Eastern
J Broughton	WQ Scientist	?	NE
A Brewster	Scientific Support Officer	Env. Planning	Anglian Northern
S Newstead	Head of National Centre	National Centre	Lancaster
N Smith	Team Leader Monitoring/Tech	Env. Protection	NE Ridings
K Harsham	Env. Planning Manager	Env. Planning	Thames West
P Proctor	Team Leader Data Services	Field Data Services	NE Dales
S Jackson	IPC/RAS Team Leader	IPC	NE Dales
J Goddard	Team Leader Scientific Invest.	Env. Protection	Thames
I Adamson	Technical Officer	?	NE Dales
J Dolby	Area Env. Protection Manager	Env. Protection	Midlands Trent
C Greenhall	Principal Data Resources	Env. Protection	NW
R Stevens	Collaborative completion		

Annex D1: Business Needs Questionnaire

R & D Project to Determine Data Transmission Requirements for Operator Self Monitoring

Business Needs QUESTIONNAIRE

QuantiSci Ltd, with Aspinwall & Company Ltd, have been commissioned to undertake this project by the Environment Agency. A range of information about the collection and use of self monitoring data related to releases or ambient conditions is required to assess the Agency's needs and develop options for data transmission. A synthesis of this information will be included in our report to aid in the assessment of the options identified and to provide a background for those involved in further development and implementation of data transmission strategies. Completed questionnaires, *excluding section G*, will be included as an annex to the report.

This questionnaire covers a wide range of aspects of data transmission and it will be an exception for any one respondent to be in a position to answer all of it. If in any section of the questionnaire you can provide a source of more information or a fuller answer (perhaps more technical, or more authoritative) please attempt to extract this information rather than offer names of individuals or functions within the Agency.

If your answers cover a wide range of situations, you may find it useful to make copies of this form, or append additional information.

Please remember that these questions relate *only* to operator self-monitoring.

A. Respondent

A.1 Name: _____

Address: _____

_____

_____

_____

Telephone: _____

A.2 Job Title: _____

A.3 Environment Agency Region/Area/Function

_____

Please return to: Dr David Waters, QuantiSci Ltd, Chiltern House, 45 Station Road, Henley-on-Thames, Oxon, RG9 1AT

A.4 **In what way are you involved with self-monitored data?**

- specification of policy
- specification of need for self-monitored data
- reception of self-monitored data
- checking data for compliance with self-monitoring requirement
- checking data against prescribed limits
- summarising for internal reporting
- data management or archiving
- management of officers carrying out one or more of the above
- data distribution (including external distribution)
- other, please specify

	Yes	No
A.3 Are you in a position to sketch out the flow of monitoring data with which you are involved within or between Agency offices and functions? If so, please append.		

B. Legislation requiring self-monitoring

B.1 Please mark against the attached list (Annex A) of 'enforced legislation' that legislation where:

- You or your part of the Agency are actively involved in self-monitoring
- You are aware that self-monitoring is or can be a requirement
- There is a requirement for self-monitoring data to be placed on a public register

	Yes	No
B.2 Can you direct us to any more detailed tabulations of statutory instruments, orders, regulations etc. related to any of this legislation that would aid in describing the scope of self-monitoring ? If so please list them - or append copies if they are to hand.		

	Yes	No
B.3 Are you aware of any proposed regulatory changes that would increase the scope of self-monitoring if they came into effect? If so, what are they?		

C. The primary interface with the Operator of the regulated activity

C.1 In relation to what activities do you or your office have direct contact with the operator?

C.2 How is the self-monitoring requirement specified to the operator (e.g. extract from regulation, standard letter etc.)? Please include any procedures prepared but not yet in use.

C.3 Please list (and attach copies of) any documents that are good examples of the way in which the self-monitoring requirement is specified.

	Yes	No
Can you provide further examples if requested?		

C.4 How do you reach agreement over, specify, or approve the technical adequacy (instrumentation, accuracy, precision, validation) of the self-monitoring procedure?

C.5. Can you provide references to, or copies of, standard specifications, or documentation related to specific agreements, that illustrate how these technical details are approved.
Is so please attach and list here:

Yes	No

C.6 What happens when there is a change in the nature of data received (e.g. due to an equipment upgrade) from the operator.

C.7 Are data provided:

- mainly in hard-copy form
- mainly in electronic form
- neither predominantly hard-copy or electronic

C.8 Are any data or alerts received directly from operators sites by remote telemetry. If so, please describe briefly.

Yes	No

D. Filing, data management and validation

D.1 How are the data in their 'as received' form stored?

D.2 How is it confirmed that received data meet the required specification (for example include the correct parameters, in the correct format at the required frequency and are free from blunders in copying or transmission)?

D.3 Who is responsible for this data management?

D.4 Are copies (i.e. duplicates) of the data placed in a paper filing system (e.g. an additional working copy)? If so, please expand.

Yes	No

D.5 Are the data loaded into a computer system for storage or analysis?
If so is it:

Yes	No

- a 'PC-compatible' computer?
- another type of desktop computer or workstation?
- a multi-user computer accessed through a terminal or PC?
- is the computer connected to a network in your office (LAN)?
- is the computer connected via a network to other offices (WAN)?

D.6 If a computer system is used is the program used to store the data:

- a general purpose database program (e.g. Microsoft Access, Oracle)
- a spreadsheet (e.g. Microsoft Excel)
- a program specifically designed for monitoring data, is so what programs are used

---

D.7 How is the accuracy of the data entry process confirmed:

D.8 Are these data recorded in the Agency Catalogue of Data Sources? If so, what are they called?

Yes	No	Unsure

D.9 Are any data placed on a public register in your office? If so which register?

Yes	No

	Yes	No
D.10 Are copies (not summaries) made for forwarding to, or received from, other offices or sections of the Agency? If so are they:		
• paper copies		
• computer files on disk		
• files over a computer network		
• files via a modem or network link		

D.11 Where is the most authoritative 'master copy' stored?

	Yes	No
D.12 Are computer systems used at your office for the management of data collected by or on behalf of the Agency? If so:		
• are they also used for self-monitored data		
• could they be used for self-monitored data		

E. Audit, analysis and reporting of monitoring data

E.1 How are the data compared to prescribed limits, trigger levels etc:

- by manual inspection and assessment
- with the aid of a computer program. If so what program?

	Yes	No
E.2 Is this carried by you or your staff? If not where is it carried out and by whom?		

E.3 Are statistical summaries of the data made for further use within the Agency or any other organisation? If so who makes them and to whom are they reported.

Yes	No

E.4 Are statistical summaries of the data received in your office from elsewhere in the Agency? If so, who makes them?

Yes	No

E.5 Are any data loaded into regional or national Agency databases or repositories? If so:

Yes	No

Which databases/repositories?

How are the data transferred?

Is the whole data set loaded, or part of it?

E.6 Are the data copied to or reported to one of the Agency's national centres? If so which ?

Yes	No	Don't know

F. Information systems procedures and strategy

F.1 Is the management of the self-monitored data subject to procedures or protocols specified in any Agency manuals, guidance notes or practice notes?
If so please specify.

Yes	No

F.2 Are you aware of any current Agency information systems developments (e.g. standardisation of software or hardware, introduction of new network infrastructure) that would:

- require changes in the way that any monitoring data are managed
- provide opportunities for enhancement of the management of self-monitored data

If so, what are they:

Yes	No

F.3 Are the computer hardware and software systems used for managing the data the standard systems specified by CIS. If not, please briefly describe them.

F.4 Are you aware of any Information Systems strategy decisions (e.g. GIS strategy, convergence, more open access) which would in the future affect the way in which self monitored data are managed.
If so please list.

Yes	No

F.5 Are you aware of any new initiatives under the national data policy that would affect the way in which self-monitor data are managed?
If so please list.

Yes	No

G. A Personal Overview

G.1 Please summarise from your point of view the pros and cons of the existing data management.

G.2 What changes can you suggest that would assist the management of self monitoring data.

G.3 What, in your opinion, are the operators views of the present procedures.

ANNEX A: THE ENVIRONMENT AGENCY - ENFORCED ACTS

This list is taken from 'Enforcement Practise General Guidance' Ref :OP/OP/009 V1 05/96. Please indicate:

- Those Acts that you know can require self-monitoring
- Those Acts under which you or your group handled self-monitored data

Any tabulations of statutory instruments, order and the like associated with these Acts would be of assistance to the project.

WATER MANAGEMENT

WATER RESOURCES

Water Resources Act 1991
Water Industry Act 1991

FLOOD DEFENCE

Water Resources Act 1991
Land Drainage Act 1991
Land Drainage Act 1976
[Flood Defence Bylaws]

FISHERIES

Diseases of Fish Act 1937
Sea Fisheries Regulation Act 1996
Salmon & Freshwater Fisheries Act 1975
Wildlife & Countryside Act 1981
Diseases of Fish Act 1986
Salmon Act 1986
Water Resources Act 1991

NAVIGATION

Water Act 1989
Water Resources Act 1991
Land Drainage Act 1976
Sea Fish Industry Act 1951
Pilotage Act 1987
Harbour Docks & Piers Clauses Act 1847
Anglian Water Act 1977
Upper Medway Navigation & Conservancy Act 11 & 14
Southern Water Authority Act 1982
Thames Conservancy Acts; 1932, 1950,1959,1966,1972

RECREATION

Water Resources Act 1991

CONSERVATION

Water Resources Act 1991

POLLUTION REGULATION

DISCHARGES TO WATER

Water Act 1989
Water Resources Act 1991
Water Industry Act 1991
Salmon & Freshwater Fisheries Act 1975
Environmental Protection Act 1990

WASTE REGULATION

Control of Pollution Act 1974
Control of Pollution (Amendment) Act 1989
Environmental Protection Act 1990

INTEGRATED POLLUTION CONTROL

Environmental Protection Act 1990

AIR POLLUTION

Health & Safety at Work etc Act 1974
Alkali etc Works Regulation Act 1906
Environmental Protection Act 1990

RADIOACTIVE SUBSTANCES

Radioactive Substances Act 1993

GENERAL

The Environment Act 1995
European Communities Act 1972

Annex D2: Data Collection Questionnaire

R & D Project to Determine Data Transmission Requirements for Operator Self Monitoring

Data Collection Questionnaire

QuantiSci Ltd, with Aspinwall & Company Ltd, have been commissioned to undertake this project by the Environment Agency. A range of information about the collection and use of self monitoring data related to releases or ambient conditions is required.

This questionnaire is an aid to gathering information about the data that are collected, how parameters are measured and recorded, the equipment and software used, and current methods of transmission and data management. This information will be used to assess the Agency's needs and develop options for future data transmission.

The information required is so diverse that this form can act only as a 'prompt' when seeking details. Do not be constrained by the layout, please append any additional documents or notes that clarify or replace answers to any of the questions. You may find it useful and simpler to complete several copies of the form relating to different data-sets. You may prefer to complete the questionnaire as a team effort (if so, please attempt to complete sections rather than enter names of colleagues). Completed questionnaires *excluding section G*, will be included as an annex to the report.

A. Respondent

A.1 Name: _____

Address: _____

_____

_____

_____

_____

_____

Telephone: _____

A.2 Job Title: _____

A.3 Organisation _____

A.4 If Environment Agency what Region/Area/Function _____

A5. Role in relation to self monitored data _____

Please return to: Dr David Waters, QuantiSci Ltd, Chiltern House, 45 Station Road, Henley-on-Thames, Oxon, RG9 1AT

B. The data sets that you handle

B.1 What activities are you involved in that result in a requirement to carry out the monitoring?

B.2. Please list the data-sets with which you are involved. Group them in whatever way is relevant to the way in which they are handled (e.g. legislative requirement, equipment or laboratory used, database or filing system in which they are stored).

C. Field Sampling and measurement

	Yes	No
C.1 Do any monitoring data come from in-situ, fixed equipment? If so:		
• are the measurement made automatically (without the presence of an operator)?		
• are measurements transmitted directly to an office for recording?		
• are measurements logged for periodic transmission via a communications link?		
• are measurements logged for later download by a visiting operator?		
• are measurements made only when an operator visits?		
• does the equipment take samples for later collection and analysis?		
• are measurements made in any other way - please specify?		

C.2 Please list, or give examples of, the equipment used.

C.3 Are any data collected using portable equipment? If so:

	Yes	No
• are specified monitoring points visited repeatedly?		
• are measurements made other than at repeatedly visited points?		
• are measurements stored in a logging device?		
• are measurements or observations logged manually?		

C.4 Please list, or name examples of, measurement and logging equipment used for self-monitored data.

C.5 How is the equipment calibrated?

Does it conform to any standards?

C.6 Are samples collected for laboratory analysis? If so

Yes	No

- are analyses carried out by the operator of the activity?
- are analyses carried out by a another laboratory ? If so, which laboratories?

Yes	No

- does the laboratory make results available in digital form?
- does the laboratory have NAMAS accreditation?

D. The Data

D.1 What parameters are required? (add separate sheet if necessary)

Please list or append additional notes or documentation illustrate the requirement, (e.g. example records that examples of pro-formas).

D.2 Are the limits, trigger levels etc that the measurements ar compared with based on any national or international standard? If so, which?

Yes	No

D.3 How are sampling points identified geographically? (eg National Grid co-ordinates, latitude and longitude, distance from reference point)

D.4 What quantity of data is sent? Use whatever method of quantification and time period that is most appropriate.

D.5 Do the Environment Agency receive individual measurements?

Yes	No

D.6 Do the Agency receive summaries?

Yes	No

D.7 Are the procedures used for self-monitored data also used for data that is not transmitted to the Agency. That is, is the Agency requirement met as part of monitoring procedures with a wider scope? If so, please briefly describe this:

Yes	No

D.8 Does the monitoring procedure record other parameters or measurements additional to those required by the Environment Agency?

Are other parameters sent to the Agency, in addition to those required? If so please give a brief description:

Yes	No	Don't know

D.9 What are the guidelines or conventions for reporting missing data, values below detection limits etc

E. Data processing , storage and transmission to the Agency

E.1 In your offices, are the data stored:

- mainly in paper files?
- mainly in computer/electronic files?
- using a combination of paper and electronic records?

Yes	No

E.2 Are the data transmitted to the Environment Agency:

	Yes	No
• mainly in paper files?		
• mainly in computer files on disk?		
• mainly in computer files via a modem or network link, or via the internet?		
• using a combination of paper and electronic records?		
• is there any trend towards electronic transfer?		

E.3 Are any data transmitted directly to the Agency from remote monitoring equipment (e.g. as a condition of a discharge consent) ? Who is your contact if things go wrong.

Yes	No

E.4 Are there any guidelines for the format in which data are transferred?
What are they:

Yes	No

E.5 Where data are processed for collation, review, presentation or statistical analysis how is this carried out? Please indicate any computer programs used. Is any processing carried out by third-parties (e.g. laboratories or contracting organisations).

F. Future data transmission methods

F.1 What likely changes in equipment (for measurement or logging), or computer facilities (hardware, software, network links) can you anticipate that would affect the way in which data could be transmitted to the Environment Agency.

F.2 Please give your suggestions as to changes in transmission methods that would make data transfer more efficient or effective (e.g. easier, quicker, cheaper).

F.3 What are the most relevant technical developments in the industry sector with which you are most involved in relation to monitoring?

G. A Personal Overview

G.1 Please summarise from your point of view the pros and cons of the existing data management.

G.2 What changes can you suggest that would assist the management of self monitoring data.

G.3 What, in your opinion, are the operator's/Agency officer's (ie "the other side") views of the present procedures.

Annex E: Urban Waste Water Treatment Regulations - an example of data transmission requirements

Examples of many of the issues discussed in this report can be illustrated by some of the monitoring requirements of the EC Urban Waste Water Treatment Directive (91/27/EEC) as implemented by the Urban Waste Water Treatment (England and Wales) Regulations 1994.

The Regulations primarily concern monitoring by the Agency (i.e. the 'Authority' referred to in the Regulations). However, we understand that similar specifications form part of conditions placed on operators to carry out monitoring in consents under the Water Resources Act 1991.

Regulations 10 (Samples and Records) and 11 (Monitoring), and the Schedules to the Regulations are included here. The following notes comment on these extracts in the order in which they occur, highlighting aspects relevant to the issues considered here.

An important observation relating to the Regulations as a whole is that in general they state a performance requirement rather than a detailed specification. For this reason much of the material required to perform a 'data and systems analysis' for data transmission will be found in the actual consents under the Water Act, rather than in the Regulations.

Regulation 10 - Samples and Records.

Paragraph 2 states that operators' measuring and recording equipment shall be presumed to register accurately unless the contrary is shown. Quality assurance data (e.g. calibration records) are important in tests of this presumption, and it is a 'business needs' issue at the time of granting a consent as to whether these data are to be transmitted to the Agency along with monitoring data.

Paragraph 3 implies archiving and/or transmission of records to the Agency and paragraph 4 concerns a regulatory aspect of missing data. Data transmission requirements thus depend on the 'business need' issue (legal in this case) of what form of records will be admissible in evidence.

Paragraph 5 appears to concern a single detail of the sampling procedure, demonstrating the importance of one parameter (sampling time). However, it illustrates that there is information about the sampling procedure (automatic collection in this case) that is relevant when monitoring data are assessed, and might not form part of each transmitted data set, but which both primary and secondary users must have access to.

Regulation 11 - Monitoring

Paragraph 1 places a duty on the Agency carry out or procure monitoring, referring to Schedule 3 (discussed below), and paragraph 3 places a duty on the Agency to retain this monitoring information. Although not concerning operator-monitored data the 'business needs', data transmission, data management and analysis requirements stemming from these duties will be in large part analogous to those for operator data.

Schedule 1

Parts I and II of this schedule set out criteria for identification of sensitive areas and high natural dispersion areas. Although these designations will not be part of the data sets transmitted by an operator they do affect the monitoring requirements and the standards to be achieved. Thus the data-management systems into which the operator data are imported must be able to link the data to these designations, probably through some geographic identifier or parameter. This is of particular importance to secondary users.

Schedule 3

Part I of this schedule states that treatment plants should be engineered so that it is possible to take 'representative samples' from treatment plants. It is probable that some record of how this is achieved will need to be available for interpretation of monitoring data. This Part includes the tables giving standards to be achieved. General points relevant to data transmission formats and data management systems are that additional parameters must be monitored in discharges into sensitive areas, and that other Community Directives may impose other requirements.

Table 1 requires monitoring of Biological Oxygen Demand (BOD5) and Chemical Oxygen Demand (COD) against reference concentrations in the discharge or reductions in concentration between influent and effluent. A footnote adds a requirement regarding total suspended solids for discharges from lagoons. Table 2 requires monitoring of Total Phosphorous and/or Total Nitrogen with the same alternatives as to reference concentrations or reductions in concentration. Thus a data-set carrying the determinations of the parameters covered by these tables may require up to 9 fields.

A footnote to Table 1 inserts an additional need for flexibility in the data handling systems by allowing for substitution of Total Organic Carbon or Total Oxygen Demand for BOD5 'if a relationship can be established'. Thus the data handling system must cater for the addition of one of two additional parameters - and the information detailing the relationship to BOD5.

Thus the data management system will require a record for each plant indicating which parameters should be present, and for the requirements of Table 2, the 'population equivalent (p.e.) of the plant.

Part II of the Schedule is concerned mainly with evaluation of results, but does introduce further complications into the data to be transmitted and managed. Paragraph 2(a) allows for flow-proportional or time-based sampling, and paragraph 2(b) requires "good international laboratory practice", both adding data related to the method (i.e. quality data) to the data transmission requirement. Paragraph 5, in its reference to "unusual situations such as those due to heavy rain" highlights the fact that a wide range of other data may be required in order to evaluate the results and will be required either within the data-set transmitted or by reference to other records.

Conclusions

This brief review of some of the monitoring requirements stemming from the Urban Waste Water Treatment Directive illustrates a number of issues relevant to operator self-monitored data:

- The details of the monitoring requirement that is placed on operators are to be found not in these Regulations but in the results of the Agency's 'business processes' in applying them along with other regulations. In this case via discharge consents under the Water Act.
- Where requirements are performance based, or specific but make allowance for alternative methods that can be shown to be equivalent, flexibility will be required in formats for data transmission.
- For each plant there will be a collection of data that remain constant over long periods (designations of receiving water, population equivalent, engineering-related factors that affect sampling etc.) which need to be accessed alongside monitoring data during evaluation.

An overall conclusion is that details of the data transmission requirement are dependant as much on the Agency's business processes as on the actual parameters lists and sampling methods indicated by the Regulations.

We are aware of a number of other EC Directives or Regulations (e.g. the Landfill Directive, IPPC Directive, Groundwater Regulations) which require monitoring by operators and include a requirement to report to Brussels, in some cases within 6 or 12 months of implementation. These place an increasing burden of reporting on the Agency.

As has been discovered during this project, and illustrated in this section on the Urban Water Waste Treatment Regulations, data transmission requirements do not stem directly from the specifications in a statutory instrument. Rather they result from the Agency's implementation of performance specifications in a number of instruments as part of its 'business function' and depend to a great extent on the data-management regimes set up to handle the great range of environmental data. Whilst assessment of the data management (including transmission) procedures required to fulfil each EC Directive is essential it is likely that effective overall procedures will require that the needs of related Directives are considered together.