Reliability of marine ecosystem models for operational advice: How to assess skills and sensitivity

Sigrid Lehuta - Ifremer, DECOD, Nantes





MEthods for numerical EXploration of COmplex systems



International Council for the Exploration of the Sea

l'Exploration de la Mer

An example : ISIS-Fish for the évaluation of the West-Med management plan
 Stecf context : 1 week meetings, twice a year
 Tors defined by DG MARE



Hake - Gulf of Lion Calibrated 2015-2017 Evaluation of spatial measures

An example : ISIS-Fish for the évaluation of the West-Med management plan
 Stecf context : 1 week meetings, twice a year
 Tors defined by DG MARE

Scenarios clarified on say 2 (if not 3) SISS-Fish Hake - Gulf of Lion Calibrated 2015-2017 Evaluation of spatial measures

An example : ISIS-Fish for the évaluation of the West-Med management plan
 Stecf context : 1 week meetings, twice a year

Tors defined by DG MARE



Evaluation of spatial measures







Complex ecosystem models as decision support tools ?



Marine Policy 61 (2015) 291-302



Fig. 1. Plot of articulation index vs. descriptive accuracy index for the models reviewed in this study, showing the current accuracy frontier. Complexity Costanza and Slar, 1985

 $\widehat{\mathbf{a}}$ Replicability, flexibility, automating processes



Transparency, appropriateness, robustness

☆ Communication, Appropriation

An appropriate lexel to for an ended

Replicability, flexibility, automating processes

Transparency, appropriateness, robustness



Reliability of the model never asked by managers... but first thing fishers questionned Communication, Appropriation

Replicability, flexibility, automating processes



Reliability of the model never asked by managers... but first thing fishers questionned Transparency, appropriateness, robustness

Communication, Appropriation

Best Practices for MEMs in operational advice Validation/ skill assessment/ benchmarking/ evaluation...



Journal of Marine Systems 109-110 (2013) 103-108

Dealing with uncertainty in ecosystem models: The paradox of use for living marine resource management

J.S. Link^{a,*}, T.F. Ihde^b, C.J. Harvey^c, S.K. Gaichas^d, J.C. Field^e, J.K.T. Brodziak^f, H.M. Townsend^b R.M. Peterman^g



Craig A. Aumann*

CW-405, Department of Biological Sciences, University of Alberta, Edmonton, Alta., Canada T6G 2E9

But we can't get enough! The need for paving the way with questions...

- ... that could serve as a template for reporting the exercice
- ... and accomodate the diversity of MEMs (one fits all)

inspired from ICES « key runs »

and Overview Design Details (Grimm et al., 2006)







Skill assessment ecosystem-based

Alexander Kempf^{a,°}, Valerio Bartolino^e, N Contents lists available at ScienceDirect

Fisheries Research

General questions

What is needed for a successful delivery of the advice product?

- What is the advice question?
- Has the model the right complexity?
- Can the model deliver the output needed at the right spatial and temporal scales?
- Which are the most important outputs and metrics for the advice?

What type of model is available and which skill assessment methods are appropriate?

What real world observations are available for skill assessments?

- Are sufficient observations for important outputs in relation to the advice question(s) available?
- Are the observations at the right spatial and temporal scale in relation to the advice question(s)?
- How certain and/or biased are the observations?

Hindcast

Which are the most sensitive parameters and is there room for improvement?

What is the performance of the (final) model hindcasts?

- How good is the agreement between model output and real world observations?
- How large are the estimated parameter uncertainties?
- Are there indications for major structural uncertainties (e.g. identified by sensitivity analyses)?

Are the parameterization and emerging properties from the (final) model sound accorscientific knowledge?

Are there retrospective patterns?

Forecast

What is the predictive skill of the (final) model?

- Do predictions show expected model behavior?
- How does the model perform in cross validations?
- How large are the estimated uncertainties?
- Are there indications for major structural uncertainties (e.g., identified by sensitivity analyses)?
- How is the performance of short-, medium- and long-term forecasts?



Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

- A standard pre Table 1 (continued)
 Benjamin Planque
 EVALUATION EVALUATION
- Cecilie Hansen^b, 1 Torstein Pedersen Leif Christian Stig

Question

17

METHODOLOGY

SENSITIVITIES

- 16 Are sanity checks conducted? If so, what is the method used? If not, explain why.
 - a Which data and patterns are used for this?
 - b Does this apply to patterns that are not otherwise evaluated for this model application?
 - What is the methodology used to compare ecological patterns derived from independent data with patterns from the model? a What is the rationale for choosing this method?
 - b How are observational and/or model output uncertainties handled?
 - c Does the methodology rely on specific assumptions?
 - d Were other methods experimented? If they didn't succeed, explain why.
- 18 Is there a threshold level (match between observed and modelled patterns) that can separate acceptable from unacceptable models?
- 19 How comparable are the patterns derived from the model and those derived from the independent data?
- 20 Has a model sensitivity analysis been performed? If so, how? If not, explain why.
 - a on the model structure?
 - b on the model parametrization?
 - c on other aspects of the model?
- 21 Which elements are the modelled patterns most sensitive to?
 - a input parameters
 - b priors and assumptions
 - c structural elements
 - d processes
- 22 How sensitive are the modelled patterns to the choice of initial conditions, boundary conditions, spatial and temporal resolution?
- 23 How sensitive is the model evaluation to the independent data availability and uncertainty?
- 24 How much is the model evaluation constrained by computational or theoretical limits?
- 25 How does the perceived performance of the model depend on the chosen evaluation methodology?

Model behavior



Skill assessment ecosystem-based

Alexander Kempf^{a,*}, Valerio Bartolino^e, N Contents lists available at ScienceDirect

Fisheries Research

General questions

What is needed for a successful delivery of the advice product?

- What is the advice question?
- Has the model the right complexity?
- Can the model deliver the output needed at the right spatial and temporal scales?
- Which are the most important outputs and metrics for the advice?

What type of model is available and which skill assessment methods are appropriate?

What real world observations are available for skill assessments?

- Are sufficient observations for important outputs in relation to the advice question(s) available?
- Are the observations at the right spatial and temporal scale in relation to the advice question(s)?

Emphasis on hindcast vs. Forecast, technical, Estimation models+

How good is the agreement between model output and real world observations?

- How large are the estimated parameter uncertainties?
- Are there indications for major structural uncertainties (e.g. identified by sensitivity analyses)?

Are the parameterization and emerging properties from the (final) model sound accor ______ ... scientific knowledge?

Are there retrospective patterns?

Forecast

What is the predictive skill of the (final) model?

- Do predictions show expected model behavior?
- How does the model perform in cross validations?
- How large are the estimated uncertainties?
- Are there indications for major structural uncertainties (e.g. identified by sensitivity analyses)?
- How is the performance of short-, medium- and long-term forecasts?





Leif Christian Stig

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmode

A standard pre Table 1 (continued)

SENSI

Benjamin Planque Cecilie Hansen^b, 1 Torstein Pedersen # Question

16 Are sanity checks conducted? If so, what is the method used? If not, explain why.

- a Which data and patterns are used for this?
- b Does this apply to patterns that are not otherwise evaluated for this model application?

17 What is the methodology used to compare ecological patterns derived from independent data with patterns from the model? a What is the rationale for choosing this method?

b How are observational and/or model output uncertainties

Emphasis on patterns, documentation, sensitivity

and modelled ptable models?

succeed, explain

nodel and those

Simulation models+

a on the model structure?

- b on the model parametrization?
- c on other aspects of the model?

21 Which elements are the modelled patterns most sensitive to?

- a input parameters
- b priors and assumptions
- c structural elements
- d processes

24

- 22 How sensitive are the modelled patterns to the choice of initial conditions, boundary conditions, spatial and temporal resolution?
- 23 How sensitive is the model evaluation to the independent data availability and uncertainty?
 - How much is the model evaluation constrained by computational or theoretical limits?

25 How does the perceived performance of the model depend on the chosen evaluation methodology?

A Practical Guide for Conducting Calibration and Decision-Making Optimisation with Complex Ecological Models

8 Stephanie Mahévas*, 8 Victor Picheny, 8 Patrick Lambert, 8 Nicolas Dumoulin, 8 Lauriane Rouan, 8 Jean-Christophe Soulié, 8 Dimo Brockhoff, 8 Sigrid Lehuta, 8 Rodolphe Le Riche, 8 Robert Faivre, 8 Hilaire Drouineau



- The same need dedicated to the calibration (fitting) phase
- The ODD of Optimisation protocole

1.157	Problem Formulation	Model	
		Question	
		Data	
ing		Parameters Bounds&constrainsts	
ess		Uncertainty (process and data)	
00.		Initial objective function	
īd		building	
Cobjectiv	Objective Function	reshaping	
		final	
	Exploratory Analysis	data	
		Reduction dimension	

8	Family	
	Description-Justification	
	Changes in the algorithm	
A	Settings	

Family	
Description-Justification	
Changes in the algorithm	
Settings	

ng	Convergence	
rocessi	Optimum properties including Identifiability	
5	Residual analysis	
Pos	Multicriteria	

An insightful Sensitivity analysis of an Atlantis model Contents lists available at ScienceDire

Atlantis in the Eastern English Channel



	Coments uses available at Scienceptreet	10.1 C 11/10.
	Ecological Modelling	et71
LSEVIER	journal homepage: www.elsevier.com/locate/ecolmodel	

Improving confidence in complex ecosystem models: The sensitivity analysis of an Atlantis ecosystem model

Chloe Bracis^{a,*}, Sigrid Lehuta^b, Marie Savina-Rolland^c, Morgane Travers-Trolet^b, Raphaël Girardin^a

- Morris method
- 4550 simulations 90 parameters 4h/run
- Impact on biomass (40 groups) and system stability



An exemple of insightful Sensitivity analysis of a complex model

- Atlantis in the Eastern English Channel
- Top down control : unrealistic behavior of top predateurs
- better represent competition among top predators and the effects of limited food
- Bottom-up control : growth rates of phytoplankton and zooplankton
- Improve representation, coupling/inputs from NPZ models
- Strong benthopelagic coupling in the system consistent with shallow system
- Distinction betw. parameters affecting all groups vs. their own group
- → Implication for calibration



Skill assessment for testing assumptions

ISIS-Fish for Sole in the Eastern English Channel

(Kotthaus, 1963 ; Coggan & Dando, 1988 ; Riou et al., 2001 ; Burt & Millner, 2008 ; Rochette et al., 2012 ; Le Pape & Cognez et al., 2016 ; Du Pontavice et al., 2018)





Calibation (GA) of 3 alternative models - 11 parameters - 1h/run

- Comparative model skill assessment
- Implication for MSY estimates

Skill assessment for testing assumptions

- ISIS-Fish for Sole in the Eastern English Channel
- More differences betw. variables than models

model

Sole

- Complementarity of metrics
- F and SSB less reliable than other outputs





Sole3pop

SoleMetasole

Combining sensitivity and skill assessment to calibrate

Larval drift IBM for Sole in the North Sea

MARINE ECOLOGY PROGRESS SERIES Mar Ecol Prog Ser	In press	
		-

Lessons from the calibration and sensitivity analysis of a fish larval transport model

Léo Barbut^{1,2,*}, Sigrid Lehuta³, Filip A.M. Volckaert², Geneviève Lacroix¹

- 8 parameters (1 discret) 12h/1 year r
- Observations available to assess model skills over 12 years
- Step 1 : SA (optimal design)
- → Most influent parameters
- Step 2 : MSA (full factorial design)



Combining sensitivity and skill assessment to calibrate

Larval drift IBM for Sole in the North Sea



- ⇒ Validation of model behavior
- ⇒ Identification of factors influencing recruitment success
- Preliminary calibration over 12-years at managable cost
- Dimension reduction allows for a quantitative calibration



C Replicability, flexibility, automating processes



Transparency, appropriateness, robustness

• Sharing commun knowledge, representations, tools, uncertainties

Objectives:

- Demystifying models, explain uncertainties
- Exchange knowledge and representations
- Learning by doing







• Sharing commun knowledge, representations, tools, uncertainties

- 2 workshops with fishers representatives
 - WK 1 : build your own model to predict biomass and catch in 5 years (25 pers.)
- Conclusions
- Highly complex models
- Unexpected uncertainties and detailed compartements



• Sharing commun knowledge, representations, tools, uncertainties

- 2 workshops with fishers representatives
 - WK 1 : build your own model to predict biomass and catch in 5 years (25 pers.)
- Conclusions
- Highly complex models
- Unexpected uncertainties and detailed compartements



Advice time and science time...

Vote for your prefered representation of uncertainty

Question 1

Incertitude en sortie du modèle - Quelle(s) représentation(s) comprenez-vous le mieux ?



- Need for a practical guide for thr interpretation of graphs
- With fishers, make sentences !



Tackle big challenges in small chunks...



 \bigcirc with appropriate tools :

Maths, templates, common language...

From various disciplines (IT, Social sciences, mathematics...)

Compatibility of « Advice time » ?

Thank you for your attention





Martin P. Marzloff^{a,b,*}, Craig R. Johnson^a, L. Rich Little^b, Jean-Christophe Soulié^c, Scott D. Ling^a, Stewart D. Frusher^a



rocky reefs dynamics

Journal of Marine Systems 76 (2009) 95-112

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, C08001, doi:10.1029/2006JC003852, 2007

ICES Journal of Marine Science

ICES Journal of Marine Science (2016), 73(7), 1715-1724. doi:10.1093/icesjms/fsw047

Quo Vadimus

A guinea pig's tale: learning to review end-to-end marine ecosystem models for management applications

Isaac C. Kaplan^{1*} and Kristin N. Marshall²

¹Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2725 Montlake Boulevard E., Seattle, WA 98112, USA
²University of Washington School of Aquatic and Fisheries Sciences, PO Box 355020, Seattle, WA 98195, USA

*Corresponding author: tel: +1-206-302-2446; fax: +1-206-860-3394; e-mail: isaac.kaplan@noaa.gov

Kaplan, I. C., and Marshall, K. N. A guinea pig's tale: learning to review end-to-end marine ecosystem models for management applications. – ICES Journal of Marine Science, 73: 1715–1724.

And applications...

Fisheries Research 143 (2013) 57-66

Geosci. Model Dev., 9, 59–76, 2016 www.geosci-model-dev.net/9/59/2016/ doi:10.5194/gmd-9-59-2016 © Author(s) 2016. CC Attribution 3.0 License.



1

The assessment of a global marine ecosystem model on the basis o emergent properties and ecosystem function: a case study with ERSEM

L. de Mora, M. Butenschön, and J. I. Allen Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth, PL1 3DH, UK *Correspondence to:* L. de Mora (ledm@pml.ac.uk)

Chinese Journal of Oceanology and Limnology

http://dx.doi.org/10.1007/s00343-016-5068-3

Discussion of skill improvement in marine ecosystem dynamic models based on parameter optimization and skill assessment*

SHEN Chengcheng (沈程程)^{1,2}, SHI Honghua (石洪华)^{2,**}, LIU Yongzhi (刘永志)³, LI Fen (李芬)³, DING Dewen (丁德文)²



Geoscientific

Model Development



Received 31 Ju

Assessm

models:





ICES	International Council for the Exploration of the Sea	
CIEM	Conveil International pour l'Exploration de la Plar	