Stock Assessment Data Flow

Status Update

Colin Millar, Arni Magnusson
and the ICES Data Centre

WGCHAIRS
24 Jan 2018

ICES
CIEM
Science for sustainable seas

The plan

Brief overview of TAF

Diagram - The TAF workflow

R scripts - Separation of concerns

Data flow in an ICES stock assessment

Overview - From national institutes to advice

Catches - Collation and raising to stock level

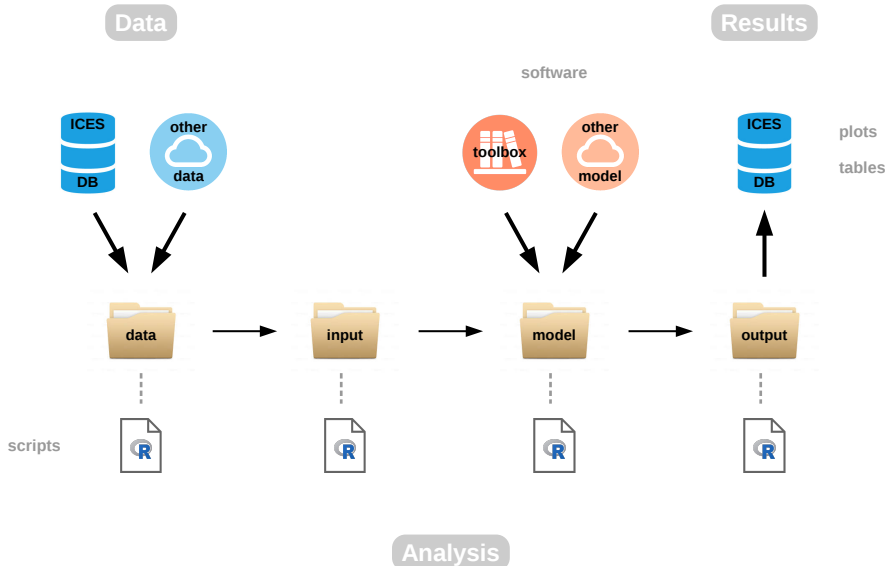
Surveys - Generation of survey indices

Modelling - Data processing and modelling

Report - Formatting and presenting results and data

Advice - High level summary, review, approval

TAF core workflow overview



Workflow scripts

Initial

upload.R - Upload/Download raw data or model files

Core

data.R - Preprocess data, write TAF data tables

input.R - Convert data to model format, write model input files

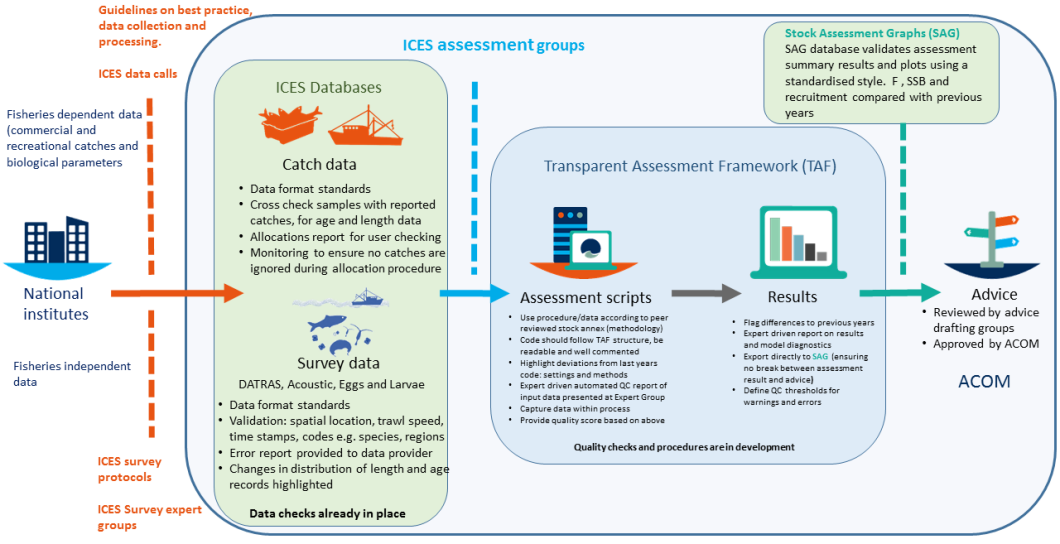
model.R - Run analysis, write model results

output.R - Extract results of interest, write TAF output tables

Also

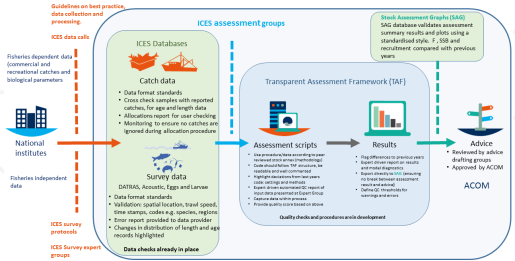
report.R - Prepare plots/tables for report

From national institutes to advice

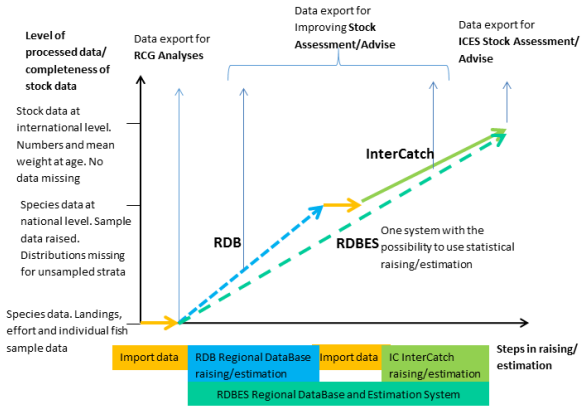


ICES Databases

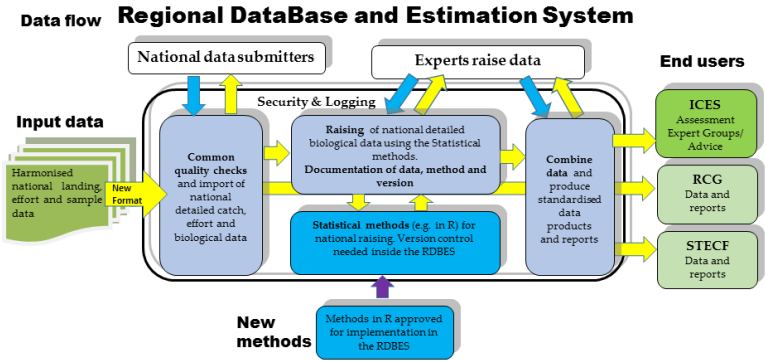
- Regional Database and Estimation System (RDBES)
- Database of Trawl Surveys (DATRAS)
- Acoustic Data Portal
- Stock Assessment Graphs (SAG)
- Eggs and Larvae, Stomach contents, Contaminants and VMS



Data flow in RDB, IC and RDBES



RDBES system overview

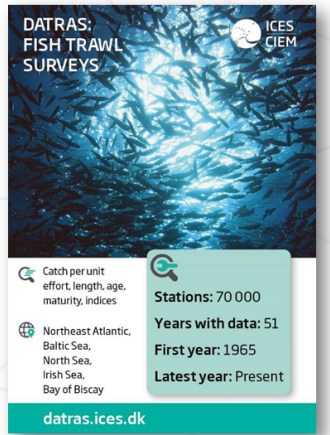


RDBES status

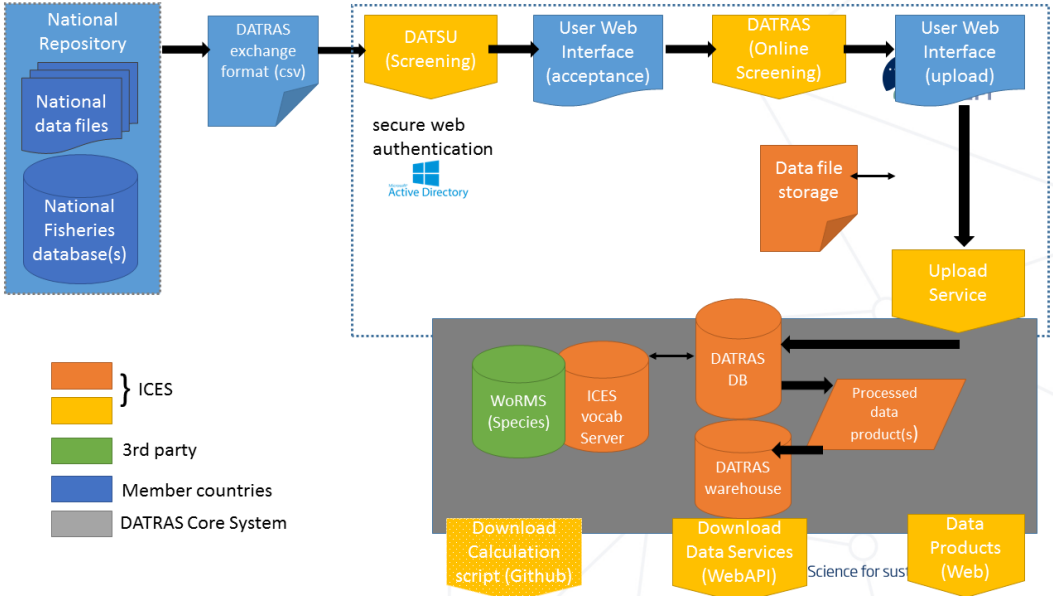
- Currently working with the RDB Development Support Group to specify the data model for the RDBES.
- When the RDBES data model is finalised it is sent to the countries for feedback. Can the countries use it? Is a change needed?
- The specifications of the RDBES should be ready in April.
- The plan is that the RDBES is ready in the spring 2019.

DATRAS (Database of Trawl Surveys)

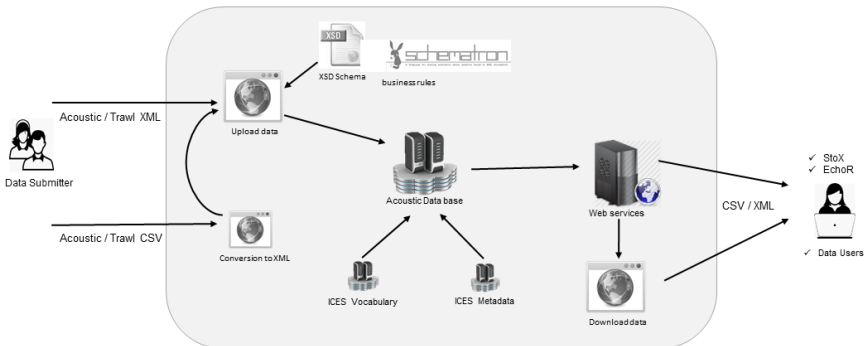
- DATRAS is an online database
- Standardised data extraction
- Updated continuously
- Standard input and quality assurance of all data
- Handles different survey designs (random and stratified)
- Outputs used for fisheries independent stock indices
- Marine litter and biodiversity indicators
- Access data product though WebAPI
- Data product calculation code on GitHub (On going..)



DATRAS dataflow

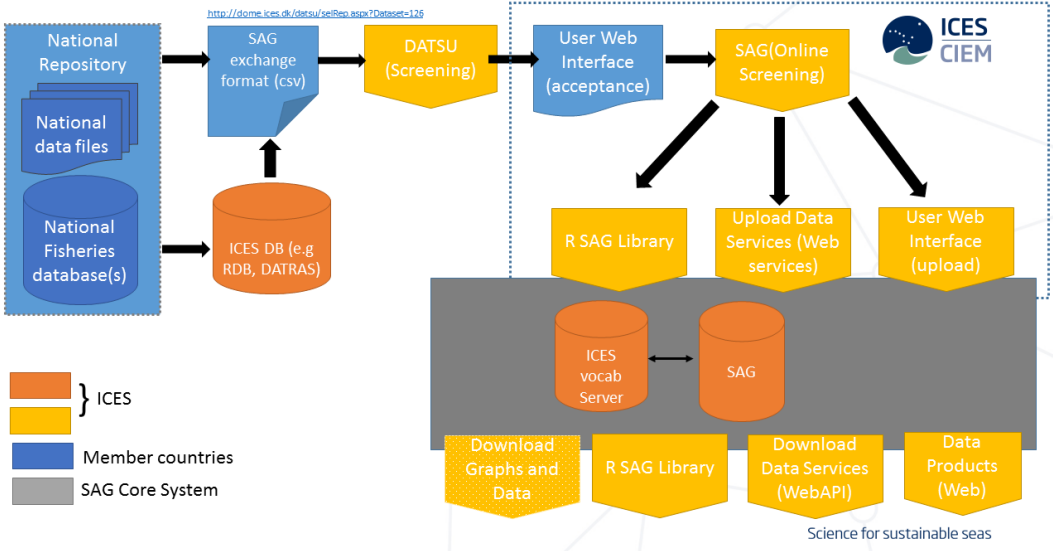


Acoustic Data Portal

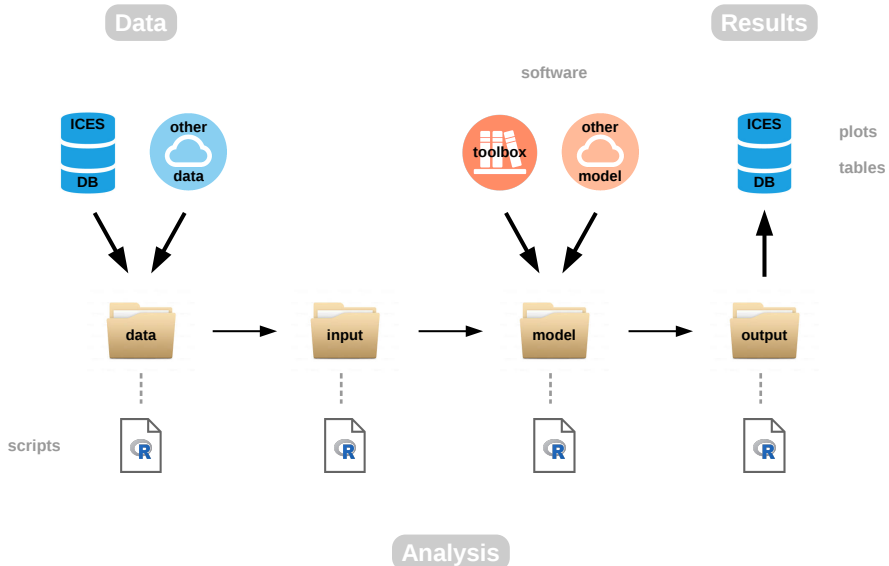


http://acoustic.ices.dk

SAG dataflow



TAF core workflow overview revisited



Workflow scripts revisited

Initial

upload.R - Upload/Download raw data or model files

Core

data.R - Preprocess data, write TAF data tables

input.R - Convert data to model format, write model input files

model.R - Run analysis, write model results

output.R - Extract results of interest, write TAF output tables

Also

report.R - Prepare plots/tables for report

Uploading data from DATRAS Web service - upload.R

```r
library(icesTAF)
library(icesDatras)

# download hh data from DATRAS web service
hh <- getDATRAS("HH", survey = "NS-IBTS", years = 2007:2017, quarter = 1)
# save locally to _raw folder
write.taf(hh, "_raw/hh.csv")

#upload to the 2017 North Sea Cod assessmet TAF repo: 2017_cod.27.2047d
upload("2017_cod.27.2047d", "raw", "_raw/hh.csv")

```

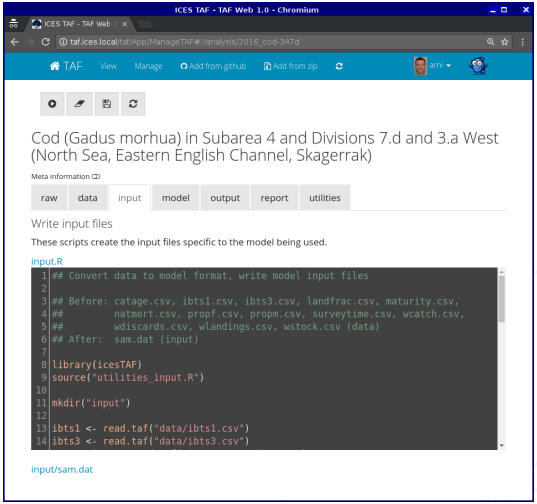
Getting data into the core workflow - data.R

```
5   library(icesTAF)
6
7   # set the download location
8   url <- "http://taf.ices.local/taf/fs/2017_cod.27.2047d/raw/"
9
0   ## Download hh data
1   hh <- read.taf(paste0(url, "hh.csv"))
2
```

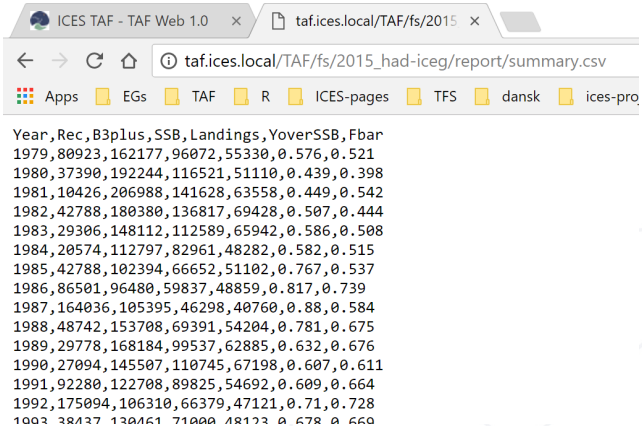
Getting data into SAG - output.R

```
# upload to SAG (with some checking by DATSU on the way
options(icesSAG.use_token = TRUE)
options(icesSAG.messages = FALSE)
key <- uploadStock(info, fishdata)

## Converting to XML format ...
## Done
## Uploading                ...
## Success: (200) OK
## Screening file          ...
## Success: (200) OK
## Importing to database    ...
## Done
## Upload complete! New assessmentKey is: 8651
## To check upload run (with 'options(icesSAG.use_token = TRUE)'):
##  findAssessmentKey('whb.27.1-91214', 2010, full = TRUE)

key

## [1] 8651
```

Viewing, editing and running R code online



Inputs and results available to view or download



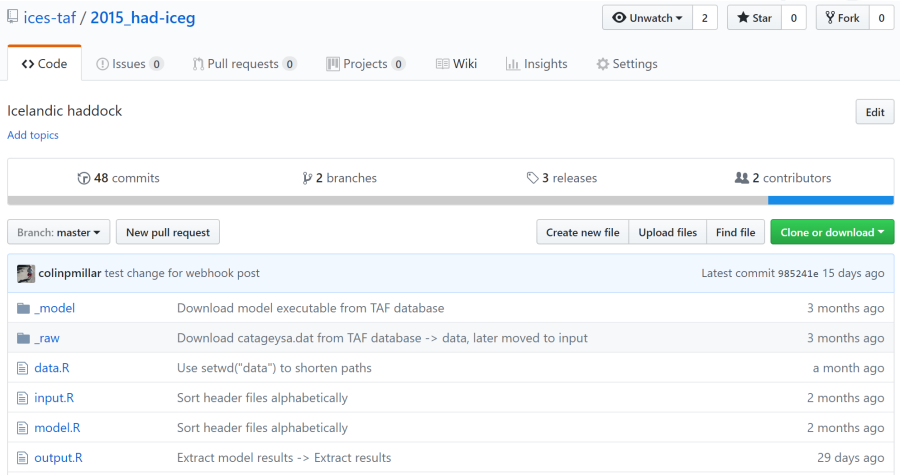
And can be easily read into R

R Console (32-bit)

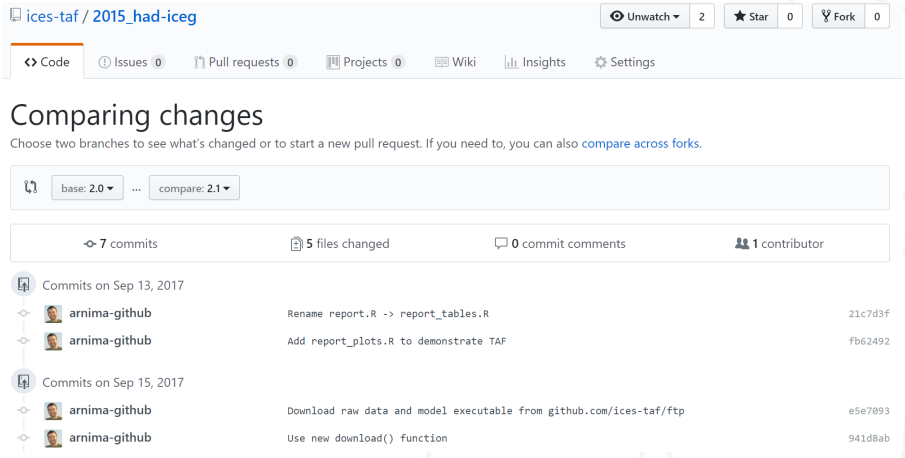
File Edit Misc Packages Windows Help

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> read.csv("http://taf.ices.local/TAF/fs/2015_had-iceg/report/summary.csv")
          Year     Rec B3plus    SSB Landings YoverSSB  Fbar
1         1979   80923 162177  96072    55330    0.576 0.521
2         1980   37390 192244 116521    51110    0.439 0.398
3         1981   10426 206988 141628    63558    0.449 0.542
4         1982   42788 180380 136817    69428    0.507 0.444
5         1983   29306 148112 112589    65942    0.586 0.508
6         1984   20574 112797  82961    48282    0.582 0.515
7         1985   42788 102394  66652    51102    0.767 0.537
8         1986   86501  96480  59837    48859    0.817 0.739
9         1987  164036 105395  46298    40760    0.880 0.584
10        1988   48742 153708  69391    54204    0.781 0.675
11        1989   29778 168184  99537    62885    0.632 0.676
12        1990   27094 145507 110745    67198    0.607 0.611
```

Published assessments accessible on GitHub



List code changes between stages or between years



View code changes between stages or between years

Showing 5 changed files with 86 additions and 10 deletions.

Unified | Split

2 data.R — View

```
@@ -9,7 +9,7 @@ source("utilities.R")
 mkdir("data")
-url <- "http://taf.ices.local/taf/fs/2015_had-iceg/raw/"
+url <- "https://raw.githubusercontent.com/ices-taf/ftp/master/nwwg/2015/had-iceg/raw/"
 ## Download data, this file will later be moved to input
 download.file(paste0(url,"catageysa.dat"), "data/catageysa.dat", quiet=TRUE)
```

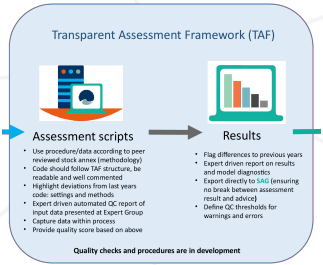
5 model.R — View

```
@@ -7,12 +7,11 @@ library(icesTAF)
 mkdir("model")
-url <- "http://taf.ices.local/taf/fs/2015_had-iceg/model/"
```

Quality Control

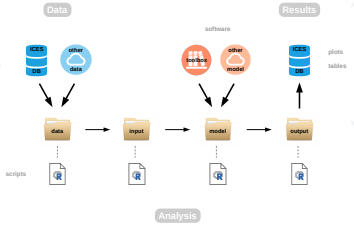
- Code should follow TAF structure, be readable and well commented
- Highlight changes from last year: code, settings, methods, data and results
- Expert driven QC reports covering inputs, results and model diagnostics
- Define QC thresholds to flag warnings
- Directly connecting DBs → assessment → SAG (advice)
- Provide quality score

TIC-TAQ



Benefits

- Easy to find data and results from final assessment
- Open and reproducible science, improved quality control
- Easy for scientists around the world to get ICES data
- Easy to run an update assessment next year
- If scientist changes jobs, next person can take over
- Existing and future tools can use TAF services



Follow ongoing development

taf.ices.dk

Main landing page

ices-taf-dev.github.io

Technical overview and design comments

ices-taf-dev.github.io/pdf/taf-flyer.pdf

TAF outline and diagram

github.com/ices-taf

Example assessments (R scripts)

Time for a demo :)

Thanks!

Questions?