Key decisions: How should ensembles be constructed and combined?

Nicholas Ducharme-Barth (NOAA PIFSC)

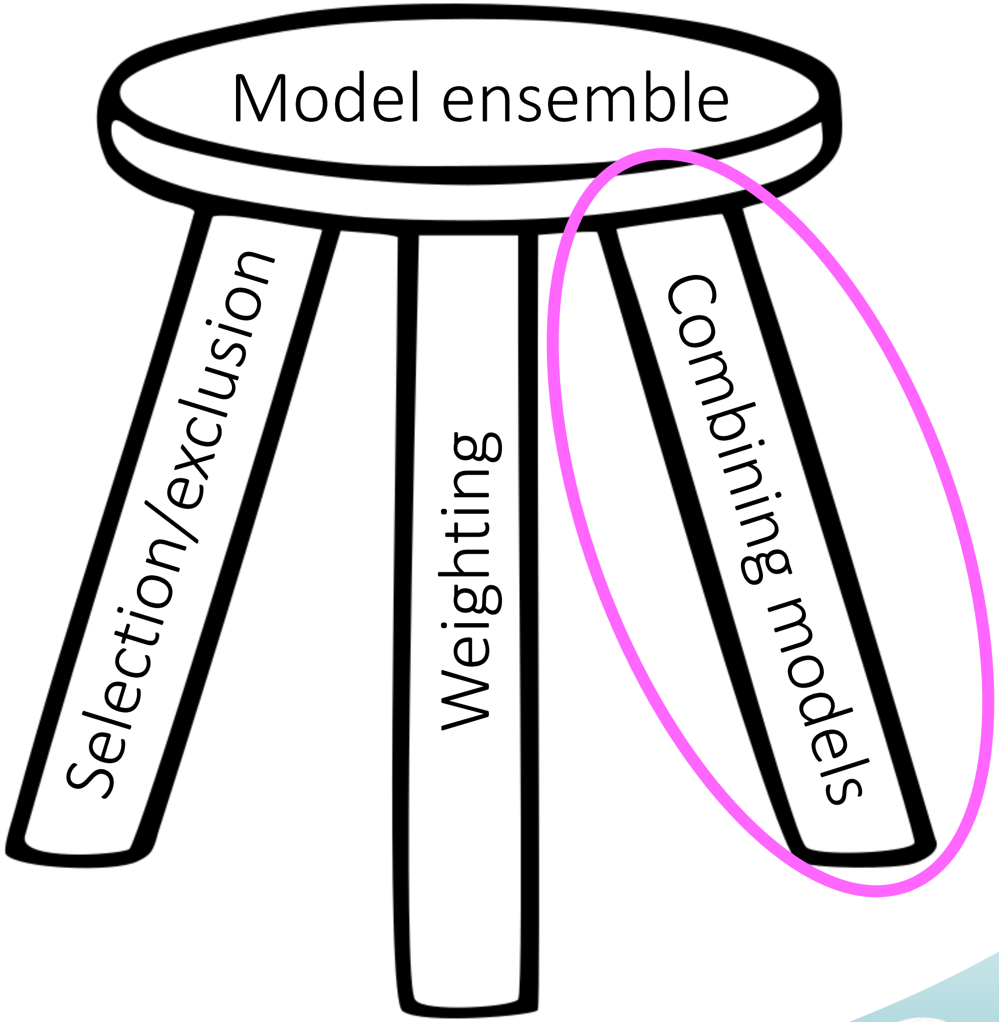
Matthew Vincent (NOAA SEFSC)

December 01, 2022

CAPAM/IATTC Model Weighting Workshop

Combining models -Key decisions

1. How are model ensembles constructed?
2. How are models in the ensemble combined to present the central tendency in stock status?
3. How was uncertainty in stock status from the ensemble presented?



Combining models -Key decisions

1. How are model ensembles constructed?
2. How are models in the ensemble combined to present the central tendency in stock status?
3. How was uncertainty in stock status from the ensemble presented?

ELSEVIER

Fisheries Research

Volume 255, November 2022, 106452

Focusing on the front end: A framework for incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments

Nicholas D. Ducharme-Barth [a, b] [1], Matthew T. Vincent [c, 1]

https://doi.org/10.1016/j.fishres.2022.106452

Combining models -Key decisions

1. How are model ensembles constructed?
2. How are models in the ensemble combined to present the central tendency in stock status?
3. How was uncertainty in stock status from the ensemble presented?

ELSEVIER

Fisheries Research

Volume 255, November 2022, 106452

Focusing on the front end: A framework for incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments

What are other practitioners doing???

CAPAM Model Weighting: Current practices

As a part of the upcoming CAPAM Model Weighting workshop we would like to characterize current practice for how model ensembles are created and combined for use in stock assessment. Please help by filling out this short survey (< 5 minutes).

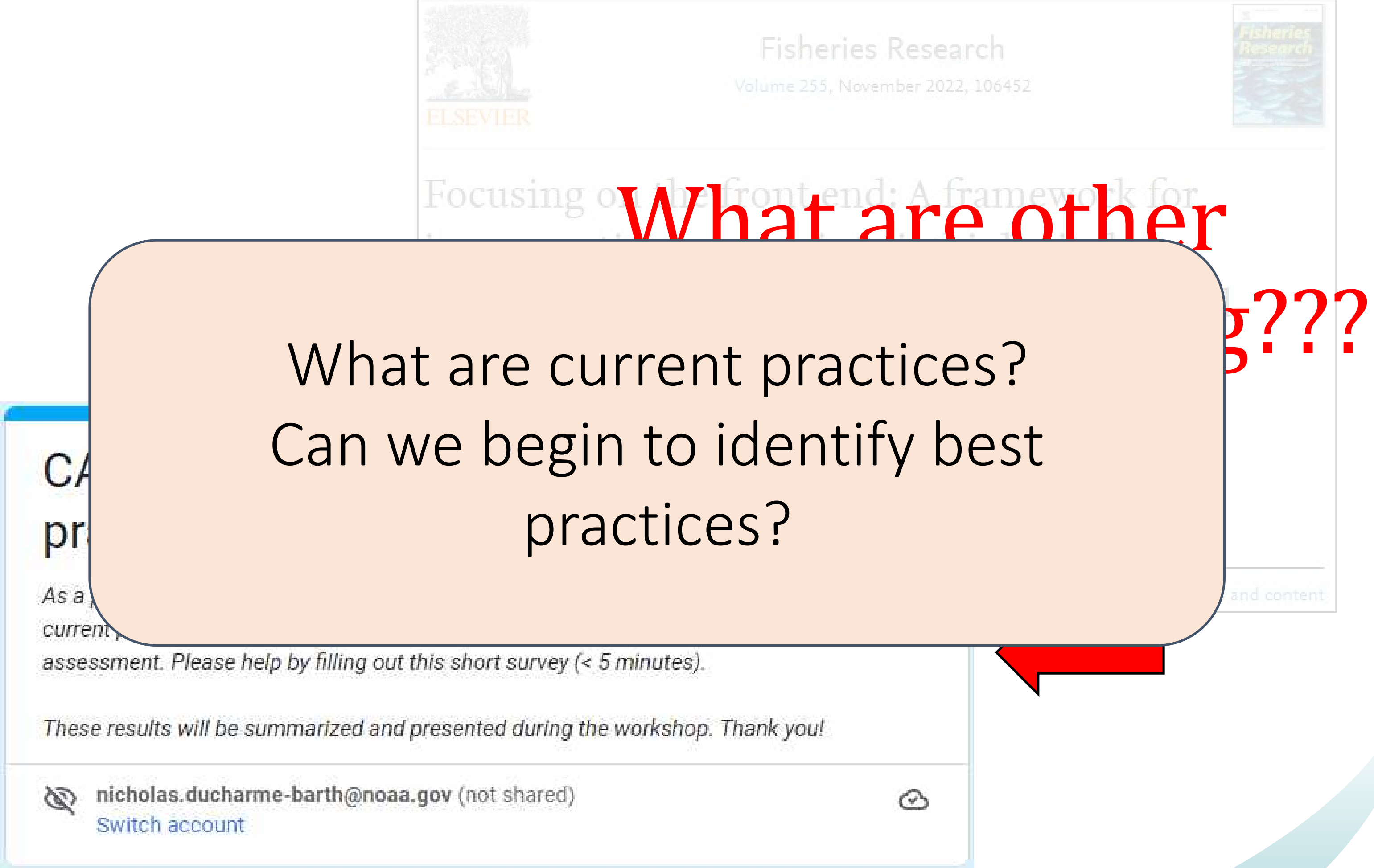
These results will be summarized and presented during the workshop. Thank you!

nicholas.ducharme-barth@noaa.gov (not shared)
Switch account

NOAA FISHERIES

Combining models -Key decisions

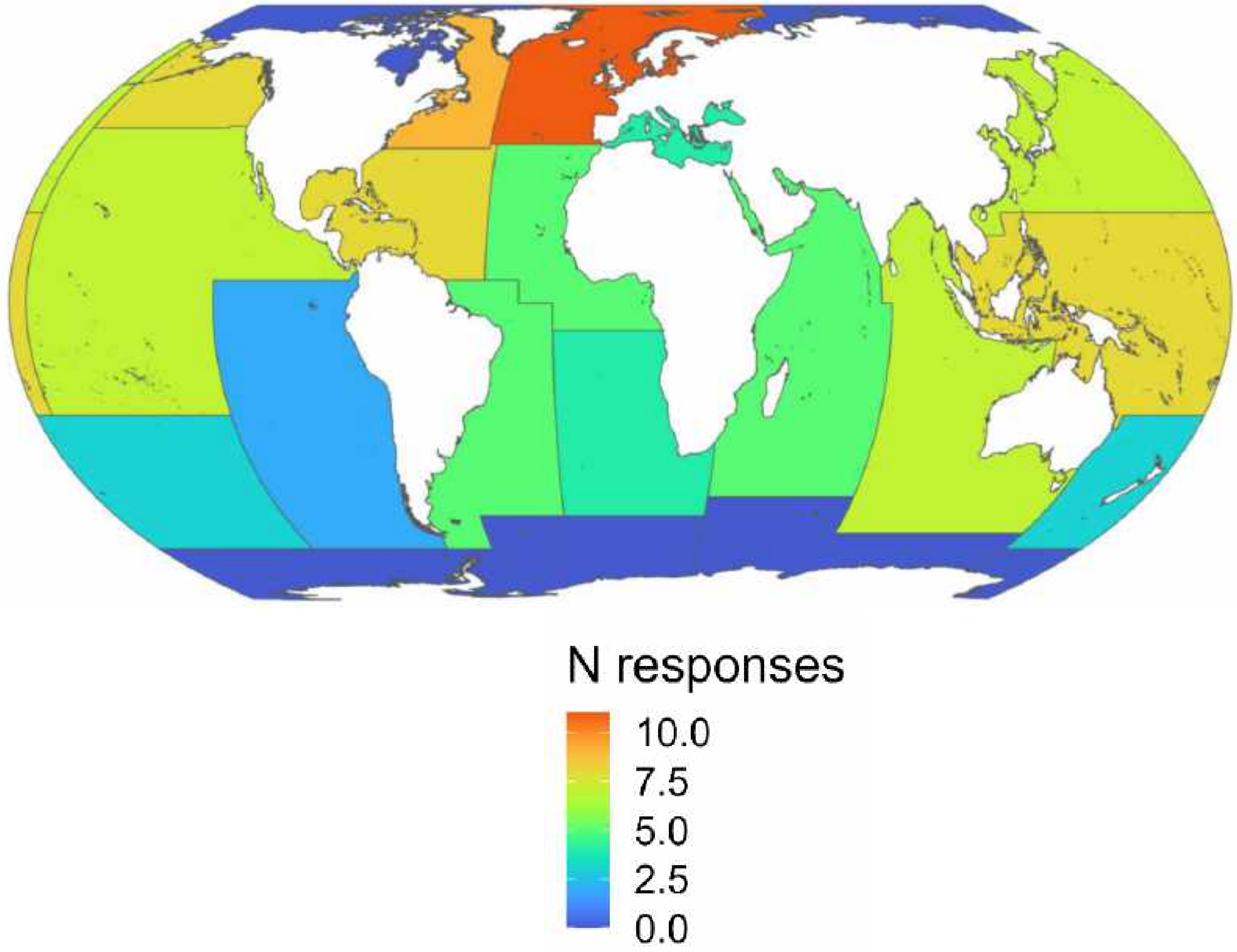
1. How are model ensembles constructed?
2. How are models in the ensemble combined to present the central tendency in stock status?
3. How was uncertainty in stock status from the ensemble presented?



What are current practices? Can we begin to identify best practices?

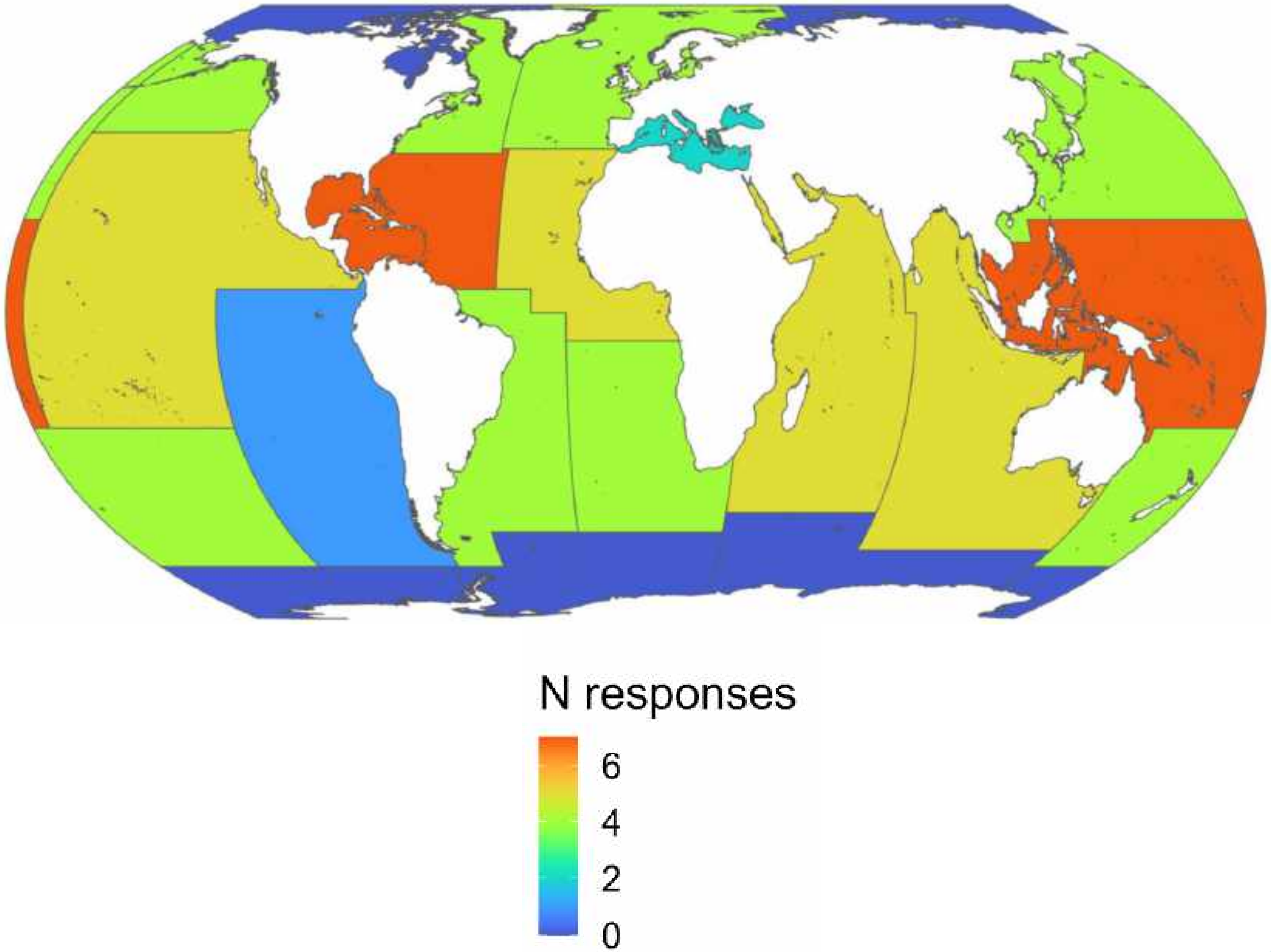
Survey overview

- 42 responses (globally)
- 25 of 42 (~60%) participants had used a model ensemble to characterize stock status
- Good response rate across experience level.
- If participants had used an ensemble, all had used one within the last 3 years.



Where are ensembles being used?

- RFMOs
  - ICCAT, IOTC, IATTC, WCPFC, IHPC, ICES, ISC, GFCM
- Domestically
  - USA (Southeast/Hawaii)
  - Canada
  - Australia (Queensland)



How are model ensembles constructed?

- Ad-hoc combination of models
- Hypothesis tree approach
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution
- Other

How are model ensembles constructed?

- Ad-hoc combination of models
- Hypothesis tree approach (Hierarchical approach)
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution
- Other

How are model ensembles constructed?

	Ensemble experience
• Ad-hoc combination of models	24%
• Hypothesis tree approach	12%
• Full-factorial combination of uncertainties	32%
• Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution	24%
• Other	8%

How are model ensembles constructed?

	Ensemble experience	No ensemble experience
Ad-hoc combination of models	24%	50%
Hypothesis tree approach	12%	19%
Full-factorial combination of uncertainties	32%	6%
Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution	24%	0%
Other	8%	25%

How are model ensembles constructed?

- Ad-hoc combination of models
- Hypothesis tree approach
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution

- Inefficient (model explosion)
- Unrealistic parameter combinations
- *Subjective* choices for parameter levels & model weighting

Natural Mortality (M)
0.2 (low)
0.3 (medium)
0.4 (high)

×

Growth (L2)
140 (low)
160 (medium)
180 (high)

Model 1: M=0.2,L2=140	Model 2: M=0.2,L2=160	Model 3: M=0.2,L2=180
Model 4: M=0.3,L2=140	Model 5: M=0.3,L2=160	Model 6: M=0.3,L2=180
Model 7: M=0.4,L2=140	Model 8: M=0.4,L2=160	Model 9: M=0.4,L2=180

Can we do better? Monte Carlo Bootstrap (MCB)

- Develop a multivariate distribution for key parameters that would be fixed in an assessment.
- Each model in the ensemble would be parametrized by parameters drawn from distribution.
- Benefits?

A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries

Issue: 90(4)

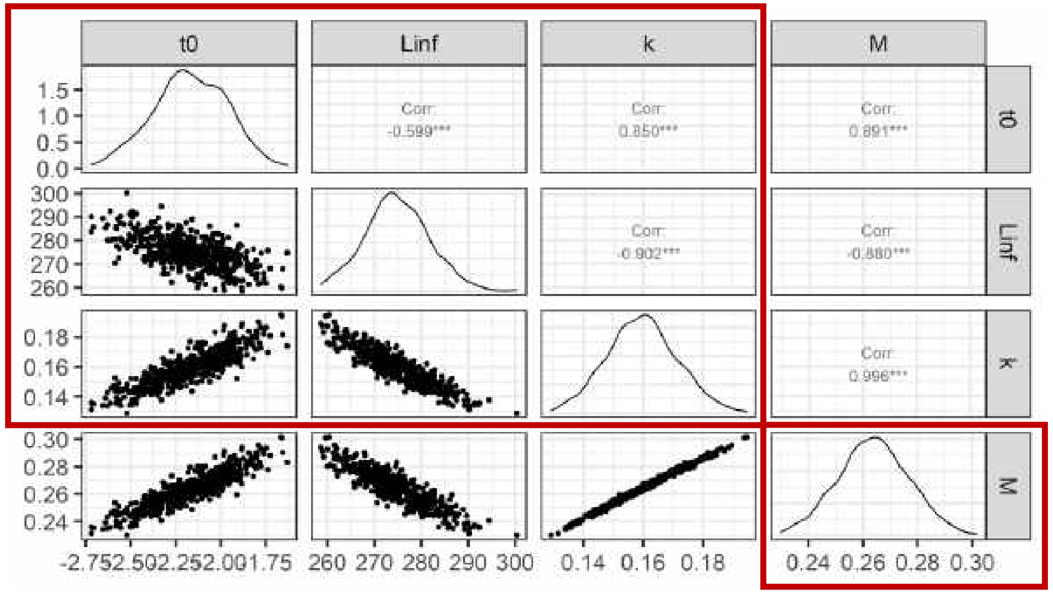
Author(s): Victor R. Restrepo, John M. Hoenig, Joseph E. Powers, James W. Baird, Stephen C. Turner

Cover date: 1992

An Applied Framework for Incorporating Multiple Sources of Uncertainty in Fisheries Stock Assessments

Finlay Scott, Ernesto Jardim, Colin P. Millar, Santiago Cerviño

Published: May 10, 2016 • https://doi.org/10.1371/journal.pone.0154922



e.g. $M = 4.118 \times k^{0.73} Linf^{-0.33}$; *Then et al. 2015*

NOAA FISHERIES

Can we do better? Monte Carlo Bootstrap (MCB)

- Develop a multivariate distribution for key parameters that would be fixed in an assessment.
- Each model in the ensemble would be parametrized by parameters drawn from distribution.
- Benefits?

A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries

Issue: 90(4)

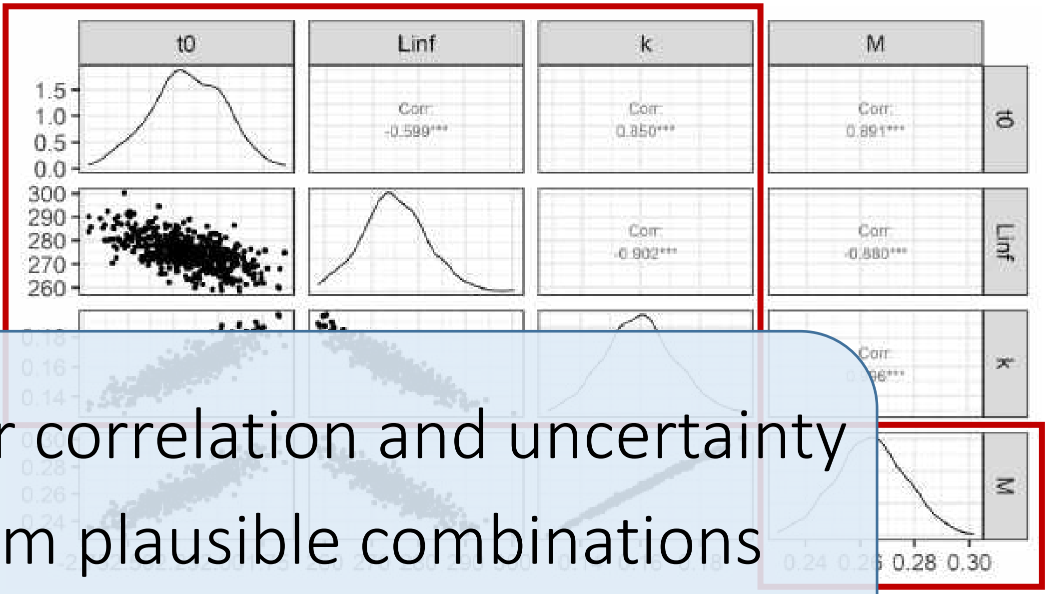
Author(s): Victor R. Restrepo, John M. Hoenig, Joseph E. Powers, James W. Baird, Stephen C. Turner

Cover date: 1992

An Applied Framework for Incorporating Multiple Sources of Uncertainty in Fisheries Stock Assessments

Finlay Scott, Ernesto Jardim, Colin P. Millar, Santiago Cerviño

Published: May 10, 2016 • https://doi.org/10.1371/journal.pone.0154922

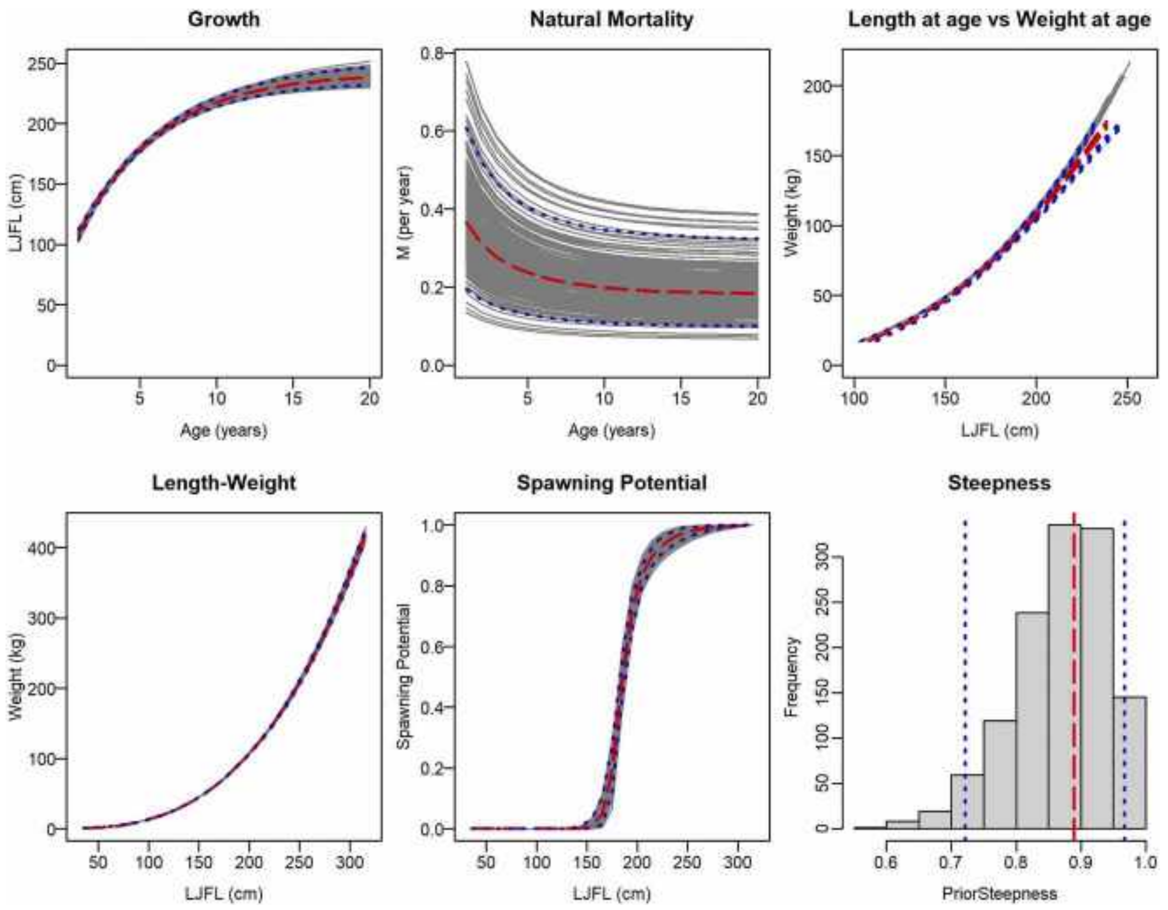


e.g. $M = 4.118 \times k^{0.73} Linf^{-0.33}$; Then et al. 2015

- Preserves parameter correlation and uncertainty
- Life history can inform plausible combinations
- Implicit weighting

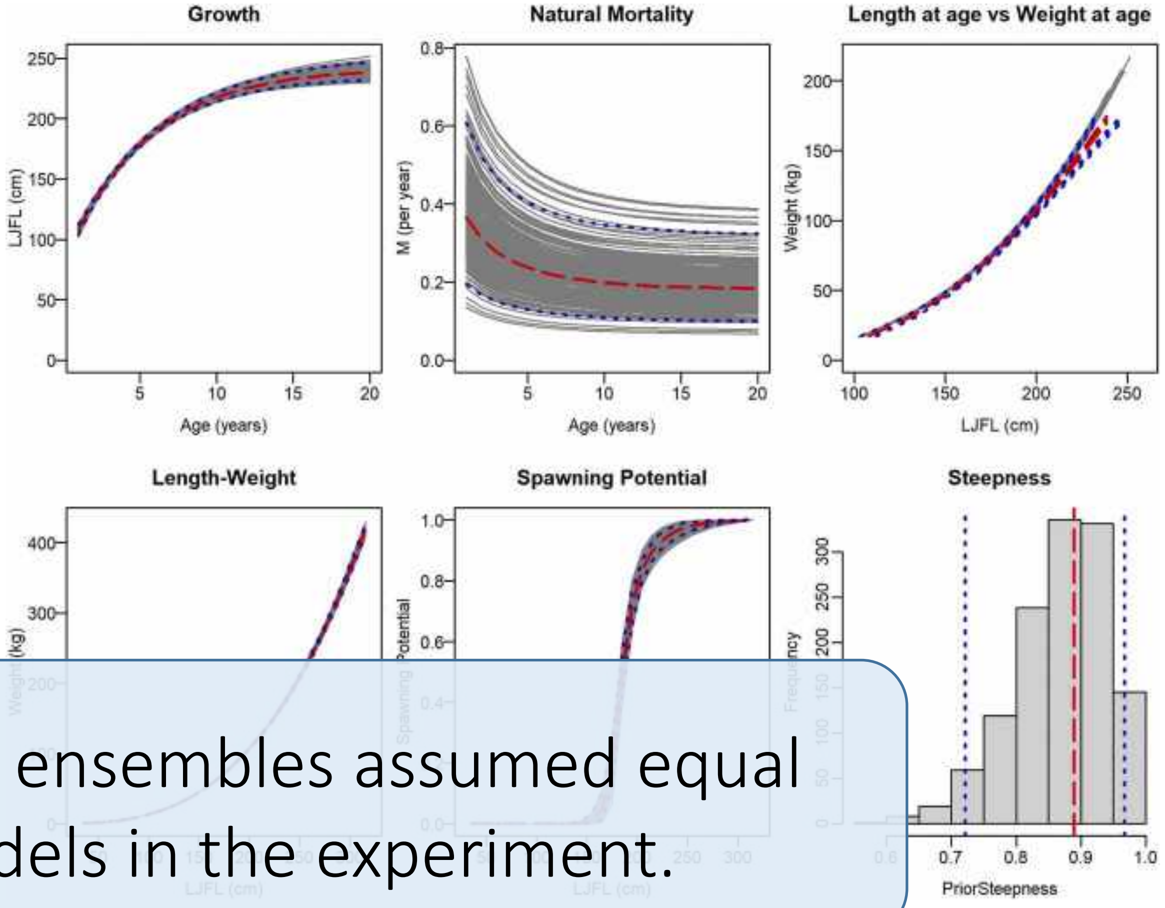
Case study: 2017 SWPO swordfish

- Develop a multivariate distribution for key biological assumptions (growth, maturity, length-weight, natural mortality & steepness).
- Conduct experiment
  - Develop full-factorial ensemble using 3 levels from 5 biological axes (243 models)
  - Compare to MCB ensembles of varying sizes: 30, 50, 75, 100, 200, 300, 500 models



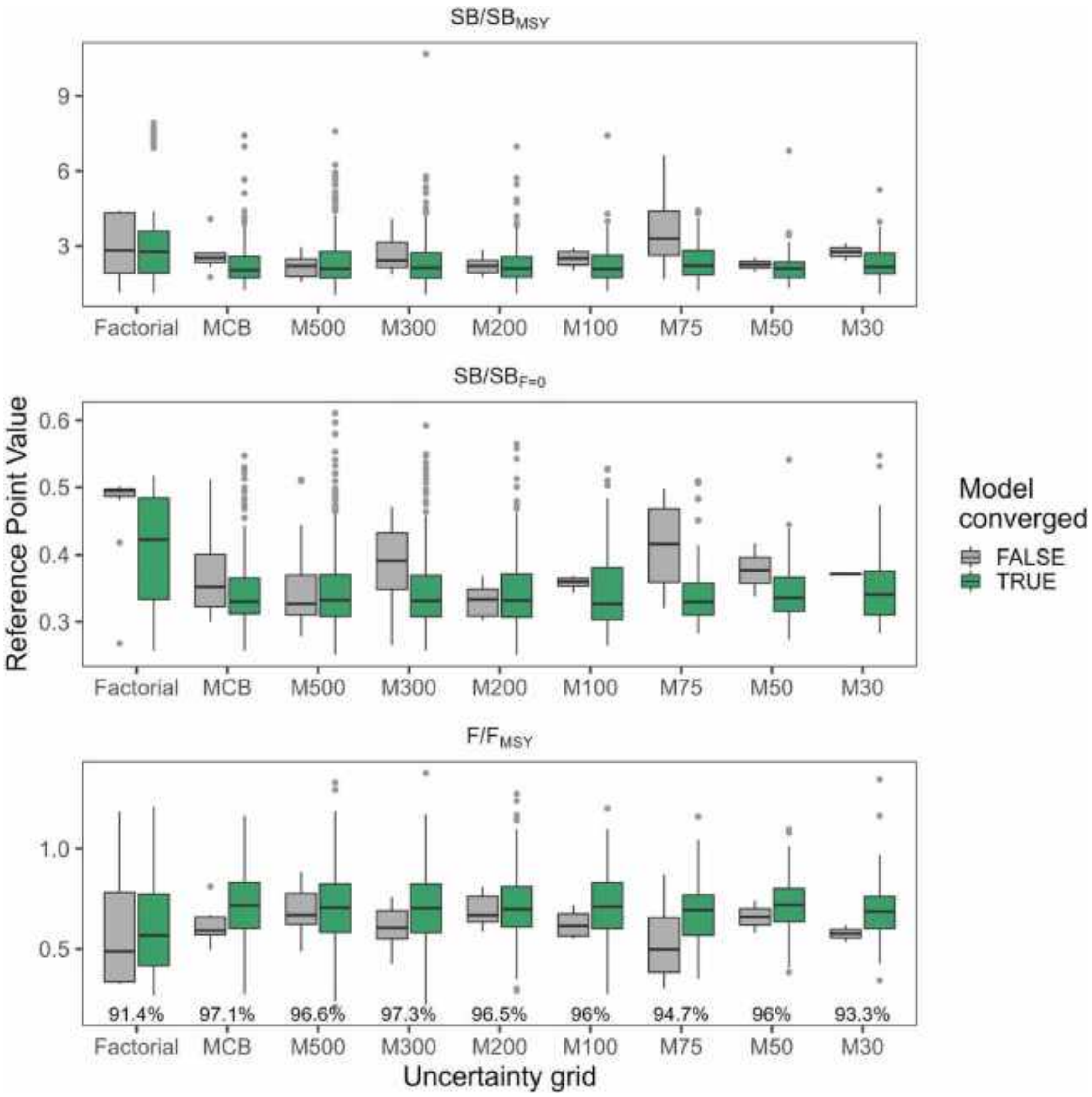
Case study: 2017 SWPO swordfish

- Develop a multivariate distribution for key biological assumptions (growth, maturity, length-weight, natural mortality & steepness).
- Conduct experiment
  - Develop full-factorial ensemble using 3 levels from 5 biological axes (243 models)
  - Compare to MCB ensembles of varying sizes: 30, 50, 75, 100, 200, 300, 500 models



Note: Both types of ensembles assumed equal weights for models in the experiment.

Case study: 2017 SWPO swordfish



- Distributions of the MLEs of management reference points differed in terms of central tendency & uncertainty between the two ensemble types
- MCB ensembles of at least 50 members functionally equivalent in terms of characterizing reference points.

How are model ensembles constructed? Summary

- Ad-hoc combination of models
- Hypothesis tree approach
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution
- Other

- Useful for continuous parameters
- Efficient alternative to full-factorial
- Can reduce uncertainty by self-censoring unlikely combinations
- Shifts scrutiny of weighting decisions from post-hoc discussion to how the multivariate distribution is developed

How are model ensembles constructed? Summary

- Ad-hoc combination of models
- Hypothesis tree approach
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution
- Other

- Useful if number of models small or if capturing uncertainty between discrete choices
- Can be combined with MCB ensemble -> this can be a special case of a hypothesis tree.
- Ensemble dispersion (emphasis on tails) dependent of how levels and weighting is selected

NOAA FISHERIES

How are model ensembles constructed? Summary

- Ad-hoc combination of models
- Hypothesis tree approach
- Full-factorial combination of uncertainties
- Monte-Carlo Bootstrap: fixing parameters to values drawn from a pre-defined distribution
- Other

- Similar pros & cons to the full-factorial approach
- Can be directly linked to conceptual model
- Hierarchy can be tailored to remove redundant and/or unlikely combinations

How are model ensembles combined?

- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED

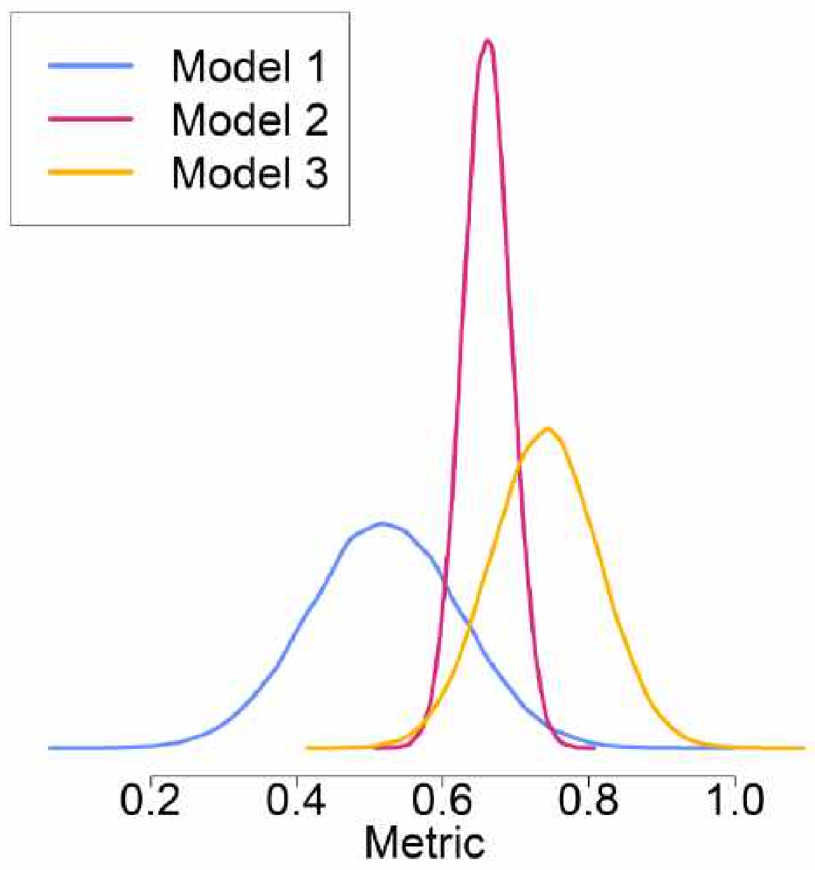
How are model ensembles combined?

& How was uncertainty in the ensemble calculated?

- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED

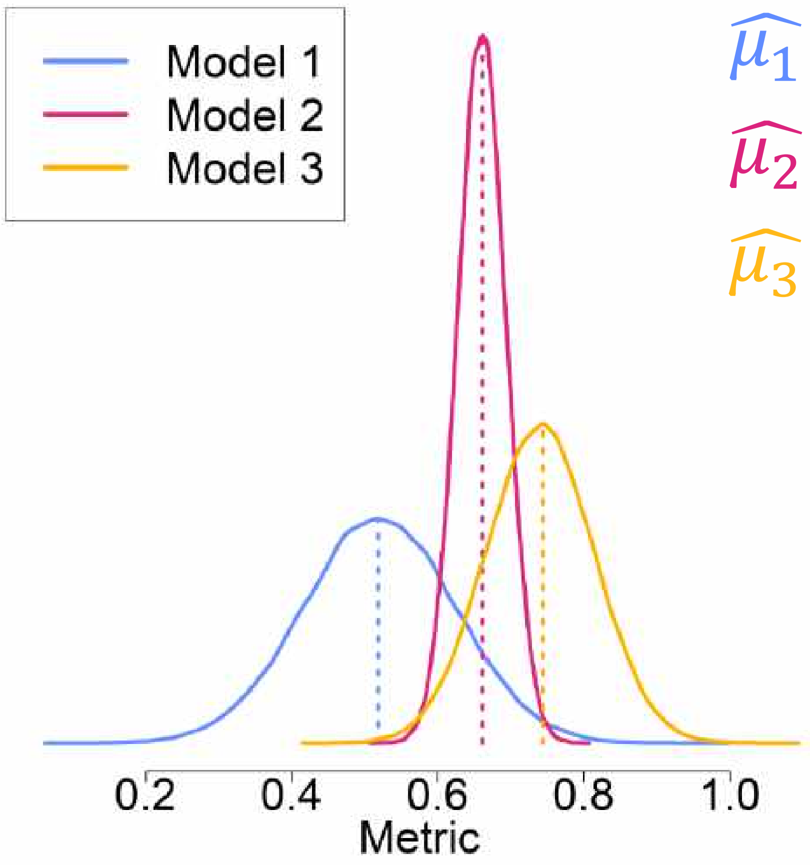
How are model ensembles combined?

- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED



How are model ensembles combined?

- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED



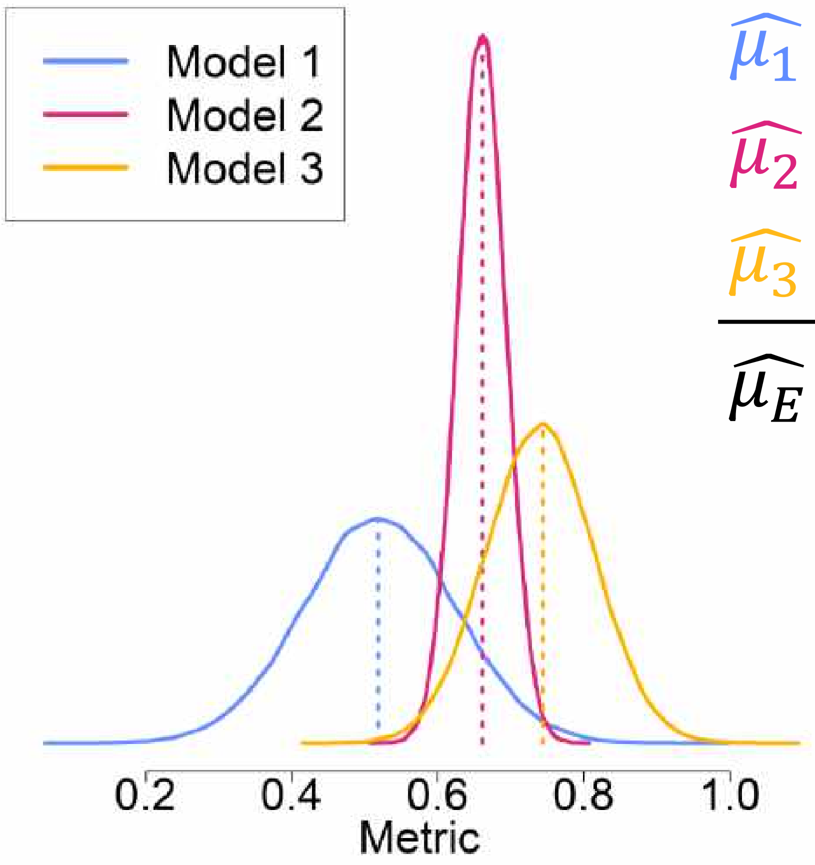
$\widehat{\mu_1} = 0.52$

$\widehat{\mu_2} = 0.66$

$\widehat{\mu_3} = 0.74$

How are model ensembles combined?

- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED



$\widehat{\mu_1} = 0.52$

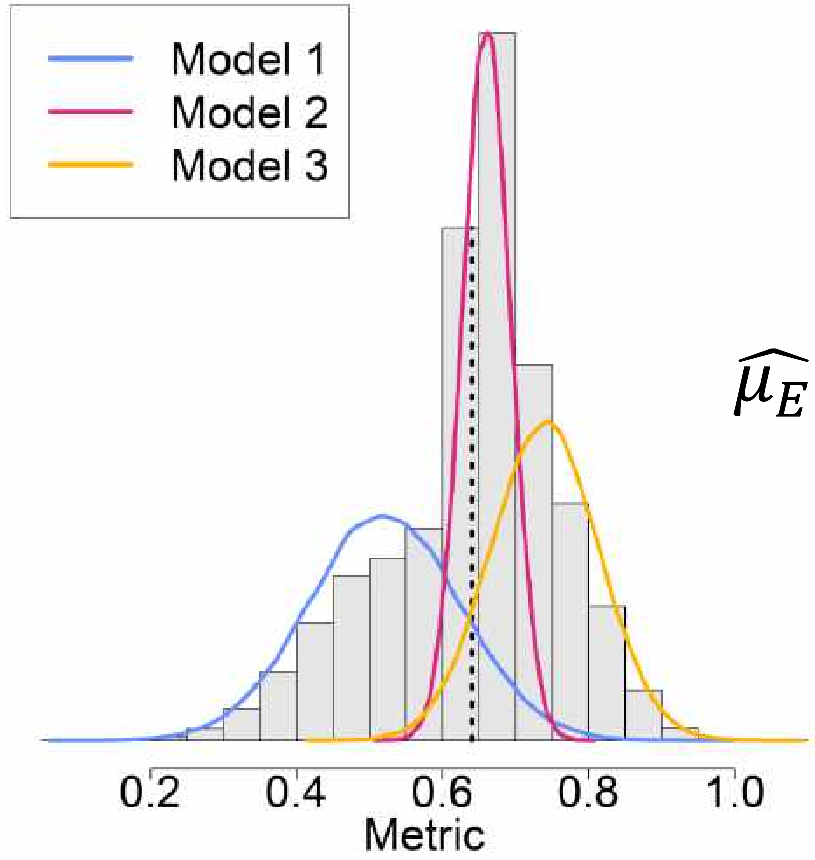
$\widehat{\mu_2} = 0.66$

$\widehat{\mu_3} = 0.74$

$\widehat{\mu_E} = 0.64$

How are model ensembles combined?

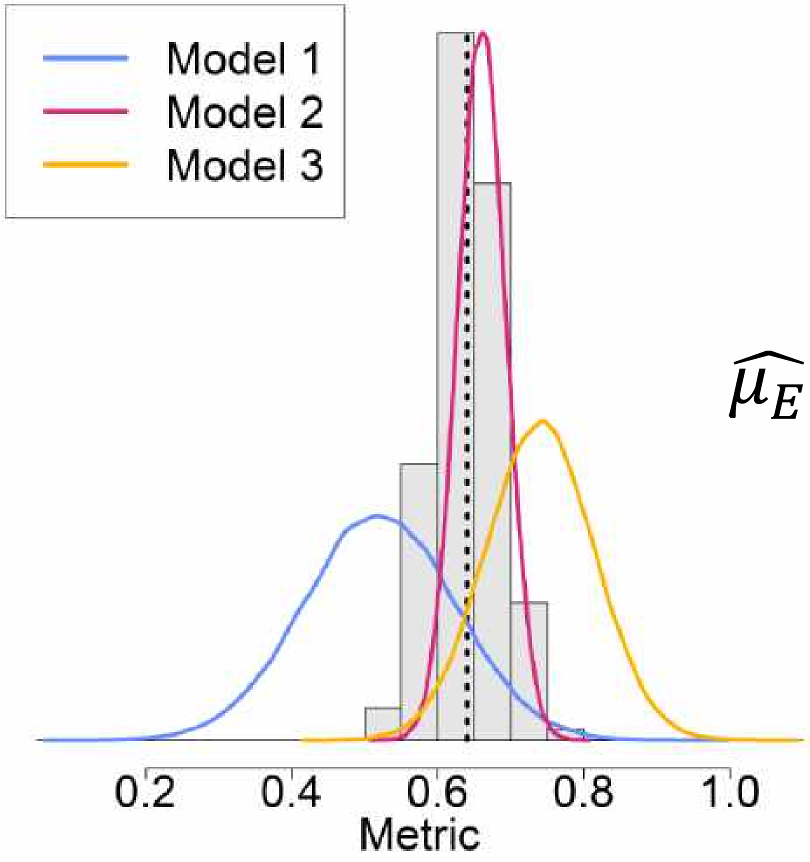
- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED



$\widehat{\mu_E} = 0.64$

How are model ensembles combined?

- Models not combined
- Model estimates AVERAGED together
- Distributions of quantities of interest COMBINED
- Distributions of quantities of interest AVERAGED



$$\widehat{\mu_E} = 0.64$$

How are model ensembles combined?

Ensemble experience	
• Models not combined	8%
• Model estimates AVERAGED together	29%
• Distributions of quantities of interest COMBINED	46%
• Distributions of quantities of interest AVERAGED	17%

NOAA FISHERIES

How are model ensembles combined?

	Ensemble experience	No ensemble experience
Models not combined	8%	27%
Model estimates AVERAGED together	29%	9%
Distributions of quantities of interest COMBINED	46%	36%
Distributions of quantities of interest AVERAGED	17%	27%

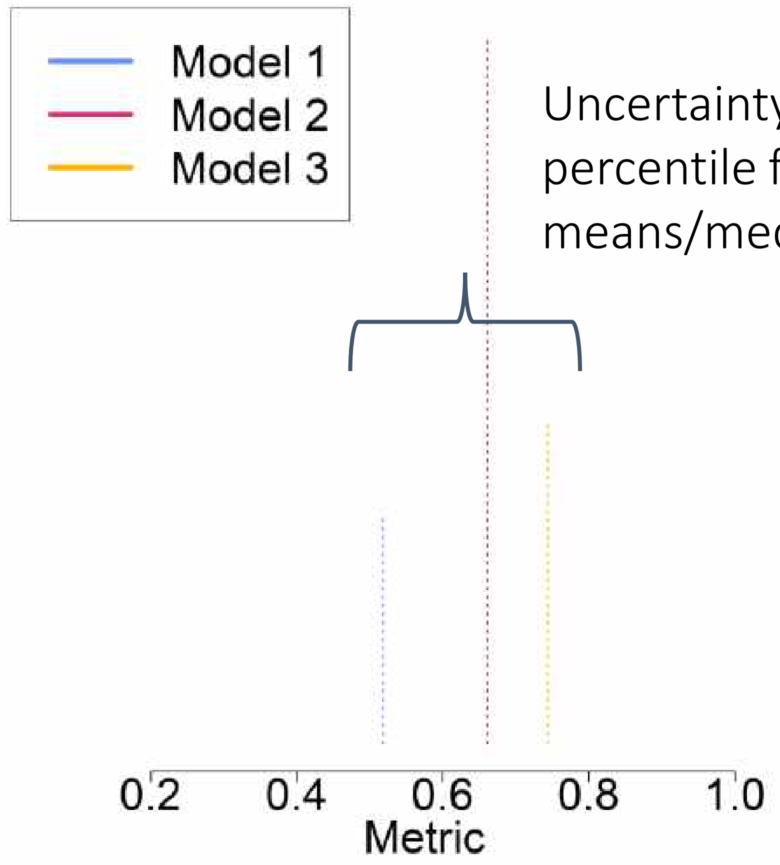
NOAA FISHERIES

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution



Uncertainty based on percentile from model means/medians

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

$$Var\left(\sum_{i=1}^{m} \breve{w}_i \, \hat{\mu}_i\right) = \sum_{i=1}^{m} \breve{w_i^2} \sigma_i^2$$

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

$$Var\left(\sum_{i=1}^{m} \breve{w}_i \hat{\mu}_i\right) = \sum_{i=1}^{m} \breve{w_i^2} \sigma_i^2$$

Ensemble variance decreases as number of models increases; "losing the tails"

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

$$Var\left(\sum_{i=1}^{m} \breve{w}_i \hat{\mu}_i\right) = \sum_{i=1}^{m} \breve{w}_i^2 \sigma_i^2 + \sum_{i=1}^{m} \sum_{j \neq i} \breve{w}_i \breve{w}_j \rho_{ij} \sigma_i \sigma_j$$

How was uncertainty in the ensemble calculated?

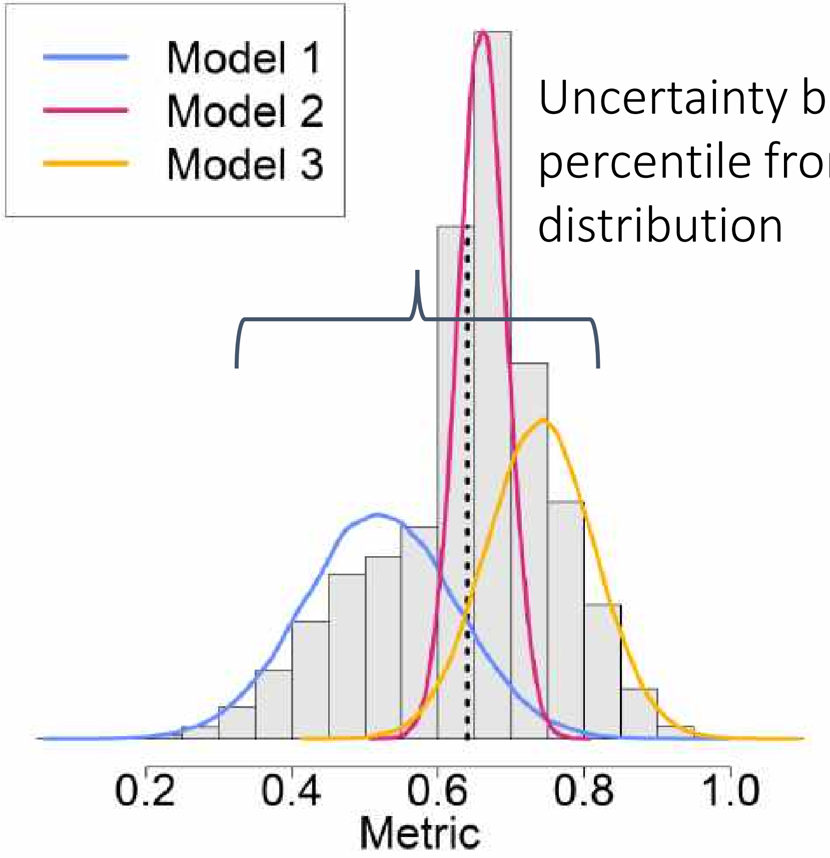
- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

$$Var\left(\sum_{i=1}^{m} \breve{w}_i \, \hat{\mu}_i\right) = \sum_{i=1}^{m} \breve{w}_i^2 \sigma_i^2 + \sum_{i=1}^{m} \sum_{j \neq i} \breve{w}_i \breve{w}_j \rho_{ij} \sigma_i \sigma_j$$

Conservatively assume ρ_{ij} = 1; still can “lose tails”

How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution

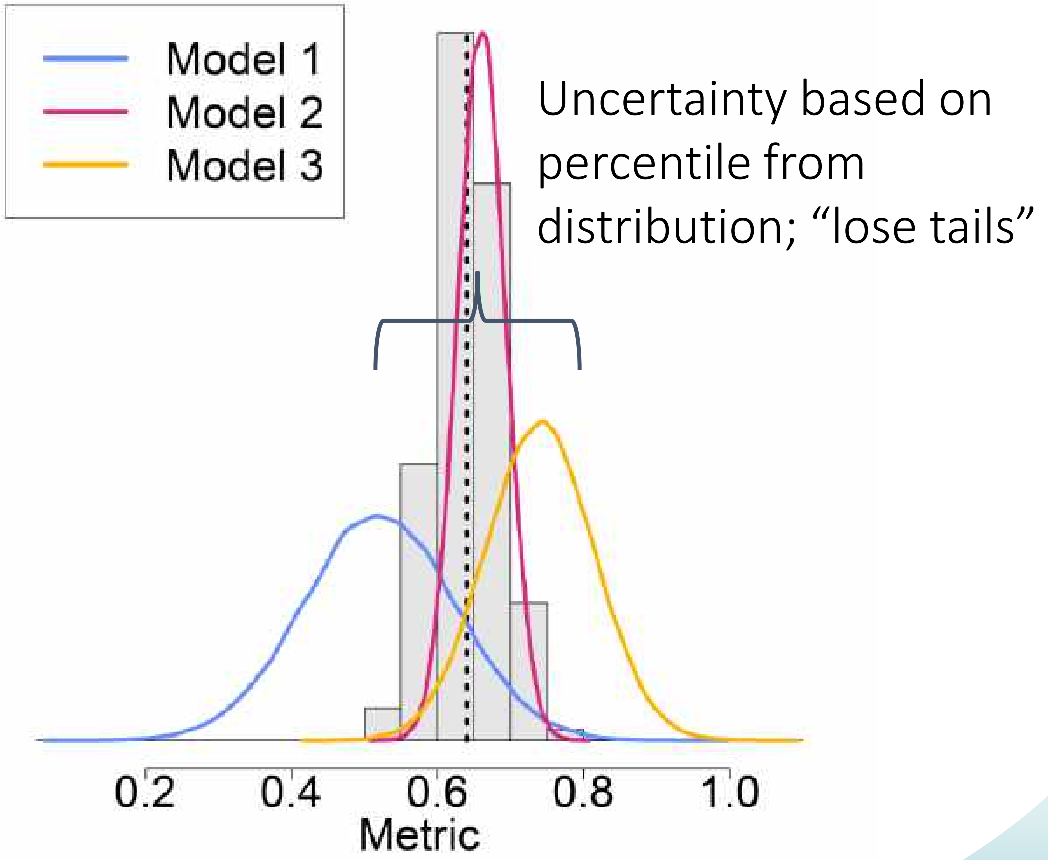


Uncertainty based on percentile from distribution

Mixture distribution, super-distribution, 'model stitching'

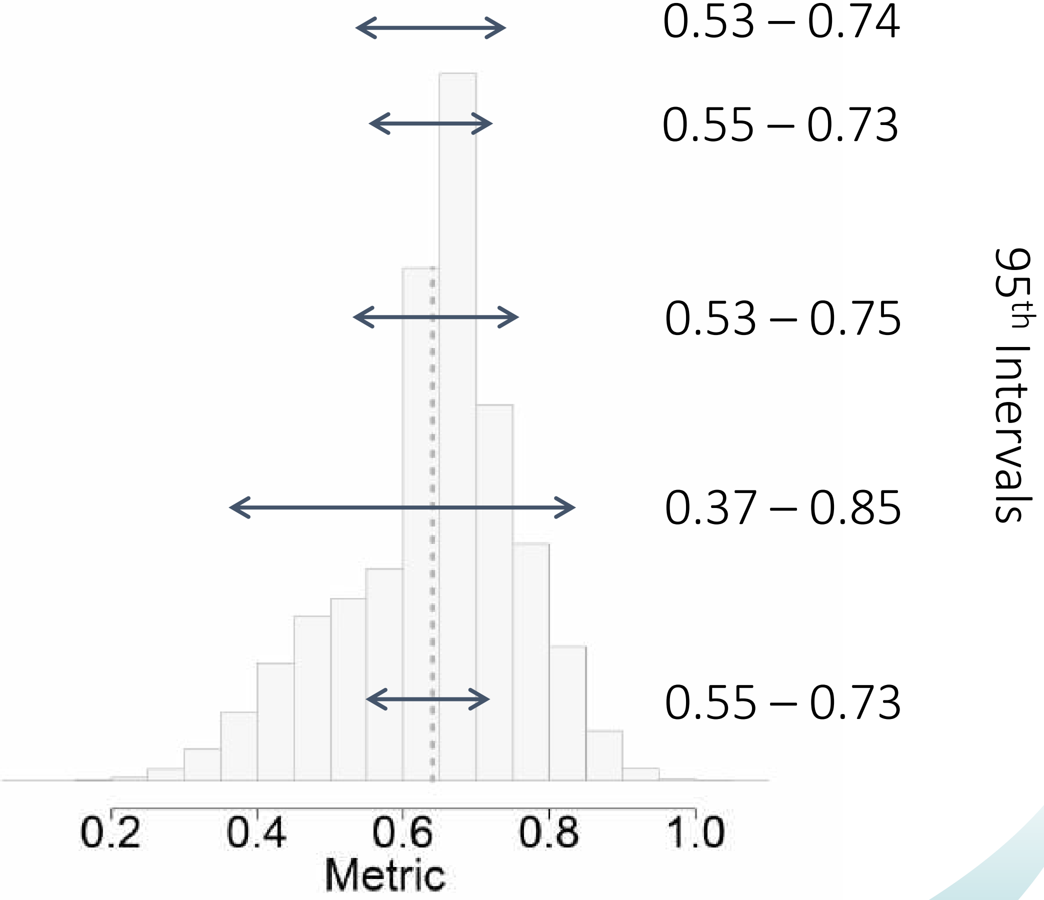
How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution



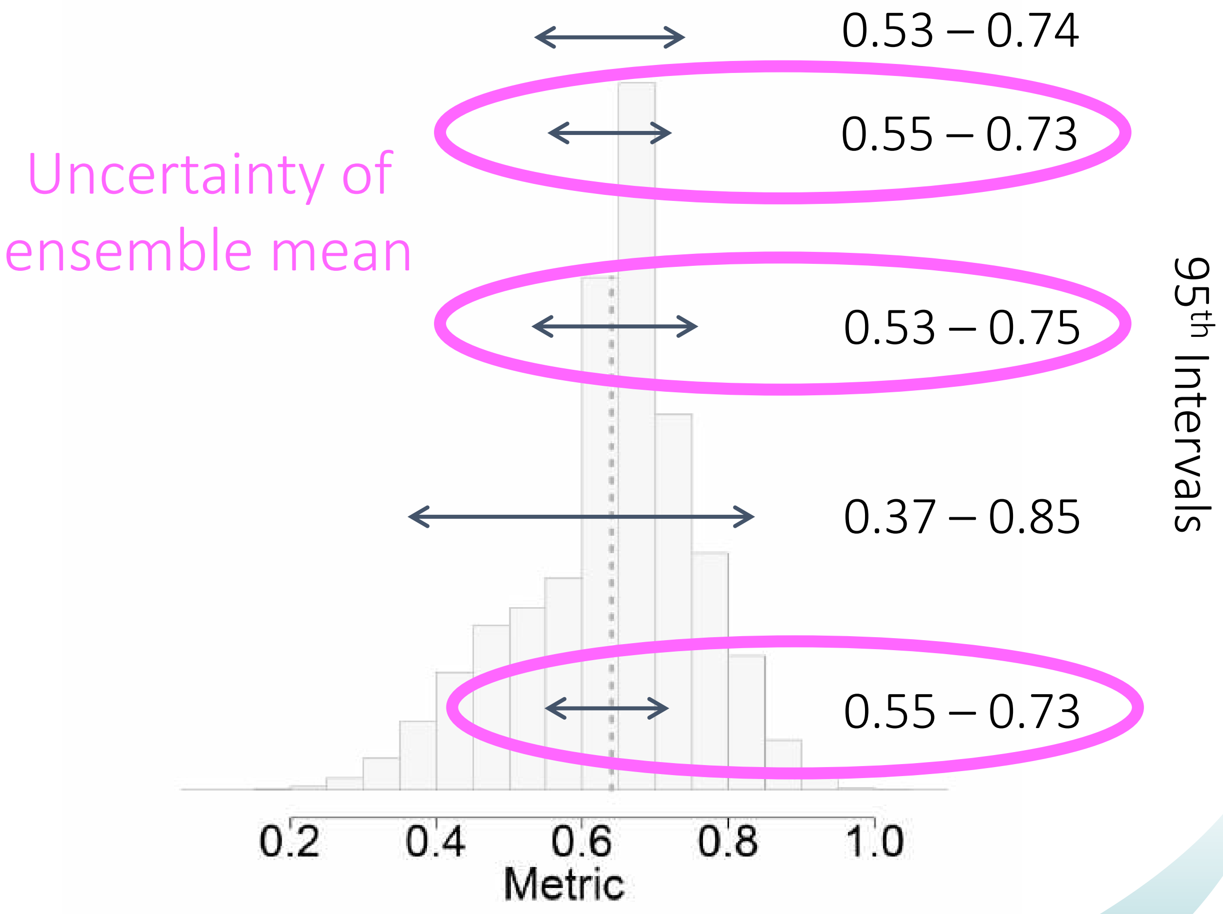
How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution



How was uncertainty in the ensemble calculated?

- Model uncertainty only
- Model + Estimation: Analytical as INDEPENDENT random variables
- Model + Estimation: Analytical as DEPENDENT random variables
- Model + Estimation: Approximated from COMBINED distribution
- Model + Estimation: Approximated from AVERAGE distribution



How was uncertainty in the ensemble calculated?

Model uncertainty only	25%
Model + Estimation: Analytical as INDEPENDENT random variables	8%
Model + Estimation: Analytical as DEPENDENT random variables	4%
Model + Estimation: Approximated from COMBINED distribution	54%
Model + Estimation: Approximated from AVERAGE distribution	8%

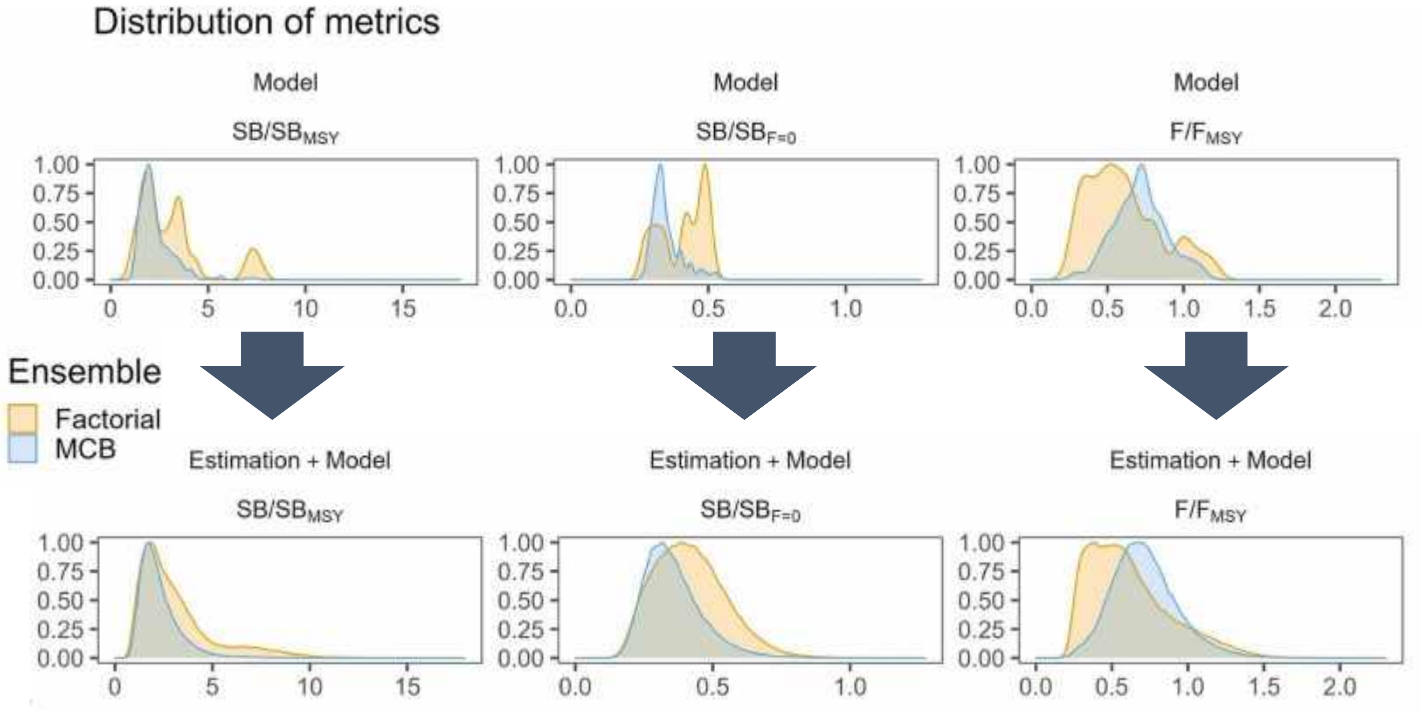
Ensemble experience

NOAA FISHERIES

How was uncertainty in the ensemble calculated?

	Ensemble experience	No ensemble experience
• Model uncertainty only	25%	9%
• Model + Estimation: Analytical as INDEPENDENT random variables	8%	36%
• Model + Estimation: Analytical as DEPENDENT random variables	4%	9%
• Model + Estimation: Approximated from COMBINED distribution	54%	27%
• Model + Estimation: Approximated from AVERAGE distribution	8%	18%

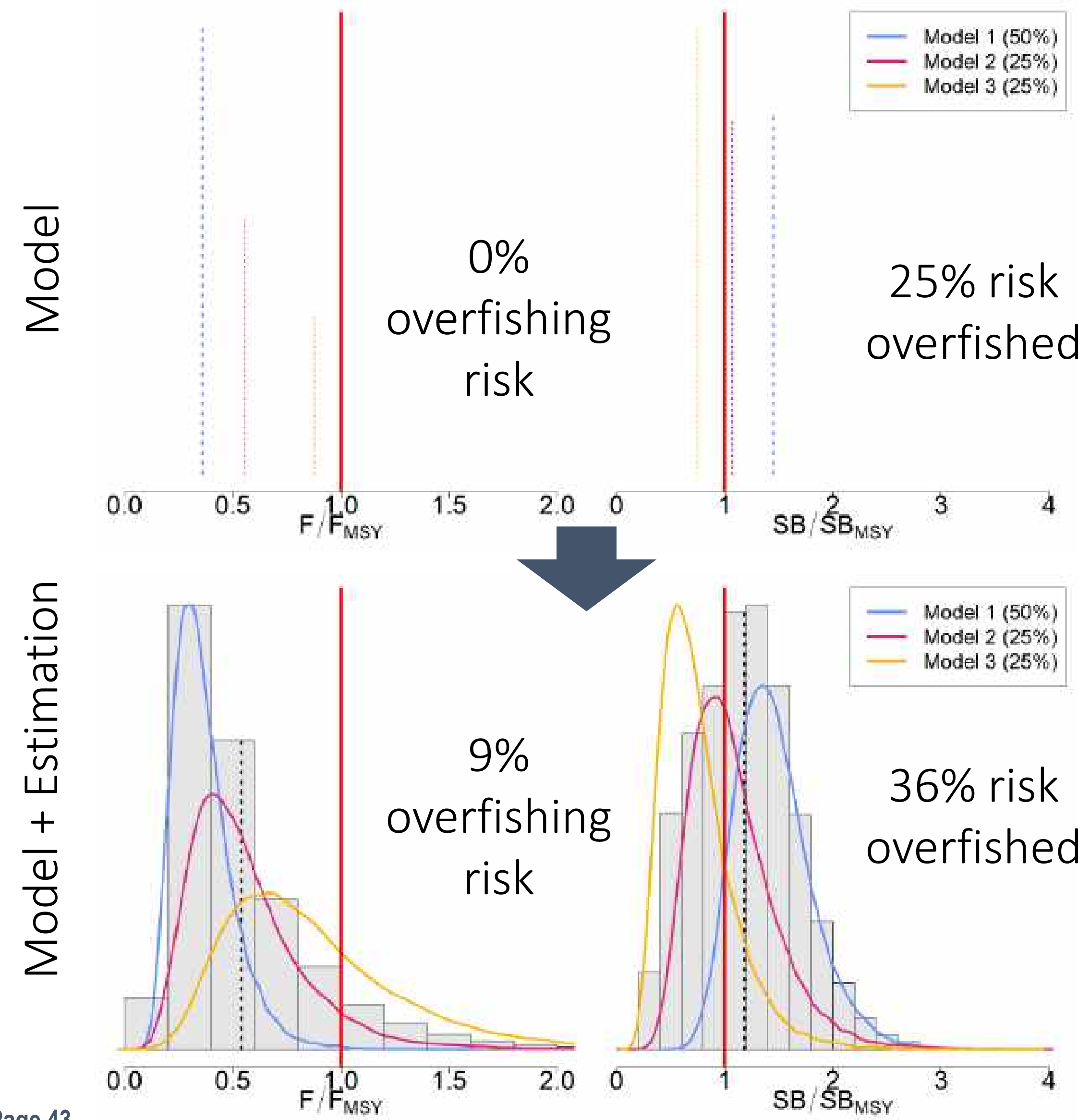
Case study: 2017 SWPO swordfish



	$SB/SB_{MSY} < 1$		$SB/SB_{F=0} < 0.2$		$F/F_{MSY} > 1$	
Error Type	**Factorial**	**MCB**	**Factorial**	**MCB**	**Factorial**	**MCB**
Model	0	0	0	0	10.4	7.2
Estimation + Model	2.6	2.3	1.9	2.4	10.7	9.8

- Mixture distribution combining model + estimation uncertainty created using delta-MVLN like approach
- Some increase in uncertainty but did not meaningfully change risk relative to reference points
- Applying sample-importance resampling (likelihood based) weighting did not change distributions

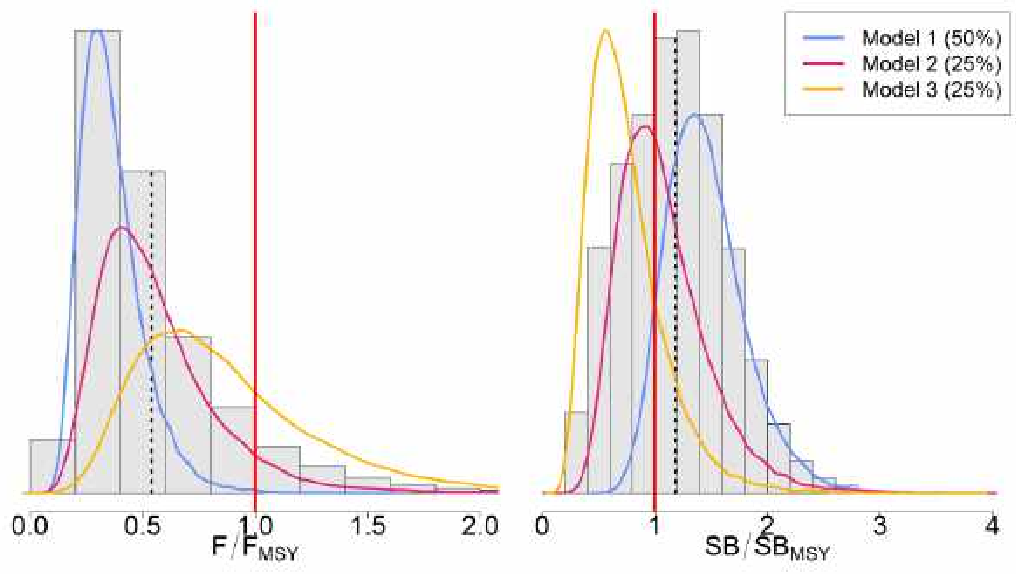
Case study: 2022 NPO blue shark



- 3 model ensemble developed using hypothesis tree approach. Weights assigned *a priori* based on plausibility of each hypothesis. delta-MVLN approach used to create mixture distribution.
- Noticeable increase in risk when estimation uncertainty accounted for.

Case study: 2022 NPO blue shark

Model + Estimation



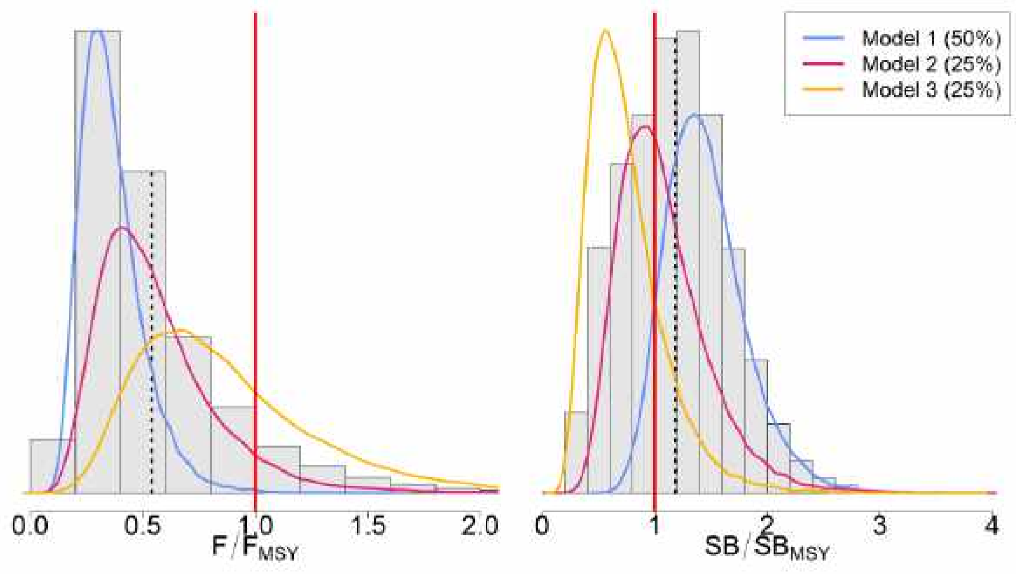
Variance partitioning

$$Var(\widehat{\mu_E}) = \sum_{i=1}^{m} \widetilde{w_i}\sigma_i^2 + \sum_{i=1}^{m} \widetilde{w_i}(\widehat{\mu_i} - \widehat{\mu_E})^2$$

Ducharme-Barth and Vincent (2022), Eq. 4

Case study: 2022 NPO blue shark

Model + Estimation



	F/F_{MSY}	SB/SB_{MSY}
Model	38%	40%
Estimation	62%	60%

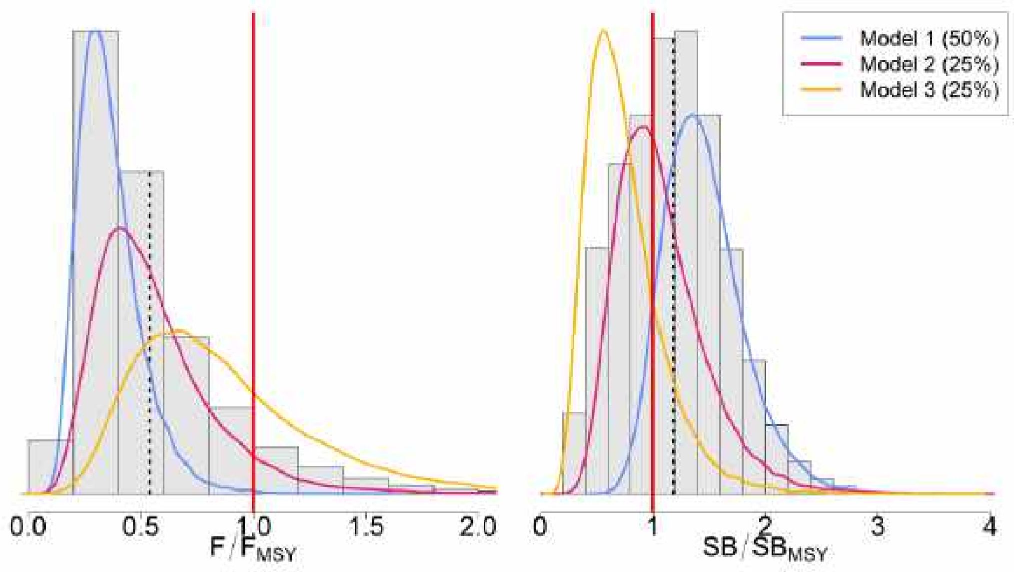
Variance partitioning

$$Var(\widehat{\mu_E}) = \underbrace{\sum_{i=1}^{m} \widetilde{w_i}\sigma_i^2}_{\text{Estimation}} + \underbrace{\sum_{i=1}^{m} \widetilde{w_i}(\widehat{\mu_i} - \widehat{\mu_E})^2}_{\text{Model}}$$

Ducharme-Barth and Vincent (2022), Eq. 4

Case study: 2022 NPO blue shark

Model + Estimation



Variance partitioning

$$Var(\widehat{\mu_E}) = \underbrace{\sum_{i=1}^{m} \widetilde{w_i}\sigma_i^2}_{\text{Estimation}} + \underbrace{\sum_{i=1}^{m} \widetilde{w_i}(\widehat{\mu_i} - \widehat{\mu_E})^2}_{\text{Model}}$$

Ducharme-Barth and Vincent (2022), Eq. 4

	F/F_{MSY}	SB/SB_{MSY}
Model	38%	40%
Estimation	62%	60%
Model 1 (50%)	22%	46%
Model 2 (25%)	15%	19%
Model 3 (25%)	63%	35%

How are model ensembles combined? Summary
& How was uncertainty in the ensemble calculated?

- Most practitioners construct “mixture distributions” which emphasizes “tail retention”. This appears different to some other fields where variance reduction is the focus.
- Some ambiguity in terminology? Model averaging implies some level of variance reduction which is not what is usually achieved.
- Accounting for model + estimation uncertainty will always increase uncertainty when creating “mixture distributions”. More transparent but can change perceived risk levels.
- Individual model variance (and number of models) related to the importance of including estimation uncertainty.
- Stakeholder education and buy-in are critical as decisions related to ensemble construction & combination change perceptions of risk.

Thank you & Questions?

Nicholas Ducharme-Barth, Ph.D.

nicholas.ducharme-barth@noaa.gov

Matthew Vincent, Ph.D.

matthew.vincent.@noaa.gov

NOAA FISHERIES

End slides