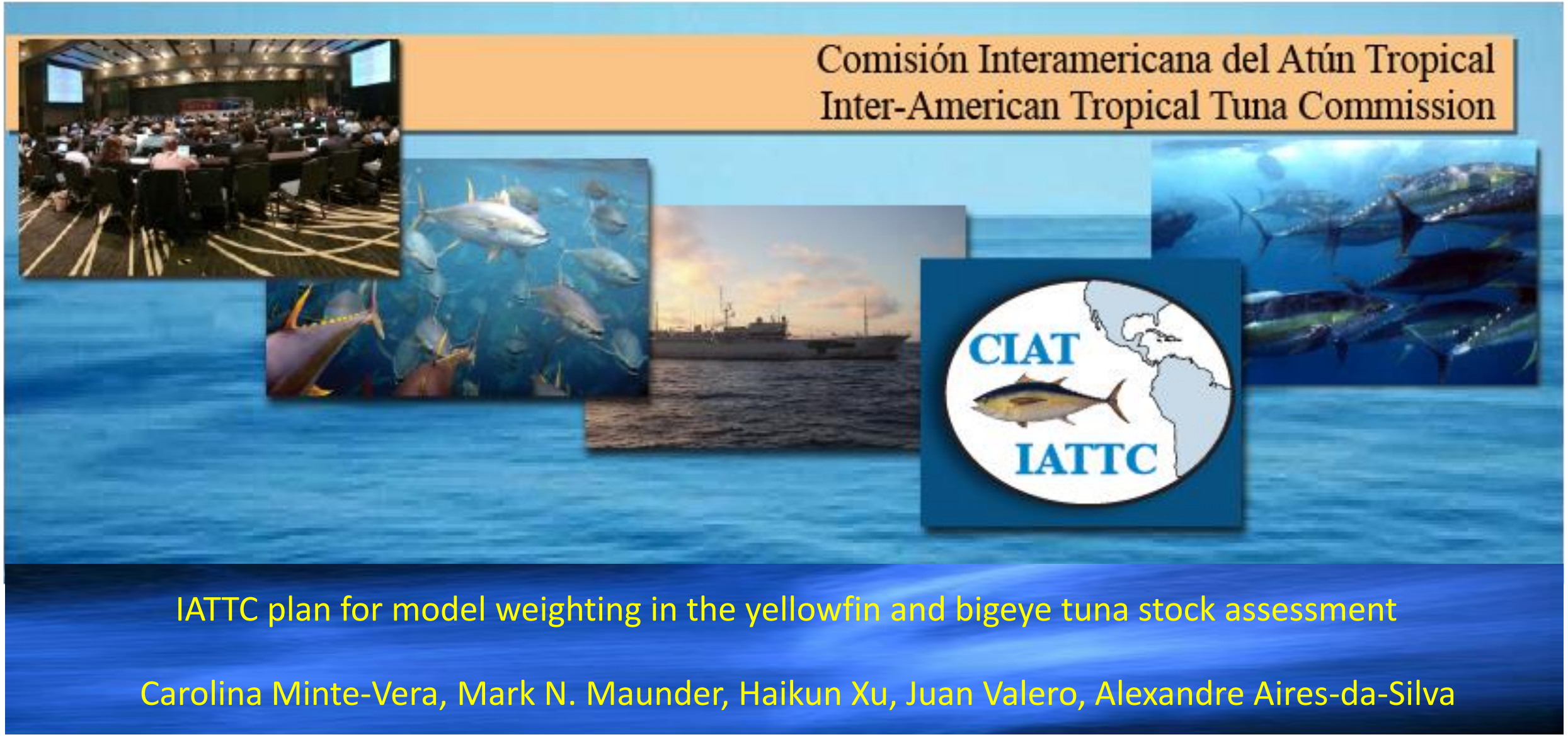


IATTC plan for model weighting in the yellowfin and bigeye tuna stock assessment

Carolina Minte-Vera, Mark N. Maunder, Haikun Xu, Juan Valero, Alexandre Aires-da-Silva

2nd Workshop on improving the risk analysis for the tropical tunas in the eastern Pacific Ocean: model weighting for integrated stock assessments
(by videoconference)

Outline

- Recap 2020 risk analysis approach
- Proposal for updated risk analysis approach

The 2020 risk analysis approach

Described in Maunder et al. 2020 (SAC-11- INF-F):

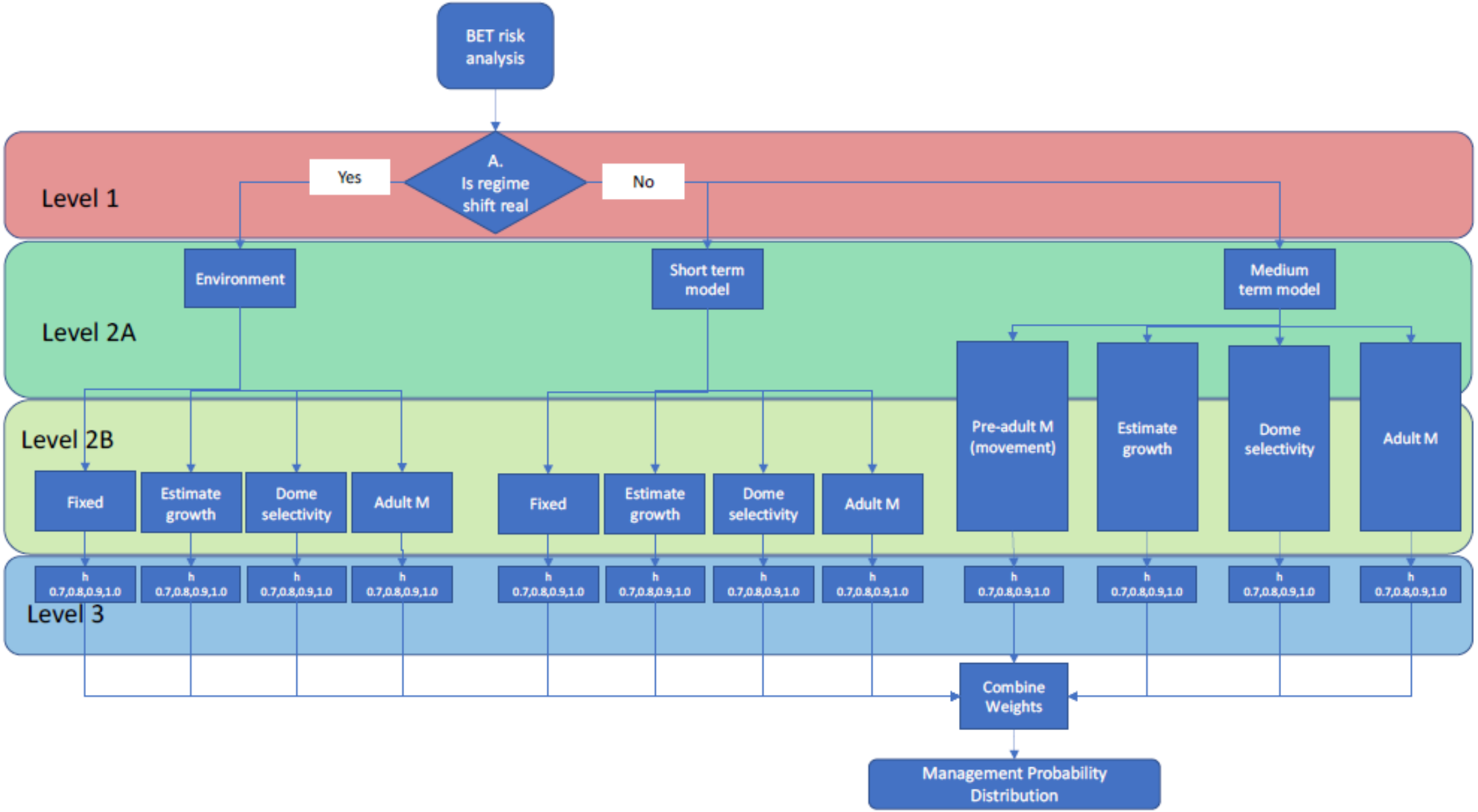
1. **Identify alternative hypotheses** (‘*states of nature*’) about the population dynamics of the stock that address the main issues in the assessments

2. **Implement stock assessment models representing alternative hypotheses**

3. **Assign relative weights to each hypothesis (model)**

4. **Compute combined probability distributions for management quantities using model relative weights**

The 2020 flow chart for bigeye tuna



The 2020 method for assigning model weights

- Level 1 (regime shift hypothesis) is weighted independently solely on experts opinion
- Level 2 is weighted based on several criteria:
  - Expert opinion
  - Convergence
  - Fit to data
  - Plausible parameter estimates
  - Plausible model results
  - Model diagnostics
  - Recruitment shift metric
  - Empirical selectivity vs. estimated selectivity
- Level 3 (steepness hypothesis) is weighted independently solely on experts opinion

The 2020 method for assigning model weights

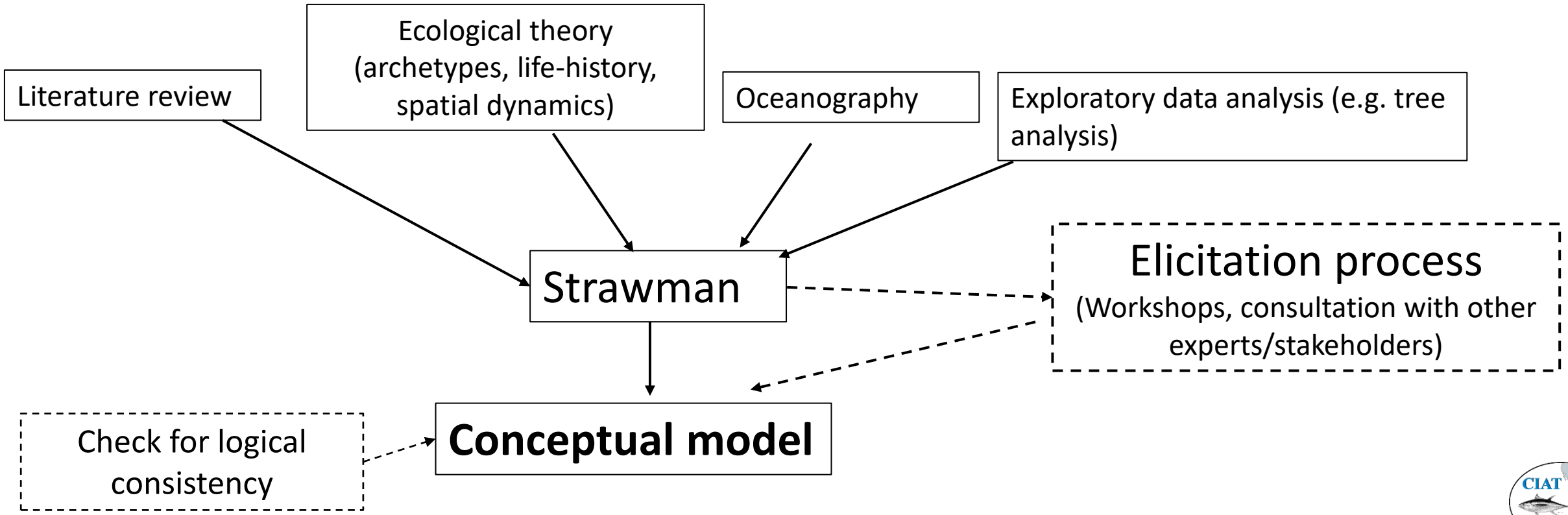
- Level 1 (regime shift hypothesis) is weighted independently solely on experts opinion
- Level 2 is weighted based on several criteria:
  - Expert opinion
  - Convergence
  - Fit to data
  - Plausible parameter estimates
  - Plausible model results
  - Model diagnostics
  - Recruitment shift metric
  - Empirical selectivity vs. estimated selectivity
- Level 3 (steepness hypothesis) is weighted independently solely on experts opinion

Proposal for new approach

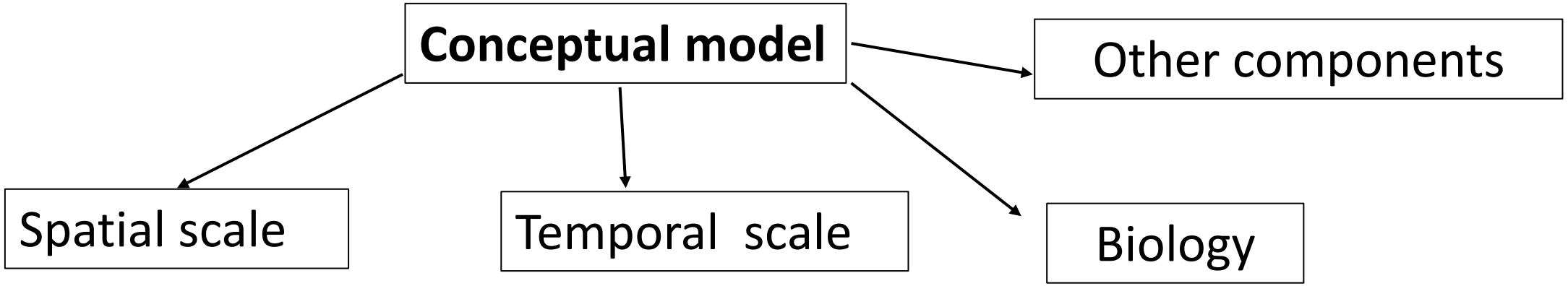
1. Develop a **conceptual model** for the fisheries system:

Proposal for new approach

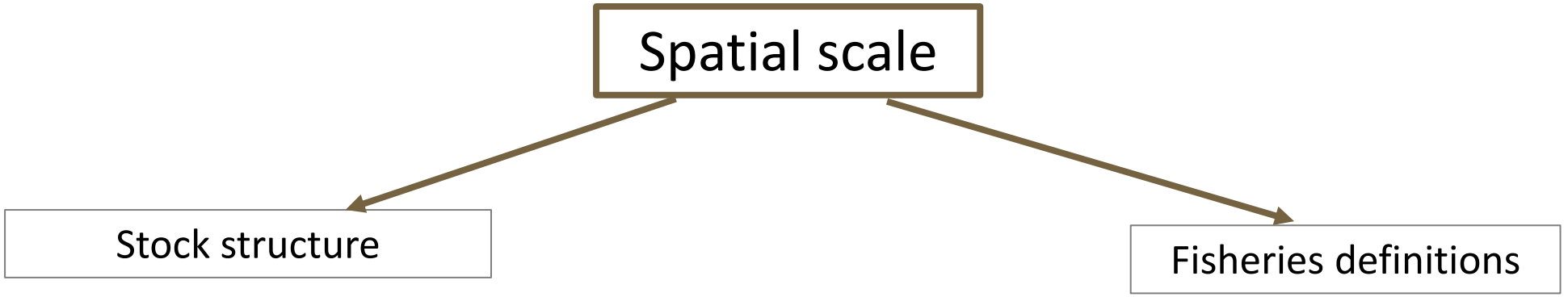
1. Develop a **conceptual model** for the fisheries system:



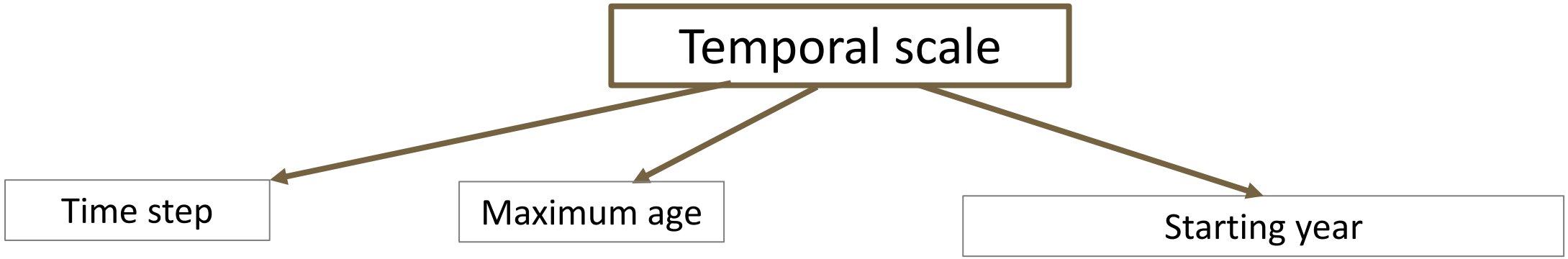
Develop a conceptual model



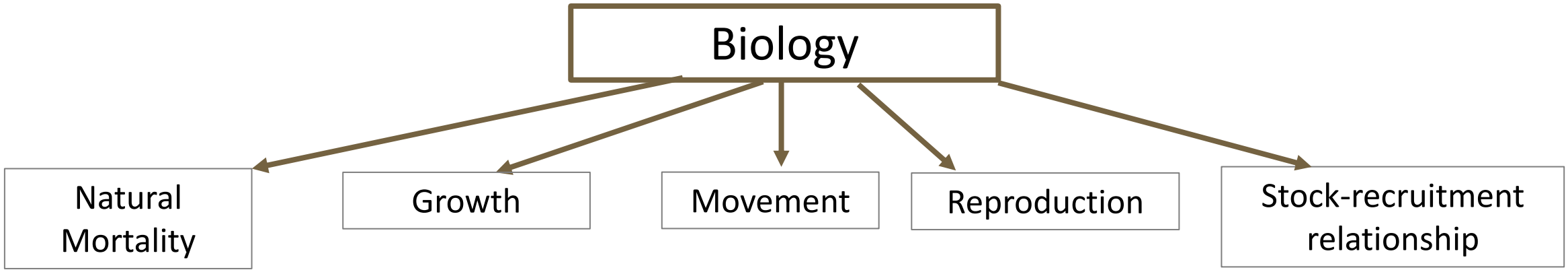
Develop a conceptual model



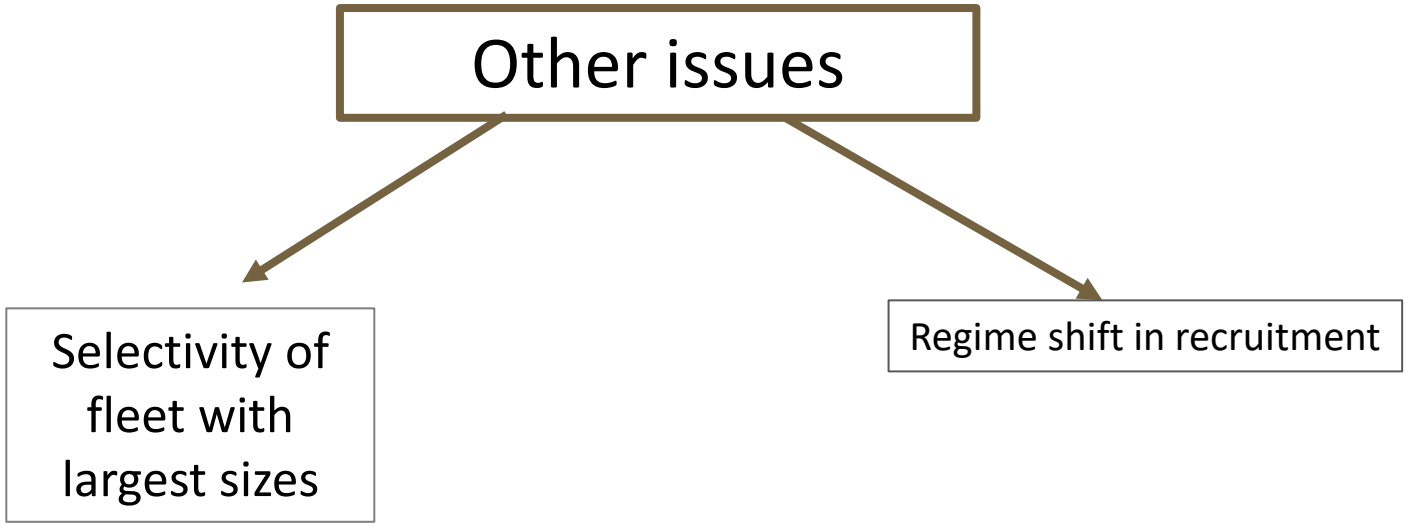
Develop a conceptual model



Develop a conceptual model



Develop a conceptual model



Proposal for new approach

1. Develop a **conceptual model** for the fisheries system:
   - Include justifications for assumptions/hypotheses
   - Use an elicitation / consultation process if needed
   - Organize hypotheses in a hierarchical way
   - Identify independent (orthogonal) uncertainty axes
     - Parameters that will have different fixed values (Level 3 hypotheses) tried for other hypotheses
   - Identify processes (parameters) with no information in the data used to fit the model
     - Use knowledge from published simulation studies to understand whether a parameter is estimable (e.g. Lee et al on M and h)
     - Perform simulation to investigate estimability of the parameter on the fishery system under study
     - Identify processes (parameters) that can be informed by external data:
       - Include prior knowledge as i) fixed value, ii) parameter range, iii) prior distribution, iv) joint prior distribution for correlated processes (e.g. Natural mortality and growth)
       - Decide whether to include this info in the estimation or in the weighting
   - Identify practical and impractical hypotheses (can they be implemented given the available data/tools?)
     - Impractical hypotheses will have weight of 0, and considered in the future when resource become available
   - Check the conceptual model for logical consistency

Proposal for new approach

2. Implement stock assessment models representing alternative hypotheses

- Implement a base or “ancestral” model(s) from which other models will be developed:
  - Models can use different data or have completely different structure according to the hypothesis they represent
- Run a set of diagnostics to learn about the base /“ancestral” model(s) (decide in advance what those diagnostics will be)
  - Improve base or “ancestral” models/change conceptual model if issues arise that merit new hypotheses
- Decide on a candidate set of reference models (and decide in advance which diagnostics to used, noting some diagnostics are for model understanding while others are for model validity *)

A. Large set of models that hopefully covers all the models to improve diagnostics

B. Smaller set of curated models to represent specific hypotheses:
   - Use diagnostics to improve the models or discard them (how or how much to improve models?)
   - Fit all reference models to data:
     - Data will need to have its own quality control rules to be able to be included in the model
- Arrive at a final list of models:
  - The final list will be more or less inclusive according to the diagnostics used

Proposal for new approach

3. Assign relative weights to each hypothesis (model)

- Decide in advance the weighting scheme based on the objectives of the assessment, which may include in this order:
  - Use equal weight as default or as alternative
  - Decide prior weight of overarching hypotheses that cannot be evaluated by the application of models (e.g. stock structure)
  - For independent processes not informed by the data (e.g. steepness, movement rates):
    - use prior distributions (e.g. from meta-analyses, expert opinion or other) as weights , if they where not used in the estimation
  - Use predictive ability (hindcast) to weight models:
    - Ideally the predicting ability should be estimated on the quantity of interest
    - Quantities of interest in stock assessment are in general not directly observable
    - Observable quantities are used and should be chosen carefully (prefer those that are proxies for quantities of interest)
  - Weight nested hypotheses/models with conditional weights (sum to 1 within a branch)
- Final weights consider the prior weight of overarching hypotheses, independent process weights, weight based on hindcast and the position of the model in the hierarchy

Proposal for new approach

4. Compute combined probability distributions for management quantities using model relative weights

- Use probability staking rather than model averaging to estimate the tails of the distribution
- To compute the final probability distribution by:
  - Normal approximations based on the estimate and standard error
  - Some standard errors are approximated
  - The resulting distribution is rescaled to obtain P(Quantity|Model=m).
  - Works well when the data is very informative
  - Probability distribution may be asymmetrical
  - Evaluate appropriateness of the approximation using Posteriors derived from limited MCMC analyses

List of model diagnostics (Pass/Fail can be used to improve models as well)

Standard diagnostics

- **Pass/Fail** Model convergence: e.g. size of the gradient at the MLE, Hessian matrix, jitter analysis.
- **P/F** Check for parameters in the bounds
- **P/F run test** Evaluation of residuals: need to take into account the assumed distribution, use PIT residuals for composition data
- **Improve** Effective sample sizes and variances
- **Weighting** Cross validation and **hindcasting**
- **P/F** Bayesian model checking
- **P/F** Retrospective analysis: The interpretation is context dependent. When the data is more informative at the end of the time-series may result in pattern. If we have no retrospective pattern there is no learning from the data in the tails.

Stock Assessment specific

- **P/F** R0 likelihood component profile
- **Improvement** Age-structured Production Model (ASPM)
- **P/F** Catch curve analysis
- **P/F** Empirical selectivity
- **Improvement** Patterns of process error (*e.g.* recruitment residuals, but other process error may need to be checked too such as time-varying selectivity)

Plausibility

- **P/F** Parameter values (e.g. smaller F for highly exploited stocks)
- **P/F** Results
- **Not used** Projection with current catches (if the system cannot support current catches in the future, the production function may be implausible)

Notes

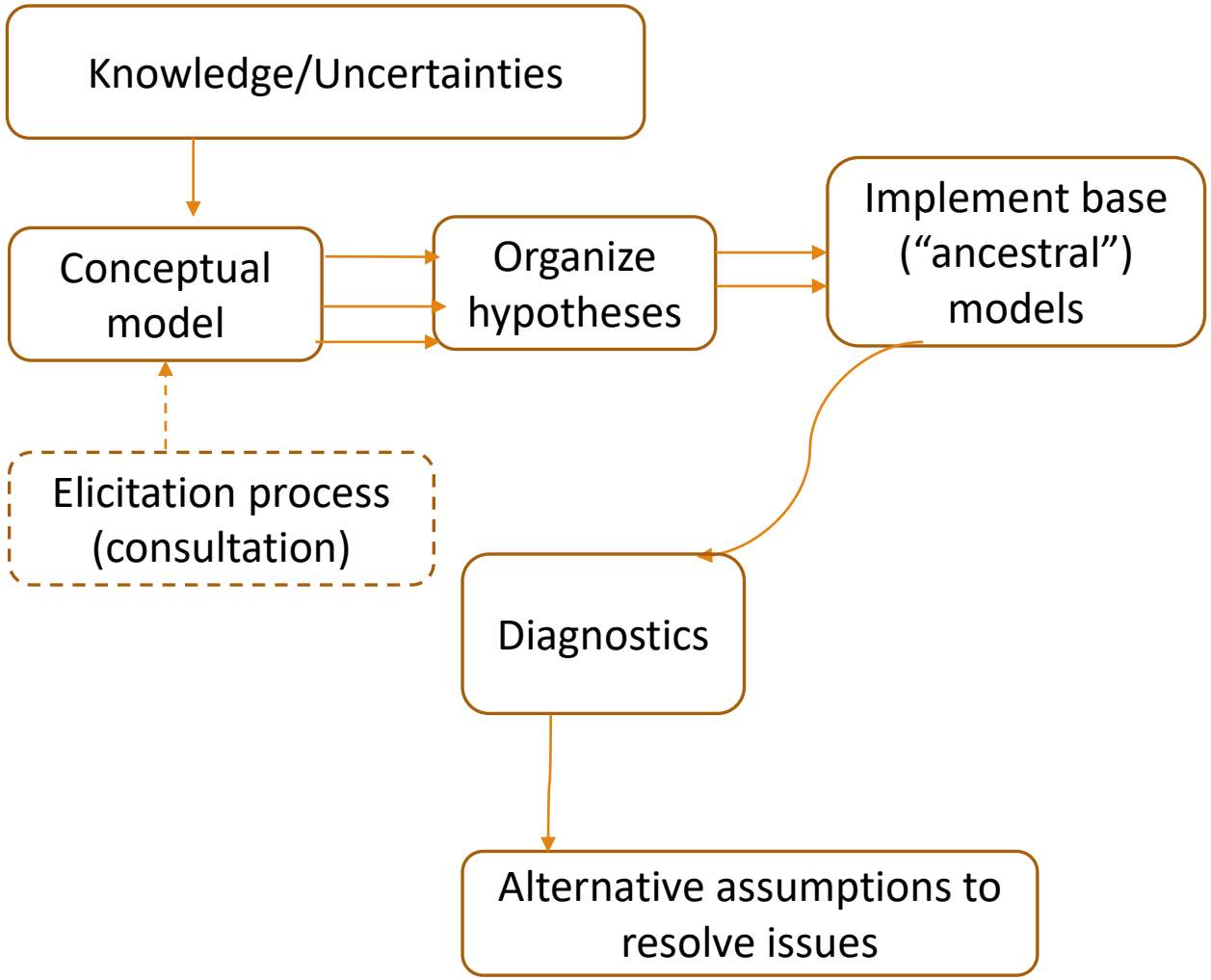
- Values for pass/fail or criteria for each diagnostics need to be worked out
- If the model fails one of the P/S diagnostics, it will be excluded
- Investigate emergent proprieties, what are the combination of parameters that cause the model to break

Hindcasting

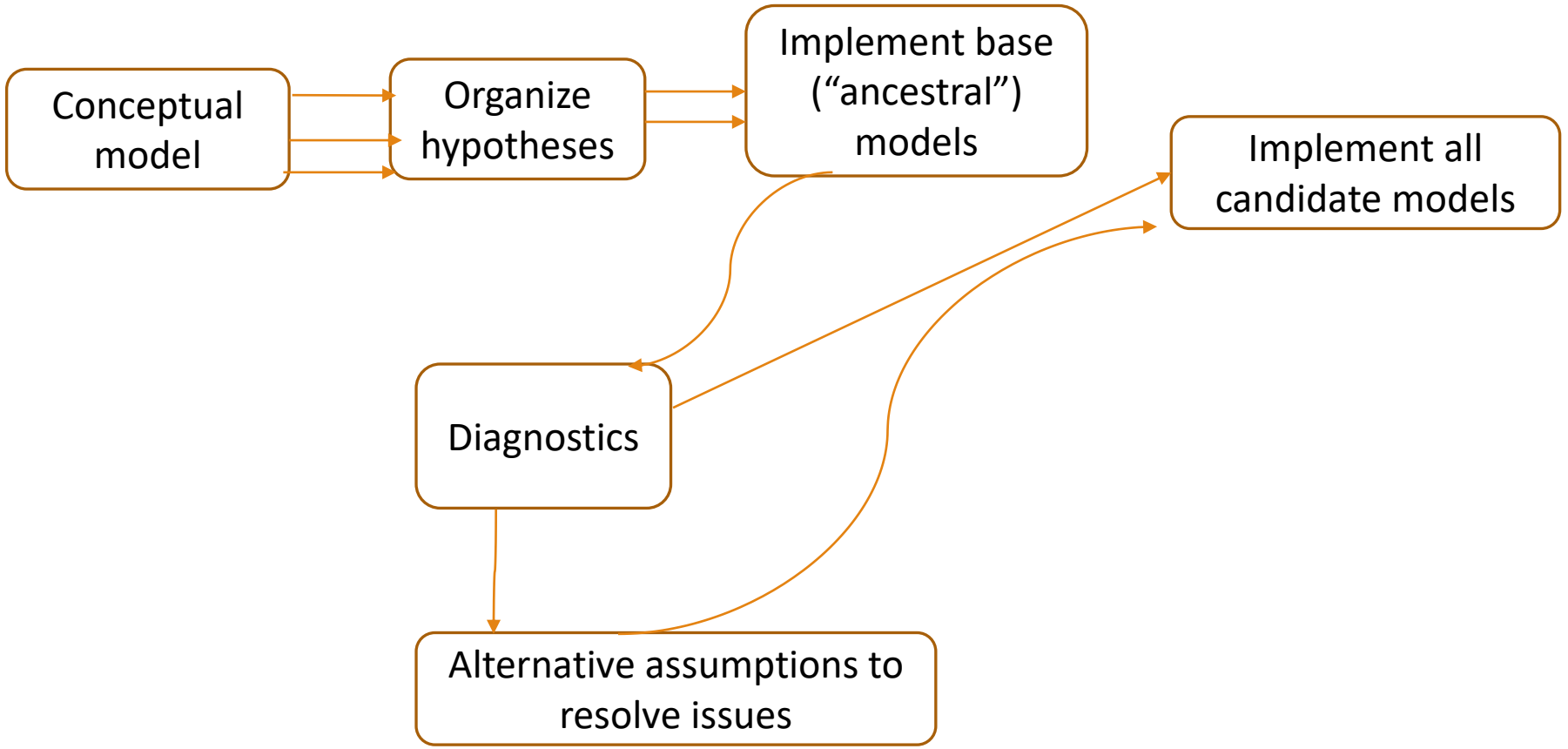
Decisions to make (in advance):

- Need some simulation studies to make informed decisions
- What to predict:
  - index of abundance
  - average size/quantile in the survey/index (if available)
- For how long?
  - One set ahead in the time step of the model (quaterly)
- What data to exclude?
  - All the data in the year except the catches or effort
- How many peels?
  - Minimum of 10 quarters (2.5 years)

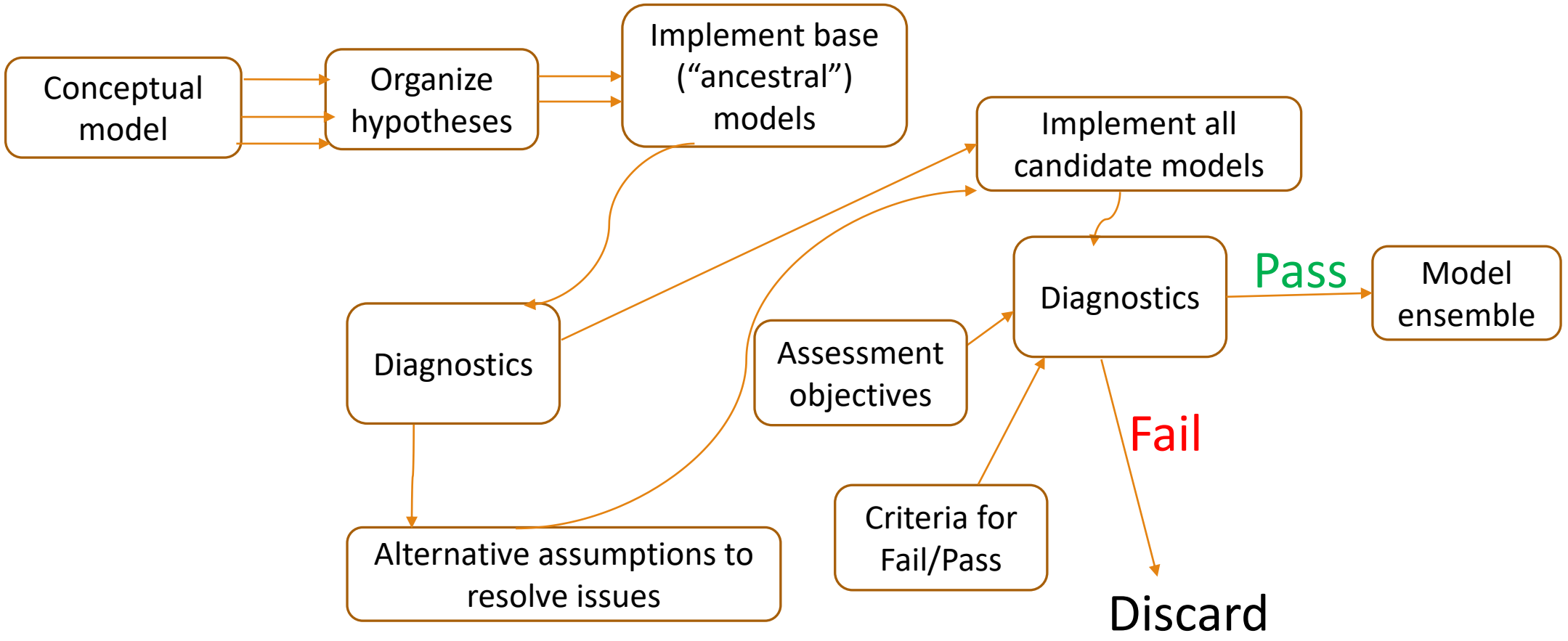
Summary: Ensemble model approach



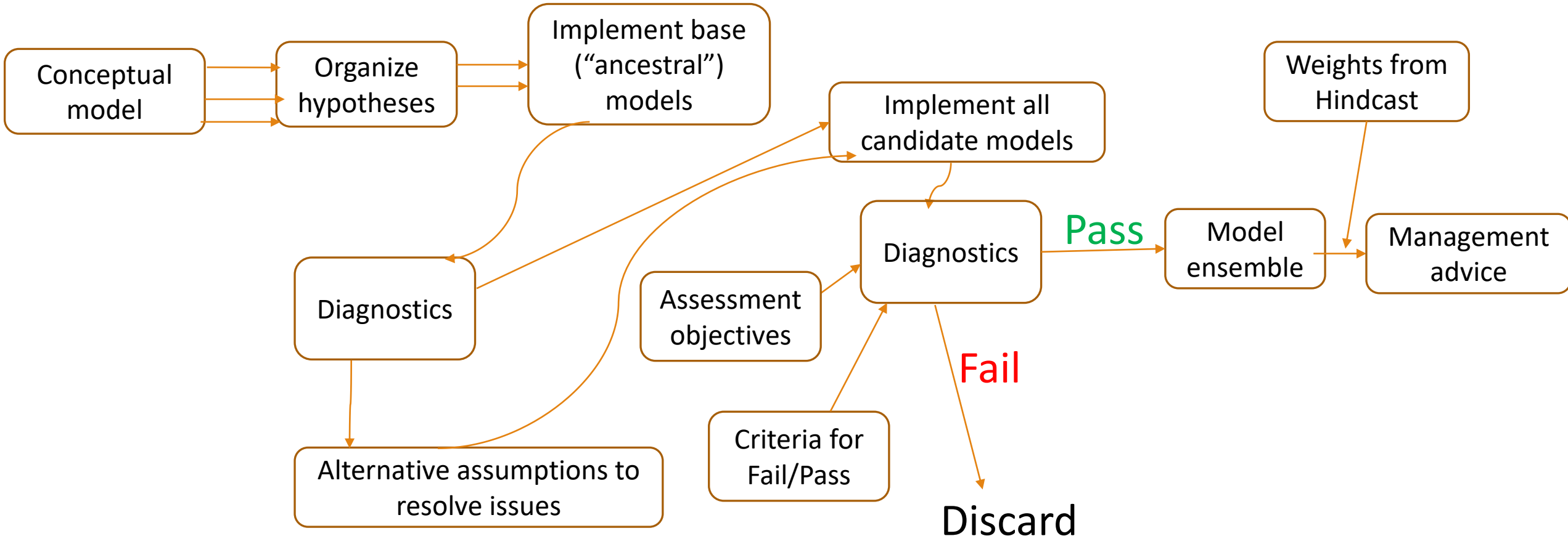
Summary: Ensemble model approach



Summary: Ensemble model approach



Summary: Ensemble model approach



CIAT

IATTC

Questions