

Comisión Interamericana del Atún Tropical
Inter-American Tropical Tuna Commission



IATTC Risk Analysis

Mark N. Maunder, Haikun Xu, Cleridy E. Lennert-Cody, Juan L. Valero, Alexandre Aires-da-Silva, Carolina Minte-Vera

Virtual meeting, 28 Nov – 2 Dec 2022

Introduction: Why we need a risk analysis

- Assessments are uncertain
- IATTC HCR for tropical tunas (Resolution C-16-02) addresses uncertainty through probability statements
 - “if the probability that F will exceed the limit reference point (FLIMIT) is greater than 10%, as soon as is practical management measures shall be established that have a probability of at least 50% of reducing F to the target level (FMSY) or less, and a probability of less than 10% that F will exceed FLIMIT.”
- Evaluations
 - Current status relative to reference points
 - Status under different management scenarios
- Transition from single base-case assessment to set of reference models

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
- Data should not be solely used to weight models

Introduction: Main features

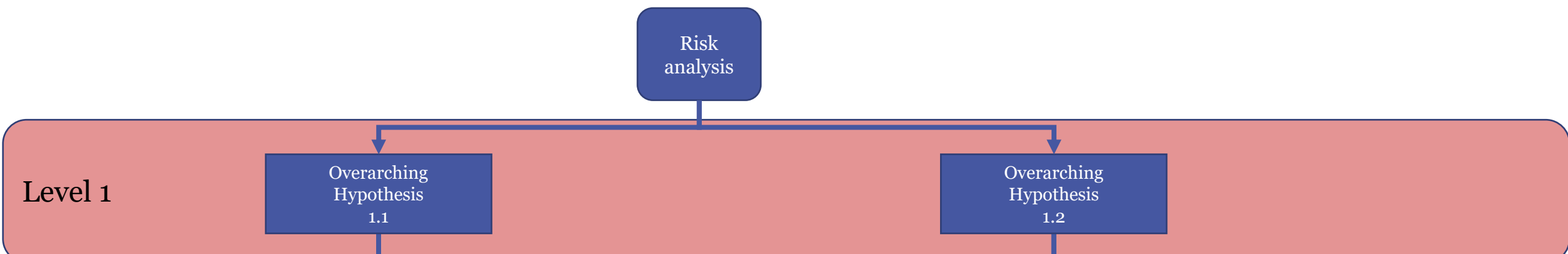
1. Hypotheses developed to address issues
2. Hypotheses represented by stock assessment models
3. Hypotheses are grouped into a hierarchical framework
 - Avoids any hypothesis dominating
 - Facilitates model development and weight assignment
4. Sub-hypotheses represent models with parameters that cannot be reliably estimated
5. Multiple metrics to evaluate plausibility of the hypotheses
6. Model fit only plays a limited role
7. Efficient approach to eliminate unlikely hypotheses

Introduction: 5 main steps

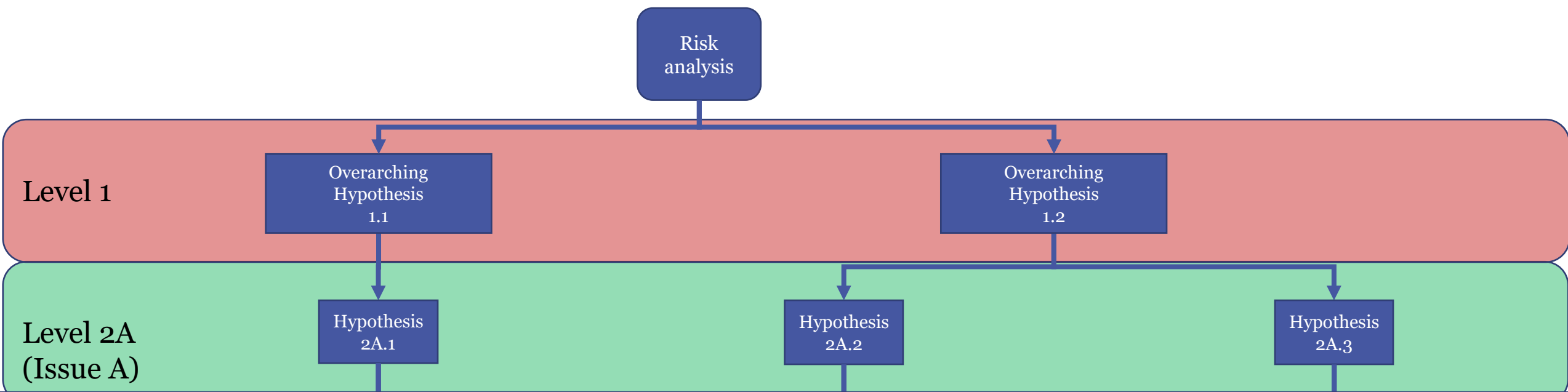
1. Establishing a hierarchy of hypotheses and models
2. Define a weighting system for hypotheses and models
3. Calculate the probability distributions for quantities of interest for a model
4. Combine probability distributions across models
5. Present the results in the form of a risk analysis

1. Hierarchy of hypotheses and models

Risk
analysis

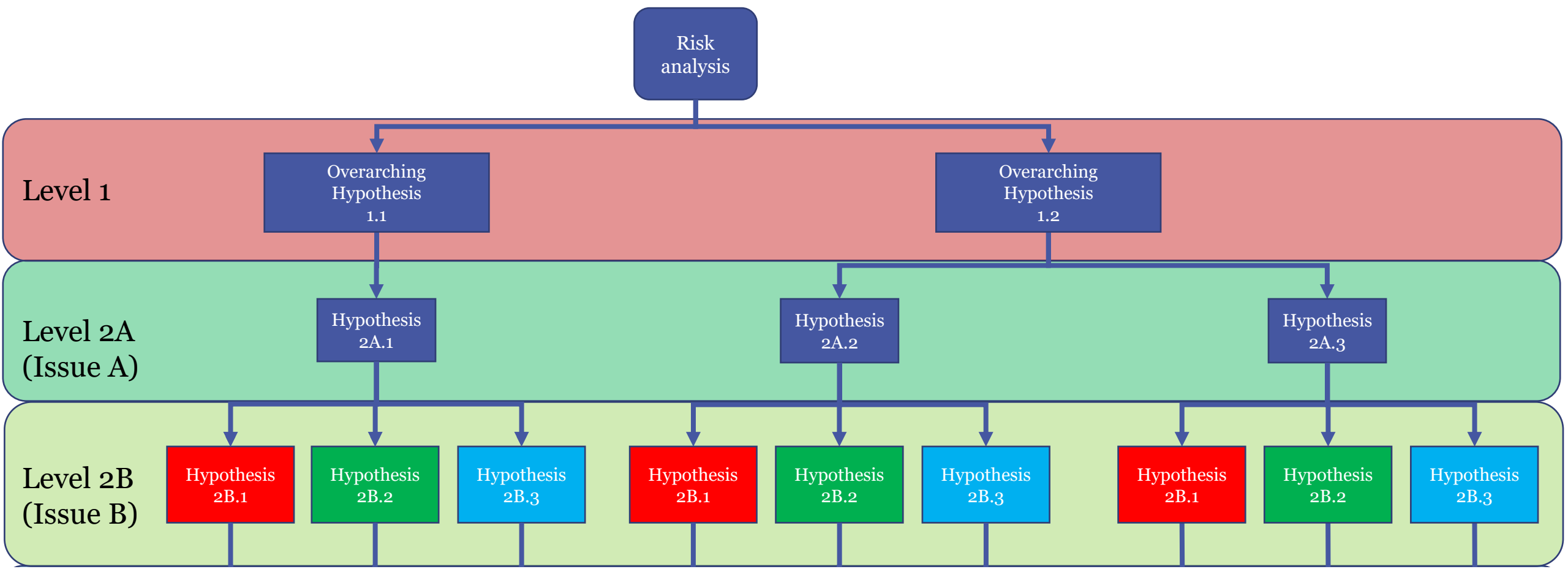


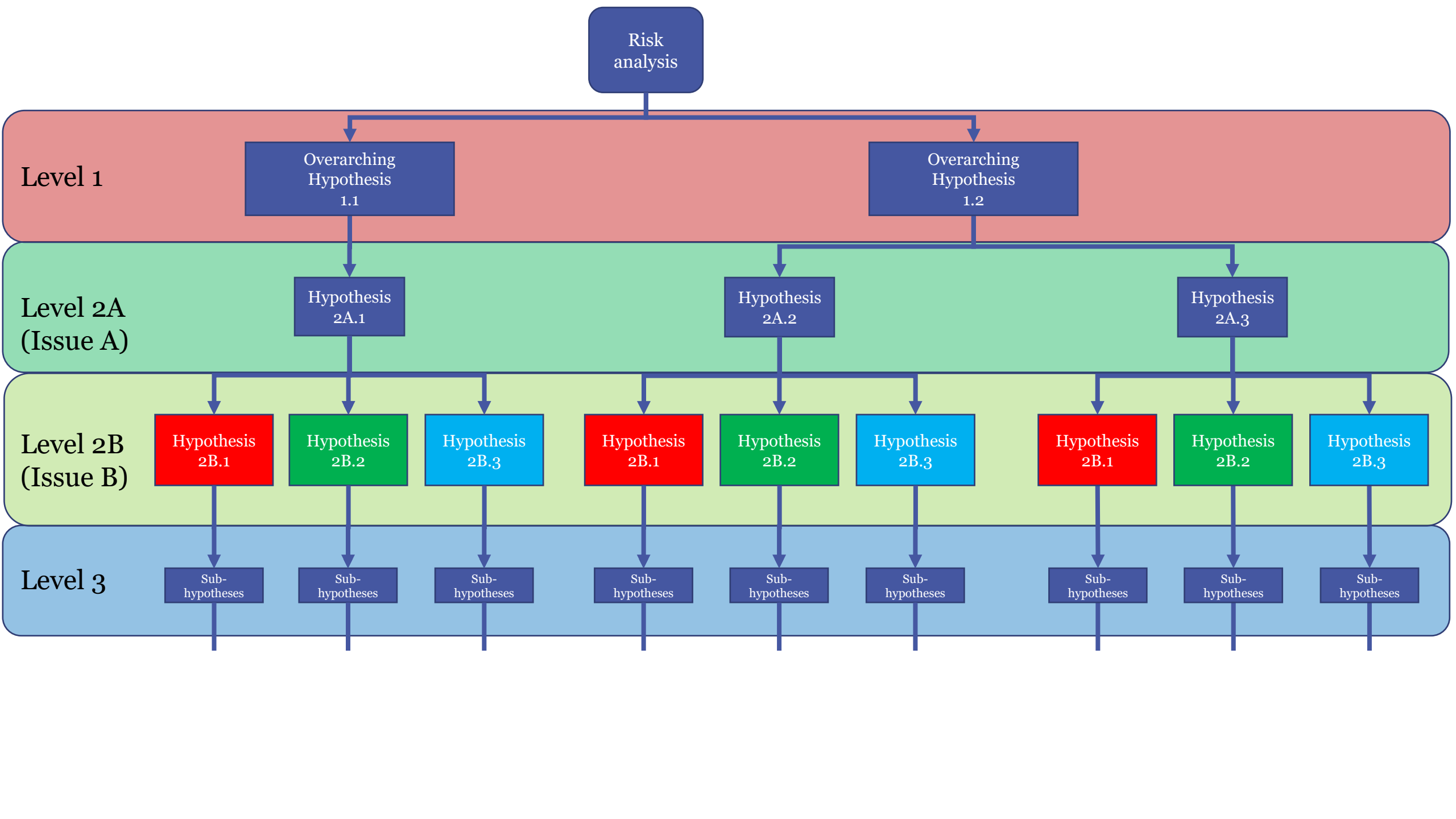
- Level 1: Overarching hypotheses
 - Broad states of nature (e.g. the number of stocks)
 - Represented by a variety of models and data
 - Not evaluated by fit to data
 - Expert opinion for weights

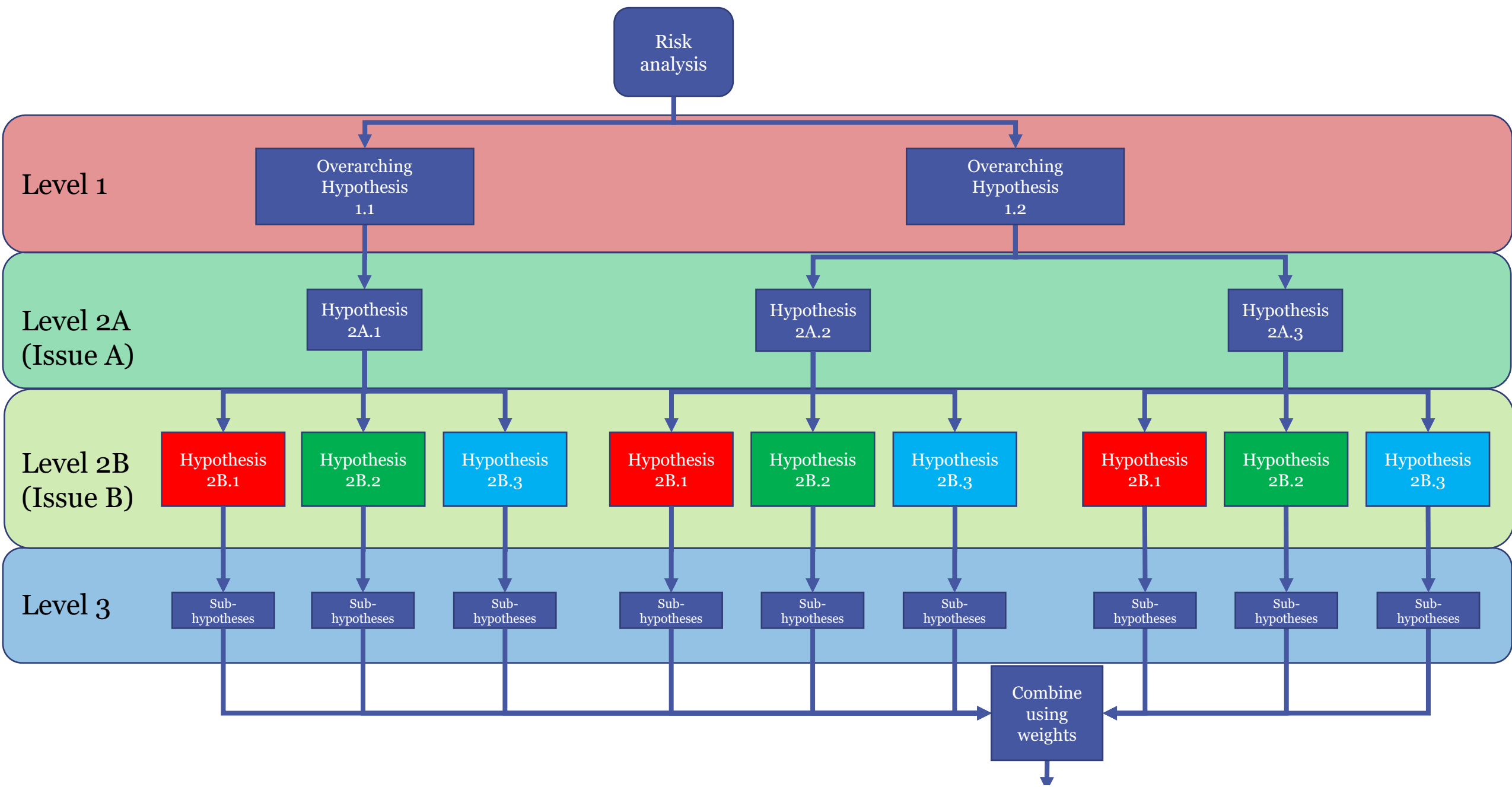


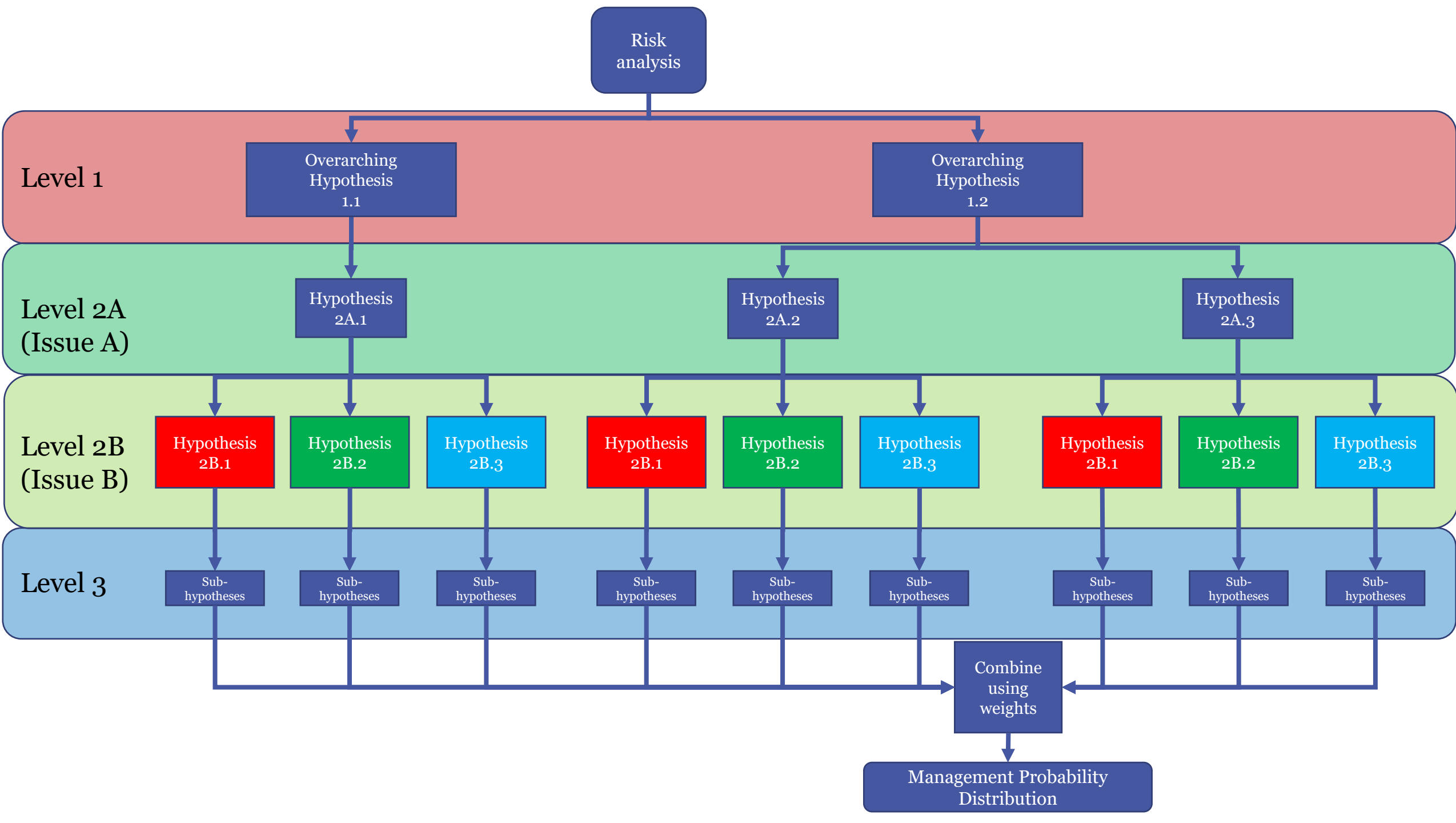
- **Level 2: Hypotheses**

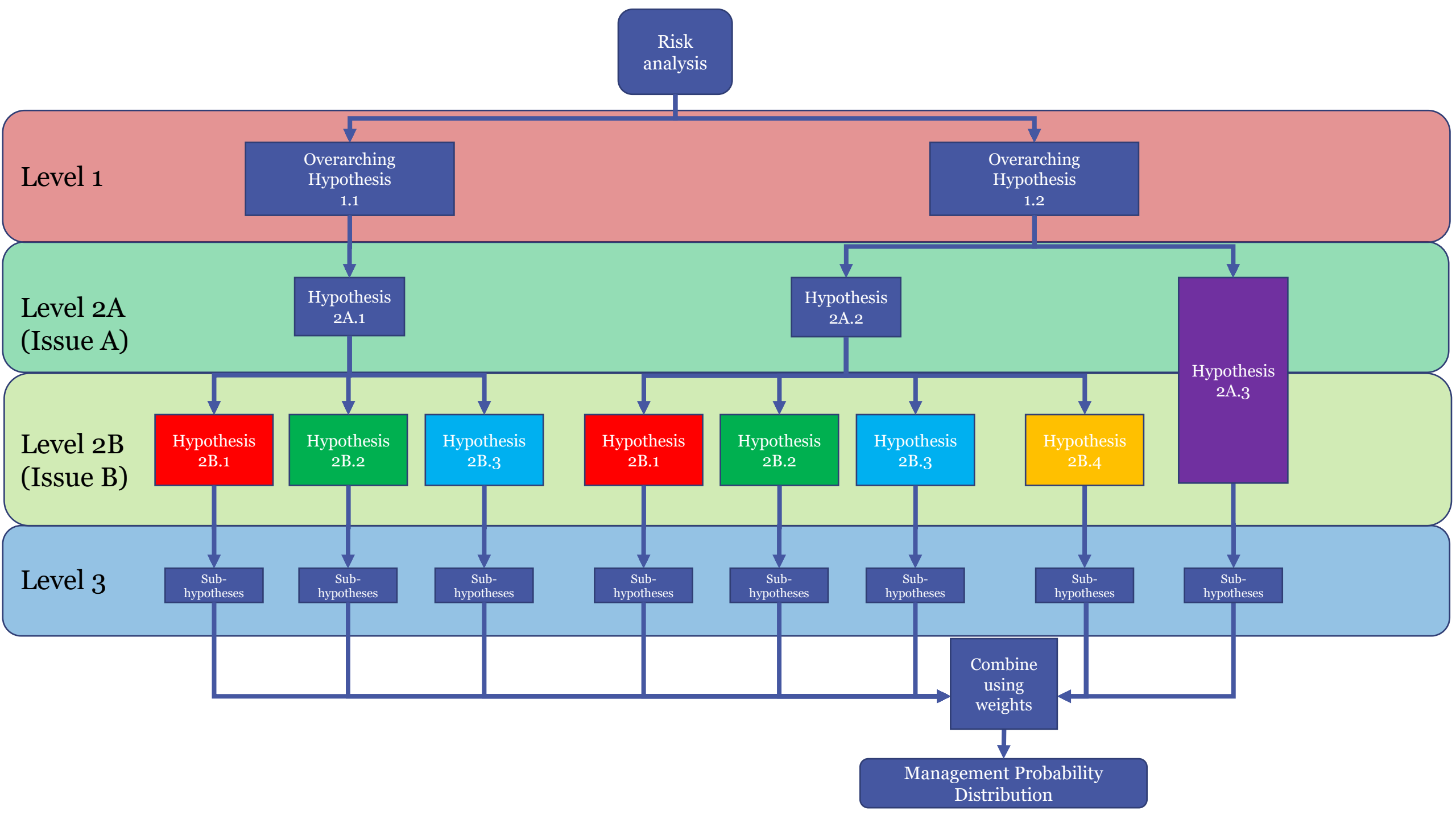
- Represented by a model
- Divided into sub-levels (A, B, ...) where each sub-level addresses an issue in the assessment
- Sub-levels are typically used in combination to solve all the assessment issues
- Aid in assigning weights











2. Defining a weighting system for hypotheses and models

- a) Establish weight categories
- b) Select weight metrics
- c) Assign weights and rescale to be used in a probabilistic framework
- d) Ensure the number of hypotheses is practical

Weighting system: weight categories

- Weighting is subjective
- Use general weight categories
- Assign each category a numeric value

Weight Category	Value
None:	0
Low:	0.25
Medium:	0.5
High:	1.0

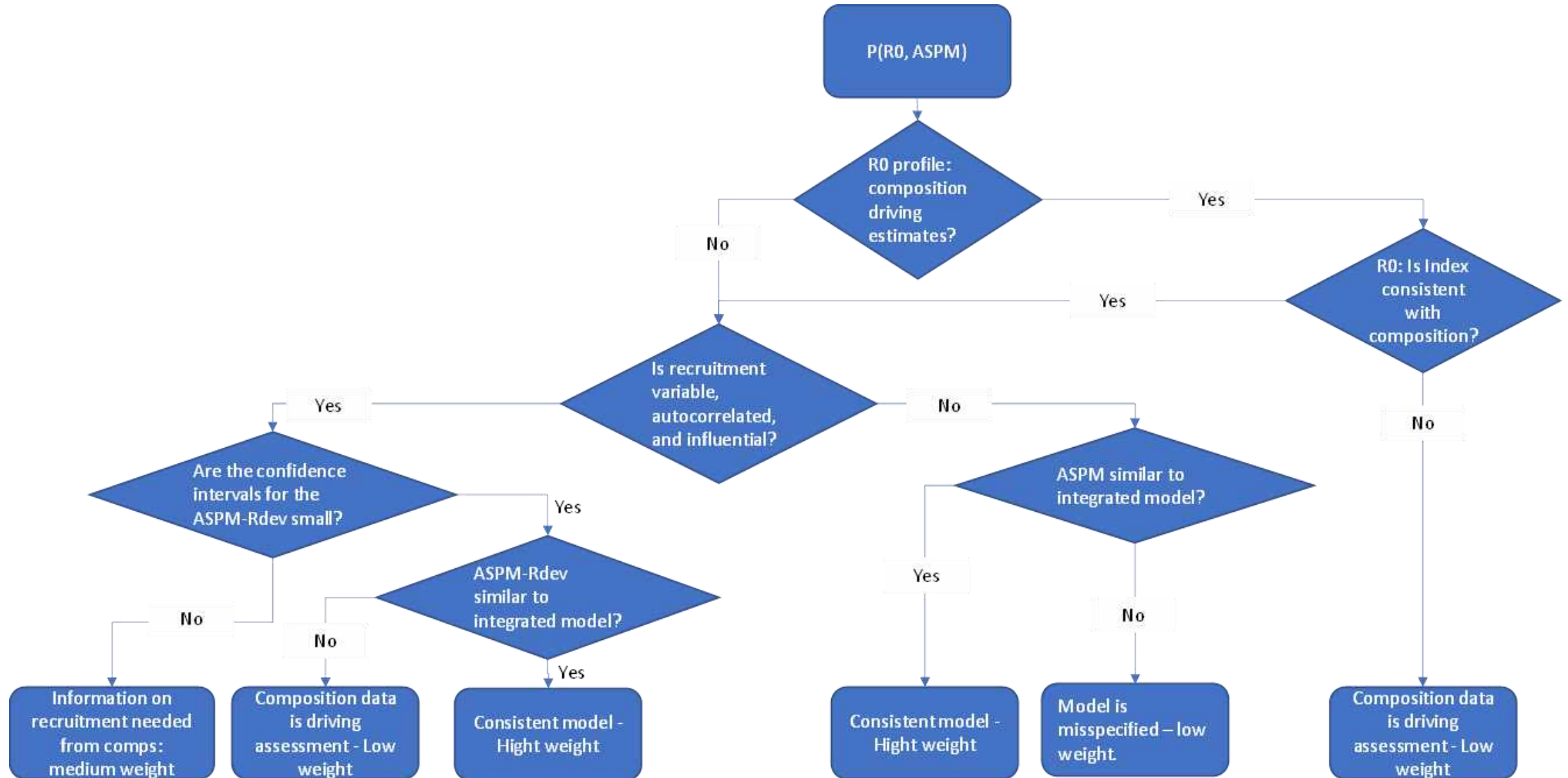
Weighting system: Weight metrics

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

Weighting system: Diagnostics

- $W(\text{ASPM, } R_0, \text{ Catch curve})$
- $W(\text{Retrospective analysis})$
- $W(\text{Composition residuals})$
- $W(\text{Index residuals})$
- $W(\text{Recruitment residuals})$

Weighting system: R0 profile and ASPM diagnostic



Weighting system: Assigning and rescaling weights

- When should the weights be rescaled to sum to one
 - Level 1
 - Rescale across all overarching hypotheses
 - Weights will then be multiplied by the weights from the other levels.
 - Level 2
 - Rescale within each sub-level (e.g. A, B, ...) within a branch of the hierarchy
 - Exception is model fit with different or down-weighted data.
 - Rescale within groups of models with the same data
 - Level 3
 - Rescale to sum to one within a branch of the hierarchy (i.e. for a given Level 2 hypothesis).

Weighting system: Assigning and rescaling weights

- How to assign the weights for a specific model relative to the other models
 - Level 1
 - $W(\text{Expert})$ relative to all overarching hypotheses.
 - Level 2
 - $W(\text{convergence})$, $W(\text{Plausible parameters})$, $W(\text{Plausible results})$ and $W(\text{Diagnostics})$ relative to all models and hypotheses.
 - $W(\text{Fit})$ relative to models that use the same data independent of branches in the hierarchy
 - $W(\text{Expert})$ relative to models in the same branch of the hierarchy (e.g. assuming a Level 1 overarching hypothesis is true).
 - Level 3
 - Relative to models in the same branch of the hierarchy (i.e. for a given Level 2 hypothesis).

Calculating probability distributions for quantities of interest for a model

- Normal approximations based on the estimate and standard error
- Some standard errors are approximated
- The resulting distribution is rescaled to obtain $P(\text{Quantity} | \text{Model}=m)$.
- Works well when the data is very informative
- Probability distribution may be asymmetrical
- Posteriors derived from limited MCMC analyses used to evaluate appropriateness of the approximation

Combining probability distributions across models

- a) Determine the weight of each model: $W(\text{model})$
- b) Rescale the values from (a): “ $P(\text{Model} = m)$ ”
- c) Calculate the probability of the quantity of interest for each model, rescaled so that they sum to one: $P(\text{Quantity} | \text{Model}=m)$.
- d) Multiply (b) and (c) for each model in the collection and sum across models: $P(\text{Quantity})$.
- e) Evaluate (d) for all management quantities.

$$P(\text{Quantity}) = \sum_{m \in \{\text{Models}\}} P(\text{Quantity} | \text{Model} = m) P(\text{Model} = m)$$

Presenting the results in the form of a risk analysis

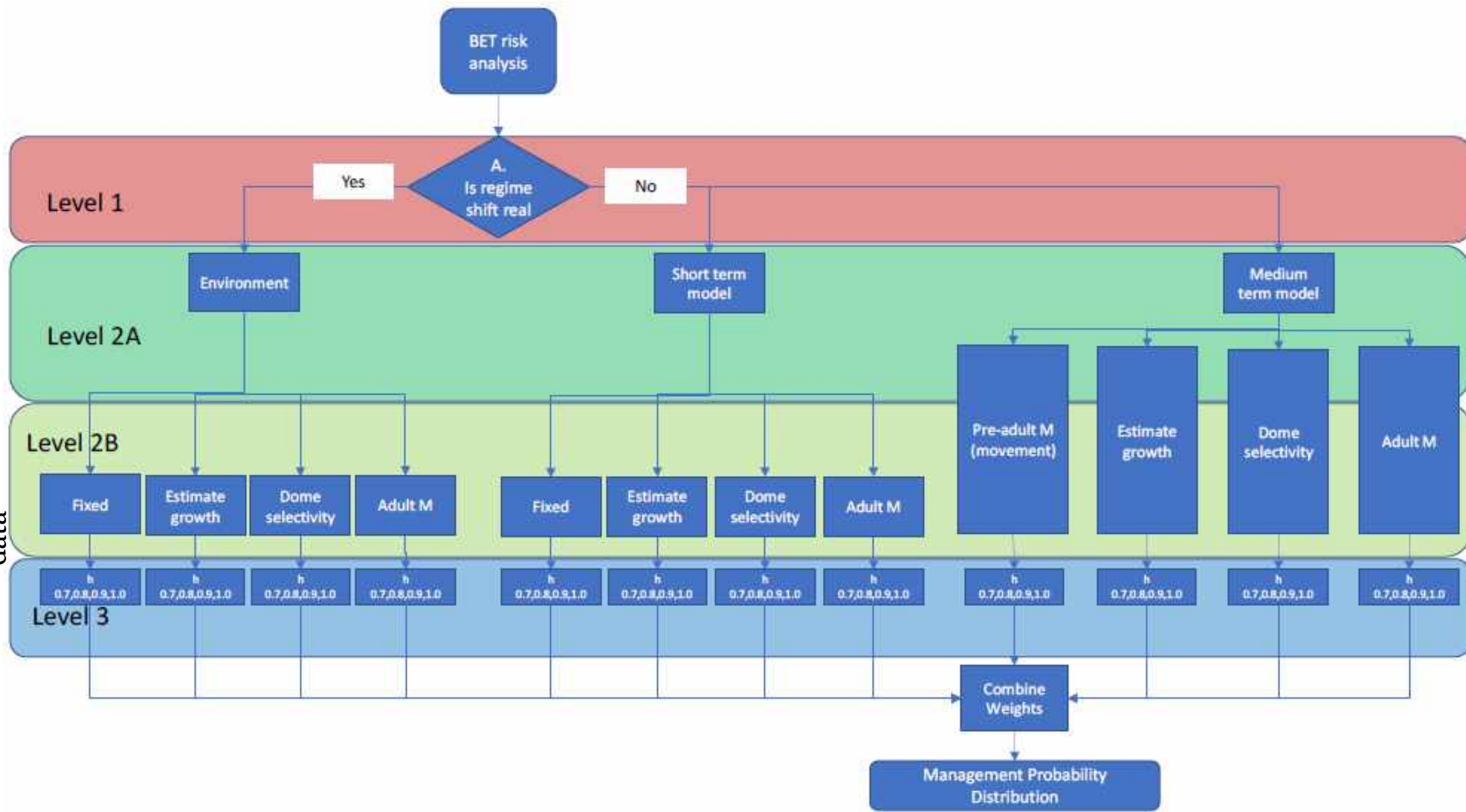
- Plot distributions by components (e.g. hypotheses at level 2A and 2B)
- Cumulative density functions (CDFs) can be used to determine the probability of exceeding the reference points.
- Decision tables
 - Outcome of specific management action under different states of nature.
 - The states of nature could be the individual models, combinations of models, or a derived quantity (e.g. biomass).
 - The probability of each state of nature is also included
- Risk curves
 - Probability of outcome versus management action

Application: bigeye tuna

- Conducted in Stock Synthesis
- Many fisheries
- CPUE and length composition data
- Overarching hypothesis: is regime shift in recruitment when fishery on juveniles expanded real
- Issues
 - A. Regime shift
 - B. Misfit to large fish in composition data from asymptotic fishery
- Panel of experts that subjectively assign weights

Flow chart for bigeye tuna

Misfit comp
Regime shift

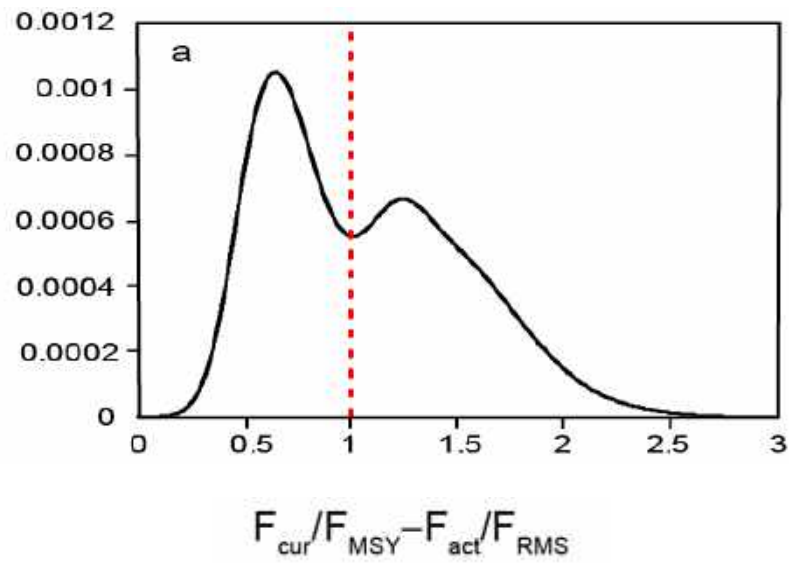


BET: F_{cur} probability distributions relative to RPs

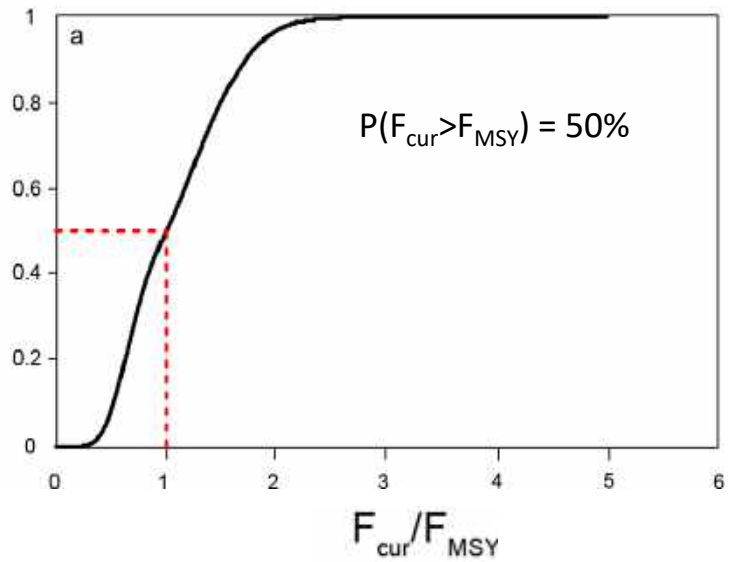


TARGET

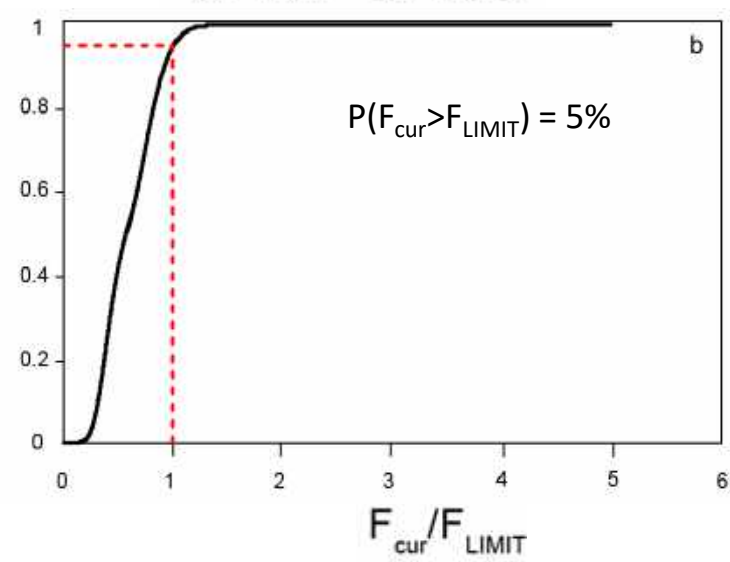
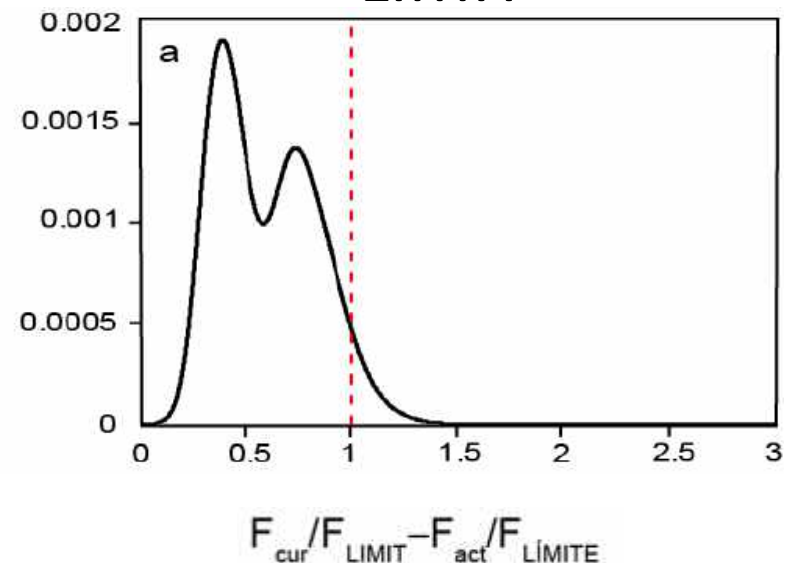
Probability density



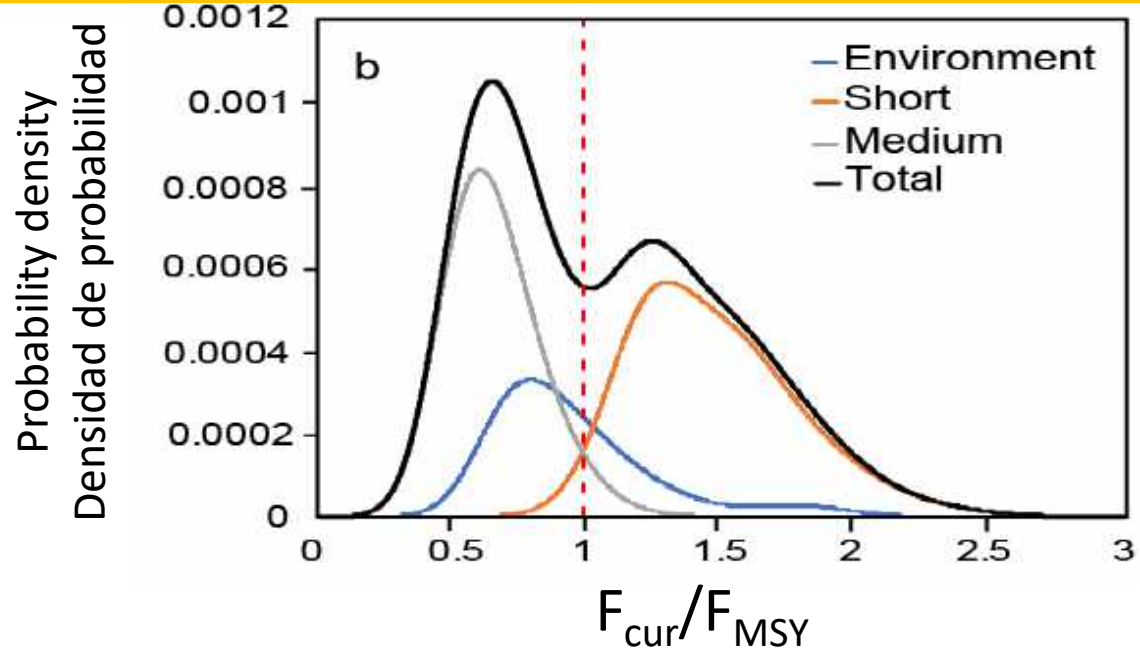
Cumulative probability



LIMIT

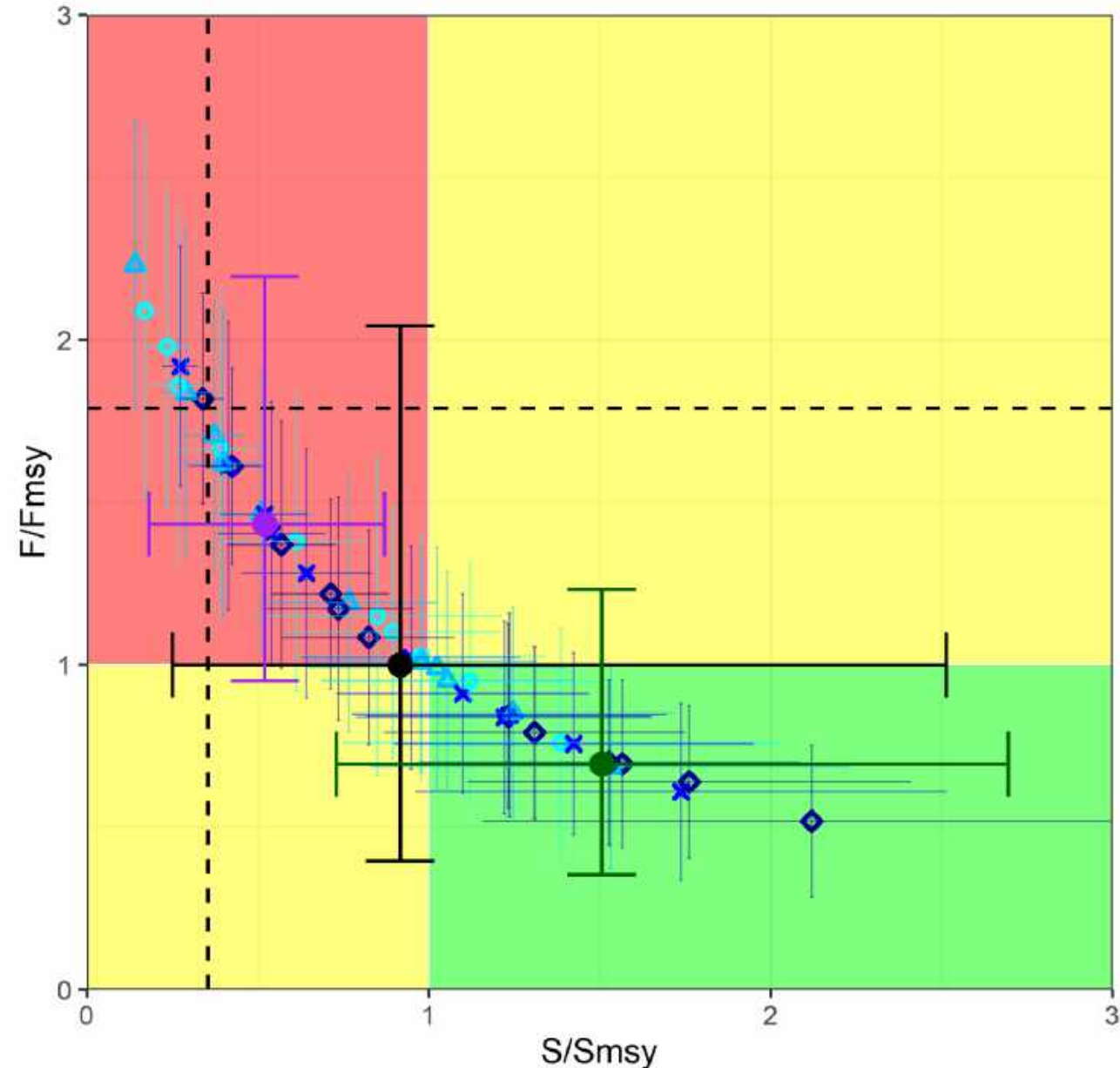


BET: Composition of F_{cur}/F_{MSY} prob. distribution



BET

BET: Current stock status (Kobe plot)



- TARGETS

- 50% probability that F_{MSY} has been exceeded: $P(F_{cur} > F_{MSY}) = 50\%$
- 53% probability that S_{cur} is below S_{MSY} : $P(S_{cur} < S_{MSY}) = 53\%$

Steepness

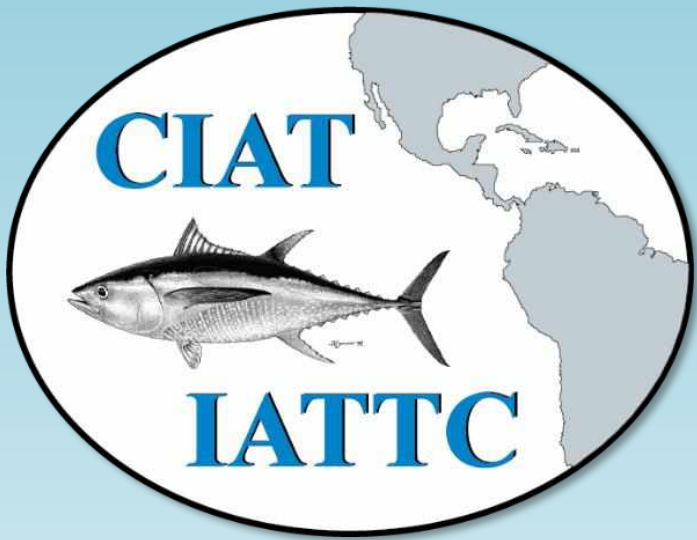
- ◉ 0.7
- ◈ 0.8
- ✖ 0.9
- ◊ 1

- LIMITS

- There probability that either S and F limit reference points have been exceeded is not negligible:
 $P(S_{cur} < S_{LIMIT}) = 6\%$; $P(F_{cur} > F_{LIMIT}) = 5\%$

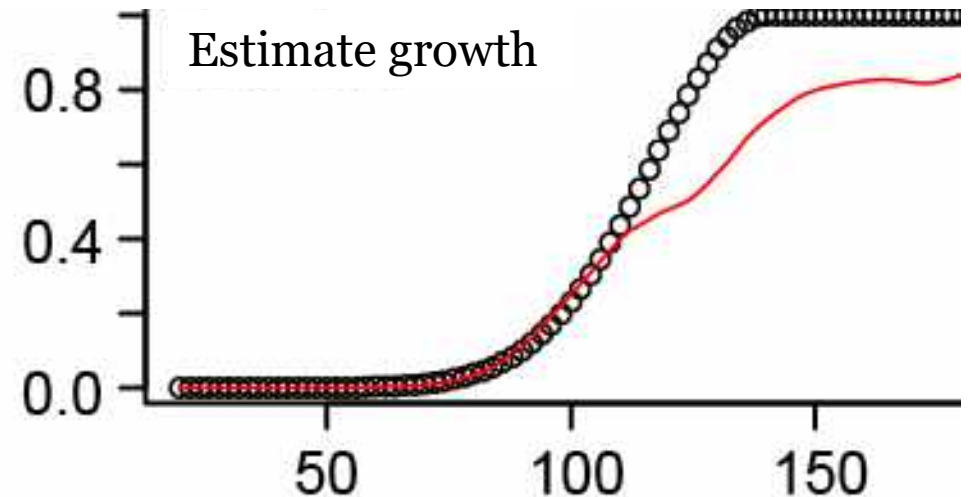
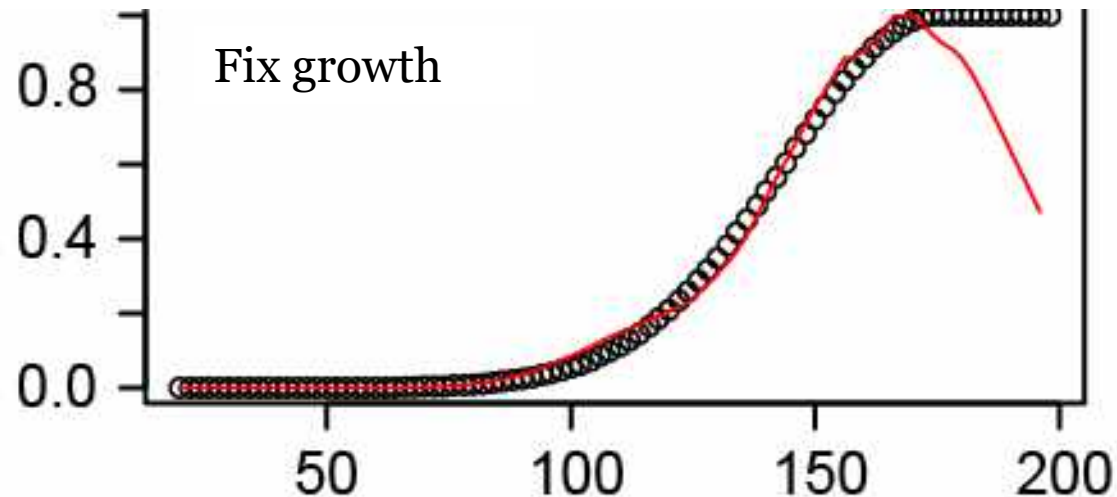
Summary: main concepts

- Develop alternative hypotheses to address issues with assessment
- Hierarchical structure to represent hypotheses
- Model weighting based on a set of metrics to assign model probabilities (e.g. diagnostics) not just fit (e.g. AIC)

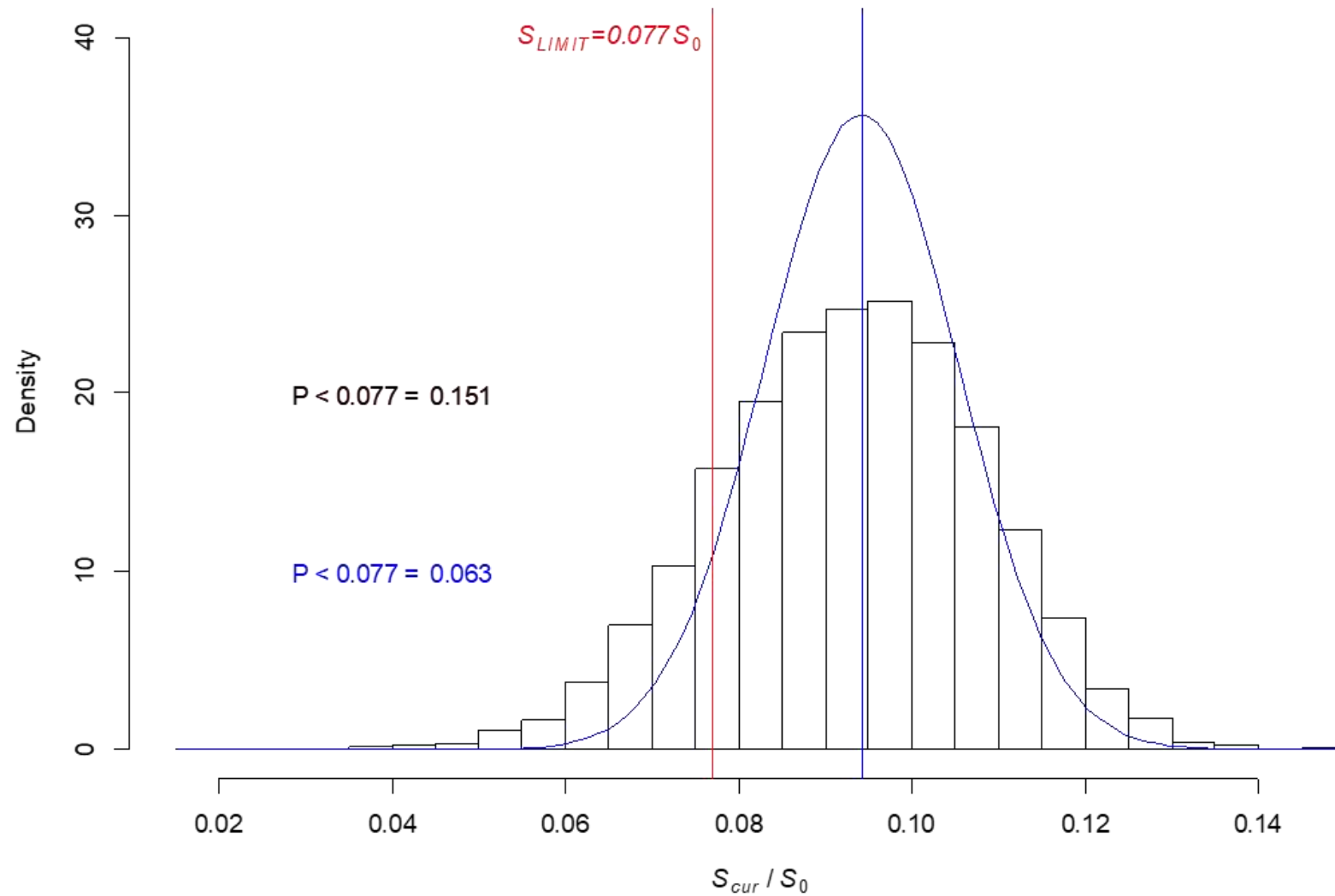


Weighting system: W(“Empirical” selectivity)

- Compares “Empirical” selectivity with estimated selectivity
- “Empirical” is the catch at length in numbers divided by the estimated abundance at length in numbers
- Focusses on larger fish which are more influential



Probability distributions: MCMC comparison



Other presentations and documents

1. Identify alternative hypotheses

- YFT: SAC-11-J; BET: SAC-11 INF-F

2. Implement stock assessment models representing alternative hypotheses

- YFT: SAC-11-07; BET: SAC-11-06

3. Assign relative weights to each hypothesis (model)

- YFT: SAC-11 INF-J; BET: SAC-11 INF-F

4. Compute combined probability distributions for management quantities

- SAC-11-08

Documents

<https://www.iattc.org/Meetings/Meetings2020/SAC-11/IATTC%20Scientific%20Meeting%20and%20Working%20Groups%202020ENG.htm>

Presentations

<https://www.youtube.com/playlist?list=PLKeH-azh54PVfbUDbePSLcZvIozGXSHRa>



Discussion

- Are we doing ensemble modelling or just model development/selection?
- Need more objective and transparent scoring
- Other diagnostics
 - Posterior predictive checks and Frequentist equivalents
 - One-step-ahead predictions
- We use data for parameter uncertainty but diagnostics for model uncertainty

Introduction: Assessment uncertainty

- Parameter uncertainty
 - Standard practice in stock assessment
 - Confidence intervals on quantities of interest
- Model structure uncertainty
 - Sensitivity analysis
 - Multiple models
 - Combine models
 - Model weights
- Uncertainty about the future (e.g. process variation)
 - E.g. recruitment variation
 - Not implemented yet
 - Can't evaluate biomass reference points

Presenting results: Decision tables

	State of nature	Total
	Probability	
Management action	Outcome	Outcome

Presenting results: Decision tables

	Model, group of models, derived quantity	Total
	Probability	
Catch, Effort, Closure days	Catch, Biomass, $P(F > F_{LIMIT})$	Outcome

Presenting results: IATTC Risk curves and decision tables

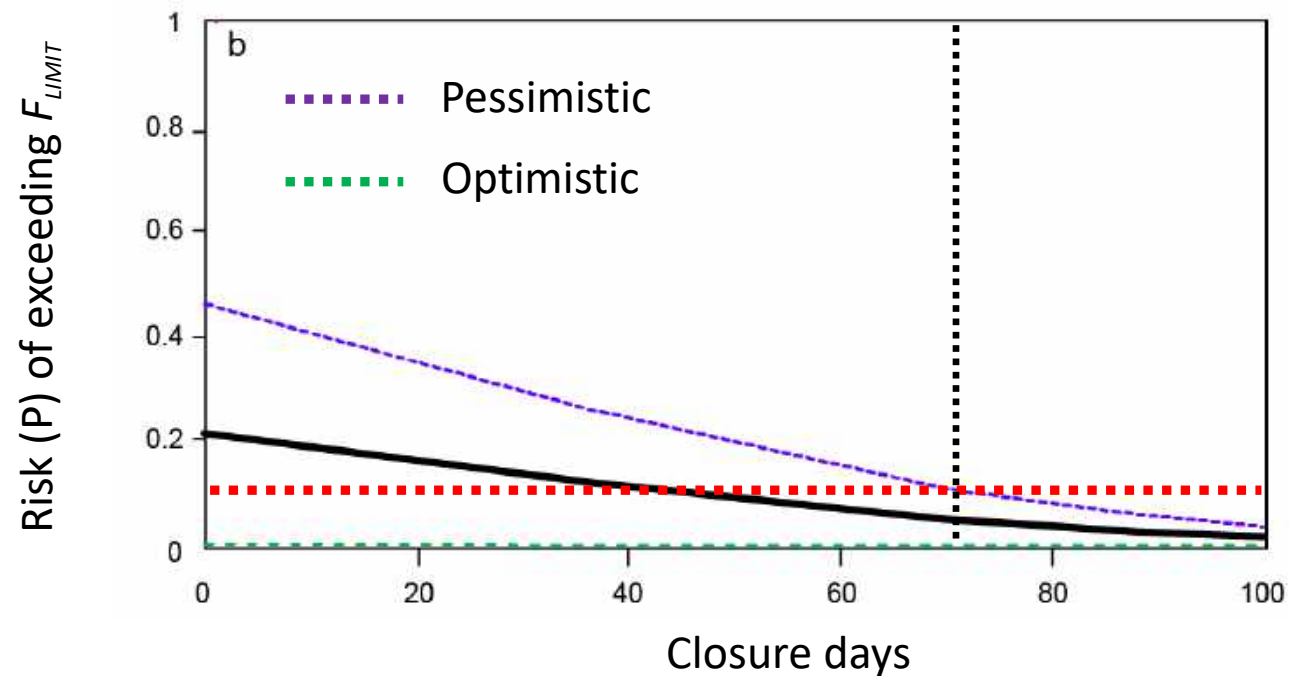
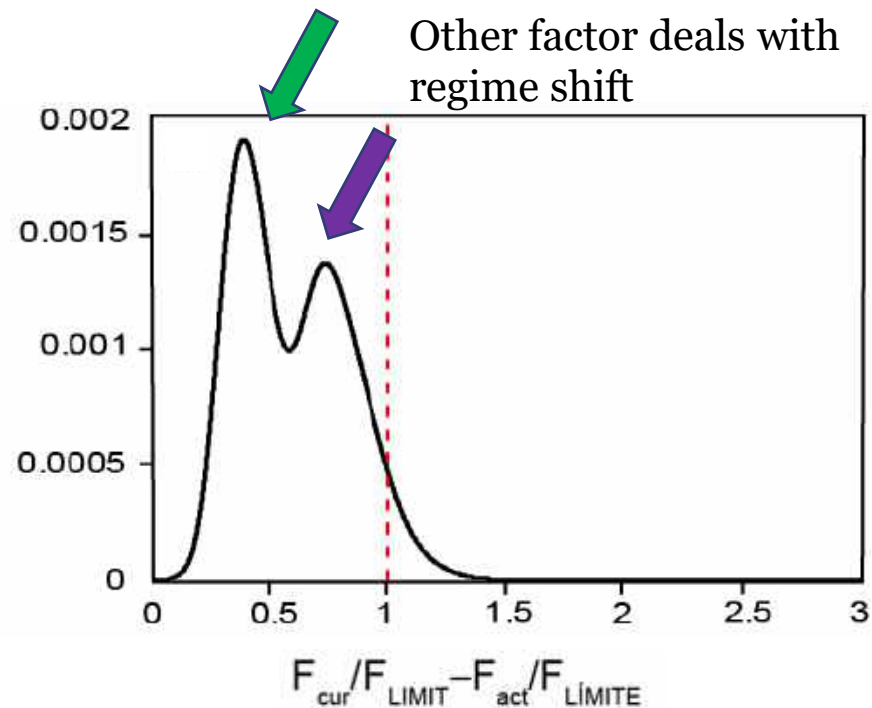
- Outcome of different levels of fishery closures
- Assumes fishing mortality is proportional to the days the fishery is open
 - 365 – days of closure
 - Adjusted for changes in fishing capacity and the Corralito
- $P(F > F_{MSY})$ and $P(F > F_{LIMIT})$
- Need to do projections for spawning biomass so not provided

BET: Risk curves for exceeding F_{LIMIT}



High biomass
removes regime shift

Other factor deals with
regime shift



BET: Decision table



Closure days	Env-Fix	Env-Gro	Env-Sel	Env-Mrt	Srt-Fix	Srt-Gro	Srt-Sel	Srt-Mrt	Mov	Gro	Sel	Mrt	Comb
$P(\text{model})$	0.01	0.13	0.05	0.02	0.04	0.22	0.11	0.07	0.01	0.24	0.09	0.02	
$P(F > F_{MSY})$	Probability $\leq 50\%$ $> 50\%$												
0	1.00	0.48	0.78	0.98	1.00	1.00	0.99	1.00	0.47	0.09	0.31	0.65	0.62
36	1.00	0.32	0.63	0.93	1.00	0.99	0.97	1.00	0.30	0.03	0.17	0.45	0.56
70	1.00	0.19	0.44	0.84	1.00	0.97	0.92	0.99	0.15	0.01	0.07	0.25	0.50
72	1.00	0.18	0.43	0.83	1.00	0.96	0.91	0.98	0.14	0.01	0.06	0.24	0.49
88	1.00	0.13	0.35	0.75	1.00	0.93	0.87	0.97	0.09	0.00	0.04	0.17	0.46
100	1.00	0.09	0.28	0.67	1.00	0.88	0.81	0.95	0.06	0.00	0.02	0.11	0.43
$P(F > F_{LIMIT})$	Probability $\leq 10\%$ $> 10\%$												
0	0.97	0.00	0.04	0.17	0.89	0.39	0.37	0.57	0.00	0.00	0.00	0.00	0.21
36	0.79	0.00	0.01	0.06	0.67	0.19	0.18	0.33	0.00	0.00	0.00	0.00	0.12
70	0.33	0.00	0.00	0.01	0.38	0.07	0.06	0.14	0.00	0.00	0.00	0.00	0.05
72	0.30	0.00	0.00	0.01	0.36	0.06	0.06	0.13	0.00	0.00	0.00	0.00	0.05
88	0.11	0.00	0.00	0.00	0.25	0.03	0.03	0.08	0.00	0.00	0.00	0.00	0.03
100	0.04	0.00	0.00	0.00	0.17	0.02	0.02	0.04	0.00	0.00	0.00	0.00	0.02

Weighting system: Reducing the number of models

- All model combinations is impractical
- Some diagnostics are computationally intensive
- Metrics assigned zero eliminate a model
- Eliminating groups of models
 - Define a “base” model
 - The base model is the simpler model
 - If base model is eliminated, then the other models derived from this model are also eliminated
 - Need to consider the reason for the elimination because other models may correct for the reason the base model was eliminated

1. Hierarchy of hypotheses and models

- **Level 1: Overarching hypotheses**
 - Broad states of nature (e.g. the number of stocks)
 - Represented by a variety of models and data
 - Not evaluated by fit to data
 - Expert opinion for weights
- **Level 2: Hypotheses**
 - Represented by a model
 - Divided into sub-levels (A, B, ...) where each sub-level addresses an issue in the assessment
 - Sub-levels are typically used in combination to solve all the assessment issues
 - Aid in assigning weights

Introduction - Hierarchy of hypotheses and models

- Level 3: Sub-hypotheses
 - Evaluated differently
 - Avoid the influence of data
 - Reduce the number of analyses
 - Convenience
 - Typically encompassed by a single hypothesis
 - Can be represented by restricting a model (e.g. fixing the value of a parameter, such as steepness)
 - Applied to most, if not all, models on Level 2.

2. Defining a weighting system for hypotheses and models

Weighting system: W(Fit)

- Does not use standard AIC rules
- $W(\text{Fit}) = \text{Low} + (\text{High} - \text{Low}) \times (1 - [\Delta \text{AIC} / \max(\Delta \text{AIC})])$
- Needs same data and same data weighting
- For models with data specific to a parameter (e.g. age at length data for growth), calculate AIC without those data
- Otherwise, models with different data evaluated separately
- Modelled process variability brings additional complications

Outline

- Introduction
- Hierarchy of hypotheses and models
- Weighting system
- Probability distributions for quantities of interest
- Combining probability distributions across models
- Presenting results
- Summary

Summary

- Assessments are uncertain
- IATTC HCR for tropical tunas (Resolution C-16-02) addresses uncertainty through probability statements
- Transition from single base-case assessment to set of reference models
- Hierarchy of hypotheses to define models
- Rigorous statistical framework is not applicable
- Set of metrics to assign model probabilities
- Decision table to present outcome of alternative management actions