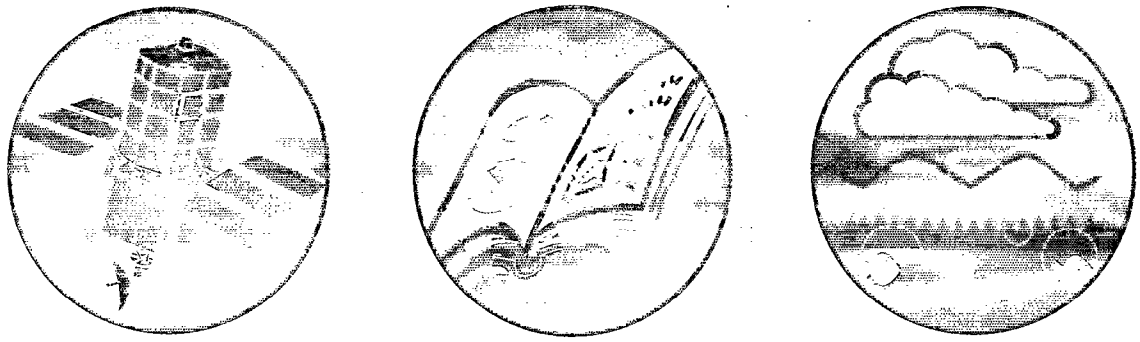
Fisheries Assessment Software within the Environment Agency



Research and Development

Technical Report
W176

ENVIRONMENT AGENCY

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

Fisheries Assessment Software within the Environment Agency

Technical Report W176

Steve Barnard and Robin Wyatt

Research Contractor:
WRc plc

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: 1 85 705043 6

© Environment Agency 1999

All rights reserved. No part of this document may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the Environment Agency.

The views expressed in this document are not necessarily those of the Environment Agency. Its officers, servant or agents accept no liability whatsoever for any loss or damage arising from the interpretation or use of the information, or reliance upon views contained herein.

Dissemination status
Internal: Released to regions
External: Released to public domain

Statement of use
Implementation of the recommendations contained within this report will, in the first instance, be through uptake by a Fisheries function group which is currently developing the fisheries assessment software from Midlands region as a best interim solution to satisfy corporate I.S. requirements.

This document will be useful in the longer term in helping to guide further development of this software or, if appropriate, production of a robust and practical alternative.

Completed questionnaire returns are documented in R&D Project Record W2/i719/1 and R&D Project Record W2/i719/2.

Research contractor
This document was produced under R&D Project i719 by:

WRc plc
Henley Road
Medmenham
Marlow
Buckinghamshire
SL7 2HD

Tel: 01491 571531 Fax: 01491 579094

WRc Report No.: EA 4581

Environment Agency's Project Manager
The Environment Agency's Project Manager for R&D Project i719 was:
Irven Forbes - Environment Agency, Anglian Region

CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

A sound knowledge of the status of fish populations is a pre-requisite to informed fisheries and environmental management. Within the Environment Agency, fish population surveys are afforded a high priority, with a large number of sites being surveyed each year. The data obtained from these surveys are many, complex and varied, and require substantial statistical analysis, presentation and referral before being used in management decisions. There is a consequent need for appropriate software to be available for use by staff throughout the Fisheries Function. The current project was commissioned in order that the software requirements of those tasks associated with fisheries stock assessment be identified, and that the extent to which existing software fulfils these requirements be assessed.

Information on software (current for the period 1996/7) was collected by means of a detailed questionnaire, forwarded to each of eight regional contacts. Subsequently, regional meetings were held with each of the eight contacts to enable both the reporting requirements and the perceived software requirements of each region to be more closely assessed.

Through discussion with the regional contacts it appears that, as a general principle, the range of analysis tools that are used or required by any given region are effectively dictated by the methods that are employed to collect the fishery data. In turn, the nature of data that are collected can be affected by: local conditions; the extent of the proposed survey programme; staff availability; and material resources.

The information that was provided by the regions was used to assess the functionality and capability of currently available software. In addition, a range of idealised functional requirements was drawn up. It was apparent that most (if not all) of the software currently available to staff of the Agency's Fisheries Function failed to meet the functional requirements that were identified. In addition, it was found that available software was, in many cases, outmoded and made use of inappropriate methods.

At present, there is no agreed national standard for stock assessment analysis methods or for reporting requirements. In conjunction with the poor quality of the stock assessment software that is currently available within the Agency, this represents a potential limiting factor to the provision of informed and effective fisheries management.

The availability of fisheries software from sources outside of the Agency was examined. However, no suitable software packages were identified from within key research bodies within the UK or from the American Fisheries Society in the United States.

Consequently, it is recommended that a new national fisheries stock assessment package is developed, to fully service the needs of the Agency in a nationally consistent manner. The development of new software, under a national R&D initiative, represents the best option for achieving the objective of providing software that fulfils the requirements identified during the course of this project, and for ensuring that such software is subsequently available throughout the Agency.

It is proposed that new software is developed such that a standard survey protocol is not imposed upon the regions, but rather that the use of raw data which may have been generated

by any one of a variety of means is facilitated. However, the proposed software would nevertheless allow the analysis and reporting of fisheries data to be undertaken in a nationally consistent and standardised manner. The development of new software should be undertaken on a modular basis, with the potential for the development of individual modules to be prioritised and phased.

The production of a new software is broken down to a series of interrelated work items. Indicative costs and comments on the scheduling of each of the range of work items are provided. It is recommended that the development of new software should not be an isolated proposition. There are several potential links between the proposed development of new fisheries software and current or proposed R&D initiatives and these are briefly discussed.

KEY WORDS

Fisheries; stock assessment; software; population estimation.

1. INTRODUCTION

1.1 Objectives

1.1.1 Overall project objective

The Fisheries Function of the Environment Agency makes use of computer software in undertaking several of its duties. The current project was commissioned in order that the software requirements of those tasks associated with fisheries stock assessment be identified, and that the extent to which existing software fulfils these requirements be assessed.

The overall objective of the project was:

- to evaluate current Environment Agency fish assessment computer software and comparable commercially available external software, in order to determine the best option for implementation and integration into routine fishery survey management.

1.1.2 Specific objectives

The specific objectives of the project were as given below:

(i) to determine the requirements of both the Environment Agency's Fisheries (particularly end-users) and national I.S. functions, and to specify the criteria against which to assess software, with respect to success in satisfying these Functions' business needs across the Environment Agency;

(ii) to evaluate the software existing in the Environment Agency, as well as that commercially available externally, for the planning of fishery surveys, determination of fisheries statistics and management information output, e.g. length-weight relationships, growth and age analyses, yield assessment, stock management and other data analysis, archiving (and retrieval) and presentation needs of the Fisheries Function across the Environment Agency;

(iii) to advise the Environment Agency on the current usage, functionality and performance of its fish stock assessment and other software used in fisheries management;

(iv) to rank the fisheries software by the criteria established in (i);

(v) to assess the adaptability of the existing software and specify how the findings from R&D Note 292, plus other final outputs from R&D Project 325 ('Fish stock assessment methodology') and the end-user requirements may be incorporated, where appropriate, into the existing software;

(vi) to assess the value of developing new fisheries computer software, which would include the best principles from the existing software and incorporate the findings from R&D Note 292, plus the other outputs from R&D Project 325, and the end-user requirements;

(vii) to present the Environment Agency with options for fisheries software, including costs and benefits for up to four options;

(viii) to produce an outline functional specification for the Environment Agency's preferred option, as agreed with the Project Board; this option shall cover the Fisheries and IS functions' requirements for fisheries software;

(ix) to advise the Environment Agency on interim measures to ensure that fish stock assessment is supported by software of a standard acceptable to the Environment Agency;

(x) to report, accurately and clearly, the findings and recommendations of the Project and produce final reports which present the business case for the preferred option and meet the requirements of the Environment Agency's Fisheries and I.S. functions.

1.2 Method of working

The project objectives were addressed through the completion of a work programme that was, in turn, split into four basic elements:

- the determination of the requirements of both the Fisheries and I.S. functions
- the assessment of existing software
- the assessment of (development) options
- the production of recommendations.

The first of these elements is addressed in Section 3 of this report. The assessment of existing software, both within and outside of the Agency, is reported in Section 4, whilst development options are discussed in Section 5. Recommendations for the development of fisheries assessment software within the Agency are provided in Section 6.

2. METHODS

2.1 Software within the Agency

2.1.1 Collection of information

Information on existing software was collected by means of a detailed questionnaire (an example of which is reproduced as Appendix A). These questionnaires were forwarded to each of eight regional contacts, a list of who are given as Appendix B. The questionnaires were designed so as to facilitate the collation of information on a range of subject areas, covering:

1. the requirements and use of stock assessment software;
    - requirements of, and applications for, software.
    - necessary software characteristics or features of software
    - general specifications of software that is currently employed within the region (i.e. what is used, and for what general purpose); and
2. technical specifications of software that has been produced, or modified, within a given Agency region (i.e. providing more detailed information on individual pieces of software).

For both elements of the questionnaire, distinction was made between four (functional) categories of software - namely:

- software developed for the assessment of data pertaining to individuals (e.g. length-weight and length-at-age relationships)
- software developed for the assessment of data at the site, or population, level (e.g. population estimates, biomass estimates)
- software developed for the assessment of data at the river reach level (e.g. the derivation of river fishery performance classes) and
- software developed to assist with the planning and design of fishery surveys.

Subsequently, regional meetings were held with each of the eight contacts to enable both the reporting requirements and the perceived software requirements of each region to be more closely assessed. In addition to the regions' reporting and processing requirements, the techniques used for the collection of fisheries data were also discussed. Where appropriate, regional contacts ensured that other staff from the Agency's Fisheries Function (e.g. the area fishery officers) were present at these meetings.

2.1.2 Subsequent reporting

Following discussion with the Project Leader, it was felt that a formal ranking of the available software (Specific Objective (iv), Section 1.1.2) was inappropriate. Discussion as to the relative merits or benefits of alternative systems was largely confined to their functionality

and, subsequently, increased attention was focused on the specification of recommended development options.

During the course of the project it became apparent that the range of software in use within the Agency's regions was very limited. Furthermore, there had been relatively few recent developments in software and, consequently, little software was available for use in the 'Windows' operating environment. In addition, the information provided in response to the 'Technical Specification' section of the set of questionnaires disseminated to regional contacts was very limited in its detail.

2.2 Software use outside of the Agency: collection of information

Short questionnaires (see Appendix C) were sent out to a group of commercial organisations and academic institutes to assess the extent and nature of fisheries software currently in use outside of the Environment Agency. It was intended that this would help indicate those stock assessment software products that are potentially available to, but not currently used by, the Agency.

A list of those organisations/institutes contacted is given as Appendix D. The information derived from these questionnaire returns are discussed in Section 5.

In addition, the American Fisheries Society's Computer User Group was contacted for information regarding commercially available software within the United States.

3. SOFTWARE REQUIREMENTS OF THE FISHERIES AND I.S. FUNCTIONS

3.1 Operational practices within the regions

3.1.1 Introduction

Through discussion with the regional contacts it appears that, as a general principle, the range of analysis tools that are required by any given region are effectively dictated by their operational practices (i.e. the methods that are initially employed to collect the fishery data). In turn, the nature of data that are collected can be affected not only by local conditions (e.g. the physical nature of the water bodies that are being sampled) but also, by the extent of the proposed survey programme and the availability of staff and material resources. Software development should be mindful of the survey methodologies employed by the regions and should provide adequate facility for the analysis of the range of data that are routinely produced. It is appropriate, therefore, to briefly consider the range of survey methodologies that are applied by the regions.

The following section (Section 3.1.2) outlines the range of survey methodologies that are currently in use within the regions. In line with the specific objectives for the project (which direct effort to software relating to the determination of fisheries statistics from fishery surveys, and to the planning and presentation needs of the Agency) and with early discussions with the Project Leader, the bulk of this report deals with software for use with data from those fishery surveys that are undertaken by 'traditional' survey methods such as electric fishing and netting. It effectively excludes software relating to the processing of information that is derived from other sources (e.g. fish counters, hydroacoustic surveys, licence returns, angling match result analysis, etc.). Sections 3.2.1 *et seq.* (which document the perceived software requirements of the regions, the nature of input data that are currently derived from fishery surveys, and the current reporting requirements of the regions) are similarly restricted to software for use in the analysis of data from site-specific (i.e. electric fishing and netting) surveys. Finally, Section 3.3 outlines the Agency's requirements in terms of the software's technical specification, and is based on discussions with the Agency's Corporate Information Systems (C.I.S.) Section.

3.1.2 Survey methodologies

Whilst it is not surprising to discover that survey methodologies vary considerably between regions, there are examples of a range of varying survey techniques being employed even within the same region. In Anglian region, for example, the Eastern area makes extensive use of electric fishing (either by wading or using a boat) whilst sites in the Fenland drains are predominantly surveyed by netting.

Similarly, within the Midlands region, multiple catch depletion fishing with stop nets may be used on smaller watercourses, whilst on larger rivers a qualitative (or, at best, semi-

quantitative) approach (e.g. employing a single-run with a boom-boat) may be the most appropriate technique. On canals, the use of boom-boats is generally restricted to semi-quantitative surveys where fish are not actually caught, but all observations are recorded, with all fish that are seen during the course of a survey being assigned to a species and size-class. Seine netting may be employed, but is qualitative and often use simply for presence/absence information or for checks on disease status or parasite loadings.

In the North-East, salmonid nursery and juvenile habitats are usually sampled by single run electric fishing; extensive quantitative surveys are not feasible due to resource limitations. In addition, there is now a general move towards single run semi-quantitative surveying (without stop nets). It should be stressed, however, that the data collected by these methods is nevertheless meaningful, and plays an important role in guiding management decisions. The physical nature of those rivers that support important coarse fisheries tends, generally, to render them unfishable by quantitative methods. Consequently, coarse fishery survey work has not, historically, made use of quantitative stock assessment methodologies. For the few stillwaters that are present in the region, sampling tends to be reactive and relies mainly on the collection of presence/absence data.

Fishery surveys in the North-West are largely semi-quantitative (the North and South areas are exclusively semi-quantitative, whilst only about 10% of surveys in the Central area are fully quantitative). The region's philosophy is that surveys are undertaken to identify 'problem areas' and to identify the 'general health' of fisheries in a given catchment, and not to produce detailed, site-specific, information. Fully quantitative survey methodologies tend to be reserved for reactive or strategic surveys (e.g. impact assessment). Whilst netting techniques may be used on stillwaters, the resultant data are often not analysed other than for cursory information (e.g. species presence/absence). In many cases such surveys are looked upon more as public relation exercises than as part of monitoring programme.

In Southern region, only about a third of the surveys are fully quantitative. The remaining two thirds are semi-quantitative 'strip' surveys, which are based on a single run. As for the single-run fishings employed for canal surveys in North-West region, fish that are picked up by the electric fishing gear are not removed from the water but instead are assigned to a species and to one of four size classes. Increasing use is being made of this methodology (especially on narrow channels) as it effectively allows the length of river being surveyed to be increased, and so decreases the influence that fish mobility can have on results. As with many of the other regions, very few netting exercises are undertaken. When netting exercises are carried out it is often on a 'look-see' basis - with no population estimates being derived. The data that are collected may be used for growth analysis, but there is little information available for historic comparison.

Whilst widespread use is made of quantitative electric fishing methods in South-West region, certain areas (e.g. Cornwall area) undertake a high proportion of semi-quantitative (single run) surveys. In addition, netting methodologies are also occasionally used.

In general, fisheries assessment in Thames region is concerned with river sites. For these, both standard electric fishing methods (multiple run, depletion fishing) and semi-quantitative methods (employing boom-boat electric fishing and hydro-acoustic surveying) are used (approximately 75% and 25% of surveys respectively). By way of contrast, whilst netting

methods (employing mark-recapture techniques) may be used lakes in the region, such exercises represent a small proportion (probably less than 1%) of overall fisheries survey effort.

In Welsh region, about 90% of electric fishing sites are semi-quantitative (e.g. 25-50m single run - with or without stop-nets) or, where the river is very wide, 5 minute fishing of a 'riffle' sites. In general, fisheries staff would aim to undertake one fully-quantitative site within a set of surveys (e.g. 1 site from 5 done on any given day). Other sources of data that are used for fisheries assessment in Welsh region include: log books, automatic fish counters, commercial netting returns, routine fish trapping data and radio tracking studies. Currently there are no definitive packages for analysing such data. However, in the short term, all that would be required is a means of summarising and visualising the data. Statistical analyses are generally applied on an *ad hoc* basis and there is a general perceived need to be able to apply 'standard' statistics to such data.

3.1.3 Current regional reporting

Discussions with the regional contacts provided some information on the regional policies on reporting that have been adopted. It was apparent that the Agency's views on reporting requirements vary considerably between the regions - a summary of main points is provided below. No distinction is made between those requirements that are the result of accepted regional policy and those that are the result of pragmatic, resource driven policy decisions.

Within Anglian region, reported fishery statistics are limited to coloured maps that show the lengths of river which fall within each of four biomass classes. These classes are based on the total fish biomass present, and represent the four ranges 0-5g.m^{-2}; 5-10g.m^{-2}; 10-20g.m^{-2}; and >20g.m^{-2}. It is felt that these satisfy the region's limited reporting requirements, which mainly involve replying to enquiries (e.g. from an angling club) regarding the state of a given fishery. Detailed site reports are not generally produced; they are held to be too labour intensive to produce and are not used enough to justify the required effort.

Notwithstanding the belief that fish are able move relatively freely and rapidly within sections of river between (apparently) impounding structures, and that fish are also able to move upstream over weirs, etc., the boundaries that are used for the colour coding of the river classification maps nevertheless tend to relate to fixed structures (such as weirs, STWs, etc.).

For Midlands region's operational reporting there is the requirement for population estimates (as biomass for cyprinids and as density for salmonids) and for growth analyses and assessments of species richness. For population estimates the region currently requires that estimates be derived for individual species. However, the region also want to be able to use the total number of fish per run (i.e. all species combined) to calculate an overall population estimate. Subsequently, this overall estimate is used, together with information on the species composition of the total catch, to derive population estimates at the species level.

Within North-East region the primary reporting requirement is for suitable inputs to the National Fisheries Classification Scheme (i.e. biomass estimates for rheophilic and limnophilic coarse fish and density estimates for salmonid species). In addition to this

reporting requirement, surveys are also used to provide information on the population age- and size-composition of individual species (in order to permit the effects of certain environmental changes, such as droughts, to be assessed, and to enable staff to comment on the balanced development and sustainability of fish stocks). Comparisons of catches between surveys (in terms of species caught, and in both size- and age-compositions) are routinely made. Surveys are also held to be a basic requirement for the implementation of effective and informed fisheries management policies, and to identify derogations under the EC Freshwater Fisheries Directive.

Fish population estimates at the species level are not routinely made within North-East region, due mainly to the low efficiencies of capture that prevail in the large rivers of the region (and, incidentally, to the cost associated with improving the precision of population estimates compared to the benefits that may be accrued in obtaining a large-scale overview by fishing more sites). With regard to these alternative approaches, data requirements for coarse fish surveys are currently being re-assessed on the basis of results of an intensive survey programme undertaken in 1996 to address this point.

The reporting of operational surveys in North-West region is limited to catchment overviews. Colour coded markers are used to indicate site locations and fishery performance (in terms of overall density) on river network maps. Five density classes (plus an 'absent' class) are used. No additional detail is reported for the operational surveys.

In both Southern and South-West regions, the level of detail that is reported is generally inconsistent. Outputs are not produced to pre-defined formats that are common across the regions. However, within Southern region it has been suggested that a three-tier structure for fishery reports should be adopted. Such a system would have reports produced at the following levels:

- level one: a basic report on stillwater or riverine surveys, comprising one or two sheets of A4 (plus, where appropriate, associated graphs and tables for release to angling clubs, riparian owners, etc.);
- level two: a more detailed, site-specific survey report (for riverine surveys only) which would include a detailed analysis of all of the results for that site;
- level three: a catchment-level report, summarising the results of all individual surveys within a survey programme.

Within Thames region, a range of site specific data are presented in reports of the operational fishery monitoring. These data include biomass and density estimates (both at the species level) together with species composition information and an indication of the size frequency structure of the population (again at the species level). In addition, some habitat data are presented to assist in data interpretation. As with Anglian region, coloured maps summarising fish biomass are produced within Thames region.

Operational surveys within Welsh region are largely confined to juvenile salmonids. Fishery performance is reported using both densities and an abundance classification scheme. This scheme, which was originally derived within Welsh region and is currently used for the reporting of the Regional Juvenile Salmonid Monitoring Programme (RJSMP), will be replaced from 1998 onwards by the use of the National Fisheries Classification Scheme.

Results from the RJSMP are reported annually as catchment summaries with tabulated densities and with site locations and abundance classes marked on maps.

3.1.4 National reporting - the National Fisheries Classification Scheme

Whilst the use of the National Fisheries Classification Scheme (NFCS) and associated software was generally supported by the regions, its actual operation was felt to be hindered by the lack of compatibility between the NFCS software and the software that is used to derive population estimates from survey data. Whilst software used to derive population estimates invariably produces species-specific estimates, the inputs that are required by the Classification software are mainly combined population estimates for groups of species.

Because of this incompatibility, the use of the Classification software inevitably requires fishery statistics to be reprocessed, with species groupings either being combined (manually or by means of a third software package, such as a spreadsheet system). Consequently, whilst it was felt that the Classification system was a useful reporting tool, its use would be greatly enhanced if software were available that produced population estimates for the species groupings directly.

3.2 Software requirements

3.2.1 Regional requirements for data analysis

Introduction

The regions' analysis requirements are not only related to the survey methods that are employed, but also to their reporting requirements. This is likely to be the reason why the range of analysis requirements that were put forward by the regions seemed disparate. However, in many cases, the perceived need for software appeared often to be related more to the particular software that was currently in use than to the actual requirements of the staff.

The software (analysis) requirements that were identified, either through the completion of the questionnaires or in the subsequent regional meetings, are reproduced below (see Table 3.1 to Table 3.5).

Additional requirements

In addition to software for fisheries data analysis, the regions also voiced the need to consider the provision of software to fulfill other aspects of their work. In particular, mention was made of software for the production of graphics/report outputs, and the benefits that may potentially be derived from the adoption of a national standard GIS (geographic information system).

Although it would be useful for software to have a standardised graphical output format for the production of final graphics, Thames region suggested that routines to produce 'interim'

graphics, both as aids to data interpretation and to facilitate data checking, should be incorporated into new software. For example, length-frequency plots could be shown after the entry of length data to provide an initial indication of the structure of the population as well as highlighting the presence of outliers. Such interim graphics should be capable of being modified interactively (e.g. by altering the bin size on frequency plots) so that the clarity of final graphics is effectively optimised. Whilst supporting the notion of enabling the production of interim graphics, South-West region suggested that the production of 'final' graphics should be hired out to dedicated graphics software (which could link more effectively with Agency standard report producing packages). In this sense, the fisheries assessment software would not produce graphics itself, but would produce data files for import to other reporting packages.

Several regions discussed the possibilities of linking reporting with outputs from a GIS. A common view expressed was that a GIS derived location map (e.g. river network) with site locations marked (and colour coded according to some performance metric) would be a very useful reporting output. In addition, it was felt that many useful variables could be derived from a GIS (reach gradient, catchment areas, etc.).

3.2.2 Software functionality

The information that was collated on regional operational practices (see above, Section 3.1) was used to derive a generalised overview of the range of fishery operations that stock assessment software should be able to address. This overview, together with information on the range of analyses that need to be supported (see above, Section 3.2.1) was used to derive the functionality of required software, as presented below.

Data input

Whatever software package is to be used for data analysis, its use will necessarily involve the input of field data. Within the Agency both of the basic alternative practices regarding the input of the data obtained from surveys into fisheries assessment software are employed, namely:

- manual input to software from hard-copy;
- direct electronic input to software from data logger.

Manual input of data, from hard-copy recorded in the field, is the method currently used by most regions, including: Anglian; Midlands; North-West; South-West; Welsh; and most of the North-East and Southern.

Data loggers have been used extensively by Thames region. Currently, data loggers are being phased out in Thames region and are being replaced with 'ruggedised' laptop Pcs. Loggers are used to a lesser extent by staff in the Northumbrian area of North-East region and in Southern region (see Table 3.6, below). Although they do not currently make use of loggers, both South-West and Welsh regions would want the option to be input data from field loggers within any newly developed fishery stock assessment software.

Table 3.1 Software requirements for information relating to individual fish

Requirement	Region							
	Anglian	Midlands	NE	NW	Southern	SW	Thames	Welsh
Estimation of length-weight relationships	✓	✓	✓	✓	✓	✓	✓	✓
Comparison of length-weight relationships	✓	✓						
Interrogation or application of length-weight relationships	✓		✓		✓		✓	
Length-frequency histograms as an output			✓	✓	✓	✓	✓	✓
Identification of cohorts by analysis of length-frequency data	✓	✓		✓		✓	✓	✓
Assessment of condition factor		✓	✓	✓			✓	✓
Estimation of mean length-at-age (MLA)	✓			✓		✓	✓	✓
Identification of cohorts by analysis of MLA & length-frequency data	✓			✓				
Age-length frequency output as stacked histogram	✓		✓		✓		✓	
Scale- and body-proportional back-calculation of fish length to produce historical MLA information for estimating (individual) growth rates	✓		✓	✓	✓	✓		✓
Fitting growth models to length-at-age data	✓	✓	✓		✓			✓

Table 3.2 Software requirements for site specific information

Requirement	Region							
	Anglian	Midlands	NE	NW	Southern	SW	Thames	Welsh
Population estimation by fully quantitative removal methods, including: Seber & LeCren; Moran-Zippin; Carle & Strub	✓	✓	✓	✓	✓	✓	✓	✓
Population estimation from semi-quantitative data		✓	✓	✓	✓	✓	✓	✓
Population estimation by mark recapture methods	✓	✓				✓	✓	✓
Use of minimum estimate where other methods are not possible	✓		✓					
Estimation of catch weight from derived length-weight relationships	✓		✓					
Assistance with interpretation of habitat effects on observed populations				✓				
Assessment of year-class strengths		✓						
Estimation of survival from year-class abundance	✓							✓
Species composition pie-charts as an output						✓		

Table 3.3 Software requirements for information at the reach level

Requirement	Region							
	Anglian	Midlands	NE	NW	Southern	SW	Thames	Welsh
Estimation of mean density or biomass at the reach level (with appropriate input to Fishery Classification)	✓	✓	✓		✓	✓	✓	✓
Comparison of reach means (spatial or temporal) to assess statistical significance of observed differences	✓							
Assessment of observed reach performance with some expected performance measure based on habitat measure - seen as distinct from Fishery Classification					✓			

Table 3.4 Software requirements for survey design or planning

Requirement	Region							
	Anglian	Midlands	NE	NW	Southern	SW	Thames	Welsh
Indication of number of sites required to reach a given level of precision - with assessment of current situation	✓	✓	✓	✓	✓		✓	✓
Indication of the periodicity required to optimise a rolling programme - with assessment of current situation	✓							
Identification of number of sites required to detect a given % change in the population compared to control (either temporal or spatial control)								✓

Table 3.5 Other software requirements

Requirement	Region						
	Anglian	Midlands	NE	NW	Southern	SW	Thames	Welsh
Interpretation of hydroacoustic data		✓	✓		✓			
Analysis of angler-catch (match) data or creel census information		✓	✓	✓			✓	
Estimation of salmonid run-size - general						✓		✓

Table 3.6 Hardware used for field-based data logging

Region	Loggers used
Thames	Huskey Hunters or Huskey Hunter 'ruggedised' laptop PCs
North-East	Huskey Hunters
Southern	Toshiba laptops & Psion organisers

Interpretation of length and weight data - operational practices

The currently available software makes use of a range of types of data, a reflection of the diversity of operational practices. Simple programs, such as those that produce population estimates from data on the number of fish caught on each of a series of removals, may only require two or three numbers to be input (e.g. the numbers of fish caught on each of two runs, plus an estimate of the site area). More complex programs (such as those that produce length-frequency histograms, or can be used to assess growth rates) will require additional information. Inevitably this is provided in the form of length and/or weight data for either the full population that was caught or for a subsample of individuals.

Within the Agency, length data are currently recorded either by reading off the fork lengths of individual fish from a measuring board, or by marking fork lengths on a waxed sheet with a pointed seeker. Data are input to software (or, in the case of Thames region, the field data logger) manually. Whilst there is the possibility of using a modified digitising tablet to automate this step (either in the field, using the live fish; or in the office, using pricked wax sheets) no regions currently make routine use of such facilities.

In most regions, lengths of all the fish (of major species) that are caught are recorded (see Table 3.7). However, the software use by Anglian region allows for a sub-sample of lengths to be entered, the length-frequency distribution of which is then used to generate estimated lengths for those fish whose lengths were not recorded. In the North-West and North-East regions, less use is made of length data that are gathered on an operational basis, and it is therefore not always recorded. In these regions, further use of length data may be made when it is collected as part of non-operational (i.e. strategic or reactive) survey work.

Weight data are not currently recorded by all regions (again, see Table 3.7). In particular, areas where fisheries are salmonid dominated (e.g. Welsh region) make little use (if any) of weight data.

Table 3.7 Recording of length and weight data

Region	Length measurements	Weight recording (see notes)
Anglian	Usually all fish measured	Bulk weighing [1]
Midlands	Usually all fish measured	No weighing [2]
North-East	Usually all fish measured	Occasional bulk weighing [3]
North-West	Some lengths recorded	Some weights recorded [4]
Southern	Usually all fish measured	Subset weighed [5]
South-West	Usually all fish measured	Subset may be weighed [6]
Thames	Usually all fish measured	Subset weighed [7]
Welsh	Usually all fish measured	No weighing [8]

1 no individual weights are routinely recorded;

2 weights estimated by application of standard condition factors;

3 weights estimated by application of standard length-weight relationships;

4 occasional weight recording -but would prefer to use standard length-weight relationships to derive weights from lengths;

5, 6 & 7 site-specific relationship derived from length-weight pairs and applied to recorded lengths;

8 little (if any) use made of weight data - therefore not generally recorded.

Interpretation of length and weight data - software requirements

Ideally, software needs to be able to accept length data for fish and produce length-frequency histograms as an output. There should be facilities to identify cohorts; most effectively by allowing for the 'manual' on-screen division of length-frequency histograms into data sets relating to distinct cohorts (if the actual age-length data for individual fish are known then this may be used to assist in this operation). Whilst routines are available that go part way to automating this process, the assumptions that they entail are legion. Overall, the benefits that may accrue from the adoption of such methodologies are likely to be insignificant.

Given weight information for individuals that have length data, software should be able to allow the calculation of length-weight relationships (assuming the underlying relationship $W=aL^b$; where 'a' and 'b' are constants). Under these conditions, 'a' and 'b' are estimated from the linear regression of log-transformed data. Derived length-weight relationships can be used to predict the weights of fish of known length that are taken from the same population but which were not weighed in the field. It should be possible for software to use estimates of the parameters that have been derived from site-specific survey data or, alternatively, to use archived 'standard' parameter estimates. The need to produce weight estimates for un-weighed fish is important as it facilitates the subsequent estimation of population biomass.

The ability to subsequently compare parameter estimates with regional 'standards' or with the (archived) parameter estimates from other sites was highlighted as a useful and desirable software feature. However, formal comparison of the regression parameters 'a' and 'b'

between different populations should be avoided, as variation in either or both parameters can suggest significant differences between populations and, unless both parameters are considered simultaneously, valid interpretation of the results is difficult (Bolger and Connolly, 1989).

Software should also allow the estimation of condition factors (e.g. Fulton, 1911) for all fish that have both length and weight data. Although the use of Fulton's condition factor (where the condition factor is estimated as the ratio of fish weight to the cube of the length) assumes isometric growth (which is often not the case) many fish do have length weight relationships with regression coefficients very close to three. Moreover, the exponent value of three can be considered simply as a method of transforming the linear dimensions of length to the cubic dimensions appropriate for the discussion of weight (again, see Bolger and Connolly, 1989).

Interpretation of age data - operational practices

The other main type of fish data that are derived from routine surveys relates to the ages of those individuals that are caught. With current Agency policy being for ageing to be carried out centrally (at the Brampton laboratory in Anglian region) there is an inevitable lag between the production of raw survey data and corresponding age data. Generally, software that relies on the input of age information must therefore be able to archive 'incomplete' data records that can be completed at a later date, once the age data have become available. A principal exception to this is where software is applied to data from salmonid populations where the main distinction between different age classes can be between fry and parr. Such differentiations can often be made by eye, so removing the need for the otherwise inevitable delay in the process of age data production.

Where age data are used by software, it is usually input by hand.

Interpretation of age data - software requirements

Software should allow for age information obtained from 'direct' ageing (e.g. analysis of scale or otolith samples) to be used to confirm or correct the cohort identification that is initially undertaken with the length-frequency data (see above). Subsequently, options should be available to produce stacked length-frequency histograms that effectively distinguish between the different cohorts present in a population (for example, see Figure 3.1, below). If presented on-screen, such a figure could be used interactively to help divide up the observed size-frequency distribution into successive age-classes. Inferred ages for those fish that were not aged directly (e.g. by reading scale or otolith samples) could subsequently be written back out to the database that holds the raw survey data.

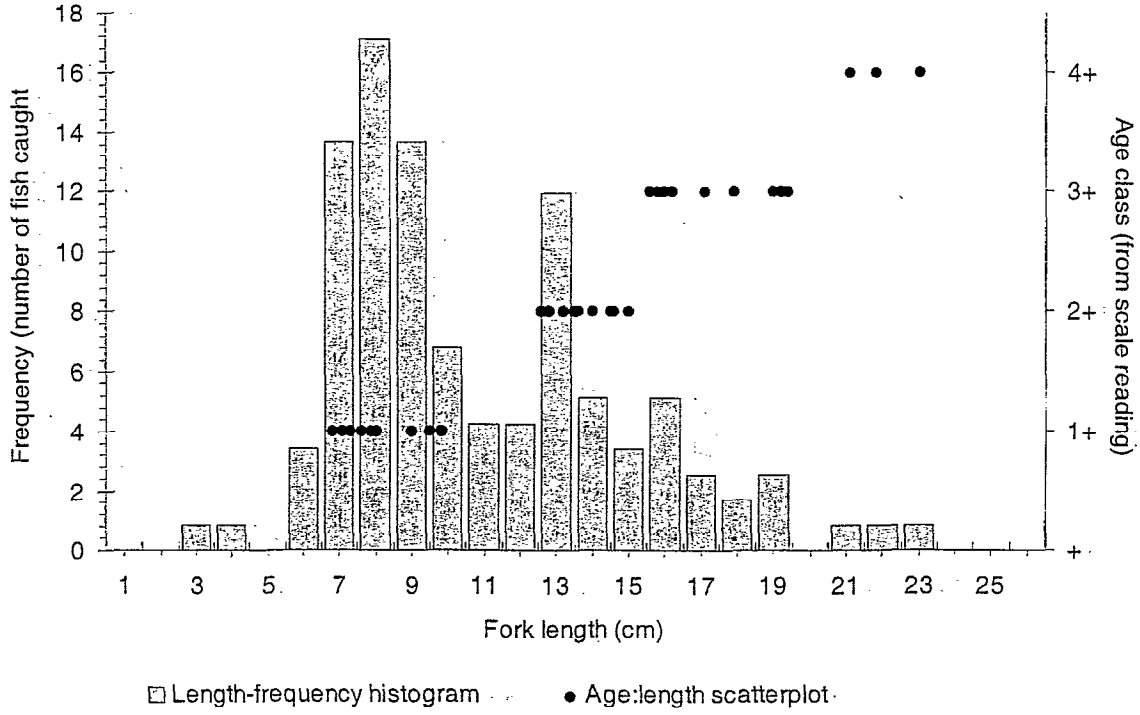


Figure 3.1 Example of length-frequency plot overlaid with length-at-age scatterplot

If age data are available (either from the direct ageing of scale samples, or from a combination of direct ageing and inference from length-frequency distribution information) then software should allow the calculation of mean length-at-age (MLA) for each species that is represented.

Options should be available to allow growth models to be fitted to:

- length-at-age data for the entire sample population (as obtained at the time of the survey) - effectively combining data from several cohorts and to
- back-calculated lengths - i.e. allowing the growth rates of individual cohorts to be identified.

Software requirements for population estimation - fully quantitative removal methods

For data from surveys based on the removal method, software should allow fully quantitative estimates to be made of the population size, at the species level (and for identified sub-populations such as specific age-classes). For the purposes of deriving population estimates from removal data, the NRA Interim Report 325/5/A (Lacey *et al.*, 1992) recommended that the Carle and Strub MWL method be adopted for routine use within the NRA. Furthermore, the same report suggested that the 'asymptotic' formulae for variance estimates were unreliable and should be replaced by variances estimated by simulation. Should it be decided that it should still be possible for the user to choose between alternative methodologies for population estimation, then succinct and unambiguous advice should be provided to enable the user to make the necessary choice of methodology in an informed and confident manner.

Software requirements for population estimation - semi-quantitative (CPUE) methods

Software should be capable of handling data that have been derived by semi-quantitative fixed-effort) sampling methods. The use of semi-quantitative methods (in terms of the statistical aspects of the subsequent analysis of the data) is currently receiving attention under the national R&D programme (R&D Project 7716). Consequently, it is not appropriate at this point to formally produce a functional specification for this aspect of stock assessment software. Instead, this aspect of software functionality should be left open and should ultimately be directed by the findings of this ongoing R&D.

Software requirements for population estimation - mark recapture methods

Although the requirement for the assessment of data derived from mark-recapture methods is limited, software should have the option to deal with such data. A range of statistical methodologies are used in connection with the analysis of such data (e.g. Petersen, Schnabel, Bailey, etc.). Whilst there are certain circumstances where one may technique might be used in preference to the others (e.g. in relation to the number of recapture exercises undertaken) there is currently no formal guidance as to the optimum analysis methodology to undertake under any particular conditions. As with the use of semi-quantitative methodologies, it is suggested that recommendations as to the required functionality of national stock assessment software should not be produced at this stage, but should follow from an in-depth appraisal of the suitability and performance of the different available estimator functions.

Software requirements for population estimation - minimum estimates

Where it is not possible to apply either of the foregoing classes of analysis to field data (e.g. where the assumptions underlying the removal methodology are known to have been violated) then software should allow the production of minimum estimates.

Additional population parameters

In addition to deriving estimates of population size, software should subsequently provide further information on the overall fishery structure such as:

- species composition by number (this can be represented graphically, e.g. by pie charts)
- estimation of (species specific) biomass (from population size estimates, known length frequency distributions and known length-weight relationships) and
- species composition by biomass (as above, this can be represented graphically).

Software for whole fishery assessments

Data from individual (site specific) surveys can be combined and assessed either spatially or temporally. Spatial combination of data will produce statistics that relate to greater areas of river than individual data from individual sites, and are therefore more suitable for reporting at a relatively coarse scale. However, the process of combination will necessarily entail a number

of assumptions regarding, for example, the degree to which the individual sites are representative of the whole river reach.

Spatial (reach) assessments

Software should facilitate the combination of site-specific data from a series of sites to estimate the mean density or biomass at the reach level. Whilst such combinations should be undertaken at the species level, reach estimates should also be provided for species groupings so as to provide an appropriate input to the National Fishery Classification Scheme.

Software should either be able to undertake a (spatial) comparison of reach means to allow an assessment to be made of the statistical significance of any observed differences, or should produce the information necessary to calculate the appropriate test statistic(s) manually.

Temporal assessments

The requirements for temporal assessments of fisheries data discussed at regional meetings covered two principal topics:

- the calculation of year-class strengths and
- the estimation of year-on-year survival rates from abundance data for a given year-class over successive years.

The calculation of relative year-class strength requires extensive information, especially when a large number of surveys are being considered. It is defined by Mills and Mann (1985) as the percentage of the population composed of a given year-class of fish, divided by the mean percentage of the population composed of fish of the same age-group (as calculated over the period in question). For example, a given year-class of roach that makes up 25% of the total roach population in a given year, with roach of the same age-group making up, on average, only 20% of the population over a pre-defined period, would have a relative year-class strength of 1.25 for that particular year.

Whilst it is not likely that software would be required to report relative year-class strengths on a regular, routine basis, where software is capable of providing the information that is necessary for a manual calculation, it would not be a far more complex task to facilitate the automated calculation of year-class strengths.

Similarly, the information that is required for the estimation of year-on-year survival rates could also be produced by any newly developed software. However, the assumptions that would, necessarily, be inherent in any automated routines that would need to be developed for the estimation of year-on-year survival rates are likely to be legion.

Habitat information

It has been suggested that software should provide a means by which observed reach performance can be compared with some objective expected performance measure based on a

combination of habitat measurements. That is, software should (ideally) facilitate the interpretation of estimated population size in relation to the prevailing habitat. This was seen as a distinct, and more detailed, requirement than that which is currently supported by the National Fishery Classification Scheme. Again, attention is drawn to current R&D on habitat inventories (Wyatt and Barnard, 1997). The development of such options would necessarily follow the reporting of completed R&D in this field.

Planning

Software should be available that facilitates the planning of a survey programme, in particular allowing:

- the estimation of number of sites required to reach a given level of precision
- the assessment of the effect of periodicity on the suitability of a rolling programme's periodicity
- the estimation of the number of sites that are required to detect a given % change in the population compared to temporal or spatial control.

Any such applications would require information on spatial and temporal population variability. This information should be available through the (automated) interrogation of the main fishery data archive, although it would be prudent to also facilitate the interactive (i.e. keyboard) entry of such data.

Other applications

The specification of additional applications, for example covering the requirements currently addressed by the software outlined in Section 4.2.2, is not covered by this report. However, within the range of requirements for fisheries stock-assessment there remains the necessity for software to be compatible with other commonly-used fisheries software products. In particular, the use of hydro-acoustic data; angler-catch (match) data/creel census information; or salmonid counter/trapping data will continue as important sources of population information for fishery managers within the Agency. However, the nature of the methodologies used are such that it is not reasonable to expect that they would fall within the same software framework as those routines that are employed for stock assessment and reporting using data that are derived by removal or mark-recapture methods.

What should be addressed, however, is the commonality of outputs. It is appropriate to direct the development of software for alternative stock assessment methods such that the software-user interfaces that are presented are broadly similar to those for other stock-assessment software. It is important to direct resources to unifying and simplifying the approaches taken to fisheries stock assessment. The proliferation of PCs presents a challenge to the Agency in that it is all too easy to produce a range of independent software products, but far harder to ensure that different products share a common 'look and feel' and that archive information in a similar manner (such that the subsequent interrogation of stored data can be accomplished without the need for recourse to a range of different analysis packages).

Habitat data

Currently, habitat data are recorded in an inconsistent manner. There is, however, a range of habitat information recorded, ranging from simple site dimensions through to more detailed habitat assessments (including, for example, instream substrate, instream and riparian vegetation, land-use, channel gradient, etc.). Little formal use is currently made of these data; its prime use is a (subjective) aid to data interpretation.

3.2.3 Underlying considerations

General

Software development embodies a range of decision making processes, which can be split to at least two distinct areas. Those which relate to the selection of appropriate routines for data handling (input and output) and storage, and those which relate to the selection of algorithms or models that are to be applied to the raw data during the implementation of the software's analysis or interrogation routines. In effect, this can be thought of as: 'how do we handle the data?' and 'how do we do the calculations?'.

The choice of inappropriate data handling routines (or, more fundamentally, the selection of an inappropriate software platform) may tend to make the software less efficient or less user-friendly. For example, a spreadsheet package may be being used for stock assessment software which relies heavily on data archiving and retrieval, and where the use of a database package would be more appropriate. However, the choice of inappropriate methodologies for the analysis of raw data has more far-reaching consequences, in that it may result in software that is sub-optimal and (at worst) inaccurate. Consequently, whilst it is important to pay attention to the data handling elements of software design, it is of vital importance that the correct routines for the analysis of stock assessment data be applied.

R&D Note 292 and associated reports

National R&D on the development of stock assessment methodologies and methods was reported as R&D Note 292 ('Guidance notes on the design and analysis of river fishery surveys'; Wyatt and Lacey, 1994). The specific objectives for this work included:

- the review of current statistical methodologies for within-site population estimation and analysis of survey data
- the evaluation and development of methods for the analysis of survey data and
- the provision of recommendations on best practice for survey design and data analysis.

One of the main recommendations to come out of the associated research was that the Maximum Weighted Likelihood (MWL) method of Carle and Strub (Carle and Strub, 1978) be adopted for routine use within the Agency. Furthermore, it was found that the available formulae for estimating the variances of removal estimates were unreliable. It was recommended that the use of such formulae should be abandoned in favour of variances estimated by simulation.

R&D Note 292 also provides guidance on the design of river fishery surveys, including:

- the estimation of spatial, temporal, spatial-temporal interaction and error variances
- methods for the quantification of survey objectives and
- the design and analysis of each of a range of survey types (e.g. surveys for estimating abundance, estimating spatial and temporal changes in population, and for undertaking impact assessment).

The outputs from this work are obviously central to both the assessment of currently available software and to any future software developments.

3.3 C.I.S. Function's requirements regarding software development

3.3.1 General

This section addresses several key questions regarding the future development environment and the C.I.S. Function's requirements for software production.

In particular, current proposals regarding the Agency's standard computing environment (in terms of standard hardware and software specifications) and future development requirements are outlined.

3.3.2 The standard computing environment

A phased migration from a 16- to 32-bit platform, together with the installation of Windows 95, is currently underway within the Agency. This process is scheduled for completion by mid-1999. Concomitant with this migration is the installation of the standard Microsoft Office package (including Microsoft Word, Excel and Access) onto Agency PC's.

Any new software developments would benefit from being able to interface with these standard packages. In particular, when dealing with fisheries data, it is likely that input from, and output to, a standard (Excel) spreadsheet format would provide a robust interface. The use of standard packages in this way would help increase the degree of user-friendliness of any new applications, and facilitate a rapid and straightforward uptake by the end user.

3.3.3 Development languages / platforms

At present, for new applications that are developed in-house, the Agency is considering a shift from the Delphi programming environment to Visual Basic. Whilst existing Delphi developments tend to utilise the InterBase package for data archiving, it is likely that if future developments are undertaken using Visual Basic, they will employ either Access or Oracle for their data archiving requirements. The choice between these two, will in turn, be dictated by the size of the application. Where smaller applications will make use of Access, larger applications that have larger databases, or that may need to be scaled to the client-server environment, are more are likely to developed using Oracle.

It is unclear as to how the Agency's preferred choice of development environment will spill over to development work that is contracted out. It would be reasonable to assume that, for smaller applications that do not need to communicate with Oracle databases, the choice of development environment would be at the discretion of the developer. However, for larger applications, or systems that require extensive communication with Oracle, the Agency may stipulate that a specific development environment is used.

3.3.4 Software and data storage

Where it is appropriate, the Agency would encourage the use of regional database structures. However, it has been recognised that this ideal may be compromised by practical constraints such as the efficiency of network communication systems. Careful thought therefore needs to be applied to the geographic scale of software and data storage, i.e.:

- the 'standalone' choice of software permanently loaded on individual PC's and data being saved to local (hard) drives
- software and data being held on LANs (local area networks)
- software and data being held at the regional scale on WANs (wide area networks).

The choice between these various options is likely to be very case specific, and account will need to be taken of both the need for data to be available to other (remote) users and for software implementation and data access to be acceptably rapid. It is likely, however, that current Agency policy would favour the second of these three options.

3.3.5 The need for external contractors for software development

The ability of the Agency to produce large applications - either alone or in tandem with an external consultant responsible for development guidance - is currently unclear. However, even where development is undertaken in-house, the specific technical requirements of software for fishery data archiving and analysis are such that the input of an external consultant is likely to be necessary. Under such circumstances it could be argued that it would be beneficial to contract-out the majority of the development work, with Agency C.I.S. function retaining an advisory role to oversee matters of software and hardware compatibility.

Where software is produced outside of the Agency, the degree to which it will be supported by the Agency's C.I.S. following its delivery to the Agency will need to be agreed in advance. Such agreements should necessarily form part of the Terms of Reference for any such development project.

4. FISHERIES SOFTWARE AVAILABLE WITHIN THE AGENCY

4.1 Current use of stock assessment software within the Agency

4.1.1 Software use within the Agency.

The range of stock assessment software (for data derived from electric fishing or netting surveys) that was recorded as currently being in use within the Agency is indicated on the following page (Table 4.1) together with an indication of the extent of their usage.

An arbitrary differentiation has been made between identifiable software products (programs or other recognised, formalised software routines) and other *ad hoc* software. In general terms, the latter category includes those routines or macros that are used largely in isolation and which have not been formally documented.

Main identified software

FINS: the Fisheries INformation System. This is the analysis part of the FLAPS package (Fish Logging And Processing System), which itself is often referred to as 'FLAPS & FINS'. This package, developed within Thames region, is based on the FoxPro and dBase database packages and provides a virtually integrated package for data analysis, reporting and archiving.

FDPS: the Fishery Data Processing System. This was developed within Anglian region, and was written using the BASIC programming language. Like FINS, it too provides an integrated package for the analysis, reporting and archiving of fisheries data.

FSAS: the Fisheries Science Application System. A series of routines written in the BASIC programming language, based on the routines presented in Saila *et al.* (1988). They are principally used for growth analysis.

Smart: includes software systems produced within Midlands region using both SmartWare II and Smart 3.1 software. These are different versions of an integrated PC package, which includes spreadsheet and database functions. The functionality of the two versions is essentially the same. The software routines that have been developed form an integrated system, fulfilling most of the regions current analysis, reporting and archiving requirements.

Autosurvey: is a package developed within North-East region using the Lotus 1-2-3 spreadsheet package. It performs only a limited range of functions, including the production of length-frequency plots and age-length tables.

Carle & Strub: is a simple program for generating population estimates from removal data using the Carle and Strub (1988) MWL methodology. It was written using the BASIC programming language.

SurvForm: is a template for data analysis and archiving that has been developed by Southern region using the Lotus 1-2-3 spreadsheet package. It requires population estimates to be supplied from an external source but, once entered, will convert these to estimates of population density. It further serves as a data storage system for raw fishery and habitat data, and facilitates the production of certain pre-defined graphics (such as pie charts indicating biomass contribution by species).

Zippin: is a simple program written using the BASIC programming language that is used in conjunction with 'SurvForm' within Southern region . As its name suggests, it calculates population estimates from removal data using the Zippin methodology. Apart from this it has no additional functions.

Fish Population Estimation program: a simple program, written using the BASIC programming language (but operating under a VAX/VMS mainframe operating system) for producing population estimates from removal data. It was produced within South-West region, and applies the Carle and Strub approach to population estimation. It does not perform any further analysis functions.

Remove: is a simple program, written using the BASIC programming language, for calculating a population estimate from removal data. It was written by staff at IFE (Clarke, 1992). It does not support any further functions.

RS 1, routines within **SPSS**, and macros within **SuperCalc**: These are routines written by staff within Welsh region, using either RS1 (a statistics package), SPSS (Statistics Package for the Social Sciences - a statistics package) or SuperCalc (a spreadsheet package). No formal descriptions or documentation of these routines have been made available for this project. However, it is understood that, in combination, they are used by staff throughout the Welsh region for the production of population estimates from routine survey data.

Additional *ad hoc* software

In addition to the range of software listed above, Agency staff also identified occasions where they made *ad hoc* use of additional software (see entries in the final row of Table 4.1).

The examples that have been identified represent a wide range of types software, including: spreadsheet packages (Excel, Lotus 1-2-3, SuperCalc); statistics/data analysis packages (Unistat, Minitab); database (DataEase) and graphics/reporting (Freelance) packages.

Reference to such *ad hoc* software indicates where the software is used but, unlike (for example) the routines within RS1 and SuperCalc that are used in Welsh region, no formalised procedures are available. These entries therefore include 'one-off' pieces of software that may be produced by individuals and are intended for only limited use.

Table 4.1 Range of stock assessment software currently used within the Environment Agency

	Region where software is used							
Software title	Anglian	Midlands	North-East	North-West	Southern	South-West	Thames	Welsh
FINS			✓		✓	✓	✓	
FDPS	✓				✓			
FSAS	✓							
SmartWare II / Smart 3.1 ('Smart')		✓						
Autosurvey			✓					
Carle & Strub program ('C&S')			✓	✓				
'Zippin'			✓		✓			
SurvForm					✓			
Fish population estimation program ('FPE')						✓		
Remove						✓		
RS 1								✓
(Routines within) SPSS								✓
Macros within SuperCalc V ('SC5')								✓
Other *ad hoc* software (see notes in text)	Excel DataEase Lotus 1-2-3	Lotus 1-2-3 SuperCalc	Lotus 1-2-3 Freelance	DataEase Minitab	Unistat	Lotus 1-2-3 Freelance Unistat Minitab	Freelance	

4.1.2 Nature of software used

For ease of comparison, the range of programming languages or software platforms that are employed by the main stock assessment software currently used in the Agency are reproduced in the following table (Table 4.2).

Table 4.2 Programming languages or software platforms

Stock assessment software	Software platform / programming language
FINS	FoxPro / dBase
FDPS	BASIC
FSAS	BASIC
Smart	SmartWare II / Smart 3.1
Autosurvey	Lotus 1-2-3
'C&S'	BASIC
SurvForm	Lotus 1-2-3
'Zippin'	BASIC
'FPE'	VAX / VMS
Remove	BASIC
RS 1	RS 1
SPSS	SPSS
SC5	SuperCalc

4.2 Functionality of software used within the Agency

4.2.1 Stock assessment software

The following tables (Table 4.3 to Table 4.5) indicates the range of features, functions and methodologies employed by the main pieces of software that are currently used by the Agency.

Table 4.3 Functionality of software used within the Environment Agency for the analysis of fishery population structure

	Software								
Features or methodologies supported	FINS	FDPS	FSAS	'Smart'	Autosurvey	Survform	RS1	SPSS	'SC5'
Regression of length-weight data	✓	✓				✓		✓	
Prediction of weight from length	✓	✓							
Prediction of length from weight		✓							
Statistical comparison of two length-weight relationships		✓							
Visual assessment of length-frequency data	✓			✓	✓	✓			✓
Extraction of length-frequency data		✓							
Condition factor calculation (Fulton)	✓								
Visual assessment of age-frequency data	✓								
Length-at-age - back-calculation or modelling of non-linear growth			✓	✓			✓	✓	
Reporting of length-at-age data				✓	✓				
Weight-at-age - non-linear growth									
Estimation of survival rates		✓						✓	

Table 4.4 Functionality of software used within the Environment Agency for population estimation

	Software									
Methodologies supported	FINS	FDPS	'Smart'	Autosurvey	'C&S'	Survform	'Zippin'	'FPE'	Remove	'SC5'
Removal methods										
Carle & Strub - EML and/or MWL	✓				✓				✓	✓
Zippin	✓	✓	✓				✓	✓	✓	✓
Seber & LeCren	✓	✓	✓							
Mark-release-recapture methods										
Bailey	✓									
Lincoln		✓								
Petersen	✓									
Schnabel	✓									
Semi-quantitative methods										
Simple minimum estimate	✓	✓		✓						
Product of minimum estimate and estimated probability of capture	✓									
General										
Conversion of population estimates to density or biomass	✓	✓	✓		✓	✓			✓	
Storage of associated habitat data	✓		✓	✓						
Storage of associated water quality data			✓							
Species composition summary data				✓						

Table 4.5 Functionality of software used within the Environment Agency for the analysis of 'whole fishery' data

Features or methodologies supported	Software			
	FINS	FDPS	Autosurvey	SPSS
Mean population density (by species) for a selected series of sites (each sampled on only one date)	✓	✓		✓
Mean population density at a single site (by species) for a selected series of dates		✓		
Mean population biomass (by species) for a selected series of sites	✓	✓		
Mean population biomass (by species) for a selected series of dates		✓		
Summary of whole catch (e.g. average total biomass and relative contribution by different species) from range of sites			✓	

4.2.2 Ancillary software

In addition to the software that is used for fishery stock assessment and the derivation of related statistics, the Agency also use a number of other software packages for associated purposes. These include:

- packages to analyse or collate and report information on commercial catch returns, data from fish counters, and data from angling matches
- demonstration software for use as an aid to survey design
- software to implement the National Fishery Classification Scheme (NFCS)
- software implementing the HABSCORE habitat assessment and salmonid population models.

The software packages involved make use of a variety of software platforms (see Table 4.6).

Table 4.6 Ancillary software used in the regions

Software	Regions where used	Software platform
Commercial catch returns	North-West	Oracle
Fish counter database	North-West	Oracle
Match returns	North-East	Access
Survey design	North-West South-West Welsh	Excel
NFCS	Anglian Midlands North-East North-West South-West Welsh	Access (run-time application)
HABSCORE	North-East South-West Welsh	SuperCalc / Delphi †

†: originally developed to run under SuperCalc, the software has been recently re-developed as a stand-alone package using Delphi.

4.3 Assessment of the suitability and performance of current stock assessment software

4.3.1 General

In general terms, each region has access to software that fulfils its core requirements for day-to-day data analysis and reporting of fishery survey data. However, there are three principal areas of concern that have become apparent during the course of this project. These relate to:

- the mis-use or mis-application of software (e.g. the use of inappropriate or sub-optimal methods for obtaining population estimates)
- the ease of use of software (both in terms of the operation of the software and the interpretation of the outputs) and
- the degree to which information derived from different software systems can be combined (compatibility problems are even evident within regions; e.g. Midlands region where fisheries data are analysed on different systems - SmartWare II, Smart 3.1 and SuperCalc - in different areas).

Of these three areas, the first is the most important as it dictates the quality of information that is made available for subsequent reporting. Errors or inaccuracies in the statistics that are

derived from the raw survey data will be propagated through the subsequent reporting and decision-making process which represents the fisheries management activity.

However, both of the remaining areas need to be addressed. For example, the degree to which the software is user-friendly will inevitably dictate its uptake and use within the regions - with software that is complex or difficult to use, and which produces outputs that are hard to interpret, being used less widely and perhaps being viewed unfavourably by those staff that need to make use of it. Also, the compatibility of software obviously dictates both the potential for information interchange between areas and regions, and the potential for fisheries information produced by the analysis of stock assessment data to be used at a national level (e.g. for national R&D work).

4.3.2 Technical appraisal

The appropriateness of the stock-assessment software that is available within the Agency is discussed under four headings, reflecting the broad classes of applications that are covered by stock assessment software, *viz.*: population structure; (site specific) population estimation; whole fishery statistics; and planning.

Population structure

The assessment of population structure (in terms of length-frequency distributions, growth rates, condition factors, etc.) is undertaken by several of the software packages currently in use. Where such options are available, they are sufficient for reporting purposes.

Site specific population estimation

NRA R&D Project 325 made recommendations that, for the removal method of estimating population size, the Carle and Strub Maximum Weighted Likelihood (MWL) method should be adopted for routine use within the NRA. As shown in Table 4.4, of the eight software systems for producing population estimates from removal data that are commonly in use in the Agency, only four are able to implement the Carle and Strub MWL methodology. Of these, three are programs or routines that are dedicated to the task of providing a population estimate (the 'Carle and Strub' program used in North-West region, the 'Remove' program used in South-West region - which also allows for both the Exact Maximum Likelihood, EML, and the MWL methods - and routines written under SuperCalc - 'SC5' - in Welsh region). With the exception of calculating population densities, these programs do not facilitate any further data analysis, interpretation or archiving. The remaining program, FINS, not only supports the implementation of the Carle and Strub MWL methodology but is also one of the few packages that fulfils the wider range of additional functional requirements of flexible stock assessment software.

Of the software packages in use in the Agency, only the FINS package addresses the assessment of semi-quantitative data by means other than the simple reporting of 'minimum estimates'. However, the methods employed by FINS rely on the provision of an estimate of the probability of capture by the user. The software does not have the facility to provide the

user with a choice of likely values. It may be possible to provide a 'most likely value' based on data from other (fully quantitative) sites and information on the species being fished for; the size and nature of the site being fished; and the fishing method being used. Such possibilities are currently being explored under R&D Project 7716.

Various mark recapture methodologies are employed by the currently available software. For simple two-catch estimates, the Petersen (1896) method may be employed, whilst for multiple recapture methods the Schnabel (1938) method can be used. Neither of these methods make allowance for open populations (i.e. populations with potential immigration or emigration between samples). This situation is accommodated by the Bailey (1951) method which is itself simply a special case of the generalised Jolly-Seber models (Jolly, 1965; Seber, 1965). Beyond the information contained within the original papers, there is little guidance available on the relative merits and weaknesses of the alternative mark-recapture methods. Until the findings of a critical review of the alternative practices are available, it will not be possible to provide definitive guidance on the methodologies that should be embodied within any newly developed stock assessment software.

Whole fishery statistics

There is no stock assessment software used by the Agency that is totally dedicated to the derivation of 'whole fishery' statistics from electric fishing survey data. All whole fishery statistics that are produced come from packages which deal primarily with site-specific information but which permit the combination and subsequent reporting of data from more than one site.

Notwithstanding this, the only whole fishery statistics produced by currently available software are simple averages of population density or biomass over (pre-specified) groups of sites. This procedures that are used are not optimal, in that they take no account of the relative contribution of each site's area to the overall area of the target reach. Appropriate methodologies for deriving mean population density/biomass estimates are discussed in R&D Note 292 (Wyatt and Lacey, 1994).

Planning

Other than the demonstration software produced under R&D Project 325 there is no software product used within the Agency to assist in the planning of survey programmes. The demonstration software is not widely available and, although it has been used by some Agency staff, it was intended solely for demonstration purposes and not for operational use.

4.3.3 Ease of use

In general terms, the strongest messages that came across from the regional meetings was that software should be as user-friendly as possible, and that it should be flexible enough to allow changes in data collection methods or statistical techniques, or the incorporation of additional analyses to be undertaken, with a minimum of disruption.

Whilst all of the software discussed above is in regular use, and can therefore be thought of as being adequate in terms of its ease of use, there was a general feeling that currently available software was idiosyncratic and often required a degree of familiarisation before it could be used effectively. The available stock assessment software is based on a range of commercial software platforms, and generally fails to make optimum use of the facilities provided by the software and hardware used by the Agency.

4.3.4 Compatibility

It was recognised that the current situation, with several distinct pieces of software being used across the Agency for effectively the same purpose was not ideal.

Whilst there are obvious regional differences regarding survey methods, the statistical processes that should underlie the analysis of the raw data that are produced are relatively straightforward. However, the use of sub-optimal or inappropriate analytical methods has been seen to be a potential problem with existing software. The consequent application of a range of alternative (and often inappropriate) methods considerably reduces the compatibility of outputs. On a more fundamental level, data that have been archived using one software system is effectively unavailable to other software packages, and so effectively perpetuates the use of inappropriate software.

4.4 Adaptability of current software

4.4.1 Current Agency software

Software that is currently available within the Agency has evolved in an apparently piecemeal manner over many years. Only a very restricted number of packages, capable of dealing with a range the Agency's reporting requirements, have been produced and are still in use. These are:

- FINS
- FDPS
- Survform.

Of these, the only system currently available within the Agency that has a sufficiently broad functionality, and that is therefore potentially worth adapting for use at a national level, is the FINS package.

However, the software platform for FINS is outmoded and the software would require a substantial re-write to bring it into the Windows operating environment. In addition, some users have expressed the feeling that FINS is not particularly intuitive or user friendly, and there are recognised limitations to the algorithms that are used for the calculation of variance estimates for population statistics. Consequently although, in functional terms, FINS embodies the majority of features required by a national stock assessment package, the adoption of the current version of FINS as a national package would not be an effective solution to the regions' operational requirements.

Given that, within the short- to medium-term future, new standard software needs to be made available for fisheries stock assessment within the Agency, it would not be appropriate to expend effort on the development of 'new' interim software through the modification or improvement of current software. Although it may be possible to modify or upgrade one or more existing software products such that they are able to fulfil the requirements of the Fisheries Function (whilst employing the most appropriate analysis methodologies) there would inevitably be a degree of inertia in the uptake and implementation of such software. This inertia would be linked both to the inherent delay in the implementation of newly modified software (distribution, installation, etc.) and the requirement for staff to gain familiarity with the software (i.e. the need for staff to climb a 'learning curve').

Where such inertia was associated with the implementation of what was effectively only an 'interim' product, the advantages that may potentially be conferred in the short-term may be lost. Consequently, it is not recommended that such adoption is undertaken at the national level, as it is felt that this would be counter-productive over anything other than the short term.

4.5 New software development initiatives

4.5.1 Background

At present, Midlands region make intensive use of stock assessment software based on the SmartSuite II and Smart 3.1 packages. However, IS recently indicated that they intended to withdraw support to this software platform. This withdrawal of support is to be associated with the removal of these software packages from the region's computer systems. Due to the obvious consequences that this would have (with respect to the ability of fishery managers to analyses and report routine, strategic or reactive stock assessment data) a development programme was initiated with the aim of producing a new regional software package for fisheries stock assessment.

The following sections outline the functionality of this new software. The full specification for this initiative as drawn up by the Regional Fisheries Officer, Paul Lidgett, is reproduced as Appendix E. The software is currently being developed as an Access run-time application by Dave Martland (from Welsh region's IS section).

4.5.2 Functionality

Overall, the software that is developed through this initiative should process and store data obtained from fish population surveys. More specifically, the software should be capable of:

- recording various types of raw survey information, including individual fish lengths, bulk fish weights, species numbers and numerous descriptive data
- analysing these data according to standard fisheries science procedures
- archiving various raw data and processed results in a format which is accessible for future retrieval and reprocessing

- producing a range of standard survey reports.

The reporting functions would be available at four different levels, producing what are effectively described as: 'jobsheet reports'; 'habitat reports'; 'population reports'; and 'length distribution reports'.

Jobsheet reports

One jobsheet report would be produced for each survey, and would consist of information on: the survey date and location; riparian ownership details; and personnel present and equipment used.

Habitat reports

Similarly, one habitat report would be produced for each survey, and would consist of information on: the survey date and location; the site dimensions; various habitat data.

Population reports

One population report would be produced for each survey, although three distinct types (type 'a', 'b' or 'c') would be available. Type 'a' would be produced for population estimates derived using depletion fishing methods, and would consist of information on:

- the survey location and date
- the numbers and weights of each species caught on each run
- the species composition of the total catch (percentage composition by both numbers and weight) including the option of a pie chart as a graphical output
- population estimates (as produced by the Seber & LeCren; Zippin; or Carle & Strub methodologies)
- biomass estimates (total weight per species)
- density estimates (number.$100m^{-2}$)
- standing crop estimates (g.$100m^{-2}$)
- species richness
- fishing efficiency
- other comments (i.e. a 'memo' field).

Type 'b' would be produced for minimum estimates and would consist of information on:

- the survey location and date
- the numbers and weights of each species caught
- density estimates (number.$100m^{-2}$)
- standing crop estimates (g.$100m^{-2}$)

- the species composition of the total catch (percentage composition by both numbers and weight) including the option of a pie chart as a graphical output
- species richness
- other comments (i.e. a 'memo' field).

Type 'c' would be produced for surveys that were only able to produce presence/absence data and would consist of information on:

- the survey location and date
- a list of the range of species caught
- subjective estimates of the relative abundance of the different species
- species richness
- other comments (i.e. a 'memo' field).

Length distribution reports

Finally, for each species caught on a given survey, a length distribution report would be produced. These reports would consist of information on:

- the survey date and location
- the total catch of the species (i.e. the effective sample size)
- an estimate of the species population size
- an estimate of the species biomass
- the length range of the fish that were caught
- the length-weight coefficient for the species
- the length-frequency distribution of the species, including a histogram as an output
- other comments (i.e. a 'memo' field).

4.5.3 Underlying methodologies

In terms of the formulae necessary to derive population estimates and associated variance statistics, the development of the new Midlands' software was intended to be in line with Appendix 1 of the NRA Interim Report for R&D Project 325 (Lacey *et al.*, 1992). However, despite outputs from R&D Project 325 recommending that Seber & LeCren and the Zippin methods of population estimation be rejected in favour of the adoption of the Carle & Strub (Maximum Weighted Likelihood) methodology, the proposed software allows for the generation of population estimates by any (selected) one of the three methods. In addition, there is no evidence to suggest that variances for population estimates are to be estimated by simulation methods rather than deterministic formulae - the former being the option recommended by R&D Project 325.

From its description, it would appear that the proposed Midlands software is (not unreasonably) biased towards operational practice within Midlands region. Consequently, its functional specification does not cover all of the requirements that have become apparent at

the national level. In particular, there is no provision for estimating population size from mark-recapture data.

5. USE AND AVAILABILITY OF SOFTWARE OUTSIDE OF THE AGENCY

5.1 Software applications in use outside of the Agency

5.1.1 The range of software used

Basic information on the nature of fisheries software currently in use outside of the Environment Agency was obtained from those questionnaires that were returned by commercial organisations and academic institutes. A précis of this information, indicating the extent of software usage, is given below, as Table 5.1.

5.1.2 Software for population estimation

The stock assessment software that is used by those groups that responded appears to be limited to essentially the same range that is used within the Agency. Those groups that did not identify a specific package for stock assessment tend to make use of 'standard' population estimate predictors (e.g. Zippin, Carle and Strub, etc.) to process their removal data. To implement these procedures they make use of simple routines written using whatever software they employ on a day-to-day basis (e.g. the 'Excel' spreadsheet package, the Minitab statistics package).

As indicated in Table 5.1, IFE, DANI and SOAEFD make use of the 'Remove' program (Clarke, 1992). The Fisheries Conservation Board for Northern Ireland currently use the FINS package to fulfil their analysis and archiving requirements, whilst the Fisheries Department at UWCC (University of Wales, College of Cardiff) rely on an interactive computer program (based on Higgins, 1985) for population estimation using the Zippin method.

Several of the groups that responded undertake semi-quantitative surveys, most subsequently using known efficiency factors or estimates of the probability of capture to derive population estimates. As a slight variation to this, the Salmon Research Agency is intending to incorporate new strategies for relating fishing effort and fishing efficiency that have recently been developed (Connolly, 1996). The details of this technique or its application are not yet clear.

5.1.3 Software for additional stock assessment applications

IFE Windermere identified the FSAS series of analysis programs as being of use in their stock-assessment analyses. IFE Wareham make use of the 'Backcalc' program (written in the BASIC programming language) for generating back-calculated length-at-age data from scale or otolith measurements.

Table 5.1 Use of software outside of the Agency

Organisation	Is software used for survey design ?	Name of stock assessment software that is in use
DANI - River Bush Salmon Station	-	'Remove'
Fisheries Conservation Board for Northern Ireland	✓	FINS
IFE - Eastern Rivers Laboratory	-	'Remove'
IFE - River Laboratory	-	'Remove'
IFE - Windermere Laboratory	-	'Remove'
Powergen Freshwater Biology Group	-	-
Salmon Research Agency - Co.Mayo	-	-
SOAEFD - Freshwater Fisheries Laboratory - Pitlochry	-	'Remove'
University of Aberdeen - Zoology Department	-	-
University of Liverpool - Dept. of Env. and Evol. Biol.	✓	-
University of Westminster	✓	-
UWCC - School of Pure and Applied Biology	✓	'Zippin'

5.1.4 Software for survey design

Although four of the groups that responded said that they made use of software in the survey design process, no specific packages were identified. In each, respondents stated that 'standard software' (e.g. Excel spreadsheets) was used to fulfil this function. No detail was supplied of the range of considerations that were taken into account during the planning procedure.

5.2 Other available software

5.2.1 General

In addition to those packages used by the Agency or by UK research organisations, other fisheries analysis software is readily available for use within the UK. Such software is however, generally limited to that which has been developed for marine fisheries assessment. Three such packages that have been identified by this study (ELEFAN, LFDA and CEDA) are discussed briefly below (Section 5.2.2).

The range of software that is available for freshwater fishery applications is supplemented by a range of products from the United States, a brief appraisal of which is given in Section 5.2.3.

5.2.2 Software marine fishery oriented applications

Many fishery workers are familiar with the ELEFAN system for length-frequency analysis (see, for example, Pauly, 1987). The original ELEFAN software, produced by ICLARM (the International Centre for Living Aquatic Resource Management), does not support a wide range of functions - its application being essentially restricted to the estimation of growth and mortality parameters from pooled length-frequency data. However, despite the apparent familiarity that fishery staff have with ELEFAN it is not in routine use, probably because of its limited functionality.

The LFDA (Length-Frequency Distribution Analysis) package that is produced by MRAG (the Marine Resources Assessment Group) facilitates not only the ELEFAN approach to the estimation of growth and mortality parameters for length-frequency data, but also allows alternative methodologies to be applied. For example, in addition to the ELEFAN method, the parameters of the von Bertalanffy growth curve may be estimated from length-frequency data by Shepherd's Length Composition Analysis ('SCLA' - Shepherd, 1987) or by the project matrix method ('Projmat' - Basson *et al.*, 1988). The LFDA package also allows for mortality rate estimates to be derived using the Beverton-Holt method (Beverton and Holt, 1956), a method based on the use of the projection matrix, a method that allows the rate to be estimated directly from the derived age-frequency distributions, or by the Powell-Wetherall method (Powell, 1979; Wetherall *et al.*, 1987). Despite this range of approaches however, the LFDA package can (like the original ELEFAN package) be thought of as a 'specialist' system, dealing as it does with only one aspect of the overall stock-assessment process that Agency staff undertake.

The CEDA (Catch Effort Data Analysis) package (also produced by MRAG) is PC-based system for analysing catch, effort and abundance data, giving estimates of current and unexploited stock size, catchability and associated population dynamic parameters. Whilst presenting a series of options that deal quite comprehensively with the analysis requirements of large (marine) dataset, the package does not represent a useful analysis tool for the assessment of fisheries data as generated by the Agencies routine or strategic fisheries monitoring programme.

5.2.3 Software from the United States

Source of material

The American Fisheries Society (A.F.S.) has a Computer User Section, that produces a listing of their program library. The January 1994 listing includes over 20 software packages, covering most aspects of fisheries management. However, only four packages deal with the analysis of fisheries survey data and are of relevance to this study. Two of these (Microfish and FISHPROG) produce population estimates from survey data, whilst the remaining two packages (FISHPARM and the FishCalc89-DisBCal89 combined package) provide means of assessing length and age data.

In addition to the material listed by the A.F.S., one other package for estimating fish population size (and, incidentally, production rates) has been identified. This is the Pop/Pro modular suite of software (Kwak, 1992).

Microfish

Microfish is a program for generating population estimates for removal data based on maximum likelihood estimation theory.

FISHPROG

This package estimates fish population sizes and annual production rates in small streams from multiple pass sampling data. Input is in units of length (as centimetres) and of weight (as grams). In addition to providing total population estimates for each species, the outputs also consist of catch (by length class) for each sample-species combination, catch for each pass, average weight per length class, and production for each species by site.

FISHPARM

FISHPARM performs non-linear parameter estimation for 13 separate statistical models commonly used in fisheries. These models are: the von Bertalanffy growth function; Gompertz growth function; Beverton-Holt recruitment function; Gamma function; Shepherd recruitment function; Allometry equation, Power function; Exponential growth function; Logistic growth function; Quadratic function; Weibull cumulative distribution; Mesh selection; and LD-50 estimation by the logistic model.

DisBCal89-FishCalc89

The DisBCal89 portion of this package is used to measure linear projections of bony fish parts (i.e. scales or otolith samples) with a digitizer, to explore the relationships between fish body length and the sizes recorded from the bony parts, and to back-calculate body lengths.

The FishCalc89 portion performs cross-tabulation on fish length, weight, sex and age. It also produces (high-resolution) tables and charts of: length structure and population density; CPUE and percent of sample by length interval; age composition; mean lengths at age; and the year-class contributions to each length class.

The software does not analyse survey data to produce population estimates, but rather provides a standardised means of presenting length and age data in combination with information on population size.

PopPro

This is a series of modular routines covering population estimation (four modules) and production estimation (a further four modules) and is briefly described by Kwak (1992).

The population estimation modules cover estimation by the removal method (a maximum-weighted likelihood method, after Seber, 1982 and Bohlin *et al.*, 1989) or by a single-census mark-recapture method (a modified Peterson method, after Ricker, 1975). In both cases the software will perform calculations by age class or (if age data are not available) by size class. The production estimation modules use either the instantaneous growth rate method or the increment summation method of analysis.

The software is not fully interactive, in that the majority of inputs to the software are mediated through input files.

5.3 Assessment of software from outside of the Agency

Several important functions (involving population estimation, length-frequency data analysis and growth analysis) can be effectively undertaken using combinations of the software from outside of the Agency that is outlined in the preceding sections. However, amongst the software that has been identified there would not appear to be any particular package that would provide new facilities to the Agency's Fisheries Function.

6. OPTIONS FOR FISHERIES SOFTWARE DEVELOPMENT

6.1 The need for software development

6.1.1 Introduction

A sound knowledge of the status of fish populations is a pre-requisite to informed fisheries and environmental management. Within the Environment Agency, fish population surveys are afforded a high priority, with a large number of sites being surveyed each year. The data obtained from these surveys are many, complex and varied, and require substantial statistical analysis, presentation and referral before being used in management decisions. There is a consequent need for appropriate software to be available for use by staff throughout the Fisheries Function. Such software should satisfy a range of criteria, *viz.*:

- it should be suitable for its intended use (e.g. making use of valid and appropriate statistical methods)
- it should be easy to use (in terms of actual software implementation)
- its outputs/results should be unambiguous, and appropriate to the intended audience or use
- it should make optimal use of available computer resources
- it should be widely available and
- it should be standardised across all regions of the Environment Agency.

As noted within Section 4.3, much (if not all) of the software systems that are currently in use within the Environment Agency fail to meet these criteria. In general terms, the available software is outmoded and, in many cases, has been found to make use of inappropriate methods. At present, there is no agreed national standard for stock assessment analysis methods or for reporting requirements. In these terms, the poor quality of the stock assessment software that is available represents a potentially serious limiting factor to the provision of informed and effective fisheries management.

The shortcomings of the software within the Agency cannot be readily addressed by the adoption of software from the range that has been identified outside of the Agency (either within the UK or, in the guise of the American Fisheries Society, from the United States). No software has been identified from such sources that would be able to confer any advantages over the range of software which is currently in use within the Agency.

Over recent years the national R&D sponsored by the Agency (and previously the NRA) has provided several important tools that help satisfy the needs of fishery managers regarding both data analysis and interpretation. There is an increasing need for the outputs of such R&D (not only in the form of new tools and methodologies, but also in terms of the elucidation of optimal methods in situations were several alternatives are available) to be made available throughout the Agency. However, at present, there are only limited resources being directed

towards the development or modification of fisheries stock assessment software. Software modification (and minor development) is likely to be currently ongoing in many regions, but only in a reactive and *ad hoc* manner.

The only current example of 'proactive' software development (i.e. software development to an agreed programme and specification) is in Midlands Region. The development of this new software, which will support a range of data input, analysis and reporting functions (as outlined and discussed in Section 4.5) should result in a software product that will be of potential value to several of the Agency's regions.

6.1.2 Possible development options

In addition to the 'do nothing' option, there are three further options available to the Agency that would potentially satisfy the need for the provision of improved software:

- the modification or improvement of existing software
- the adoption of current development initiatives
- the initiation of new national R&D relating to the development of a suite of stock assessment software.

Do nothing

The 'do nothing' option is not viable in this instance. A failure to develop the present position, where disparate and inappropriate systems are being used to analyse and archive fisheries data, will perpetuate the current unacceptable situation where reported data are not directly comparable across the regions, and will potentially hamper future attempts to formulate national policies. In addition, the absence of a national standard for data archiving leads to a situation where large quantities of data, that are of great potential value in many national R&D initiatives, are effectively inaccessible.

Modification / improvement

From the software that currently in use within the Agency, there is none that can be easily adopted for effective national use. Furthermore, none of the software that has been identified as being in use outside of the Agency is any more sophisticated than, or can offer any advantages over, the software that is currently used within the Agency.

It is not thought that the modification or improvement of existing software represents an effective or appropriate means of ensuring the provision of stock assessment software that meets the criteria outlined in Section 6.1.1.

Adoption of current development initiatives

The production of new stock assessment software for Midlands region (currently in preparation) represents an important development initiative. However, whilst the functional

specification is reasonably robust, it does not fully satisfy the range of requirements that have been outlined by the regions.

The software does, however, have the distinct advantage that it is being developed as a stand-alone (run-time) application running within the Windows environment. Consequently its use should be largely intuitive, more especially if it has been developed with reference to the Agency's (draft) document on 'Guidelines for computer human interaction' (Environment Agency, 1995).

For those regions or areas who feel that they have only limited access to suitable software, adoption of the Midlands software should be considered as an interim solution. However, whilst it may be possible for the Midlands software to be adopted by those regions that currently have limited access to appropriate software this option is not recommended as anything other than a short-term solution.

Development of new software

The development of new software, under a national R&D initiative, represents the best option for achieving the objective of providing software that fulfils the requirements listed under Section 6.1.1, and for ensuring that such software is subsequently available throughout the Agency. It is therefore recommended that new software be developed to satisfy the Agency's needs regarding the analysis and archiving of fishery stock assessment data.

Whilst it is reasonable to expect an increase in the degree to which data analysis, reporting and archiving are standardised across the Agency's regions, it is unlikely that the detail of surveying strategies (e.g. general methodologies, field practices, etc.) would ever be dictated to (and hence standardised across) the regions. Software development under a national initiative would help produce a situation where the archiving, processing, and reporting (i.e. outputting) of information that is collected by a wide range of survey strategies can be achieved in a manner that is consistent throughout the Agency.

6.2 Recommendations for software development

6.2.1 General

The development of new software should be undertaken on a modular basis, with the potential for the development of modules to be prioritised and phased. In this context, the term 'module' refers to one or more software routines that accomplish a pre-defined set of related functions. Under this definition, separate modules could be simply a series of groups of routines within a larger program, or may be a distinct programs that utilise common input and output data. In either case, it is envisaged that the data used and produced by the proposed modules would be stored as a series of tables within a database. It is intended that such a database would, in addition to being an intrinsic element of the proposed software, be capable of being interrogated by commercially available software (e.g Access). In this way, the subsequent use of information that is archived within the database would not be limited to

those analyses that are undertaken by the proposed 'dedicated' software that is outlined in these recommendations.

The prioritisation of module development would permit the allocation of R&D resources to be better focused, such that the most pressing (software) needs of the Agency could be met ahead of less pressing requirements, whilst ensuring that appropriate software is made available to staff with the minimum of delay. This has the advantage of saving time at the initial development stage (in that, if an analysis routine has a low priority then it is not produced ahead of other systems). For example, it may be decided that, along with the data input systems, only the basic elements of the analysis routines should initially be produced. For whatever reason, it may be decided that the development of (for example) additional modules for the 'intermediary analysis' of raw data and for the automated production of survey reports could be delayed.

Modular development would also allow redundancy in the software to be restricted. For example, features that were subsequently found not to be required or that were not used, would not be 'hidden' amongst other more useful routines, a situation that may easily occur if all of the required routines were to be produced within a single software package. Consequently, redundant modules could be easily identified and (if necessary) removed, without danger of corrupting the coding for other required routines. In addition, a modular approach would only need the component analysis routines to be developed once, and would enable the straightforward redevelopment or replacement of methodologies should alternatives or improvements become available.

Separate modules could be developed for each of the areas identified within Section 3.2. For example, a series of modules would initially be required to handle the inputting and archiving of raw data. Further 'intermediary analysis' modules would then be required for the processing of length, weight and age data.

Subsequently, further modules would be required to deal with (for example):

- population estimation
- interpretation of population estimates (with the incorporation of additional information where appropriate)
- reporting
- planning.

The adoption of a modular structure would also facilitate the subsequent development of (potentially) automated linkages to other analysis, interpretation or reporting systems (such as the National Fisheries Classification Scheme, HABSCORE, etc.).

Finally, where it was considered appropriate to report fishery survey information at a series of different levels (e.g. as suggested by Southern region - see Section 3.1.3) this would be facilitated by a modular approach. For example, a range of separate reporting modules could be developed, each of which facilitates the implementation of the various analyses and outputs that are required at the relevant reporting level.

6.2.2 Overview of specification

As well as providing a means of standardising the analysis and reporting of fisheries statistics ahead of the production of new stock assessment software, the derivation of national analysis and reporting guidelines (see Section 6.5) would effectively form part of the technical specification for the new software.

The choice of software platform for new software will, effectively, be directed by the Agency's C.I.S. function. Close discussion with the C.I.S. function will be required on this matter when the Terms of Reference are drawn up. It is recommended that a nominated contact point for the C.I.S. function be identified and sit on the Project Board associated with any new development.

Figure 6.1 indicates the inter-relationships between the various sets of data relating to fisheries stock assessment (Tables A to F) and the software modules that would be involved in their calculation or manipulation. The various elements of Figure 6.1 are discussed in detail below.

In Figure 6.1, the two blocks to the left of the figure (plain text in single, solid boxes) represent data archives. The upper block represents the proposed fishery data archiving system, whilst the lower block represents additional data archives holding supplementary fishery data (such as that derived from hydro-acoustic surveys, from fish counters or traps and from catch-effort data).

The central column of elements in Figure 6.1 (all double boxes) represent the inputs to the archiving system. Text in solid boxes represents 'raw' data, whilst text in broken boxes represents 'derived' data.

The elements to the right of the figure (text within bold boxes) represent data manipulation, analysis and interpretation modules. The upper set of elements (text with broken bold borders) represent modules that produce derived data which is subsequently written back to the data archive system. The lower set (those with solid bold borders) represent modules that encompass routines for interpreting data.

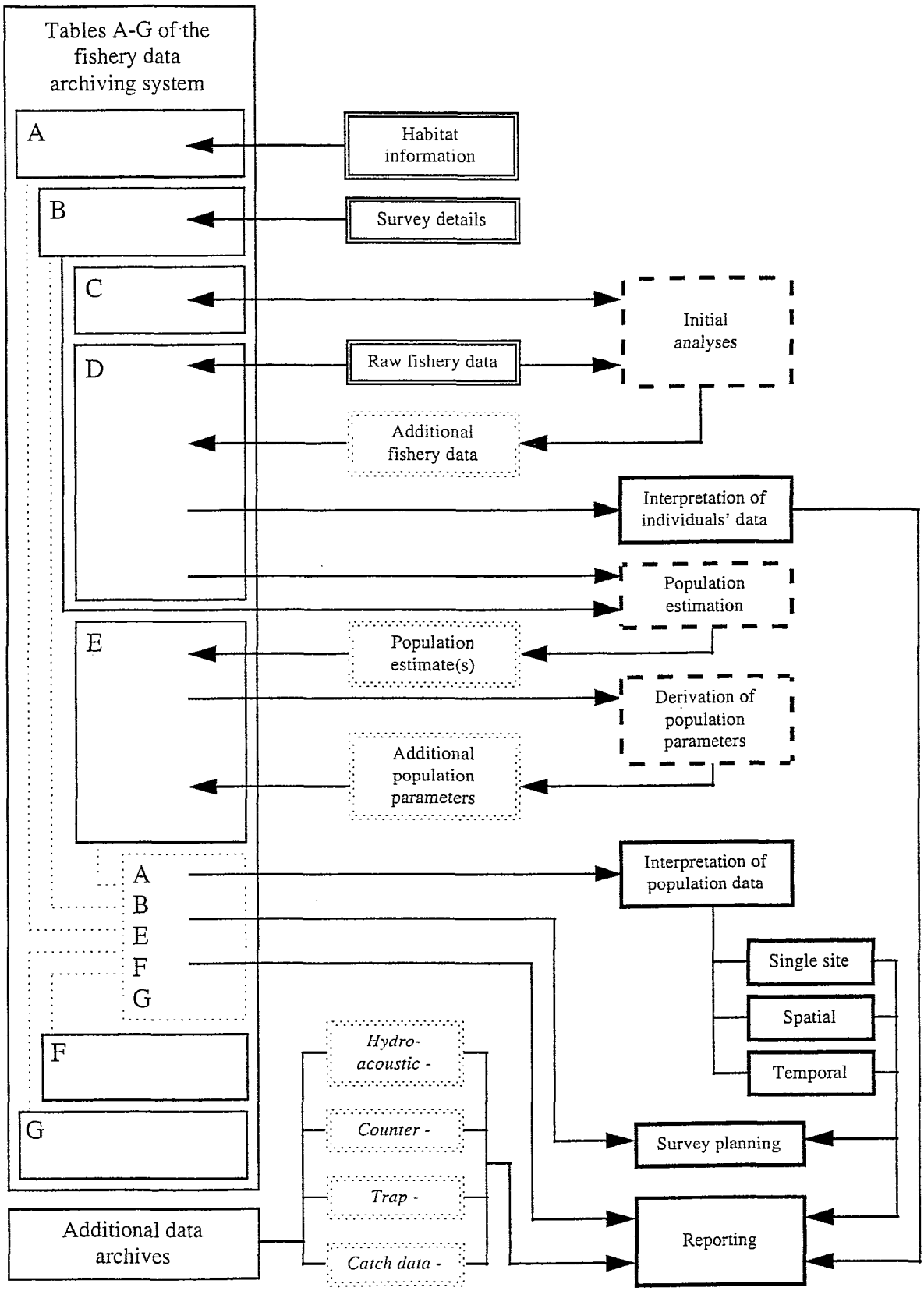


Figure 6.1 Relationships between elements of proposed software system - see text for detail

6.2.3 The fishery data archiving system

The block to the upper left of Figure 6.1 represents the fishery data archiving system. This should take the form of a relational database, consisting of a series of tables (labelled A to F in Figure 6.1). These tables would hold all of the raw data that is produced by routine fishery survey work, in addition to supplementary information that may be required by analyses that may be subsequently applied. Detailed descriptions of the fields that would be contained within each table are given in Appendix F.

Whilst ratification is required from the Agency, it is proposed that there should be separate tables holding data on site information (Table A) and on survey details (Table B). Each survey that is undertaken would result in a new record being generated in each of these tables. Raw fishery data would be stored in a third table (labelled D in Figure 6.1).

In the first instance, data would be entered to these three tables via three separete input modules (i.e. one module each for site information, survey details and raw fishery data). In addition, further data would be appended to the raw fishery data table by the first of three intermediary analysis modules (e.g. using predetermined length-weight relationships stored in Table C to generate estimates of weight for individuals that have only length data).

Derived population data (e.g. estimates of population size and biomass) would be written to another table (labelled E in Figure 6.1) within the data archive. As for site and survey data, derived population data would be stored as new record(s) appended to a single table.

A fifth table (Table F) would be a 'read-only' table holding information on species codes and full (English and Latin) species names.

Finally, a sixth table (Table G) would hold definitions of age codings.

Linking between Table A and B would be by means of a 'site code' field, common to both tables.

Each of Tables B, D and E would contain a 'survey code' field. Most cross-referencing between tables would be facilitated by linking this field.

Linking between tables containing fishery data (Tables C, D and E) and the tables holding species code information and age code definitions (Tables F and G respectively) would be facilitated by using the common 'species code' and 'age code' fields.

6.2.4 Direct (raw) data inputs

The required raw data inputs are represented in the centre portion of Figure 6.1 as text within double boxes.

Each of the three sets of inputs shown would require a data input module to facilitate the entry of information to the fishery data archive. All three of these input modules should allow inputs

to be made directly (i.e. from the PC keyboard) or indirectly (i.e. utilising information down-loaded from a data-logger or a portable PC).

Site information

The precise nature of the habitat elements of the fishery data archive are currently unclear, and decisions on the nature and detail of information that should be entered should await the completion of the proposed (second phase) R&D on the river fishery habitat inventory. Nevertheless, it is possible to indicate the nature of some of the likely fields to be contained within this table, e.g.:

- site name
- NGR
- site dimensions - length & width, or area
- etc.

It is likely that the range of information required by this input module would be largely independent of that which is entered in the 'survey details' module. Nevertheless, there is the potential for some relationships between these two modules to be exploited. For example, certain habitat data (such as gradient, altitude, etc.) could be automatically entered to the fishery data archiving system through the interrogation of a GIS using the site location (NGR) details previously entered under the 'survey details' input module.

Survey details

The data entry module for survey details would include facilities for inputting data on:

- survey date
- survey method - plus (where appropriate) the number of runs; time spent fishing; and the number of repeat fishings
- survey staff and other notes.

It is proposed that the selection of survey method should automatically enable/disable the requirement for other information covered by this input module. For example, where the survey method is 'CPUE' an estimate of the probability of capture would be required, whilst for surveys based on 'removal' methods, the number of runs would have to be supplied. Similarly, for surveys that are identified as being based on 'mark-recapture' methods, the number of removals would be required as an input.

Raw fishery data

The data entry module for raw fishery data would cover the input of information on:

- species
- individual fish length
- individual fish weight

- fish age
- (for surveys based on constant effort) the estimated probability of capture for each species/age combination and
- (for mark-recapture surveys) the occurrence of marks.

The raw fishery data input screen would require an input indicating the site code and survey date to facilitate cross-referencing between the raw data table that is generated and the appropriate record in the 'survey details' table.

In addition, the form of the fishery data input screens that are presented to the user should be controlled by the information entered in the previous (survey details) module. For example, unlike surveys that are undertaken by the removal method, situations where the survey method has been identified as 'mark-recapture' would require supplementary information on whether a given fish was marked and (if so) when such a mark had been obtained. Similarly, where the survey was 'single run' based on a known efficiency, then this efficiency would need to be entered with the raw fishery data. Where not required, the entry boxes would for these data be disabled.

6.2.5 Derived inputs from intermediary analysis modules

These inputs are shown in the centre portion of Figure 7.1 as text within boxes edged with double broken lines. The intermediary analysis modules that are required are represented on the upper right of Figure 6.1 as text within boxes edged with broken bold lines. Three such modules would be required - one for the initial analysis of raw data, a second for the generation of population estimates and a third for the production of population statistics.

Implementation of the three intermediary analysis modules (described below) would effectively complete the input of information to the fishery data archiving system. It is not envisaged that any further analysis or interpretation of the archived data would produce additional information that would be stored within the fishery data archiving system.

Initial analyses

The first of the required intermediary analysis modules would facilitate the initial analysis of fishery data (e.g. processing of bulked data, routines for the estimation of weights from length data, or of age from length data, etc.). Outputs from this module would be written back to the raw data table (i.e. Table D in Figure 6.1).

For the purposes of generating weights from length data, this module should contain the option to either default to known length-weight relationships (as stored in Table C) or to estimate relationships from the observed data. This module should also allow the details of any newly derived relationships to be archived back to Table C, as new records.

Population estimation

The second intermediary analysis module would be responsible for the generation of population estimates, using the data previously entered to Table D, and would write the results out to a fourth table (Table E) within the fishery data archiving system. A series of options would be available within this analysis module, according to the nature of the raw data that are available.

For estimates based on CPUE methods, it may eventually be possible to derive estimates of the probability of capture from habitat data. Although the likelihood of such methods being available will be clearer following the completion of the current R&D initiative on semi-quantitative methods this module may, in addition to the necessary links with Table D, require additional links to the table containing habitat data (Table B).

Where appropriate, additional information (e.g. estimated probability of capture, variance of the population estimate, etc.) would also be produced and written to the table of derived population data (Table E).

It is envisaged that the population statistics derived by this module would be produced independently for all age classes of fish represented in the fish caught during the survey.

Table 6.1 Population estimates estimable from different survey data

	Possible population estimates:		
Data derived from:	**fully quantitative**	**semi-quantitative**	**minimum estimate**
Mark-recapture methods	✓ *	✓ *	✓
Removal methods	✓	✓ †	✓
CPUE methods	n/a	✓ †	✓

Notes

*: requires R&D to establish optimum methodology

†: development of appropriate methodology being addressed by current R&D

n/a: not applicable

Derivation of population parameters

The final intermediary analysis module would use data on population estimates (at the species and age levels of detail) to produce more general population statistics. It is likely that these statistics would include, for example, estimates of total fish biomass, and the density/biomass within redefined species groupings (such as the predatory, rheophilic and limnophilic groupings that are used within the NFCS).

The information used to derive these statistics would be taken from Table E and would be written back to Table E as new records.

6.2.6 Further analysis and interpretation

The remaining analysis modules shown in Figure 6.1 are in the right portion of the figure and are represented by text within boxes edged with solid bold lines. They represent the analysis and interpretation of the data that are stored within the fishery data archiving system and, rather than being stored back within the fishery data archiving system, their outputs would be utilised by the reporting or planning modules.

Interpretation of individuals' data

This module would allow (for example) the production of size-frequency plots, size-frequency plots by age (as stacked bars), the results of analysis of growth rates, etc. As this module would be concerned with the production of statistics or graphics for subsequent reporting, rather than the generation of data for archiving, the exact functions that it would support would be driven primarily by the proposed development and subsequent acceptance of standards for reporting fishery survey information within the fisheries function.

Interpretation of population data

It is likely that population data will need to be reported in three distinct manners, and interpretation of the data that is undertaken by this particular module will need to reflect this by allowing:

- interpretation of the data obtained from a single survey at a single site (e.g. absolute values for density and biomass - given by species - together with the relative percentage contribution by each species to the total estimated density or biomass)

- temporal analyses (i.e. the same site(s) but analysed through time) - e.g. year-class strengths, year-on-year growth rates, survival rates for specific year-classes, etc.

- spatial analyses - e.g. estimating mean density or biomass values across groups of sites (i.e. the production of reach averages).

Where appropriate, the data for the separate age-classes of fish (as held in Table E) would be combined, to produce overall estimates for the species.

It is envisaged that for the purposes of data interpretation, site data and survey details for each site (as stored in Tables A and B, respectively) would be used in conjunction with the population data that are held in Table E.

As for the production of statistics based on individuals' data, the interpretation of population data will, necessarily, be driven by the agreed reporting requirements of the Agency.

6.2.7 Reporting

It is proposed that the reporting module would not contain any routines to analyse or otherwise interpret the population or individuals' data that are passed to it. The reporting module would simply provide routines by which pre-determined sets of information can be taken from either the preceding interpretative modules or from the fishery data archiving system itself.

In addition to facilitating the reporting of different groups of information (i.e. single site; spatial and temporal assessments) the reporting module should also permit levels of detail to be reported. For example, for a single site analysis, a series of reports at different levels of detail may be required, in the manner described by Southern region (Section 3.1.3) or specified by Midlands region for their current software development (Section 4.5.2).

6.2.8 Survey planning

The survey planning module would facilitate the interrogation of the fishery data archive to provide information required for the effective planning (and design) of fishery surveys. Although, with the exception of the demonstration software developed under R&D Project 325, there is no software currently available within the Agency to assist in survey design and planning, the necessary methodologies have been assessed.

It should be possible to have what would effectively be an autonomous module dealing with survey planning and design, sitting alongside the fishery data archiving system. It would not provide an input to the archive, but would interrogate the database that the archive represents, together with (where appropriate) outputs from the two interpretative modules. In this way, many of the data that are required for the implementation of the routines developed for survey design may be produced directly from appropriate data sources, rather than having to be supplied from other sources by the user. Nevertheless, to increases the value of this module, it would be prudent to allow for estimates of the required data to be input to the planning module interactively (and to allow values derived from the archived data to be overwritten).

6.2.9 Incorporation of other data

The inclusion in Figure 6.1 of a representation of the additional data sources that may be incorporated into the reporting module is for illustrative purposes only. The use of fisheries data generated by means other than stock-assessment surveys undertaken by electric-fishing or netting (i.e. those methods other than the removal method, fixed effort sampling or mark-recapture techniques) lies outwith the terms of reference for this project. However, it is important to recognise that other methods of fishery assessment (e.g. hydro-acoustic data, counter data, trapping data, and angler catch data) may be routinely used within the Agency.

Whilst Figure 6.1 indicates the proposed structure of software for the processing, archiving and reporting of stock-assessment data generated by electric-fishing or netting surveys, no detailed consideration is made of the contribution from alternative methods. Nevertheless, the data from such assessment methods may often need to be reported alongside, for example, the results of a series of surveys that were based on the removal method.

It is appropriate, therefore, that attention should be paid to the potential contributions that could be made from data generated by alternative methods. To facilitate reporting it would be appropriate to move towards the situation where the archiving of data generated by other methods is undertaken in a similar flexible manner to that which has been proposed for data derived from electric-fishing and netting surveys.

6.3 Priorities, timescales and costs

6.3.1 Introduction

The current understanding of the methodologies that underlie the various elements of the proposed software system is not equable across the range of methodologies. Consequently, the development of certain elements or modules of the proposed software could be undertaken ahead of others. For example, the development of those modules that rely on methodologies that are, at present, poorly understood, could be delayed relative to those that rely on methods that are well defined and understood.

However, an additional factor that should drive the prioritisation of development is the intrinsic importance of each module. For example, the fact that one routine within a module (say, population estimation from mark-recapture data) relies on methods that are poorly understood and perhaps need to be investigated through new R&D work should not delay the development of the rest of the 'population estimation module'. Such a module would be central to the processing and interpretation of fisheries data, and many of the data handling routines that are used within the module would be equally valuable whatever the method of population estimation that is ultimately used.

The following paragraphs describe, on a module by module basis, what new work would need to be completed before the module could be developed to its final state - assuming that the overall software system follows the layout proposed in Section 6.2. In addition, an indication is given of the importance of the module.

The table at the foot of this section (Table 6.2) provides a summary of the modules' status - in terms of their recommended priority and whether or not further work needs to be undertaken to complete their development. The subsequent section (Section 6.3.3) provides a proposed schedule for development, together with an indication of the associated costs.

6.3.2 Prioritisation considerations

Input of survey details

This module one of the central modules of the proposed software system, and as such must be in place (at least in prototype form) for the software to be viable.

Whilst there are no potential R&D initiatives associated with the completion of this module, there are two areas that need to be considered. Firstly, the range of information that needs to

be entered (e.g. see Section 6.2.4) must be formally agreed upon. Secondly, the potential for down-loading data from a logger (or similar) needs to be addressed. This second task would necessarily follow completion of the first and would, in the simplest terms, require a standardisation for the format of down-loaded data.

Input of habitat information

This module is another of the central modules of the proposed software system. It will need to be in place (at least in prototype form) for the software to be viable.

At this stage, the range of habitat information that needs to be stored by the fishery data archive system has not been finalised. Current R&D initiatives on the development of a river fisheries habitat inventory and on the use of semi-quantitative methods need to be completed, and their recommendations used to help identify the range of habitat parameters that should be recorded.

However, it would be preferable to develop a 'prototype' module for handling habitat information, whilst accepting that modifications (most likely in the form of the addition of further habitat parameters) may need to be made following the completion of the R&D mentioned above.

Other than this (and, again, the use of loggers - as outlined above) there is no reason to delay the development of this module.

Input of raw fishery data

As it deals with the raw survey data this module is probably the key module as regards data input. It will obviously need to be in place (at least in prototype form) for the software to be viable.

Other than agreement on the format to be used for down-loaded data, there are no obstacles to the development of this module.

Initial analyses

This module is of particular importance where bulked data have been entered, or where the raw data are incomplete (e.g. where only lengths, and no weights, have been entered). Given that the software could produce basic population estimates based only the numbers of fish that have been entered, this module is not indispensable. Nevertheless, it does have a very high importance.

The allocation of individual estimates of length or weight from bulked data (for example individual weights from a bulked weight) would require the frequency distribution seen in a representative sub-sample to be applied to the bulked data, and the derived data pertaining to individual fish would then be written back out to the table of raw fishery data (table D). Routines should be set up to allow the provision of additional information to improve this

procedure. For example, upper and lower limits for lengths within the bulked sample could be entered - with a subsequent restriction of the sub-sample of measured fish that are used to derive the imputed lengths. All data that estimated by such means should be flagged accordingly.

This module would also facilitate the estimation (and subsequent application and archiving) of length-weight relationships from existing data or the straightforward application of 'standard' relationships that have been previously archived. Estimates would be flagged as such within the data table.

The facility to split up on-screen length-frequency plots into successive age-classes would be used to enable estimates to be made of the age of all those fish not formally aged by reference to scales or otoliths. Again, instances where the age of a fish has been estimated would be flagged as such within the data table.

It is not thought that there are any real barriers to the development of the routines that are necessary to implement any of the procedures required by this module.

Population estimation

This module is of prime importance within the overall software system.

However, as discussed above (Section 6.2.5), not all of the methodologies that may need to be applied within this module have been formally assessed. In particular, the use of mark-recapture and semi-quantitative data have not been fully addressed. Although the latter is the subject of current R&D, the various methodologies that are available for the interpretation of the data from mark-recapture studies remain un-assessed. Nevertheless, it is recommended that the module be developed as a priority.

For data obtained using depletion fishing (i.e. the removal method) the module should have routines that allow the Carle and Strub MWL methodology to be implemented (with variances estimated by simulation).

To prevent redundancy, a framework for using CPUE data should also be established within the module (e.g. with the user supplying estimates of the probability of capture). After the current R&D on semi-quantitative methods has reported, the relative merits of including new methodologies for the (automatic) estimation of the probability of capture that may subsequently be available should be assessed. Modifications to the software should then be made as required, with population estimates being recalculated where necessary.

Similarly, there will (initially) be little or no guidance as to the optimum methods that should be adopted for processing mark-recapture data. Rather than produce a population estimation module that is incapable of handling mark-recapture data, it is suggested that it be developed such that an agreed 'default' methodology is automatically implemented, allowing population estimates to be generated. As for semi-quantitative methods, the module can then be modified at a later date, allowing the findings of the proposed R&D initiative in this area to be incorporated.

Derivation of population parameters

Whilst the functions performed by this module would be important (in that they would provide information at a level which is easily communicated) the functions that it carries out could effectively be undertaken by hand.

Its development would effectively be dependent only on the agreement of which species are to be grouped together. As there are no other significant barriers to the development of this module its should be produced alongside the initial population estimation module.

Interpretation of individuals' data

The importance of this module can only be assessed following the development of guidelines for the reporting of fisheries survey data. However, it is unlikely that this module will be felt to be critically important in the development of the overall software system, although it would need to be completed before development of the reporting module could be successfully accomplished. Notwithstanding this, the information produced by this module may be regarded as 'supplementary' to the information that is reported at the population level by the second interpretation module (below) and as such, its development may be delayed relative to certain other modules.

In its proposed form (in which it is largely responsible for the production of graphical representations of data which relate to individual fish - e.g. length frequency plots, pie-charts of species composition by biomass, etc.) there are no significant barriers to the development of this module. It would, however, be necessary to agree in advance the form of the growth models that should be applied to length-at-age data.

Interpretation of population data

The development of this module has a high priority within the overall software system that has been proposed. It is this module that will take the processed data and convert it to a form that can be readily reported.

It is envisaged that there may up to three distinct elements within this module, relating to the interpretation of data from:

- a single survey at a single site
- an amalgamation of several sites (e.g. to derive an overall population estimate for a reach)
- one or more sites over a period of time.

The first of these is relatively straightforward (involving, for example, the calculation of densities from population estimates and site dimensions), although the contribution of habitat data would need to be addressed following the completion of current R&D on river fishery habitat inventories.

As regards the other elements of this module, guidance is available within R&D Note 292 (Wyatt and Lacey, 1994) on appropriate methodologies for the spatial or temporal

combination of site- and survey-specific information. In addition, methodologies are outlined that facilitate the production of test statistics for the comparison of datasets (e.g. for the temporal comparison of reach averages derived from successive surveys).

Additional temporal-based analyses (e.g. the calculation of year-class strengths, year-on-year growth rates, survival rates for specific year-classes, etc.) would also be facilitated by this module.

In general terms, it is proposed that little additional work is required in order to be able to develop this module.

Survey planning

With the exception of the demonstration (or interpretative) software produced under R&D Project 325 there is no software currently available within the Agency to assist with the planning of fishery surveys. Whilst the opinions expressed during the consultation phase of this current project suggested that the provision of such software would be welcomed throughout the Fisheries Function, it was noted that surveys were invariably resource limited, and that the detailed planning of survey programmes was therefore often academic. This, together with the fact that this module would be entirely independent of the reporting function supported by the proposed software, leads to the conclusion that the development of this module should attract a relatively low priority.

However, much (if not all) of the theory was reported under R&D Project 325 and it should be possible to develop routines within a 'survey planning' module that are able to interrogate the data that are stored in the 'derived fishery data' table within the fishery data archiving system. It should therefore be possible to provide routines within a 'survey planning' module that will provide guidance to fisheries staff at the survey design stage, without the need for further investment in R&D in planning topics.

Reporting

As the module that provides the final output from the software, the development of the reporting module should receive a high priority.

Once all of the modules that precede the reporting stage have been completed, the only item that needs to be addressed ahead of the production of the reporting module is agreement of the extent of information that is to be reported. As commented on in Sections 3.1.3 and 4.5.2, a range of different reporting 'levels', ranging from detailed, site specific reports to catchment overviews, may be appropriate. Such a range of reporting levels would be addressed by different routines within the one module, different data being taken from the fishery data archiving system as appropriate for the desired reporting level. Obviously, to facilitate this approach, it is first necessary to have agreement on the content of each reporting level together with agreed formats for report production. This, however, is likely to be the sole factor that needs to be addressed before the development of the reporting module.

Table 6.2 Summary of module development

Module	Importance within the overall software structure	Further work required to develop the necessary routines within the module
Input of survey details	High	Agree format for down-loaded data
Input of habitat information	High	Agree format for down-loaded data Agree nature of habitat data to be recorded [1]
Input of raw fishery data	High	Agree format for down-loaded data
Initial analyses	Moderate	Minimal
Population estimation	High	Minimal to become operational [2]
Derivation of population parameters	Moderate	Minimal
Interpretation of individuals' data	Low	Agreement of form of growth models to apply to length-at-age data
Interpretation of population data	High	Minimal to become operational [3]
Survey planning/design	Low [4]	Minimal
Reporting	High	Agree standard national format for reports

Notes

[1] the nature of data that is entered under this module may need to be reviewed and revised following the reporting of the second phase R&D on the development of a river fisheries habitat inventory.

[2] the module should be developed as a priority, although it must be recognised that modifications will subsequently need to be made.

the routines for 'semi-quantitative' estimation will need to be revised (possibly with the incorporation of routines to estimate probability of capture from habitat data) following the completion of current R&D on semi-quantitative methods.

R&D to assess the available methodologies for population estimation from mark-recapture data should be initiated, with routines for 'mark-recapture' estimation being revised following the reporting of the R&D.

[3] again, although the module can be developed, it must be recognised that modifications may subsequently need to be made (i.e. following the completion and reporting of the second phase R&D on the development of a river fisheries habitat inventory.

[4] although the importance of survey planning within the overall software structure has been assigned a low priority, it should nevertheless retain a high level of importance with respect to the overall fisheries management programme.

6.3.3 Timescales scheduling and costs

Software development

The anticipated effort involved in undertaking each of the work items associated with the proposed software development initiatives, together with guidance as to their likely approximate costs and an indication of the scheduling of the work, is shown overleaf in Figure 6.2.

The work items marked with an asterisk (the 'core items') are those that would need to be completed in order to produce an initial working software system. Such an initial (or 'core') system would not only allow the production and reporting of population estimates from raw survey data but, due to its modular structure, would effectively form the basis of a system that could be easily modified and extended in the future. Consequently, the additional modules that are specified in this report (i.e. those work items in Figure 6.2 that are given as bold text, but which are not marked with an asterisk - items 13, 15 & 17) could be completed and integrated as needed. In addition, the development of this modular software would allow new methodologies or analysis routines that are developed in the future to be incorporated as necessary and, through the modification of the reporting module, would permit new reporting requirements to be met.

In Figure 6.2, no differentiation has been made between those work items that would be undertaken by external contractors and those that could be undertaken by Agency staff (from either the Fisheries or the IS Functions). It is entirely feasible that all of the work items involved in the development of a 'core' system could be undertaken internally. It is more likely that, should some of the work items be undertaken internally, then this practice would be reserved for those items involved in the production of agreements on the functionality or detailed specification of the software (e.g. items 1, 2, 3, 5, and 11).

The final work item shown in Figure 6.2 relates to the production of a final report and full documentation for the complete software system. Obviously the scope of the outputs from this work item would effectively be set by the range of other items that had been addressed during the course of the software development. Should a manual be produced it is envisaged that it would follow the modular structure of the software, such that the upgrade or modification of a module would be matched by the production of one or more replacement sections rather than eliciting the revision of the entire manual.

Work item - items marked with an asterisk are 'core items' (see text) - proposed software modules are shown in bold text	Anticipated effort (weeks)	Cost guide (£)	Scheduling (time in months)
* 1. Agreement on national standards for assessment methodologies and formats for reporting	5	8k	
* 2. Agreement on detailed specification of all constituent elements of proposed software	4	7k	
3. Agreement on the format for down-loaded data	5	8k	
* 4. **Development of module for the input of survey details**	1	7k	
5. Agreement on nature of habitat data to be recorded	4	7k	
* 6. **Development of module for the input of habitat information**	1	7k	
* 7. **Development of module for the input of raw fishery data**	1	7k	
* 8. **Development of module for initial analyses**	5	7k	
* 9. **Development of module for population estimation**	6	10k	
10. Initiate R&D to assess available methodologies for mark-recapture estimates and incorporate findings as appropriate	30	50k	
11. Agreement of form of growth models to apply to length-at-age data	4	7k	
* 12. **Development of module for the derivation of population parameters**	3	3k	
13. **Development of module for the interpretation of individuals' data - no inclusion of habitat data**	5	8k	
* 14. **Development of module for the interpretation of population data - no inclusion of habitat data**	8	14k	
15. **Development of module for survey planning**	10	17k	
* 16. **Development of module for report generation**	4	14k	
17. **Final reporting and documentation of complete system**	8	15k	

Figure 6.2 Scheduling and costs of development tasks

Note that no allowance is made, in either the anticipated effort or in the cost guide figures, for the installation of the software, for software training or for post-installation support. The required effort and potential costs for these additional items would be dependent on a range of factors, including:

- the number and range of parties responsible for the software production;
- the software platform adopted;
- the extent of software development (i.e. the range of modules that are developed at the initial stage).

Notwithstanding the above, it is important that items such as installation, training and support are considered in any software development initiatives that follow from these recommendations.

Additional R&D

In addition to the proposed R&D involved in the production of software, the need for further R&D in selected areas has been highlighted by this project. In particular, methodologies for the use of mark-recapture data should be reviewed, and the derivation and use of semi-quantitative estimates examined.

It is likely that, following from the reporting of Phase Two of the Agency R&D on the development of a river fisheries habitat inventory, modifications would need to be made to certain areas of the proposed software. Most obviously, the module involved in the input of habitat data may need to be revised, although the development of methods to incorporate habitat information into the analysis of population data (i.e. as part of the 'interpretation of population data' module) should be considered.

Also, links between the proposed fishery data archiving system (and its associated reporting functions) and the National Fisheries Classification Scheme (NFCS) should be examined. Not only should the reporting requirements of the NFCS be taken into account (in terms of the production and reporting of population data) but also the potential for having software that runs the NFCS interfacing directly with the proposed fishery data archiving system. As such, the NFCS would become an additional module under the proposed stock assessment software system.

Finally, the ability of the fisheries data archive system to interface with other databases should also be examined. In particular, possible links with an Agency GIS should be considered.

6.4 Links with other Agency R&D

The development of new software would not be an isolated proposition. There are several potential links between the proposed development of new fisheries software and current or proposed R&D initiatives. The most apposite of these is probably the development of new stock assessment software by Midlands region. However, close links would need to be maintained with other recent or current Agency R&D, including:

- work on stock assessment methodologies
- the development of the National Fisheries Classification Scheme
- the use of semi-quantitative methods
- the development of a river fisheries habitat inventory.

In addition, it is recommended that new R&D is initiated to assess the use of mark-recapture methodologies, in a similar way to that used for removal methods (Lacey *et al.*, 1992; Wyatt and Lacey, 1994).

6.4.1 Stock assessment methodologies

The findings of R&D Project N° 325 (Lacey *et al.*, 1992; Wyatt and Lacey, 1994) should be taken on board in relation to the development of new stock assessment software. In particular, the recommendation that the Carle and Strub Maximum Weighted Likelihood (MWL) methodology be adopted as the Agency's operational standard for population estimation by the removal method should be noted. Furthermore, the observation that the variance of the population estimate calculated by the Carle and Strub MWL method should be estimated by simulation rather than direct empirical calculation using standard formulae should also be taken into account.

Outputs from the same R&D Project supported the use of standard CPUE methods, but recommended further work on the derivation of estimates for the probability of capture.

Under R&D Project N° 325 no assessment was made of the suitability of alternative population estimation techniques based on mark recapture survey methodologies. This area remains one that should be addressed by a targeted R&D Project, the findings of which should be used to formulate stock assessment methodologies, and hence software specification.

6.4.2 National Fisheries Classification Scheme

The current National Fisheries Classification Scheme (NFCS) is being used throughout the Agency. However, as noted in several of the regional meetings held during the course of this study its use would be facilitated by the provision of processed data in a form that can be input straight into the software. For example, software should produce separate combined estimates for the population densities of rheophilic, limnophilic, predatory and minor species groupings.

Whether the NFCS continues to be used for the standard reporting of fishery performance in its current or a modified form (see Section 6.4.4, below) requirements for classification purposes should be used to drive the design of stock assessment software.

6.4.3 Semi-quantitative methods

The deployment of CPUE methodologies is currently addressed only by the FINS software package, although the (manual) application of CPUE methods may be more widespread. When applying CPUE methodologies, the estimate of probability of capture that is specified is

of critical importance, and (currently) there is little or no formal guidance available to help Agency fishery staff select appropriate values.

However, robust strategies for deriving suitable estimates for probability of capture in semi-quantitative survey scenarios are being developed as part of the R&D Project on 'Semi-quantitative methods' (R&D Project N° 7716). The specification of stock assessment software should pay close attention to the findings and recommendations of this R&D. In addition, as it may be possible to produce default values for the probability of capture that apply to semi-quantitative surveys from information on the physical nature of the watercourse, the methods developed in the second phase of the fisheries habitat inventory project (see below) are also likely to be of key importance.

6.4.4 Development of a river fisheries habitat inventory

Phase One of the R&D on the development of a river fisheries habitat inventory (Wyatt and Barnard, 1997) has laid the foundations for the development of an effective and integrated assessment protocol whereby fisheries data can be used in conjunction with information on the general reach suitability and the instream habitat quality to produce measures of fishery performance.

As such a methodology is likely to become central to the reporting of fisheries information, its requirements should be considered when the final functional specification for new software is drawn up.

6.5 Software development - interim measures

Whether regions continue to make use of the (sub-optimal) software that they currently possess, or adopt the new software from Midlands region, it would be prudent to address the implications of the subsequent implementation of new software developed through a national R&D initiative.

It is proposed that stock assessment software would be developed such that a standard survey protocol is not imposed upon the regions, but rather that the use of raw data which may have be generated by any one of a variety of means is facilitated. At the same time the proposed software would allow the analysis and reporting of fisheries data to be undertaken in a nationally consistent and standardised manner.

It is therefore not envisaged that regions would need to reassess their operational practices regarding surveying protocol. However, as it is intended that analysis and reporting should be standardised, it is suggested that ahead of the production of new software, standard analysis and reporting methodologies should be agreed upon and (as far as possible) adopted throughout the Agency. This would help to move the Agency to a position where fisheries information is compatible across the regions, and would achieve this independently to the production of the proposed new software. In this way, the benefits of national compatibility could be realised sooner, i.e. ahead of the software production.

In some cases there are definitive procedures that should be adopted. For example, the analysis of removal (depletion) data should be undertaken using the Carle and Strub MWL methodology, with variances estimated by simulation. Storage of raw data should, where possible, be undertaken such that as little information is lost as possible.

In addition, it is suggested that the region's opinions as to their fishery reporting requirements (see Section 3.1.3) should be used by the Agency as a basis for developing a standard reporting protocol. For example, a range of standard report 'types' (c.f. Southern regions tiered reporting structure) should be agreed upon. It should be recognised that the reporting protocols that are produced may need to be reviewed, subsequent to the completion of current or proposed fishery R&D initiatives, with regard to the integration of new or modified reporting requirements.

7. REFERENCES

BAILEY, N.T.J. (1951) On estimating the size of mobile populations from capture-recapture data, *Biometrika*, **38**, 293-306.

BASSON, M., ROSENBERG, A.A. and BEDDINGTON, J.R. (1988) The accuracy and reliability of two new methods for estimating growth parameters from length-frequency data, *Journal du Conseil internationale pour l'exploration de la Mer*, **44**, 277-285.

BEVERTON, R.J.H. and HOLT, S.J. (1956) A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling, *Rapports et proces-verbaux du Conseil internationale pour l'exploration de la Mer*, **17A**, 1-153.

BOHLIN, T., HAMRIN, S., HEGGBERGET, T.G. RASMUSSEN, G. and SALTVEIT, S.J. (1989) Electrofishing - theory and practice with special emphasis on salmonids, *Hydrobiologia*, **173**, 9-43.

BOLGER, T. and CONNOLLY, P.L. (1989) The selection of suitable indices for the measurement and analysis of fish condition, *Journal of Fish Biology*, **34**, 171-182.

CARLE, F.L. and STRUB, M.R. (1978) A new method for estimating population size from removal data, *Biometrics*, **34**, 621-630.

CLARKE, R. (1992) *REMOVE - a program for fish population estimation.* IFE internal Publication.

CONNOLLY, P.J. (1996) *Resident cutthroat trout in the Central Coast Range of Oregon: logging effects, habitat associations and sampling protocols.* Unpublished doctoral thesis. Oregon State University: Corvallis, OR.

ENVIRONMENT AGENCY (1995) *Guidelines for computer human interaction (first draft).* Environment Agency Corporate Information Systems, Westbury-on-Trym. Document Reference N° IS/DS02/00001.01.0.

FULTON, T.W. (1911) *The sovereignty of the sea.* London: Edinburgh.

HIGGINS, P.J. (1985) An interactive computer program for population estimation using the Zippin method, *Aquaculture and Fisheries Management*, **16 (1)**, 287-295.

JOLLY, G.M. (1965) Explicit estimates from capture-recapture data with both death and immigration; stochastic model, *Biometrika*, **52**, 225 247.

KWAK, T.J. (1992) Modular microcomputer software to estimate fish population parameters, production rates and associated variance, *Ecology of Freshwater Fish*, **1(1)**, 73-75.

LACEY, R.F., WYATT, R.J., van DIJK, P.A.H. and HAY, S.V. (1992) *An evaluation of removal and catch-effort sampling methods for river fisheries.* Interim report to NRA N°. 325/5/A; WRc.

LINCOLN, A.G. (1930) *The Lincoln Index for biological applications.* Oxford Publications.

MILLS, C.A. and MANN, R.H.K. (1985) Environmentally-induced fluctuations in year-class strength and their implications for management, *Journal of Fish Biology*, **27 (supplement A)**, 209-226.

PAULY, D. (1987) A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic vertebrates. In: D.PAULY and G.R.MORGAN, (eds.) *Length-based methods in fisheries research*, pp7-34. ICLARM: Manila, Philippines and KISR: Safat, Kuwait. 468p.

PETERSEN, C.G.J. (1896) The yearly immigration of young plaice into the Limfjord from the German Sea, *Report of the Danish Biological Station*, **6**, 1-48.

POWELL, D.G. (1979) Estimation of mortality and growth parameters from the length-frequency data in the catch, *Rapports et proces-verbaux du Conseil internationale pour l'exploration de la Mer*, **175**, 167-169.

RICKER, W.E. (1975) Computation and interpretation of biological statistics of fish populations, *Bulletin of the Fisheries Research Board of Canada*, **191**, 382p.

SAILA, S.B., RECKSIEK, C.W. and PRAGER, M.H. (1988) *Basic fishery science programs - a compilation of microcomputer programs and manual of operation. Developments in Aquaculture and Fisheries Science*, **18**; Elsevier Science Publishers B.V., Amsterdam.

SCHNABEL, Z.E. (1938) The estimation of the total fish population of a lake, *American Mathematical Monographs*, **45**, 348-352.

SEBER, G.A.F. (1965) A note on the multiple-recapture census, *Biometrika*, **52**, 249-259.

SEBER, G.A.F. (1982) *The estimation of animal abundance and related parameters.* Charles Griffin: London. 645p.

SEBER, G.A.F. (1973) *The estimation of animal abundance.* Griffen: London.

SHEPHERD, J.G. (1987) A weakly parametric method for the analysis of length composition data. In: D.PAULY and G.R.MORGAN, (eds.) *Length-based methods in fisheries research*, pp113-120. ICLARM: Manila, Philippines and KISR: Safat, Kuwait. 468p.

WETHERALL, J.A., POLOVINA, J.J. and RALSTON, S. (1987) Estimating growth and mortality in steady-state fish stocks from length-frequency data. In: D.PAULY and G.R.MORGAN, (eds.) *Length-based methods in fisheries research*, pp53-74. ICLARM: Manila, Philippines and KISR: Safat, Kuwait. 468p.

WYATT, R. and BARNARD, S. (1977) *River fisheries habitat inventory (Phase 1).* Draft technical report to the Environment Agency; WRc.

WYATT, R.J. and LACEY, R.F. (1994) *Guidance notes on the design and analysis of river fishery surveys.* Report to NRA N°.292; WRc.

APPENDIX A QUESTIONNAIRE USED TO COLLECT INFORMATION ON FISHERIES SOFTWARE FROM WITHIN THE ENVIRONMENT AGENCY

Requirements and use of stock assessment software within the Agency's Fisheries Function

Introduction

This questionnaire is split into two parts. Part One is primarily designed to provide a means by which the software requirements of the Fishery Function can be recorded, whilst Part Two is designed to record detailed information on the software that is currently in use.

In Part One, information is requested on:

- the perceived need for software (i.e. those applications for which the fisheries staff in the Region require software);
- the requirements of that software (i.e. the particular characteristics or features of software that you would, ideally, wish to have available for use for a particular application);
- the names of those pieces software currently, or recently, used within the Region, together with an indication of the general availability of suitable software.

This section of the questionnaire provides a means for you to identify the nature of the software that you need to have access to. It should be stressed that this section relates to what you and your staff require, and not simply how you view what is currently available.

In Part Two, information is sought on the use of specific pieces of software, including (for example) those features that are used; features that are not used; the ease of use; quality of outputs; etc.).

It is intended that this questionnaire should be completed by, or with the assistance of, a member of Agency staff who is familiar with the needs of the Fishery Function across the Region. It possible that, in order to obtain the requisite information (especially for completion of the copies of Part Two), you will need to disseminate copies of sections from this questionnaire to other staff within the Region, and to collate their responses.

Return of questionnaires

Please return the completed Part One of this questionnaire, together with completed copies of Part Two, to Steve Barnard at the address given below. Please ensure that all completed questionnaires are returned by 31 July, 1996. If you have any queries regarding the completion of this questionnaire, please contact Steve Barnard at WRc (tel. 01 491 571 531).

Return address:

FAO - Steve Barnard,
WRc plc.,
Henley Road,
Medmenham,
Marlow,
Bucks., SL7 2HD.

Part One - Software requirements

A1-General information

Name of respondent .....

Area .....

Region .....

A2 - The Regional software needs

Please identify your Regional requirements for stock assessment software (i.e. those applications that you feel need to be supported or addressed by software). This listing should not be restricted to the range of software that is currently available, but should effectively be a 'wish-list' covering the full range of software that would, ideally, be available to you and/or your staff.

Please record these requirements by assigning each to one of the following four categories (the nature of which are discussed further, see below):

- analysis of individual-based data [enter information into Table 1];
- analysis of within-site (i.e. single visit, site-specific) population-based data [enter information into Table 2];
- analysis of whole fishery data (several sites and/or several occasions) [enter information into Table 3];
- survey design / planning [enter information into Table 4].

Enter details into the first column of either Table 1, 2, 3 or 4 (as appropriate) - where necessary, making copies of the tables to increase the space that is available for your responses.

Analysis of data to derive individual-based information (Table 1)

This category would include all applications concerned with the production of information relating to individual fish (e.g. ages, condition factors, growth rates, weights, etc.). Note that no distinction is made between whether the raw data used in such applications is derived from either a single survey at a single site, or from several surveys (i.e. on more than one occasion and/or at more than one site).

Examples of software functions from within this category might include:

- length-frequency analysis to facilitate the identification of discrete cohorts;
- length-at-age estimation through the use of back-calculation techniques applied to scale sample data;
- estimation of (individual) growth rates;
- estimation of length-weight relationships.

Derivation of within-site population-based data (Table 2)

This category of potential applications would include all functions relating to the overall population at a single site as recorded on a single occasion. This category would include, for example, the estimation of both population abundance and population biomass.

As indicated above, within-site applications would, by definition, utilise data from only a single site. Your survey methods (which may include netting; electric fishing; or hydroacoustic survey) are likely to reflect the population estimation techniques that you employ (e.g. the removal method or mark-recapture techniques).

Within this category, required software functions might include:

- population estimation by fully quantitative removal methods;
- population estimation from semi-quantitative data (e.g. derived from single pass electric fishing);
- population estimation from data obtained using mark-recapture techniques.

Derivation of whole fishery population information (Table 3)

In addition to using previously derived within-site population parameters from several discrete sites, the production of whole fishery population statistics (from several sites and/or several occasions) may also make use of data obtained by trapping; automatic counters; or catch statistics (e.g. angler census techniques and logbook schemes; match returns; licence returns).

Applications for which software is required may include:

- estimates of the average biomass or mean abundance of a given species within a river fishery;
- the classification of river reaches (covering several survey sites) using the National Fishery Classification Scheme;
- estimation of salmonid migration run-size.

Survey design / planning (Table 4)

This final category would include any software which is used to help in the design or planning of proposed fishery surveys. Such software may be concerned with either the spatial or the temporal requirements of a proposed survey programme and could, for example, be used to:

- indicate how many sites should be surveyed along a reach of river in order to obtain an overall fishery population estimate with a known degree of precision; or to
- indicate the desired periodicity of a rolling programme.

Note that any software requirements concerned with data archiving (i.e. the storage and retrieval of site- or reach-specific information) should be classified under either the 'within-site' or the 'overall fishery' option, as appropriate.

A3 - Your requirements of software for use in specific applications

For each of the applications you have identified in Section A2, list the characteristics or features which you feel are required in a software package in order for it to adequately fulfil your requirements. In particular, consider your requirements in terms of data input, data analysis and results output.

Enter this information into the second column of Table 1, 2, 3 or 4, as appropriate.

For example, if one of your requirements is:

- 'analysis of growth rates';

then you may consider the following characteristics to be important:

- 'software should allow *either* known length-at-age data *or* information on fish length and scale dimensions (e.g. radii of the scale annuli) to be used as input values';
- 'software should permit data for several species to be input at the same time, whilst allowing the results to be accessed independently, i.e. on a species-by-species basis';
- 'software should allow a choice of the growth curve which is fitted to the raw data';
- 'software should allow the production of graphs showing raw data together with fitted relationship and confidence limits';
- 'software should allow graphical outputs to be saved as Lotus 1-2-3 files, to facilitate subsequent inclusion in reports'.

You may also have specific requirements for the data handling capacity of software. For example, you may require software that is capable of taking input data (e.g. lengths and weights) for up to a maximum of (say) 1000 fish at a time.

A4 - Use and availability of specific software

For each of the requirements that you have listed in Tables 1-4, identify any software that is currently used, or has recently (e.g. within the last five years) been used to address your needs.

In addition, indicate the current availability of appropriate software within the Region. Rate this availability on a scale from '1' (software available within the Region but only on a very limited basis) through to '3' (software freely available to all, or most, fisheries staff). Where no software is available to fulfil one of the particular requirements that you have identified, rate the availability as '0'.

Table 1

Requirements for stock assessment software for analysis of data to derive individual-based information

A2 - Specific requirement / application:	A3 - Necessary characteristics or features of software:	A4 - Software used and general availability availability (0 to 3)
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Table 2
Requirements for stock assessment software for derivation of within-site population information

A2 - Specific requirement / application:	A3 - Necessary characteristics or features of software:	A4 - Software used and general availability availability (0 to 3)
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Table 3

Requirements for stock assessment software for derivation of whole fishery population information

A2 - Specific requirement / application:	A3 - Necessary characteristics or features of software:	A4 - Software used and general availability availability (0 to 3)
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....
....	....	....

Table 4
Requirements for stock assessment software for survey design / planning

A2 - Specific requirement / application:	A3 - Necessary characteristics or features of software:	A4 - Software used and general availability availability (0 to 3)
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Part Two - Software use

Note that a separate copy of this questionnaire should be completed for each piece of software currently, or recently, used within the Region.

Section A - General information

Name of respondent (if different from Part One).....

Area.....

Region.....

Section B - Software use within the Region

B1 - Software identification

Please supply:

- the software title:.....
- and version number: .....

B2 - Software applications

Indicate the category of applications for which the software is used (note that, in some instances, you may need to tick more than one box).

The software is used for:

- analysis of individual-based data ..... ❑
- analysis of within-site (i.e. single visit, site-specific) population-based data ..... ❑
- analysis of whole fishery or site-specific temporal population-based data..... ❑
- survey design / planning ..... ❑

Please provide more detail on the applications addressed by this particular piece of software by listing the specific functions that are employed (e.g. applying the 'Zippin' approach to data obtained by the removal method; estimates of mean length-at-age; etc.).

If a particular piece of software is no longer used, write 'Not used' in the 'functions performed' section, but continue to complete Sections B3 to B7 to the best of your ability. In Section B7 you will have the opportunity to indicate why the software is not used.

Functions employed

1. .....
2. .....
3. .....
4. .....
5. .....
6. .....
7. .....
8. .....
9. .....
10. .....

B3 - Extent of use

Indicate the extent of your use of this software by listing those features or capabilities of the software that are utilised, and those which are not used and are, effectively, redundant.

For example, software that calculates population estimates from removal data may allow the use of several statistical methods, and may permit data to be entered either direct from the keyboard or by means of an import file. In practice, however, you may use only the Carle and Strub statistical approach and always enter the data direct from the keyboard.

Features / methodologies used

1. .....
2. .....
3. .....
4. .....
5. .....
6. .....
7. .....
8. .....
9. .....
10. .....
11. .....
12. .....

Features or methodologies *not* used

1. .....
2. .....
3. .....
4. .....
5. .....
6. .....
7. .....
8. .....
9. .....
10. .....
11. .....
12. .....

B4 - Source of software

What was the original source of the software?

- borrowed / copied from within same Region ..... ❑
- borrowed / copied from outside of Region but from within the Agency .... ❑
- purchased / copied from outside of the Agency (please give details) ..... ❑

  .....

  .....

B5 - Subsequent modifications to the software

Has the software been modified from its original form ?

- no, the software has remained unmodified..... ❑
- yes, the software has been modified
  from its original form (please give details below) ..... ❑

  .....

  .....

B6 - Ease of use / quality of output

Ease of use

Is the software highly complex, requiring an extended period of familiarisation before a user is able to operate it with a reasonable level of competence - or is the implementation of the software straightforward and intuitive?

Please tick one of the following boxes:

- complex software requiring an extended period of familiarisation..... ❑
- moderately complex software; reasonable demands regarding the level of familiarisation that is required ..... ❑
- straightforward and intuitive software..... ❑

Quality of outputs

This section deals with both the clarity of the output (i.e. how easy the output is to interpret) and the usefulness of the output (for example, whether the output can be readily used for reporting or presentation purposes).

For example, the outputs may be complex and difficult to interpret - or their interpretation may be relatively straightforward.

Equally, the outputs may be in a form that cannot, subsequently, be readily used - or they are in a form that lends itself to other purposes (e.g. graphics that can be easily transferred to other packages for use in reporting).

Please tick the following boxes as appropriate:

Interpretability:

- complex outputs; difficult to interpret..... ❑
- moderately easy to interpret..... ❑
- straightforward, easily interpreted outputs ..... ❑

Usefulness for other applications (e.g. reporting):

- outputs cannot be used directly for other, subsequent, applications ..... ❑
- Moderately useful outputs ..... ❑
- useful outputs, easily used within other applications ..... ❑

B7 - Further comments

Finally, please use the space below if you have any further comments regarding the use of this piece of software.

.....

.....

.....

.....

.....

.....

.....

.....

Thank you for your co-operation. Please return all completed questionnaires to your Regional contact point for collation and return to WRc.

Technical specification of software for use within the Agency's Fisheries Function

Introduction

This information recorded on this questionnaire is intended to enable WRc to produce a technical appraisal of the fisheries software that has either:

- been produced within your region of the EA; or
- although originally produced outside of your region, has subsequently been modified by staff within your region.

It is intended that this questionnaire should be completed with the assistance of those Agency staff who are familiar with the development, modification or implementation of fisheries software within the region. It is possible that, in order to obtain the requisite information, you will need to disseminate copies of sections from this questionnaire to other staff within the region and collate their responses.

Only one copy of Section A should be completed: this should be based on information collated from throughout the region.

However, a separate copy of Section B (and, where appropriate, copies of one or more of Sections C; D; and E) should be completed for each piece of fisheries software that has been produced or modified by your region.

Please return all completed sections of this questionnaire to Steve Barnard, at the address given below, by 31 July, 1996.

If you have any questions regarding the completion of this questionnaire, please contact Steve Barnard at WRc (tel. 01 491 571 531).

Return address: FAO - Steve Barnard,
WRc plc.,
Henley Road,
Medmenham,
Marlow,
Bucks., SL7 2HD.

Section A : Nature of software used in your region

A1-General information

Name of regional contact .....
Area .....
Region .....

A2-Software used within the region

Use Table 1 (page 4 of this section) to list all of the fishery software packages that are currently used by staff in the region, or which have been used in the recent past (e.g. in the last five years or so). Make copies of the table if there is insufficient space for full completion. For each piece of software, indicate:

- the software title;
- the software 'source' (see below);
- the nature of the applications that are performed by the software (again, see below).

Software source

The 'source' of the software should be categorised as follows:

- software originally developed within your region (including instances where your region has been responsible for the subsequent modification of the software);
- software originally developed by another region of the Agency, but subsequently modified within your region;
- software originally developed by an organisation or individual external to the Agency, but subsequently modified within your region for use by the Agency;
- software, produced either by another region or by an organisation or individual external to the Agency, which is used in an unmodified form.

Software applications

For the purposes of this questionnaire, a distinction is made between different categories of fisheries software applications. Four broad categories are employed:

- analysis of individual-based data;
- analysis of within-site (i.e. single visit, site-specific) population-based data;
- analysis of whole fishery or site-specific temporal population-based data;
- survey design / planning.

The nature of each of these categories is detailed further, overleaf.

Analysis of data to derive individual-based information

This category would include all software applications concerned with the production of information relating to individual fish (e.g. ages, condition factors, growth rates, weights, etc.). Note that no distinction is made between whether the raw data used in such applications is derived from either a single survey at a single site, or from several surveys (i.e. on more than one occasion and/or at more than one site).

Examples of software functions from within this category might include:

- length-frequency analysis to facilitate the identification of discrete cohorts;
- length-at-age estimation through the use of back-calculation techniques applied to scale sample data;
- estimation of (individual) growth rates;
- estimation of length-weight relationships.

Derivation of within-site population data

This category of potential applications would include all functions relating to the overall population at a single site as recorded on a single occasion. This category would include, for example, the estimation of both population abundance and population biomass.

As indicated above, within-site applications would, by definition, utilise data from only a single site. Your survey methods (which may include netting; electric fishing; or hydroacoustic survey) are likely to reflect the population estimation techniques that you employ (e.g. the removal method or mark-recapture techniques).

Within this category, required software functions might include:

- population estimation by fully quantitative removal methods;
- population estimation from semi-quantitative data (e.g. derived from single pass electric fishing);
- population estimation from data obtained using mark-recapture techniques.

Derivation of whole fishery population information

In addition to using previously derived within-site population parameters from several discrete sites, the production of whole fishery population statistics (relating to several sites and/or to several separate sampling occasions) may also make use of data obtained by trapping; automatic counters; or catch statistics (e.g. angler census techniques and logbook schemes; match returns; licence returns).

Applications for which software is required may include:

- estimates of the average biomass, mean abundance, or survival rate of a given species within a river fishery;
- the classification of river reaches (covering several survey sites) using the National Fishery Classification Scheme;
- estimation of salmon or sea-trout adult run size.

Survey design / planning

This final category would include any software which is used to help in the design or planning of proposed fishery surveys. Such software may be concerned with either the spatial or the temporal requirements of a proposed survey programme and could, for example, be used to:

- indicate how many sites should be surveyed along a reach of river in order to obtain an overall fishery population estimate with a known degree of precision; or to
- indicate the desired periodicity of a rolling programme;
- number of individual fish to sample.

Software concerned solely with data archiving (i.e. the storage and retrieval of site- or reach-specific information) should be classified under either the 'within-site' or the 'overall fishery' option, as appropriate.

Software title	Source of software (tick only one box per software title)				Application (tick all applicable boxes)			
	Developed within your region	Modified within your region but *developed within another region*	Modified within your region but developed outside of Agency	In use, but neither developed nor modified by your region	Derivation of individual-based parameters	Derivation of within-site population parameters	Derivation of overall fishery population parameters	Survey planning and design
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐
----------------------	☐	☐	☐	☐	☐	☐	☐	☐

The remainder of this questionnaire relates solely to those pieces of software either originally developed, or subsequently modified, within your region (i.e. software in the first three categories in the above table).

As stated above, a separate copy of Section B of the questionnaire (including, where appropriate, copies of one or more of Sections C; D; E; or F) should be completed for each piece of fisheries software which has been produced or modified by your region.

However, please note that information regarding the use of unmodified software from external sources is of importance. At a later date, we will need to assess the suitability and applicability of all fisheries software which is currently available, regardless of its original source.

Section B - General information on regional software products

This, and subsequent, sections should be completed only for software that was produced, or has been modified, within the region.

A separate copy of this section (and, as appropriate, one or more of sections C, D, E or F) should be completed for each piece of software.

Before returning to the regional contact, ensure that Section B is attached firmly to completed copies of sections C, D, E or F.

B1-General information

Name of respondent (if not regional contact) .....
Area .....
Region .....

Software title and version number .....
.....

What was the date of the original development / latest modification of this software? .....

B2-Technical information - hardware and software requirements

What is the required computer base ? Mainframe .....❑
or PC .....❑

If PC-based, please indicate the minimum specification that is required:

Processor (e.g. 386; 486) .....
RAM (Mb) .....
Monitor (e.g. monochrome only, colour VGA, etc.) .....
Hard disk capacity (Mb) .....
(other) .....
(other) .....

Please list any operating system (e.g. DOS; OS/2; VMS) or software 'platform' requirements (e.g. Windows; Lotus 1-2-3; Excel; SuperCALC; dBASE) necessary for the implementation of the software.

Alternatively, if the software is a 'stand-alone' package which does not require a commercial software package to support its operation, please state this clearly, and indicate which language the programme has been written in (e.g. Fortran; Basic; Visual Basic).

Where appropriate, also include the version number of the required operating system, software 'platform' and / or programming language.

Operating system .....

Software platform - if applicable .....

Programme language .....

	Yes	No
Are the original 'installation' disks still available?	❑	❑

We do not require these disks but, should it be decided that the software should be *modified, access to the original programme disks would be advantageous.*

If available, please supply a copy of the installation manual.

B3-Nature of software applications

Indicate the nature of those applications with which the software is concerned (if applicable, please tick more than one category).

- analysis of individual-based data..... ❑
  (complete section C)
- analysis of within-site (i.e. single visit, site-specific) population data..... ❑
  (complete section D)
- analysis of whole fishery (or site-specific, temporal) population data..... ❑
  (complete section E)
- survey design / planning..... ❑
  (complete section F)

B4-General data requirements

Site identification data

	Required	Optional	No facility
Site name	☐	☐	☐
NGR	☐	☐	☐
Survey date	☐	☐	☐

Physical site information

	Required	Optional	No facility
Length	☐	☐	☐
Width	☐	☐	☐
Area	☐	☐	☐
Depth	☐	☐	☐
Other (please state)	☐	☐	
Other (please state)	☐	☐	

Site habitat data

	Required	Optional	No facility
Instream data	☐	☐	☐
Riparian data	☐	☐	☐
Catchment data	☐	☐	☐

... continued overleaf

For each of the classes of habitat data that you have indicated as being required or optional, please supply a list of the type data that is entered into the software (e.g. flow types; substrate types; degree of shading by riparian vegetation; land-use of catchment; area of catchment; etc.).

.....

.....

.....

.....

.....

.....

.....

.....

NB: further information on data requirements, specific to the software application, is requested in Sections C, D, E and F.

B5-Data input

How is data input undertaken?.....directly, via the keyboard .....❑

indirectly, via an input file .....❑

either, i.e. there is the option of keyboard or input file .....❑

If possible, please supply a hardcopy (printout) of the data input screen(s).

Where possible, please supply print-outs of examples of data input files.

B6-Similarities in input between applications

Where the software supports more than one type of application, are the options discussed in Sections B4 and B5 available for all of the applications, or only for a selection?

All data input options available to all of the software's applications ..... ❑

Input options not common to all of the software's applications ..... ❑

If the input options are not common to all of the software's applications, but apply only to a selection, then please give details below.

.....

.....

.....

.....

.....

.....

B7-Help and guidance to software operation

Is there a user manual available to accompany the software? ..... Yes ..... No

❑ ❑

If 'yes', please enclose a copy of the manual with your questionnaire return.

Does the software have any associated 'help' or 'read-me' text files which the user is able to either read or print-off before the software is used?

	Yes	No
Help files	❑	❑
'Read-me' files	❑	❑

If 'yes', please enclose print-outs of the help file with your questionnaire return.

	Yes	No
Does the software have an on-line 'help' facility?	❑	❑

Section B is now complete for this piece of software.

Please continue and complete Sections C, D, E and/or F, as appropriate (i.e. as identified in question B3, above).

Thank you for your co-operation.

Section C : Software for the analysis of individual-based data

We may need to contact the person who was responsible for the completion of this section for further details on specific aspects of this software package.

To facilitate this, please give your name here: .....

C1-Applications

Please list the applications supported by the software

These may include, for example, the derivation of length-at-age or weight-length relationships, the estimation of growth rates, or the identification of discrete cohorts through the analysis of length-frequency data.

.....
.....
.....
.....
.....
.....
.....
.....

C2-Data requirements

Please indicate whether the inputs listed below are required or are optional, or whether there is no facility to input such data.

Data relating to individual fish

	Required	Optional	No facility
Species	☐	☐	☐
Length	☐	☐	☐
Weight	☐	☐	☐
Age	☐	☐	☐
Sex	☐	☐	☐
Batch identification mark(s)	☐	☐	☐
Unique identification mark(s) (e.g. tag numbers)	☐	☐	☐
Other (please state)	☐	☐	
Other (please state)	☐	☐	

C3-Basis of operation

Please indicate the statistical methods that the software employs for each of the applications that you have identified in C1 (above). Where applicable, include a literature reference for each of the (statistical) methods that are used.

Application:	Method employed:	Reference:
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Assumptions, approximations and selection from alternative methods

For each method listed above, please detail any assumptions and approximations that are made (e.g. regarding error structures; random nature of sampling; etc.).

In certain cases, the software may have a choice of potential methods (e.g. the use of either back-calculated length-age data for a single cohort of fish or observed length-at-age data for a range of cohorts to calculate growth rates). Where there is such a choice please describe the basis of any subsequent selection procedure (e.g. user-defined selection, automatic selection based on subsequent fit of models, etc.).

Method:	Assumptions, approximations or selection mechanism:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

C4-Detail of operation

Limits of operation :

Is the software able to cope with missing data?..... Yes ..... No
❑ ❑

If 'yes', how is this accomplished? .....
.....
.....
.....
.....

What other limits apply to the operation of the software? Please list these below.

For example, when analysing a weight-length relationship for a species, there may be a maximum number of individuals that the software is capable of handling.

.....
.....
.....
.....
.....

C5-Data output and displays

How are results shown?

on-screen / with graphical results output ..... ❑
on-screen / with numerical results output ..... ❑
hard-copy / with graphical results output ..... ❑
hard-copy / with numerical results output ..... ❑

What is the potential for linking to other software products (i.e. can the software outputs be used as an input for other software packages)? Note that a second software package may either accept data directly from those output files produced by the original package or may require the manual entry of the results derived from the first package. Please give details below:

.....

.....

.....

.....

Can data be exported in an electronic format? ..... Yes ..... No

❑ ❑

If 'yes', what formats can be used?
(e.g. .CSV; .WK*; .DBF; .TXT; .XLS; .CAL; .MTB; .PRN; .PLT; etc.)

.....

.....

Where you have indicated that fishery data is required or optional, please list in the table below those data types that can be input (e.g. 'numbers of fish per run for each site on the river reach'; 'population estimate for each site'; 'total biomass of fish >10m fork length for each site'; etc.).

Required inputs:	Optional inputs:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

E3-Basis of operation

Please indicate the statistical methods that the software employs for each of the applications that you have identified in E1 (above). Where applicable, include a literature reference for each of the (statistical) methods that are used.

Application:	Method employed:	Reference:
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Section D : Software for the analysis of within-site (i.e. single visit, site-specific) population data

We may need to contact the person who was responsible for the completion of this section for further details on specific aspects of this software package.

To facilitate this, please give your name here: .....

D1-Applications

Please list the applications supported by the software

These may include, for example, population estimation from electric fishing or netting data, biomass estimation.

.....

.....

.....

.....

.....

.....

.....

.....

D2-Data requirements

Please indicate whether the inputs listed below are required or are optional, or whether there is no facility to input such data.

Data relating to individual fish

	Required	Optional	No facility
Species	☐	☐	☐
Length	☐	☐	☐
Weight	☐	☐	☐
Age	☐	☐	☐
Sex	☐	☐	☐
Batch identification mark(s)	☐	☐	☐
Other (please state)	☐	☐	
Other (please state)	☐	☐	

Derived relationships

	Required	Optional	No facility
Length-weight models	❑	❑	❑
Length-at-age (growth) models	❑	❑	❑
Other (please state)	❑	❑	
Other (please state)	❑	❑	

Data relating to fish sampling/capture:

	Required	Optional	No facility
Numbers caught per run	❑	❑	❑
Run identifier	❑	❑	❑
Trap identification	❑	❑	❑

D3-Basis of operation

Please indicate the statistical methods that the software employs for each of the applications that you have identified in D1 (above). Where applicable, include a literature reference for each of the (statistical) methods that are used.

Application:	Method employed:	Reference:
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Assumptions, approximations and selection from alternative methods

For each method listed above, please detail (in the table overleaf) any assumptions and approximations that are made (e.g. regarding error structures; random nature of sampling; etc.).

In certain cases, the software may have a choice of potential methods (e.g. the use of either 'Seber-LeCren' or 'Zippin' for deriving population estimates). Where there is such a choice please describe the basis of any subsequent selection procedure (e.g. user-defined selection, automatic selection based on magnitude of estimated variance, etc.).

Assumptions, approximations and selection from alternative methods

Method:	Assumptions, approximations or selection mechanism:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

D4-Detail of operation

Resolution

Does the software simply operate at whatever level of detail is inherent in the original input data or does it allow the user to select an appropriate level of resolution?

For example, software which is used to generate population estimates, in addition to producing estimates for individual age-classes, may permit the production of estimates at a lower resolution (e.g. for all ages of fish combined).

Application	Output detail only at level of input data	Option of lower resolution
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑

Limits of operation

Is the software able to cope with missing data?..... Yes ..... No

❑ ❑

If 'yes', how is this accomplished? .....

.....

.....

.....

.....

What other limits apply to the operation of the software? Please list these below.

For example, when calculating population estimates based on the removal technique, there may be a maximum number of removals which the software is capable of handling.

.....
.....
.....
.....
.....

D5-Data output and displays

How are results shown?

on-screen / with graphical results output ..... ☐
on-screen / with numerical results output ..... ☐
hard-copy / with graphical results output ..... ☐
hard-copy / with numerical results output ..... ☐

What is the potential for linking to other software products (i.e. can the software outputs be used as an input for other software packages)? Note that a second software package may either accept data directly from those output files produced by the original package or may require the manual entry of the results derived from the first package. Please give details below:

.....
.....
.....
.....

Can data be exported in an electronic format? ..... Yes ..... No

❑ ❑

If 'yes', what formats can be used?
(e.g. .CSV; .WK*; .DBF; .TXT; .XLS; .CAL; .MTB; .PRN; .PLT; etc.)

.....

.....

Does the software have facilities for
archiving information (i.e. data storage & retrieval facilities)? ..... Yes ..... No

❑ ❑

Does the software have data logging options for field use? ..... Yes ..... No

❑ ❑

Where possible, please supply print-outs of examples of all of the outputs that are produced by the software, including printed reports (tables, graphs, etc.) and output files.

Section D is now complete for this piece of software.

Please attach this section to the completed Section B, together with completed copies of Sections C, E and/or F (as appropriate) and then return the full questionnaire to your regional contact for collation.

Thank you for your co-operation.

Section E : Software for the analysis of whole fishery population data (i.e. data from several sites or several visits)

We may need to contact the person who was responsible for the completion of this section for further details on specific aspects of this software package.

To facilitate this, please give your name here: .....

E1-Applications

Please list the applications supported by the software

These applications may include, for example, population estimates for river reaches based on data from a series of discrete sites.

.....

.....

.....

.....

.....

.....

.....

.....

E2-Data requirements

Nature of input data

Is the fishery data required in its raw form (e.g. the numbers of fish caught per electric-fishing run at each of a series of sites) or in a summary, derived form (e.g. as estimates of population size - with associated variances - for each of a series of sites)?

Please indicate whether such fishery information is required or is optional, or whether there is no facility to input such data.

	Required	Optional	No facility
Raw data	❑	❑	❑
Derived data	❑	❑	❑

Where you have indicated that fishery data is required or optional, please list in the table below those data types that can be input (e.g. 'numbers of fish per run for each site on the river reach'; 'population estimate for each site'; 'total biomass of fish >10m fork length for each site'; etc.).

Required inputs:	Optional inputs:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

E3-Basis of operation

Please indicate the statistical methods that the software employs for each of the applications that you have identified in E1 (above). Where applicable, include a literature reference for each of the (statistical) methods that are used.

Application:	Method employed:	Reference:
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Assumptions, approximations and selection from alternative methods

For each method listed above, please detail any assumptions and approximations that are made (e.g. regarding error structures; random nature of sampling; etc.).

In certain cases, the software may have a choice of potential methods. Where there is such a choice, please describe the basis of any subsequent selection procedure (e.g. user-defined selection, automatic selection based on magnitude of estimated variance, etc.).

Method:	Assumptions, approximations or selection mechanism:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

E4-Detail of operation

Resolution

Does the software simply operate at whatever level of detail is inherent in the original input data or does it allow the user to select an appropriate level of resolution?

For example, software which is used to generate species-specific estimates of fish biomass for a reach of river may also permit the production of a biomass estimate for the reach at a lower resolution (e.g. for all species of fish combined).

Application	Output detail only at level of input data	Option of lower resolution
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑
.....	❑	❑

Limits of operation

Is there a maximum number of sites that the software is capable of handling? ..... Yes ❑ ..... No ❑

If 'yes', what is this maximum? .....

Is the software able to cope with missing data? ..... Yes ..... No
❑ ❑

If 'yes', how is this accomplished? .....
.....
.....
.....
.....

What other limits apply to the operation of the software? Please list these below.

For example, there may be a maximum number of separate species which the software is capable of handling at any one time.

.....
.....
.....
.....
.....

E5-Data output and displays

How are results shown?
on-screen / with graphical results output ..... ❑
on-screen / with numerical results output ..... ❑
hard-copy / with graphical results output ..... ❑
hard-copy / with numerical results output ..... ❑

What is the potential for linking to other software products (i.e. can the software outputs be used as an input for other software packages)? Note that a second software package may either accept data directly from those output files produced by the original package or may require the manual entry of the results derived from the first package. Please give details below:

.....
.....
.....
.....

Can data be exported in an electronic format? ..... Yes ..... No

❑ ❑

If 'yes', what formats can be used?
(e.g. .CSV; .WK*; .DBF; .TXT; .XLS; .CAL; .MTB; .PRN; .PLT; etc.)

.....

.....

Does the software have facilities for
archiving information (i.e. data storage & retrieval facilities)? ..... Yes ..... No

❑ ❑

Does the software have data logging options for field use? ..... Yes ..... No

❑ ❑

Where possible, please supply print-outs of examples of all of the outputs that are produced by the software, including printed reports (tables, graphs, etc.) and output files.

Section E is now complete for this piece of software.

Please attach this section to the completed Section B, together with completed copies of Sections C, D and/or F (as appropriate) and then return the full questionnaire to your regional contact for collation.

Thank you for your co-operation.

Section F : Software for survey design or planning

We may need to contact the person who was responsible for the completion of this section for further details on specific aspects of this software package.

To facilitate this, please give your name here: .....

F1-Applications

Please list the applications supported by the software. These may include, for example, 'identifying the minimum number of survey sites required to derive a population estimate (with a known level of precision) for a given river reach'.

.....
.....
.....
.....
.....
.....
.....
.....

F2-Data requirements

Fishery data

Please list the types of fishery data (i.e. the range of data relating to fish stocks) that can be input to the design software. Such data types may include, for example, site- and species-specific population estimates; estimates of the temporal variability of fish abundance; etc. Please categorise your responses according to whether each input is required or is optional.

Note that information on other data inputs (e.g. information relating to the site or reach which is under examination) is covered in previous sections.

Required inputs:	Optional inputs:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

F3-Basis of operation

Please indicate the statistical methods that the software employs for each of the applications that you have identified in F1 (above). Where applicable, include a literature reference for each of the (statistical) methods that are used.

Application:	Method employed:	Reference:
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....

Assumptions, approximations and selection from alternative methods

For each method listed above, please detail any assumptions and approximations that are made (e.g. regarding error structures; random nature of sampling; etc.).

In certain cases, the software may have a choice of potential methods. Where there is such a choice, please describe the basis of any subsequent selection procedure (e.g. user-defined selection, automatic selection based on magnitude of estimated variance, etc.).

Method:	Assumptions, approximations or selection mechanism:
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

F4-Detail of operation

Limits of operation

Is there a maximum number of sites
that the software is capable of handling data for? ..... Yes ..... No
❑ ❑

If 'yes', what is this maximum? .....

Is there a maximum number of years-worth of data
that the software is capable of handling? ..... Yes ..... No
❑ ❑

If 'yes', what is this maximum? .....

Where site- or reach-specific information (for example, an estimate of temporal variability in population size) is not available, is the
software able to supply default values? ..... Yes ..... No
❑ ❑

If 'yes', what data are supported in this way? Please list these data types below.

.....

.....

If there is missing data for which the software is unable to provide a default estimate, is the software still able to
produce a functional output? ..... Yes ..... No
❑ ❑

What other limits apply to the operation of the software? Please list these below.

For example, there may be a maximum number of separate species which the software is capable of handling at any one time.

.....
.....
.....
.....
.....

F5-Data output and displays

How are results shown?

on-screen / with graphical results output ..... ☐
on-screen / with numerical results output ..... ☐
hard-copy / with graphical results output ..... ☐
hard-copy / with numerical results output ..... ☐

What is the potential for linking to other software products (i.e. can the software outputs be used as an input for other software packages)? Note that a second software package may either accept data directly from those output files produced by the original package or may require the manual entry of the results derived from the first package. Please give details below:

.....
.....
.....
.....

Can data be exported in an electronic format? ..... Yes ☐ ..... No ☐

If 'yes', what formats can be used?
(e.g. .CSV; .WK*; .DBF; .TXT; .XLS; .CAL; .MTB; .PRN; .PLT; etc.)

.....
.....

Does the software have facilities for archiving information (i.e. data storage & retrieval facilities)? ..... Yes ..... No

❑ ❑

Does the software have data logging options for field use? ..... Yes ..... No

❑ ❑

Where possible, please supply print-outs of examples of all of the outputs that are produced by the software, including printed reports (tables, graphs, etc.) and output files.

Section F is now complete for this piece of software.

Please attach this section to the completed Section B, together with completed copies of Sections C, D and/or E (as appropriate) and then return the full questionnaire to your regional contact for collation.

Thank you for your co-operation.

APPENDIX B REGIONAL CONTACTS

The assistance of the following regional contacts from the Environment Agency is gratefully acknowledged.

Region	Contact	Date of regional meeting
Anglia	Robin Burrough	19 November, 1996
Midlands	Paul Lidgett	12 November, 1996
North East	Steve Chambers	13 November, 1996
North West	Miran Aprahamian	25 November, 1996
South West	David Bird	11 November, 1996
Southern	Chris Lee	23 January, 1997
Thames	Alan Butterworth	14 November, 1996
Welsh	Dave Mee	13 January, 1997

APPENDIX C QUESTIONNAIRE USED TO COLLECT INFORMATION ON FISHERIES SOFTWARE USAGE OUTSIDE OF THE ENVIRONMENT AGENCY

Fax Transmission from WRc plc
Henley Road, Medmenham, Marlow, Bucks, SL7 2HD, UK

Telephone: 01491 571531 International + 44 1491 571531
Fax: 01491 579094 International + 44 1491 579094

To (organisation):			
Fax No.:		**Date:**	3 January 1997
For attention of:		**No. of pages: (including this one)**	5
From:	Steve Barnard	**Telephone extension No:**	4161

Dear Sir

The Environment Agency has recently commissioned WRc to assess the requirements for, and availability of, computer software used in:

- the design of fishery surveys; and
- the subsequent analysis and interpretation of fisheries (stock assessment) data.

Information has been supplied by staff throughout the Agency's regions to allow an assessment to be made of the suitability of fisheries software currently used by Agency staff.

However, it will be important not to view the information that has been collated from the Agency completely in isolation, but rather to be in a position where the use of alternative fisheries software (developed or used by commercial organisations and academic institutes) may be recognised.

Accordingly I would be extremely grateful if you, or one of your colleagues, could set aside a few minutes to complete the attached questionnaire and to return it to me at your earliest convenience.

If you have any further questions, please do not hesitate to contact me.

Thank you in advance.

Yours faithfully

Steve Barnard

Ecological Assessment Group

For WRc Internal use only:

Priority:				
				X
	Within 1 hour	0830-1300	1300-1730	Overnight

Confirmation of transmission: X Internal Mail Immediate

MEDMENHAM

Introduction

The Environment Agency has commissioned WRc to assess the requirements for, and availability of, computer software used in both the design of freshwater fishery surveys and in the subsequent analysis and interpretation of fisheries (stock assessment) data. Information has been supplied by staff throughout the Agency's regions which will allow an assessment to be made of the suitability of fisheries software which is currently used by Agency staff.

It will be important, however, not to view the information that has been collated from the Agency in isolation, but rather to be in a position where the use of alternative software (such as that developed or used by commercial organisations and academic institutes) may be recognised.

Accordingly, I would be extremely grateful if you, or one of your colleagues, could set aside a few minutes to complete and return the following few questions.

Part 1: Respondent details

1. Name of respondent .....

   Organisation / institution .....

   Contact telephone number .....

Part 2: Design and planning of fishery surveys

2.a If you are planning a (freshwater) fishery survey programme, do you make use of any computer software to assist in the experimental design programme ?

   Yes ❑

   No ❑

   If 'no', proceed to Part 3.

   If 'yes', does this software simply take the form of a set of routines that you run within a 'standard' software package (e.g. a spreadsheet or a statistics package), or is it a piece of software that has been written specifically to perform this task ?

   Routines within standard software packages ..... ❑

   Software written specifically for the task..... ❑

3.b What stock-assessment methods do you employ ?

Please tick those methods that are used, and indicate their relative frequency of use, e.g. fully-quantitative catch-removal methods 20%;
semi-quantitative, single-run electric fishing 80%

Fully quantitative - catch removal (e.g. by electric fishing) ❑ .....

- mark/recapture ❑ .....

Semi-quantitative (e.g. single-run electric fishing) ❑ .....

Hydro-acoustic ❑ .....

Counter data ❑ .....

Angler census ❑ .....

Others (please specify) .....

.....

.....

3.c What population estimation methods do you use ?

For example, all of your population estimates may be fully quantitative and may use a single method (e.g. Carle and Strub), or you may have the option of choosing between two or more methods (such as Carle and Strub; Zippin; Seber and LeCren; etc.). Alternatively, all of your surveys may be semi-quantitative, with minimum estimates being converted to population estimates by multiplying up by a known efficiency factor (i.e. an estimate of the probability of capture).

Please list the range of methods that you use, irrespective of whether or not they are implemented by means of specially written software. Where applicable, indicate the factors that influence the choice between alternative methods. For example, you may only make use of a given method 'if there is better than 50% depletion between successive electric fishing runs' or 'if the data is derived from two runs'.

Method	Conditions for use - if applicable
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

If software has been produced specifically to perform this task, what is its name, and was it produced internally (i.e. within your organisation/institution) or externally ? If external, please indicate its original source.

Name of software.....

Internal..... ❑

External..... ❑

Source of software.....

Part 3: Analysis of stock-assessment data

3.a When you undertake the analysis of (freshwater) fishery stock-assessment data, do you use software that has been written specifically to undertake the task, or do you simply perform a series of routines within a 'standard' software package (e.g. a spreadsheet or a statistics package) ?

Software written specifically for the task..... ❑

Routines within standard software packages ..... ❑

If software has been produced specifically to perform this task, what is its name, and was it produced internally (i.e. within your organisation/institution) or externally ? If external, please indicate its original source.

Name of software.....

	Internally	Externally
Produced.....	❑	..... ❑

Source of software.....

3.d In the course of your research or routine monitoring programmes, approximately how many sites do you survey annually ?

Number of sites per year .....

3.e Finally, do you use computer software to archive your fishery stock-assessment data or do you rely on hard-copy archives ?

Computer software archiving system..... ❑

Hard-copy only..... ❑

Thank you for your help.

Please post or fax the completed questionnaire to:

Steve Barnard
WRc plc, Henley Road
Medmenham, Marlow
Bucks., SL7 2HD

Tel (01491) 571531
Fax (01491) 579094

APPENDIX D NON-AGENCY BODIES CONTACTED FOR INFORMATION ON SOFTWARE USAGE

- DANI - River Bush Salmon Station
  Gersham Kennedy
- Fisheries Conservation Board for Northern Ireland
  Bill Smith
- HIFI - University of Hull
  no response
- IFE - Eastern Rivers Laboratory
  Richard Mann
- IFE - River Laboratory
  Bill Beaumont
- IFE - Windermere Laboratory
  Ian Winfield
- Institute of Hydrology
  no response
- Powergen Freshwater Biology Group
  Sally McGuire
- Renewable Resources Assessment Group
  no response
- Salmon Research Agency - Co.Mayo
  Phil McGinnity
- SOAEFD - Freshwater Fisheries Laboratory - Pitlochry
  Ross Gardiner
- University of Aberdeen - Zoology Department
  Lindsay Laird
- University of Buckingham - Department of Life Sciences
  no response
- University of Liverpool - Dept. of Environmental and Evolutionary Biology
  Phillip Smith
- University of Stirling - Institute of Aquaculture
  no response
- University of Westminster
  Brian Knights
- UWCC - School of Pure and Applied Biology
  Dave Bowker
- UWCS - School of Biological Sciences
  no response

APPENDIX E SPECIFICATION FOR DEVELOPMENT OF MIDLANDS REGION'S NEW STOCK ASSESSMENT SOFTWARE

FISHERIES SURVEY DATA PROCESSING APPLICATION

SOFTWARE SPECIFICATION

PAUL LIDGETT

REGIONAL FISHERIES OFFICER

17 September 1996

1 PROJECT DETAILS

Project Title : Fish Population Survey Data Base

Function : Fisheries

Project Leader : Paul Lidgett
Post Title : Regional Fisheries Officer, Midlands Region

2 PROJECT DATES

Start Date: August 1996
End Date: 29 November 1996

3 OVERALL PROJECT OBJECTIVE

3.1 To develop an application on Agency standard software to process and store data obtained from fish population surveys. This must be capable of -

- recording various types of raw survey information, including individual fish lengths, bulk fish weights, species numbers and numerous descriptive data;
- analysing this data according to standard fisheries science procedures;
- archiving various raw data and processed results in a format which is accessible for future retrieval and reprocessing;
- producing a range of standard survey reports;

3.2 The application will be operated on IBM-compatible PCs, connected to the Agency LAN and in a multi-user Windows environment.

4 OVERALL APPROACH

4.1 The project has two specific requirements -

1. To produce a software application to record, process and present fish survey data and an associated user guide to the specification described below.

2. Subject to user acceptance, to install the application on networked PCs at each area office. This must take account of required user access and information transfer.

4.2 The highly technical nature of the project will require the contractor to liaise extensively with regional fisheries staff to determine and fulfil the exact requirements of the system.

NB. Sites are single sample points where it is possible to undertake depletion estimates, minimum estimates, presence/absence or hydro-acoustic sampling. Reaches are longer sections of rivers or canals where only minimum estimates, presence/absence or hydro-acoustic information may be obtained.

5.3 From the raw data obtained from each of the first three sampling techniques, it is possible to calculate length-frequency distributions and length-weight coefficients for individual species. All-species length-frequency distributions can be calculate from the results of hydro-acoustic surveys.

5.4 Fish survey data are currently recorded and analysed on three different software applications in the Midlands Region -

- Upper Severn area uses an application written in SmartWare II v1.02, integrating both the spreadsheet and database modules and controlled by a project processing file;

- Lower Severn area uses a SuperCalc V spreadsheet template, incorporating standard formulae and macros;

- Upper and Lower Trent areas use an application written in Smart 3.1, again integrating the spreadsheet and database modules.

5.5 The incompatibility of these applications with each other has hampered data transfer and regional reporting in the past. In an attempt to gain regional consistency, an attempt at developing a single application in SmartWare II was made. However, this was found not to be suitable and rejected following a trial period.

5.6 It is currently a NIS requirement that unsupported software be removed from Agency PCs and any required systems be replaced using supported software. In order that Fisheries Science staff can continue to provide expert advice on the status of fish stocks, it is essential that a single replacement application is developed as a matter of priority.

6 REQUIREMENTS

6.1 General requirements

6.1.1 The Agency has adopted MicroSoft Access, Lotus 1-2-3 and Lotus Freelance, operating in the Windows environment, as its standard desktop PC software for archiving, processing and presenting data. The new application must be written using one or more of these packages.

4. Fish population - recording raw data of fish caught (species numbers, individual fish lengths or frequency of species at length-range caught, bulk weights, etc.) and other species details. This might require one record per species caught and different record types depending on the survey method.

6.2.2 All data will need to be archived to allow recalculation. Each survey's data-set must be coded according to a specific coding system to allow it to be cross-referenced to a data base of rivers, canals and stillwaters.

6.2.3 To aid data entry and consistency, data for selected fields will entered from combo- or list-boxes and using formulae/queries. Where appropriate (and if possible), the application will replicate data in the current record from the most recent record of data from the same site.

6.3 Processing

6.3.1 Central to the application is its ability to calculate fish population estimates and a number of other measures from raw data. The statistical procedure used will vary according to the survey method and/or the quality/quantity of data obtained.

6.3.2 Using details of the number of fish caught on each occasion, population estimates can be calculated for individual and all species using four principal methods -

- Seber & LeCren method - this is applicable when the site has been sampled on only two runs.

- Moran-Zippin estimators - this calculates population estimates using two or more runs.

- Carle & Strub (Maximum Weighted Likelihood) method - this can be used in the same way as Zippin, but is considered more robust.

(Historically, Seber & LeCren and Zippin have been used in this region. Carle & Strub is becoming more widespread in its application - the application should provide the option to choose the desired method.)

- minimum estimate - this is simply the total count of fish caught/observed whilst fishing (usually from one occasion)

6.3.3 For each estimate (except minimum estimate), the appropriate confidence limits and variances will need to be calculated and archived.

6.3.4 Appendix 1 (NRA R&D Project 325 - Interim Report) details the various population estimate and confidence limits formulae to be incorporated.

and -

- habitat report (one report per survey), comprising -

  - survey date and location;
  - site dimensions;
  - input habitat data.

and -

- population estimate report (one report per survey) - for depletion estimates, comprising -

  - survey location and date;
  - numbers and weight of each species caught on each run;
  - percentage species composition of total catch (numbers and weight), including option of pie chart;
  - population estimate* (numbers); } - for each and all species
  - biomass estimate* (weight); }
  - density* (numbers per $100m^2$); }
  - standing crop* (grams per $100m^2$); }
  - species richness;
  - efficiency;
  - comments (memo) field.

  * defaults to min. estimate if data obtained does not allow calculation of population estimate

or -

- population estimate report (one report per survey) - for minimum estimates, comprising -

  - survey location and date;
  - number/weight of fish caught; } - for each and all species
  - density (numbers per $100m^2$); }
  - standing crop (grams per $100m^2$); }
  - percentage species composition, plus option of pie chart;
  - species richness;
  - comments (memo) field.

or -

- presence/absence report (one report per survey), comprising -

6.8 Additional requirements/outputs

6.8.1 The Fisheries Classification Scheme (FCS) is a means by which the results of fish population surveys are classified against set of nationally consistent class boundaries. A software version of the system has already been developed as a run-time application of MicroSoft Access. The data entered into the FCS are derived from survey results, either directly or as combined statistics. The application should be capable of producing the various statistics required by the FCS and, if possible, transfer them directly into it.

6.8.2 A user support guide and on-line help facility are required to assist users and introduce new users. These need detailed explanation of installation, the application process and comprehensive trouble-shooting. In addition, details of the system structure must be documented to enable future development.

6.8.3 Raw and analysed fish survey data are currently stored on existing applications. The possibility of migrating some or all of these data to the new application needs to be explored and if feasible, undertaken.

6.9 Support

6.9.1 Whilst users will support the application in the long-term, the contractor will support the application for the first 12 months.

APPENDIX F STRUCTURE OF PROPOSED TABLES WITHIN FISHERY DATA ARCHIVING SYSTEM

The following provides a break-down of the fields that are required within each of Tables A to G within the proposed fishery data archiving system (see Figure 6.1).

The fields marked in bold would be essential to the minimal operation of the system, in terms of producing population estimates from raw field data. Note that, as all of the fields within the 'derived fishery data' table are produced automatically, none of the field descriptors within that table are given in bold typeface.

Fields within Table A, 'site information' (one record per site):

Field	Description
Site identifier	Computer generated index that would uniquely identify any given site. Where details recorded for a site were found to have changed, a new record would be generated with the concomitant production of a new, unique site identifier.
Site name	Full site name.
NGR	National Grid Reference - as an eight digit code (e.g. TL 456 934).
Site length (m)	Length of site surveyed, in metres.
Mean site width (m)	Mean site width, in metres
Site area (m^2)	Where data are entered to the 'site length ' and 'site width' fields, the 'area' field would be calculated automatically, but there would be the option to enter an area estimate directly.
Site gradient ($m.km^{-1}$)	To ensure future compatibility with (for example) the National Fisheries Classification Scheme there would be the option to enter an estimate of the site gradient (in $m.km^{-1}$).

N.B. To aid the precise identification of site limits, it may be prudent to include two 'free form' fields, to permit the entry of descriptions of the upper and lower site limits. In addition, further fields may be included in this table (such as, for example: conductivity; mean depth; percentage area of submerged vegetation; distance from source of river). There may also be scope for storing Bitmap images from scanned site photographs within a field in this table.

Fields within Table B, 'survey details' (one record per survey):

Field	Description
Survey identifier	Computer generated index that would uniquely identify any given survey. This field would be used susbequentlyt to link between the tables within the database.
Site identifier	Field to link table B to the site information table (Table A).
Date	Entered in a standard manner - e.g. dd/mm/yy.
Method	A letter code used to indicate the type of survey undertaken (e.g. removal, constant effort or CPUE, or mark-recapture). This field would also flag whether the 'probability of capture' field (see below) was calculated from removal data, derived from habitat information or entered as a subjective estimate.
Number of runs/samples	The total number of runs or removals undertaken during the survey.
Staff	'Memo' field for recording names of staff involved in survey.
Notes	'Memo' field for recording any further information.

Fields within Table C, 'length-weight relationships' (one record per relationship):

Field	Description
Survey identifier	Linking field to identify the survey from which the data were derived.
Species code	A simple coding for species, linking to Table F.

Additional fields would be required to indicate the nature of the relationship and the values of the coefficients used.

Fields within Table D, 'raw fishery data' (one table per survey, one record per fish):

Field	Description
Survey identifier	Linking field to identify the survey from which the data were derived.
Fish identifier	This field (which would be a sequential record counter) would identify the specific fish for which the data pertains.
Species code	A simple coding for species, linking to Table F.
Run/sample number	For surveys that entailed only a single run (e.g. CPUE fishing) the value of this field would default to one; in the case of surveys based on removal methods this field would indicate the run number, whereas for mark-recapture it would indicate the sample number.
Mark present	This field would only be used when the survey type was 'mark-recapture' and would indicate whether a fish was marked or not, and if so from which run the mark was obtained.
Number of fish	This field would have a default value of one (i.e. the data held in the other fields of the same record would refer only to one single fish) - values of greater than one would be used for the bulk entry of data (e.g. bulk weights).
Length (cm)	Standard fork length of fish in centimetres.
Method of length estimation	A single letter field indicating whether the value entered into the 'length' field was: a direct measurement; an estimate from weight data; an inferred value from a sub-sample of length-frequency data; etc.
Weight (g)	Weight of fish in grammes.
Method of weight estimation	A single letter field indicating whether the value entered into the 'weight' field was: a direct measurement; an estimate from length data; an inferred value from a sub-sample of weight-frequency data; etc.
Age code	Link to defined age-classes or age-groups as detailed in Table G.
Method of age estimation	A single letter field indicating whether a fish was aged directly (i.e. from analysis of scales or otolith samples) or indirectly (e.g. inferred from fork length).

Fields within Table E, 'derived fishery data' (one record per species per survey or per species group per survey):

Field	Description
Survey identifier	Field to identify the survey from which the data were derived.
Species code	A simple coding for species, linking to Table C.
Age code	Link to defined age-classes or age-groups as detailed in Table G.
N	The calculated population estimate.
Var (N)	The variance of the population estimate.
P	The estimate of the probability of capture.
Mean length	Mean length.
Mean weight	Mean weight.

Fields within Table F, 'species codes' (one record per species or combination of species):

This table would be 'read-only', in that it is envisaged that the records within the table would not be added to or changed but only used to convert species codes to species names in any outputs that are produced.

Field	Description
Species code	This would be a simple code, allowing linkage between Table F and Tables C, D, and E.
Species name (i)	This field would contain the full standard (English) version of the species name. Alternatively this field would, where appropriate, contain an indication of species nature (e.g. as used for the NFCS; *predatory* species, *rheophilic* species, *limnophilic* species, etc.).
Species name (ii)	This field would contain the full standard (Latin) version of the species name or species grouping.

Fields within Table G, 'age groups' (one record per age-class or age group):

As for Table F, Table G would be 'read-only', in that it is envisaged that the records within the table would not be added to or changed but only used to convert age codes to defined age-classes or age groups in any outputs that are produced.

Field	Description
Age code	Identifier for age-class or age-group.
Age definition	The estimated age-class (0+, 1+, etc.) or age-group (fry, juvenile, all ages, etc.) to which the derived estimates refer. This field would be used for linking to Tables D and E.

N.B. The records held within this table would need to code for a wide range of age-classes and age-groups. It would be prudent to have the option to store data for quite old fish (e.g. possibly up to 15+ in order to accommodate long-lived species such as carp) although a top limit would be best decided by consultation with the Agency during the initial phases of development. In addition, and as noted above, it would be appropriate to have a set of age codes available for combined data such as age-groups.