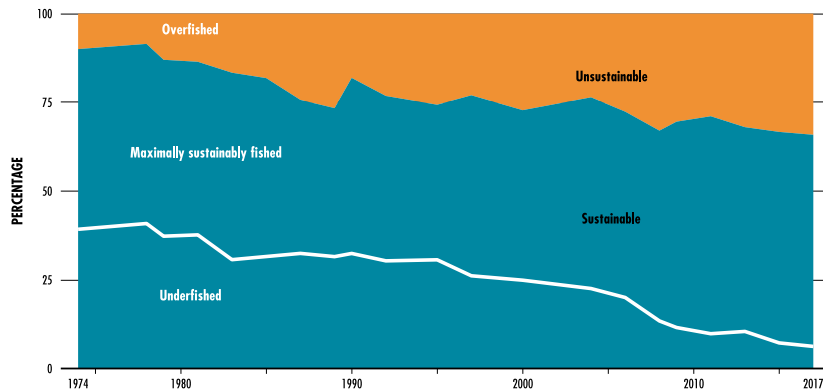
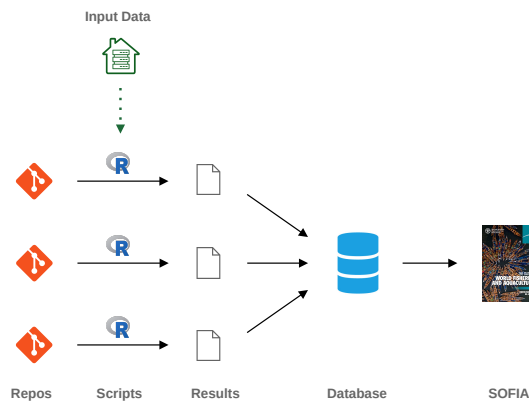


Transparent Analytical Framework for Evaluating the Status of Stocks

SOFIA-TAF Design and Development Progress



Arni Magnusson

December 2024

Contents

1	Executive summary	3
2	Introduction	4
2.1	Background	4
2.2	Objectives	5
3	Design	6
3.1	SOFIA-TAF repositories	7
3.2	Input data pathway	9
3.3	SOFIA R package	10
3.4	Database	10
4	Development status and next steps	11
4.1	SOFIA-TAF repositories	11
4.2	Input data pathway	13
4.3	SOFIA R package	14
4.4	Database	16
5	Acknowledgements	17
6	References	17
A	Arni’s contributions in 2024	18

1 Executive summary

This report gives a brief overview of the design and development progress of the new SOFIA Transparent Analytical Framework (SOFIA-TAF) for evaluating the status of stocks. The framework consists of four components:

- SOFIA-TAF repositories, where each repository contains one analysis, calculating the status of stocks in a given area.
- Input data pathway, with fisheries data for all areas and stocks.
- SOFIA R package, a collection of utilities that are commonly used in SOFIA-TAF analyses.
- Database, storing the results from all SOFIA-TAF analyses.

This report uses a similar structure as the previous reports on SOFIA-TAF design and development progress report (Magnusson 2021, 2022, 2023), updating each section to reflect the current state of SOFIA-TAF. The main activities and milestones achieved in 2024 are listed in Table 1.

Table 1. Overview of main activities and milestones in 2024.

Apr 2024	Development update of the R package <code>fishstat</code> to incorporate the newest FAO FishStat data release. — Workshop with experts in Area 51 (W Indian Ocean).
May 2024	Workshop with experts in 77 (E Pacific).
Oct 2024	New version 2.1.3 of the R package SOFIA released, with improved plots and a new color scheme. — Updated standard demos WorkshopEffortShared , WorkshopEffortByStock , WorkshopIndexByStock , and WorkshopPriorsByStock to use SOFIA 2.1.3. — Workshop with experts in Area 61 (NW Pacific).
Dec 2024	SOFIA-TAF repositories developed in 2024 are in Areas 51, 77, and 87. — Progress report summarizes the overall design and the current development status.

At the end of this report is an appendix, listing Arni Magnusson’s contributions in 2024 to SOFIA-TAF design and development.

2 Introduction

2.1 Background

To provide a frame of reference behind the activities and milestones from SOFIA-TAF development in 2024, it is worthwhile to quickly review the timeline from 2021–2023 (Table 2). See Magnusson (2021, 2022, 2023) for details.

Table 2. Review of key activities and milestones in 2021–2023.

Mar 2021	Initial design discussions.
Nov 2021	Initial development of the SOFIA package.
Dec 2021	Total of 12 SOFIA-TAF repositories on GitHub. — First design and development progress report.
Jan 2022	SOFIA package version 1.0 released.
May 2022	First workshop with experts in Area 37 (Mediterranean and Black Sea).
Oct 2022	SOFIA-TAF launch event at CAPAM Good Practices conference in Rome.
Dec 2022	Total of SOFIA-TAF 32 repositories and 3 regional workshops.
Apr 2023	Development of four standard demos that are used in regional workshops and in routine tests of the virtual research environment (VRE): WorkshopEffortShared , WorkshopEffortByStock , WorkshopIndexByStock , and WorkshopPriorsByStock .
Jun 2023	Initial development of R package <code>fishstat</code> that provides FAO FishStat catch data in R format.
Dec 2023	Total of 44 SOFIA-TAF repositories and 8 regional workshops.

Regional workshops (Table 3) are of central importance for introducing the new framework for evaluating the SOFIA status of stocks. The workshops provide an opportunity to reach out to regional experts, demonstrate the new analytical approach, and identify fisheries where additional data could be submitted to FAO.

Table 3. List of regional workshops.

Area	Short name	Time	Venue
31	W Atlantic	2022 Nov	Miami
34	E Atlantic	2023 May	Banjul
37	Med & Black Sea	2022 May	Rome
41	SW Atlantic	2022 Nov	Mar del Plata
47	SE Atlantic	2023 Dec	Cape Town
51	W Indian Ocean	2023 Mar	Dar es Salaam
—	W Indian Ocean II	2024 Apr	Kochi
57	E Indian Ocean	2023 Jan	Bangkok
61	NW Pacific	2024 Oct	Rome
71	W Pacific	2023 Nov	Bangkok
77	E Pacific	2024 May	Mexico City
87	SE Pacific	2024 Nov	Santa Marta

2.2 Objectives

It is not for this author to define or prioritize the objectives of SOFIA-TAF to support the overall analysis behind SOFIA. However, the following topics are worth mentioning, as a context for some of the design decisions and features that are being developed.

Efficiency is the ability to edit code in a single place to modify a large number of analyses, and to calculate top-level summary statistics from a large number of analyses.

Clarity is the ability to easily navigate to a specific part of the analysis of a particular stock group and area, and to look up a specific result from one or more analyses.

Traceability is the ability to backtrack exactly how a specific result was calculated, such as the status of a particular stock group in a given area.

Open science is the ability to make the R scripts available online, along with the input data required for the scripts to run, inviting peer review of methodology and scientific collaboration.

Reproducibility is the ability to run analyses on a variety of computers, e.g., a personal Windows laptop or a high-performance Linux cluster, to get the exact same result — also when the analysis is rerun months or years later.

Quality assurance is the design and adoption of a workflow that reduces the probability of making human mistakes when preparing, modifying, running, and postprocessing the results from analyses.

Quality control is the ability to identify where a human mistake has been made in a given analytical process, so the mistake can be located and corrected.

The initial development of SOFIA-TAF has focused especially on Tier 2 cases of SOFIA analyses, where official stock assessments are not available, but catch, effort, and survey index data exist as a basis for estimating stock status using data-limited methods.

The SOFIA-TAF design also aims to serve as a platform to organize Tier 1 analyses (deriving stock status from official stock assessments) as well as Tier 3 analyses (deriving stock status estimates from expert elicitation). These tiers will require less R code than Tier 2 but use the same structure for R scripts and data provenance to document exactly how the stock status was calculated.

3 Design

The SOFIA-TAF design (Figure 1) is based on repositories containing R scripts that can read input data from files or queries, to estimate the status of stocks. These results are then stored in a database, which serves as the foundation for calculating summary statistics for the final SOFIA report.

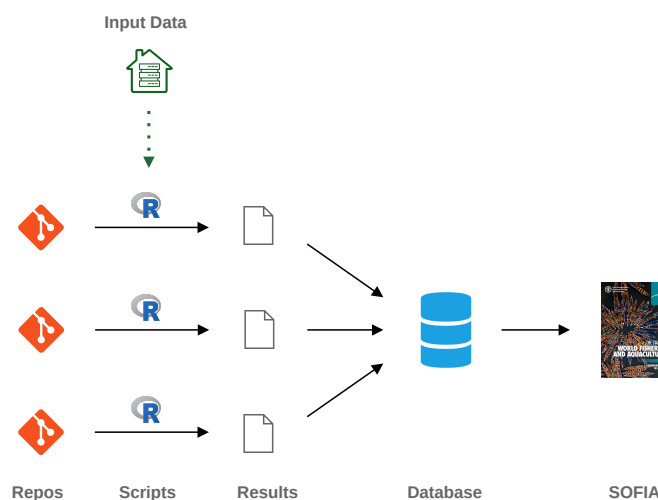


Figure 1. SOFIA-TAF diagram, showing the flow of information from individual repositories (analyses of stocks and areas) to the final SOFIA report.

The scripts use a dedicated R package called SOFIA. The next sections describe each component of the SOFIA-TAF design in some detail: repositories, input data pathway, SOFIA R package, and the database.

3.1 SOFIA-TAF repositories

3.1.1 Repository features

Each SOFIA-TAF repository is a unit of analysis, corresponding to a specific area and a set of stocks. A GitHub repository, sometimes abbreviated as ‘repo’, is an online directory that is especially convenient for organizing text files, such as scripts and data files. GitHub repository features relevant for SOFIA-TAF include:

- Ability to make scripts available online, for browsing and downloading, along with the input data required for the scripts to run.
- Automatic backup of all files with the ability to return to previous saved states.
- Tracked changes showing who changed what and when, supporting online teamwork.
- Ability to tag specific saved states of the analysis and give them descriptive names, such as ‘starting point’, ‘2021 data’ or ‘results imported to database’.
- Ability to upload large attachments (>100 MB) to accompany tagged states.
- Online facilities to compare text files and view changes, line by line.

3.1.2 R scripts

The analysis inside each repository consists of a set of R scripts that are organized in TAF format (Magnusson and Millar 2023). This means there are four standard scripts (Table 4) that conduct and document the analysis:

Table 4. Standard TAF scripts for a given analysis.

Script	Purpose
<code>data.R</code>	Preprocess data, write TAF data tables
<code>model.R</code>	Run analysis, write model results
<code>output.R</code>	Extract results of interest, write TAF output tables
<code>report.R</code>	Prepare plots and formatted tables

The TAF scripts are run sequentially, each reading files that were created in a previous step. The first script, `data.R`, reads data files that were declared and documented in a `DATA.bib` text file. A similar `SOFTWARE.bib` file can be used to declare specific versions of software used in the analysis, to strengthen reproducibility.

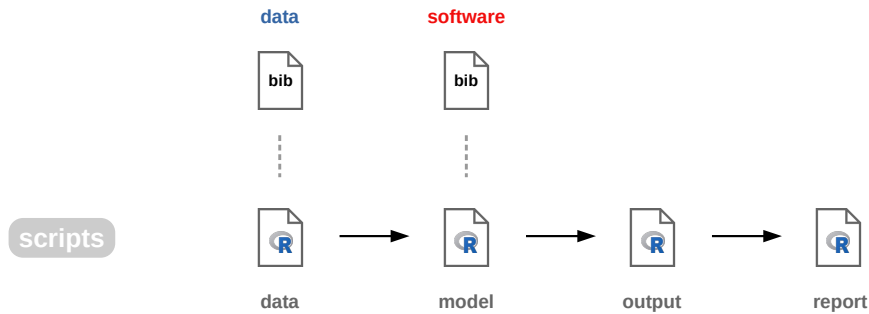


Figure 2. TAF scripted workflow. Each SOFIA-TAF repository/analysis contains four standard R scripts that are run sequentially. The initial data and software are declared in so-called bib files.

The R scripts conducting SOFIA-TAF analyses rely especially on three R packages:

- **SOFIA** – a dedicated package to support SOFIA-TAF (Sharma and Magnusson 2024)
- **TAF** – utilities to manage scripts, data files, metadata, and R data objects (Magnusson and Millar 2023)
- **sraplus** – biomass dynamics model with Bayesian priors (Ovando 2022)

3.1.3 Repository names and directory structure

GitHub repositories are given descriptive names, such as

<https://github.com/sofia-taf/2021Area37Coastal>

for the analysis started in 2021 of coastal stocks in Area 37.

On a personal laptop or a high-performance cluster, the SOFIA-TAF repositories are organized in a hierarchical directory structure (Figure 3), similar to the repository name as year/area/stock group.

This hierarchical directory structure is practical to navigate and run the analyses, access results, and run top-level summary calculations across a large number of analyses.

Remote repository

<https://github.com/sofia-tsaf/2021Area37Coastal>

Local directory structure

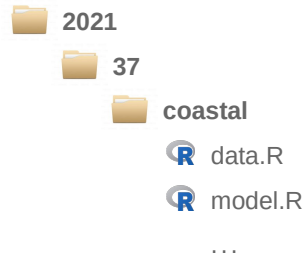


Figure 3. SOFIA-TAF local directory structure, showing the similarity between the remote repository name and the local directory names.

3.2 Input data pathway

3.2.1 Data collection templates

SOFIA-TAF input data are collected and prepared in a collaborative process between FAO and regional experts. The data used within a Tier 2 analysis are organized in files such as `catch.csv`, `effort.csv`, and `index.csv`, containing multiple stocks in each file. Data provided by regional experts, however, are termed *primary* data files and are organized as one file per stock, using filenames such as `Yellowtail_snapper_Mexico.csv`. The conversion from primary data files to Tier 2 input data files is handled by the SOFIA package.

3.2.2 The fishstat R package

As of December 2024, all SOFIA-TAF analyses start with catch data uploaded as a `catch.csv` text file in CSV format.

An alternative input data pathway could be used for the catch data by reading directly from the FAO FishStat database. To facilitate this pathway, development was started in 2023 to create an R package called `fishstat` that provides FishStat catch data in R format as data frames.

Reading catch data directly from FishStat, rather than uploading handcrafted text files, would be an improvement and result in a more streamlined workflow that is traceable and quality-controlled.

The development of the `fishstat` R package will be continued in 2025.

3.2.3 Input data warehouse

Dedicated GitHub repositories can be used to host data that are used across many analyses:

- <https://github.com/sofia- taf/catches> for catch data collections other than FishStat.
- <https://github.com/sofia- taf/effort> for effort data by year and area.

These repositories offer a central place for SOFIA-TAF input data where they can be updated, documented, and quality checked.

3.3 SOFIA R package

The SOFIA package (Sharma and Magnusson 2024) contains utilities that are commonly used in SOFIA-TAF analyses. It is developed in a dedicated GitHub repository:

<https://github.com/sofia- taf/SOFIA>

The SOFIA package provides a single place to modify a large number of SOFIA-TAF analyses. Incremental improvements become more manageable, e.g., changing the format of a specific plot, without having to edit each and every SOFIA-TAF analysis. It also makes the R scripts for each analysis shorter and thus easier to read, write, and maintain.

3.4 Database

The database will contain the results from all the individual SOFIA-TAF analyses. The results describe the status of stocks by FAO Major Fishing Areas, both in numerical terms and in categorical terms: underfished, fully fished, and overfished.

The only way to enter stock status results into the database is via SOFIA-TAF analyses, as indicated in the design diagram (Figure 1), and when analyses of a specific areas and stocks are updated, the database is automatically updated. Furthermore, the top-level analysis for the final SOFIA report, aggregating a large number of SOFIA-TAF analyses, should be based on queries to the database.

The above design guarantees the traceability of SOFIA results, all the way from the individual datasets and analyses to the final published report.

The database is also a convenient stage in the pipeline to apply quality control. Examples of quality checks could include referential integrity of species and stock names, summary statistics at different levels of aggregation, counting stocks in each area, plotting the distribution of numerical stock status, etc. This will ensure that all stocks are accounted for, and reveal any inconsistencies or issues that should be checked in the underlying analyses.

4 Development status and next steps

4.1 SOFIA-TAF repositories

4.1.1 Current status

As of end of December 2024, the GitHub site <https://github.com/sofia- taf> contains 49 SOFIA-TAF repositories, 12 that were created in 2021, 20 were created in 2022, 12 were created in 2023, and 5 were created in 2024 (Table 5).

Table 5. Overview of SOFIA-TAF repositories.

Year	Area	Number of analyses	Specifically
2021	Area 31	1	Test
	Area 37	11	ClamsOK, Coastal, Cods, Demersal, Flounder, Herring, Other, Shads, Shimps, Squid, Test
2022	Area 31	10	CatchEffortGlobalv2JM, CatchOnly, DemoEffortByStock, DemoEffortShared, DemoIndexByStock, EffortJMstcks, effortlocalcatchlocalv2JM, IndexMethodv2JM, Indexstocks, Testnewstocks
	Area 37	2	Clams, Demo
	Area 41	3	DemoPriorsByStock, RelevantStocks
	Area 57	4	Efforttest, EffortTestNoPriornew70yrstocks, Tier2FAOMonitored31withdepPrior, Tier2FAOMonitorednodepPrior
	Deep Seas	1	CCAMLR
2023	Area 34	4	Decapterus, DemersalsSouthLocalEffort, GlobalEffortFishstatstocks, Loligo
	Area 51	2	Demo, ShortDemo
	Area 57	5	33MonitoresStockssraplusSathia, Demo, FAOnonmonitoredstcks70yrs, NonMonStocks-25-50yrs, NonMonStocks50-69yrs
	Area 71	1	FishStat30
2024	Area 51	1	Fishstat
	Area 77	3	Fishstat, Tier2Manual, Tier2FishstatJ
	Area 87	1	Tier2

The status of all of these analyses is exploratory for development purposes, as opposed to production analyses for a final SOFIA report. Each SOFIA-TAF analysis declares the version of the SOFIA package used, so the analysis can be rerun at a later point and produce consistent results.

In addition to the main GitHub site <https://github.com/sofia- taf>, another site <https://github.com/sofia- taf- dev> has been created, as an alternative place to organize development. One approach to manage SOFIA-TAF would be

store experimental analyses on the ‘dev’ site and production analyses on the main site.

Four demo repositories are used in regional workshops and in routine tests of the virtual research environment (VRE):

- [WorkshopEffortShared](#)
- [WorkshopEffortByStock](#)
- [WorkshopIndexByStock](#)
- [WorkshopPriorsByStock](#)

These four demos serve as a reference for SOFIA-TAF, demonstrating standardized scripts for analyzing various combinations of effort data, survey index data, and priors. As reference analyses, they will be updated and synchronized when changes and improvements are introduced in the SOFIA package and other components of the SOFIA-TAF framework.

4.1.2 Next steps

Read catch data from the fishstat package

The current development state of the `fishstat` R package is ready for testing. A test repository would start from an existing analysis and then replace the `catch.csv` file with an R script that extracts the catch data from the `fishstat` package. If the `catch.csv` file contained catch data from the FishStat database, then the two input pathways should produce the same input data. The `fishstat` package offers a more efficient, traceable, and quality-controlled workflow.

Tier 1 and 3

All the current SOFIA-TAF repositories developed so far correspond to Tier 2 analysis. Tests of Tier 1 and/or Tier 3 analyses are likely to be scheduled in 2025, involving numerical and categorical data from official stock assessments and expert elicitation. This will further increase the clarity and traceability of the overall SOFIA analysis of stock status.

Managing repositories

During the development of SOFIA-TAF, a number of experimental analyses have been created as repositories on the main <https://github.com/sofia-taf> site. Some analyses are likely to become production analyses for a final SOFIA report, whose results should be imported into the database of SOFIA-TAF results, while other analyses could be put aside and migrated to the development area on <https://github.com/sofia-taf-dev>. Some older experimental analyses could be deleted, if they have been superseded by more recent analyses.

When importing results from analyses into the database of SOFIA-TAF results, the relevant analyses could be defined as all repositories on <https://github.com/sofia-taf> whose name begins with year and area.

Managing output files

A significant challenge in SOFIA-TAF is that each analysis takes considerable time to run (>1 hr) and produces large output files (>100 MB). Every time a small update is made to the scripts or underlying data, a new run is required and the output files are likely to change. SOFIA-TAF development so far has explored two approaches to store the output files.

Approach 1. Initial development (e.g., `area37`) kept the output files outside of the repository. Instead of uploading the output files along with the R scripts, output files were uploaded as GitHub ‘release assets’. The advantage of this approach is that the repository remains very light and easy to work with, and takes much less space on the hard drive of a personal laptop.

Approach 2. Later development (e.g., `2021Area37Coastal`) has the output files stored inside the repository. The advantage of this approach is that it reduces the need to manage tags and GitHub releases, and makes it slightly less likely to have mismatching R scripts and output files.

Unfortunately, neither of the above approaches can guarantee a correct match between the R scripts and the output files. In other words, when a change is made to an R script and uploaded to the repository, it’s easy to forget or omit running the entire analysis and uploading new output files.

A 3rd approach worth exploring would be not to upload output files to the GitHub repository at all. Instead, the database server would run all analyses locally. Specifically, a GitHub webhook could be developed, so that whenever a change is uploaded to the GitHub repository, the database server pulls the changes, runs the analysis and imports the results into the database. This would guarantee a strong linkage between the SOFIA-TAF repositories and the database.

4.2 Input data pathway

4.2.1 Current status

The `fishstat` R package is being developed in a public repository, <https://github.com/sofia-taf/fishstat>. It is now ready for testing (see Section 4.1.2: Read catch data from the `fishstat` package).

The input data warehouse <https://github.com/sofia-taf/catches> repository currently contains one file `cap_2021-10-17_193245.csv` with catch data. The <https://github.com/sofia-taf/effort> repository has been created but is still empty.

4.2.2 Next steps

Once testing has completed, the `fishstat` R package will be announced, possibly with a short journal communication.

Effort data by year and area can be added to the `effort` repository in CSV format. These general effort series are used when stock-specific effort data are not available.

4.3 SOFIA R package

4.3.1 Current status

Version 2.1.3 of the `SOFIA` package was released on 8 October 2024. The package help page that comes with `SOFIA` lists the following functions, categorized by functionality.

Prepare data:

<code>addDriors</code>	add driors column to stocks object
<code>addEffort</code>	add effort column to catch data
<code>addIndex</code>	add index column to catch data
<code>convertData</code>	convert primary data to combined data
<code>groupData</code>	group primary data in subdirectories

Calculate:

<code>calcCat</code>	stock status categories
----------------------	-------------------------

Plot:

<code>plotCat</code>	summary of stock status categories
----------------------	------------------------------------

Repositories:

<code>gitRepos</code>	list GitHub repositories
<code>gitClone</code>	clone GitHub repository
<code>gitCloneAll</code>	clone all <code>SOFIA-TAF</code> repositories

The current status of the package is stable and fully documented, passing a strict R CMD `--as-cran` quality check. Table 6 shows the complete release history.

Table 6. SOFIA package release history.

Version	Date	New functions
2.1.3	2024-10-08	
2.1.2	2023-05-19	
2.1.1	2022-11-15	
2.1.0	2022-11-11	<code>addIndex</code>
2.0.0	2022-08-08	<code>gitRepos</code> , <code>gitClone</code> , <code>gitCloneAll</code>
1.2.1	2022-05-21	
1.2.0	2022-02-25	<code>convertData</code>
1.1.0	2022-02-20	<code>groupData</code>
1.0.3	2022-02-14	
1.0.2	2022-01-24	<code>addDriors</code> , <code>addEffort</code> , <code>calcCat</code> , <code>plotCat</code>

A more detailed list of changes introduced in each version of the SOFIA package can be found online at <https://github.com/sofia-taf/SOFIA/blob/main/NEWS.md>.

4.3.2 Next steps

Prior on carrying capacity

With the accumulated experience of applying the `sraplus` model to a wide variety of stocks, it seems that in some cases it might be helpful to use a prior distribution on carrying capacity K . This could keep the estimated value of K within a realistic range. The SOFIA package could assist the user in setting a reasonable prior on K , based on the historical catches, specifically the `format_driors` arguments `k_prior` and `k_prior_cv`. Such a prior could either be applied by default for all stocks or enabled for specific stocks as necessary.

Generalized prior input file

Currently, the prior input file is parsed expecting certain content to be found in the `priors.csv` file, specifically the columns `initial_state`, `initial_state_cv`, `terminal_state`, and `terminal_state_cv`. An improved interface would allow the user to specify any priors, including `k_prior` and `growth_rate_prior`.

New plots

New plots could be implemented, inspired by the JARA package (Winker et al. 2020), showing B/B_{MSY} and F/F_{MSY} time series as dots, median, and confidence regions.

Fixed random seed

When many workshop participants run the same example, they seem to get slightly different results. This could be caused by different random seeds used in the Markov chain Monte Carlo (MCMC) analysis. However, the `sraplus` function `fit_sraplus` has a default `seed = 42` that all workshop participants were probably using. It would be worth looking into how the random seed is handled by the `sraplus` package, as using a fixed random seed would strengthen the reproducibility of SOFIA-TAF analyses.

Quality control

One area of the SOFIA package that can be developed is the creation of quality control functions for SOFIA-TAF analyses, identifying possible mistakes in the analyses that can then be fixed.

4.4 Database

4.4.1 Current status

A new database has been developed by FAO to store SOFIA-TAF results. The presentation by Anton Ellenbroek at regional workshops from 2023 onward gives an outline of the objectives, design, and functionality of this database.

4.4.2 Next steps

Strong link between repositories and database

The link between SOFIA-TAF repositories and the database of results (Figure 1) is essential for traceability and the fundamental purpose of SOFIA-TAF. The very basis of the SOFIA-TAF design is that the results found in the database should match exactly the results produced by the SOFIA-TAF repositories.

The design and development of this linkage is still ongoing, including automatic detection of which analyses contain results that should be imported into the database. The separation of repositories into production analyses on `sofia-taf` and experimental analyses on `sofia-taf-dev` (Section 4.1.2 on ‘Managing repositories’) is related to designing a strong link between repositories and the database of SOFIA-TAF results.

Database server managing SOFIA-TAF output files

One design possibility is to have a dedicated SOFIA-TAF database server as the main platform to run SOFIA-TAF analyses and manage output files, in addition to importing the results into the database.

The development goal ‘Managing output files’ (Section 4.1.2) elaborates on this possible approach, which could involve a GitHub webhook to establish a reliable pipeline of information from SOFIA-TAF repositories to the database.

5 Acknowledgements

Working with Rishi Sharma on this project is a privilege and joy. Our domains of expertise complement each other, and we share a common vision and enthusiasm to enhance the FAO infrastructure and analytical workflows for estimating the state of the world’s fisheries. I would like to acknowledge Colin Millar for our collaboration in creating TAF (Magnusson and Millar 2023), which has served as an inspiration and basis for the SOFIA-TAF design. I thank Anton Ellenbroek for leading the development of a database storing SOFIA-TAF results. Last but not least, I am grateful to Pedro Barros and colleagues at FAO for their guidance and vote of confidence for this technical development project.

6 References

- FAO (Food and Agriculture Organization of the United Nations). 2024. The State of World Fisheries and Aquaculture 2024: Blue Transformation in action. Rome. 232 pp.
<https://doi.org/10.4060/cd0683en>
- Magnusson, A. 2021. SOFIA Transparent Analytical Framework: Design and development progress.
<https://arni-magnusson.github.io/pdf/2021-sofia- taf.pdf>
- Magnusson, A. 2022. SOFIA Transparent Analytical Framework: Design and development progress.
<https://arni-magnusson.github.io/pdf/2022-sofia- taf.pdf>
- Magnusson, A. 2023. SOFIA Transparent Analytical Framework: Design and development progress.
<https://arni-magnusson.github.io/pdf/2023-sofia- taf.pdf>
- Magnusson, A. and C. Millar. 2023. TAF: Functions to Support the ICES Transparent Assessment Framework. R package version 4.2.0.
<https://cran.r-project.org/package=TAF>
- Ovando, D. 2022. sraplus: Run sraplus Assessments. R package version 3.7.5.
<https://github.com/DanOvando/sraplus>
- Sharma, R. and A. Magnusson. 2024. SOFIA: Tools to Work with SOFIA-TAF Analyses. R package version 2.1.3.
<https://github.com/sofia- taf/SOFIA>
- Winker, H., N. Pacoureaux, and R.B. Sherley. 2020. JARA: Just Another Red-List Assessment. bioRxiv.
<https://doi.org/10.1101/672899>

A Arni's contributions in 2024

SOFIA-TAF design

The work on this project is divided between design and development. The design part takes place largely during online technical meetings of the SOFIA-TAF development team, and is the product of dynamic teamwork and discussions. As the SOFIA-TAF design borrows both ideas and technical components from TAF (Magnusson and Millar 2023), Arni has served in a lead role in the design of many aspects of SOFIA-TAF, especially the structure of SOFIA-TAF repositories, the fishstat R package, and the SOFIA R package.

SOFIA-TAF repository development

On the development front, Arni has created and maintained four demo repositories that serve as a reference, demonstrating the current and recommended format for SOFIA-TAF analyses.

Teaching and documenting

For the regional workshops in 2024, Arni gave presentations and participated in discussions to explain the overall design of SOFIA-TAF and how data files and R scripts perform the analysis of stock status. The goal is to enable regional experts to contribute the best data available to SOFIA-TAF.

All functions of the fishstat and SOFIA R packages are fully documented and updated with each release. Arni has also analyzed the extent of package dependencies of the sraplus package, which is especially relevant for the reproducibility aspect of SOFIA-TAF analyses.

R packages

Arni maintains the SOFIA package, acting as a single place of analytical methods used in all SOFIA-TAF analyses. This greatly improves the ability to manage and maintain the large number of SOFIA-TAF analyses behind SOFIA.

He also maintains the fishstat package, which may improve the input data pathway for catch data in terms of efficiency, traceability, and quality control.

Contributions to project management

With this report, updated annually, Arni has aimed to describe the current status and next steps in the development of SOFIA-TAF, especially the aspects relating to the analyses (data, scripts, and results) and the underlying SOFIA package. This conveys a design manifesto and has helped to track progress, set objectives, and provide an up-to-date documentation of how the SOFIA-TAF components work together.

Links to deliverables

- Demo analyses:
 - [WorkshopEffortShared](#)
 - [WorkshopEffortByStock](#)
 - [WorkshopIndexByStock](#)
 - [WorkshopPriorsByStock](#)
- Presentations:
 - Area 51 ([overview](#), [demo](#))
 - Area 77 ([overview](#), [demo](#))
 - Area 87 ([overview](#), [demo](#))
- [Documentation](#) page, including [sraplus dependencies](#) and [history](#)
- R packages:
 - [SOFIA](#)
 - [fishstat](#)
- [This](#) current report