# UKTAG River Assessment Method

# Fish Fauna

# Fish Classification Scheme 2 (FCS2) Scotland

**by**

**Water Framework Directive – United Kingdom Technical Advisory Group (WFD-UKTAG)**

**Health and safety statement**

WARNING.  Working in or around water is inherently dangerous; persons using this standard should be familiar with normal laboratory and field practice. This published monitoring system does not purport to address all of the safety problems, if any, associated with its use. It is the responsibility of the user to establish appropriate health and safety practices and to ensure compliance with any national regulatory guidelines.

It is also the responsibility of the user if seeking to practise the method outlined here, to gain appropriate permissions for access to water courses and their biological sampling.

# UKTAG Guide to Fish in Rivers
# Fish Classification Scheme 2 (FCS2) Scotland)

## 1 Introduction

This classification method enables the assessment of fish in rivers according to the requirements of the Water Framework Directive (WFD). It encompasses fish abundance, taxonomic composition and age structure. It replaces the interim fish classification tool used for the classifications in the first river basin planning cycle.

### 1.1 Method summary

The FCS2 Scotland classification tool relies on predictive statistical models which are based on the Environment Agency Fisheries Classification Scheme 2 (FCS2) method. The models have been adapted specifically for Scottish fish species and sites.

The tool relates the number of salmon and trout caught in a survey to the predicted abundance and prevalence of the species at the specific site. Salmon and trout are separated into age groups (i.e. fry (0+, or young-of the year individuals) and parr (1 year and older individuals)) and comparisons are therefore made for four groups in total.

The expected catch for each group is modelled on the survey area, a range of site-specific predictive variables (see Appendix 1) and a parameter that describes the probability distribution of the predicted fish catch for a location. The EQR is equal to the probability of observing the same or a lower number of fish at the site under reference conditions. EQR values are calculated hierarchically for each species and age group, for the whole fish survey, then by site (where multiple surveys exist for single site), and finally by waterbody. A classification result is then provided, based on boundary values derived from EQR values for idealised fish populations conforming with the normative definitions provided by the Water Framework Directive.

### 1.2 Environmental pressures to which the method is sensitive

Environmental pressures are intrinsically incorporated within the final EQR outputs from FCS2, because the model uses pressure variables to compare observed results with expected results under reference condition. These are variables which are expected to be related to human activity, and which also have a significant effect on fish populations in the model fitting process. For the final model fit, the following pressure variables were selected: *Artificial impassable barriers*, *Salmon stocking*, *Ammonium*, *SRP/MRP* and *Coniferous forests percentage*. Reference values for these covariates were taken from median values from sites falling within the range

specified by the relevant WFD standards for high status sites, or from REFCOND guidance where appropriate.

Specifically, the values used were:

Barriers: No artificial barriers

Salmon stocking: no stocking

Ammonium: Upland low alkalinity- 0.236mg/l, lowland high alkalinity- 0.662mg/l

SRP/MRP: Low alkalinity sites- 0.01 mg/l, high alkalinity sites, 0.03 mg/l

Coniferous forest in catchment upstream: maximum of 0.58%

The response of this assessment method to environmental pressures has also been separately tested during the intercalibration process by comparing EQR output with pressure data held in the intercalibration validation database. This showed that the tool responds as expected to changes in water quality and physical habitat.

## *1.3  Geographic application*

This assessment method is appropriate for all rivers which occur in Scotland and which are assessed in accordance with the Water Framework Directive. The range of values for each of the environmental covariates used in the model building datasets should be considered when making assessments (see Appendix 1). Classification of sites with environmental variable values which lie outwith this range should be treated with caution. As an example, there were very few high altitude rivers, or wide rivers in the dataset, so classification results from these sites are likely to have a low confidence.

## *1.4  Inter-calibration*

This is a process whereby all European Member States were required to compare WFD class boundary values for each biological quality element (e.g. fish) to ensure similar levels are set across all countries.  The process involved some adjustments of class boundary values for many of the classification tools in use and this process has had a minor influence on the Good/Moderate boundary values used in the FCS2 method.  Once a classification method has been inter-calibrated, the method must be adhered to by Member States for the purposes of WFD assessment and reporting.

Inter-calibration focussed on the EQRs which define the class boundaries between High and Good and Good and Moderate, the H/G and G/M class boundaries respectively.  For this reason the method tends to refer to class *boundary* values rather than EQRs which define a particular class.

## 2  Input data collection

This method statement should be used in conjunction with the associated file bundle that is located on the UKTAG website.

### *2.1  Fish survey method*

FCS2 relies on an accurate estimate of the number of salmon and trout at a survey site, which should be determined by electric fishing. Electric fishing methods should conform to the CEN standard BS EN 14011:2003 Water Quality- Guidance standard on sampling fish with electricity.

The Scottish Fisheries Coordination Centre (SFCC) has also produced protocols for electric fishing surveys in Scotland[1], to improve the quality and consistency of fish survey data in Scotland. These conform with the CEN standards, but are more detailed. They are publicly available, are supported by a specific training programme and have been adopted by almost all organisations monitoring fish in Scottish rivers. The data used to produce the models on which the FCS2 method depends were collected using these protocols, so any data used for WFD classification should have been collected to these standards. In addition to providing a level of data quality, this will also ensure that the relevant field variables required by the model are also collected.

The model on which the Scottish tool is based was built using data from wadable sites, and did not incorporate any boat-based electric fishing data. Information from other member states suggests that this type of data is not directly comparable, so **FCS2 (Scotland) should not be used to provide a classification for boat-based electrofishing sites.**

FCS2 (Scotland) requires data from area-delimited surveys, because the model assesses observed catches against predicted catches based on numbers per unit area. **Time-delimited data cannot therefore be used.** Data from single-run, area-delimited sites can be used, although multiple-run data will reduce the level of uncertainty in the model, as catch efficiency can be calculated rather than incorporated as an unknown variable. Where multiple run data are used, catches must be separated by run.

A sample input file "Scotland_db_subset_example_file.xls" is provided in the file bundle associated with this method to help users enter data in the required format.

Sample sites should be chosen to broadly represent the conditions in the waterbody being classified, in terms of habitat type and likely pressures.

---

[1]http://www.scotland.gov.uk/Topics/marine/science/sfcc/Protocols

A minimum of one site per waterbody is required for classification, but it is recommended that additional sites are used to provide greater certainty in the classification.

## 2.2 Other input data requirements

In addition to fish counts, information on the predictive environmental variables is necessary to enable FCS2 to calculate the expected fish abundance at the site. These variables are listed in Appendix 1.

The layout and format of this information can also be seen in the example input file, "Scotland_db_subset_example_file.xls". The "Field descriptions" worksheet explains the format and detail required for each parameter.

The usual standard for water quality data is expected to be 3 years mean values from monthly sampling. Lower quality data may be used with caution, and users should be aware that this may result in an incorrect classification.

# 3 Procedures for calculating EQRs

The following sections describe how to generate EQRs for individual surveys, sites and waterbodies using the FCS2 Scotland tool. Each section comprises a step by step process which will provide EQRs and overall WFD status results at each of the possible levels (individual surveys, sites, and waterbodies).

The process of developing the model used to calculate these is complex, and it is not intended to describe this stage in detail here. Full details are available in the development project final report, available through SNIFFER[2], and also included in the FCS2 Scotland file bundle.

## 3.1 Software requirements

The calculation of EQRs and classification results for FCS2 Scotland requires a specially developed add-on software package ("fcs2") to be run in the statistical programme 'R'.

### 3.1.1 Software installation

#### 3.1.1.1 Installing R

R is open-source and freely available from for example http://www.r-project.org/.

The 'fcs2' add-on package was developed on R versions 2.11 and 2.12. At the time of writing this method statement, the current version of R was 3.0.2. **Version 2.12.0**

---

[2] Sniffer (2011), *River fish classification tool: science work phase 3 report*. Available online at http://www.sniffer.org.uk/files/4413/4183/7993/Fisheries_Classification_Phase3v4.0.pdf

**must be used to successfully produce a fish classification.** More recent versions will cause problems with software compatibility.

This archive version should be available by following links to older versions of R on the various websites from which R can be downloaded.

R is available for Windows, Mac and Linux but fcs2 has only been tested on Windows.

### 3.1.1.2 Using R

This method statement is not intended to provide detailed instruction in the general use of the R software. For users unfamiliar with R, a thorough introduction is available from http://www.stats.bris.ac.uk/R/doc/manuals/R-intro.html from which the sample session (Appendix A) may be a good place to start. This introduces the user to the R command line into which commands must be typed and executed.

### 3.1.1.3 Installing R Packages

One of the benefits of R is that additional functionality may be added by using add-on packages such as 'fcs2'.

The fcs2 package depends upon a number of other packages (these are called "dependencies"), most of which are available on the Comprehensive R Archive Network (CRAN).

Packages on CRAN may be simply downloaded and installed from within R by selecting the menu item **Packages -> Install package(s)**. This is the recommended installation method.Note that direct installation through the R menus requires R's access to the internet to not be blocked by a firewall.

The fcs2 package requires the CRAN package **R2WinBUGS** and **RODBC**. Particular functions also make use of **maptools** and **rpanel**. All four of these packages must be installed using the method above. Each of these packages depends upon other CRAN packages that are automatically installed if using the recommended installation method.

The fcs2 package also requires the R package **INLA** which is not available on CRAN but is provided with the FCS2 Scotland file bundle. The INLA package can be found in the "INLA.zip" file. This zip archive is the R package and should be passed directly to R using the menu item '**Packages' ->'Install package(s)' from local zip file**… . Newer versions of INLA are not compatible with fcs2 so this version must be used.

### 3.1.1.4 Installing 'fcs2' package

Once the additional required packages have been installed, the **fcs2** package itself should be installed via a local zip file. The fcs2 package can be found in the

**fcs2_1.0.zip** file. This should be installed using **Packages -> Install package(s) from local zip file** as with the INLA step described above.

## *3.2 Other required files*

The following **six files** are also required for the calculation of EQRs. These are all included with the FCS2 Scotland file bundle. They must all be copied and stored in one folder, which will be used as the working directory within R:

- Scotland_db_subset.xls (Excel database with survey data, <u>which must besaved as excel 1997-2003 '.xls' file type</u>). This is the file which will contain the input data, and it must be named exactly in order for the R packages to correctly identify the input data. A sample excel file "Scotland_db_subset_example_file.xls" is provided in the suite of files available from the FCS2 Scotland download site. This provides details of the input file format, which must be saved with exactly the same column headings in the correct order. Any unexpected formats e.g. text in a number column, are likely to result in an error message and failure to produce a final classification output.
- Calculate_EQRs_Sco.R (R script file).
- All model fit files (SalmonFryFit.RData; SalmonParrFit.RData; TroutFryFit.RData;TroutParrFit.RData).

## *3.3 Procedure for running the EQR calculation*

1. Open R and set the working directory to the correct folder (where the appropriate input files from section 3.2 above have been saved) using the menu item **'File' >- 'Change dir'**

2. Run the 'Calculate_EQRs_Sco' R script using the menu item **'File'>- 'Source R code'**.

This creates the following files within the working directory folder:

'ScoDataAtRef.RData' (the data at reference conditions),

'EQRs.RData' (EQRs saved in the R workspace for reviewing and plotting)

'EQRSummary.RData' (mean EQRs and probability of class in the R workspace)

'EQRSummary.csv' (mean EQRs and probability of class in a .csv table which will be viewable in excel)

**Note**: Calculating EQRs for large sample sizes can take a long time to run (as much as many hours).The R console displays progress to allow this to be checked.

## 3.4 Applying class boundaries to generate a classification

The procedure in section 3.3 above will generate an EQRSummary.csv output file. This contains an overall EQR for each waterbody in Colum M. This value should be used to generate a classification by using the following class boundary values:

| High | Good | Moderate | Poor/Bad |
|------|------|----------|----------|
| EQR>0.85 | EQR between 0.60-0.85 | EQR between 0.11 and 0.59 | <0.11 |

A final step is required for differentiation between status classes of Poor and Bad. Fish survey data should be re-examined for all waterbodies provisionally at Poor/Bad status, and only those waterbodies containing a survey showing a complete absence of all fish species should be classified as Bad. Waterbodies with an EQR of less than 0.11, but with fish present should be classified as Poor.

## 3.5 Worked example

The FCS2 file bundle includes an example input file waterbody_db_subset.xls, which includes data from 19 sites. **This data has been specifically collated for describing the FCS2 method and should not be used for other purposes.** When the steps above are followed, the results in the file example_EQRSummary.csv are generated.

The classification outcomes for the sites are included in Table 1 below. Direct classifications from FCS2 are provided for the High-Moderate sites. For the water-bodies that classify as Poor/Bad using FCS2 a further manual classification step is required to discriminate between these classes. Examination of the original dataset shows that at least one year class of either trout or salmon was present at sites 3, 6, 11, 12, 13 & 14, so applying the rule in step 3.4 above means that these should all be classified as Poor. No salmonids were caught during the survey at site19. The raw data should then be checked for the presence of other species. In this case, no fish of any species were caught, so a classification of Bad is appropriate.

*Table 1: Results generated by running the example input file.*

| Site | Waterbody EQR (all species) | EQR Classification | Final classification |
|---|---|---|---|
| Site 1 | 0.322 | Moderate | Moderate |
| Site 2 | 0.198 | Moderate | Moderate |
| Site 3 | 0.101 | Poor/Bad | Poor |
| Site 4 | 0.613 | Good | Good |
| Site 5 | 0.528 | Moderate | Moderate |
| Site 6 | 0.093 | Poor/Bad | Poor |
| Site 7 | 0.864 | High | High |
| Site 8 | 0.754 | Good | Good |
| Site 9 | 0.754 | Good | Good |
| Site 10 | 0.606 | Good | Good |
| Site 11 | 0.083 | Poor/Bad | Poor |
| Site 12 | 0.058 | Poor/Bad | Poor |
| Site 13 | 0.058 | Poor/Bad | Poor |
| Site 14 | 0.058 | Poor/Bad | Poor |
| Site 15 | 0.277 | Moderate | Moderate |
| Site 16 | 0.277 | Moderate | Moderate |
| Site 17 | 0.277 | Moderate | Moderate |
| Site 18 | 0.980 | High | High |
| Site 19 | 0.003 | Poor/Bad | Bad |

## Appendix 1: Input variables required for FCS2 Scotland

| Field | Description | Source | Range |
|---|---|---|---|
| DataHeldBy | Data owner descriptor | Field data | N/A |
| SiteCode | Unique site code e.g. SEPA location code | Field data | N/A |
| SiteName | Site name | Field data | N/A |
| SurveyDate | Date of survey | Field data | N/A |
| WBId | SEPA Water body ID | SEPA GIS system/ SEWeb | N/A |
| WBName | SEPA Waterbody Name | SEPA GIS system/ SEWeb | N/A |
| NumberOfRuns | Number of separate runs in electrofishing survey | Field data | N/A |
| SurveyArea | Wetted area of survey site in square metres | Field data | 5-1628 |
| WetWidth | Wetted width of survey site in metres | Field data | 0.3-41 |
| Slope | Slope of survey site, in metres per kilometre | GIS or paper maps | 0.01-153 |
| BarrierType | Type of impassable barrier downstream, if present. 3 possible values: Artificial, Manmade or Blank | Field data and/or analysis of barrier data on SEWeb | N/A |
| ImpassableBarriers | Is the site above an impassable barrier? 2 possible values; 1= present, 0= absent | Field data and/or analysis of barrier data on SEWeb | N/A |
| CatchmentAreaUpstream | Total catchment area upstream in square km | GIS | 0.05-2550 |
| CatchmentDrainageDirection | Catchment drainage direction. Where does catchment enter the sea? Seven possible categories: SE= Tweed estuary to Fife Ness E=Fife Ness to Fraserburgh NE= Fraserburgh to Duncansby Head N= Duncansby Head to Cape Wrath NW= Cape Wrath to Mallaig W= Mallaig to Mull of Kintyre SW= Mull of Kintyre to Solway Firth | GIS or field data | N/A |

| Field | Description | Source | Range |
|---|---|---|---|
| **GeologyClass** | Predominant geology type in catchment. 3 possible categories: S=Siliceous, C=Calcareous, O=Organic | GIS | N/A |
| **Altitude** | Site altitude (metres above sea level) | Field data  or GIS | 1-499 |
| **DistanceFromSource** | Distance from the site to the furthest upstream source (metres) | GIS | 115-125066 |
| **DistanceToSea** | Distance downstream from the site to the tidal limit (metres) | GIS | 2-160267 |
| **AnnualMeanFlow** | Annual mean flow in cubic metres per second | Hydrological modelling (Low Flows Enterprise of similar) | 0.003-59.9 |
| **AlkalinityValue** | Alkalinity as CaCO3 (mg/L) | Chemistry monitoring data | 0.54-263.11 |
| **TotalPValue** | Total Phosphorus as P (mg/L) | Chemistry monitoring data | 0.006-0.505 |
| **DOCValue** | Dissolved organic carbon<0.45µm (mg/L) | Chemistry monitoring data | 2.023-24.2 |
| **SuspendedSolidsValue** | Suspended solids (mg/L) | Chemistry monitoring data | 0.63-32.36 |
| **HydrometricAreaNo** | The SEPA-defined hydrometric area which the site lies in | SEPA GIS layer | N/A |
| **LandUse.AgriculturalAreas** | % of catchment with Agricultural landuse. Defined as the following Corrine codes: Non-irrigated arable land (211), Pastures (231), Complex cultivation patterns (242), Land principally occupied by agriculture with significant areas or natural vegetation (243) | GIS analysis of Corrine data | 0-99.9 |
| **LandUse.ConiferousForests** | % of catchment with coniferous forest landuse (Corrine Code 312) | GIS analysis of Corrine data | 0-92.9 |
| **LandUse.Wetlands** | % of catchment with wetland landuse. Defined as the following Corrine codes: Inland marshes (411), Peat bogs (412), | GIS analysis of Corrine data | 0-95.7 |

| Field | Description | Source | Range |
|-------|-------------|--------|-------|
|  | Intertidal flats (423) |  |  |
| **Substrate.Small** | Total % of survey bed substrate composed of "High Organic", "silt", "sand" and "gravel", from SFCC survey protocol | Field data | 0-100 |
| **Substrate.Large** | Total % of survey bed substrate composed of "Pebble", "cobble" and "boulder" from SFCC survey protocol | Field data | 0-100 |
| **Substrate.Bedrock** | Total % of survey bed composed of "Bedrock" from SFCC survey protocol | Field data | 0-90 |
| **Salmon_fry.Run1Total** | Total salmon fry (0+) caught in Run 1 | Field data | 0-1014 |
| **Salmon_fry.Run2Total** | Total salmon fry (0+) caught in Run 2. Leave blank if no 2nd run fished | Field data | 0-352 |
| **Salmon_fry.Run3Total** | Total salmon fry (0+) caught in Run 3. Leave blank if no 3rd run fished | Field data | 0-251 |
| **Salmon_fry.Run4Total** | Total salmon fry (0+) caught in Run 4. Leave blank if no 4th run fished | Field data | 0-12 |
| **Salmon_parr.Run1Total** | Total salmon parr (1++) caught in Run 1 | Field data | 0-133 |
| **Salmon_parr.Run2Total** | Total salmon parr (1++) caught in Run 2. Leave blank if no 2nd run fished | Field data | 0-68 |
| **Salmon_parr.Run3Total** | Total salmon parr (1++) caught in Run 3. Leave blank if no 3rd run fished | Field data | 0-34 |
| **Salmon_parr.Run4Total** | Total salmon parr (1++) caught in Run 4. Leave blank if no 4th run fished | Field data | 0-15 |
| **Trout_fry.Run1Total** | Total salmon fry (0+) caught in Run 1 | Field data | 0-508 |
| **Trout_fry.Run2Total** | Total salmon fry (0+) caught in Run 2. Leave blank if no 2nd run fished | Field data | 0-225 |
| **Trout_fry.Run3Total** | Total salmon fry (0+) caught in Run 3. Leave blank if no 3rd run fished | Field data | 0-96 |
| **Trout_fry.Run4Total** | Total salmon fry (0+) caught in Run 4. Leave blank if no 4th run fished | Field data | 0-59 |
| **Trout_parr.Run1Total** | Total salmon parr (1++) caught in Run 1 | Field data | 0-120 |
| **Trout_parr.Run2Total** | Total salmon parr (1++) caught in Run 2. Leave blank if no 2nd run fished | Field data | 0-70 |

| Field | Description | Source | Range |
|---|---|---|---|
| **Trout_parr.Run3Total** | Total salmon parr (1++) caught in Run 3. Leave blank if no 3rd run fished | Field data | 0-43 |
| **Trout_parr.Run4Total** | Total salmon parr (1++) caught in Run 4. Leave blank if no 4th run fished | Field data | 0-25 |