

# Schedule 5 data request – data notes

---

Cefas 29/10/2020

## Schedule 5 folder layout

This folder contains all data pertaining to the R outputs for TR339. It is organised in the following format.

**Raw data** – Pisces data from 2009-2013 only, Cefas raw data (2014-2017) is held within Cefas' FSS (Fishing Survey Systems) database, which is the central repository of all Cefas' research vessel and charter vessel survey data. The Pisces data is split into the representative years (2009 -2013) with each visit saved in an excel workbook e.g. '*Sizewell\_10.02.18 QAd*'.

**Data manipulation workbooks** – The workbooks are split into folders for Pisces and Cefas monitoring data.

The 'Pisces 2009\_2013' folder holds five annual sub-folders with yearly fish and invertebrate totals compiled. The full compilation of Pices data is contained in the spreadsheet '*SZ CIMP 2009\_2012 Pisces SW compilation spreadsheet MASTER*'. This spreadsheet contains the workings of all Pisces data between 2009 and 2013.

The 'Cefas 2014\_2017' folder holds the file '*TR339 Data compilation workbook Cefas 2014-2017*'. This file contains the workings of all Cefas data collected between 2014 and 2017.

Both the Cefas and Pisces data were combined into data compilation workbooks for weights and numbers. These compilation spreadsheets are located in the 'Data manipulation workbooks' folder, in the following files '*TR339 Combined data Pisces and Cefas - weights 5 for R*' and '*TR339 Simplified combined data Pisces and Cefas\_w\_blenny*'.

**R Code** – The fully annotated R code is provided.

Figure 1 Provides a flow diagram of the process of data handling from raw data files to the master spreadsheets required to run R code to reproduce results tables in TR339.

## Pisces impingement data: 2009 to 2013

The Comprehensive Impingement Monitoring Programme (CIMP) was carried out by Pisces Ltd over a 49-month programme.

This data is saved by programme year:

- CIMP1\_2009-2010;
- CIMP2\_2010-2011;
- CIMP3\_2011-2012;
- CIMP4\_2012-2013; and,
- CIMP5\_2013.

The raw data provided by Pisces for each programme year includes individual data files - QA'd excel spreadsheets for each 24hr period. These data files detail the number, weight (kg) and length range of fish (5mm categories), the weight (kg) of crustaceans plus the weight (kg) of any debris including algae species impinged on the screens for the 18hr bulk sample and up to 6 hourly samples (where applicable).

Each sampling visit has its own corresponding Excel data file, saved in an example format ('Sizewell\_10.02.17 QAd'). In each of these sampling visits the data are presented in three forms: 1) the raw data collected, 2) data representing a complete 24-hour sample with the station as it was actually operating, and, 3) the data estimated at full station capacity.

Details are provided including the date, start and end time, no. of hourly samples, no. of pumps and screens operating, water temperature and salinity, and weather conditions. The QA process was completed by Pisces, any suspected errors in the data were investigated by Cefas. This data was compiled into a spreadsheet containing all impingement data collected by Pisces between 2009 and 2013.

## Cefas impingement data: 2014 to 2017

On-site data was captured using a portable Electronic Data Capture (EDC) system. The benefit of using this system is that it minimises the need to manually key in and checking data on return to the laboratory, thus reducing the potential for errors. During this process the 'deckmaster' EDC machine conducts checks to ensure that all species that were weighed have been measured. Thus, if the incorrect species code is used there will be a mismatch between the two units, and this can be corrected before uploading to the database. Once back in the lab, the sample data was transferred to Cefas' FSS (Fishing Survey Systems) database via a USB stick.

All data were stored in Cefas' FSS database. This database is the central repository of all Cefas' research vessel and charter vessel survey data and is also where the BEEMS offshore survey data are stored. All data were cross-checked for quality.

All manually input metadata were independently checked to ensure accurate inputting. The data from the sampling visit were then uploaded directly into the FSS from the USB stick. Again, the system performs several checks against the data to ensure that the correct sample process codes (weighed and measured WM / weighed and counted WC / weighed only WO) were used to capture that data. Once the data are uploaded, the FSS system automatically calculates any raising factors for sub-sampled WC species, to raise up the sampled numbers to total numbers, and in the case of sub-sampled fish, the system automatically calculates the numbers at length raised to the total weight of fish present, using the measured numbers at length, the total weight and sub-sample weight.

For analysis, all data were extracted from FSS to an Access database, which was interrogated using a series of queries to produce 1) the complete set of sample data, 2) the number, weight and lengths of all weighed and measured (WM) fish and crabs, 3) the number and weight of all weighed and counted (WC) invertebrates, and 4) the weight of any weighed only (WO) components. These were combined to a single dataset ('TR339 Data compilation workbook Cefas 2014-2017') which consisted of, for each sample, the sample details and the weight and number of all the catch components. Having the data in a single dataset ensured that any raising would be uniformly applied to all the components of a given sample.

As described in the sampling methods, samples were of either 1-hr (hourly) or 18-h (bulk) duration, but in some cases, the sampling time was reduced because of high volumes of material entering the net. Therefore, each sample was raised to its intended sampling duration by multiplying by either (60/sampling duration) for hourly samples, or (1080/sampling duration) for bulk samples, where sampling duration was the number of minutes sampled based on the start and end times.

## Combined Pisces and Cefas impingement data: 2009 to 2017

Both the Pisces and Cefas data from 'SZ CIMP 2009\_2012 Pisces SW compilation spreadsheet MASTER' and 'TR339 Data compilation workbook Cefas 2014-2017' respectively were combined into data compilation workbooks for weights and numbers.

This is held within the 'Data manipulation workbooks' folder, in the following files 'TR339 Combined data Pisces and Cefas - weights 5 for R' and 'TR339 Simplified combined data Pisces and Cefas\_w\_blenny'.

## R code

The files 'TR339 Combined data Pisces and Cefas - weights 5 for R' and 'TR339 Simplified combined data Pisces and Cefas\_w\_blenny' are used to run the R code to reproduce the results in TR339.

The R code is saved in the 'R code' folder as: 'Sizewell impingement bootstrap numbers v02 EA'. The R code is fully annotated to aid use.

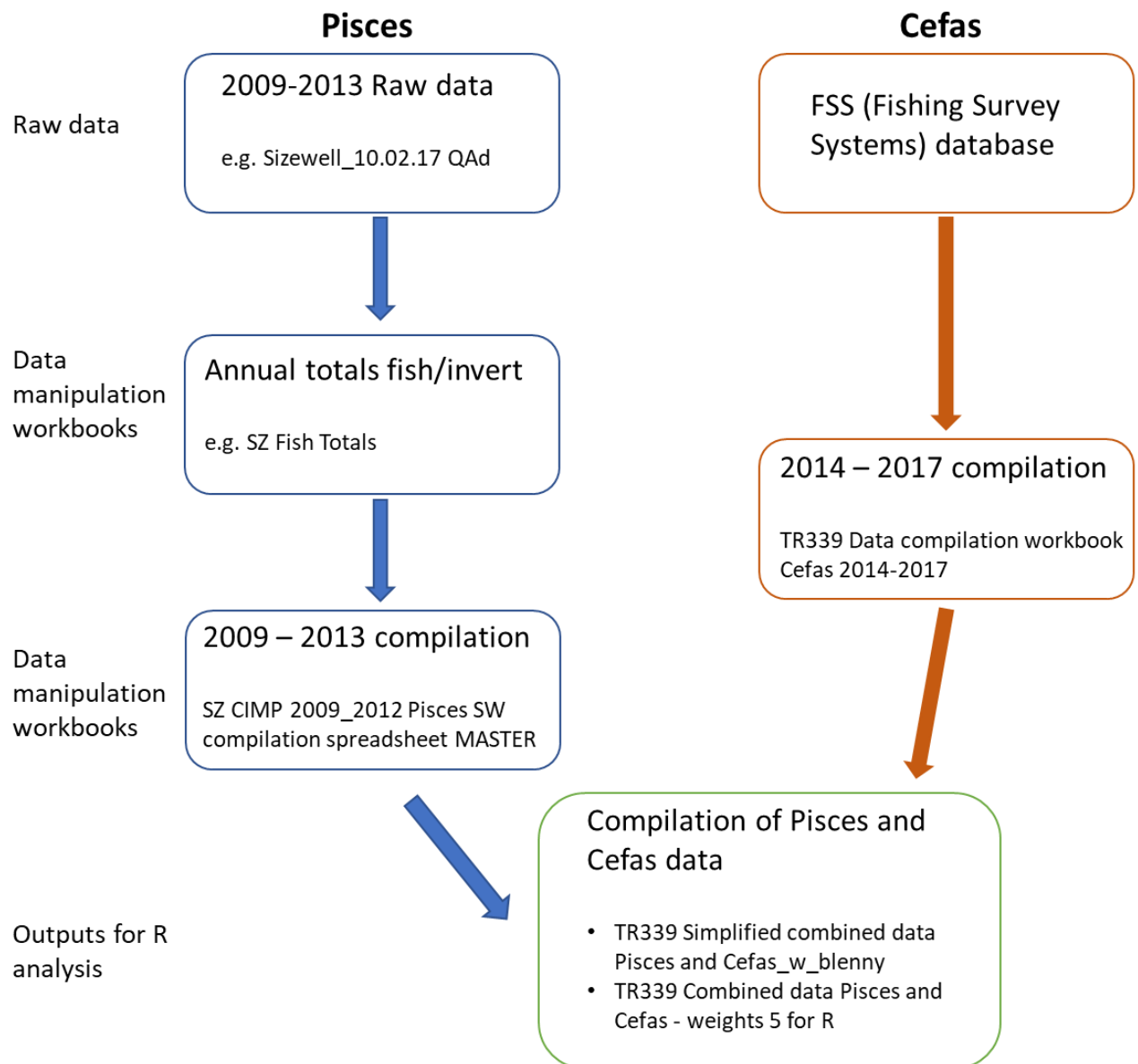


Figure 1 Flow diagram of both Pisces and Cefas data for TR339