Center for the Advancement of Population Assessment Methodology

Comisión Interamericana del Atún Tropical
Inter-American Tropical Tuna Commission

CAPAM

CIAT
IATTC

Workshop on Model Diagnostics in Integrated Stock Assessments

Virtual meeting, Jan 31-Feb 3, 2022 (9am to 1pm - San Diego)

Workshop goals

- Design an approach to diagnostics that is more objective, transparent, and automated
  - Identifying and defining candidate diagnostics
  - Create quantitative criteria for using each diagnostic
  - Evaluate diagnostics
    - Identify model misspecification is present
    - Identify what is misspecified
    - Recommend how to fix the misspecification
  - Automate the diagnostics

Topics

- General approaches
  - Residuals and effective sample size
  - Hind casting
  - Bayesian model checking
- Stock assessment specific
  - Retrospective analysis
  - R0 likelihood component profile
  - ASPM and catch-curve analysis
- Other
  - Process error (Random effects/state-space models) diagnostics
- Diagnostics in applications (e.g. Stock Synthesis)
- Simulation testing
- Automation

CAPAM Workshop series and special issues

- Natural Mortality: 2021
- Next Generation Stock Assessment Models: 2019
- Spatial stock assessment models: 2018
- Spatio-temporal modelling: 2018
- Recruitment: 2017
- Data Weighting: 2015
- Growth: 2014
- Selectivity: 2013

http://www.capamresearch.org/

CAPAM Stock Assessment Good Practices Workshop

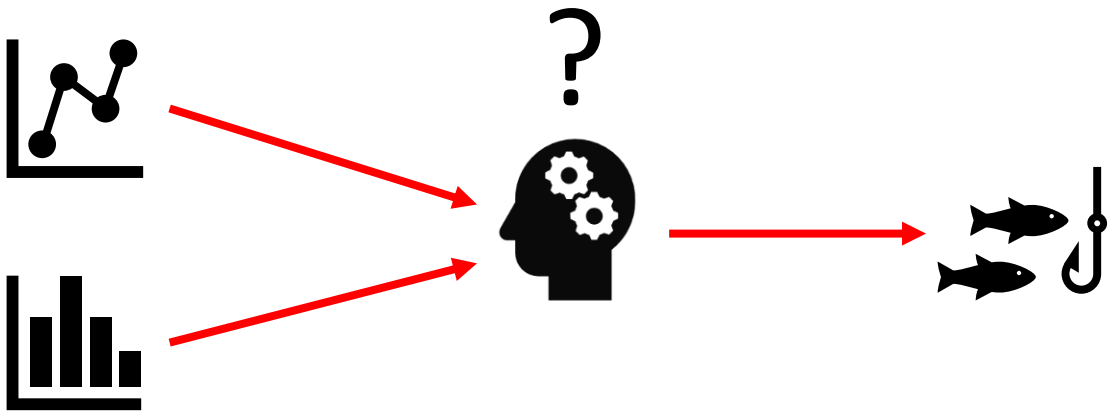
- **Location:** Rome (San Diego or Virtual as a backup)
- **Date:** Oct 31-Nov 4, 2022
- **Special issue:** Fisheries Research
- **More info:** http://www.capamresearch.org/GPG-Workshop

Key notes and commenters

- Why are good practices needed: Rick Methot
- A Good Practice Guide: Andre Punt
- Stock and fishery structure: Keynote-Steve Cadrin; Commenter-Ernesto Jardin
- CPUE standardization: Keynote-Simon Hoyle; Commenter-Shannon Cass-Calay
- Growth: Keynote-Kevin Piner/Hui-Hua Lee; Commenter-Kai Lorenzen
- Selectivity: Keynote-Rick Methot; Commenter-Ana Parma
- Natural mortality: Keynote-Owen Hamel; Commenter-John Hoenig
- Recruitment: Keynote-Liz Brooks; Commenter Carl Walters
- Data weighting: Keynote-Jim Thorson; Commenter Richard Hillary
- Random effects/state-space models: Keynote-Anders Nielsen; Commenter-Tim Miller
- Spatial stock assessment models: Keynote-Dan Goethel/Aaron Berger; Commenter- Collin Millar
- Diagnostics: Keynote-Felipe Carvalho/Henning Winker; Commenter-Laurie Kell
- Data limited: Keynote: Jason Cope; Commenter: TBA
- Integrated Population Models: Michael Schaub/Marc Kery

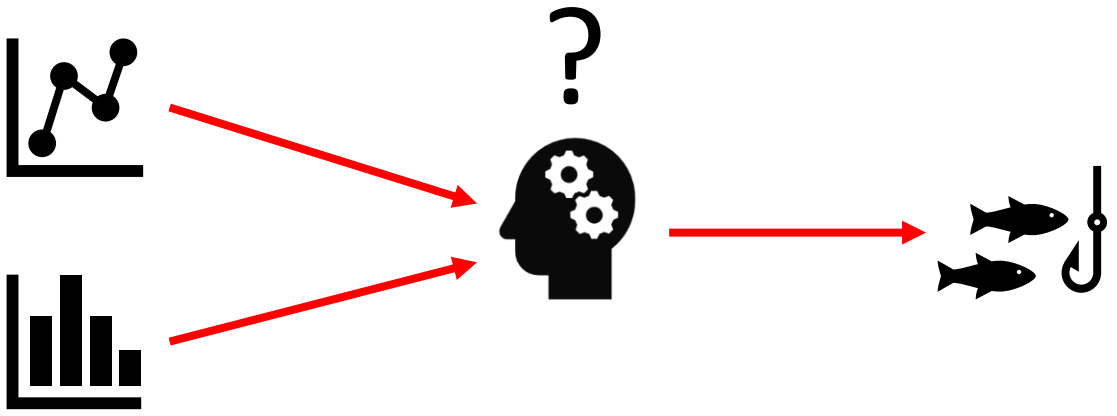
What is CAPAM’s purpose

Stock assessment

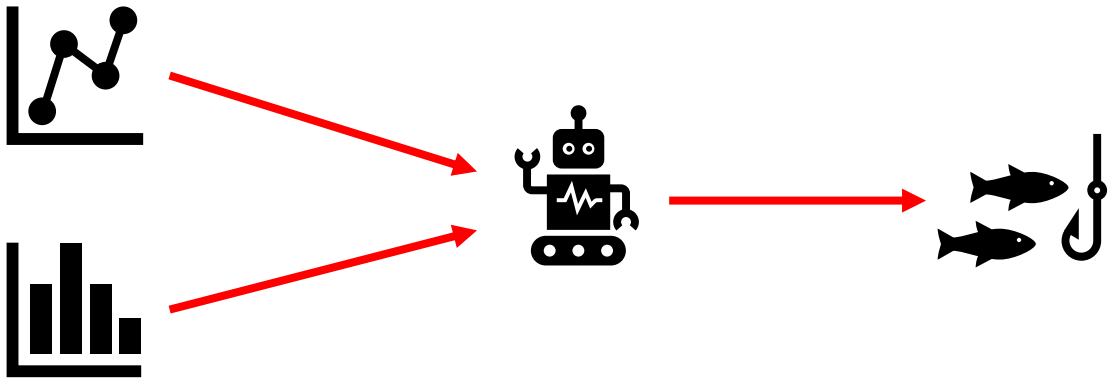


Stock assessment

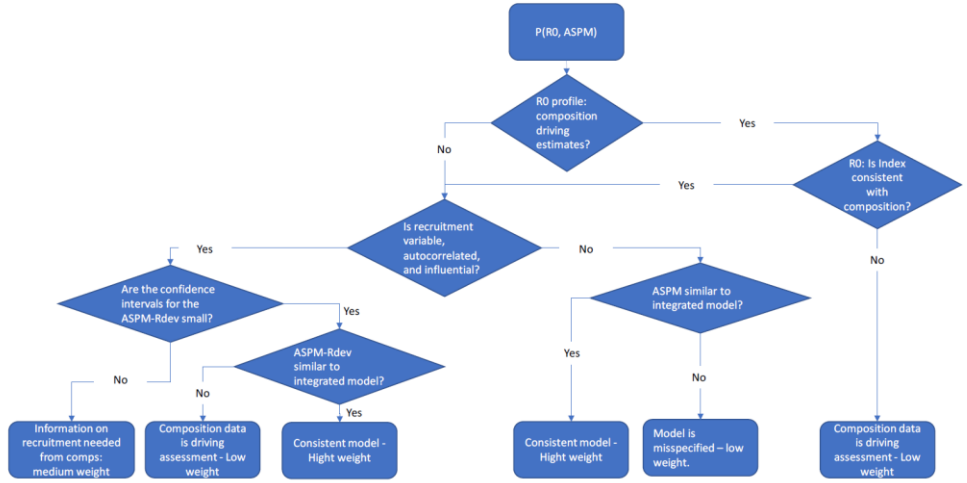
- High demand for assessments
- Not enough resources



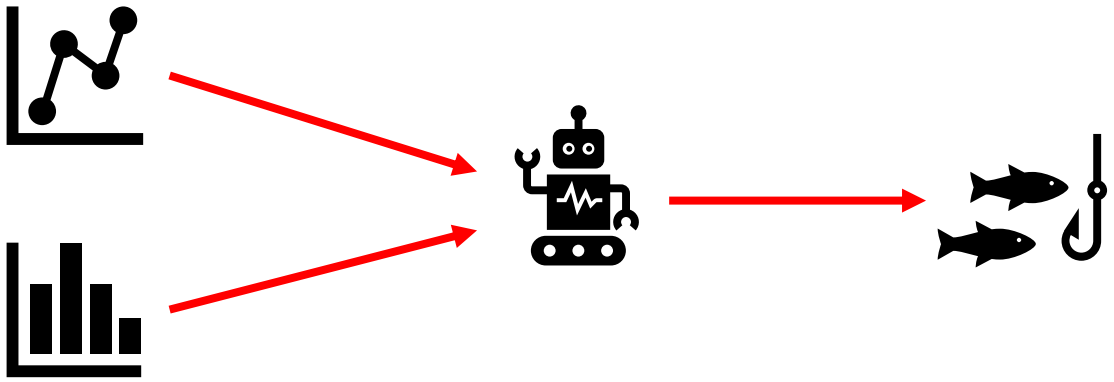
Automation



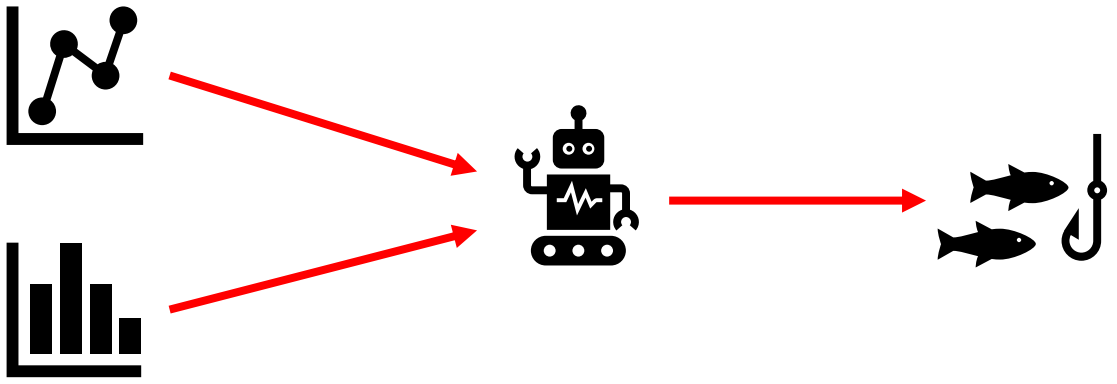
Expert System



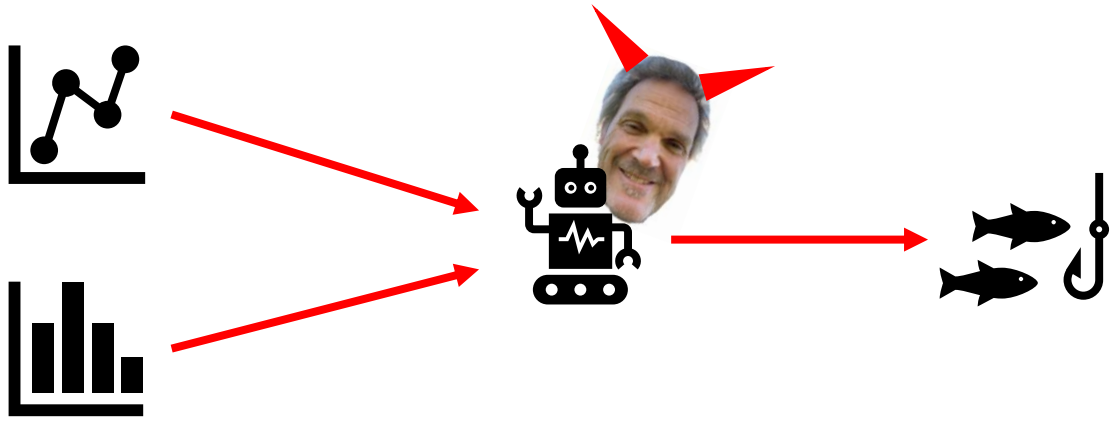
- All the rage in the early 1990s
- Basically, a decision tree
- Need to estimate the decision probabilities
- Did not have the automated learning (Machine learning)
- Can we do it for stock assessment?
- 30 years later we are still working on it



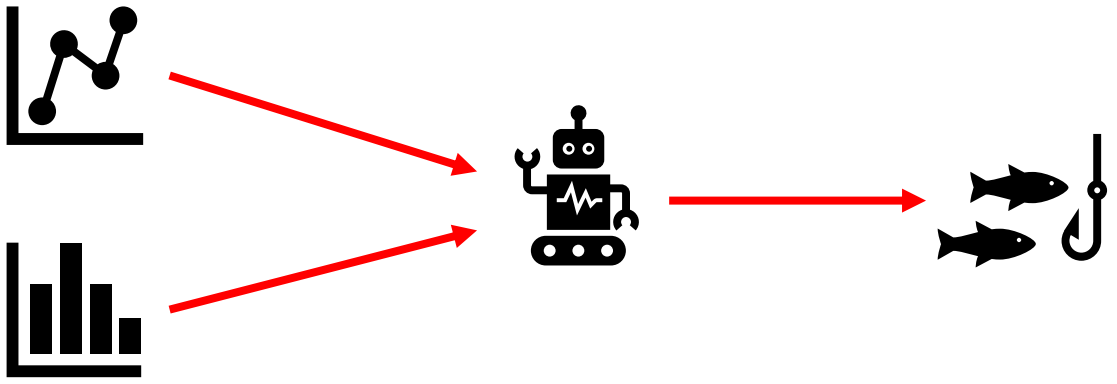
Expert System



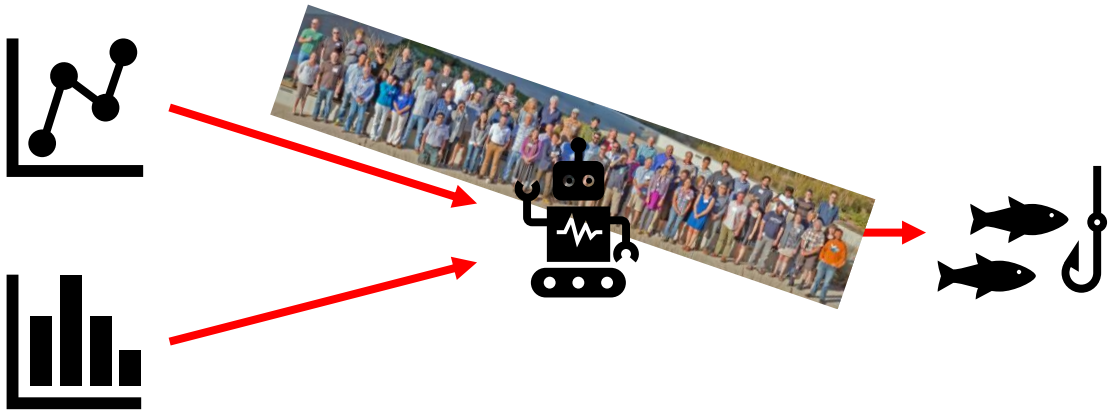
Expert System: Evil Dictator



Expert System

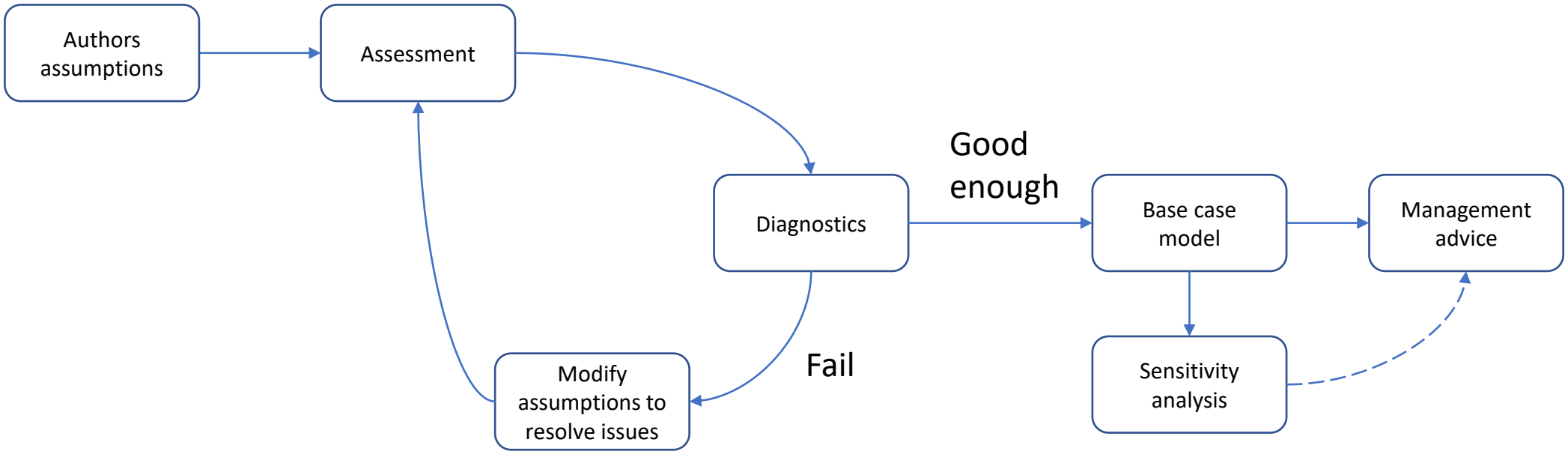


Expert System: CAPAM

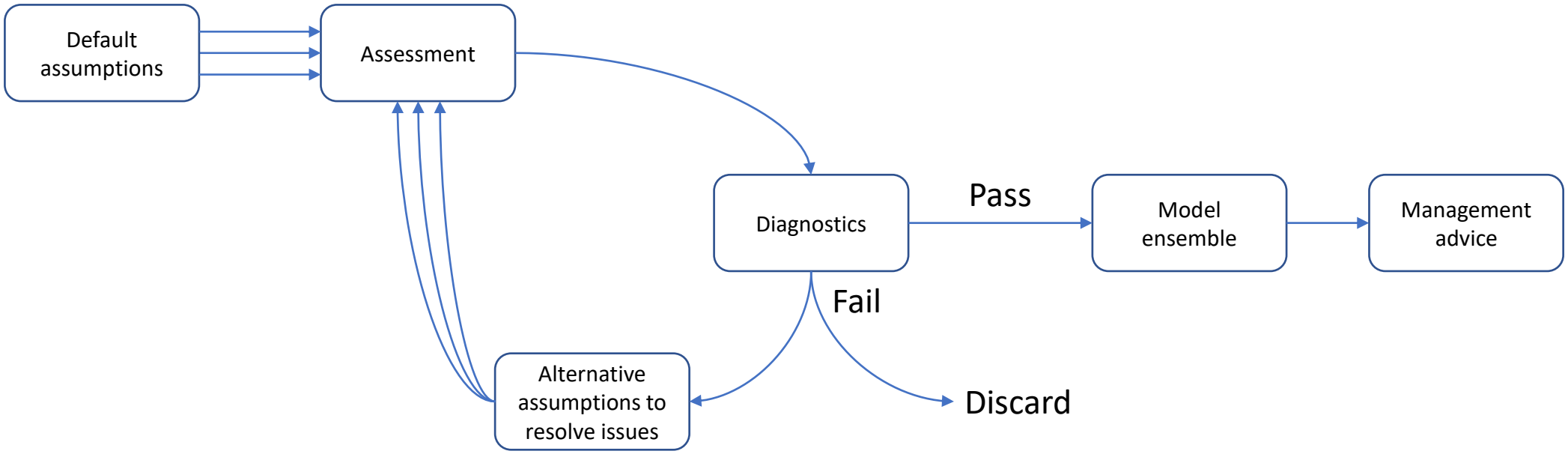


Where do diagnostics fit in

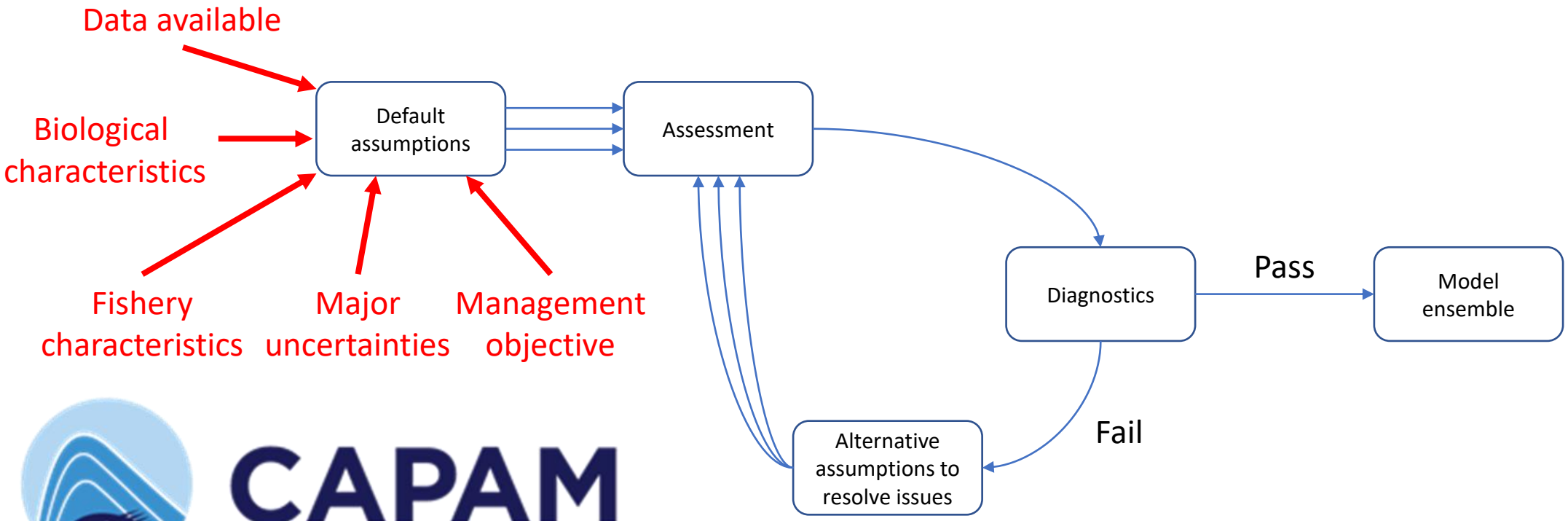
Typical current approach



Expert system to construct an ensemble of models for fisheries stock assessment

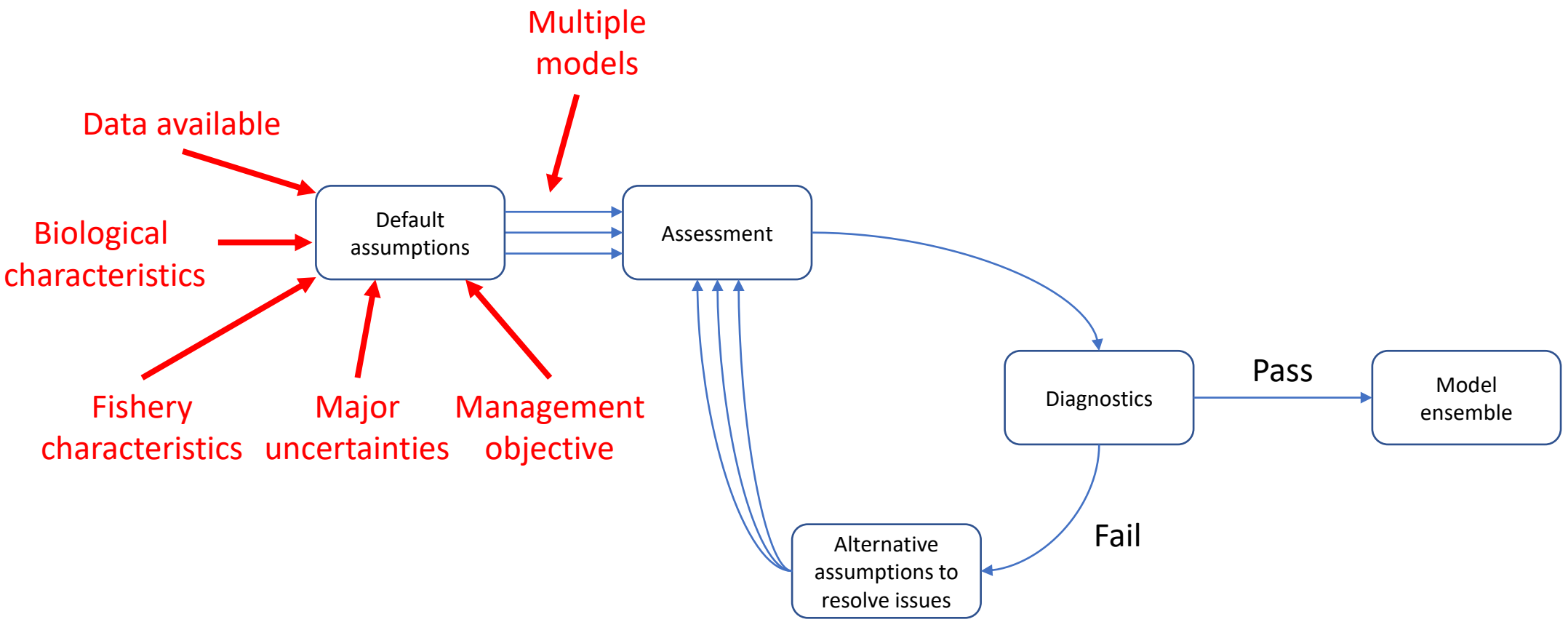


Default assumptions

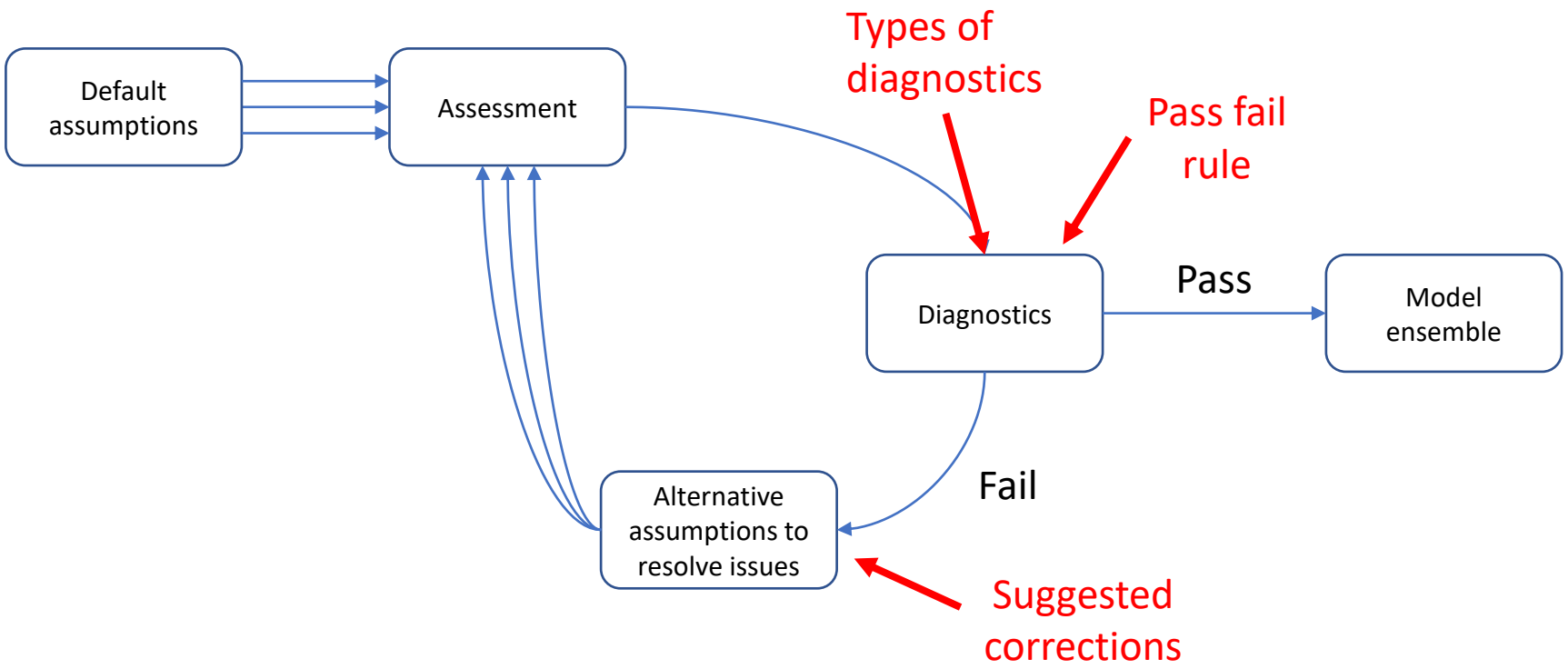


CAPAM

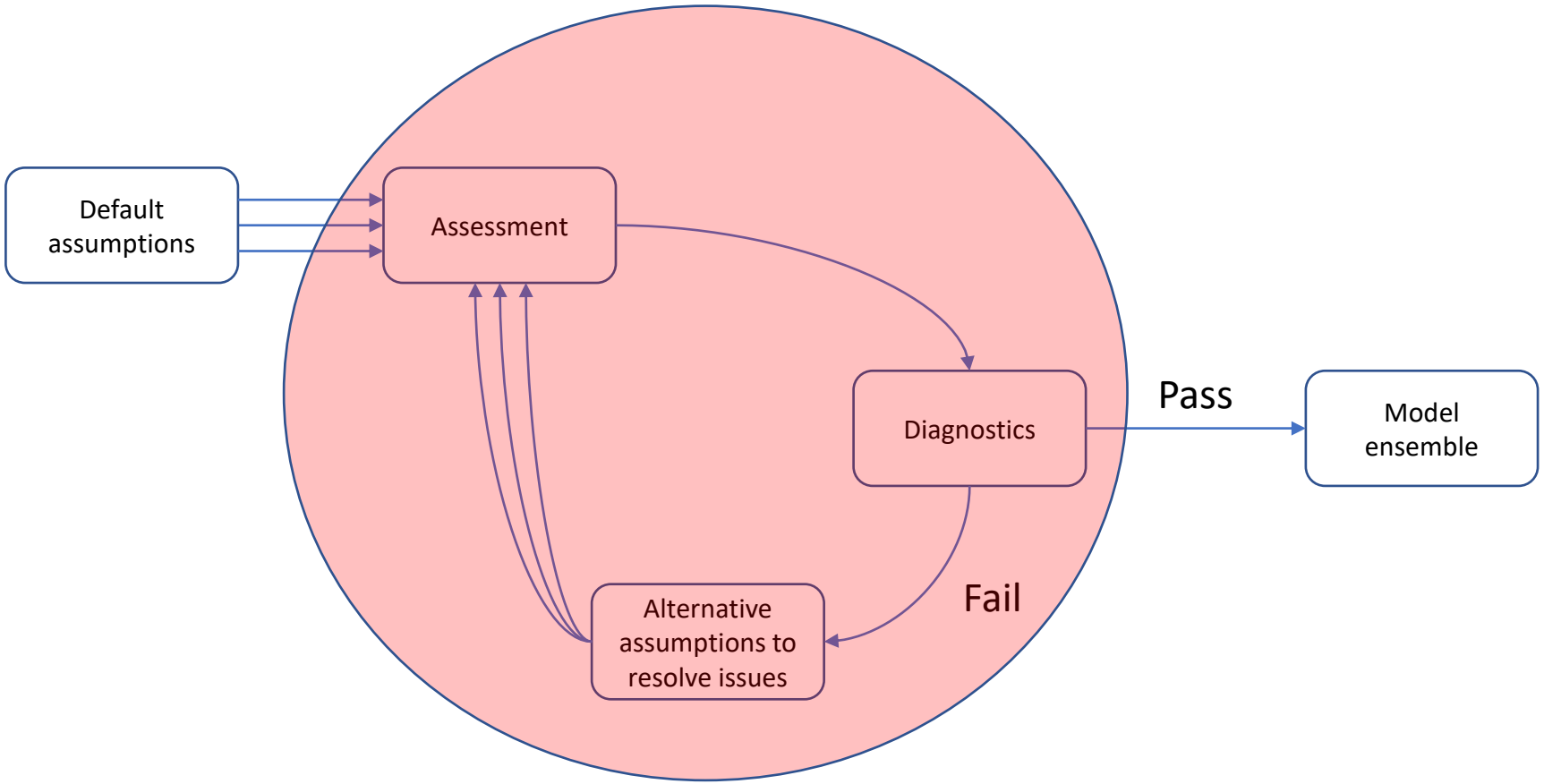
Default assumptions



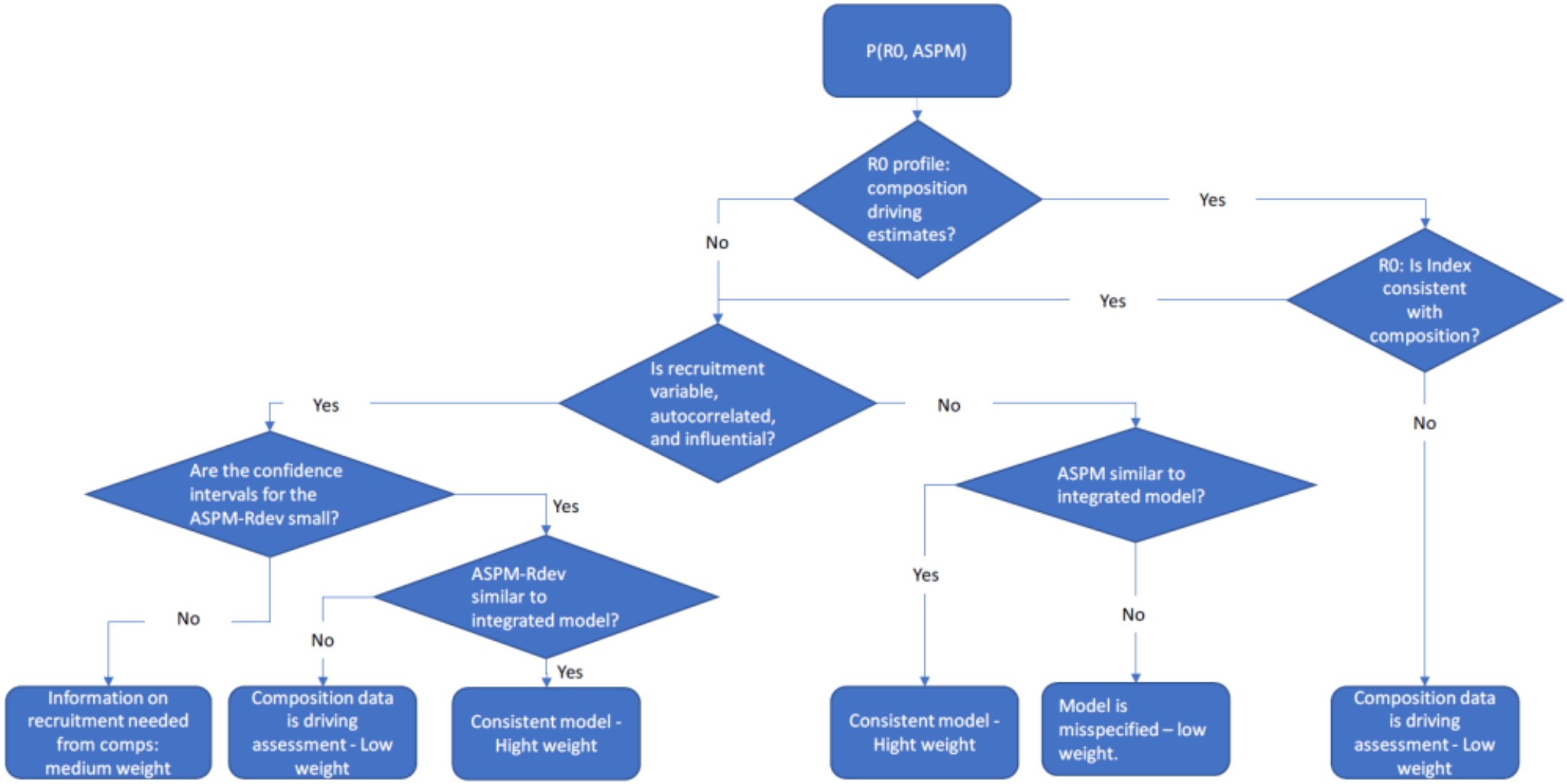
Diagnostics



Diagnostics



A decision tree is an expert system



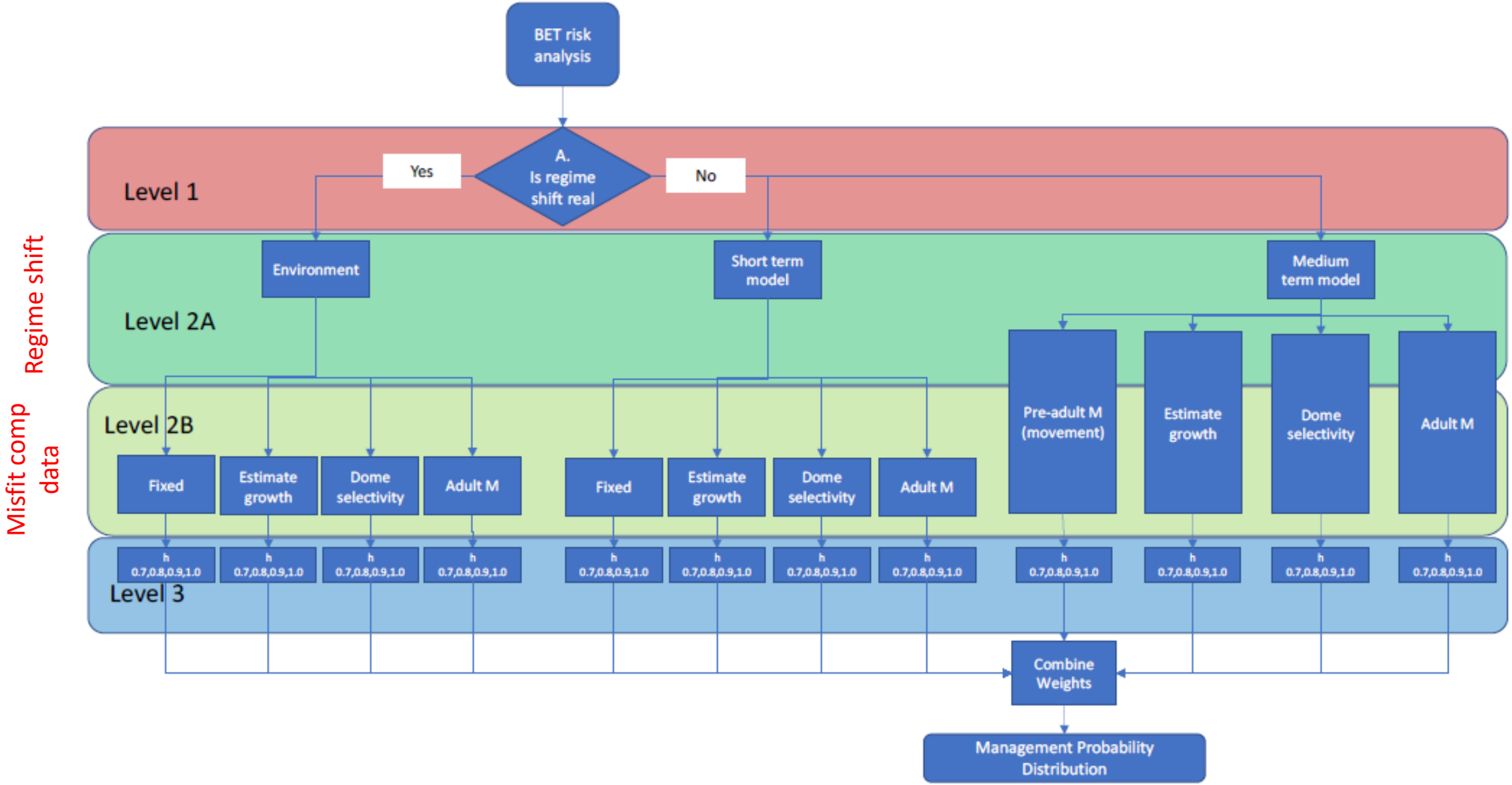
Maunder et al. 2020. Implementing reference point-based fishery harvest control rules within a probabilistic framework that considers multiple hypotheses. IATTC Document SAC-11 INF-F REV

Motivation: IATTC Decision analysis

Application: bigeye tuna decision analysis

- Conducted in Stock Synthesis
- Many fisheries
- CPUE and length composition data
- Issues
  A. Regime shift in recruitment when fishery on juveniles expanded
  B. Misfit to large fish in composition data from asymptotic fishery

Flow chart for bigeye tuna



Introduction: Main concept

- A rigorous statistical framework is not applicable
  - Multiple model assumptions are possible
  - Stock assessment models are complex and highly parameterized
  - Models are misspecified
  - Process variation is ignored
  - Data are not weighted appropriately
- Fit to data (e.g. AIC) should not be solely used to weight models

Weighting system: Weight metrics

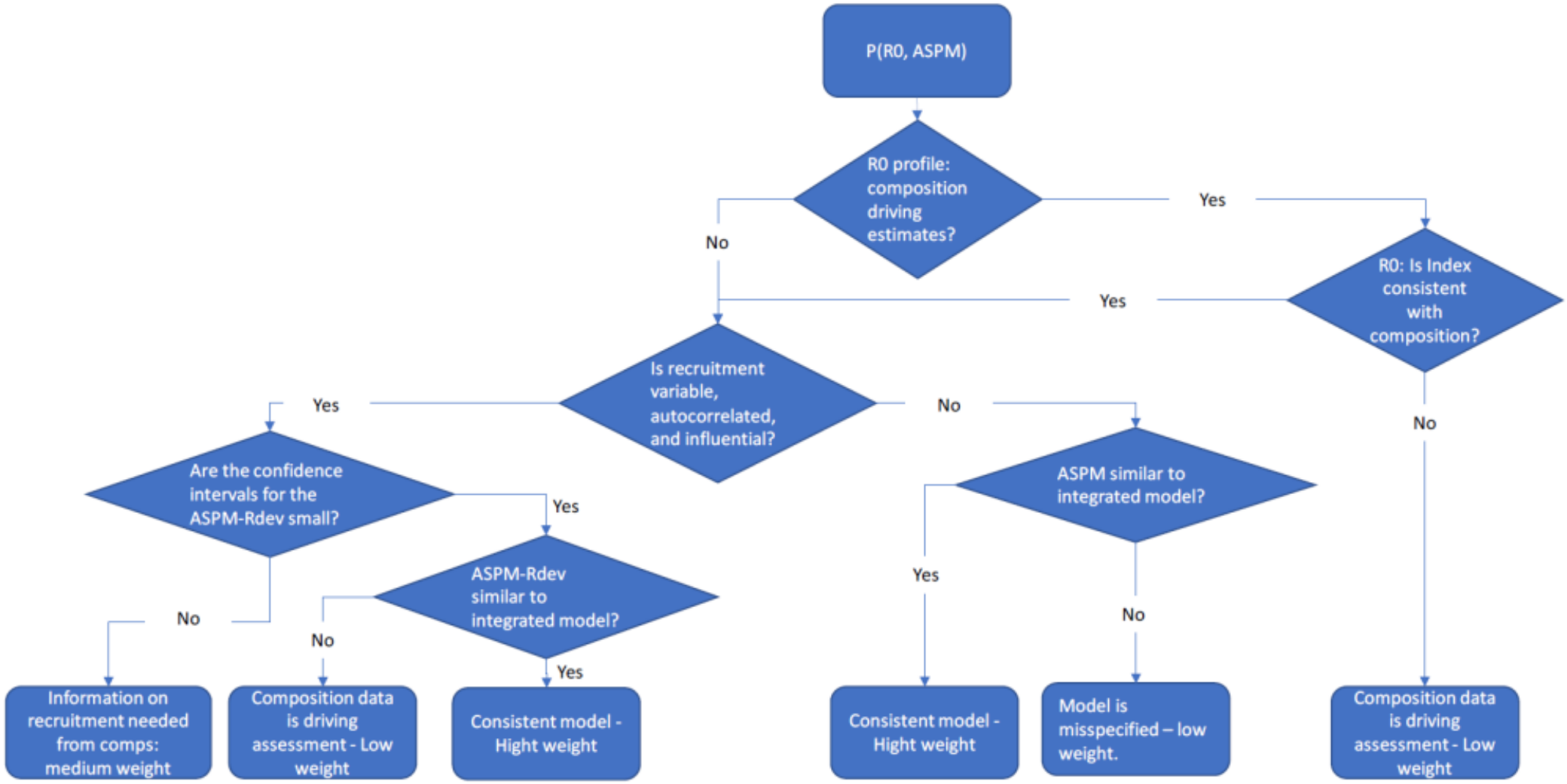
- W(Expert): Assigned “a-priori”, without consideration of model fit
- W(Convergence): Model convergence criteria of the estimation algorithm
- W(Fit): Fit of model to data
- W(Plausible parameters): Plausibility of estimates of parameters representing the hypothesis
- W(Plausible results): Plausibility of model results
- W(Diagnostics): Reliability of the model based on diagnostics

https://www.iattc.org/Meetings/Meetings2020/SAC-11/Docs/_English/SAC-11-INF-F_Implementing%20risk%20analysis.pdf

Diagnostics

- ASPM, R0, Catch curve
- Retrospective analysis
- Empirical selectivity
- Residual analysis
  - Composition residuals
  - Index residuals
  - Recruitment residuals

R0 and ASPM Diagnostics



Maunder et al. 2020. Implementing reference point-based fishery harvest control rules within a probabilistic framework that considers multiple hypotheses. IATTC Document SAC-11 INF-F REV

Weighting

- Subjective
- Panel of experts
- Differences in interpretation and scoring

Workshop goals

- Design an approach that is more objective, transparent, and automated
  - Identify and define candidate diagnostics
  - Create quantitative criteria for using the diagnostic
  - Evaluate diagnostics
    - Identify model misspecification is present
    - Identify what is misspecified
    - Recommend how to fix the misspecification
  - Automate the diagnostics

What is a diagnostic

Data weighting: The Law of Conflicting Data

Axiom

Data are facts

Implication

Conflicting data implies model misspecification

Caveat

Data conflict needs to be interpreted in the context of random sampling error

Significance

Down weighting or dropping conflicting data is not necessarily appropriate because it may not resolve the model misspecification

---

Maunder, M. N. and Piner, K. R. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fisheries Research. 192: 16-27.

Diagnostics: The Law of conflict between model and data

~~Data weighting~~: ~~The Law of Conflicting Data~~

Axiom

Data are facts

and conflicts between the model and the data

Implication

Conflicting data implies model misspecification

Caveat

~~Data~~ conflict needs to be interpreted in the context of random sampling error

Significance

Down weighting or dropping conflicting data is not necessarily appropriate because it may not resolve the model misspecification

---

Maunder, M. N. and Piner, K. R. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fisheries Research. 192: 16-27.

Agenda

- **Monday**
- Introduction
- Residuals and effective sample size
- **Tuesday**
- Retrospective analysis, hindcasting, Bayesian Model Checking
- **Wednesday**
- R0 profile, ASPM, catch curve analysis, empirical selectivity
- Application
- **Thursday**
- Automation and simulation testing
- Summary

Agenda: Monday

- 9:00 Welcome and instructions, Alex Aires-da-Silva and Mark Maunder
- 9:10 Introduction, Mark Maunder
- 9:30 The value of diagnostics in stock assessment, Felipe Carvalho
- 10:30 “and he saith unto them, Follow me, and I will make you fishers of [data]” (i.e., how to sort and weigh data). James Thorson.
- 11:00 Break
- 11:15 The Logistic-normal as a tool to diagnose model misspecification? The proposed idea, its comparison to common diagnostics, and some initial considerations. Nicholas Fisch
- 11:45 Guidelines to validating generalized linear mixed models in Template Model Builder using quantile residuals. Andrea Havron
- 12:15 Discussion

Key concept of CAPAM workshops

Discussion