



SOFIA-TAF

SOFIA analyses organized as TAF repositories on GitHub

Arni Magnusson, Rishi Sharma

FAO-GFCM Workshop

Rome, 25 May 2022

Introduction *Background, Objectives, Open science, Reproducible analyses*

Design *Modular, Data flow, Ongoing development*

Repositories *Browse, Download, Clone, Hierarchy*

TAF *Standard structure, Sequential R scripts, How to run*

R packages *SOFIA, TAF, srapius*

SOFIA analyses *Tier 2, Catch, Effort/Index, Priors, Demo*

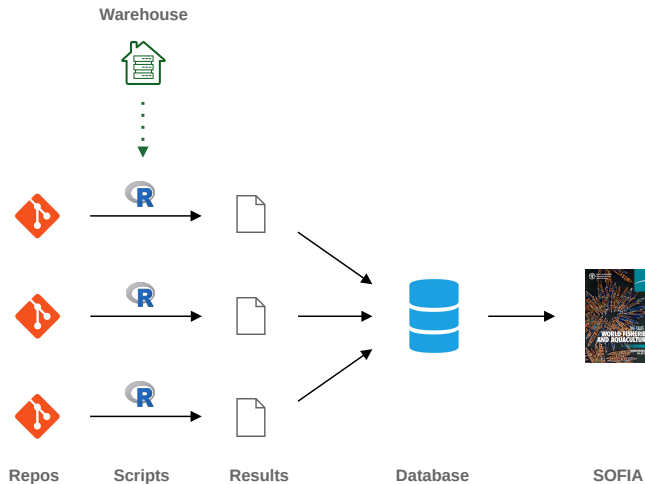
Objectives

1. **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.
2. **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.
3. **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.

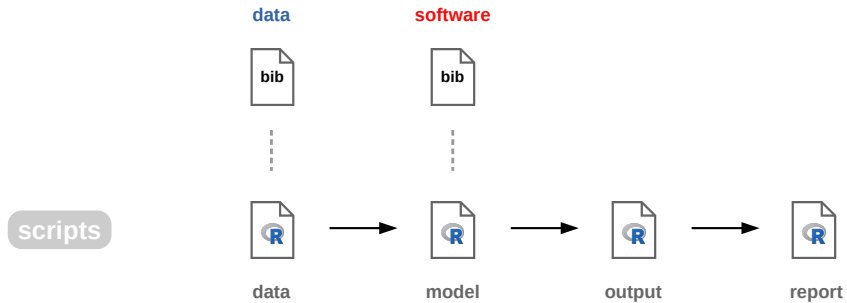
Objectives (cont)

4. **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.*
5. **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.*
6. **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.*
7. **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.*

SOFIA-TAF Diagram



TAF Diagram

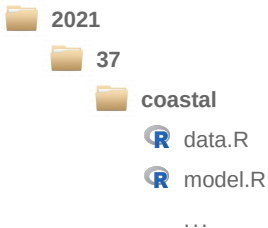


SOFIA-TAF Directory Structure

Remote repository

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

Local directory structure



TAF Scripts

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