Dealing with uncertainty in complex models an application to the OSMOSE ecosystem model of the northern Humboldt current system

Doris Criscely Luján Paredes

▶ **To cite this version:**

Doris Criscely Luján Paredes. Dealing with uncertainty in complex models an application to the OSMOSE ecosystem model of the northern Humboldt current system. Ecosystems. Université Paris-Saclay, 2022. English. NNT : 2022UPASB013 . tel-04860630

HAL Id: tel-04860630

https://theses.hal.science/tel-04860630v1

Submitted on 1 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

Dealing with uncertainty in complex models: an application to the OSMOSE ecosystem model of the northern Humboldt current system

Faire face à l'incertitude dans les modèles complexes: une application au modèle écosystémique OSMOSE dans l'écosystème nord du courant de Humboldt

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 567, Sciences du Végetal: du gène à l'ecosystème (SEVE),
Spécialité de doctorat: Biologie,
Graduate School: BioSphERA - Biologie, Société, Ecologie Environnement, Ressources, Agriculture Alimentation,
Référent : Faculté des sciences d'Orsay

Thèse préparée dans l' unité de recherche Ecologie Systématique et Evolution (Université Paris-Saclay, CNRS, AgroParisTech)
sous la direction de **Paul LEADLEY**, professeur,
la co-direction de **Yunne-Jai SHIN**, directrice de recherche,
le co-encadrement de **Ricardo OLIVEROS-RAMOS**, chercheur

Thèse soutenue à Paris-Saclay, le 24 mars 2022, par

Doris Criscely LUJÁN PAREDES

Composition du jury

Philippe CURY Directeur de recherche, IRD, Université de Montpellier	Président
Verena TRENKEL Chercheuse, HDR, IFREMER Centre Atlantique, Nantes	Rapportrice & Examinatrice
Eva PLAGANYI Chercheuse, HDR, CSIRO Queensland Biosciences Precinct	Rapportrice & Examinatrice
Dimitri GUTIERREZ Directeur de recherche, IMARPE, UPCH	Invité
Dorleta GARCIA Chercheuse, AZTI	Invitée
Paul LEADLEY Professeur, Université Paris-Saclay	Directeur de thèse

THESE DE DOCTORAT

NNT : 2022UPASB013

Abstract

Models of increasing complexity are being developed to improve the realism in the representation of natural systems. However, as these complex models are often perceived as highly uncertain, it becomes critical to characterize and quantify their sources of uncertainty and explore the degree of confidence that can be placed in their predictions. Techniques such as sensitivity analyses and uncertainty analyses are frequently used for this purpose, both addressing uncertainty in different ways but sharing the need to run large ensembles of model simulations. This can be technically limiting when dealing with complex models due to their intrinsic characteristics. In this context, this research work addresses two main questions: the first focused on how to deal with uncertainty in complex models, and the second focused on the impact of uncertainty on the diverse outcomes produced by these models. These questions will be explored from a theoretical to a practical angle, by studying uncertainty in the OSMOSE marine ecosystem model applied to the northern Humboldt current ecosystem (the NHCE OSMOSE model).

First, this thesis presents a literature review dedicated to the search for methods and tools applied to the study of uncertainty with potential applications to biodiversity models. In this context, these models are defined as any mathematical model (with elements such as state variables, forcing, mathematical equations, and parameters) that covers a broad spectrum of ecological organization levels (from individuals to ecosystems).

We then implemented an uncertainty analysis on the NHCE OSMOSE model, intending to quantify the effect of parameter uncertainty on a set of ecological indicators derived from this model. For this, we worked with two uncertainty scenarios for the model parameters that describe the modelled species, corresponding to a 10% and 20% coefficient of variation around the reference values of the parameters. This study revealed that the uncertainty in the parameters of a few species has strong repercussions on the rest of the food web; this was the case for euphausiids and Humboldt squid. In addition, this work allowed to identify a

set of ecological indicators that were more robust to input parameter uncertainty, such as the slope of the size spectrum, the marine trophic index and the trophic level of fish communities.

The results of the uncertainty analysis opened up a set of questions that we addressed by proposing a protocol dedicated to the realization of a sensitivity analysis of a complex model. We show the drawbacks of using arbitrary variability ranges to represent uncertainty in model inputs. This is a frequent problem when working with models with a large number of parameters, for which there is a lack of information to determine their uncertainty. To address this issue, we propose the "parameter reliability criterion". This criterion serves to classify model parameters according to the source of information used to estimate their values and assign an uncertainty level to model parameters. We illustrate the use of this protocol by implementing a sensitivity analysis using the NHCE OSMOSE model.

Finally, we conclude that the inherent complexities of an ecosystem model should not be used as an argument to avoid studying uncertainty appropriately. Computational methods and tools are available in the literature. However, they should be selected carefully considering the potential limitations that the use of a model may generate. The uncertainty level of a complex model depends strongly on the uncertainty in model inputs, its propagation in time, but also on the outputs studied. In this research, we highlight the difficulties of studying uncertainty in a complex model (including but not limited to ecosystem models) and propose possible solutions and guidelines for future studies addressing this research problem. Our results are only a first step, starting with the study of parameter uncertainty, but leaving as a perspective the study of other sources of uncertainty such as in model forcing, initial conditions, and uncertainty due to model structure.

Keywords: complex models, uncertainty analysis, sensitivity analysis, ecosystem modelling, OSMOSE, northern Humboldt current ecosystem.

Résumé

Des modèles de plus en plus complexes sont développés dans le but d'améliorer le réalisme de la représentation des systèmes naturels. Cependant, comme ces modèles complexes sont souvent perçus comme très incertains, il est nécessaire de caractériser et de quantifier leurs sources d'incertitude, ainsi que d'explorer le degré de confiance que l'on peut accorder à leurs prédictions. À cette fin, des techniques telles que l'analyse de sensibilité et l'analyse d'incertitude sont fréquemment utilisées. Les deux techniques abordent l'incertitude de différentes manières, mais elles ont en commun l'exécution de multiples simulations du modèle, ce qui peut être techniquement limitant dans le cas de modèles complexes. Dans ce contexte, ce travail de recherche aborde deux questions principales: dans la première nous explorons quelles méthodes permettent de traiter au mieux l'incertitude dans les modèles complexes, et la seconde question porte sur la quantification de l'impact de l'incertitude sur les différents résultats produits par ces modèles. Ces questions seront abordées d'une manière théorique et pratique, en étudiant l'incertitude du modèle d'écosystème marin OSMOSE appliqué à l'écosystème nord du courant de Humboldt (le modèle NHCE OSMOSE).

En premier lieu, cette thèse propose une revue de la littérature dédiée à la recherche de méthodes et d'outils appliqués à l'étude de l'incertitude avec des applications potentielles aux modèles de biodiversité. Dans ce contexte, ces modèles sont définis comme tout modèle mathématique (avec des éléments tels que des variables d'état, des forçages du modèle, des équations mathématiques et des paramètres) qui couvre un large spectre de niveaux d'organisation écologique (allant des individus aux écosystèmes).

Ensuite, nous mettons en œuvre une analyse d'incertitude du modèle NHCE OSMOSE, dans le but de quantifier l'effet de l'incertitude de ses paramètres sur un ensemble d'indicateurs écologiques en sortie de ce modèle. Pour cela, nous avons travaillé avec deux scénarios d'incertitude utilisant les paramètres du modèle et correspondant à un coefficient de variation de 10% et 20% autour des valeurs de référence des paramètres décrivant chaque espèce modélisée. Les résultats de

cette étude révèlent que l'incertitude des paramètres de certaines espèces a de fortes répercussions sur le reste du réseau trophique; ce fut le cas des euphausiacées et du calmar géant. En outre, ce travail nous a permis d'identifier les indicateurs écologiques les plus robustes à l'incertitude des paramètres d'entrée, tels que la pente du spectre de taille, l'indice trophique marin et le niveau trophique des communautés de poissons.

Les résultats de l'analyse d'incertitude ont ouvert un ensemble de questions que nous proposons d'aborder à travers l'élaboration d'un protocole guidant la réalisation d'une analyse de sensibilité d'un modèle complexe. De plus, nous montrons les biais potentiels introduits par l'utilisation de gammes de variation arbitraires pour représenter l'incertitude dans les paramètres du modèle. C'est un problème fréquemment rencontré avec les modèles comportant un grand nombre de paramètres, qui ne disposent pas toujours d'informations relatives à leur incertitude. Pour y remédier, nous proposons le "critère de fiabilité des paramètres". Ce critère sert à classer les paramètres du modèle en fonction de la source d'information utilisée pour estimer leurs valeurs, et d'attribuer un niveau d'incertitude aux paramètres d'un modèle. Nous illustrons l'utilisation de ce protocole en mettant en œuvre une analyse de sensibilité utilisant le modèle NHCE OSMOSE.

Enfin, nous concluons que les complexités inhérentes aux modèles complexes ne doivent pas être utilisées comme une excuse pour éviter d'étudier leur incertitude de manière appropriée. Des méthodes et des outils informatiques sont disponibles dans la littérature. Cependant, ils doivent être sélectionnés avec soin en tenant compte des limitations potentielles que l'utilisation d'un modèle peut générer. Le niveau d'incertitude d'un modèle complexe dépend fortement de l'incertitude sur les entrées du modèle, de sa propagation dans le temps, mais aussi des sorties étudiées. Dans cette recherche, nous mettons en évidence les difficultés des analyses d'incertitudes pour les modèles complexes (incluant les modèles écosystémique mais ne s'y limitant pas) et proposons des solutions possibles ainsi que des lignes directrices pour les études futures traitant de ce problème de recherche. Nos résultats ne constituent qu'une première étape, en commençant par l'étude de l'incertitude des paramètres, mais en laissant comme perspective l'étude d'autres sources d'incertitude telles que dans les forçages du modèle, les conditions initiales, et la structure du modèle.

Mots-clés: modèles complexes, analyse d'incertitude, analyse de sensibilité, modélisation des écosystèmes, OSMOSE, écosystème nord du courant de Humboldt.

Resumen

Modelos cada vez más complejos se desarrollan con el objetivo de mejorar el realismo en la representación de los sistemas naturales. Sin embargo, dado que estos modelos complejos suelen percibirse como muy inciertos, es necesario caracterizar y cuantificar sus fuentes de incertidumbre y explorar el grado de confianza que puede depositarse en sus predicciones. Para ello, se utilizan con frecuencia técnicas como los análisis de sensibilidad y los análisis de incertidumbre. Ambas técnicas abordan la incertidumbre de manera distinta, pero comparten la necesidad de ejecutar múltiples simulaciones del modelo en uso, lo que puede resultar técnicamente limitante cuando se utilizan modelos complejos debido a sus características intrínsecas. En este contexto, este trabajo de investigación aborda dos preguntas principales: la primera se centra en investigar cómo tratar la incertidumbre en modelos complejos, y la segunda en cómo estudiar el impacto de la incertidumbre en los diversos resultados producidos por estos modelos. Estas preguntas serán exploradas desde un contexto teórico hasta uno práctico, estudiando la incertidumbre en el modelo ecosistémico marino OSMOSE aplicado al ecosistema norte de la corriente de Humboldt (el modelo NHCE OSMOSE).

Primero, esta tesis se centra en una revisión de literatura dedicada a la búsqueda de métodos y herramientas aplicadas al estudio de la incertidumbre con potenciales aplicaciones a modelos de biodiversidad. En este contexto, estos modelos son definidos como cualquier modelo matemático (con elementos como variables de estado, forzantes, ecuaciones matemáticas y parámetros) que cubre un amplio espectro de niveles de organización ecológica (desde individuos hasta ecosistemas). Luego, nosotros implementamos un análisis de incertidumbre en el modelo NHCE OSMOSE, con el objetivo de cuantificar el efecto de la incertidumbre de sus parámetros en un conjunto de indicadores ecológicos derivados de este modelo. Para esto, nosotros trabajamos con dos escenarios de incertidumbre usando los parámetros del modelo y correspondientes a un coeficiente de variación del 10% y 20% en torno a los valores de referencia de los parámetros que describen cada especie modelada. Los resultados de este estudio revelaron que la incer-

tidumbre en los parámetros de algunas especies tienen fuertes repercusiones en el resto de la red trófica; este fue el caso de los eufáusidos y la pota. Además, este trabajo nos permitió identificar los indicadores ecológicos que fueron más robustos a la incertidumbre de los parámetros de entrada; como por ejemplo, el índice trófico marino y el nivel trófico de las comunidades de peces.

Los resultados del análisis de incertidumbre abrió un conjunto de preguntas que abordamos en el capítulo tres. Aquí, nosotros proponemos un protocolo dedicado a la realización de un análisis de sensibilidad de un modelo complejo. Además, en este capítulo mostramos los inconvenientes de utilizar rangos de variabilidad arbitrarios para representar la incertidumbre en las entradas de un modelo. Este es un problema frecuente en el uso de modelos que presentan gran cantidad de parámetros, los cuales no siempre tendrán información relacionada a su incertidumbre. Para hacer frente a esto, nosotros proponemos el "criterio de fiabilidad de los parámetros". Este criterio sirve para clasificar los parámetros del modelo según la fuente de información utilizada para estimar sus valores, y para asignar un nivel de incertidumbre a los parámetros de un modelo. Nosotros ilustramos el uso de este protocolo implementando un análisis de sensibilidad usando el modelo NHCE OSMOSE.

Finalmente, nosotros concluimos que las complejidades inherentes a los modelos complejos no deben utilizarse como excusa para evitar estudiar la incertidumbre de forma apropiada. Existen métodos y herramientas computacionales en la literatura. Sin embargo, deben seleccionarse cuidadosamente teniendo en cuenta las posibles limitaciones que puede generar el uso de un modelo. El nivel de incertidumbre de un modelo complejo depende en gran medida de la incertidumbre en las entradas del modelo, de su propagación en el tiempo, pero también de las salidas estudiadas. En esta investigación, destacamos las dificultades de estudiar la incertidumbre en modelos complejos (incluyendo, pero no limitándose, a los modelos de ecosistemas) y proponemos posibles soluciones y directrices para futuros estudios que aborden este problema de investigación. Nuestros resultados son sólo un primer paso, empezando por el estudio de la incertidumbre en los parámetros, pero dejando como perspectiva el estudio de otras fuentes de incertidumbre como en los forzantes del modelo, las condiciones iniciales y la estructura del modelo.

Palabras clave: modelos complejos, análisis de incertidumbre, análisis de sensibilidad, modelización de ecosistemas, OSMOSE, ecosistema norte de la corriente de Humboldt.

Acknowledgements

This PhD was at the beginning only a professional challenge, but along the way it became a personal one that changed my life in a number of ways.

I want to start this section by thanking my two main supervisors (**Paul Leadley** and **Yunne Shin**) who did a huge job giving me all their support and help through the development of this PhD thesis.

Yunne meant to me the opportunity to witness the big picture of working as researcher. In this context, I harness the space to express my admiration for her research work, her investment in all the things she engage, and for being a very generous human being. All this is very motivating for young researchers around her. On the other hand, I thank Yunne for all the freedom she gave me in this thesis, letting me explore my research problem which lead to develop more skills (from theoretical skills until technical ones). Also, I acknowledgement the criticisms and rigor Yunne putted on me. All this pushed me to be better day by day.

I thank Paul for accepting me as a student, for having trusted in my work and for had being there when I needed him. Paul, I have admired your positivism and kindness every time we have a discussion. Also, thanks for all the administrative work you had to deal with. I imagine that having a student with a Peruvian scholarship (and outside Paris) had not been easy at all, and thank you for all the extra work that this may have generated. Finally, I acknowledge all the meaningful suggestions and the editions you both provided for the generation of this manuscript.

Un agradecimiento especial a **Ricardo Oliveros** por su ayuda durante este trabajo. Ricardo, gracias por todos los comentarios y consejos que han hecho que yo me haga más sólida como investigadora. Además, tengo el gusto de haber trabajado contigo en el transcurso de mi formación profesional y te agradezco por haber estado en los momentos más motivadores y más tristes de mi vida profesional ¡Muchas gracias Ricardo, promesa cumplida y tesis terminada!

I am particularly grateful with my jury members: **Philippe Cury**, **Verena Trenkel**, **Eva Plagányi**, **Dimitri Gutiérrez** and **Dorleta García**, for having accepted to carry out the evaluation of my thesis manuscript and oral defense. In addition, I ap-

preciated all your interesting comments and suggestions that had enriched this research work.

I sincerely thank all researches that participate during my thesis committees: **Mark Rounsevell**, **William Cheung**, **Didier Gascuel**, **Morgane Travers** and **Sigrid Lehuta**, for the constructive and interesting discussions we had. But also, for the support and motivation they provided me along the development of this work.

I thank all the administrative staff of Paris-Saclay and the doctoral school SEVE (ex-SdV) for all the facilities they provided me during this work. Specially to **Jacqui Shykoff** for all her kindness from the day one that I started this thesis (and even before, when we started the preparation of the thesis project). Jacqui, all your advice during the autumn school at Paimpont I will never forget, it was a nice experience. I also thank **Marianne Delarue**, with who we entered in contact at the end of the project. Marianne, thanks for your support and all the administrative help you provided to me.

I thank **FONDECYT** (Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica) for the Peruvian individual doctoral research grant (contract n° 134-2016-FONDECYT). This funding made possible my stay in France throughout the thesis development but also allowed me to organize diverse missions and attend several conferences around the world. To my administrative advisors from FONDECYT, Carmen Castañeda and Aldo de la Torre, thank you for all your support during the execution of this project. Thanks to you I learned the administrative requirements of managing a research project.

I had to deal with the use of complex models within the development of my thesis. In this regard, I spent most of the time using the OSMOSE model and taking many reflections around this tool. I am deeply grateful to all developers and users of OSMOSE who were available for any question and discussion I had, starting with the team leader (**Yunne Shin**) but also with **Nicolas Barrier** for being of great support to the OSMOSE team. Nico, thanks for had been helping me with every question I had regarding OSMOSE, Datarmor or whatever, your help was huge! I also thank **Ricardo Oliveros** for all the discussions about the NHCE OSMOSE model. In general, thanks to all the OSMOSE community with whom I had the pleasure to discuss: **Fabien Moullec**, **Laure Velez**, **Morgane Travers**, **Alaia Morell** and **Ghassen Halouani**.

Toute la période de ma thèse s'est déroulée dans la ville de Montpellier. Elle a été réalisée au sein de l'Unité Mixte de Recherche MARBEC (MARine Biodiversity Exploitation and Conservation), sur le site de l'Université de Montpellier. Je remercie toute la direction de **Marbec**, ainsi que le personnel administratif et de

recherche qui m'ont permis de rester dans ce laboratoire. Je remercie toutes les personnes de ce laboratoire avec qui j'ai pu discuter et échanger, mais surtout à mes collègues du bureau 64 (au premier étage du bâtiment 24 de l'Université de Montpellier). Ils ont été mes professeurs de français, mais aussi de culture, d'administration et de gastronomie française. Si je peux écrire ce paragraphe dans cette langue, c'est grâce à beaucoup d'entre vous: **Laure Velez**, **Fabien Moullec**, **Nicolas Loiseau**, **Romain Pete**, **Gaël Mariani**, et **Béatrice Bec**; mes meilleurs colocs et meilleurs compagnons d'aventures à Marbec. Ainsi que ceux qui se sont joints à nous au cours de ce projet: **Juliette Langlois**, **Camille Magneville**, **Ignacio Pita**, **Valentine Fleure**, **Alaia Morell**, **Adrien Chevallier**, **Fabrice Houngnandan**, etc. Merci pour toutes les expériences vécues avec vous, et pour votre soutien moral à tout instant, mais surtout dans la dernière partie de la thèse, la plus difficile. J'ai passé de bons moments avec vous, des débats intéressants (scientifiques, mais ouverts à tous les sujets), et de grands moments sportifs (à Runup, la bonne époque). J'espère que vous n'oublierez pas cette Péruvienne. Je me souviendrai toujours de vous avec un grand sourire.

Mi agradecimiento profundo y sincero a quienes fueron mis colegas de IMARPE (Instituto del Mar del Perú). Instituto de investigación donde inició mi carrera científica. Aquí realicé mis prácticas pre-profesionales, realicé mis estudios de maestría, encontré mi primer trabajo, y empezó la aventura de realizar un doctorado. Agradezco a cada uno de los profesionales que pude encontrar en este centro de investigación, porque de cada uno de ustedes aprendí algo que fue necesario durante esta tesis. A los colegas del ex-CIMOBP (actual LMOECC). Pero de manera especial, a mis colegas y amigos, de la Dirección General de Investigaciones en Recursos Pelágicos. De manera muy especial, a **Miguel Ñiquen** por haberme alentado constantemente en cada decisión que tome en mi vida profesional y sobre todo por haber sido para mí un ejemplo de profesionalismo y perseverancia. También agradezco a **Erich Díaz**, **Rocío Joo**, **Ricardo Oliveros**, **Marilú Bouchon**, **Teobaldo Dioses**, **Ana Alegre**, **Cecilia Peña**, entre muchos otros. De cada uno de ustedes aprendí cosas invaluables y fue un placer compartir oficina y espacios de trabajo juntos. A los que pude considerar mis colegas-amigos, y nos formamos casi en el mismo "cardumen": **Wencheng**, **Josymar**, **Enrique**, **Pablo** y **Giancarlo**, gracias por todo lo compartido.

A special thanks to all the people who made my life in Montpellier (especially outside my country) more enjoyable, beautiful and charismatic. To my first "french family" in Montpellier, **Stephanie Egloff**, thank you for your friendship and kindness without limits. To **Elvira Hernandez** (mi hermana, amiga and best partner

ever), thank you for all the metaphorical discussions about life and science. **Alejandro Ariza** and **Valentina Leoni**, you are also part of the family, los quiero mucho chicos, esta aventura fue genial con ustedes cerca! To the nice girls, **Yolanda Boyer**, **Raquel Marques** and **Marie Charlotte Cheutin**, for those shared laughs and beautiful moments of good energy (in Marbec but also outside the laboratory). To all **R-Ladies Montpellier**, in particular **Cécile Sauder**, **Lexie Sauder** and **Jean Delmotte**, for the R-discussions and good anecdotes shared in this chapter. Those who joined at the end of the adventure, **Emma Patrone** and **Olivier Martin** for the trekkings, picnics, and soirées jeux de société; but also to **Yacine Seghir** and **Clémence Coutier** for showing me the beautiful things that France has, merci beaucoup pour tout les gars!

During my doctoral training, I had many challenges to face, planned challenges and others that arose unexpectedly. However, in every moment in which I felt that “I couldn’t continue", I always had a word of motivation from a friend. Those friends that we cross through life, and no matter where you are, they are there to let us know that everything will be fine. Thanks to **Lucero**, **Carolina**, **Blanca Rosa**, **Maite**, **Kristen**, **Claudia J.**, **Fiorella M.**, **Claudia T.**, **Jair V.** (y UNALMos en general), **Mayra Raquel**, **Rocio Joo**, **Danny Engel**. Los abrazo chicos!

Mon séjour en France a changé ma vie à bien des égards, intégrant de nouveaux collègues, amis, et autres appelés “famille". Dans cette rubrique, je remercie tout particulièrement **Daniel Moreno**, pour avoir fait partie de ma famille française. Dany, merci pour ta patience (parfois infinie), pour avoir enduré mes moments critiques de stress, pour m’avoir soutenu sans relâche jusqu’à la fin de cette thèse. Dany, si cela se termine c’est en grande partie grâce à toi. Merci d’avoir fait partie de cette belle aventure. De même, je remercie toute la **famille Moreno**, et d’une manière très spéciale à **Christine M.**, pour tous les moments agréables partagés, et pour m’avoir fait sentir que je n’étais pas seule en France, et que je pouvais compter toujours sur vous.

Gracias de manera muy especial a mis padres (**Doris** y **Santiago**) por toda su ayuda, a mis hermanos (**Marlon** y **Rodolfo**) por ser mi apoyo incansable, pero también a **Diana** y **Santiaguin** por alegrarme con sus noticias. A mi madrina **Bety** (y familia), a mi familia del ”**Rímac**" (en especial a mi tía **Ceci** y **Mamalena**), y todos aquellos que me apoyaron en mi formación como persona y profesional. Sepan que todo lo que puedo ser ahora es también gracias a ustedes. Se los agradezco de corazón.

Dedicado al biólogo Miguel Ñiquen, por todas sus enseñanzas y por todo su apoyo durante mi formación como científica.

Contents

Introduction

The world's population lives today in the Decade of the Oceans, the period 2021-2030 declared by the UN as the "Decade of Ocean Sciences for Sustainable Development" (Decade of the Oceans). This decade whose vision is to "develop scientific knowledge, build infrastructure and foster relationships for a sustainable and healthy ocean", aims to achieve societal outcomes, as for example: (1) a predictable ocean that enables society to understand current and future ocean conditions; and (2) a transparent ocean, giving citizens equitable access to data, information, and technologies. In parallel, the UN has also declared the same period (2021-2030) as the "Decade for Ecosystem Restoration". Thus, these two agreements give the scientific community an unique opportunity and an imperative to work towards a sustainable future for the ocean (Heymans et al., 2020). With the need for quantitative approaches to better understand and predict ecosystem structure and functioning under global change, the scientific community faces great challenges where ecosystem models could play a significant role (Heymans et al., 2020; Ryabinin et al., 2019).

Ecosystem models attempt to incorporate multiple ecosystem components (e.g., species, populations, functional groups) and processes (e.g., predation-prey interactions, causes of natural mortalities, species spatial distribution, growth, reproduction, metabolic responses to climate change etc) into a unique modelling framework (Geary et al., 2020). These models can integrate various disciplines (e.g., physical oceanography, biochemistry, biology, social and economic sciences) and have been extensively used in the marine field (Steenbeek et al., 2021). There is currently an extensive set of model types for marine ecosystems, from the qualitative (and conceptual) ones to coupled and end-to-end models (Fulton 2010; Tittensor et al., 2018). In general, all of the ecosystem models share common objectives: to contribute to improving our understanding of natural phenomena, but also of the impacts of human activities, or climate change on marine food webs (Steenbeek et al., 2021; Heymans et al., 2020; Travers 2007). Nevertheless,

there are many limitations that ecosystem models face, one of which is addressed through this thesis, namely the impact of uncertainty in model predictions.

Models are built under uncertainties, due to e.g., their parameters, the equations describing the processes and assumptions, their initial conditions (Cariboni et al., 2007). In the literature, several review works explain the diverse sources of uncertainty in ecological models; for example, Regan et al. (2002) proposed a classification of various sources of uncertainty accompanied by its most common treatments in ecology and conservation biology. Also, Link et al. (2012) presented the major sources of uncertainty encountered by ecosystem modellers and that affect the entire process of ecosystem-based fisheries management: natural variability, observation error, inadequate communication, outcome uncertainty, unclear or nonspecific management objectives, and structural uncertainty. Besides, more recent studies have focused on five main sources of uncertainty, which are described as: the structural (model) uncertainty, the initialization, the internal variability uncertainty, the parametric uncertainty, and the scenario uncertainty (Payne et al., 2016; Cheung et al., 2016; Lehuta 2016; IPBES 2016). The impact of these sources of uncertainty on ecosystem models' predictions needs to be more systematically studied, but also adequately communicated from modellers to decision-makers, supporting decision-making based on model outcomes in a more transparent and informative way (Rounsevell et al., 2021; Steenbeek et al., 2021).

A variety of marine ecosystem models have been developed worldwide especially in the last two decades (Plaganyi et al., 2007; Tittensor et al. 2018). In the northern Humboldt Current Ecosystem (NHCE), mainly two different ecosystem models have been applied, i.e. Ecopath with Ecosim (EwE) (e.g., in Tam et al., 2008; Taylor et al., 2008) and OSMOSE (e.g., in Marzloff et al., 2009; Oliveros-Ramos et al., 2017). The Humboldt Current ecosystem is a case study of particular interest as it is one of the largest and most productive marine ecosystems globally, extending along the west coast of Peru and Chile (Serra et al., 2012). In addition, it is one of the four major Eastern boundary upwelling systems in the ocean (with Canary, Benguela, and California), characterised by dominant equatorial coastal wind stress, offshore Ekman transport, and the presence of coastal upwelling of cold, nutrient-rich subsurface waters that fuel high levels of plankton and fish production and the highest fishery production in the world (Thiel et al., 2007).

In particular, the NHCE produces more fish per unit area than any other region in the world, accounting for up to 10% of the global fish catches, and has supported a fish production 20 times bigger than in Canary or Benguela (Chavez et al., 2008; Bakun and Broad 2003). The NHCE exhibits high climatic and oceanographic variability at several scales (e.g., seasonal, interannual, and decadal) (Gutierrez et al., 2016). This system does not only vary in terms of environmental conditions but also in fishing activity, which depends on the abundance and accessibility of the primary fishery resources. The NHCE hosts main fishery resources like, for example, the Peruvian anchovy (*Engraulis ringens*), sardine (*Sardinops sagax*), jack mackerel (*Trachurus murphyi*), pacific mackerel (*Scomber japonicus*) (Ñiquen and Bouchon 2004, Oliveros et al., 2021). Given the intertwined dynamics of the environment and marine life, the representation and improved understanding of the behaviour of the NHCE, it is necessary to use holistic frameworks such as ecosystem models, in particular end-to-end models that allow to couple the dynamics of the environment to that of the living components of the ecosystem.

In this thesis, we addressed the uncertainty related to the use of the NHCE OSMOSE model (Oliveros-Ramos et al., 2017). This is a complex model resulting from the coupling of three disciplinary models: ROMS (regional ocean physical model), PISCES (biogeochemical model), and OSMOSE (fish community model). The challenge of this thesis work was due to the combination of the strong technical issues related to the uncertainty quantification besides the use of a complex ecosystem model. The NHCE OSMOSE model is characterized by a large number of inputs (e.g., parameters, forcing, initial conditions) and outputs (multiple simulated variables) in addition to the high computational requirements (simulation time, memory space required). This work contributes to developing methodologies to study uncertainty in complex models considering the limitations raised. It addresses two main research questions: the first focused on understanding how to deal with uncertainty in complex models, and the second on the impact of uncertainty on the diverse output variables produced by a complex ecosystem model. The main outcomes of this work are presented in three chapters, and some general conclusions and perspectives are finally drawn to pave the way for future work.

In chapter 1, we review methods dedicated to studying uncertainty with potential applications to biodiversity models. In the context of this chapter, we define a biodiversity model as any mathematical model (with elements such as state

variables, forcing, mathematical equations, and parameters) that covers a broad spectrum of levels of ecological organization (from individuals to ecosystems). Here we present a set of methods dedicated to applying uncertainty and sensitivity analyses and a list of selected computational tools developed in various programming languages that allow the implementation of both types of analyses. In addition, this chapter includes a review of the different sources of uncertainty in scenarios and models of socio-ecological systems in support of decision-making.

In chapter 2, we implemented an uncertainty analysis with the main objective of quantifying the effect of parameter uncertainty on a set of ecological indicators derived from a complex model. We worked on the OSMOSE modelling platform and its application to the northern Humboldt Current ecosystem (NHCE OSMOSE). In this study, we worked with two uncertainty scenarios corresponding to ranges of 10% and 20% variability around the reference values of the parameters describing each modelled species. Besides, we analyzed the responses of a set of ecological indicators to different levels of uncertainty. In this work, we hypothesized that: 1) the uncertainty coming from one species' parameters could propagate through the food web, 2) the response of ecological indicators to uncertainty is not homogeneous, and that 3) depending on the indicators chosen among the outputs of the model, these could amplify or dampen the level of uncertainty in the model inputs.

In chapter 3, we propose a protocol dedicated to implementing a sensitivity analysis of a complex model. This chapter shows the drawbacks of using arbitrary variability ranges to represent the uncertainty of model inputs. We also propose the "parameter reliability criterion". It serves to classify the model parameters according to the source of information used to estimate their values, and to estimate the ranges of variability for each parameter to be used in a sensitivity analysis. As the main results of this chapter, we developed a protocol for implementing a sensitivity analysis using complex models. By applying this protocol to the NHCE OSMOSE model, we showed that the use of arbitrary ranges could lead to biased sensitivity analyses.

In this thesis, we have faced many computational challenges. The development of a tool that allows us to automatize the process of running large simulation batches and analyzing the results obtained has been fundamental. As a complementary work to this thesis, we reviewed how R packages dedicated to computer

simulations have been developed in CRAN (The Comprehensive R Archive Network) and developed a series of guidelines on how these tools can be improved to facilitate their adoption by future users. We also took the lessons learned in this study to enhance the OSMOSE package (implemented in R), which is currently used for simulation and analysis of the model outputs. As a result, an additional product of this thesis was the creation of a package that allowed us to run uncertainty and sensitivity analyses automatically, taking into account the peculiarities of complex models. While the NHCE OSMOSE model illustrates the functionality of this package, this tool has been built flexibly so it can be used with any model and it is available on GitHub as an open-source code.

Finally, we conclude with some insights brought out by our work on how we could deal with uncertainty using complex ecosystem models, with the many technical limitations associated to its study and quantification. However, these difficulties should not be used as an excuse for not tackling uncertainty, and this thesis is a small contribution to illustrate that rigorous analyses of the behaviour of models, even complex ones, can be undertaken.

References

Bakun, A. and K. Broad. (2003). Environmental loopholes and fish population dynamics: comparative pattern recognition with focus on El Niño effects in the Pacific. Fisheries Oceanography. 12(4/5):458-473.

Cariboni, J., Gatelli, D., Liska, R., and Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. Ecological modelling, 203(1-2), 167-182.

Chavez F. P., A. Bertrand, R. Guevara-Carrasco, P. Soler and J. Csirke. (2008). The northern Humboldt Current System: Brief history, present status and a view towards the future. Progress in Oceanography 79:95-105.

Cheung, W. W., Frölicher, T. L., Asch, R. G., Jones, M. C., Pinsky, M. L., Reygondeau, G., et al. (2016). Building confidence in projections of the responses of living marine resources to climate change. ICES Journal of Marine Science, 73(5), 1283-1296.

Fulton, E. A. (2010). Approaches to end-to-end ecosystem models. Journal of Marine Systems, 81(1-2), 171-183.

Geary, W. L., Bode, M., Doherty, T. S., Fulton, E. A., Nimmo, D. G., Tulloch, A. I., et al. (2020). A guide to ecosystem models and their environmental applications.

Nature Ecology Evolution, 4(11), 1459-1471.

Gutiérrez, D., Akester, M., and Naranjo, L. (2016). Productivity and sustainable management of the Humboldt Current large marine ecosystem under climate change. Environmental Development, 17, 126-144.

Heymans, J. J., Bundy, A., Christensen, V., Coll, M., De Mutsert, K., Fulton, E. A., et al. (2020). The ocean decade: a true ecosystem modeling challenge. Frontiers in Marine Science, 7, 766.

IPBES (2016). The methodological assessment report on scenarios and models of biodiversity and ecosystem services (IPBES).

Lehuta, S., Girardin, R., Mahévas, S., Travers-Trolet, M., and Vermard, Y. (2016). Reconciling complex system models and fisheries advice: practical examples and leads. Aquatic Living Resources, 29(2), 208.

Link, J. S., Ihde, T. F., Harvey, C. J., Gaichas, S. K., Field, J. C., Brodziak, J. K. T., et al. (2012). Dealing with uncertainty in ecosystem models: the paradox of use for living marine resource management. Progress in Oceanography, 102, 102-114.

Marzloff, M., Shin, Y. J., Tam, J., Travers, M., and Bertrand, A. (2009). Trophic structure of the Peruvian marine ecosystem in 2000–2006: insights on the effects of management scenarios for the hake fishery using the IBM trophic model Osmose. Journal of Marine Systems, 75(1-2), 290-304.

Ñiquen, M., and Bouchon, M. (2004). Impact of El Niño events on pelagic fisheries in Peruvian waters. Deep sea research part II: topical studies in oceanography, 51(6-9), 563-574.

Oliveros-Ramos, R., Verley, P., Echevin, V., and Shin, Y. J. (2017). A sequential approach to calibrate ecosystem models with multiple time series data. Progress in Oceanography, 151, 227-244.

Oliveros-Ramos, R., Ñiquen, M., Csirke, J., and Guevara-Carrasco, R. (2021). Management of the Peruvian anchoveta (Engraulis ringens) fishery in the context of climate change. Adaptive management of fisheries in response to climate change, 237.

Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., et al. (2016). Uncertainties in projecting climate-change impacts in marine ecosystems. ICES Journal of Marine Science, 73(5), 1272-1282.

Plagányi, É. E. (2007). Models for an ecosystem approach to fisheries.

Regan, H. M., Colyvan, M., and Burgman, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological applications, 12(2), 618-628.

Rounsevell, M. D., Arneth, A., Brown, C., Cheung, W. W., Gimenez, O., Hol-

man, I., et al. (2021). Identifying uncertainties in scenarios and models of socio-ecological systems in support of decision-making. One Earth, 4(7), 967-985.

Ryabinin, V., Barbière, J., Haugan, P., Kullenberg, G., Smith, N., McLean, C., et al. (2019). The UN decade of ocean science for sustainable development. Frontiers in Marine Science, 6, 470.

Serra, R., Akester, M., Bouchón, M., Gutierrez, M. Sustainability of the Humboldt Current Large Marine Ecosystem. Frontline observations on climate change and sustainability of large marine ecosystems United Nations Development Programme, New York. 2012;p. 112–134.

Steenbeek, J., Buszowski, J., Chagaris, D., Christensen, V., Coll, M., Fulton, E. A., et al. (2021). Making spatial-temporal marine ecosystem modelling better–A perspective. Environmental Modelling & Software, 145, 105209.

Tam, J., Taylor, M.H., Blaskovic, V., Espinoza, P., Ballón, R.M., Díaz, E., Wosnitza-Mendo, C., Argüelles, J., Purca, S., Ayón, P., Quipuzcoa, L., Gutiérrez, D., Goya, E., Ochoa, N., Wolff, M. (2008). Trophic modeling of the Northern Humboldt Ecosystem, Part I: comparing trophic linkages under La Niña and El Niño conditions. Progress in Oceanography 79, 352–365.

Taylor, M.H., Tam, J., Blaskovic, V., Espinoza, P., Ballón, R.M., Wosnitza-Mendo, C., Argüelles, J., Díaz, E., Purca, S., Ochoa, N., Ayón, P., Goya, E., Gutiérrez, D., Quipuzcoa, L., Wolff, M. (2008). Trophic modeling of the Northern Humboldt Current Ecosystem, Part II: elucidating ecosystem dynamics from 1995–2004 with a focus on the impact of ENSO. Progress in Oceanography 79, 366–378.

Thiel, M., Macaya, EC., Acuna, E., Arntz, WE., Bastias, H., Brokordt, K., et al. (2007). The Humboldt Current System of northern and central Chile: oceanographic processes, ecological interactions and socioeconomic feedback. Oceanography and Marine Biology, 45:195–344.

Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., ... Walker, N. D. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1. 0. Geoscientific Model Development, 11(4), 1421-1442.

Travers, M., Shin, Y. J., Jennings, S., and Cury, P. (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. Progress in oceanography, 75(4), 751-770.

Chapter 1

Addressing uncertainty in biodiversity models: from methods to tools

Biodiversity models are key tools to explore the current and future impacts of global change drivers on biodiversity. To be useful in decision and policy-making, the uncertainty linked to models' projections needs to be addressed and quantified. Here we review how uncertainty of these biodiversity impact models can be handled according to three uncertainty types: parameter, initialization and internal variability uncertainty. We present a panel of methods for uncertainty analysis and sensitivity analysis as well as online tools available in the literature. Most methods presented here were found in the statistical and mathematical literature with applications across different fields. Seeding new ideas and approaches from various disciplines is a good step forward to advance uncertainty quantification of biodiversity models and their predictions.

Keywords: biodiversity models, uncertainty analysis, sensitivity analysis, uncertainty quantification.

Addressing uncertainty in biodiversity models: from methods to tools

Criscely Luján [1,2], Paul Leadley [1], Mark D.A. Rounsevell [3,4], Sigrid Lehuta [5], Isabelle Maréchaux [6], Peter H. Verburg [7,8], William W.L. Cheung [9], Almut Arneth [3], Romain Pete [2], Yunne-Jai Shin [2]

[1] *Laboratoire d'Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91400, Orsay. France*
[2] *MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Montpellier. France*
[3] *Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research/Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, and the Department of Geo-ecology (IFGG), Karlsruhe, Germany*
[4] *School of GeoSciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK*
[5] *IFREMER Ecologie et modèles pour l'halieutique, Rue de l'île d'Yeu, BP 2011, 44311 Nantes Cedex 03, France*
[6] *AMAP Lab, University of Montpellier, CIRAD, CNRS, INRAE, IRD, 34000 Montpellier, France*
[7] *Institute for Environmental Studies, VU University Amsterdam, de Boelelaan 1087, 1081 HV Amsterdam, the Netherlands*
[8] *Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland*
[9] *Changing Ocean Research Unit, Institute for the Oceans and Fisheries, the University of British Columbia, Vancouver, BC V6T 1Z4, Canada*

*Corresponding author: criscelylujan@gmail.com

1. Introduction

Quantitative models have been important tools in studying biodiversity (Mouquet et al., 2015). Models simplify complex ecological systems to help understand their functioning and predict their behavior or responses underpinned by a set of explicit assumptions (Smith and Smith 2012). Quantitative biodiversity models consist of a set of elements, namely state variables, external variables, mathematical equations and parameters (Jørgensen and Fath 2011). Most existing quantitative biodiversity models are composed of three main components that should be identified: the biophysical components or characteristics of an ecosystem represented by state variables, the drivers impacting the state variables, and the ecological processes determining changes in the biological components (Ferrier et al., 2016).

Because of the direct relevance to managing the effects of human activities on natural systems, modelling biodiversity responses to global changes is particularly common in supporting conservation policies and decision making. Biodiversity dynamics are determined by both external drivers, mostly human-induced, and internal biological processes acting within the ecosystem (Lavergne et al., 2010; Leung et al., 2012). These drivers and processes have been integrated into scenarios and models to elucidate the new challenges that biodiversity is exposed to (Pereira et al., 2010). Diverse quantitative biodiversity modelling approaches have been developed for terrestrial (Alkemade et al., 2009; Sala et al., 2000), marine (Tittensor et al., 2018), and freshwater ecosystems (Holland, Darwall, and Smith 2012; Hammen and Settele 2011); and also across the three realms, evaluating for example the impacts of climate change on the future of global biodiversity (Bellard et al., 2012). However, the utility of these biodiversity models in informing biodiversity status mostly depends on the predictability of these models which depends on the capacity to quantify different sources of uncertainty.

Diverse typologies of uncertainty can be found in the scientific literature with different sources and forms of treatment (Charles 1998; Regan, Colyvan, and Burgman 2002; Link et al., 2012; Skinner et al., 2014). Despite this, uncertainty *perse* in biodiversity impact models has not been examined and compared systematically (Ludwig, Mangel, and Haddad 2001). Partitioning uncertainty into its sources is essential to understand the various contributions to uncertainty in the use of models (Cheung et al., 2016). Without such a partitioning, it is hard to target efforts to reduce this uncertainty or interpret the uncertainties of models' outcomes.

The present work focuses on the methods aiming at addressing the challenges related to uncertainty quantification with a three-tier approach guided by the following questions: i) how to characterize uncertainty?, ii) how to quantify (or measure) uncertainty?, and finally iii) how to reduce uncertainty? We first provide definitions of each uncertainty source. We then provide a range of methods for quantifying uncertainty. We also pinpoint tools developed for dealing with uncertainty that could be useful when using biodiversity models. We finally present gaps, future directions and guidelines for a step improvement in uncertainty quantification of biodiversity models and their predictions.

2. Sources of uncertainty in biodiversity models

We identified three main sources of uncertainty in using biodiversity models: parameter, internal variability and initialization uncertainty (see Table 1). Other sources of uncertainty, such as scenario and decision uncertainty, have been reviewed recently (Rounsevell et al., 2021; Ferrier et al., 2016) and will not be addressed here.

Table 1. *Sources of uncertainty in biodiversity models. The description of the uncertainty and its source are based on Payne et al., (2015) and Cheung et al., (2016).*

Uncertainty type	Description	Source
Parameter uncertainty	It is the uncertainty associated with the parameter values used in a model, and that influences its behavior.	Many factors can give rise to parameter uncertainty, including imperfect measurements, inadequate coverage of the range of natural variability, or natural variability in biological parameters.
Initialization uncertainty	It is the uncertainty associated with the initial conditions of a model, meaning the state of the system from which the model then integrates forward in time.	This uncertainty arises from the inability to fully and accurately observe or characterize the state of a complex system at a given point in time.
Internal variability uncertainty	It is the uncertainty associated with the natural variability of the modeled systems. It occurs in systems that change in ways that are difficult to predict.	This uncertainty arises in both temporal and spatial dimensions and even in the absence of any model forcing. This uncertainty is due to stochasticity, feedbacks, non-linearities, or periodicities.

3. Quantifying the uncertainty

3.1 Uncertainty analysis and sensitivity analysis

Both uncertainty analysis (UA) and sensitivity analysis (SA) are widely used to quantify model uncertainty across many fields (e.g. physics, environment, economics) (Saltelli, Chan, and Scott 1999). SA determines how the uncertainty in model outputs can be apportioned to different sources of uncertainty in model inputs (Saltelli et al., 2019), in other words looking for how model output responds to relative changes in the inputs. While UA characterizes the uncertainty in model outputs, without necessarily identifying which assumptions, incomplete knowledge or misspecification of the modeled system are primarily responsible for this uncertainty (Cariboni et al., 2007). Although most modellers distinguish SA from UA, sometimes practitioners use both terms interchangeably. Nevertheless, for the purposes of this review we remain with the definitions provided above.

For example, when applied to biodiversity models (See **Figure 1**), the UA should aim to assess the uncertainty propagation in the model outputs from the sources (e.g., model parameters, initial values and the internal variability of the system). Different methods were proposed to estimate this uncertainty with the aim to generate a distribution of possible model results (the gray area in **Figure 1**). The characterization of such distribution, including summary statistics, is part of the UA. On the other hand, the decomposition of the uncertainty among the uncertainty sources (represented by the pie chart) is part of the SA.

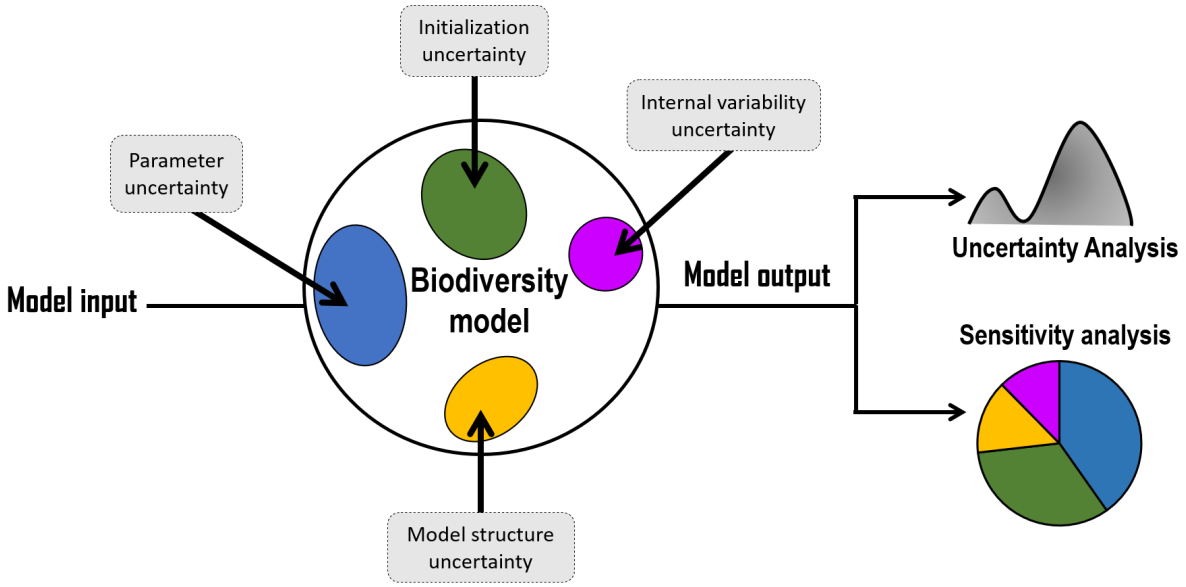

Figure 1. Schematic representation of the differences between UA and SA*. Uncertainty coming from different sources (parameter, structure, initialization and internal variability uncertainty) propagates through the model, then the UA aims to quantify all these uncertainty sources (black curve). The uncertainty in the model output captured e.g., by its variance (or other statistical metrics), is then decomposed into the different sources thanks to a SA (Adapted from Saltelli et al., 2019).*

3.2 Methods for uncertainty analyses

Summarizing and displaying the uncertainty associated with model outputs address many questions, such as: what are the mean and median of *Y*? what is the variance of *Y*? are there any discontinuities associated with the distribution function of *Y*? (Iman and Helton 1988). Methods to estimate the uncertainty in models under different assumptions and procedures were proposed in the scientific literature. The following methods are based on statistical techniques which could be applied directly to computer simulations and to biodiversity models. In **Table 2** we summarize the principal advantages and limits of each of these methods.

Table 2. *Methods for UA. This table summarizes the main advantages and limits of common methods used for UA.*

Method	Advantages	Limits
Statistical inference: This approach is based on the use of data to deduce some properties of an underlying probability distribution of model's parameters. In the quantification of uncertainty two inference paradigms are commonly used: bayesian and frequentist.		
a. Bayesian methods	Allows the integration of prior information about the system parameters	The use of prior information that is not well estimated can lead to subjectivity issues
b. Frequentist methods	Do not require prior distribution for all parameters in the model	Do not offer a structured framework to incorporate prior information and multiple model structures
Sampling-based approaches: These form a broad class of algorithms which involve the generation and exploration of a probabilistic mapping from inputs to outputs. Here we only focus on the Monte Carlo analysis and the latin hypercube sampling. They are indeed commonly used to quantify random uncertainty simulated distribution for outputs as a function of input distribution.		
c. Monte Carlo analysis	Formal probabilistic approach which does not require intermediate surrogates or intermediate models to obtain uncertainty measures	Multiple model evaluations are required producing computational burden
d. Latin hypercube sampling	Requires a reduced number of model evaluations demanding less computational time than Monte Carlo analysis	Expensive optimization and problems of dimensionality. Space filling becomes difficult with large numbers of model inputs
Modelling approach: Through this approach a model will be approximated by a new surrogate one, and this will be used in the subsequent uncertainty studies. Here we focus on the response surface and differential analysis approaches. The first approach uses classical experimental designs to select points to be used in developing a response surface in replacement of a model. While differential analysis involves approximating a model with Taylor series.		
e. Response surface method	Reduces the complexity of the model under study replacing it by surrogate models. UA becomes inexpensive once the response surface approximation is done	Requires the construction of surrogate models restricting the analysis to few input variables and limited number of outputs where the input-output relationships are simple ones (e.g linear, quadratic)
f. Differential analysis	Reduces the complexity of the model under study replacing it by surrogate models. UA becomes	Poor technique in the estimation of distribution functions of model outputs.

	inexpensive once the series are implemented	

UA based on statistical inference: Bayesian methods provide a formal framework for calculating the probability of model parameter values given: i) data sets, ii) one or more structural models, and iii) information about models or parameters from external sources which are specified as prior distributions (Patterson et al., 2001). Under Bayesian methods observations are considered as known quantities that might have been generated by a variety of processes, then distributions are assigned to the parameters of these processes to enable inferences to be drawn about them (Harwood and Stokes 2003). Bayesian methods conventionally involve reporting results in terms of the expected value of a posterior distribution for each parameter. This estimated probability distribution reflects the major uncertainties about the system under study. The advantages of Bayesian methods is the integration of prior information which is the prior uncertainty based on experiences in similar simulations (e.g. measurement of errors, distribution of measurements). Although a strength of Bayesian methods, this also introduces complexity due to the requirement of the prior distribution for all the model parameters. Specifying a prior distribution leads to subjectivity issues. Not all the priors can lead to credible intervals, which results in a complex problem even for simple models (Lele 2020). A more extended discussion about the benefits of Bayesian methods to quantify uncertainty is presented by van Oijen (2017).

In Bayesian methods, parameters are considered to be random variables while in the frequentist methods parameters are treated as unknowns that possess "true" values. By using frequentist methods, data are used to estimate the values of model parameters (including errors of measurements) based on assumptions about the statistical distributions from which the data were drawn. These methods aim to calculate a probability model describing the sampling distribution of estimators derived from the data, given particular values which are assumed to be true. Then this model can be used to construct a confidence distribution for parameters, and a confidence level is determined to describe the probability that a calculated confidence interval will

encompass the true values of the parameters (Patterson et al., 2001). Assumptions related to frequentist methods change the way the uncertainty is interpreted.

Frequentist methods provide a confidence interval that under repeated sampling of new data sets under identical conditions, a correctly defined 95% confidence interval should contain the true value in 95% of the cases. By contrast, Bayesian methods quantify uncertainty via a posterior distribution which can be summarized by a Bayesian credible interval. A credible interval of, for example 95 % would be interpreted as displaying a 95% certainty for the true value of the parameter to be contained in the interval. Dormann et al., (2018) provide more illustrations about the differences of both approaches (Bayesian and frequentist) during the estimation of uncertainty.

UA based on sampling-based approach: The Monte Carlo analysis is a sampling-based approach to quantify uncertainty. It is based on performing multiple model evaluations with probabilistically selected model input. Using this approach, the inputs are randomly drawn from probability distributions for each individual input and as a result probability distributions of the outputs are produced (Katz 2002). Using Monte Carlo analysis the results of model evaluations are used to determine both the uncertainty in model predictions and the input variables that give rise to this uncertainty (Helton 1993). The advantage of this analysis is that the UA is obtained directly from model predictions without the use of an intermediate surrogate model. Sampling-based UA via Monte Carlo approaches plays a central role in the characterization and quantification of uncertainty (Janssen 2013). However, the major drawback of Monte Carlo procedures is the fact that multiple model evaluations are required. If the model under evaluation is computationally expensive or if many model evaluations are required, then the cost of implementing a Monte Carlo analysis may be very large (Helton 1993). To reduce the computational cost inherent to Monte Carlo analyses, the minimisation of the number of simulations by the use of sampling techniques is a good solution, for example the Latin hypercube sampling has shown advanced sampling efficiency (McKay et al., 1979).

The Latin hypercube sampling (LHS) was created with the aim to improve the efficiency of Monte Carlo analysis regarding the coverage of the input space (Iman 2014). By using the LHS,

each probability distribution function of the k input parameters (X_1, X_2, ..., X_k) are stratified by dividing them into n non overlapping intervals of equal length, where n is the number of model simulations to be run. One value from each interval is selected at random with respect to the probability density in the interval. The n values thus obtained for X_1 are paired in a random manner with the n values of X_2. These n pairs are combined in a random manner with the n values of X_3 to form n triplets, and so on, until a set of n k-tuples is formed. This set of k-tuples is the Latin hypercube sample (Iman and Helton 1988). LHS is an efficient sampling design because each value of each parameter is used only once in the analysis. The structure of LHS is also helpful for simulations that are computationally expensive, as the method reduces the number of model evaluations based on the input stratification. LHS designs have good uniformity with respect to each input dimension individually. However, desirable properties such as space filling, or column-wise orthogonality have a very expensive optimization. In addition, this method also involves dimensionality issues. While the uniformity in each dimension is preserved, the space filling properties become questionable. As the number of variables increases, it becomes harder to fill the design space (Viana 2013).

UA based on a modelling approach: The response surface method is based on using an experimental design to select a set of specific values and pairings of the input parameters that are used to run a model. The resulting model outputs and inputs are then used to estimate the parameters of a new estimated model (Helton 2005). This estimated model is a fitted response surface that is used in place of the original model as a surrogate, and from which all the inferences with respect to UA and SA are derived (Iman and Helton 1988). Principal characteristics of the response surface method is that there is a complete control over the structure of model input through the experimental design selected for use. In addition, the UA and SA are inexpensive and straightforward once the necessary response surface approximation has been constructed (Helton 1993). However the main drawbacks are that this method will work when there are only few (typically less than 10) input parameters, a limited number of outputs, and the relationship between the inputs and output variables are basically linear or quadratic or involve few cross-products (Helton 1993).

The differential analysis is based on using a Taylor series as a surrogate for the original model for undertaking UA and SA. The dependent variable of interest is treated as a function *f* of the independent variables. Then a first-order Taylor series is constructed for the function *f* using a vector of base-case values (derivatives) for the independent variables (Iman and Helton 1988). An important step is the generation of the partial derivatives required for the series. For the UA, the Taylor series approximation can be used in conjunction with Monte Carlo simulations to estimate the distribution functions of model outputs (Iman and Helton 1988). Under this approach, the quality of the results will depend on how well this series approximates the original model, and the UA will be straightforwardly implemented once the Taylor series is developed (Helton 1993). Although differential analysis has been widely applied for simple models, this approach is a poor technique in the estimation of distribution functions of model outputs. This method also does not provide information on the possible existence of thresholds or discontinuities in the relationships between input and output variables (Helton 1993).

3.3 Methods for sensitivity analyses (SA)

There are many techniques to perform a sensitivity analysis, the aim of which is to quantify the relation between a model's factors and its responses, considering both parameters and initial values of state-variables as factors. Experimental design and screening methods are key elements related to SA. Normally we apply them before performing a SA. In Box 1 we describe these elements.

Box 1. What to do before a SA?

A. Experimental design: Working with a model of *k* factors, the ranges of possible factor values form a k-dimensional factor space. The factor space is also known as the design space. The experimental design is the description of how the design space has to be sampled (Pantus 2006). The objective is to characterize the sensitivity of a model to the factors, where the design space can also include combinations of factors called *interactions*. However, when these exercises are performed for complex models, like ecosystem models, there are computational limitations due to the potentially large number of factors and the multidimensional input spaces with the requirement of a huge

number of evaluations to estimate the response's sensitivity to the factors. To reduce the dimensional problem screening methods have been proposed.

B. Screening methods: Screening methods allow for efficient estimation of effects using a minimal number of simulation evaluations. These techniques are appropriate when there is little knowledge on which factors have an influence on the model responses and there is a need to identify, among a large number of potentially significant factors, those which are effectively influential within a fixed experimental domain. The statistics literature contains different screening methods. The resolution III and IV Hadamard developed by Plackett and Burman (Plackett and Burman 1946) is one of the first screening methods that can be found in the literature. The supersaturated method was firstly developed in the 1950's by Satterthwaite (1959) as a random balance and then by Booth and Cox (1962) as a systematic manner. Later the group screening was developed by G. Watson (1961) who suggested the idea of grouping factors to reduce the number of parameters and considered each group as a single factor, hence the name of the method. Multiple group screening proposed by Morris (1987) where each factor is assigned to more than one group, a factor being potentially influential if all groups containing this factor are active. Finally, the sequential bifurcation is a more recent group-screening technique (Bettonvil and Kleijnen 1997) that includes sequences of steps in the screening. Dupuy et al., (2014) presented an extended comparison of screening methods.

3.3.1. Brief overview of the SA methods

Important SA settings were presented in Cariboni et al., (2007). The first of them focused on the aim of SA methods, which can be the identification of the most important factor(s) which, if fixed to their values, would lead to the greatest reduction in the variance of the output. This setting is known as the factors prioritization setting. Another important setting of SA methods is the identification of those factors that can be fixed at any given value in their domains without significantly reducing the output variance. But in some cases, the aim of SAs can be the reduction of the output variance to a lower threshold (variance cutting setting) by simultaneously fixing the smallest number of input factors.

The selection of the best SA method for a particular study can be guided by several criteria according to Cariboni et al., (2007). The selected method should be sensitive to the range of variation of each input factor and to the shape of its probability distribution. Then, the method

should also operate simultaneously on all uncertainty inputs, so that interactions among factors can be detected. Thus, an ideal method should allow to treat groups of factors as a single factor, allowing for a synthesis of the results and an easier interpretation.

Methods of SA can be divided into qualitative and quantitative methods regarding the type of the SA measure to be estimated. Qualitative methods provide sensitivity measures that, as its name suggests, are qualitative, i.e. measures capable of ranking the input factors in order of importance but not of quantifying how much a given factor is more important than others. For example, these methods are able to detect influential factors within a system among many other non-influential ones. While quantitative methods are designed to provide quantitative measures like for example the exact percentage of the total output variance that each factor (or group of factors) is accounting for (Cariboni et al., 2007; Campolongo, Tarantola, and Saltelli 1999).

SA methods can also be classified into local and global methods regarding the way the design space is sampled. Local SA methods are based on computing the local response of the output(s) obtained by varying input factors and holding the other factors fixed to a central (nominal) value (Saltelli, Tarantola, and Chan 1999). They imply the assumption (rarely satisfied) that the input-output relationship is linear and that the input parameters have a small interval of fractional variation around a nominal value (Campolongo, Tarantola, and Saltelli 1999). Elementary OAT method (one-factor-at-a-time) is considered as a local SA method that provides qualitative sensitivity analysis measures (Campolongo, Tarantola, and Saltelli 1999). In this approach the effect of the variation of a single factor is estimated keeping all the other factors fixed at their nominal values; and the simplest and most intuitive way to obtain a local sensitivity index is to compute derivatives, however this is alternatively done by normalizing the derivatives by factors' standard deviations (Cariboni et al., 2007). Providing qualitative sensitivity measures as well, Morris OAT (Morris 1991) and Extended Morris (Campolongo and Braddock 1999) are methods computationally more expensive than Elementary OAT but they can already be considered global methods (Campolongo, Tarantola, and Saltelli 1999).

Global SA methods are generally used when the model is nonlinear and the input variables are affected by uncertainties at different orders of magnitude (Campolongo, Tarantola, and Saltelli

1999). They focus on the output uncertainty over the entire range of values of the input parameters (Homma and Saltelli 1996). As part of global quantitative approaches, methods like standardized regression coefficients (SRCs) can be viewed as an attempt to overcome the limitations of local methods based on Monte Carlo simulations. They reflect the shape of the probability distribution of each factor allowing the estimation of the model coefficient of determination which represents the fraction of the output variance explained by the regression model itself (Cariboni et al., 2007). The most used global methods to decompose the output variance into the contributions imputable to each input factor are the FAST (Fourier Amplitude Sensitivity Test) and Sobol's methods. FAST (Cukier et al., 1973) uses spectral analysis for the variance decomposition and Sobol' method (Sobol 1990) uses Monte Carlo methods for the variance decomposition. FAST and Sobol's methods allow to estimate two measures which summarize the model behavior: the main effect index (or first-order effect of factor) which is estimated with respect to each factor individually, and the total effect index which is estimated taking account of all the contributions to the output variation due to the factor (first-order effects plus all its interactions). The above described SA methods can be summarized and compared in terms of computational cost and the information required (**Figure 2**).

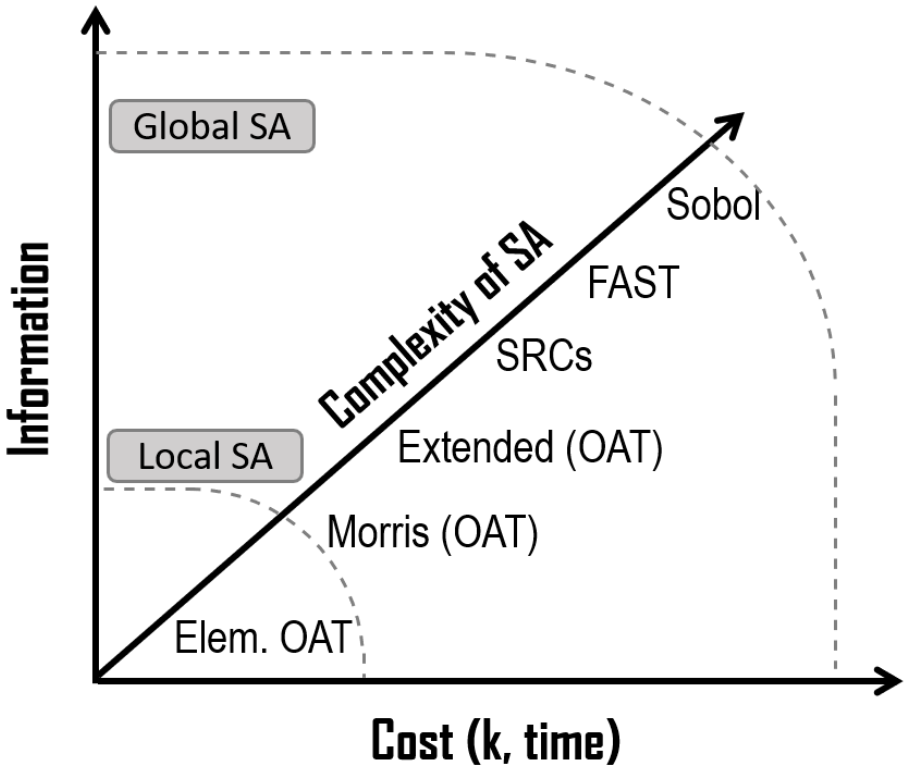

Figure 2. Sensitivity analysis methods*. The selection of the suitable SA method depends on the computational cost in terms of the number of model factors to evaluate (k) and of the CPU time required to run the model (time). On the other hand, the simplest methods (like local SA: Elementary OAT) require less amount of data than most complex methods (like global SA: Morris OAT, Extended Morris OAT, Standardized regression coefficients (SRCs), FAST and Sobol). (Adapted from Cariboni et al., 2007)*

3.3.2. Application of SA to quantify uncertainty

SA in combination with other specific methods have been used to quantify different sources of uncertainty.

SA for parameter uncertainty: sensitivity analyses are commonly applied to analyze parameter uncertainty. For example, the Ecopath model software explicitly acknowledges the need for sensitivity analyses through the inclusion of the ECORANGER module, which allows for the consideration of different levels of uncertainty associated with the input parameters (Fulton 2001; Christensen, Walters and Pauly 2000). For handling parameter uncertainty when the

models are characterized by a large number of parameters, approaches like functionality filters (Pantus 2006) can be used to identify the most sensitive parameter subsets or system components. Nonetheless, the combination of SA with UA is a good option to address the uncertainty quantification in cases where models have a large number of parameters but when other approaches are not allowed. In those cases SA allows identification of the parameters that influence model output and UA allows the uncertainty quantification on the results conditional on an assumed level of uncertainty on these parameters. Example of this approach is presented in Lehuta et al., 2010. Models limited by observational data (this can be the case for both complex integrated ecosystem models or simpler models focused on certain components of the system), it is difficult to use data-demanding statistical methods like Bayesian approaches. In these cases, the use of sensitivity analyses is to date the only option to explore the robustness of the model outputs to uncertainties associated with specific input parameter values (Payne et al., 2015).

SA for initialization and internal variability uncertainty: Perturbation analysis through the use of SA techniques can be applied to explore the initialization uncertainty. With this technique, the components of the system are perturbed to explore the flow of consequences to all system components (Fulton 2010). For example, Thompson et al., (2015) attempted to separate the variability of a climate model outputs into that due to internal variability and that due to external forcing. They estimated the uncertainty in future climate due to internal climate variability from large ensembles of climate change simulations, each of which resulting from a small perturbation in the initial atmospheric state. Since the same model and the same forcing data were used in all ensemble simulations, the differences in the climate output variables could be attributed to the internal variability uncertainty.

3.4 Software tools for uncertainty quantification

The collective scientific effort to achieve the quantification of uncertainty has generated the creation of software tools (also called “libraries” or “packages”) to address sensitivity analyses and uncertainty analyses. Here we present a set of libraries, which are most commonly used and are representative of the different features existing in the landscape of tools for uncertainty quantification (**Table 3**). These software programs provide open source code. They also adopt

diverse design philosophies as they have been developed in different disciplines and are fit to diverse purposes. The range of existing tools allows to address various model complexities, demanding the application of specific methods and were developed using diverse programming languages (including C, C++, Julia, Python, Matlab, R).

Table 3. Software tools for UA and SA*. Overview of principal tools available in the literature to perform UA and SA. They were developed across diverse programming languages (including C, C++, Python, R, Julia, and Matlab) but also for diverse purposes and applications.*

Software tools	Language	Principal characteristics
1. PSUADE: "Problem solving environment for Uncertainty analysis and Design Exploration" (Gan et al., 2014) (Tong, 2015)	C++	- Open source software developed to perform uncertainty quantification for large complex system models; - It has been developed principally for performing uncertainty quantification, UA, SA, and numerical optimization; - This software comprises three major components: a suite of sampling methods, a job execution environment, and a collection of analysis (including numerical optimization) tools.
2. UQLab: "Uncertainty Quantification in Matlab" (Marelli and Sudret, 2014)	Matlab	- Open source code with the objective of creating a powerful, modular and simple-to-extend software framework for uncertainty quantification (UQ) in engineering applications; - It is focused principally on uncertainty propagation through Monte Carlo sampling, SA, and reliability analysis; - It is designed to encourage both academic researchers and field engineers to use and develop algorithms for uncertainty quantification.
3. OpenTURNS (Baudin et al., 2017)	Python, C++	- Open source initiative to treat uncertainties, risks and statistics; - The main goal of this software is to provide all functionalities needed to treat uncertainties studies with industrial applications; - Targeted users are engineers who want to introduce a probabilistic dimension in their deterministic studies.
4. SALib (Herman and Usher 2017)	Python	- Open source software for performing SA; - Provides a decoupled workflow. It does not directly interface with mathematical or computational model but it is responsible for generating model inputs, using one of the proposed sample functions, and compute the sensitivity indices from the model outputs; - This software is useful in system modeling when looking for quantifying the effects of model inputs or exogenous factors on outputs of interest.
5. sensitivity: "Global Sensitivity Analysis of Model Outputs" (Iooss et al., 2020)	R	- Open source software available in the Comprehensive R Archive Network (CRAN) for performing global SA (GSA) of model outputs; - This tool includes methods for SA but also for screening analysis; - It works using the decoupled approach, designed to work with any model written in R or in another language, including heavy computational codes.
6. VARS-Tool (Razavi et al., 2019)	Matlab, C++	- Open source for UA and SA . It was developed primarily round the Variogram Analysis of Response Surfaces (VARS); - It adopts a multi-method approach that enables generation of sensitivity

		indices, including those based on derivative, variance, and variogram concepts from a simple sample; - This software is applicable to the full range of computer simulation models including Earth and Environmental System Models (EESMs), and includes a set of mathematical test functions using a dynamical hydrologic model.
7. MADS (Vesselinov et al., 2019)	Julia, C, C++	- Open source high-performance computational framework for data and model-based analysis; - It is focused principally on SA, parameter estimation, uncertainty quantification, surrogate modeling, machine learning and blind source separation, and decision analysis; - It can be internally or externally coupled with any existing model simulator; - This tool has been successfully applied to perform analyses related to environmental management.
8. Dakota (Adams et al., 2020)	C++	- Open source code which provides a framework for design optimization, parameter estimation, uncertainty quantification and SA; - This software is a flexible, extensible interface between simulation codes and a variety of iterative analysis methods; - It contains algorithms for optimization, uncertainty quantification, deterministic / stochastic calibration, and parameter SA.
9. SAFE: "Sensitivity for everybody" (Pianosi et al., 2015; Pianosi et al., 2020)	Matlab, R, Python	- Open source software developed for the application of GSA; - It has been developed for non-specialist users, with only a basic knowledge of GSA and Matlab; - This software includes three principal objectives: the application of multiple GSA, the analysis of the robustness of the SA indices, and also the use of tools to visualize the results.
10. SobolGSA (Kucherenko and Zaccheus, 2020)	C++, Matlab, Python	- It is a general purpose GUI (graphical user interface) driven global SA and metamodeling software; - This software provides diverse metamodeling techniques: quasi random sampling-high dimensional model representation (QRS-HDMR) method (with regression and projection methods) and radial basis function method; - This software can be applied to both static and time-dependent problems.

This set of libraries are diverse in terms of the methods included. *SALib* (Herman and Usher 2017) provides good user documentation with reproducible examples to implement SA, a feature which is useful for users who are not too familiar with routines implemented in Python. *sensitivity* (Iooss et al., 2020) is a collection of functions for factor screening, global sensitivity analysis and robustness analysis. It is the only exclusive library implemented in R. *SAFE* (Pianosi, Sarrazin, and Wagener 2015; Pianosi, Sarrazin, and Wagener 2020), originally developed in Matlab, is now available in R and Python. This tool includes screening methods and was developed for users with basic background in Global SA. *PSUADE* (Gan et al., 2014;

Tong 2015) is a library that was developed to perform uncertainty quantification for large complex system models in C++. This tool includes methods for UA, SA and numeric optimization. *UQLab* (Marelli and Sudret, 2014), a library developed exclusively in Matlab, focuses on uncertainty propagation through Monte Carlo sampling and reliability analysis. *OpenTURNS* (Baudin et al., 2017), *VARS-Tool* (Razavi et al., 2019), *MADS* (Vesselinov et al., 2019), *Dakota* (Adams et al., 2020), and *SobolGSA* (Kucherenko and Zaccheus, 2020) were all developed in C++ but also extended to other programming languages. *OpenTURNS* (Baudin et al., 2017) also available in Python treats uncertainties and performs risk analysis with industrial applications. *VARS-Tool* (Razavi et al., 2019) available also in Matlab for UA and SA extends its use to Earth and Environmental System Models (EESMs) including hydrological models. *MADS* (Vesselinov et al., 2019) is the only tool available in Julia but its code is also extended to C and C++. It includes methods for SA and uncertainty quantification, surrogate modeling, machine learning with application to environmental management. *Dakota* (Adams et al., 2020), available exclusively in C++ like PSUADE (Gan et al., 2014; Tong 2015), provides a framework for parameter estimation in face of uncertainty. Finally *SobolGSA* (Kucherenko and Zaccheus, 2020), available in Matlab and Python, provides methods to perform global SA and implement metamodelling analysis.

4. Challenges, gaps and perspectives

Achieving the quantification of uncertainty in models depends on both the characterization of the sources of uncertainty and also the methods and tools used for the uncertainty quantification. The perception of uncertainty in biodiversity models has clearly evolved since it now appears as a requisite step when using models in projections and scenarios (Ferrier et al., 2016) but also as an essential step for modelling and policy support across many fields including earth system modelling, water quality modeling, hydrology, among others (Razavi et al., 2021).

With this review, we provided insights and examples describing the different sources of uncertainty in the use of models. The methods and tools developed for sensitivity and uncertainty analyses are however not always adapted to complex biodiversity models with a large number of variables and parameters, often large errors and gaps in observations, and requiring high

computational costs. The increasing number of collaborations between scientists from different disciplines (modelers, programmers, statisticians with different areas of applications) is helping to move forward, identify gaps, and adapt methods to complex models.

The difficulties of quantifying uncertainty in the use of biodiversity models is not only due to model complexity. Methods and applications of uncertainty quantification are indeed presented for a range of simple to complex models. The challenge of addressing uncertainty quantification is also related to the shortage of software guidelines, best practice options and comprehensive case studies that could be used as templates. For example, there are still relatively few examples and methods dealing with the quantification of internal and initial conditions uncertainty. Nonetheless, there is a growing literature in the field of hydrological, physical and climate models where methods and tools could be beneficially transposed to biodiversity models and eventually improved to adapt to the specificities of biodiversity models.

A diversity of programming languages has been used for the development of libraries dedicated to run SA and UA. These are freely accessible tools and provide users with a range of options to implement the analyses according to their programming skills. Further development of the tools, the inclusion of more SA and UA methods and increase of the generality of the tools would attract more users and motivate more implementations of uncertainty quantification in the use of biodiversity models. Nevertheless, to be successfully and widely used, these tools must include reproducible examples of varying levels of complexity, and provide detailed practical user guidelines. This is a first key step towards the achievement of uncertainty quantification.

5. Acknowledgements

This work is part of the DUOS project funded by the Pew marine fellows programme and the USBIO project funded by the LabEx CeMEB (ANR-10-LABX-04-01). CL was supported by a peruvian individual doctoral research grant from the "Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica" (contract N° 134-2016-FONDECYT). CL and YJS were partially funded by the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817578 (TRIATLAS project), N° 869300 (FutureMARES

project), the Biodiversa and Belmont Forum project SOMBEE (BiodivScen programme, ANR contract N°ANR-18-EBI4-0003-01).

6. References

1. Adams, B.M., Bohnhoff, W.J., Dalbey, K., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Hooper, R., Hough, P.D., Hu, K., Jakeman, J.D., Khalil, M., Maupin, K.A., Monschke, J.A., Ridgway, E.M., Rushdi, A., Seidl, D.T., Stephens, J.A., Swiler, L.P., Winokur, J., 2020. Dakota A Multilevel Parallel Object-Oriented Framework for Design Optimization Parameter Estimation Uncertainty Quantification and Sensitivity Analysis: Version 6.12 User's Manual. (No. SAND2020-5001). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); Sandia National Lab. (SNL-CA), Livermore, CA (United States). https://doi.org/10.2172/1630694.
2. Alkemade, Rob, Mark van Oorschot, Lera Miles, Christian Nellemann, Michel Bakkenes, and Ben Ten Brink. 2009. "GLOBIO3: A Framework to Investigate Options for Reducing Global Terrestrial Biodiversity Loss." *Ecosystems* 12 (3): 374–90. https://doi.org/10.1007/s10021-009-9229-5.
3. Baudin, Michaël, Anne Dutfoy, Bertrand Iooss, and Anne-Laure Popelin. 2017. "Open Turns: An Industrial Software for Uncertainty Quantification in Simulation." In, 6:2001–38. Springer. https://hal.archives-ouvertes.fr/hal-01107849v2.
4. Bellard, Céline, Cleo Bertelsmeier, Paul Leadley, Wilfried Thuiller, and Franck Courchamp. 2012. "Impacts of Climate Change on the Future of Biodiversity." *Ecology Letters* 15 (4): 365–77. https://doi.org/10.1111/j.1461-0248.2011.01736.x.
5. Bettonvil, Bert, and Jack PC Kleijnen. 1997. "Searching for Important Factors in Simulation Models with Many Factors: Sequential Bifurcation." *European Journal of Operational Research* 96 (1): 180–94. https://doi.org/10.1016/S0377-2217(96)00156-7.
6. Booth, Kathleen HV, and David R Cox. 1962. "Some Systematic Supersaturated Designs." *Technometrics* 4 (4): 489–95. https://doi.org/10.2307/1266285.
7. Campolongo, Francesca, and Roger Braddock. 1999. "The Use of Graph Theory in the Sensitivity Analysis of the Model Output: A Second Order Screening Method." *Reliability Engineering & System Safety* 64 (1): 1–12. https://doi.org/10.1016/S0951-8320(98)00008-8.
8. Campolongo, F, S Tarantola, and A Saltelli. 1999. "Tackling Quantitatively Large Dimensionality Problems." *Computer Physics Communications* 117 (1-2): 75–85. https://doi.org/10.1016/S0010-4655(98)00165-9.
9. Cariboni, J, D Gatelli, R Liska, and A Saltelli. 2007. "The Role of Sensitivity Analysis in Ecological Modelling." *Ecological Modelling* 203 (1-2): 167–82. https://doi.org/10.1016/j.ecolmodel.2005.10.045.
10. Charles, Anthony T. 1998. "Living with Uncertainty in Fisheries: Analytical Methods, Management Priorities and the Canadian Groundfishery Experience." *Fisheries Research* 37 (1-3): 37–50. https://doi.org/10.1016/S0165-7836(98)00125-8.
11. Cheung, William WL, Thomas L Frölicher, Rebecca G Asch, Miranda C Jones, Malin L Pinsky, Gabriel Reygondeau, Keith B Rodgers, et al. 2016. "Building Confidence in Projections of the Responses of Living Marine Resources to Climate Change." *ICES Journal of Marine Science* 73 (5): 1283–96. https://doi.org/10.1093/icesjms/fsv250.
12. Christensen, V., C.J. Walters and D. Pauly 2000. ECOPATH with ECOSB4: A user's guide. October 2000 edition. Fisheries Centre, University of British Columbia, Vancouver, Canada and International Centre for Living Aquatic Resources Management, Penang, Malaysia.
13. Cukier, RI, CM Fortuin, Kurt E Shuler, AG Petschek, and JH Schaibly. 1973. "Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I Theory." *The Journal of Chemical Physics* 59 (8): 3873–8. https://doi.org/10.1063/1.1680571.
14. Dormann, Carsten F, Justin M Calabrese, Gurutzeta Guillera-Arroita, Eleni Matechou, Volker Bahn, Kamil Bartoń, Colin M Beale, et al. 2018. "Model Averaging in Ecology: A Review of Bayesian, Information-Theoretic, and Tactical Approaches for Predictive Inference." *Ecological Monographs* 88 (4): 485–504. https://doi.org/10.1002/ecm.1309.
15. Dupuy, D., Corre, B., Claeys-Bruno, M., & Sergent, M. (2014). Comparison of different screening methods. *Case Studies In Business, Industry And Government Statistics*, *5*(2), 115-125.

16. Ferrier, S., Ninan, K.N., Leadley, P., Alkemade, R., Acosta, L.A., Akçakaya, H.R., Brotons, L., Cheung, W., Christensen, V., Harhash, K.A., et al. (2016). IPBES (2016): The Methodological Assessment Report on Scenarios and Models of Biodiversity and Ecosystem Services. Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services Bonn, Germany.
17. Fulton, Elizabeth A. 2010. "Approaches to End-to-End Ecosystem Models." *Journal of Marine Systems* 81 (1-2): 171–83. https://doi.org/10.1016/j.jmarsys.2009.12.012.
18. Fulton, Elizabeth Ann. 2001. "The Effects of Model Structure and Complexity on the Behaviour and Performance of Marine Ecosystem Models." PhD Thesis, Australia: University of Tasmania.
19. Gan, Yanjun, Qingyun Duan, Wei Gong, Charles Tong, Yunwei Sun, Wei Chu, Aizhong Ye, Chiyuan Miao, and Zhenhua Di. 2014. "A Comprehensive Evaluation of Various Sensitivity Analysis Methods: A Case Study with a Hydrological Model." *Environmental Modelling & Software* 51: 269–85. https://doi.org/10.1016/j.envsoft.2013.09.031.
20. Hammen, VC, and J Settele. 2011. "Biodiversity and the Loss of Biodiversity Affecting Human Health." *Encyclopedia of Environmental Health*, 353–62. https://doi.org/10.1016/B978-0-444-52272-6.00655-3.
21. Harwood, John, and Kevin Stokes. 2003. "Coping with Uncertainty in Ecological Advice: Lessons from Fisheries." *Trends in Ecology & Evolution* 18 (12): 617–22. https://doi.org/10.1016/j.tree.2003.08.001.
22. Helton, Jon C. 1993. "Uncertainty and Sensitivity Analysis Techniques for Use in Performance Assessment for Radioactive Waste Disposal." *Reliability Engineering & System Safety* 42 (2-3): 327–67. https://doi.org/10.1016/0951-8320(93)90097-I.
23. Helton, Jon C, FJ Davis, and Jay D Johnson. 2005. "A Comparison of Uncertainty and Sensitivity Analysis Results Obtained with Random and Latin Hypercube Sampling." *Reliability Engineering & System Safety* 89 (3): 305–30. https://doi.org/10.1016/j.ress.2004.09.006.
24. Helton, Jon C, and Freddie Joe Davis. 2003. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems." *Reliability Engineering & System Safety* 81 (1): 23–69. https://doi.org/10.1016/S0951-8320(03)00058-9.
25. Herman, Jon, and Will Usher. 2017. "SALib: An Open-Source Python Library for Sensitivity Analysis." *The Journal of Open Source Software* 2 (9). https://doi.org/10.21105/joss.00097.
26. Holland, RA, WRT Darwall, and KG Smith. 2012. "Conservation Priorities for Freshwater Biodiversity: The Key Biodiversity Area Approach Refined and Tested for Continental Africa." *Biological Conservation* 148 (1): 167–79. https://doi.org/10.1016/j.biocon.2012.01.016.
27. Homma, Toshimitsu, and Andrea Saltelli. 1996. "Importance Measures in Global Sensitivity Analysis of Nonlinear Models." *Reliability Engineering & System Safety* 52 (1): 1–17. https://doi.org/10.1016/0951-8320(96)00002-6.
28. Iman, Ronald L. 2014. "Latin Hypercube Sampling." *Wiley StatsRef: Statistics Reference Online*.
29. Iman, Ronald L, and Jon C Helton. 1988. "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models." *Risk Analysis* 8 (1): 71–90. https://doi.org/10.1111/j.1539-6924.1988.tb01155.x.
30. Iooss, Bertrand, Sebastien Da Veiga, Alexandre Janon, Gilles Pujol, with contributions from Baptiste Broto, Khalid Boumhaout, Thibault Delage, et al. 2020. *Sensitivity: Global Sensitivity Analysis of Model Outputs*. https://CRAN.R-project.org/package=sensitivity.
31. Janssen, Hans. 2013. "Monte-Carlo Based Uncertainty Analysis: Sampling Efficiency and Sampling Convergence." *Reliability Engineering & System Safety* 109: 123–32. https://doi.org/10.1016/j.ress.2012.08.003.
32. Jørgensen, Sven Erik, and Brian D Fath. 2011. *Fundamentals of Ecological Modelling*. Fourth edition. Elsevier.
33. Katz, Richard W. 2002. "Techniques for Estimating Uncertainty in Climate Change Scenarios and Impact Studies." *Climate Research* 20 (2): 167–85. https://doi.org/10.3354/cr020167.
34. Kucherenko, S., Zaccheus, O., 2020. SobolGSA. http://www.imperial.ac.uk/a-z-research/process-systems-engineering/research/free-software/sobolgsa-software/.
35. Lavergne, Sébastien, Nicolas Mouquet, Wilfried Thuiller, and Ophélie Ronce. 2010. "Biodiversity and Climate Change: Integrating Evolutionary and Ecological Responses of Species and Communities." *Annual Review of Ecology, Evolution, and Systematics* 41: 321–50. https://doi.org/10.1146/annurev-ecolsys-102209-144628.

36. Lehuta, Sigrid, Stéphanie Mahévas, Pierre Petitgas, and Dominique Pelletier. 2010. "Combining Sensitivity and Uncertainty Analysis to Evaluate the Impact of Management Measures with Isis-Fish: Marine Protected Areas for the Bay of Biscay Anchovy (Engraulis Encrasicolus) Fishery." *ICES Journal of Marine Science* 67 (5): 1063–75. https://doi.org/10.1093/icesjms/fsq002.
37. Lele, Subhash R. 2020. "How Should We Quantify Uncertainty in Statistical Inference?" *Frontiers in Ecology and Evolution* 8: 35. https://doi.org/10.3389/fevo.2020.00035.
38. Leung, Brian, Nuria Roura-Pascual, Sven Bacher, Jaakko Heikkilä, Lluis Brotons, Mark A Burgman, Katharina Dehnen-Schmutz, et al. 2012. "TEASIng Apart Alien Species Risk Assessments: A Framework for Best Practices." *Ecology Letters* 15 (12): 1475–93. https://doi.org/10.1111/ele.12003.
39. Link, Jason S, TF Ihde, CJ Harvey, Sarah K Gaichas, JC Field, JKT Brodziak, HM Townsend, and RM Peterman. 2012. "Dealing with Uncertainty in Ecosystem Models: The Paradox of Use for Living Marine Resource Management." *Progress in Oceanography* 102: 102–14. https://doi.org/10.1016/j.pocean.2012.03.008.
40. Ludwig, Donald, Marc Mangel, and Brent Haddad. 2001. "Ecology, Conservation, and Public Policy." *Annual Review of Ecology and Systematics* 32 (1): 481–517. https://doi.org/10.1146/annurev.ecolsys.32.081501.114116.
41. Marelli, Stefano, and Bruno Sudret. 2014. "UQLab: A Framework for Uncertainty Quantification in Matlab." In *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, 2554–63. American Society of Civil Engineers, Liverpool,UK. https://doi.org/10.1061/9780784413609.257.
42. McKay, Michael D, Richard J Beckman, and William J Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21 (2): 239–45. https://doi.org/10.2307/1268522.
43. Morris, Max D. 1987. "Two-Stage Factor Screening Procedures Using Multiple Grouping Assignments." *Communications in Statistics-Theory and Methods* 16 (10): 3051–67. https://doi.org/10.1080/03610928708829557.
44. Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., et al. (2015). Predictive ecology in a changing world. J. Appl. Ecol. 52, 1293–1310. https://doi.org/10.1111/1365-2664.12482.
45. Pantus, Francis John. 2006. "Sensitivity Analysis for Complex Ecosystem Models." PhD Thesis, Australia: Department of Mathematics - The University of Queesland.
46. Patterson, Kenneth, Robin Cook, Chris Darby, Stratis Gavaris, Laurence Kell, Peter Lewy, Benoît Mesnil, et al. 2001. "Estimating Uncertainty in Fish Stock Assessment and Forecasting." *Fish and Fisheries* 2 (2): 125–57. https://doi.org/10.1046/j.1467-2960.2001.00042.x.
47. Payne, Mark R, Manuel Barange, William WL Cheung, Brian R MacKenzie, Harold P Batchelder, Xochitl Cormon, Tyler D Eddy, et al. 2015. "Uncertainties in Projecting Climate-Change Impacts in Marine Ecosystems." *ICES Journal of Marine Science* 73 (5): 1272–82. https://doi.org/10.1093/icesjms/fsv231.
48. Pereira, Henrique M, Paul W Leadley, Vânia Proença, Rob Alkemade, Jörn PW Scharlemann, Juan F Fernandez-Manjarrés, Miguel B Araújo, et al. 2010. "Scenarios for Global Biodiversity in the 21st Century." *Science* 330 (6010): 1496–1501. https://doi.org/10.1126/science.1196624.
49. Pianosi, Francesca, Fanny Sarrazin, and Thorsten Wagener. 2015. "A Matlab Toolbox for Global Sensitivity Analysis." *Environmental Modelling & Software* 70: 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009.
50. Pianosi, Francesca and Sarrazin, Fanny and Wagener, Thorsten. 2020. "How Successfully Is Open-Source Research Software Adopted? Results and Implications of Surveying the Users of a Sensitivity Analysis Toolbox." *Environmental Modelling & Software* 124: 104579. https://doi.org/10.1016/j.envsoft.2019.104579.
51. Plackett, Robin L, and J Peter Burman. 1946. "The Design of Optimum Multifactorial Experiments." *Biometrika* 33 (4): 305–25. https://doi.org/10.2307/2332195.
52. Razavi, Saman, Razi Sheikholeslami, Hoshin V Gupta, and Amin Haghnegahdar. 2019. "VARS-Tool: A Toolbox for Comprehensive, Efficient, and Robust Sensitivity and Uncertainty Analysis." *Environmental Modelling & Software* 112: 95–107. https://doi.org/10.1016/j.envsoft.2018.10.005.
53. Razavi, Saman, Anthony Jakeman, Andrea Saltelli, Clémentine Prieur, Bertrand Iooss, Emanuele Borgonovo, Elmar Plischke, et al. 2021. "The Future of Sensitivity Analysis: An Essential Discipline for Systems Modeling and Policy Support." *Environmental Modelling & Software* 137: 104954. https://doi.org/10.1016/j.envsoft.2020.104954.

54. Regan, Helen M, Mark Colyvan, and Mark A Burgman. 2002. "A Taxonomy and Treatment of Uncertainty for Ecology and Conservation Biology." *Ecological Applications* 12 (2): 618–28. https://doi.org/10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2.
55. Rounsevell, M. D., Arneth, A., Brown, C., Cheung, W. W., Gimenez, O., Holman, I., ... & Shin, Y. J. (2021). Identifying uncertainties in scenarios and models of socio-ecological systems in support of decision-making. *One Earth*, *4*(7), 967-985.
56. Sala, Osvaldo E, F Stuart Chapin, Juan J Armesto, Eric Berlow, Janine Bloomfield, Rodolfo Dirzo, Elisabeth Huber-Sanwald, et al. 2000. "Global Biodiversity Scenarios for the Year 2100." *Science* 287 (5459): 1770–4. https://doi.org/10.1126/science.287.5459.1770.
57. Saltelli, A, K Chan, and M Scott. 1999. "Sensitivity Analysis of Model Output-Samo 98-Second International Symposium on Sensitivity Analysis of Model Output-Venice, 19-22 April 1998-Preface." ELSEVIER SCIENCE BV PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS.
58. Saltelli, Andrea, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Li Sushan, and Qiongli Wu. 2019. "Why so Many Published Sensitivity Analyses Are False: A Systematic Review of Sensitivity Analysis Practices." *Environmental Modelling & Software* 114: 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012.
59. Saltelli, Andrea, Stefano Tarantola, and KP-S Chan. 1999. "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output." *Technometrics* 41 (1): 39–56. http://doi.org/10.2307/1270993.
60. Satterthwaite, FE. 1959. "Random Balance Experimentation." *Technometrics* 1 (2): 111–37. https://doi.org/10.2307/1266466.
61. Skinner, Daniel JC, Sophie A Rocks, Simon JT Pollard, and Gillian H Drew. 2014. "Identifying Uncertainty in Environmental Risk Assessments: The Development of a Novel Typology and Its Implications for Risk Characterization." *Human and Ecological Risk Assessment: An International Journal* 20 (3): 607–40. https://doi.org/10.1080/10807039.2013.779899.
62. Smith, TM, and RL Smith. 2012. *Elements of Ecology*. Eighth edition. Pearson Benjamin Cummings.
63. Soboĺ, IM. 1990. "Quasi-Monte Carlo Methods." *Progress in Nuclear Energy* 24 (1-3): 55–61. https://doi.org/10.1016/0149-1970(90)90022-W.
64. Thompson, David WJ, Elizabeth A Barnes, Clara Deser, William E Foust, and Adam S Phillips. 2015. "Quantifying the Role of Internal Climate Variability in Future Climate Trends." *Journal of Climate* 28 (16): 6443–56. https://doi.org/10.1175/JCLI-D-14-00830.1.
65. Tittensor, D. P., T. D. Eddy, H. K. Lotze, E. D. Galbraith, W. Cheung, M. Barange, J. L. Blanchard, et al. 2018. "A Protocol for the Intercomparison of Marine Fishery and Ecosystem Models: Fish-Mip V1.0." *Geoscientific Model Development* 11 (4): 1421–42. https://doi.org/10.5194/gmd-11-1421-2018.
66. Tong, C., 2015. PSUADE Reference Manual (v1.7). Lawrence Livermore National Laboratory, Livermore, CA. https://computing.llnl.gov/projects/psuade-uncertainty-quantification
67. Van Oijen, Marcel. 2017. "Bayesian Methods for Quantifying and Reducing Uncertainty and Error in Forest Models." *Current Forestry Reports* 3 (4): 269–80. https://doi.org/10.1007/s40725-017-0069-9.
68. Vesselinov, Velimir V, Boian S Alexandrov, and Daniel O'Malley. 2019. "Nonnegative Tensor Factorization for Contaminant Source Identification." *Journal of Contaminant Hydrology* 220: 66–97. https://doi.org//10.1016/j.jconhyd.2018.11.010.
69. Viana, Felipe AC. 2013. "Things You Wanted to Know About the Latin Hypercube Design and Were Afraid to Ask." In *10th World Congress on Structural and Multidisciplinary Optimization*. Vol. 19. 24.05.
70. Watson, GS. 1961. "A Study of the Group Screening Method." *Technometrics* 3 (3): 371–88. https://doi.org/10.2307/1266728.

Chapter 2

Key species and indicators revealed by an uncertainty analysis of the OSMOSE model

Systematic analyses that study the impact of uncertainty in model outputs are essential for assessing the credibility of models, especially when they are used to support decision-making. In this work, we implemented an uncertainty analysis that quantifies the effect of parameter uncertainty on a set of ecological indicators in outputs of the OSMOSE model applied to the northern Humboldt Current ecosystem (NHCE OSMOSE). We worked under simple uncertainty assumptions corresponding to ranges of 10% and 20% variability around the reference values of the parameters describing the dynamics of the species modelled in NHCE OSMOSE. The results show that the uncertainty in the parameters of some species, i.e. euphausiids and Humboldt squid, have substantial repercussions on the rest of the food web. This work also allowed us to identify some ecological indicators that can dampen the uncertainty in the model's inputs, such as the marine trophic index, the trophic level of fish communities, and the slope of the size spectrum. We also highlight the difficulties of studying the uncertainty in a complex model, where the uncertainty could propagate differently over time and across different output variables. We suggest that more guidelines and methods could be useful to develop in the future for addressing the difficulties linked to working with complex models and the limitations linked to data availability.

Keywords: uncertainty analysis, parameter uncertainty, ecosystem model, northern Humboldt Current ecosystem, OSMOSE model.

Key species and indicators revealed by an uncertainty analysis of the OSMOSE model

Criscely Luján [1,2], Ricardo Oliveros-Ramos [3], Nicolas Barrier [4], Paul Leadley [5], Yunne-Jai Shin [2,5]

[1] *Laboratoire d'Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91400, Orsay. France*

[2] *MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Montpellier. France*

[3] *Instituto del Mar del Perú, Callao, Perú*

[4] *MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Sète. France*

[5] *University of Cape Town, Cape Town, South Africa*

* Corresponding author: criscelylujan@gmail.com

1. Introduction

Advances towards the sustainable development of fisheries worldwide have been made during the past two decades (Garcia-Lorenzo et al., 2021; Farmery et al., 2019; Lynch et al., 2017), particularly considering the role of marine biodiversity in providing multiple ecosystem services (Barbier 2017; Shin et al., in press and the urgency of implementing the Ecosystem Approach to Fisheries (EAF) (Jennings and Rice, 2011; Serpetti et al., 2017). Among these developments, ecosystem modelling has emerged as an useful approach for the practical implementation of the EAF (Plaganyi et al., 2007), while an increasing scientific community has been improving the scope and the performance of ecosystem models worldwide (e.g., Steenbeck et al. 2021, Heymans et al. 2020, Tittensor et al. 2018). However, one of the main criticisms of these models is that their potentially large complexity can make ecosystem model predictions highly uncertain (Fulton et al., 2003).

Uncertainty analysis (UA) characterizes the uncertainty in model predictions (Saltelli et al., 2019) by quantifying the variability in model outputs (Cariboni et al., 2007). Implementing an UA of ecosystem models is a complicated task due to their high computational requirements (in

terms of simulation time and memory space) and the multiple model evaluations that a UA requires (Fulton, 2010). For these reasons and despite the abundant literature explaining the importance of dealing with uncertainty in the use of ecosystem models (Steenbeek et al., 2021; Lehuta et al., 2016; Payne et al., 2016; Link et al., 2012), few practical applications have been published using UA (e.g., McGregor et al., 2020). However, to increase the credibility of these models, we urgently need to strengthen our capacity in achieving the quantification of the uncertainty in model results.

In this context, this work aims to understand the impact of uncertainty in a complex ecosystem model by using an UA. The main objective is to quantify the effect of parameter uncertainty on the OSMOSE model applied to the northern Humboldt Current ecosystem (the NHCE OSMOSE model, Oliveros-Ramos et al., 2017). The northern Humboldt Current ecosystem (NHCE), one of the four eastern boundaries upwelling systems, provides the highest fish production in the world, around 10% of global fish catches, and supports the fishery of the Peruvian anchovy (Chavez et al., 2008; Bakun and Weeks, 2008), the most significant single-species fishery around the world (Chavez et al., 2003). This ecosystem is also characterized by a high environmental variability at diverse scales (i.e., at seasonal, interannual, and decadal scales, Gutierrez et al., 2016). For the study of the NHCE and given its characteristics, it is necessary to use holistic frameworks that integrate the effects of fisheries and climate, among which ecosystem end-to-end models are key tools (Tittensor et al. 2018). Thus, several ecosystem models have already been developed for the NHCE. Some examples of this are the Ecopath with Ecosim (EwE) platform (Christensen et al., 2005; Walters et al., 1997), used in the works of Tam et al. (2008) and Taylor et al. (2008), and the OSMOSE modelling platform (Shin and Cury, 2001; 2004) used in Marzloff et al. (2009) and more recently in Oliveros-Ramos et al. (2017). However, none of these applications have studied the impact of uncertainty in the models.

We chose the NHCE OSMOSE model for this study because of its ability to consider the complexity and high stochasticity of the NHCE. Besides, this model provides a variety of ecological indicators as model outputs, e.g., size-based, species-based, trophic-level based (Shin et al., 2018; Fu et al., 2019), which allows us to analyze and compare the responses of a set of ecological indicators to different levels of uncertainty. In this work, we hypothesized that: 1) the uncertainty coming from one species' parameters could propagate through the food web, 2) the

response of ecological indicators to uncertainty is not homogeneous, and that 3) depending on the indicators chosen in output of the model, these could amplify or dampen the level of uncertainty in the model inputs.

2. Material and methods

To run the UA of the NHCE OSMOSE model, we first 1) selected the model input parameters to be used in this study, 2) which values were sampled using a range of variability, to then 3) run multiple model simulations using the sampled parameter values, and finally 4) characterize the uncertainty in the model outputs.

2.1 The NHCE OSMOSE model

We performed an UA on the NHCE OSMOSE model (Oliveros-Ramos et al., 2017). OSMOSE (Object-oriented Simulator of Marine ecOSystEms) is a size-based trophic model that represents the life story and spatio-temporal dynamics of fish and macro-invertebrate species (Shin and Cury, 2001; 2004). It is an individual-based model which assumes size-based opportunistic predation based on the spatial co-occurrence of a predator and its prey. OSMOSE models the major life cycle processes (i.e., growth, reproduction, predation, natural and starvation mortalities) and fisheries impact. Also, it is forced by physical and biogeochemical models, enabling to simulate the impacts of climate change and variability. OSMOSE applications simulate the complexity and stochasticity of marine ecosystems, integrating physical, biogeochemical, and biological processes. This model has been applied across diverse marine ecosystems (e.g., Travers et al., 2009, 2014; Fu et al., 2013; Oliveros-Ramos, 2014; Halouani et al., 2016; Moullec et al., 2019a). Complementary information about OSMOSE is on its official website (see http://www.osmose-model.org).

The NHCE OSMOSE model covers an extension of the northern Humboldt Current and the Peruvian Upwelling Ecosystem, between the ranges from 20°S to 6°N and 93°W to 70°W with 1/6° spatial resolution. Fitted to data time series from the years 1992 and 2008, the NHCE OSMOSE model includes 13 species (see Table 1): nine are explicitly modelled in OSMOSE (also called focal species) and four plankton groups (also called biotic resources) are represented in the PISCES biogeochemical model coupled to the ROMS physical model (Aumont and Bopp,

2006; Echevin et al., 2008). Additional information about the NHCE OSMOSE model can be found in Oliveros-Ramos (2014) and Oliveros-Ramos et al. (2017).

Table 1. *Species or functional groups included in the NHCE OSMOSE model.*

Types of model species	Species abbreviation	Group	Species or functional group	Scientific name	Model
Biotic resources	pg_1	Phytoplankton	Nanophytoplankton	-	ROMS-PISCES
	pg_2	Phytoplankton	Diatoms	-	ROMS-PISCES
	pg_3	Zooplankton	Microzooplankton	-	ROMS-PISCES
	pg_4	Zooplankton	Mesozooplankton	-	ROMS-PISCES
Focal species	sp_1	Small pelagics	Anchovy	*Engraulis ringens*	OSMOSE
	sp_2	Demersal	Peruvian hake	*Merluccius gayi peruanus*	OSMOSE
	sp_3	Small pelagics	Sardine	*Sardinops sagax*	OSMOSE
	sp_4	Medium pelagics	Jack mackerel	*Trachurus murphyi*	OSMOSE
	sp_5	Medium pelagics	Chub mackerel	*Scomber japonicus*	OSMOSE
	sp_6	Other pelagics	Mesopelagics	*Vinciguerria sp.*	OSMOSE
	sp_7	Other pelagics	Munida	*Pleuroncodes monodon*	OSMOSE
	sp_8	Other pelagics	Humboldt squid	*Dosidicus gigas*	OSMOSE
	sp_9	Zooplankton	Euphausiids	*Euphausia mucronata*	OSMOSE

2.1.1 Selection of model parameters

From the many possible NHCE OSMOSE model input parameters (related to the nine focal species), we focused on predation, growth, reproduction parameters, and the starvation, natural, fishing, and larval mortality rates. This encompassed 237 parameters used for the UA (and 18 parameter types, Table 2). See Appendix A and Table A.1 for the complete list of parameters values used in the UA. Besides, to ensure the compliance of mathematical constraints of the

parameters (e.g., parameters standardized between 0 and 1) we performed a re-parametrization of the NHCE model before implementing the UA (see Appendix B).

Table 2. *NHCE OSMOSE parameters used in the UA. The UA was implemented using baseline values which are outcomes of model re-parametrization. The f function describes the model re-parametrization. See Appendix A for additional information about baseline values of model parameters and Appendix B for details about model re-parametrization.*

N°	Parameter type	Species	Reparameterization for the UA	Parameter scale	Number of parameters
1	predation accessibility of prey to predators *A(predator, prey)*	for the 9 focal species	$x = f(A) = A$ (identity)	logit	72
2	minimum predator-prey size ratio for each species stage (θ_{stage})	for the 9 focal species	$x = f(\theta_{stage}) = \frac{\theta_{stage}}{(\pi/2)}$	logit	17
3	maximum predator-prey size ratio for each species stage (α_{stage})	for the 9 focal species	$x = f(\alpha_{stage}) = \frac{\alpha_{stage}}{(\pi/2)}$	logit	17
4	predator-prey size threshold (s_{thr} in cm)	for 7 focal species (sp_1, sp_2, sp_3, sp_4, sp_5, sp_8, sp_9)	$x = f(s_{thr}) = \frac{s_{thr}}{L_{\infty}}$	logit	8
5	maximum starvation mortality rate ($M_{\varepsilon_{max}}$ in y^{-1})	for the 9 focal species	$x = f(A) = A$ (identity)	logarithmic	9
6	von Bertalanffy threshold (a_{thr} in y)	for the 9 focal species	$x = f(a_{thr}) = \frac{a_{thr}}{a_{max}}$	logit	9
7	egg size (cm)	for the 9 focal species		logarithmic	9
8	critical threshold of predation efficiency (ε_{crit})	for the 9 focal species		logit	9
9	maximum rate of predation ingestion (I_{max} in $g\ body\ g^{-1} y^{-1}$)	for the 9 focal species		logarithmic	9

10	natural mortality rate (M in y^{-1})	for the 9 focal species	$x = f(A) = A$ (identity)	logarithmic	9
11	larval mortality rate (M_0 in $month^{-1}$)	for the 9 focal species		logarithmic	9
12	fishing mortality multiplier (f_m)	for 6 focal species (sp_2, sp_3, sp_4, sp_5, sp_8, sp_9)		logarithmic	6
13	sex ratio ($Frac_{fem}$)	for the 9 focal species		logit	9
14	$L_{t=0}$ (von Bertalanffy growth parameter in cm)	for the 9 focal species		logarithmic	9
15	k (von Bertalanffy growth parameter in y^{-1})	for the 9 focal species		logarithmic	9
16	L_∞ (von Bertalanffy growth parameter in cm)	for the 9 focal species		logarithmic	9
17	size at maturity (s_{mat} in cm)	for the 9 focal species	$x = f(s_{mat}) = \frac{s_{mat}}{L_\infty}$	logit	9
18	constant of proportionality of the allometric length-weight relationship (c in $g\ cm^{-3}$)	for the 9 focal species	$x = f(A) = A$ (identity)	logarithmic	9

2.2 Simulation design

When information about the parameter distributions is available, it is used to specify the probability density functions (e.g., in Brown et al., 2015; Engström et al., 2016). In the absence of information, another alternative is to use ranges of variability, i.e. using a numeric value and increase and decrease it on either side of the parameter reference value. This way, we create an extensive range of parameter values where the sampling will occur. We worked with 10% and 20% ranges in this study, which are standard ranges found in uncertainty studies (e.g., Lehuta et al., 2010; Ciric et al., 2012; Morris et al., 2014; Dantec-Nedelec et al., 2017).

We used the Monte Carlo method for the UA, undertaking multiple model simulations using sampled parameter values. This method performs numerous model evaluations using random samples. Then, the results of model evaluations are used to determine the uncertainty in model

outputs. However, a drawback of this method is the high number of simulations required during the sampling process. We combined the Morris method (Morris 1991) with the Monte Carlo simulation approach to deal with this problem. Morris is a screening method composed of several randomized one-factor-at-a-time (OAT) experiments. This method is recommended for models with long-run times and is typically used to perform global SA (e.g., Bracis et al., 2020; Specka et al., 2015). Nevertheless, in this work, we applied Morris only to sample the parameters' values and produce the design of experiments to be used in the Monte Carlo simulations.

Using the Morris method and considering a model with n parameters x_i (where $i = 1, 2, \ldots, n$), each x_i will be scaled to take on values in the interval [0,1]. It creates a discrete parameter space (Ω also called the n-dimensional unit hypercube) by dividing the parameter ranges into p discrete levels. Then, the model is evaluated for r replicates within the parameter space, each of them building a trajectory inside Ω. The starting point of a trajectory is selected randomly. Only a single parameter is changed for each trajectory, taking an element of the parameter space as a new value. In each trajectory, each parameter (x_i) is only modified once, so it results in $n+1$ simulations. This procedure is repeated r times (r trajectories), resulting in a computation cost of $r\ (n + 1)$ simulations. In this work, since each parameter was on a particular scale (see Table 2), the sampling was done on the corresponding transformed scale.

2.3 UA experiments

Using the NHCE OSMOSE model, we executed 18 experiments for the UA (see Table 3). The first nine correspond to the uncertainty scenario of 10% and the last nine to 20%. Inside each uncertainty scenario, the experiments evaluated the effect of the uncertainty from each species separately (called species uncertainty). For example, in experiment $sp_{1,10\%}$ we addressed the uncertainty arising from the parameters related to the species sp_1 using the scenario 10%; experiment $sp_{1,20\%}$ relates to sp_1 parameters using scenario 20%; and so on. We only perturbed the parameters' values related to the species under study in each experiment, leaving the rest of the model configuration unchanged (i.e., their corresponding baseline values).

***Table 3**. Eighteen experiments were run as part of the UA of the NHCE OSMOSE model. Scenario uncertainty 10% corresponds to the experiments using a range of variability of 10% for the parameters of the nine NHCE OSMOSE focal species, and scenario uncertainty 20% uses 20% of variability for the same species' parameter set. The complete list of parameters' values related to each species is in Table A.1 (Appendix A). The number of parameters related to each species is also indicated.*

Scenario uncertainty	Species uncertainty								
10%	anchovy ($sp_{1,10\%}$)	hake ($sp_{2,10\%}$)	sardine ($sp_{3,10\%}$)	jack mackerel ($sp_{4,10\%}$)	chub mackerel ($sp_{5,10\%}$)	mesopelagics ($sp_{6,10\%}$)	munida ($sp_{7,10\%}$)	humboldt squid ($sp_{8,10\%}$)	euphausiids ($sp_{9,10\%}$)
20%	anchovy ($sp_{1,20\%}$)	hake ($sp_{2,20\%}$)	sardine ($sp_{3,20\%}$)	jack mackerel ($sp_{4,20\%}$)	chub mackerel ($sp_{5,20\%}$)	mesopelagics ($sp_{6,20\%}$)	munida ($sp_{7,20\%}$)	humboldt squid ($sp_{8,20\%}$)	euphausiids ($sp_{9,20\%}$)
Number of parameters	27	27	27	27	27	23	23	30	26

We arbitrarily selected the Morris parameters. In each experiment, we performed r=200 trajectories (i.e., Morris replicates) by dividing the corresponding parameter range into eight levels (p = 8), including upper and lower bounds with values uniformly distributed between them, and using the grid jump (Δ) of 4/7 when p = 8 ($\Delta = p/[2(p-1)]$ that is recommended in Morris (1991). However, given the stochasticity of the OSMOSE model, ten simulation replicates were executed (i.e., OSMOSE replicates) per Morris run. We thus performed 984000 simulations for this work, considering the nine species and their parameters (Table 3), the two uncertainty scenarios, the 200 Morris trajectories, and the 10 OSMOSE replicates. Finally, we characterize the uncertainty in model outputs for a set of selected indicators.

2.4 Uncertainty characterization

We used ecological indicators as model outputs for uncertainty characterization (Table 4). Most of them were chosen from the IndiSeas program (http://www.indiseas.org/). This program aimed to analyze a set of ecological indicators to assess the ecosystem effects of fishing in the context of environmental change and provide decision support for fisheries management. This set of

indicators was tested against several performance criteria, namely sensitivity, specificity, and responsiveness (Shin et al. 2018; Fu et al. 2019), and have been already applied in studies using OSMOSE models (e.g., Halouani et al., 2019, Moullec 2019b).

***Table 4**. Ecological indicators used for the uncertainty characterisation using the NHCE OSMOSE model.*

Indicator	Calculation	Description	Suggested references
Biomass (by species)	B (tons)	Biomass for each focal species.	-
Mean length (ML) of fish in the community	$ML = \frac{\sum_i L_i N_i}{\sum_i N_i}$	Where L_i is the average length and N_i is the mean abundance of species i.	Shin et al. (2005)
Mean trophic level (MTL) in the community	$MTL = \frac{\sum_i TL_i B_i}{\sum_i B_i}$	Where TL_i is the trophic level and B_i the biomass (in tons) of species i.	Reed et al. (2017); Shin et al. (2018)
Mean lifespan (MLS) of community	$MLS = \frac{\sum_i (A_i B_i)}{\sum_i B_i}$ (years)	Where A_i is the lifespan (in years) and B_i the biomass (in tons) of species i. A_i is defined for each species in the model input configuration.	Shin et al. (2018); Fu et al. (2019)
Biomass over Yield	B/Y	Where B and Y are the total biomass and catch of all modelled species	Shin et al. (2010)
Marine trophic index (MTI)	$MTI = \frac{\sum_{i(TL>3.25)} TL_i Y_i}{\sum_{i(TL>3.25)} Y_i}$	Where TL_i is the trophic level and Y_i the catch (in tons) for species i with TL > 3.25.	Pauly and Watson (2005); Shannon et al. (2014)
Slope of the size spectrum (SSP) of community	$Ln(N_k) = \beta_0 + \beta_1 Ln(\frac{L_k + L_{k-1}}{2})$ $SSP = \beta_1$	Where N_k is the total fish numbers in the length interval $[L_{k-1}, L_k]$ and $k \in [1, ..., k_{max}]$ is the index of length class.	Rice and Gislason, (1996); Shin and Cury (2004)

Large fish index (LFI_x): LFI_{20}, LFI_{30}, LFI_{40} of community	$LFI_x = \frac{\sum_i B_{i_{(L>x)}}}{\sum_i B_i}$	Where B_i is the biomass of species i larger than x cm ($L > x$) and B the total biomass of the community. This indicator is calculated for $x = \{20cm, 30cm, 40cm\}$.	Greenstreet et al. (2011)

For each level of uncertainty (i.e., species (sp) and scenario (sc) uncertainty) we estimated the relative change (RC_{Ind}) in the indicator (Ind), which is calculated as:

$$RC_{ind} = \left(\frac{Ind_{sp,sc} - Ind_{baseline}}{Ind_{baseline}}\right) \quad (1)$$

The baseline represents the corresponding indicator estimated using the parameter reference values of the NHCE OSMOSE model with ten OSMOSE replicates. Ten OSMOSE replicates were also run for each $Ind_{sp,sc}$. Additionally, we estimated the coefficient of variations (CV) of the indicators, where the standard deviation and mean were calculated over each Monte Carlo simulation by time step. All simulations and analyses were performed on a supercomputer (Datarmor, hosted by IFREMER, https://wwz.ifremer.fr/pcdm/Equipement) using R 3.6.1 (R Core Team 2019) and the runup R package. This package is an open-source tool dedicated to implementing uncertainty and sensitivity analyses of complex models (https://github.com/CriscelyLP/run_up).

3. Results

3.1. The effects of species uncertainty on the modelled ecosystem

Under the two uncertainty scenarios (i.e., ranges of 10% and 20%), species uncertainty propagated through the modelled ecosystem (Fig. 1). For species like anchovy, munida, and humboldt squid, apart from the greatest uncertainty that arose from their own biological parameters, other focal species also caused uncertainty. For example, the uncertainty on anchovy biomass arises from the uncertainty in its own parameters and those of euphausiids. While for species like sardine, jack mackerel, chub mackerel, and mesopelagics, the main source of uncertainty comes from other modeled focal species. Only in the case of euphausiids, the uncertainty in its biomass comes directly from the uncertainty in its own parameters. Besides, in

general, the uncertainty due to the parameters of humboldt squid and especially euphausiids affected the whole set of focal species.

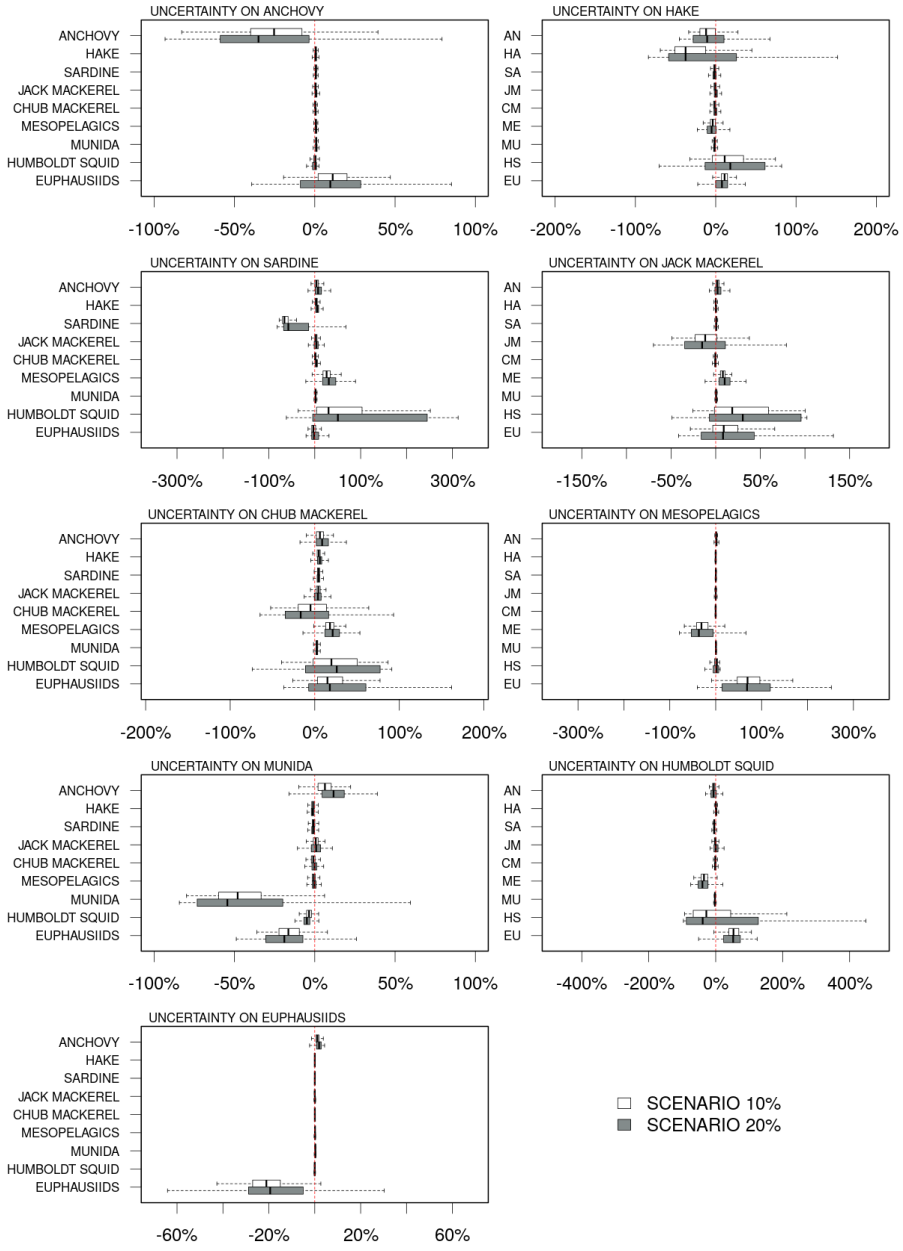

Figure 1. *Relative change (%) in biomass under uncertainty scenarios of 10% (white boxplots) and 20% (grey boxplots). Each panel shows the resulting uncertainty on the biomass of one focal*

species of the NHCE OSMOSE model (x-axis) due to the uncertainty in the nine species' parameters (y-axis). Changes were compared to a baseline simulation (dashed red line).

3.2. Ecological indicators behaviour under uncertainty

Uncertainty in ecological indicators across scenarios (Fig. 2) showed that in general, for all evaluated indicators, scenario 20% produced a higher level of uncertainty than scenario 10%. Though only a limited range of uncertainty was explored, this suggests that the level of uncertainty in the model outputs depends on the level of uncertainty in the inputs. However, the level of uncertainty captured in model outputs was heterogeneous across all evaluated indicators. In addition, the impact of the uncertainty of individual species biomass differed greatly depending on the indicators considered, except for the mean length and the slope of the size spectrum which showed similar responses to individual species uncertainties (Fig. 2). Of the nine indicators, the mean trophic level and the marine trophic index had the lowest levels of uncertainty (with CVs less than 3%), dampening the level of uncertainty in input parameters. They are followed by the slope of the size spectrum (with CVs less than 4% in absolute value), mean lifespan (with CVs less than 10%). Mean length uncertainty was very low (CV<1%) in response to all species' uncertainties, except in response to euphausiids' uncertainty (but still dampening the input uncertainty). Biomass over yield had a relatively high uncertainty (with a CV up to 70% in response to 20% anchovy uncertainty). Finally, LFI_{20}, LFI_{30}, and LFI_{40} were the three indicators that displayed the highest levels of uncertainty, amplifying the uncertainty in input parameters (with CVs up to 80%).

Since the model produces time-varying outputs, we analyzed the propagation of uncertainty through time and explored whether looking at mean uncertainty only could mask some important temporal features. For this, we focused on the indicator with the highest uncertainty, the LFI_{40}, under the 20% scenario (Fig. 3). We find that the uncertainty in LFI_{40} is reduced during the El Niño event (the light grey area in Fig. 3, between the years 1997-1998), whereas it sharply increases after the El Niño event. This pattern is particularly marked for simulations that implement parameter uncertainty in low trophic levels species such as mesopelagics, anchovy, euphausiids, but also other higher trophic level species such as humboldt squid. The case of

munida is particular, because this species was observed in the NHCE after 1996 and was introduced in the model application after the El Niño event.

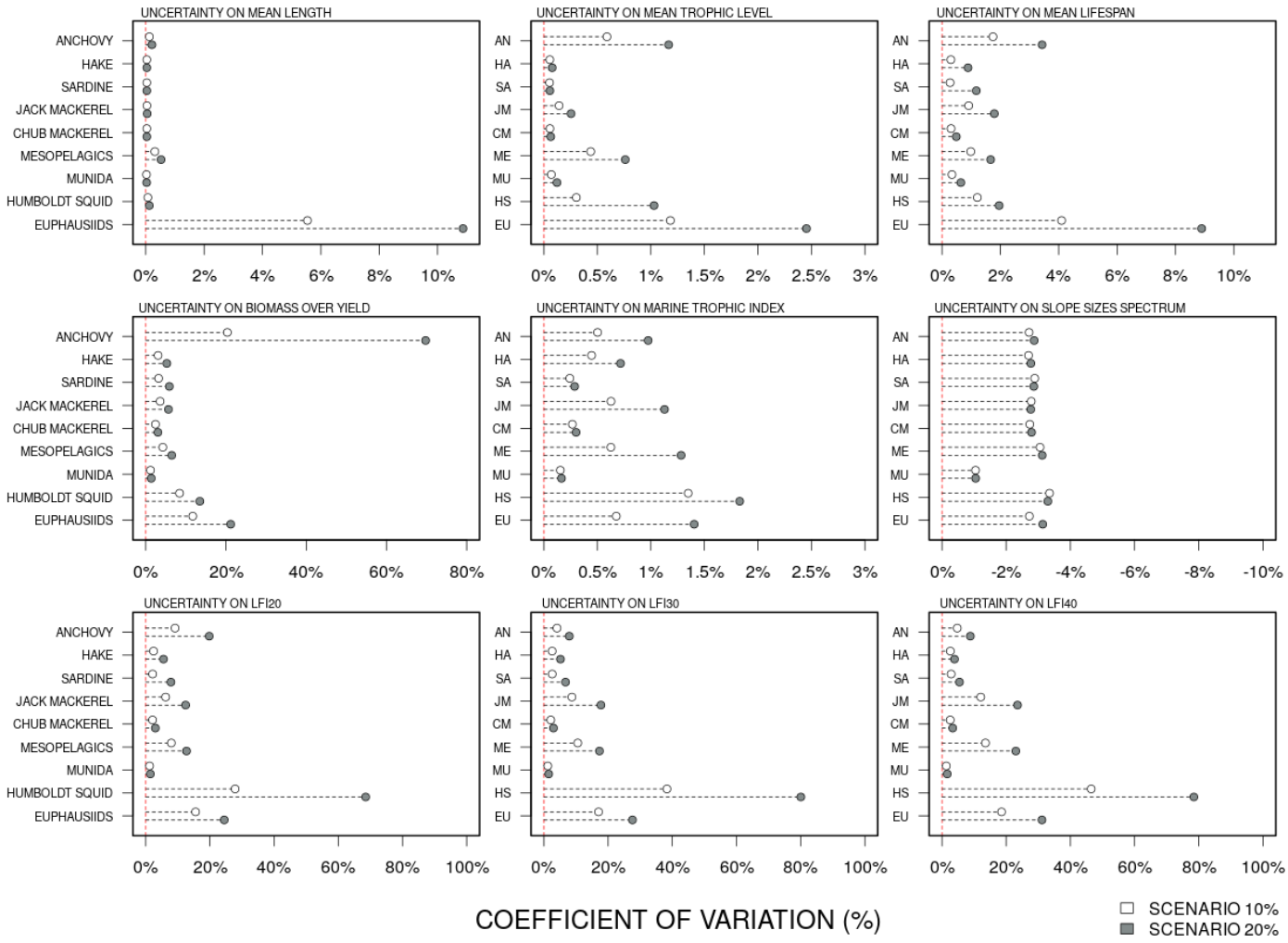

Figure 2. *Coefficient of variation of ecological indicators in output of NHCE OSMOSE (x-axis) under uncertainty scenarios of 10% (white points) and 20% (grey points). Each panel shows the uncertainty of an indicator due to the uncertainty in the 9 species' parameters (y-axis).*

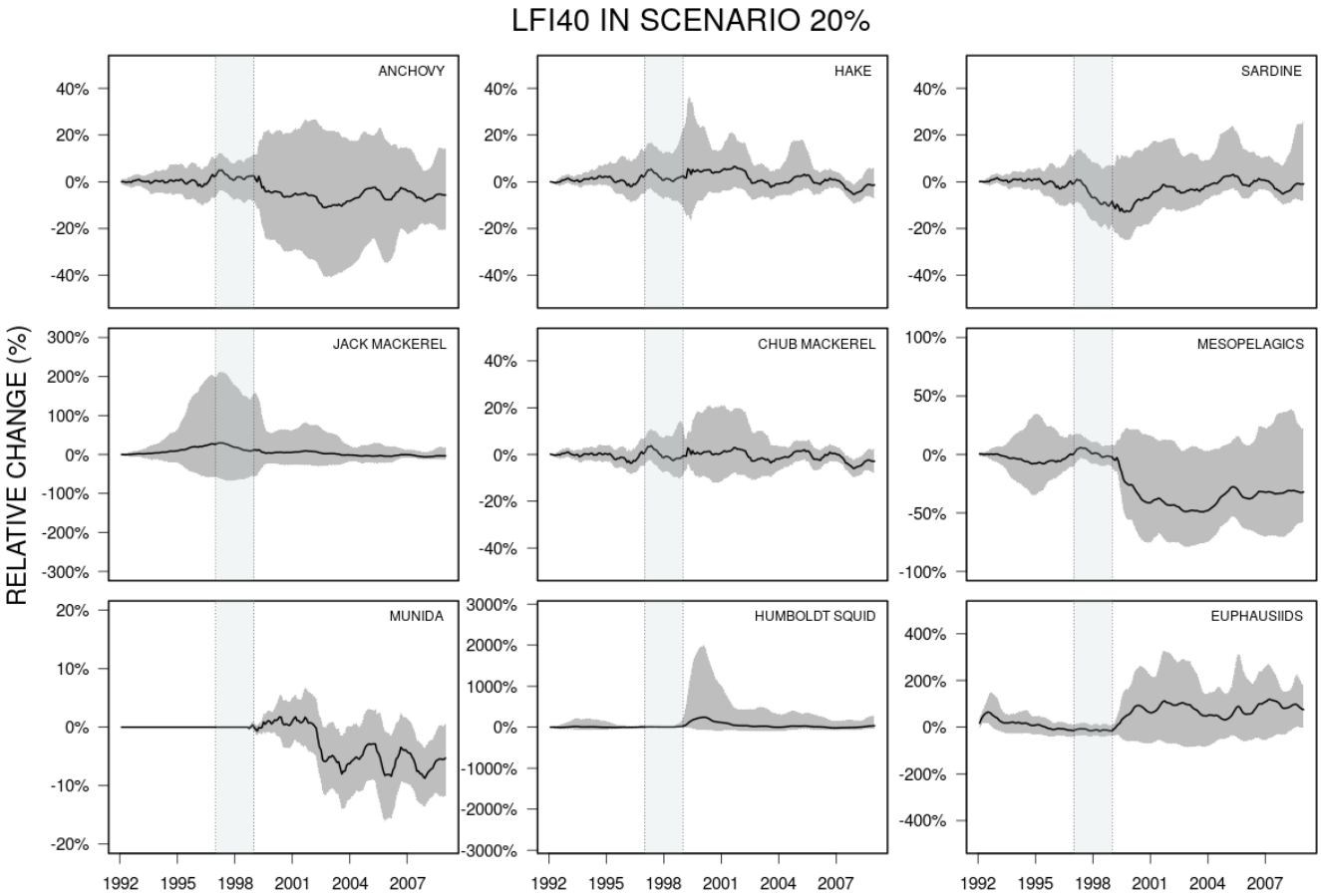

Figure 3. *Projected uncertainty in the LFI40 indicator under the scenario of 20% uncertainty in species' parameters. The dark grey area represents the uncertainty between the lower quantile (2.5%) and upper quantile (97.5%), and the median of the distribution is in black. The vertical area (in light grey) represents the duration of the El Niño event in the northern Humboldt Current ecosystem (Dewitte et al., 2012).*

We also compared the relative change of species biomass under the 20% uncertainty scenario before and after the El Niño event (Fig. 4). The results show that the uncertainty in species biomass is larger after El Niño than before; this was observed for almost all species except for euphausiids (and obviously munida which was included in the modelled system after El Niño). We also analyzed the behaviour of some output ecological indicators before and after the El Niño event (Fig. 5). The temporal changes of these indicators in response to species' parameter uncertainty are heterogeneous, but overall with more uncertainty after the El Niño event.

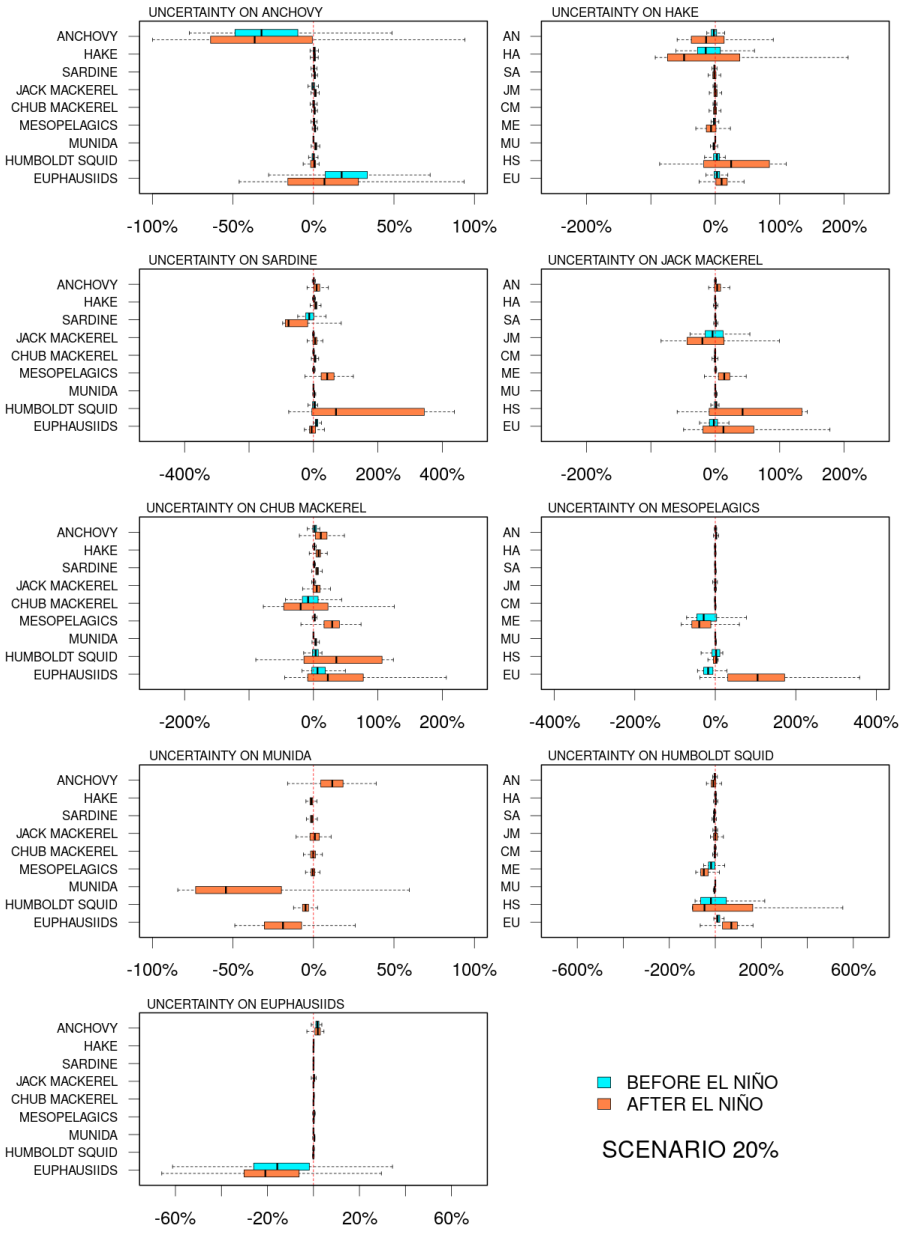

Figure 4. *Relative change in species biomass under the scenario of 20% uncertainty. Each panel shows the uncertainty on the nine focal species' biomass of the NHCE OSMOSE model due to the uncertainty in the parameters of one species. The relative changes were compared to a baseline simulation (dashed red line). In turquoise the uncertainty before the El Niño event (from January 1992 up to December 1996) and in orange, the uncertainty during and after the El Niño (from January 1997 until December 2008).*

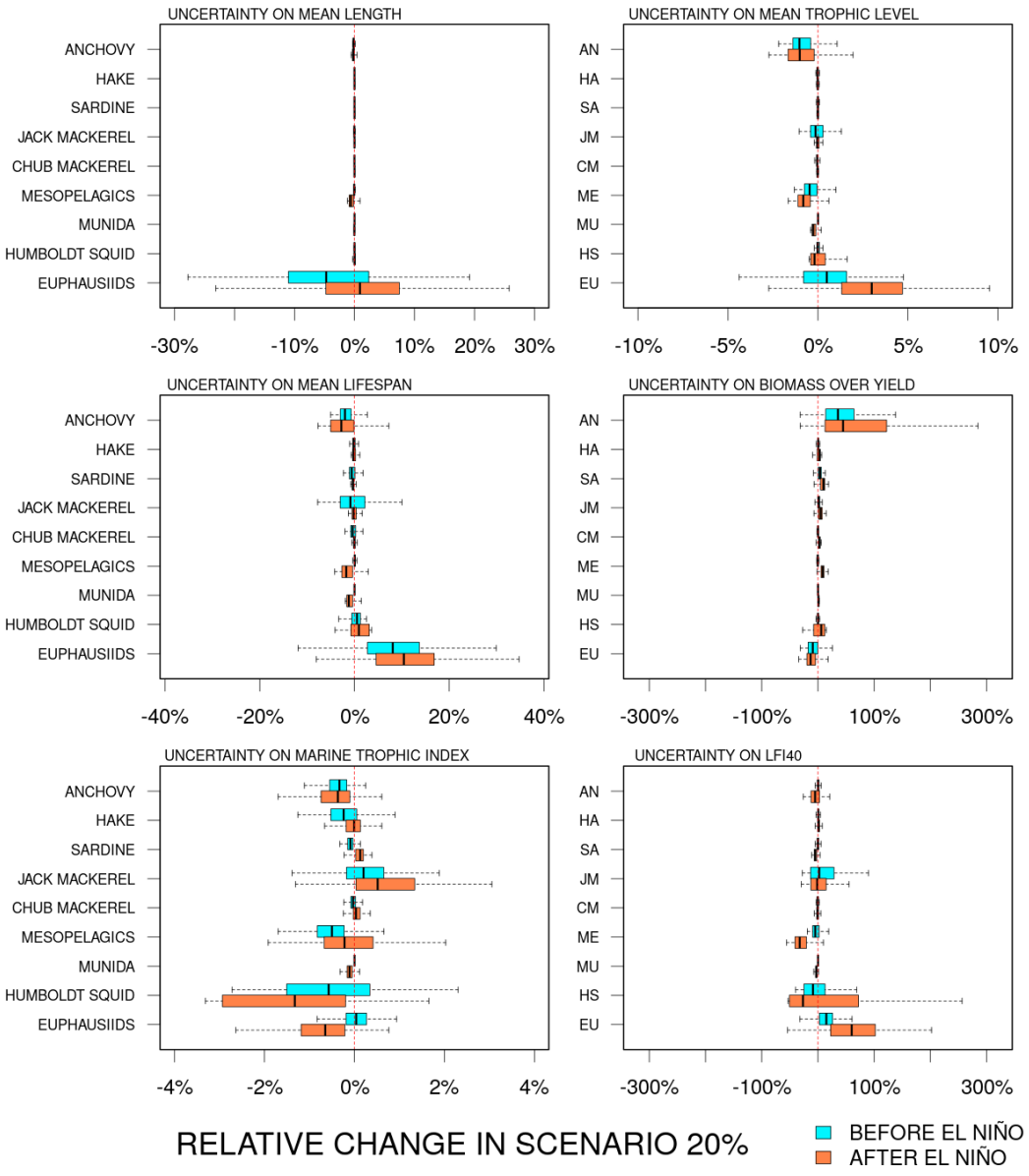

Figure 5. *Relative change in ecological indicators under the scenario of 20% uncertainty. Each panel shows the uncertainty on one of the indicators due to the uncertainty in species' parameters. The relative changes were compared to a baseline simulation (dashed red line). In turquoise, the uncertainty before the El Niño event (from January 1992 up to December 1996) and in orange, the uncertainty during and after the El Niño (from January 1997 until December 2008).*

4. Discussion

The simulation experiments allowed us to evaluate the uncertainty propagation from one species to the modelled ecosystem. The results showed that as an outcome of the complex interactions in the NHCE OSMOSE model, the uncertainty from one species could impact the rest of the food web. In particular, the uncertainty in euphausiids' and Humboldt squid's parameters substantially affected the entire food web, suggesting that both species have an essential role in the modelled ecosystem. Thus, these two species need to be studied further to ensure a realistic representation of their dynamics in the modelled system, and the information used to parameterize them must rely on high data quality. However, poor data quality and the resulting poor representation in the model could also explain the large contribution of some species to the uncertainty in the model output. Although we lacked the evidence to fully understand the roles of euphausiids and Humboldt squid in the NHCE, it should be noted that these two species lacked reliable time series of abundance indices (Oliveros-Ramos et al., 2017).

The relatively low contribution of the uncertainty in anchovy's input parameters to the other species' biomass was unexpected (Fig. 1) since the anchovy is considered a key species in the NHCE (Chavez et al., 2008; Bertrand et al., 2008; Guttierrez et al., 2012). In addition, of the nine species modelled, only two supplied uncertainty to anchovy (euphausiids and the anchovy itself). A possible explanation is that in the NHCE OSMOSE model, anchovy predators have not been represented explicitly (e.g., birds and other predators, Bertrand et al., 2012) but only as mortality parameters. We thus highly recommend improving the representation of anchovy in the NHCE, and one way to do so would be to represent the dynamics and the life cycle of their main predators explicitly. Likewise, the dynamics of hake, a demersal species, could be improved in the model by better representing the dynamics of their main prey euphausiids and munida (Orrego and Mendo, 2012; Castillo, 2012). Munida was modelled through a biomass immigration flux in the NHCE OSMOSE model, which probably simplifies too much the complex life history dynamics of this species for which small individuals have a pelagic behaviour and adults a demersal one (Guttierrez, 2002).

Unlike what was done in this study, it would have been preferable to use data to construct probability distributions for each model input parameter and design the uncertainty analyses based on these distributions (e.g., in Yegnan et al., 2002; Brown et al., 2015; Engström et al.,

2016). Unfortunately, given the high number of parameters that characterizes complex models of natural systems, there is not often enough observations, experiments and available data to construct probability distributions for each parameter. The use of fixed ranges of variability for input parameters is an option that is used in such cases of relatively poor data situations. However, a minimum level of information is needed to select appropriate parameter ranges and avoid using arbitrary ones that could underestimate or overestimate the level of uncertainty in model inputs (Saltelli et al., 2004; Lujan et al. 2022. arXiv.). Here, we used arbitrary ranges as a first step in the UA of the NHCE OSMOSE model, with range values commonly adopted in other UA of ecosystem models (Lehuta et al. (2010), Ciric et al. (2012), Dantec-Nedelec et al. (2017), Zheng et al. (2012)). By using two different ranges of parameter variability (uncertainty scenarios of 10% and 20%), we were able to confirm that larger levels of uncertainty in the NHCE OSMOSE model inputs produced a larger level of uncertainty in the model outputs.

The use of ecological indicators allows monitoring the state of ecosystems under fisheries impact and climate change, supporting decision making for an EAF (Fulton et al., 2005b). In this regard, several studies using the OSMOSE modelling platform have conducted indicators analyses to test their performance and usefulness (Halouani et al., 2019; Fu et al., 2019; Shin et al., 2018) but never considered the effect of parametric uncertainty. We found that some indicators were able to strongly mitigate the uncertainty associated with species' parameters, with especially three indicators emerging as having interesting properties: the slope of the size spectrum, the marine trophic index and the mean trophic level. The |CV| of the slope of the size spectrum was very homogeneous across all 10% and 20% uncertainty scenarios, and always less than 4%, whatever the species considered in the uncertainty scenarios. The responses of the marine trophic index and the mean trophic level were more heterogeneous, depending on which species' uncertainty was considered, but their CVs were always very low, less than 2% and 2.5%, respectively. However, further studies using more uncertainty scenarios are needed to explain these results and test their robustness. On the other hand, the large fish indicators (LFI_{20}, LFI_{30}, LFI_{40}) had relatively high levels of uncertainties in output of the NHCE OSMOSE model. This result should be taken into account when using this indicator in the model projections, especially as previous studies have shown that the LFI was performing well for detecting changes in ecosystem structure (Halouani et al., 2019; Moullec, 2019b).

An additional complexity for analyzing uncertainty in complex models is when these tools are dynamic, potentially propagating uncertainty over time. This temporal dynamic of uncertainty should be considered more often than is currently the case, when using models for both hindcast and forecast simulations. In our uncertainty experiments, the level of uncertainty rose strongly for most species after the El Niño (especially for anchovy, mesopelagics, and euphausiids). This event had a massive impact on the NHCS, disrupting the ecosystem's structure (Chavez et al., 2002; Ñiquen and Bouchon 2004; Gutierrez et al., 2011) and adding significant variability to the system. This was also detected in our results, with significant changes in uncertainty in the model's output before and after El Niño. We thus emphasize that uncertainty studies using time-averaged model outputs could mask important features for characterizing model uncertainty.

This work involved some technical challenges that we report here. One major impediment was the computational cost of running UA of complex models since these tools are characterized by many inputs (e.g., parameters, initial and forcing conditions) and outputs (multiple simulated variables). To deal with this, we decided to focus the UA on the parameters associated with focal species, but even with this restriction, we were already using over 200 parameters. Furthermore, we limited the study to the analysis of only ten ecological indicators among the long list of model outputs. UA, especially of complex models, also have high computational requirements in terms of simulation time and disk space. The results obtained for this work occupied a memory space around 7 Tb, which was very limiting for working with a larger number of uncertainty scenarios. Moreover, each species scenario required about 264 hours (11 days approximately), thus the 18 UA experiments (10% and 20% uncertainty scenarios for each of the 9 species) required a total of 198 days of simulation (about 6.6 months). To shorten the simulation time, we could run the simulations in parallel on the DATARMOR supercomputer, a high-performance computing system (HPC, with the possibility to use up to 56 cores depending on cluster availability).

The use of Morris in combination with Monte Carlo reduced the computational cost of our analyses. However, complementary studies related to Morris parameters (e.g., number of trajectories and levels, see Bracis et al., (2020) and Morris et al., (2014)) are needed. On the other hand, future studies could explore other alternative methods (Marino et al. (2008)). For

example, we could use the quasi-random sequence method (Jansen, 1999; Sobol et al., 2007), a Monte Carlo resampling procedure implemented in the R sensitivity package (Iooss et al., 2021) and with a computational cost of N(k+2) simulations, where N is the sample size (i.e. the parameter distribution is divided into N intervals that are sampled) and k is the number of parameters included in the analysis. Another possible method is the Latin hypercube sampling (Mckay et al., 1979), which is also considered low cost (computational cost of N(k+1) simulations).

In this paper, we studied the effect of parameter uncertainty on the outputs of the NHCE OSMOSE model. However, we highlight the importance of analyzing other sources of uncertainty, such as those due to initial conditions and model forcing (Payne et al., 2015; Cheung et al., 2016; Rounsevell et al., 2021). Finally, we call for an increase in interdisciplinary work, which could synergize the efforts and experience from other scientific fields in the study of uncertainty in complex models.

5. Acknowledgments

The authors acknowledge the Pôle de Calcul et de Données Marines (PCDM, https://wwz.ifremer.fr/pcdm/Equipement) of Ifremer for providing DATARMOR storage and computational resources. CL was supported by an Peruvian individual doctoral research grant from the "Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica" (contract n° 134-2016-FONDECYT). YJS acknowledges funding support from the Biodiversa and Belmont Forum project SOMBEE (BiodivScen ERA-Net COFUND programme, ANR contract n°ANR-18-EBI4-0003-01), and the Pew marine fellows programme.

6. References

1. Aumont, O., & Bopp, L. (2006). Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles*, *20*(2).
2. Bakun, A., & Weeks, S. J. (2008). The marine ecosystem off Peru: What are the secrets of its fishery productivity and what might its future hold?. *Progress in Oceanography*, *79*(2-4), 290-299.
3. Bertrand, A., Gerlotto, F., Bertrand, S., Gutiérrez, M., Alza, L., Chipollini, A., ... & Chavez, F. (2008). Schooling behaviour and environmental forcing in relation to anchoveta distribution: An analysis across multiple spatial scales. *Progress in Oceanography*, *79*(2-4), 264-277.
4. Bertrand, S., Joo, R., Arbulu Smet, C., Tremblay, Y., Barbraud, C., & Weimerskirch, H. (2012). Local depletion by a fishery can affect seabird foraging. *Journal of Applied Ecology*, *49*(5), 1168-1177.
5. Bracis, C., Lehuta, S., Savina-Rolland, M., Travers-Trolet, M., & Girardin, R. (2020). Improving confidence in complex ecosystem models: The sensitivity analysis of an Atlantis ecosystem model. *Ecological Modelling*, *431*, 109133.
6. Brown, C., Brown, E., Murray-Rust, D., Cojocaru, G., Savin, C., & Rounsevell, M. (2015). Analysing uncertainties in climate change impact assessment across sectors and scenarios. *Climatic Change*, *128*(3), 293-306.

7. Caddy, J.F., and R.C. Griffiths. (1995). Living marine resources and their sustainable development. Some environmental and institutional perspectives. FAO Fisheries Technical Paper, Vol. 353, 174pp.
8. Cariboni, J., Gatelli, D., Liska, R., & Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. Ecological modelling, 203(1-2), 167-182.
9. Castillo, D. (2019). Dieta de la merluza peruana (Merluccius gayi peruanus Ginsburg, 1954) en otoño 2012. *Boletin Instituto del Mar del Perú*, *34*(2), 355-368.
10. Chavez, F. P., Pennington, J. T., Castro, C. G., Ryan, J. P., Michisaki, R. P., Schlining, B., ... & Collins, C. A. (2002). Biological and chemical consequences of the 1997–1998 El Niño in central California waters. *Progress in Oceanography*, *54*(1-4), 205-232.
11. Chavez, F. P., Ryan, J., Lluch-Cota, S. E., & Ñiquen, M. (2003). From anchovies to sardines and back: multidecadal change in the Pacific Ocean. *science*, *299*(5604), 217-221.
12. Chavez, F. P., Bertrand, A., Guevara Carrasco, R., Soler, P., & Csirke, J. (2008). The northern Humboldt Current System: Brief history, present status and a view towards the future. *Progress in Oceanography*, *79*(special issue 2-4), 95-105.
13. Christensen, V., & Walters, C.J. (2004). Ecopath with Ecosim: methods, capabilities and limitations. Ecological Modelling 172, 109–139.
14. Christensen, V., Walters, C.J. & Pauly, D. (2005). Ecopath with Ecosim: A User's Guide. Fisheries Centre, University of British Columbia, Vancouver. 154 pp.
15. Ciric, C., Ciffroy, P., & Charles, S. (2012). Use of sensitivity analysis to identify influential and non-influential parameters within an aquatic ecosystem model. *Ecological Modelling*, *246*, 119-130.
16. Dantec‐Nédélec, S., Ottlé, C., Wang, T., Guglielmo, F., Maignan, F., Delbart, N., ... & Jouzel, J. (2017). Testing the capability of ORCHIDEE land surface model to simulate Arctic ecosystems: Sensitivity analysis and site‐level model calibration. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1212-1230.
17. Dewitte, B., Vazquez-Cuervo, J., Goubanova, K., Illig, S., Takahashi, K., Cambon, G., ... & Ortlieb, L. (2012). Change in El Niño flavours over 1958–2008: Implications for the long-term trend of the upwelling off Peru. *Deep Sea Research Part II: Topical Studies in Oceanography*, *77*, 143-156.
18. Duarte, C. M. (2000). Marine biodiversity and ecosystem services: an elusive link. Journal of experimental marine Biology and Ecology, 250(1-2), 117-131.
19. Echevin, V., Aumont, O., Ledesma, J., & Flores, G. (2008). The seasonal cycle of surface chlorophyll in the Peruvian upwelling system: A modelling study. *Progress in Oceanography*, *79*(2-4), 167-176.
20. Engström, K., Olin, S., Rounsevell, M. D., Brogaard, S., Van Vuuren, D. P., Alexander, P., ... & Arneth, A. (2016). Assessing uncertainties in global cropland futures using a conditional probabilistic modelling framework. *Earth System Dynamics*, *7*(4), 893-915.
21. FAO. (2003). World Summit on Sustainable Development 2002 and its implications for fisheries. Url: https://www.fao.org/3/y8294e/y8294e.htm
22. Ferrier, Simon, Karachepone N Ninan, Paul Leadley, Rob Alkemade, Lilibeth A Acosta, H Resit Akçakaya, Lluis Brotons, et al. (2016). "IPBES (2016): The Methodological Assessment Report on Scenarios and Models of Biodiversity and Ecosystem Services." Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services. Bonn, Germany, 348.
23. Fu, C., Perry, R. I., Shin, Y. J., Schweigert, J., & Liu, H. (2013). An ecosystem modelling framework for incorporating climate regime shifts into fisheries management. *Progress in Oceanography*, *115*, 53-64.
24. Fu, C., Xu, Y., Bundy, A., Grüss, A., Coll, M., Heymans, J. J., ... & Shin, Y. J. (2019). Making ecological indicators management ready: assessing the specificity, sensitivity, and threshold response of ecological indicators. *Ecological Indicators*, *105*, 16-28.
25. Fulton, E.A., Fuller, M., Smith, A.D.M., & Punt, A.E. (2005a). Ecological Indicators of the Ecosystem Effects of Fishing: Final Report. Australian Fisheries Management Authority Report, R99/1546. 239 pp.
26. Fulton, E. A., Smith, A. D., & Punt, A. E. (2005b). Which ecological indicators can robustly detect effects of fishing?. *ICES Journal of Marine Science*, *62*(3), 540-551.
27. Fulton, E. A. (2010). Approaches to end-to-end ecosystem models. *Journal of Marine Systems*, *81*(1-2), 171-183.
28. Garcia, S. M. The ecosystem approach to fisheries: issues, terminology, principles, institutional foundations, implementation, and outlook. Food & Agriculture Org., 2003.
29. Garcia, S. M., & Cochrane, K. L. (2005). Ecosystem approach to fisheries: a review of implementation guidelines. ICES Journal of Marine Science, 62(3), 311-318.
30. Greenstreet, S. P., Rogers, S. I., Rice, J. C., Piet, G. J., Guirey, E. J., Fraser, H. M., & Fryer, R. J. (2011). Development of the EcoQO for the North Sea fish community. *ICES Journal of Marine Science*, *68*(1), 1-11.

31. Gutiérrez, M. (2002). A technical and scientific approach about needs of new sustainable fisheries. In *Proceedings of Seoul Ocean Seminar. 1st APEC Ocean-related Ministerial Meeting. Korea Maritime Institute* (pp. 75-90).
32. Gutiérrez, D., Bertrand, A., Wosnitza-mendo, C., Dewitte, B., Purca, S., Peña, C., ... & Grados, C. (2011). Climate change sensitivity of the Peruvian upwelling system and ecological implications. *Rev. Peru. Geo-Atmosférica*, *3*, 1-26.
33. Gutiérrez, M., Castillo, R., Segura, M., Peraltilla, S., & Flores, M. (2012). Trends in spatio-temporal distribution of Peruvian anchovy and other small pelagic fish biomass from 1966-2009 v. *Latin American Journal of Aquatic Research*, *40*(3), 633-648.
34. Gutiérrez, D., Akester, M., & Naranjo, L. (2016). Productivity and sustainable management of the Humboldt Current large marine ecosystem under climate change. *Environmental Development*, *17*, 126-144.
35. Hansen, C., Drinkwater, K. F., Jähkel, A., Fulton, E. A., Gorton, R., & Skern-Mauritzen, M. (2019). Sensitivity of the Norwegian and Barents Sea Atlantis end-to-end ecosystem model to parameter perturbations of key species. *PLoS One*, *14*(2), e0210419.
36. Halouani, G., Lasram, F. B. R., Shin, Y. J., Velez, L., Verley, P., Hattab, T., ... & Romdhane, M. S. (2016). Modelling food web structure using an end-to-end approach in the coastal ecosystem of the Gulf of Gabes (Tunisia). *Ecological Modelling*, *339*, 45-57.
37. Halouani, G., Le Loc'h, F., Shin, Y. J., Velez, L., Hattab, T., Romdhane, M. S., & Lasram, F. B. R. (2019). An end-to-end model to evaluate the sensitivity of ecosystem indicators to track fishing impacts. *Ecological Indicators*, *98*, 121-130.
38. Iooss, B., Da Veiga, S., Janon, A., Pujol, G., Iooss, M. B., Rcpp, L., ... & DiceDesign, D. (2021). Package 'sensitivity'.
39. Jennings, S. (2005). Indicators to support an ecosystem approach to fisheries. Fish and Fisheries, 6(3), 212-232.
40. Jansen, M. J. (1999). Analysis of variance designs for model output. *Computer Physics Communications*, *117*(1-2), 35-43.
41. Lehuta, S., Mahévas, S., Petitgas, P., & Pelletier, D. (2010). Combining sensitivity and uncertainty analysis to evaluate the impact of management measures with ISIS–Fish: marine protected areas for the Bay of Biscay anchovy (Engraulis encrasicolus) fishery. *ICES Journal of Marine Science*, *67*(5), 1063-1075.
42. Lehuta, S., Girardin, R., Mahévas, S., Travers-Trolet, M., & Vermard, Y. (2016). Reconciling complex system models and fisheries advice: practical examples and leads. *Aquatic Living Resources*, *29*(2), 208.
43. Link, J. S., Ihde, T. F., Harvey, C. J., Gaichas, S. K., Field, J. C., Brodziak, J. K. T., ... & Peterman, R. M. (2012). Dealing with uncertainty in ecosystem models: the paradox of use for living marine resource management. *Progress in Oceanography*, *102*, 102-114.
44. Marzloff, M., Shin, Y. J., Tam, J., Travers, M., & Bertrand, A. (2009). Trophic structure of the Peruvian marine ecosystem in 2000–2006: insights on the effects of management scenarios for the hake fishery using the IBM trophic model Osmose. *Journal of Marine Systems*, *75*(1-2), 290-304.
45. Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of theoretical biology*, *254*(1), 178-196.
46. McGregor, V. L., Fulton, E. A., & Dunn, M. R. (2020). Addressing initialisation uncertainty for end-to-end ecosystem models: application to the Chatham Rise Atlantis model. *PeerJ*, *8*, e9254.
47. Mckay, M.D., Beckman, R.J., Conover, W.J. (1979). Comparison of 3 methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.
48. Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161-174.
49. Morris, D. J., Speirs, D. C., Cameron, A. I., & Heath, M. R. (2014). Global sensitivity analysis of an end-to-end marine ecosystem model of the North Sea: Factors affecting the biomass of fish and benthos. *Ecological Modelling*, *273*, 251-263.
50. Moullec, F., Velez, L., Verley, P., Barrier, N., Ulses, C., Carbonara, P., ... & Shin, Y. J. (2019a). Capturing the big picture of Mediterranean marine biodiversity with an end-to-end model of climate and fishing impacts. *Progress in Oceanography*, *178*, 102179.
51. Moullec, F. (2019b). *Impacts du changement global sur la biodiversité en mer Méditerranée: une approche par modélisation End-to-End* (Doctoral dissertation, Université Montpellier).
52. Ñiquen, M., & Bouchon, M. (2004). Impact of El Niño events on pelagic fisheries in Peruvian waters. *Deep sea research part II: topical studies in oceanography*, *51*(6-9), 563-574.
53. Oliveros-Ramos, R. (2014). End-to-End Modelling for an Ecosystem Approach to Fisheries in the Northern Humboldt Current Ecosystem. Ph.D. Thesis, France: University of Montpellier 2.

54. Oliveros-Ramos, R., Verley, P., Echevin, V., & Shin, Y. J. (2017). A sequential approach to calibrate ecosystem models with multiple time series data. *Progress in Oceanography*, *151*, 227-244.
55. Orrego, H., & Mendo, J. (2012). Variación interanual de la dieta de la merluza Merluccius gayi peruanus (GUITCHENOT) en la costa peruana. *Ecología Aplicada*, *11*(2), 103-116.
56. Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., ... & Paula, J. R. (2016). Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES Journal of Marine Science*, *73*(5), 1272-1282.
57. Pauly, D., & Watson, R. (2005). Background and interpretation of the 'Marine Trophic Index'as a measure of biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1454), 415-423.
58. Plagányi, É. E. (2007). Models for an ecosystem approach to fisheries.
59. R Core Team, 2019. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
60. Reed, J., Shannon, L., Velez, L., Akoglu, E., Bundy, A., Coll, M., ... & Shin, Y. J. (2017). Ecosystem indicators accounting for variability in species' trophic levels. *ICES Journal of Marine Science*, *74*(1), 158-169.
61. Rice, J., & Gislason, H. (1996). Patterns of change in the size spectra of numbers and diversity of the North Sea fish assemblage, as reflected in surveys and models. *ICES Journal of Marine Science*, *53*(6), 1214-1225.
62. Rounsevell, M. D., Arneth, A., Brown, C., Cheung, W. W., Gimenez, O., Holman, I., ... & Shin, Y. J. (2021). Identifying uncertainties in scenarios and models of socio-ecological systems in support of decision-making. *One Earth*, *4*(7), 967-985.
63. Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models* (Vol. 1). New York: Wiley.
64. Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., et al. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental modelling & software*, *114*, 29-39.
65. Shannon L., Coll M., Bundy A., Gascuel D., Heymans J.J., Kleisner K., Lynam C.P., Piroddi C., Tam J., Travers-Trolet M., Shin Y.-J. 2014. Trophic level-based indicators to track fishing impacts across marine ecosystems. MEPS 512: 115-140.
66. Shin, Y.-J., & Cury, P. (2001). Exploring Fish Community Dynamics Through Size-Dependent Trophic Interactions Using a Spatialized Individual-Based Model. Aquatic Living Resources 14 (2): 65–80.
67. Shin, Y.-J., & Cury, P. (2004). Using an individual-based model of fish assemblages to study the response of size spectra to changes in fishing. Canadian Journal of Fisheries and Aquatic Sciences 61, 414–431.
68. Shin Y.-J., M.-J. Rochet, S. Jennings, J. Field, H. Gislason. (2005). Using size-based indicators to evaluate the ecosystem effects of fishing. ICES Journal of marine Science 62(3): 394-396.
69. Shin, Y. J., Shannon, L. J., Bundy, A., Coll, M., Aydin, K., Bez, N., ... & Cury, P. M. (2010). Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. Setting the scene. *ICES Journal of Marine Science*, *67*(4), 692-716.
70. Shin, Y. J., Houle, J. E., Akoglu, E., Blanchard, J. L., Bundy, A., Coll, M., ... & Velez, L. (2018). The specificity of marine ecological indicators to fishing in the face of environmental change: a multi-model evaluation. *Ecological Indicators*, *89*, 317-326.
71. Shin YJ, Midgley GF, Archer E, Arneth A, Barnes DKA, Chan L, Hashimoto S, Hoegh-Guldberg O, Insarov G, Leadley P, Levin LA, Ngo HT, Pandit R, Pires APF, Pörtner HO, Rogers AD, Scholes RJ, Settele J, Smith P. In press. Actions to halt biodiversity loss generally benefit the climate. Global Change Biology.
72. Snelgrove, P. V. (1999). Getting to the bottom of marine biodiversity: sedimentary habitats: ocean bottoms are the most widespread habitat on earth and support high biodiversity and key ecosystem services. BioScience, 49(2), 129-138.
73. Sobol IM, Tarantola S, Gatelli D, Kucherenko SS, Mauntz W (2007) Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliab Eng Syst Saf 92(7):957–960.
74. Specka, X., Nendel, C., & Wieland, R. (2015). Analysing the parameter sensitivity of the agro-ecosystem model MONICA for different crops. *European Journal of Agronomy*, *71*, 73-87.
75. Steenbeek, J., Buszowski, J., Chagaris, D., Christensen, V., Coll, M., Fulton, E. A., et al. (2021). Making spatial-temporal marine ecosystem modelling better–A perspective. *Environmental Modelling & Software*, *145*, 105209.
76. Tam, J., Taylor, M.H., Blaskovic, V., Espinoza, P., Ballón, R.M., Díaz, E., Wosnitza-Mendo, C., Argüelles, J., Purca, S., Ayón, P., Quipuzcoa, L., Gutiérrez, D., Goya, E., Ochoa, N., Wolff, M. (2008). Trophic modeling of the Northern Humboldt Ecosystem, Part I: comparing trophic linkages under La Niña and El Niño conditions. Progress in Oceanography 79, 352–365.

77. Taylor, M.H., Tam, J., Blaskovic, V., Espinoza, P., Ballón, R.M., Wosnitza-Mendo, C., Argüelles, J., Díaz, E., Purca, S., Ochoa, N., Ayón, P., Goya, E., Gutiérrez, D., Quipuzcoa, L., Wolff, M. (2008). Trophic modeling of the Northern Humboldt Current Ecosystem, Part II: elucidating ecosystem dynamics from 1995–2004 with a focus on the impact of ENSO. Progress in Oceanography 79, 366–378.
78. Travers, M., Shin, Y. J., Jennings, S., Machu, E., Huggett, J. A., Field, J. G., & Cury, P. M. (2009). Two-way coupling versus one-way forcing of plankton and fish models to predict ecosystem changes in the Benguela. *Ecological modelling*, *220*(21), 3089-3099.
79. Travers-Trolet, M., Shin, Y. J., & Field, J. G. (2014). An end-to-end coupled model ROMS-N2P2Z2D2-OSMOSE of the southern Benguela foodweb: parameterisation, calibration and pattern-oriented validation. *African Journal of Marine Science*, *36*(1), 11-29.
80. Walters, C.J, Christensen, V., Pauly, D. (1997). Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments. Reviews in Fish Biology and Fisheries 7, 139–172.
81. Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., etal. (2006). Impacts of biodiversity loss on ocean ecosystem services. science, 314(5800), 787-790.
82. Xing, L., Chen, Y., Zhang, C., Li, B., Tanaka, K. R., Boenish, R., & Ren, Y. (2020). Evaluating impacts of imprecise parameters on the performance of an ecosystem model OSMOSE-JZB. *Ecological Modelling*, *419*, 108923.
83. Zheng, W., Shi, H., Fang, G., Hu, L., Peng, S., & Zhu, M. (2012). Global sensitivity analysis of a marine ecosystem dynamic model of the Sanggou Bay. *Ecological Modelling*, *247*, 83-94.

Appendix A. Model input parameters for the UA

Table A.1 *Reparametrization and baseline values of PCE OSMOSE model parameters used for the UA.*

Parameters type	*Anchovy (sp1)*	*Peruvian hake (sp2)*	*Sardine (sp3)*	*Jack mackerel (sp4)*	*Chub mackerel (sp5)*	*Mesopelagics (sp6)*	*Munida (sp7)*	*Humboldt squid (sp8)*	Euphausiids *(sp9)*
predation accessibility of prey to predator	Reparametrization: $p_{1,sp1} = A(sp_1, sp_2)$ $p_{2,sp1} = A(sp_1, sp_3)$ $p_{3,sp1} = A(sp_1, sp_4)$ $p_{4,sp1} = A(sp_1, sp_5)$ $p_{5,sp1} = A(sp_1, sp_6)$ $p_{6,sp1} = A(sp_1, sp_7)$ $p_{7,sp1} = A(sp_1, sp_8)$ $p_{8,sp1} = A(sp_1, sp_9)$ Baseline value (x): $p_{1,sp1} = 0.10$ $p_{2,sp1} = 0.90$ $p_{3,sp1} = 0.90$ $p_{4,sp1} = 0.90$ $p_{5,sp1} = 0.55$ $p_{6,sp1} = 0.05$ $p_{7,sp1} = 0.60$ $p_{8,sp1} = 0.60$	Reparametrization: $p_{1,sp2} = A(sp_2, sp_1)$ $p_{2,sp2} = A(sp_2, sp_3)$ $p_{3,sp2} = A(sp_2, sp_4)$ $p_{4,sp2} = A(sp_2, sp_5)$ $p_{5,sp2} = A(sp_2, sp_6)$ $p_{6,sp2} = A(sp_2, sp_7)$ $p_{7,sp2} = A(sp_2, sp_8)$ $p_{8,sp2} = A(sp_2, sp_9)$ Baseline value (x): $p_{1,sp2} = 0.20$ $p_{2,sp2} = 0.20$ $p_{3,sp2} = 0.15$ $p_{4,sp2} = 0.15$ $p_{5,sp2} = 0.45$ $p_{6,sp2} = 0.02$ $p_{7,sp2} = 0.65$ $p_{8,sp2} = 0.50$	Reparametrization: $p_{1,sp3} = A(sp_3, sp_1)$ $p_{2,sp3} = A(sp_3, sp_2)$ $p_{3,sp3} = A(sp_3, sp_4)$ $p_{4,sp3} = A(sp_3, sp_5)$ $p_{5,sp3} = A(sp_3, sp_6)$ $p_{6,sp3} = A(sp_3, sp_7)$ $p_{7,sp3} = A(sp_3, sp_8)$ $p_{8,sp3} = A(sp_3, sp_9)$ Baseline value (x): $p_{1,sp3} = 0.90$ $p_{2,sp3} = 0.15$ $p_{3,sp3} = 0.90$ $p_{4,sp3} = 0.90$ $p_{5,sp3} = 0.60$ $p_{6,sp3} = 0.05$ $p_{7,sp3} = 0.60$ $p_{8,sp3} = 0.60$	Reparametrization: $p_{1,sp4} = A(sp_4, sp_1)$ $p_{2,sp4} = A(sp_4, sp_2)$ $p_{3,sp4} = A(sp_4, sp_3)$ $p_{4,sp4} = A(sp_4, sp_5)$ $p_{5,sp4} = A(sp_4, sp_6)$ $p_{6,sp4} = A(sp_4, sp_7)$ $p_{7,sp4} = A(sp_4, sp_8)$ $p_{8,sp4} = A(sp_4, sp_9)$ Baseline value (x): $p_{1,sp4} = 0.90$ $p_{2,sp4} = 0.15$ $p_{3,sp4} = 0.90$ $p_{4,sp4} = 0.90$ $p_{5,sp4} = 0.60$ $p_{6,sp4} = 0.90$ $p_{7,sp4} = 0.60$ $p_{8,sp4} = 0.50$	Reparametrization: $p_{1,sp5} = A(sp_5, sp_1)$ $p_{2,sp5} = A(sp_5, sp_2)$ $p_{3,sp5} = A(sp_5, sp_3)$ $p_{4,sp5} = A(sp_5, sp_4)$ $p_{5,sp5} = A(sp_5, sp_6)$ $p_{6,sp5} = A(sp_5, sp_7)$ $p_{7,sp5} = A(sp_5, sp_8)$ $p_{8,sp5} = A(sp_5, sp_9)$ Baseline value (x): $p_{1,sp5} = 0.90$ $p_{2,sp5} = 0.15$ $p_{3,sp5} = 0.90$ $p_{4,sp5} = 0.90$ $p_{5,sp5} = 0.90$ $p_{6,sp5} = 0.90$ $p_{7,sp5} = 0.60$ $p_{8,sp5} = 0.60$	Reparametrization: $p_{1,sp6} = A(sp_6, sp_1)$ $p_{2,sp6} = A(sp_6, sp_2)$ $p_{3,sp6} = A(sp_6, sp_3)$ $p_{4,sp6} = A(sp_6, sp_4)$ $p_{5,sp6} = A(sp_6, sp_5)$ $p_{6,sp6} = A(sp_6, sp_7)$ $p_{7,sp6} = A(sp_6, sp_8)$ $p_{8,sp6} = A(sp_6, sp_9)$ Baseline value (x): $p_{1,sp6} = 0.55$ $p_{2,sp6} = 0.45$ $p_{3,sp6} = 0.60$ $p_{4,sp6} = 0.60$ $p_{5,sp6} = 0.90$ $p_{6,sp6} = 0.50$ $p_{7,sp6} = 0.95$ $p_{8,sp6} = 0.90$	Reparametrization: $p_{1,sp7} = A(sp_7, sp_1)$ $p_{2,sp7} = A(sp_7, sp_2)$ $p_{3,sp7} = A(sp_7, sp_3)$ $p_{4,sp7} = A(sp_7, sp_4)$ $p_{5,sp7} = A(sp_7, sp_5)$ $p_{6,sp7} = A(sp_7, sp_6)$ $p_{7,sp7} = A(sp_7, sp_8)$ $p_{8,sp7} = A(sp_7, sp_9)$ Baseline value (x): $p_{1,sp7} = 0.90$ $p_{2,sp7} = 0.025$ $p_{3,sp7} = 0.90$ $p_{4,sp7} = 0.90$ $p_{5,sp7} = 0.90$ $p_{6,sp7} = 0.50$ $p_{7,sp7} = 0.50$ $p_{8,sp7} = 0.55$	Reparametrization: $p_{1,sp8} = A(sp_8, sp_1)$ $p_{2,sp8} = A(sp_8, sp_2)$ $p_{3,sp8} = A(sp_8, sp_3)$ $p_{4,sp8} = A(sp_8, sp_4)$ $p_{5,sp8} = A(sp_8, sp_5)$ $p_{6,sp8} = A(sp_8, sp_6)$ $p_{7,sp8} = A(sp_8, sp_7)$ $p_{8,sp8} = A(sp_8, sp_9)$ Baseline value (x): $p_{1,sp8} = 0.60$ $p_{2,sp8} = 0.65$ $p_{3,sp8} = 0.60$ $p_{4,sp8} = 0.70$ $p_{5,sp8} = 0.60$ $p_{6,sp8} = 0.95$ $p_{7,sp8} = 0.50$ $p_{8,sp8} = 0.95$	Reparametrization: $p_{1,sp9} = A(sp_9, sp_1)$ $p_{2,sp9} = A(sp_9, sp_2)$ $p_{3,sp9} = A(sp_9, sp_3)$ $p_{4,sp9} = A(sp_9, sp_4)$ $p_{5,sp9} = A(sp_9, sp_5)$ $p_{6,sp9} = A(sp_9, sp_6)$ $p_{7,sp9} = A(sp_9, sp_7)$ $p_{8,sp9} = A(sp_9, sp_8)$ Baseline value (x): $p_{1,sp9} = 0.60$ $p_{2,sp9} = 0.50$ $p_{3,sp9} = 0.60$ $p_{4,sp9} = 0.60$ $p_{5,sp9} = 0.60$ $p_{6,sp9} = 0.90$ $p_{7,sp9} = 0.55$ $p_{8,sp9} = 0.95$
minimum predation:prey size ratio for each species stage (θ_{stage})	Reparametrization: $p_{9,sp1} = f(\theta_{stage1})$ $p_{10,sp1} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp1} = 0.0008$ $p_{10,sp1} = 0.0032$	Reparametrization: $p_{9,sp2} = f(\theta_{stage1})$ $p_{10,sp2} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp2} = 0.0127$ $p_{10,sp2} = 0.0127$	Reparametrization: $p_{9,sp3} = f(\theta_{stage1})$ $p_{10,sp3} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp3} = 0.0006$ $p_{10,sp3} = 0$	Reparametrization: $p_{9,sp4} = f(\theta_{stage1})$ $p_{10,sp4} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp4} = 0.0021$ $p_{10,sp4} = 0.0032$	Reparametrization: $p_{9,sp5} = f(\theta_{stage1})$ $p_{10,sp5} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp5} = 0.0021$ $p_{10,sp5} = 0.0032$	Reparametrization: $p_{9,sp6} = f(\theta_{stage1})$ Baseline value (x): $p_{9,sp6} = 0.0064$	Reparametrization: $p_{9,sp7} = f(\theta_{stage1})$ Baseline value (x): $p_{9,sp7} = 0.0042$	Reparametrization: $p_{9,sp8} = f(\theta_{stage1})$ $p_{10,sp8} = f(\theta_{stage2})$ $p_{11,sp8} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp8} = 0.0182$ $p_{10,sp8} = 0.0116$ $p_{11,sp8} = 0.0091$	Reparametrization: $p_{9,sp9} = f(\theta_{stage1})$ $p_{10,sp9} = f(\theta_{stage2})$ Baseline value (x): $p_{9,sp9} = 0.0002$ $p_{10,sp9} = 0.0003$

maximum predation:prey size ratio for each species stage (α_{stage})	Reparametrization: $p_{11,sp1} = f(\alpha_{stage1})$ $p_{12,sp1} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp1} = 0.0784$ $p_{12,sp1} = 0.1020$	Reparametrization: $p_{11,sp2} = f(\alpha_{stage1})$ $p_{12,sp2} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp2} = 0.1921$ $p_{12,sp2} = 0.2295$	Reparametrization: $p_{11,sp3} = f(\alpha_{stage1})$ $p_{12,sp3} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp3} = 0.0248$ $p_{12,sp3} = 0.0041$	Reparametrization: $p_{11,sp4} = f(\alpha_{stage1})$ $p_{12,sp4} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp4} = 0.0297$ $p_{12,sp4} = 0.0392$	Reparametrization: $p_{11,sp5} = f(\alpha_{stage1})$ $p_{12,sp5} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp5} = 0.0297$ $p_{12,sp5} = 0.0392$	Reparametrization: $p_{10,sp6} = f(\alpha_{stage1})$ Baseline value (x): $p_{10,sp6} = 0.1708$	Reparametrization: $p_{10,sp7} = f(\alpha_{stage1})$ Baseline value (x): $p_{10,sp7} = 0.2909$	Reparametrization: $p_{12,sp8} = f(\alpha_{stage1})$ $p_{13,sp8} = f(\alpha_{stage2})$ $p_{14,sp8} = f(\alpha_{stage2})$ Baseline value (x): $p_{12,sp8} = 0.2241$ $p_{13,sp8} = 0.2836$ $p_{14,sp8} = 0.4909$	Reparametrization: $p_{11,sp9} = f(\alpha_{stage1})$ $p_{12,sp9} = f(\alpha_{stage2})$ Baseline value (x): $p_{11,sp9} = 0.0422$ $p_{12,sp9} = 0.0631$
predation:prey size threshold (s_{thr} in cm)	Reparametrization: $p_{13,sp1} = f(s_{thr})$ Baseline value (x): $p_{13,sp1} = 0.5128$	Reparametrization: $p_{13,sp2} = f(s_{thr})$ Baseline value (x): $p_{13,sp2} = 0.2647$	Reparametrization: $p_{13,sp3} = f(s_{thr})$ Baseline value (x): $p_{13,sp3} = 0.3358$	Reparametrization: $p_{13,sp4} = f(s_{thr})$ Baseline value (x): $p_{13,sp4} = 0.2451$	Reparametrization: $p_{13,sp5} = f(s_{thr})$ Baseline value (x): $p_{13,sp5} = 0.4926$	-	-	Reparametrization: $p_{15,sp8} = f(s_{thr})$ $p_{16,sp8} = f(s_{thr})$ Baseline value (x): $p_{15,sp8} = 0.6316$ $p_{16,sp8} = 0.5$	Reparametrization: $p_{13,sp9} = f(s_{thr})$ Baseline value (x): $p_{13,sp9} = 0.2308$
maximum starvation mortality rate ($M_{\varepsilon_{max}}$ in y^{-1})	Reparametrization: $p_{14,sp1} = f(A)$ Baseline value (x): $p_{14,sp1} = 0.10$	Reparametrization: $p_{14,sp2} = f(A)$ Baseline value (x): $p_{14,sp2} = 0.05$	Reparametrization: $p_{14,sp3} = f(A)$ Baseline value (x): $p_{14,sp3} = 0.10$	Reparametrization: $p_{14,sp4} = f(A)$ Baseline value (x): $p_{14,sp4} = 0.15$	Reparametrization: $p_{14,sp5} = f(A)$ Baseline value (x): $p_{14,sp5} = 0.05$	Reparametrization: $p_{11,sp6} = f(A)$ Baseline value (x): $p_{11,sp6} = 0.5$	Reparametrization: $p_{11,sp7} = f(A)$ Baseline value (x): $p_{11,sp7} = 0.1$	Reparametrization: $p_{17,sp8} = f(A)$ Baseline value (x): $p_{17,sp8} = 0.1$	Reparametrization: $p_{14,sp9} = f(A)$ Baseline value (x): $p_{14,sp9} = 0.5$
von Bertalanffy threshold (a_{thr} in y)	Reparametrization: $p_{15,sp1} = f(a_{thr})$ Baseline value (x): $p_{15,sp1} = 0.0179$	Reparametrization: $p_{15,sp2} = f(a_{thr})$ Baseline value (x): $p_{15,sp2} = 0.0074$	Reparametrization: $p_{15,sp3} = f(a_{thr})$ Baseline value (x): $p_{15,sp3} = 0.0129$	Reparametrization: $p_{15,sp4} = f(a_{thr})$ Baseline value (x): $p_{15,sp4} = 0.0061$	Reparametrization: $p_{15,sp5} = f(a_{thr})$ Baseline value (x): $p_{15,sp5} = 0.0123$	$p_{12,sp6} = f(a_{thr})$ Baseline value (x): $p_{12,sp6} = 0.0438$	$p_{12,sp7} = f(a_{thr})$ Baseline value (x): $p_{12,sp7} = 0.1190$	$p_{18,sp8} = f(a_{thr})$ Baseline value (x): $p_{18,sp8} = 0.0105$	Reparametrization: $p_{15,sp9} = f(a_{thr})$ Baseline value (x): $p_{15,sp9} = 0.0385$
egg size (cm)	Reparametrization: $p_{16,sp1} = f(A)$ Baseline value (x): $p_{16,sp1} = 0.10$	Reparametrization: $p_{16,sp2} = f(A)$ Baseline value (x): $p_{16,sp2} = 0.10$	Reparametrization: $p_{16,sp3} = f(A)$ Baseline value (x): $p_{16,sp3} = 0.10$	Reparametrization: $p_{16,sp4} = f(A)$ Baseline value (x): $p_{16,sp4} = 0.10$	Reparametrization: $p_{16,sp5} = f(A)$ Baseline value (x): $p_{16,sp5} = 0.10$	Reparametrization: $p_{13,sp6} = f(A)$ Baseline value (x): $p_{13,sp6} = 0.10$	Reparametrization: $p_{13,sp7} = f(A)$ Baseline value (x): $p_{13,sp7} = 0.10$	Reparametrization: $p_{19,sp8} = f(A)$ Baseline value (x): $p_{19,sp8} = 0.10$	Reparametrization: $p_{16,sp9} = f(A)$ Baseline value (x): $p_{16,sp9} = 0.10$
critical threshold of predation efficiency (ε_{crit})	Reparametrization: $p_{17,sp1} = f(A)$ Baseline value (x): $p_{17,sp1} = 0.57$	Reparametrization: $p_{17,sp2} = f(A)$ Baseline value (x): $p_{17,sp2} = 0.57$	Reparametrization: $p_{17,sp3} = f(A)$ Baseline value (x): $p_{17,sp3} = 0.57$	Reparametrization: $p_{17,sp4} = f(A)$ Baseline value (x): $p_{17,sp4} = 0.57$	Reparametrization: $p_{17,sp5} = f(A)$ Baseline value (x): $p_{17,sp5} = 0.57$	Reparametrization: $p_{14,sp6} = f(A)$ Baseline value (x): $p_{14,sp6} = 0.57$	Reparametrization: $p_{14,sp7} = f(A)$ Baseline value (x): $p_{14,sp7} = 0.57$	Reparametrization: $p_{20,sp8} = f(A)$ Baseline value (x): $p_{20,sp8} = 0.57$	Reparametrization: $p_{17,sp9} = f(A)$ Baseline value (x): $p_{17,sp9} = 0.57$
maximum rate of predation ingestion	Reparametrization: $p_{18,sp1} = f(A)$ Baseline value (x):	Reparametrization: $p_{18,sp2} = f(A)$ Baseline value (x):	Reparametrization: $p_{18,sp3} = f(A)$ Baseline value (x):	Reparametrization: $p_{18,sp4} = f(A)$ Baseline value (x):	Reparametrization: $p_{18,sp5} = f(A)$ Baseline value (x):	Reparametrization: $p_{15,sp6} = f(A)$ Baseline value (x):	Reparametrization: $p_{15,sp7} = f(A)$ Baseline value (x):	Reparametrization: $p_{21,sp8} = f(A)$ Baseline value (x):	Reparametrization: $p_{18,sp9} = f(A)$ Baseline value (x):

(I_{max} in $g\ body g^{-1} y^{-1}$)	$p_{18,sp1} = 3.5$	$p_{18,sp2} = 3.5$	$p_{18,sp3} = 3.5$	$p_{18,sp4} = 3.5$	$p_{18,sp5} = 3.5$	$p_{15,sp6} = 3.5$	$p_{15,sp7} = 3.5$	$p_{21,sp8} = 3.5$	$p_{18,sp9} = 3.5$
natural mortality rate (M in y^{-1})	Reparametrization: $p_{19,sp1} = f(A)$ Baseline value (x): $p_{19,sp1} = 0.9177$	Reparametrization: $p_{19,sp2} = f(A)$ Baseline value (x): $p_{19,sp2} = 0.2102$	Reparametrization: $p_{19,sp3} = f(A)$ Baseline value (x): $p_{19,sp3} = 0.1624$	Reparametrization: $p_{19,sp4} = f(A)$ Baseline value (x): $p_{19,sp4} = 0.1853$	Reparametrization: $p_{19,sp5} = f(A)$ Baseline value (x): $p_{19,sp5} = 0.2136$	Reparametrization: $p_{16,sp6} = f(A)$ Baseline value (x): $p_{16,sp6} = 0.7108$	Reparametrization: $p_{16,sp7} = f(A)$ Baseline value (x): $p_{16,sp7} = 0.2322$	Reparametrization: $p_{22,sp8} = f(A)$ Baseline value (x): $p_{22,sp8} = 1.8466$	Reparametrization: $p_{19,sp9} = f(A)$ Baseline value (x): $p_{19,sp9} = 0.2879$
larval mortality rate (M_0 in $month^{-1}$)	Reparametrization: $p_{20,sp1} = f(A)$ Baseline value (x): $p_{20,sp1} = 9.5821$	Reparametrization: $p_{20,sp2} = f(A)$ Baseline value (x): $p_{20,sp2} = 9.8035$	Reparametrization: $p_{20,sp3} = f(A)$ Baseline value (x): $p_{20,sp3} = 13.2393$	Reparametrization: $p_{20,sp4} = f(A)$ Baseline value (x): $p_{20,sp4} = 12.5983$	Reparametrization: $p_{20,sp5} = f(A)$ Baseline value (x): $p_{20,sp5} = 13.1210$	Reparametrization: $p_{17,sp6} = f(A)$ Baseline value (x): $p_{17,sp6} = 2.0804$	Reparametrization: $p_{17,sp7} = f(A)$ Baseline value (x): $p_{17,sp7} = 6.7572$	Reparametrization: $p_{23,sp8} = f(A)$ Baseline value (x): $p_{23,sp8} = 8.9452$	Reparametrization: $p_{20,sp9} = f(A)$ Baseline value (x): $p_{20,sp9} = 1.5779$
fishing mortality multiplier (f_m)	Reparametrization: $p_{21,sp1} = f(A)$ Baseline value (x): $p_{21,sp1} = 1$	Reparametrization: $p_{21,sp2} = f(A)$ Baseline value (x): $p_{21,sp2} = 1$	Reparametrization: $p_{21,sp3} = f(A)$ Baseline value (x): $p_{21,sp3} = 1$	Reparametrization: $p_{21,sp4} = f(A)$ Baseline value (x): $p_{21,sp4} = 1$	Reparametrization: $p_{21,sp5} = f(A)$ Baseline value (x): $p_{21,sp5} = 1$	-	-	Reparametrization: $p_{24,sp8} = f(A)$ Baseline value (x): $p_{24,sp8} = 1$	-
sex ratio ($Frac_{fem}$)	Reparametrization: $p_{22,sp1} = f(A)$ Baseline value (x): $p_{22,sp1} = 0.5$	Reparametrization: $p_{22,sp2} = f(A)$ Baseline value (x): $p_{22,sp2} = 0.5$	Reparametrization: $p_{22,sp3} = f(A)$ Baseline value (x): $p_{22,sp3} = 0.5$	Reparametrization: $p_{22,sp4} = f(A)$ Baseline value (x): $p_{22,sp4} = 0.5$	Reparametrization: $p_{22,sp5} = f(A)$ Baseline value (x): $p_{22,sp5} = 0.5$	Reparametrization: $p_{18,sp6} = f(A)$ Baseline value (x): $p_{18,sp6} = 0.5$	Reparametrization: $p_{18,sp7} = f(A)$ Baseline value (x): $p_{18,sp7} = 0.5$	Reparametrization: $p_{25,sp8} = f(A)$ Baseline value (x): $p_{25,sp8} = 0.5$	Reparametrization: $p_{21,sp9} = f(A)$ Baseline value (x): $p_{21,sp9} = 0.5$
$L_{t=0}$ (von Bertalanffy growth parameter in cm)	Reparametrization: $p_{23,sp1} = f(A)$ Baseline value (x): $p_{23,sp1} = 1.9682$	Reparametrization: $p_{23,sp2} = f(A)$ Baseline value (x): $p_{23,sp2} = 3.6483$	Reparametrization: $p_{23,sp3} = f(A)$ Baseline value (x): $p_{23,sp3} = 9.8834$	Reparametrization: $p_{23,sp4} = f(A)$ Baseline value (x): $p_{23,sp4} = 3.7278$	Reparametrization: $p_{23,sp5} = f(A)$ Baseline value (x): $p_{23,sp5} = 0.8238$	Reparametrization: $p_{19,sp6} = f(A)$ Baseline value (x): $p_{19,sp6} = 0.5333$	Reparametrization: $p_{19,sp7} = f(A)$ Baseline value (x): $p_{19,sp7} = 0.4861$	Reparametrization: $p_{26,sp8} = f(A)$ Baseline value (x): $p_{26,sp8} = 8.9544$	Reparametrization: $p_{22,sp9} = f(A)$ Baseline value (x): $p_{22,sp9} = 0.7795$
k (von Bertalanffy growth parameter in y^{-1})	Reparametrization: $p_{24,sp1} = f(A)$ Baseline value (x): $p_{24,sp1} = 0.76$	Reparametrization: $p_{24,sp2} = f(A)$ Baseline value (x): $p_{24,sp2} = 0.205$	Reparametrization: $p_{24,sp3} = f(A)$ Baseline value (x): $p_{24,sp3} = 0.22$	Reparametrization: $p_{24,sp4} = f(A)$ Baseline value (x): $p_{24,sp4} = 0.1670$	Reparametrization: $p_{24,sp5} = f(A)$ Baseline value (x): $p_{24,sp5} = 0.41$	Reparametrization: $p_{20,sp6} = f(A)$ Baseline value (x): $p_{20,sp6} = 1.15$	Reparametrization: $p_{20,sp7} = f(A)$ Baseline value (x): $p_{20,sp7} = 0.375$	Reparametrization: $p_{27,sp8} = f(A)$ Baseline value (x): $p_{27,sp8} = 1.1$	Reparametrization: $p_{23,sp9} = f(A)$ Baseline value (x): $p_{23,sp9} = 1.8$
L_∞ (von Bertalanffy growth parameter in cm)	Reparametrization: $p_{25,sp1} = f(A)$ Baseline value (x): $p_{25,sp1} = 19.50$	Reparametrization: $p_{25,sp2} = f(A)$ Baseline value (x): $p_{25,sp2} = 68$	Reparametrization: $p_{25,sp3} = f(A)$ Baseline value (x): $p_{25,sp3} = 38.71$	Reparametrization: $p_{25,sp4} = f(A)$ Baseline value (x): $p_{25,sp4} = 81.6$	Reparametrization: $p_{25,sp5} = f(A)$ Baseline value (x): $p_{25,sp5} = 40.60$	Reparametrization: $p_{21,sp6} = f(A)$ Baseline value (x): $p_{21,sp6} = 8$	Reparametrization: $p_{21,sp7} = f(A)$ Baseline value (x): $p_{21,sp7} = 4.2$	Reparametrization: $p_{28,sp8} = f(A)$ Baseline value (x): $p_{28,sp8} = 95$	Reparametrization: $p_{24,sp9} = f(A)$ Baseline value (x): $p_{24,sp9} = 2.6$
size at maturity (s_{mat} in cm)	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:	Reparametrization:

	$p_{26,sp1} = f(s_{mat})$ Baseline value (x): $p_{26,sp1} = 0.6154$	$p_{26,sp2} = f(s_{mat})$ Baseline value (x): $p_{26,sp2} = 0.5147$	$p_{26,sp3} = f(s_{mat})$ Baseline value (x): $p_{26,sp3} = 0.5425$	$p_{26,sp4} = f(s_{mat})$ Baseline value (x): $p_{26,sp4} = 0.3554$	$p_{26,sp5} = f(s_{mat})$ Baseline value (x): $p_{26,sp5} = 0.7143$	$p_{22,sp6} = f(s_{mat})$ Baseline value (x): $p_{22,sp6} = 0.3125$	$p_{22,sp7} = f(s_{mat})$ Baseline value (x): $p_{22,sp7} = 0.4524$	$p_{29,sp8} = f(s_{mat})$ Baseline value (x): $p_{29,sp8} = 0.6947$	$p_{25,sp9} = f(s_{mat})$ Baseline value (x): $p_{25,sp9} = 0.3077$
constant of proportionality of the allometric length-weight relationship (c in $g\ cm^{-3}$)	Reparametrization: $p_{27,sp1} = f(A)$ Baseline value (x): $p_{27,sp1} = 0.0065$	Reparametrization: $p_{27,sp2} = f(A)$ Baseline value (x): $p_{27,sp2} = 0.007$	Reparametrization: $p_{27,sp3} = f(A)$ Baseline value (x): $p_{27,sp3} = 0.0089$	Reparametrization: $p_{27,sp4} = f(A)$ Baseline value (x): $p_{27,sp4} = 0.0135$	Reparametrization: $p_{27,sp5} = f(A)$ Baseline value (x): $p_{27,sp5} = 0.0086$	Reparametrization: $p_{23,sp6} = f(A)$ Baseline value (x): $p_{23,sp6} = 0.0083$	Reparametrization: $p_{23,sp7} = f(A)$ Baseline value (x): $p_{23,sp7} = 0.174$	Reparametrization: $p_{30,sp8} = f(A)$ Baseline value (x): $p_{30,sp8} = 0.005$	Reparametrization: $p_{26,sp9} = f(A)$ Baseline value (x): $p_{26,sp9} = 0.0093$

Appendix B. NHCE OSMOSE model re-parametrization

Here, we briefly describe the NHCE OSMOSE parameters selected for the UA.

- Predation parameters: The predation accessibilities are used in the model to describe the accessibility of prey to predators. These parameters reflect the vertical overlap of predators and prey in the water column. Predation accessibilities take values between 0 and 1 and determine the total accessible biomass (P_{tot_s}) for the predator school (s):

$$P_{tot_s} = \sum_{p=1}^{p=N} A(pred_s, prey_p)\, B_{prey_p} \quad \text{(b.1)}$$

where p is the prey school (p from 1 to N) available for school s, $A(pred_s, prey_p)$ is the accessibility of each prey p to the predator school s and, B_{prey_p} the corresponding biomass of each prey p. In addition, the total biomass that a predator school can eat ($P_{eatable_s}$) depends on the biomass of the predator (B_{pred_s}) and the maximum rate of ingestion of the predator species (I_{max}):

$$P_{eatable_s} = B_{pred_s} I_{max} \quad \text{(b.2)}$$

We decided to run the UA on the $A(pred_s, prey_p)$ parameters (predation accessibilities parameters, in Table 2 the parameter type 1). These values are proportions $\in [0, 1]$, so we chose a logit scale for these parameters. We also run the UA on the I_{max} (Table 2 the parameter type 9). Since the only constraint is to keep positive the parameter value, we used the logarithmic scale for I_{max}.

- Feeding size ranges: In OSMOSE, the predation is controlled by the minimum (R_{min}) and maximum (R_{max}) size ratios (in Table 2, parameters type 2 and 3). Predator schools can only feed on prey schools whose lengths meet these thresholds:

$$R_{min} \leq \frac{L_{pred}}{L_{prey}} \leq R_{max} \quad \text{(b.3)}$$

where the ratios (R_{min} and R_{max}) are the threshold values for predator length (L_{pred}) over prey length (L_{prey}) ratio. When information is available, a split around the size threshold (s_{thr}) allows to specify different feeding size range between species stages (the parameter called predator-prey

size threshold, type 4 in Table 2). Predator-prey size threshold (s_{thr}) is part of the UA. However, we re-parameterized it by $s_{thr}/\ L_{\infty}$ in logit scale (with values from 0 to 1). It ensures that s_{thr} will always be smaller or equal to L_{∞}.

Likewise, R_{min} should never be higher than R_{max}. Manipulations of R_{min} and R_{max} separately for the UA are risky and should be avoided for model consistency. We thus decided to re-parametrize the equation of this process and implement the UA using angles in polar coordinates. Users should think carefully about the potential consequences of changing parameter values that are interdependent in the model equations and re-parametrize them when necessary. Using equation b.3 in cartesian coordinates (predator size in abscissa and prey size in ordinate), we transformed it into polar coordinates calculating angles between the abscissa, minimum and maximum prey length:

$$g(m1, m2) = tan^{-1}\left(\frac{m_2 - m_1}{1 + m_1 m_2}\right) \quad \text{(b.4)}$$

where $g(m1, m2)$ is the angle (in degrees) between the slopes (m_1 and m_2), m_2 being always greater than m_1. Using equation b.4, we calculated two angles: i) the angle between the horizontal axis and the minimum prey length ($\theta = g(0, m1)$), and the angle between the minimum and maximum prey length ($\alpha = g(m1, m2)$). The parameters m_1 and m_2 were re-parametrized as the angles θ and α for each species (see Table A.1, e.g., $\theta_{stage\ 1}$ and $\alpha_{stage\ 1}$, and $\theta_{stage\ 2}$ and $\alpha_{stage\ 2}$ for the first and second anchovy stages). The new parameterization physically restricts the parameter values to the first quadrant, representing all the possibilities for the predation-prey size ratios (π/4 means predator and prey size can be equal while π/2 means there is no limit on the prey size for a predator). These values (as part of the re-parameterization) are presented in Table A.1. For the UA, the angles θ and α (as fractions of π/2) use a logit scale.

- Starvation mortality: When the predation efficiency (ε_i) is below the critical threshold (ε_{crit}, parameter type 8 in Table 2), schools do not have enough food to fulfill their requirements undergoing starvation mortality (M_{ε_i} of school i, see equation b.5). This process is controlled by the maximum starvation mortality rate ($M_{\varepsilon_{max}}$) of the species as follows:

$$M_{\varepsilon_i} = -\frac{M_{\varepsilon_{max}}}{\varepsilon_{crit}} . \varepsilon_i + M_{\varepsilon_{max}} \quad \text{(b.5)}$$

We implemented the UA using $M_{\varepsilon_{max}}$ (parameter type 5 in Table 2) on a logarithmic scale since its values are positive. ε_{crit} is also part of the UA, but we use a logit scale because its values accept only numbers between 0 and 1.

- Growth: Individuals of a given school are assumed to grow in size and weight (at time *t)* only when the food ingested fulfills their maintenance requirements. A simple linear model is used below the threshold age (a_{thr}, parameter type 6 in Table 2). Above a_{thr}, the von Bertalanffy model is used to determine the average mean length increase, as follows:

$$L_t = L_\infty(1 - exp(-k(t - t_0)) \quad \text{(b.6)}$$

where L_∞ is the asymptotic size (parameter type 16 in Table 2), k the growth coefficient (parameter type 15 in Table 2), and t_0 the theoretical age when size is 0. The UA is implemented using L_∞, k, $L_{t=0}$ (the length at t = 0, parameter type 14 in Table 2). These parameters are manipulated individually and on a logarithmic scale since the only constraint is to keep the parameter value positive. a_{thr} is also part of the UA on a logit scale. However, we re-parameterized it by $a_{thr}/\ a_{max}$ on a logit scale (with values from 0 to 1), where a_{max} is the longevity (in years). It ensures that a_{thr} will always be smaller or equal to a_{max}, but never bigger than a_{max}.
The weight of a school is calculated using the allometric relationship:

$$W = cL^b \quad \text{(b.7)}$$

where b is the exponent and c the constant of proportionality of the allometric length-weight relationship. We chose to constrain the UA to the constant of proportionality (*c,* in Table 2 the p_{18}) in logarithmic scale, leaving b fixed.

- Size and weight of eggs: The life cycle of each species is modelled starting with the egg stage. Egg size is part of the UA (parameter type 7 in Table 2) and it co-varies with egg weight. For each UA simulation, we calculated the corresponding egg weight as a function of egg size, keeping the mean density constant as follows:

$$egg_{weight} = \frac{4}{3}\pi\left(\frac{egg_{size}}{2}\right)^3 mean_{density} \quad \text{(b.8)}$$

The $mean_{density}$ is calculated using the baseline egg size and weight values. Then each time that the egg_{size} takes a new value due to the UA, the egg_{weight} should be recalculated, meaning that

both parameters are changing together. The only mathematical constraint for these parameters is to obtain a positive numerical value, so we used the logarithmic scale.

- Sources of mortality: The abundance of a school i at time $t + \Delta t$, ($N_{i,t+\Delta t}$), is modeled as a function of abundance at time t ($N_{i,t}$) and the following sources of mortality:

$$N_{i,t+\Delta t} = N_{i,t}\, e^{-\Delta t(F_s(t) + M_s(t) + M_{\epsilon_i}(t) + P_i(t))} \quad \text{(b.9)}$$

where F_s is the fishing mortality rate and M_s the natural mortality rate of species s. M_{ϵ_i} is the starvation mortality rate and P_i the predation mortality rate, both related to the fish school i.

The natural mortality rate (parameter type 10 in Table 2) and starvation mortality (already described in equation b.5) are part of the UA. The fishing mortality rate used in the NHCE OSMOSE model varies with time (t) and species size. We decided to include the fishing mortality in the UA by creating a fishing mortality multiplier f_m (parameter type 12 in Table 2). It is a positive value with a baseline value equal to 1. The fishing multiplier is used in the UA by multiplying the fishing mortality rate (matrix of fishing mortalities, time vs. species size).

Another source of mortality is the larval mortality rate ($M_{0,s}$). This parameter controls the number of eggs and larvae for species s, at age 0 and at time $t + \Delta t$, ($N_{s,0,t+\Delta t}$), as follows:

$$N_{s,0,t+\Delta t} = N_{s,0,t}\, e^{-\Delta t M_{0,s}} \quad \text{(b.10)}$$

where $N_{s,0,t}$ represents the number of eggs and larvae in the system at time t, and $M_{0,s}$ is the larval mortality rate by species (parameter type 11 in Table 2). This parameter is a vector (one larval mortality rate by time step) in the NHCE OSMOSE model. The UA takes this parameter using the mean of the vector (mean over time).

Baseline values of natural mortality rate, larval mortality, and fishing mortality were estimated by a calibration process following a logarithmic scale. We chose the same scale to perform the experiments as part of the UA.

- Reproduction: any school whose length is greater than the size at sexual maturity (s_{mat}, parameter type 17 in Table 2) enters the reproduction process at the end of the time step. This allows the generation of a new school at the egg stage. For a given species:

$$N_{eggs} = Frac_{fem}\, \alpha\, season\, B_{mat} \quad \text{(b.11)}$$

where $Frac_{fem}$ is the fraction of females (also called sex ratio, parameter type 13 in Table 2), α is the relative fecundity. This parameter indicates the number of eggs per gram of mature female.

The $season$ is a file that provides the spawning distribution within a year, and B_{mat} is the sum of the biomass of the schools that reached sexual maturity.

$Frac_{fem}$ is part of the UA. We chose the logit scale because this parameter only takes values from 0 to 1. The size at sexual maturity (s_{mat}) is also part of the UA. However, we re-parameterized it by s_{mat}/ L_{∞} in logit scale (with values from 0 to 1). It ensures that s_{mat} will always be smaller or equal to L_{∞}, but never bigger than L_{∞}.

Chapter 3

A protocol for implementing sensitivity analyses in complex ecosystem models using parameter reliability

Ecosystem model projections need to be explored in the face of uncertainty. Sensitivity analysis is an essential step in understanding model uncertainty, but there are many challenges when dealing with complex ecosystem models. One of them is related to the uncertainty quantification in model input parameters, mainly when limited data leads to arbitrary fixed ranges of variation to draw the uncertainty around the parameter values. We show the drawbacks of this practice and propose an alternative methodology based on the parameter reliability criterion. This criterion serves to classify the model parameters according to the source of information used to estimate their values and calculate the ranges of variability for each parameter used in a sensitivity analysis. We illustrated our methodology by implementing a sensitivity analysis on the marine ecosystem modelling platform OSMOSE applied to the northern Humboldt current ecosystem. We compared these results with those obtained from sensitivity analysis using fixed ranges of variation. Our main result is that using arbitrary ranges of variation could produce varying conclusions depending on the used ranges. Alternative approaches, such as the one based on the reliability of the parameters, can be helpful in situations where uncertainty quantification of model inputs is not available. However, we recommend further developments of new approaches which facilitate the study of uncertainty in complex models.

Keywords: parameter uncertainty, ecosystem models, OSMOSE model.

A protocol for implementing sensitivity analyses in complex ecosystem models using parameter reliability

Criscely Luján [1,2], Ricardo Oliveros-Ramos [3], Nicolas Barrier [4], Paul Leadley [1],

Yunne-Jai Shin [5]

[1] *Laboratoire d'Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91400, Orsay. France*

[2] *MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Montpellier. France*

[3] *Instituto del Mar del Perú, Callao, Peru.*

[4] *MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Sète. France*

[5] *University of Cape Town, Cape Town, South Africa*

*Corresponding author: criscelylujan@gmail.com

1. Introduction

With growing knowledge and databases on ecosystem structure and functioning, there has been a simultaneous investment in developing more integrative ecosystem models, progressively representing additional components and ecological processes. This has resulted in more complex models, often interdisciplinary, e.g., integrating the physical, chemical and biological components of the ecosystem (Travers et al., 2007; Alkemade et al. 2009; Holland et al., 2012; Tittensor et al., 2018). Meant to increase the realism of ecosystem representation, one of the main criticisms of these models is that their complexity can render their predictions highly uncertain (Fulton, Smith, and Johnson 2003). To increase the credibility and usefulness of ecosystem model projections, we urgently need to find means to systematically address their uncertainty (Ferrier et al., 2016).

Among the different sources of uncertainty related to complex ecosystem models (Regan, Colyvan, and Burgman 2002; Link et al. 2012), we focus here on parameter uncertainty. This source of uncertainty relates to the specific parameter values used to run a model and influence its behavior (Cheung et al., 2016). Many processes can give rise to parameter uncertainty, including lack of data, imperfect measurements, inadequate coverage of the range of natural variability, or

the natural variability of biological parameters (Payne et al., 2015). In this context, sensitivity analyses techniques are a good option to explore the robustness of ecosystem model outputs to uncertainties (Payne et al., 2015).

Sensitivity analysis (SA) examines how the uncertainty in model outputs can be apportioned to different sources of uncertainty in the model inputs (Saltelli et al., 2019). Performing a SA can be helpful for a wide range of purposes (Pannel et al., 1997), for example: (i) to increase the understanding of the relationships between model input and output; (ii) to test the robustness of model outputs in the presence of uncertainty; (iii) to identify the model input parameters that have critical or threshold values which influence optimal strategies based on model outputs; (iv) to perform model simplification by fixing inputs that have low effect on the model outputs; (v) to search errors in the model by finding unexpected relationships between model inputs and outputs, and (vi) to make recommendations based on a more credible and robust model.

Far from being trivial, selecting a SA method depends on the problem addressed, the characteristics of the model under study, and the computational cost afforded (Campolongo, Tarantola, and Saltelli 1999). There are many methods to perform a SA (Saltelli et al. 2008; Cariboni et al. 2007). But in general, most procedures follow four steps (Figure 1) which consist in 1. quantifying the uncertainty in model inputs; 2. running the model several times following an experimental design; 3. identifying the model outputs to be analyzed; 4. using the model outputs to calculate sensitivity measures of interest.

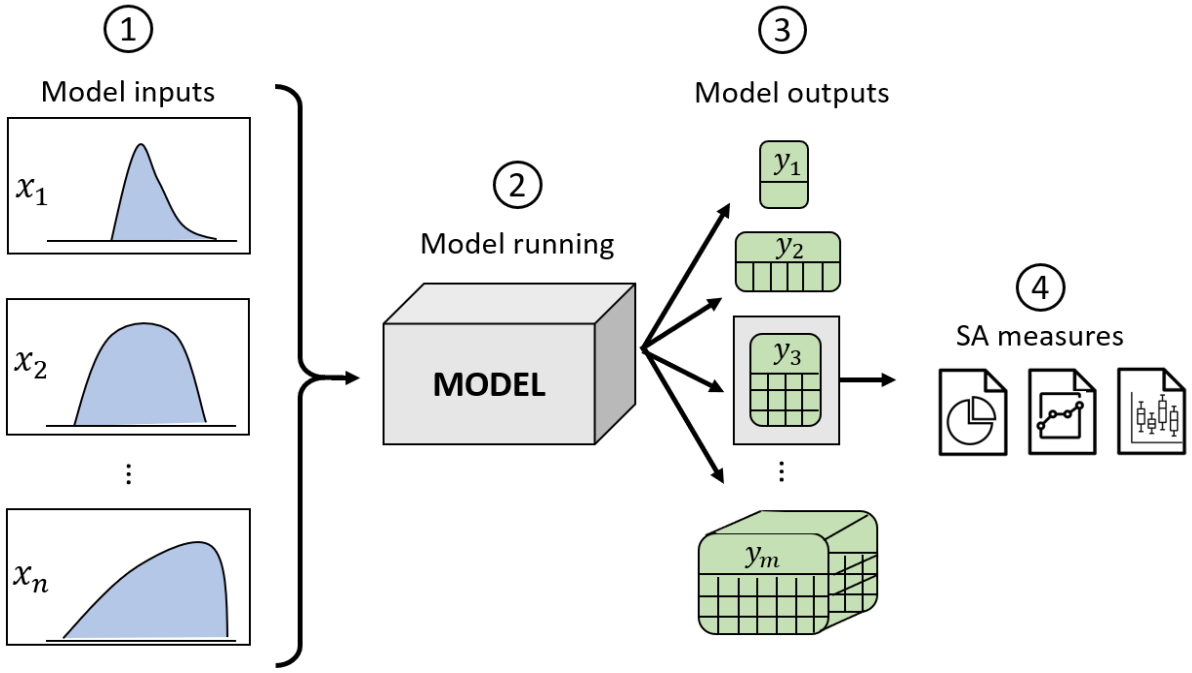

Figure 1. *Schematic representation of a SA implementation. We represent a model as a gray box that uses a set of inputs (x_1, x_2, ... , x_n; n is the total number of model inputs). This model produces m model outputs (y_1, y_2, y_3, ... ,y_m). 1. The first step in the SA is the uncertainty quantification in each model input, represented here as probability distributions (blue curves). 2. The second step consists in running the model several times following a design of experiments dictated by the SA method. 3. Based on the problem addressed, the third step is to identify one or several model outputs to estimate SA measures (here y_3). 4. Finally, sensitivity measures of interest are estimated using the resulting model outputs.*

As explained in Figure 1, SA methods will require information about the uncertainty in model input parameters. This information can be provided in the form of probability distributions, or variability ranges that observational or experimental data can support (e.g., in Tranter et al., 2021). Nevertheless, when data is not available or insufficient, the perturbation of parameter values uses a predefined range of variation to draw the uncertainty around each parameter reference value (e.g., with values between 10% to 30% in Dantec-Nédélec et al., 2017; Specka et al., 2015; Song et al., 2013). The combined use of both strategies (some parameter bounds based

on data and others using predefined ranges) is also found in the literature (e.g., in Morris et al., 2014; Bracis et al., 2020).

Complex ecosystem models typically require an extensive set of input parameters, the uncertainty of which is often poorly determined due to a lack of information and data. Therefore SA of ecosystem models commonly use arbitrary ranges of variation in model parameters, but this can have considerable impacts on the SA results (Saltelli et al., 2000; Wallach et al., 2006). To improve the assessment of the uncertainty related to input parameters, we propose an original methodology based on the parameter reliability (PR) criterion and inspired by the "pedigree" described by Funtowicz and Ravetz (1990). This pedigree system has been implemented in the Ecopath and Ecosim model (Christensen et al., 2005), allowing the assignment of confidence intervals to model inputs based on the description of their origin. These confidence intervals are used in uncertainty studies as prior probability distributions for model inputs. Following the PR criterion, instead of directly assigning confidence intervals, we first categorize the parameters of the model; this will define the reliability level of the parameters, and based on a set of rules that prioritize the use of available data, we will calculate the ranges of the variability of the parameters. Here, we illustrate our proposed methodology with the marine ecosystem modelling platform OSMOSE (Shin and Cury 2001; 2004) applied to the Northern Humboldt Current Ecosystem (NHCE) (Oliveros et al., 2017) and compare the results of the SA with those obtained using arbitrary predefined ranges of variation. This work is conceived to be generic enough to help run SAs in other complex ecological models.

2. Material and methods

2.1 Addressing parameter uncertainty

2.1.1 Categorization of model parameters

We propose the parameter reliability (PR) criterion, which serves a triple purpose by i) describing the parameter source (i.e., the information used to estimate a parameter value), ii) assigning to each model parameter a qualitative value (named hierarchy), and iii) providing a criterion for assigning uncertainty level to model parameters.

The PR criterion relies on predefined categories of parameters that depend on the sources of information used to determine parameter values (see Table 1). Those categories are hierarchized to reflect the level of parameter reliability. Based on the PR criterion, we propose a set of rules for categorizing model parameters. The first rule is that each model parameter should be classified in only one category. However, it is not mandatory to use all categories to classify model parameters. As a second rule, the hierarchy provided according to parameters' sources should be respected (Table 1). Parameter reliability takes values that range between 1 (high reliability) to 8 (low reliability). For example, the reliability of parameters derived from high-quality data is scored as 1; it is scored as 2 when low-quality data are used and so on for all the categories presented in Table 1. Once the model parameters are grouped by information sources and classified using the PR criterion, they can be ranked and scored.

Two important precisions need to be taken into account when applying the PR criterion. The first relates to the conceptualization of a model. It is associated with the entire set of assumptions in the model building, including but not limited to the set of equations and algorithms (Jorgensen and Endoricchio 2001). For example, working with a given model, if the same set of equations and algorithms are used but applied to a different study area, or time period, or modelled species, these changes provide a model that is considered different from the first one. The second remark concerns the parameters sourced from "other models". We can classify a parameter of a model ***N*** as provided by "other models", e.g. a model ***M***, when its value is only directly or indirectly estimated by the model ***M***. On the contrary, for example, if an input parameter in ***M*** (which is different from being estimated by ***M***) falls in the "guesstimate" category, then if used in another model (***N***), this parameter should be categorized in ***N*** as "guesstimate" too, and not as coming from "other models".

Table 1. *Parameter reliability criterion. This criterion considers a set of information sources that are used to estimate a parameter (parameter source) and proposes a hierarchy of these sources of information. The parameter sources are classified into four principal categories: direct and indirect estimates and quantitative and qualitative approximations. Each category contains subcategories (eight in total across all categories) that explain the different ways in which information is used to estimate a parameter. The hierarchy allows to score the parameters from 1 (high parameter reliability) to 8 (low parameter reliability).*

Hierarchy	Parameter source	Description
Direct estimates: Data are collected, possibly after a dedicated experiment or survey, to estimate the parameter value. There is a direct link between the parameter value and the data collected.		
1	High-quality data	Data is collected with the specific purpose of estimating the parameters; sampling or measure errors are low.
2	Low-quality data	Data is collected on an opportunistic basis, not specifically for the parameter, but can be used to estimate it; sampling or measure errors can be high.
Indirect estimates: Data is not collected with the purpose to estimate the parameter value however this is done opportunistically. A model is used to create an indirect link between the parameter value and the data collected.		
3	Model calibration	Parameter estimation by minimizing an error function using the model and available data.
Quantitative approximations: Data is not collected for direct or indirect parameter estimation, but there are quantitative reasons to use a number as a value for the parameter.		
4	Empirical relationships	Parameter values using estimates from empirical relationships (for different species and systems, meta-analysis).
5	Model tuning	Parameter values using manual screening that improves model outputs without a formal optimization. It should be using the same model.
6	Proxies	Proxy of the parameter calculated from data.
Qualitative approximations: Data is not collected for direct or indirect parameter estimation, but there are qualitative reasons to use a number as a value for the parameter.		
7	Other models	Indirect or direct estimates are taken from other models.
8	Guesstimate	Use of professional judgment from a relevant expert. This judgment can be based on a literature review or expert knowledge.

2.1.2 Range calculation using the PR criterion

The PR criterion enabled to classify parameters in order of reliability. This ranking is essential for determining the range of variation of the parameters, when data are not always available to do so directly, avoiding the use of predefined and arbitrary ranges. Here, we propose a two tiers approach to determine the range of variations of all parameters involved in the SA, whether data are available or not. The first step is to identify all the parameters for which data can be used to estimate their range of variation. These data are diverse, coming from direct observations or experiments, published in the scientific literature, or derived from expert judgment. The range calculated from data should vary according to the unit and the scale (e.g., linear, logarithmic, and logit) of the parameter. It should also contain the baseline parameter value used in the model, but it is not mandatory to center the baseline value in the parameter range.

The second step uses the parameter reliability ranking to set up the range of variations of the parameters for which no data were available to do so directly. For convenience, we express the range of variation as a percentage, starting from 0%, with no upper limit fixed a priori. We assume that the range of variation of the parameters is a monotonic increasing function of the parameter reliability criterion (Figure 2). If any parameter that has a data-based range of variation does not follow this monotonicity rule, its PR categorization and the range calculation must be evaluated and performed again in an iterative way. Then the parameter ranges of variation that could not be estimated due to lack of data are delimited by the other parameter ranges following the PR ranking.

Using the PR criterion, we consider that the reliability classification is related to the parameter uncertainty. For example, a parameter classified in a low hierarchy category (Table 1) is considered to have a highly reliable estimate with a lower level of uncertainty than a parameter classified in a higher hierarchy category. This relationship between reliability and uncertainty explains why we expect smaller ranges of variation for high reliability parameters than for low reliability ones.

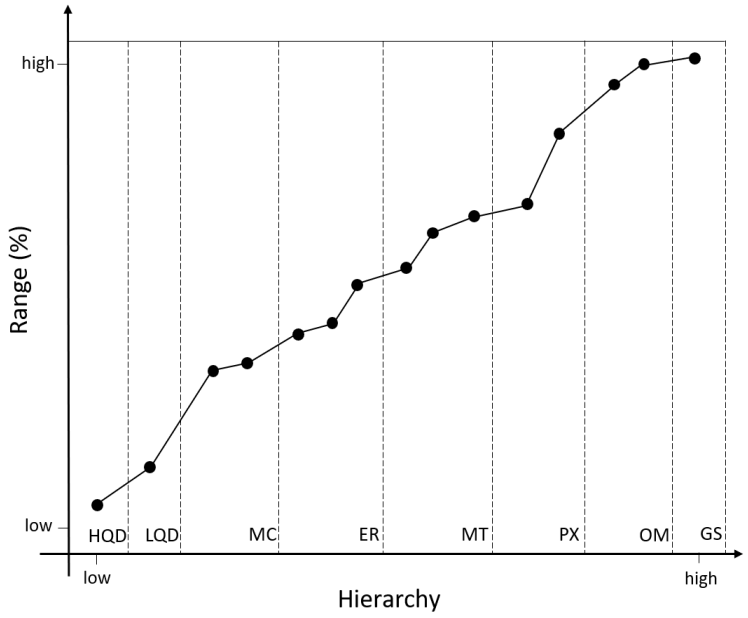

Figure 2. *Schematic representation of the monotonicity rule between parameter reliability and its range of variation. This scheme represents a model in which 15 parameters were classified (black dots) using all the categories presented in Table 1 (HQD: high-quality data, LQD: low-quality data, MC: model calibration, ER: empirical relationships, MT: model tuning, PX: proxies, OM: other models, and GS: guesstimate). The parameter ranges (in %) that could be calculated using data increase monotonically with the parameter hierarchy (reliability criterion).*

2.1.3 Illustrative cases for range calculation

To determine the range of variation of model parameters, we follow a stepwise protocol (Figure 3) which starts with classifying model parameters into the four PR categories (HQD: high-quality data, LQD: low-quality data, MC: model calibration, and ER: empirical relationships). Then several options are possible that are explained below by using six illustrative cases (Figure 4):

- An ideal case is when all the model parameter ranges can be directly calculated from available data (panel A in Figure 4).
- When there is no data to calculate the range of a parameter, we delimit all its possible values based on the monotonicity rule and on the PR category in which the parameter stands. In Panel B (Figure 4), the possible ranges (white points in the MC category) are within the gray area which is delimited by the maximum range of the LQD category and the minimum range of the ER category.

- Panel C (Figure 4) illustrates the case when there are no parameter ranges calculated from data in the same PR category (MC). Within the area with all possible ranges (area of gray color in Figure 4), alternative ranges could be used, for example, using the average range of the gray area (white points number 1) or the maximum possible ranges in the area (white points number 3). Using different range values within the same category is another possible alternative (white points number 2).
- Panel D (Figure 4) illustrates the case when data is missing to calculate the range of variation of all parameters classified in the first PR category (in Figure 4, the HQD category). Note that if the HQD is effectively the first non-empty category of the listed parameters, there is a low probability that data are missing to estimate the range of variability of the HQD parameters. Due to the monotonicity rule, the possible ranges (gray area) should be bounded by the first parameter range of the following category (LQD in this case). We represent some possible alternatives, with minimum (white points 1 - range of 0%), medium (white points 2), and maximum (white points 3) ranges.
- The opposite case, and a more probable one, is when data is missing to calculate the range of variation of all parameters in the last PR category (ER category in Panel E, Figure 4). Because no upper limit is fixed a priori for the range of variation, the only information that could be used is the maximum range of variation of the parameters in the previous PR category (last black dot in the MC category). Either this maximum percent variation could be used by default for the unknown ranges in the last PR category, or alternatively, an arbitrary percent of variation could be added to that maximum variation.
- Although we recommend calculating ranges using data and avoiding arbitrary use of ranges as much as possible, there will be cases with unavailable data for most or all parameter ranges (Panel F, Figure 4). In this case, users can only rely on the monotonicity rule. One suggestion is to start with the low hierarchy parameters (HQD category). The ranges set in this category will restrict the values of the parameter ranges in the following category (a gray area number 2 in the LQD category). And so on, the ranges selected in the MC category will restrict the set of possible parameter ranges as part of the ER category.

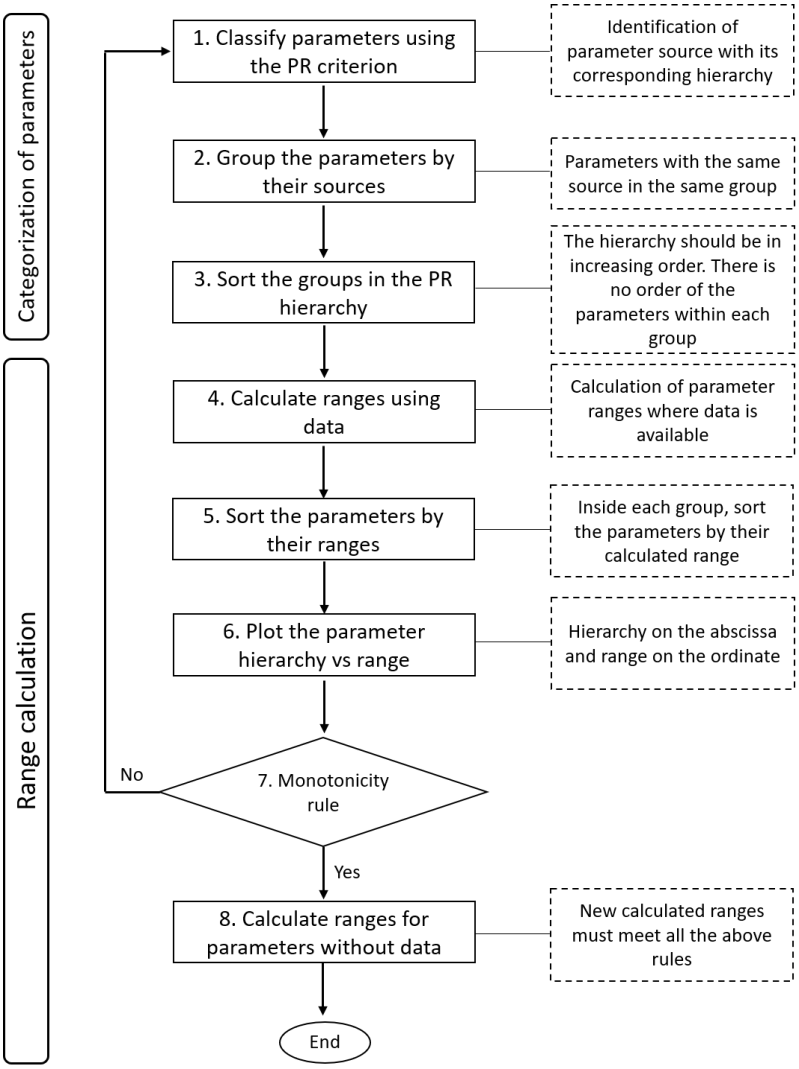

Figure 3. *A stepwise approach for parameter categorization and range calculation using the parameter reliability (PR) criterion. First, we suggest identifying the source of information used for assigning a value to the parameter and, based on Table 1, classify the parameter in its corresponding hierarchy (step 1). Parameters are grouped by source (step 2). Then the groups must be ranked using their hierarchy (step 3). For the parameters for which data are available, ranges of variation are calculated (step 4). Rank the parameters according to their ranges (step 5). Plot model parameters using their PR hierarchy (on the abscissa) and range of variation (on the ordinate) (step 6). Check that the range is an increasing monotonic function of the PR criterion (step 7). If this is not verified, parameter categorization or range estimation should be repeated (back to step 1). Otherwise, the calculation of parameter ranges when data is missing can be performed (step 8), following the recommendations in Figure 4.*

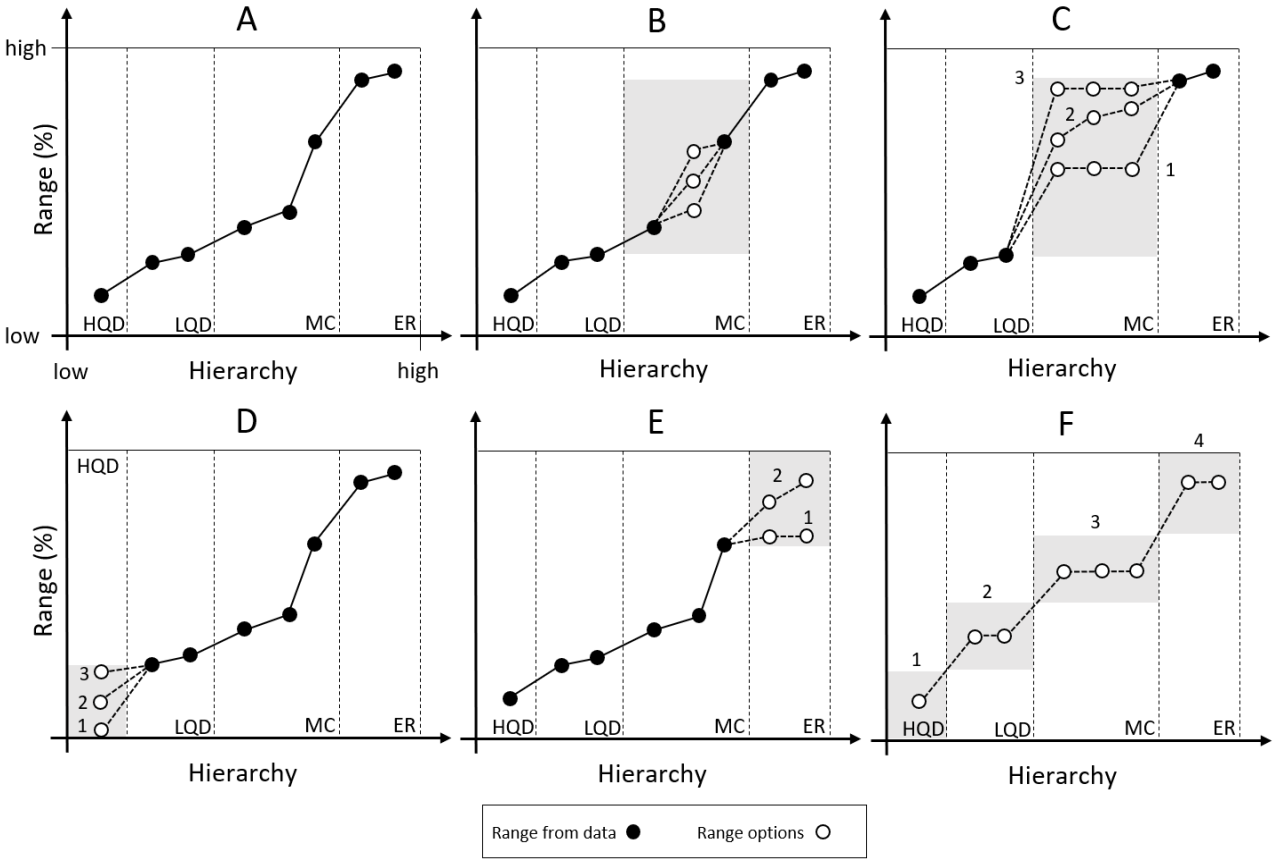

Figure 4. *Determination of parameter range of variation using the parameter reliability (PR) criterion. The panels show six illustrative cases with model parameters (dots) classified along a hierarchy comprising four PR categories (HQD: high-quality data, LQD: low-quality data, MC: model calibration, and ER: empirical relationships). Panel A shows an ideal case where ranges for all parameters are directly calculated from data (black points). When data are missing to determine the parameter ranges (white points in panels B to F), we first delimit the area with all possible ranges (the gray areas in panel B to F), which is delimited by the maximum range of the previous PR category and the minimum range of the following PR category. Then different options could be adopted to determine the range of variations, for example in Panel D taking the minimum (number 1), the median (number 2), or the maximum (number 3) range in the gray area. See main text for further explanations.*

2.2 Sensitivity analysis

2.2.1 SA method

We chose the Morris method (Morris, 1991) for the SA implementation in the NHCE OSMOSE model. Morris method is recommended for models with long-run times, which is typically the

case for many complex ecosystem models. Nevertheless, the use of other SA methods will not change the purpose of this work, namely to illustrate the PR criterion and associated protocol.

Morris is a global screening method composed of several randomized one-factor-at-a-time (OAT) experiments. Considering a model with *n* parameters x_i, ($i = 1, 2, \dots, n$), with each x_i scaled to take values in the interval [0,1], the Morris method creates a discrete parameter space (Ω) (the n-dimensional unit hypercube) by dividing the parameter ranges into *p* discrete levels. Then, the model is evaluated for *r* replicates within the parameter space, each of them building a trajectory inside Ω. The starting point of a trajectory is selected randomly. Besides, only a single parameter is changed for each trajectory, taking as a new value an element of the parameter space.

Morris proposed a sensitivity measure called Elementary Effects (EE, see equation 1). This is a simple and effective way of screening a few important parameters among the many that can be contained in the model. The EE for the *i*th parameter calculates the ratio of the variation of the model output at two different points of the parameter space to the variation of the parameter:

$$EE_i = \frac{y(x_1, \dots, x_{i-1}, x_i+\Delta, x_{i+1}, \dots, x_n) - y(x_1, \dots, x_n)}{\Delta} \quad (1)$$

where $x \in \Omega$, $x + \Delta$ are still in Ω, and Δ (called "grid jump") is a multiple of $1/(p-1)$. We followed Morris (1991), by using an even number for *p* and Δ equal to $p/[2(p-1)]$. To ensure that all the parameters (x_i) are modified once, the Morris method requires the execution of *n+1* simulations. This procedure is repeated *r* times, one per trajectory (i.e., the *r* Morris replicates). It provides *r* elementary effects for each parameter (x_i), resulting in a computational cost of $r(n+1)$ simulations to run.

The measures calculated using Morris are the mean μ_i (the average change in output across all changes in parameter *i*) and the standard deviation σ_i (the variance of the output in response to changes in parameter *i*) of the EE. μ_i assesses the overall importance of an input parameter *i* on model outputs, while σ_i reflects non-linear effects and interactions. Instead of using μ_i, we chose the alternative measure μ^*_i, which takes the absolute average change in output across all changes in parameter *i* (i.e. the absolute value is first applied to the EE before the mean is calculated, $|EE_i$

|) (Campolongo et al., 2007). μ^*_i is useful to avoid the problem of opposite signs of μ_i canceling out.

Additionally, during the SA implementation, the scale in which a parameter is distributed is relevant for the analysis and should be specified in the methodology (Saltelli et al., 2008). For example, using the Morris (1991) method, if a parameter follows a uniform distribution, the levels (p) are obtained by dividing the interval in which the parameter varies into equal parts. However, parameter values are not sampled directly if the parameter follows another distribution. In this case, the sampling will be carried out in the space of the quantiles of the n-dimensional unit hypercube (e.g., each quantile varying in [0,1] using Morris (1991)). Then, the parameter values will be derived from its known statistical distribution.

2.2.2 Case model: OSMOSE for the Northern Humboldt Current Ecosystem

OSMOSE (Object-oriented Simulator of Marine ecOSystEms) is a multispecies individual-based modelling platform that simulates the entire life of species and their trophic interactions (Shin and Cury 2001; 2004). The basic units of OSMOSE are individuals grouped into schools, which major processes of their life cycles are modelled (growth, predation, mortality (natural, starvation, and fishing mortality), reproduction, and migration). Supporting information about the OSMOSE model can be found on its official website (http://www.osmose-model.org) and user guide (https://documentation.osmose-model.org/).

In this work, we performed a SA on the OSMOSE model applied to the Northern Humboldt Current Ecosystem (NHCE) (see Appendix A for the model description). This work aims to illustrate the use of the PR criterion in addition to identifying some of the most sensitive parameters in the NHCE OSMOSE model. However, we restricted the analysis using only the set of parameters related to the Peruvian anchovy (*Engraulis ringens*) (Table 2), because this is one of the most important species, ecologically and economically speaking, in the study area where the NHCE OSMOSE model operates. For more information about the NHCE OSMOSE model configuration, we direct the reader to Oliveros-Ramos et al., (2017) and Olivero-Ramos (2014).

Table 2. *NHCE OSMOSE parameters used for the SA. The SA was implemented using the baseline values (x) of the 27 parameters related to the Peruvian anchovy (Engraulis ringens). The baseline values are the outcomes of the reparametrization described by the f function. Appendix B provides additional information about the reparametrization and parameter scales of the NHCE OSMOSE model parameters.*

Notation	Parameter names	Reparameterization	Baseline value (x)	Scale
p_1	predation accessibility $A(anchovy, hake)$	$x = f(A) = A$ (identity)	0.1	logit
p_2	predation accessibility $A(anchovy, sardine)$		0.9	
p_3	predation accessibility $A(anchovy, jack\ mackerel)$		0.9	
p_4	predation accessibility $A(anchovy, chub\ mackerel)$		0.9	
p_5	predation accessibility $A(anchovy, mesopelagics)$		0.55	
p_6	predation accessibility $A(anchovy, munida)$		0.05	
p_7	predation accessibility $A(anchovy, humboldt\ squid)$		0.6	
p_8	predation accessibility $A(anchovy, euphausiids)$		0.6	
p_9	minimum predator-prey size ratio for the first stage of anchovy ($\theta_{stage\ 1}$)	$x = f(\theta_{stage\ 1}) = \frac{\theta_{stage\ 1}}{(\pi/2)}$	$\theta_{stage\ 1} = g(0, 1/800)$ $\frac{\theta_{stage\ 1}}{(\pi/2)} = 0.0008$	logit
p_{10}	maximum predator-prey size ratio for the first stage of anchovy ($\alpha_{stage\ 1}$)	$x = f(\alpha_{stage\ 1}) = \frac{\alpha_{stage\ 1}}{(\pi/2)}$	$\alpha_{stage\ 1} = g(1/800, 1/8)$ $\frac{\alpha_{stage\ 1}}{(\pi/2)} = 0.0784$	
p_{11}	minimum predator-prey size ratio for the second stage of anchovy ($\theta_{stage\ 2}$)	$x = f(\theta_{stage\ 2}) = \frac{\theta_{stage\ 2}}{(\pi/2)}$	$\theta_{stage\ 2} = g(0, 1/200)$ $\frac{\theta_{stage\ 2}}{(\pi/2)} = 0.0032$	
p_{12}	maximum predator-prey size ratio for the second stage of anchovy ($\alpha_{stage\ 2}$)	$x = f(\alpha_{stage\ 2}) = \frac{\alpha_{stage\ 2}}{(\pi/2)}$	$\alpha_{stage\ 2} = g(1/200, 1/6)$ $\frac{\alpha_{stage\ 2}}{(\pi/2)} = 0.1020$	
p_{13}	predator-prey size threshold (s_{thr} in cm)	$x = f(s_{thr}) = \frac{s_{thr}}{L_\infty}$	$\frac{s_{thr}}{L_\infty} = \frac{10}{19.5} = 0.5128$	logit
p_{14}	maximum starvation mortality rate ($M_{\varepsilon_{max}}$ in y^{-1})		0.10	logarithmic
p_{15}	von Bertalanffy threshold (a_{thr} in y)		0.35	logit

p_{16}	egg size (cm)	$x = f(A) = A$	0.10	logarithmic
p_{17}	critical threshold of predation efficiency (ε_{crit})		0.57	logit
p_{18}	maximum rate of predation ingestion (I_{max} in $g\ body\ g^{-1}y^{-1}$)		3.50	logarithmic
p_{19}	natural mortality rate (M in y^{-1})		0.9177	logarithmic
p_{20}	larval mortality rate (M_0 in $month^{-1}$)		9.5821	logarithmic
p_{21}	fishing mortality multiplier (f_m)		1	logarithmic
p_{22}	sex ratio ($Frac_{fem}$)		0.50	logit
p_{23}	$L_{t=0}$(von Bertalanffy growth parameter in cm)		1.9682	logarithmic
p_{24}	k (von Bertalanffy growth parameter in y^{-1})		0.76	logarithmic
p_{25}	L_∞ (von Bertalanffy growth parameter in cm)		19.50	logarithmic
p_{26}	size at maturity (s_{mat} in cm)	$x = f(s_{mat}) = \frac{s_{mat}}{L_\infty}$	$\frac{s_{mat}}{L_\infty}=\frac{12}{19.5}=0.6154$	logit
p_{27}	constant of proportionality of the allometric length-weight relationship (c in $g\ cm^{-3}$)	$x = f(A) = A$	0.0065	logarithmic

We also suggest reviewing other OSMOSE applications for complementary descriptions of model equations (e.g., Travers-Trolet et al. (2014) and Halouani et al. (2016)). Besides, Appendix B briefly describes the model equations and parameters' scales used in this work. In this regard, by using a linear scale for a parameter that follows a uniform distribution, the distribution bounds can be calculated by linearly increasing or decreasing a percentage of variation to the baseline value. Other examples include the use of logit and logarithmic scales, whose choice depends on the mathematical constraints of model parameters and equations (Appendix B). Using a linear scale does not ensure a positive numerical value when a parameter is varied around its baseline value, whereas the use of logarithmic does, and the logit scale provides parameter values only between 0 and 1. Another source of information for the scale selection is the equations used for the calibration of the model. For example, in Oliveros-Ramos et al. (2017), the larval and fishing mortalities of the NHCE OSMOSE model were calibrated in logarithmic scales. Therefore, a SA

implementation involving the parameters estimated from this calibration should consider the same scales. All of these scale considerations should be taken into account and reported in any SA implementation, with careful verification of model equations, and potential reparameterization to guarantee the compliance of mathematical constraints.

The use of the PR criterion with the NHCE OSMOSE model is presented in Figure 5, where we first classified the model parameters (Table 2) using the PR categories. In this application, the parameters ended up in four categories: high-quality data (HQD), model calibration (MC), empirical relationships (ER), and guesstimate (GS). We calculated ranges for which data was available by using equations c.1 and c.2 (in Appendix C). For the parameters without data for estimating their range of variation, the ranges were calculated following the recommendations in Figure 3.

In this model application (Figure 5), the parameter $L_{t=0}$ used the average of ranges inside the HQD category. The same strategy (i.e., the average of the range inside the category), was also used for the critical predation efficiency. Because the parameters classified as MC had no data for range calculation, we decided to calculate them by taking the average of all ranges found in the HQD and ER categories. Finally, we arbitrarily set the ranges of the last category (GS) to 10% more than the last parameter range found in the ER category (egg size).

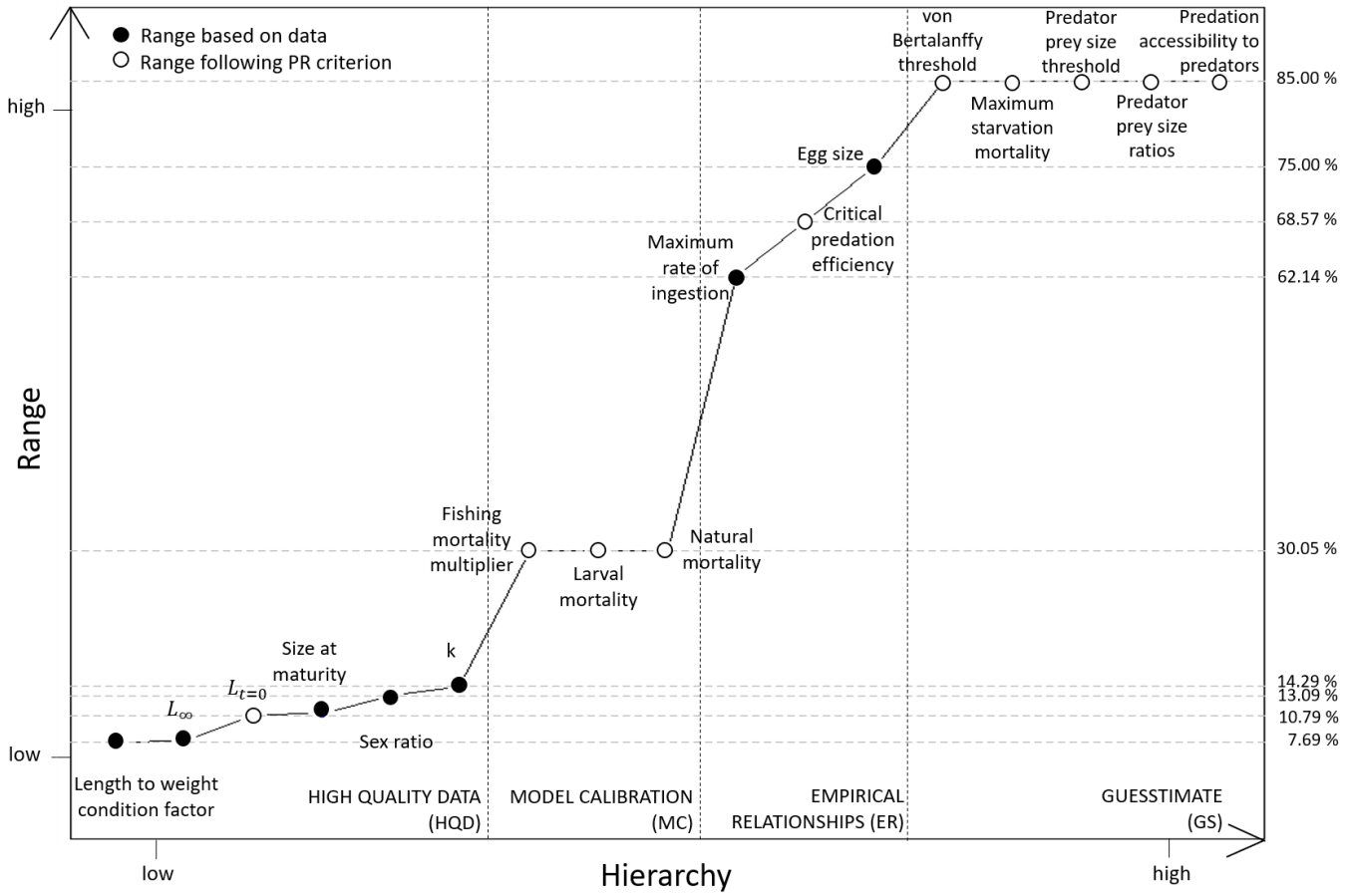

Figure 5. *Range calculation for the NHCE OSMOSE model parameters using the PR criterion. The parameters presented here are the same as in Table 2; however, we grouped some names. For example, the predation accessibilities to predators (parameters* p_1 *to* p_8 *in Table 2) are represented here with a single dot, likewise for predation size ratios (parameters* p_9 *to* p_{12} *in Table 2). Black points indicate the ranges calculated from data, and white points indicate the ranges derived from the PR criterion and associated protocol (see main text). The parameter ranges (in %) are specified on the right vertical axis.*

2.2.3 SA experiments

We ran 22 experiments for the SA of the NHCE OSMOSE model (see Table 3):

- Experiment P followed the whole PR protocol, as described in sections 2.1 and 2.2.2.
- Experiments F.x (where x takes values from 10% to 85%) assume no information on the prior distributions of any parameter (or a lack of data mining). Consequently, we used

ranges of variation arbitrarily set at 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 85% on either side of the parameters' baseline values.

- Experiments P and F.x used the parameter scales indicated in Table 2, while experiments P.Linear and F.Linear.x used linear scales for all parameters.

Table 3. *SA experiments. Twenty-two simulations were run as part of the SA experiments. Experiments P and P.Linear used ranges following the PR protocol, while experiments F.x (from F.10 to F.85) and F.Linear.x (from F.Linear.10 to F.Linear.85) used arbitrary and fixed ranges of variation. For example, in experiment F.10 all ranges were calculated at 10% on either side of parameters' baseline values. Experiments P and F.x followed the parameter scales described in Table 2, while experiments P.Linear and P.Linear.x used a linear scale for all model parameters.*

Experiments	Parameter range (in %)	Parameter scale
P	protocol	scales from Table 2
F.x	fixed to x % $x \in$ {10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%}	scales from Table 2
P.Linear	protocol	linear scale for all parameters
F.Linear.x	fixed to x %, $x \in$ {10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%}	linear scale for all parameters

We used the Morris method for each experiment (Table 3) and the transformed scale (specified in Table 2) for the sampling process. The experiments were performed using 20 trajectories (i.e., Morris replicates), dividing the corresponding parameter range into eight levels ($p = 8$) (including upper and lower bounds with values uniformly distributed between them), and using the recommended grid jump of 4/7 (where $p = 8$, and the grid jump $\Delta = p/[2(p-1)]$). Given the stochasticity of the NHCE OSMOSE model, ten model replicates were executed (i.e., OSMOSE replicates). This generated 5600 simulations for each experiment and thus a total of 123200 simulations to run. These were executed using a supercomputer (Datarmor, hosted by IFREMER, https://wwz.ifremer.fr/pcdm/Equipement).

To assess the results of the SA, we selected the total biomass of the system (i.e., the sum of focal species biomasses) among the outputs of the NHCE OSMOSE model, averaged over time and OSMOSE replicates. In addition, the total biomass was standardized, using the baseline biomass without SA perturbations.

The analysis was performed using R 3.6.1 (R Core Team 2019), and the runup R package for the SA plan and calculations of EE based on the Morris method. This package is an open-source tool available in a GitHub repository (https://github.com/CriscelyLP/run_up).

3. Results

3.1 SA in the NHCE OSMOSE model

The SA results in the NHCE OSMOSE model using the PR criterion and associated protocol show that most model parameters had σ/μ* ratios greater than 1 (Figure 6). Following the parameter classification by Sanchez et al. (2014), this means that parameters had strong interactions and/or non-linearities. The parameter p_{10} (maximum predator-prey size ratio for the first stage of anchovy) had the strongest influence, followed by p_{18} (maximum rate of predation ingestion), p_{13} (predator-prey size thresholds), and p_{17} (critical threshold of predation efficiency). Parameters with moderate influence were p_{20} (larval mortality rate), p_{9} (minimum predator-prey size ratio for the first stage of anchovy), and p_{19} (natural mortality rate). The parameters with both high mean effects and high σ/μ* ratios were classified in the guesstimate (p_{9}, p_{10}, and p_{13}) and empirical relationships (p_{17} and p_{18}) categories. Also, these parameters are related to the predation process, in particular involving the first stage of anchovy, which is one of the most critical processes in the OSMOSE modelling platform.

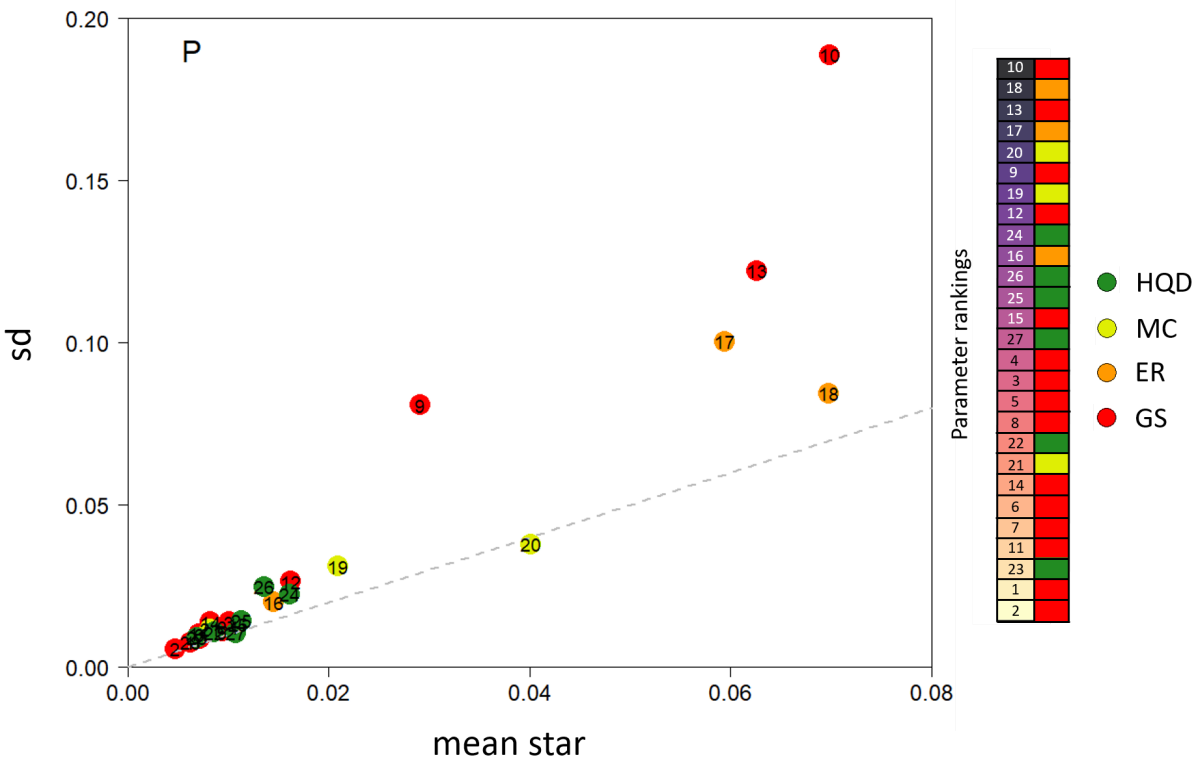

Figure 6. *Results of Morris SA of the NHCE OSMOSE model, following the PR criterion and associated protocol. Labels (inside colored dots) represent the number of the parameters (see Table 2). The point colors represent the parameter source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)). The gray dashed line represents σ/μ*=1. On the right, we classify all the parameters according to the mean star values (μ*). We rank them from the parameter with the strongest main effect (par_{10}, dark purple) to the weakest main effect (par_{2}, light yellow).*

3.2 SA: using predefined range of variations

We first compare the SA based on the PR protocol (experiment P) versus fixed parameter ranges of variation (experiments F.x, x varying from 10% to 85%). By plotting the mean μ^* (of the absolute values, called mean star) and the standard deviation σ (sd) of the Morris elementary effects, we show that in general parameters with high effects on the model outputs (high μ^*) also have strong interactions and/or nonlinear effects (high σ values) (Figure 7 and Figure D.1, Appendix D for the complete set of results). Besides, the SA results indicate notable differences between F.x experiments, with the most sensitive parameters (with the strong main effects)

changing through experiments (Figure 7). Further, the effects of parameter variations increase at a nonlinear rate, with μ^* being circumscribed to less than 0.1 for experiments F.10 to F.60 (Panel A in Figure 7) and shifting to much higher values for experiments F.70 and F.80 (panel B and Figure 7).

Using the μ^* values from the Morris SA results, the parameters tested in P and F.x (x from 10% to 85%) experiments were ranked (Figure 8) to identify the parameters with stronger effects, and verify the variability of the effects across experiments. The parameter with the strongest main effect in the PR protocol was p_{10} (the maximum predator to prey size ratio for the first life stage of anchovy). It was also found to have the strongest effect in the experiment F.50, and on average, it was ranked among the first parameters across almost all experiments (except in F.85). The parameters p_{18}, p_{19}, and p_{20} were also influential consistently through all experiments, whereas p_{13} had a strong effect only in the protocol results. By contrast, the less sensitive parameter in the protocol was p_2, the same results were found in F.30. For F.10, F.15 and F.20 experiments, parameters p_4, p_{14} and p_{15} were the least sensitive parameters, respectively. On average, p_6 was one of the less sensitive parameter for the rest of the experiments.

The protocol results (Fig. 6 and Fig. 8), where the range estimation followed the PR criterion, showed that even when parameters were classified as part of the guesstimate category (being assigned a variation range of 85%, see Figure 5), they did not all have strong main effects. The protocol results also identified parameters with strong effects (from p_{17} to p_{20}) consistently with what was found in experiments with fixed ranges. The parameters classified as high-quality data were assigned a very short range of variation (Table C.1) Despite this, some of them had a relatively high effect on the protocol experiment. Among these HQD parameters, p_{24} to p_{26} (size related parameters) had a strong effect in the F.x SA experiments as well.

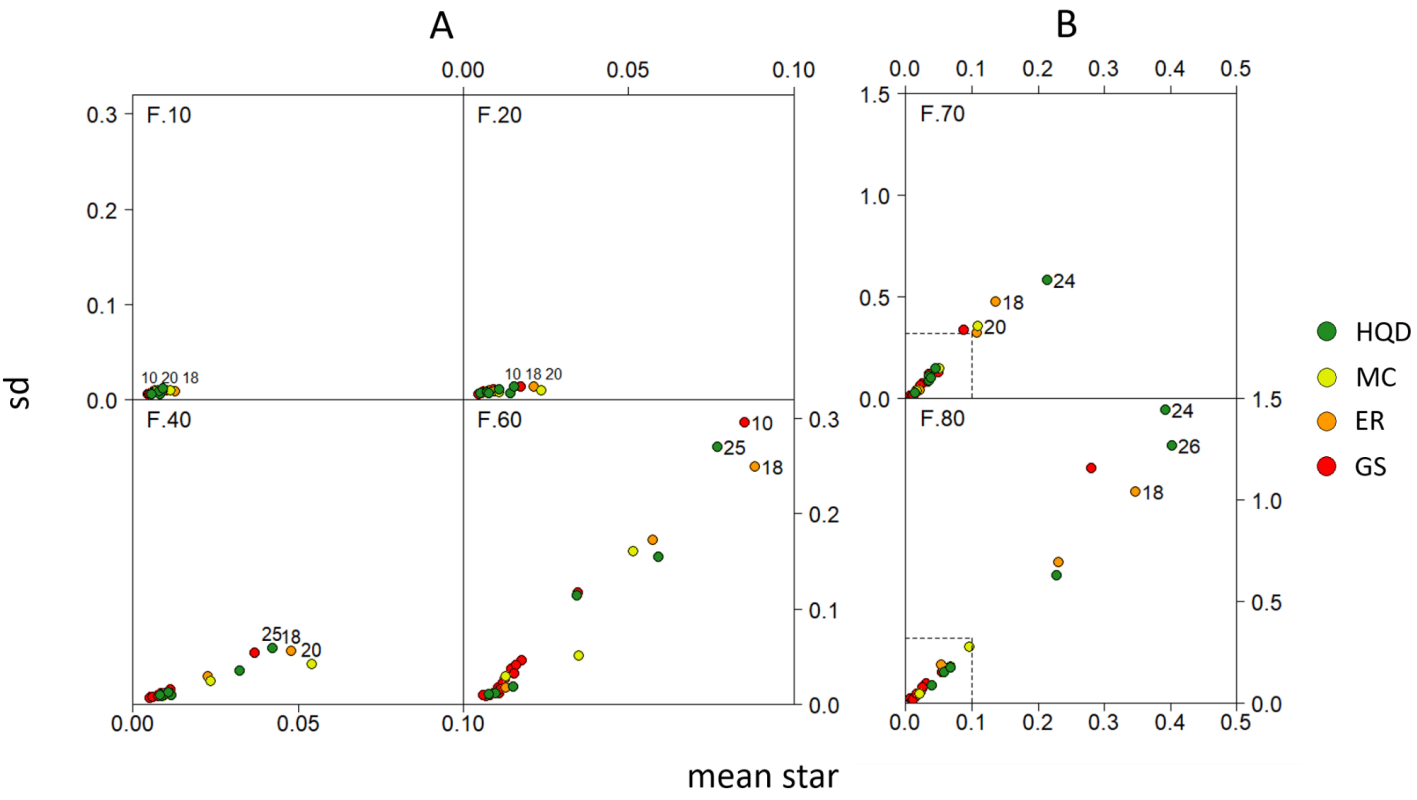

Figure 7. *Results of Morris SA of the NHCE OSMOSE model. Panel A (on the left) shows results of the experiments F.10, F.20, F.40, and F.60, while Panel B (on the right) shows results of F.70 and F.80. The point colors represent the parameter information source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)). Labels for the three most influential parameters are shown (labels from Table 2). In Panel B, gray dashed lines represent the axes used in panel A plots.*

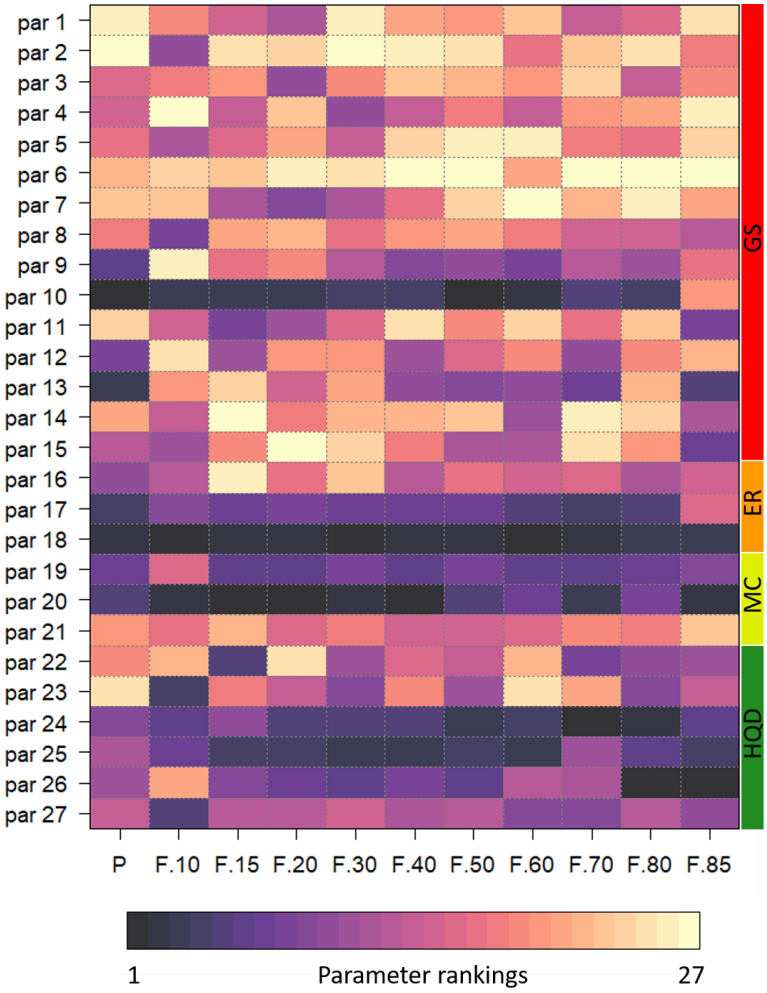

Figure 8. *Parameter rankings for P and F.x (x {10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%}) experiments. The ranking is performed using the mean star (μ^*) values from the Morris SA results. The y-axis on the left indicates the parameter reference names (used in Table 2). The vertical bar on the right y-axis indicates the parameter source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)). Rank 1 means that the parameter has the strongest main effect (with the highest μ^* value) for the corresponding experiment. On the contrary, rank 27 means that the parameter has the weakest main effect (with the lowest μ^* value).*

3.3 Effects of the scale

The SA results based on the use of linear scales for parameter sampling are presented in Figure 9 (the complete set of F.Linear.x experiments are in Figure D.2, Appendix D). These results show that the effect of the scale is highly influential for the SA implementation. Changing the parameter scale (Table 2) to a linear scale led to entirely different results for all the simulated experiments. The change in parameter scales also affects the parameters with a strong main effect (e.g., Panel A and B in Figure 9), and in general, it may incorrectly display the influence of parameters on model's outputs. For example, using the appropriate logit scale, the main effects of p_2 are moderate to weak for all experiments (Fig. 8), whereas using linear scales produce strong main effects (see experiments F.Linear.15, F.Linear.20, and F.Linear.40, in Figure D.2, Appendix D).

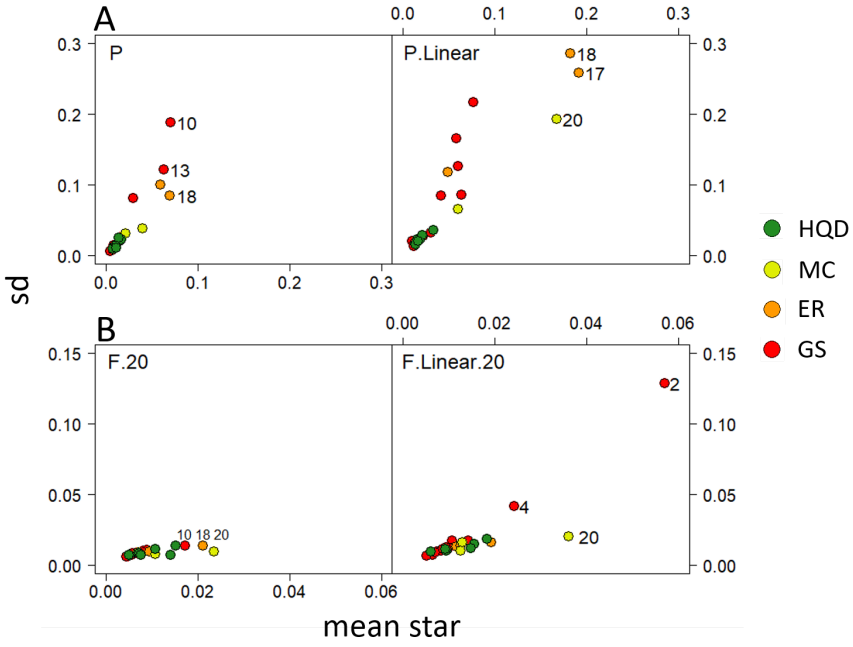

Figure 9. *Results of Morris SA evaluating the scale effect. Panel A (on the top) shows the results for P and P.Linear experiments, while Panel B (on the bottom) for F.20 and F.Linear.20 experiments. The point colors represent the parameter source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)). Labels for the three most influential parameters are shown.*

4. Discussion

4.1 The PR criterion contribution

The importance of SA for model evaluation is widely recognized (Saltelli et al., 2019, Razavi et al., 2020). However, there are still many limitations for implementing SA with complex models. In this work, we address the limitations related to the absence of data to determine the uncertainty in model inputs. In this situation, a common practice is to use fixed and arbitrary ranges of variation around the parameter reference values (e.g., in Brancis et al., 2020; Hansen et al., 2019; Specka et al., 2015). However, our results showed that this practice is not neutral at all on the outcome of the SA and could produce varying conclusions depending on the used ranges. Therefore our work aims to alert modellers to the limitations and potential biases associated with the adoption of arbitrary ranges of variation for input parameters. In addition, we provide an alternative methodology based on the parameter reliability (PR) criterion and associated protocol.

Our proposed methodology introduces the basic assumption that parameters derived from data or information with higher reliability have a smaller uncertainty than parameters derived from data with lower reliability. From this, to draw the uncertainty around the model input parameters, the PR criterion assigns smaller ranges of variation for the parameters with more reliability and wider ranges for the parameters with less reliability. From a practical angle, the PR criterion classifies the model parameters using eight predefined categories corresponding to the data and information sources that were used to determine parameters' values. The model parameters should be classified into a single category with its corresponding hierarchy that will provide insight into the reliability of the parameters. Although we strive to establish a comprehensive hierarchy of parameter categories to make the protocol as generic as possible and open it to future application in various models, new categories may be included or modified in future versions of the PR criterion.

This paper emphasizes the use of data-driven ranges of variation for SA applications. However, there are many model applications where data are missing, in which case we suggest exploring the parameter reliability criterion instead of using fixed ranges of variation for model parameters. The protocol that is described in this paper is meant to guide modellers in a stepwise manner to

avoid making arbitrary choices in the SA implementation. Although the use of the PR criterion helps the problem to a great extent, arbitrary decisions still have to be made when data is too scarce (e.g., cases E and F in Fig. 4). As part of this work, we also explain the importance of selecting parameter scales and we describe some necessary parameters' transformations which are both important steps for implementing a SA.

4.2 SA in the NHCE OSMOSE model

In this work, we used the NHCE OSMOSE model and a set of simulation experiments to illustrate the protocol based on the PR criterion. The NHCE OSMOSE model simulates the dynamics and interactions of multiple species by using a large set of input parameters, besides initial conditions and model forcing. Since this model is stochastic, it requires the execution of multiple simulations (i.e., model replicates), which increases the computational cost of its SA. Because of this, and the multiple experiments simulated in this work, we decided to use a reduced list of parameters of the NHCE OSMOSE model, by focusing on the key anchovy species, and a computationally inexpensive SA method, the Morris method (Morris 1991). However, since we used a reduced set of parameters of the NHCE OSMOSE model (27 out of hundreds), the results obtained in this study cannot be generalized as a complete sensitivity analysis of the NHCE OSMOSE model, for which the complete list of parameters should be included in the SA to identify the most influential parameters.

The Morris method is widely used in models of similar complexity to the NHCE OSMOSE model (e.g., Bracis et al., 2020). However, we also recommend that future works using Morris SA include the analysis suggested by Campolongo et al. (2007). This analysis focuses on optimizing the choice of Morris trajectories by maximizing their parsimony in the input domain. Likewise, another approach we suggest is the combined use of the Morris method as a screening analysis to identify a reduced set of the most impacting parameters, on which a numerically more expensive method like the Sobol method could be applied (Sobol, 1993). Sobol is a global SA with a higher computational cost than Morris, however this method could provide information about the amount of variance explained by each model parameter, quantifying the relative effects of parameters on model outputs and the interaction between model parameters. For example, this approach has been used in the studies of Morris et al. (2014) and Dantec-Nédélec et al. (2017).

4.3 Future directions

Discussions about the SA limitations for assessing complex ecosystem models frequently turn around the computational cost to perform SA simulations due to e.g., large number of model input parameters and the high computational requirements of the model. Although these computational issues are critical, the methodological limitations described in this work are rarely discussed. We provide some guidance to deal with data limitations to avoid the arbitrary use of parameter ranges and instead critically evaluate parameters' reliability. The set of rules proposed as part of the PR criterion and protocol are, however, only a tiny step to advance sensitivity and uncertainty analyses for complex ecosystem models.

5. Acknowledgments

The authors acknowledge the Pôle de Calcul et de Données Marines (PCDM, https://wwz.ifremer.fr/pcdm/Equipement) of Ifremer for providing DATARMOR storage and computational resources. CL was supported by an Peruvian individual doctoral research grant from the "Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica" (contract n° 134-2016-FONDECYT). YJS acknowledges funding support from the Biodiversa and Belmont Forum project SOMBEE (BiodivScen ERA-Net COFUND programme, ANR contract n°ANR-18-EBI4-0003-01), the Pew marine fellows programme, and the European Union's Horizon 2020 research and innovation programme under grant agreement No 817578 (Triatlas).

6. References

1. Alkemade, Rob, Mark van Oorschot, Lera Miles, Christian Nellemann, Michel Bakkenes, and Ben Ten Brink. 2009. "GLOBIO3: A Framework to Investigate Options for Reducing Global Terrestrial Biodiversity Loss." *Ecosystems* 12 (3): 374–90. https://doi.org/10.1007/s10021-009-9229-5.
2. Aumont, Oliver, and Laurent Bopp. 2006. "Globalizing Results from Ocean in Situ Iron Fertilization Studies." Global Biogeochemical Cycles 20 (2).
3. Bracis, Chloe, Sigrid Lehuta, Marie Savina-Rolland, Morgane Travers-Trolet, and Raphaël Girardin. 2020. "Improving Confidence in Complex Ecosystem Models: The Sensitivity Analysis of an Atlantis Ecosystem Model." Ecological Modelling 431: 109133.
4. Campolongo, F, S Tarantola, and A Saltelli. 1999. "Tackling Quantitatively Large Dimensionality Problems." Computer Physics Communications 117 (1-2): 75–85.
5. Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, *22*(10), 1509-1518.
6. Cariboni, J, D Gatelli, R Liska, and A Saltelli. 2007. "The Role of Sensitivity Analysis in Ecological Modelling." Ecological Modelling 203 (1-2): 167–82. https://doi.org/10.1016/j.ecolmodel.2005.10.045.

7. Castillo Alva, J. E. (2012). Determinación de la proporción sexual y su relación con la hora pico de desove de la anchoveta engraulis ringes (Jenyns, 1842) durante el periodo 2000-2009 en la región norte-centro del Perú.
8. Cheung, William WL, Thomas L Frölicher, Rebecca G Asch, Miranda C Jones, Malin L Pinsky, Gabriel Reygondeau, Keith B Rodgers, et al. 2016. "Building Confidence in Projections of the Responses of Living Marine Resources to Climate Change." ICES Journal of Marine Science 73 (5): 1283–96. https://doi.org/10.1093/icesjms/fsv250.
9. Christensen, V., Walters, C.J. & Pauly, D. (2005). Ecopath with Ecosim: A User's Guide. Fisheries Centre, University of British Columbia, Vancouver. 154 pp.
10. Cury, P., & Pauly, D. (2000). Patterns and propensities in reproduction and growth of marine fishes. *Ecological Research*, *15*(1), 101-106.
11. Dantec-Nédélec, Sarah, Catherine Ottlé, T Wang, F Guglielmo, Fabienne Maignan, Nicolas Delbart, V Valdayskikh, et al. 2017. "Testing the Capability of Orchidee Land Surface Model to Simulate Arctic Ecosystems: Sensitivity Analysis and Site-Level Model Calibration." Journal of Advances in Modeling Earth Systems 9 (2): 1212–30.
12. Echevin, Vincent, Olivier Aumont, Jesus Ledesma, and G Flores. 2008. "The Seasonal Cycle of Surface Chlorophyll in the Peruvian Upwelling System: A Modelling Study." Progress in Oceanography 79 (2-4): 167–76.
13. Ferrier, Simon, Karachepone N Ninan, Paul Leadley, Rob Alkemade, Lilibeth A Acosta, H Resit Akçakaya, Lluis Brotons, et al. 2016. "IPBES (2016): The Methodological Assessment Report on Scenarios and Models of Biodiversity and Ecosystem Services." Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services. Bonn, Germany, 348.
11. Fulton, Elizabeth A, Anthony DM Smith, and Craig R Johnson. 2003. "Effect of Complexity on Marine Ecosystem Models." Marine Ecology Progress Series 253: 1–16.
14. Funtowicz, Silvio O, and Jerome R Ravetz. 1990. Uncertainty and Quality in Science for Policy. Vol. 15. Springer Science & Business Media.
15. Gislason, H., & Helgason, T. (1985). Species interaction in assessment of fish stocks with special application to the North Sea. *Dana*, *5*(2), 1-44.
16. Halouani, Ghassen, Frida Ben Rais Lasram, Yunne-Jai Shin, Laure Velez, Philippe Verley, Tarek Hattab, Ricardo Oliveros-Ramos, et al. 2016. "Modelling Food Web Structure Using an End-to-End Approach in the Coastal Ecosystem of the Gulf of Gabes (Tunisia)." Ecological Modelling 339: 45–57. https://doi.org/10.1016/j.ecolmodel.2016.08.008.
17. Hansen, Cecilie, Kenneth F Drinkwater, Anne Jähkel, Elizabeth A Fulton, Rebecca Gorton, and Mette Skern-Mauritzen. 2019. "Sensitivity of the Norwegian and Barents Sea Atlantis End-to-End Ecosystem Model to Parameter Perturbations of Key Species." PloS One 14 (2): e0210419.
18. Holland, RA, WRT Darwall, and KG Smith. 2012. "Conservation Priorities for Freshwater Biodiversity: The Key Biodiversity Area Approach Refined and Tested for Continental Africa." *Biological Conservation* 148 (1): 167–79. https://doi.org/10.1016/j.biocon.2012.01.016.
19. Jørgensen, S. E., & Bendoricchio, G. (2001). Fundamentals of ecological modelling (Vol. 21). Elsevier.
20. Link, Jason S, TF Ihde, CJ Harvey, Sarah K Gaichas, JC Field, JKT Brodziak, HM Townsend, and RM Peterman. 2012. "Dealing with Uncertainty in Ecosystem Models: The Paradox of Use for Living Marine Resource Management." Progress in Oceanography 102: 102–14. https://doi.org/10.1016/j.pocean.2012.03.008.
21. Morris, David J, Douglas C Speirs, Angus I Cameron, and Michael R Heath. 2014. "Global Sensitivity Analysis of an End-to-End Marine Ecosystem Model of the North Sea: Factors Affecting the Biomass of Fish and Benthos." Ecological Modelling 273: 251–63.
22. Morris, Max D. 1991. "Factorial Sampling Plans for Preliminary Computational Experiments." Technometrics 33 (2): 161–74. https://doi.org/10.2307/1269043.
23. Ochoa Madrid, M., Marín Abanto, P., & Bouchon Corrales, M. (2020). Estimación de la pérdida de peso de la anchoveta peruana Engraulis ringens en cruceros de evaluación de recursos pelágicos.
24. Oliveros Ramos, R., Guevara Carrasco, R., Simmond, J., Csirke, J., Gerlotto, F., Peña Tercero, C., & Tam Málaga, J. (2010). Modelo de evaluación integrada del stock norte-centro de la anchoveta peruana Engraulis ringens Jenyns.
25. Oliveros Ramos, Ricardo. 2014. "End-to-End Modelling for an Ecosystem Approach to Fisheries in the Northern Humboldt Current Ecosystem." PhD Thesis, France: University of Montpellier 2.
26. Oliveros-Ramos, Ricardo, Philippe Verley, Vincent Echevin, and Yunne-Jai Shin. 2017. "A Sequential Approach to Calibrate Ecosystem Models with Multiple Time Series Data." Progress in Oceanography 151: 227–44. https://doi.org/10.1016/j.pocean.2017.01.002.
27. Pannell, D. J. (1997). Sensitivity analysis of normative economic models: theoretical framework and practical strategies. *Agricultural economics*, *16*(2), 139-152.

28. Payne, Mark R, Manuel Barange, William WL Cheung, Brian R MacKenzie, Harold P Batchelder, Xochitl Cormon, Tyler D Eddy, et al. 2015. “Uncertainties in Projecting Climate-Change Impacts in Marine Ecosystems.” ICES Journal of Marine Science 73 (5): 1272–82. https://doi.org/10.1093/icesjms/fsv231.
29. R Core Team, 2019. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
30. Regan, Helen M, Mark Colyvan, and Mark A Burgman. 2002. “A Taxonomy and Treatment of Uncertainty for Ecology and Conservation Biology.” Ecological Applications 12 (2): 618–28. https://doi.org/10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2.
31. Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., ... & Maier, H. R. (2021). The future of sensitivity analysis: an essential discipline for systems modeling and policy support. *Environmental Modelling & Software*, *137*, 104954.
32. Sanchez, D. G., Lacarrière, B., Musy, M., & Bourges, B. (2014). Application of sensitivity analysis in building energy simulations: Combining first-and second-order elementary effects methods. *Energy and Buildings*, *68*, 741-750.
33. Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity Analysis. John Wiley & Sons, Ltd.
34. Saltelli, Andrea, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. 2008. Global Sensitivity Analysis: The Primer. John Wiley & Sons.
35. Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., et al. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental modelling & software*, *114*, 29-39.
36. Shin, Yunne-Jai, and Philippe Cury. 2001. “Exploring Fish Community Dynamics Through Size-Dependent Trophic Interactions Using a Spatialized Individual-Based Model.” Aquatic Living Resources 14 (2): 65–80.
37. Shin, Yunne-Jai, and Philippe Cury. 2004. “Using an Individual-Based Model of Fish Assemblages to Study the Response of Size Spectra to Changes in Fishing.” Canadian Journal of Fisheries and Aquatic Sciences 61 (3):414–31. https://doi.org/10.1139/f03-154.
38. Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–414.
39. Song, X., Bryan, B. A., Almeida, A. C., Paul, K. I., Zhao, G., & Ren, Y. (2013). Time-dependent sensitivity of a process-based ecological model. *Ecological modelling*, *265*, 114-123.
40. Specka, Xenia, Claas Nendel, and Ralf Wieland. 2015. “Analysing the Parameter Sensitivity of the Agro-Ecosystem Model Monica for Different Crops.” European Journal of Agronomy 71: 73–87.
41. Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., ... & Walker, N. D. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1. 0. *Geoscientific Model Development*, *11*(4), 1421-1442.
42. Tranter, M., De Lucia, M., & Kühn, M. (2021). Numerical investigation of barite scaling kinetics in fractures. *Geothermics*, *91*, 102027.
43. Travers, M., Shin, Y. J., Jennings, S., & Cury, P. (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. *Progress in oceanography*, *75*(4), 751-770.
44. Travers-Trolet, M, YJ Shin, and JG Field. 2014. “An End-to-End Coupled Model Roms-N2p2z2d2-Osmose of the Southern Benguela Foodweb: Parameterisation, Calibration and Pattern-Oriented Validation.” African Journal of Marine Science 36 (1): 11–29. https://doi.org/10.2989/1814232X.2014.883326.
45. Wallach, D., Makowski, D., Jones, J.W. (Eds.), 2006. Working with Dynamic Crop Models. Elsevier.

Appendix A: Description of NHCE-OSMOSE model

The NHCE OSMOSE (Oliveros-Ramos, 2014) results from the coupling of the OSMOSE model of fish and macroinvertebrate dynamics to the ROMS-PISCES hydrodynamic and biogeochemical model (Aumont et al. 2003, Echevin et al. 2012). The NHCE OSMOSE model covers the extension of the Northern Humboldt Current Ecosystem (NHCE) and the Peruvian upwelling ecosystem, between the ranges from 20°S to 6°N and 93°W to 70°W with 1/6° of spatial resolution. It is fitted to data time series from 1992 to 2008.

Species (or groups) explicitly modeled in OSMOSE are called "focal species" (or focal groups). Additionally, the OSMOSE model is forced by a second type of functional group (called "biotic resource"), providing extra food resources to the model system. Biotic resources are data-driven or output from other models. The NHCE OSMOSE model includes 13 species (or functional groups), of which 9 as focal species, including 1 macrozooplankton, 1 crustacean, 1 cephalopod, and 6 fish species. The NHCE OSMOSE model considered four plankton groups as biotic resources represented in the PISCES model (Aumont and Bopp, 2006) coupled to the ROMS model in the NHCE (Echevin et al., 2008). The list of these species or functional groups is presented in Table A.1.

Table A.1. Species modeled in the NHCE OSMOSE model.

Types of model species	Group	Species or functional group	Scientific name	Model
Biotic resources	Phytoplankton	Nanophytoplankton	-	ROMS-PISCES
	Phytoplankton	Diatoms	-	ROMS-PISCES
	Zooplankton	Microzooplankton	-	ROMS-PISCES
	Zooplankton	Mesozooplankton	-	ROMS-PISCES
Foc0al species	Zooplankton	Euphausiids	*Euphausia mucronata*	OSMOSE
	Small pelagics	Anchovy	*Engraulis ringens*	OSMOSE
	Small pelagics	Sardine	*Sardinops sagax*	OSMOSE
	Medium pelagics	Jack mackerel	*Trachurus murphyi*	OSMOSE
	Medium pelagics	Chub mackerel	*Scomber japonicus*	OSMOSE
	Other pelagics	Mesopelagics	*Vinciguerria sp.*	OSMOSE
	Other pelagics	Red lobster	*Pleuroncodes monodon*	OSMOSE

	Other pelagics	Humboldt squid	*Dosidicus gigas*	OSMOSE
	Demersal	Peruvian hake	*Merluccius gayi peruanus*	OSMOSE

Appendix B: Parameterization of the NHCE-OSMOSE model

Here, we briefly describe the NHCE OSMOSE parameters selected for the SA.

- Predation parameters: The predation accessibilities are used in the model to describe the accessibility of prey to predators. These parameters reflect the vertical overlap of predators and prey in the water column. Predation accessibilities take values between 0 and 1, and determine the total accessible biomass (P_{tot_s}) for the predator school (s):

$$P_{tot_s} = \sum_{p=1}^{p=N} A(pred_s, prey_p)\, B_{prey_p} \quad \text{(b.1)}$$

where p is the preys school (p from 1 to N) available for the school (s), $A(pred_s, prey_p)$ is the accessibility of each prey p to the predator school s, and B_{prey_p} the corresponding biomass of each prey p. In addition, the total biomass that a predator school can eat ($P_{eatable_s}$) depends on the biomass of the predator (B_{pred_s}), and the maximum rate of ingestion of the predator species (I_{max}):

$$P_{eatable_s} = B_{pred_s} I_{max} \quad \text{(b.2)}$$

We decided to run the SA on the $A(pred_s, prey_p)$ parameters (predation accessibilities parameters, in Table 2 the parameters with references from p_1 to p_8). These values are proportions $\in [0, 1]$, so we chose a logit scale for these parameters. We also run the SA on the I_{max} (Table 2 the reference parameter p_{18}). Since the only constraint is to keep positive the parameter value, we used the logarithmic scale for I_{max}.

- Feeding size ranges: In OSMOSE, the predation is controlled by the minimum (R_{min}) and maximum (R_{max}) size ratios (in Table 2 parameters with the reference from p_9 to p_{12}). Predator schools can only feed on prey schools whose lengths meet these thresholds:

$$R_{min} \leq \frac{L_{pred}}{L_{prey}} \leq R_{max} \quad \text{(b.3)}$$

where the ratios (R_{min} and R_{max}) are the threshold values for predator length (L_{pred}) over prey length (L_{prey}) ratio. When information is available, a split around the size threshold (s_{thr}) allows to specify different feeding size range between species stages (the parameter called predator-prey size threshold, e.g., the p_{13} in Table 2). Predator-prey size threshold (s_{thr}) is part of the SA. However, we re-parameterized it by s_{thr}/L_{∞} in logit scale (with values from 0 to 1). It ensures that s_{thr} will always be smaller or equal to L_{∞}. Likewise, R_{min} should never be higher than R_{max}. Manipulations of R_{min} and R_{max} separately for the SA are risky and should be avoided for model consistency. We thus decided to re-parametrize the equation of this process and implement the SA using angles in polar coordinates. Users should think carefully about the potential consequences of changing parameter values that are interdependent in the model equations and re-parametrize them when necessary. Using equation b.3 in cartesian coordinates (predator size in abscissa and prey size in ordinate), we transformed it into polar coordinates calculating angles between the abscissa, minimum and maximum prey length:

$$g(m1, m2) = tan^{-1}\left(\frac{m_2 - m_1}{1 + m_1 m_2}\right) \quad \text{(b.4)}$$

where $g(m1, m2)$ is the angle (in degrees) between the slopes (m_1 and m_2), m_2 being always greater than m_1. Using equation b.4, we calculated two angles: i) the angle between the horizontal axis and the minimum prey length ($\theta = g(0, m1)$), and the angle between the minimum and maximum prey length ($\alpha = g(m1, m2)$). The parameters m_1 and m_2 were re-parametrized as the angles θ and α for the first ($\theta_{stage\,1}$ and $\alpha_{stage\,1}$) and second ($\theta_{stage\,2}$ and $\alpha_{stage\,2}$) stages of anchovy (Table 2). The new parameterization physically restricts the parameter values to the first quadrant, representing all the possibilities for the predation-prey size ratios (particularly, π/4 means predator and prey size can be equal while π/2 means there is no limit on the prey size for a predator). These values (as part of the re-parameterization) are presented in Table 2. For the SA, the angles θ and α (as fractions of π/2) use a logit scale.

- Starvation mortality: When the predation efficiency (ε_i) is below the critical threshold (ε_{crit}, p_{17} in Table 2), schools do not have enough food to fulfill their requirements undergoing starvation mortality (M_{ε_i} of

school i, see equation b.5). This process is controlled by the maximum starvation mortality rate ($M_{\varepsilon_{max}}$) of the species as follows:

$$M_{\varepsilon_i} = -\frac{M_{\varepsilon_{max}}}{\varepsilon_{crit}} . \varepsilon_i + M_{\varepsilon_{max}} \quad \text{(b.5)}$$

We implemented the SA using $M_{\varepsilon_{max}}$ (p_{14} in Table 2) on a logarithmic scale since its values are positive. ε_{crit} is also part of the SA, but we use a logit scale because its values accept only numbers between 0 and 1.

- Growth: Individuals of a given school are assumed to grow in size and weight (at time *t)* only when the amount of food they ingested fulfills their maintenance requirements. A simple linear model is used below the threshold age (a_{thr}, p_{15} in Table 2). Above a_{thr}, the von Bertalanffy model is used to determine the average mean length increase, as follows:

$$L_t = L_\infty (1 - exp(-k(t - t_0)) \quad \text{(b.6)}$$

where L_∞ is the asymptotic size (p_{25} in Table 2), k the growth coefficient (p_{24} in Table 2), and t_0 the theoretical age when size is 0. The SA is implemented using L_∞, k, $L_{t=0}$ (the length at t = 0, p_{23} in Table 2). These parameters are manipulated individually and on a logarithmic scale since the only constraint is to keep the parameter value positive. a_{thr} is also part of the SA on a logit scale, because this parameter accepts values only between 0 and 1.

The weight of a school is calculated using the allometric relationship:

$$W = cL^b \quad \text{(b.7)}$$

where b is the exponent and c the constant of proportionality of the allometric length-weight relationship. We chose to constrain the SA to the constant of proportionality (c, in Table 2 the p_{27}) in logarithmic scale, leaving b fixed.

- Size and weight of eggs: The life cycle of each species is modeled starting with the egg stage. Egg size is part of the SA (p_{16} in Table 2) and it co-varies with egg weight. For each SA simulation, we calculated the corresponding egg weight as a function of egg size, keeping the mean density constant as follows:

$$egg_{weight} = \frac{4}{3}\pi(\frac{egg_{size}}{2})^{3} mean_{density} \quad \text{(b.8)}$$

The $mean_{density}$ is calculated using the baseline values for egg size and weight values. Then each time that the egg_{size} takes a new value due to the SA, the egg_{weight} should be recalculated, meaning that both parameters are changing together. The only mathematical constraint for these parameters is to obtain a positive numerical value, so we used the logarithmic scale.

- Sources of mortality: The abundance of a school *i* at time $t + \Delta t$, ($N_{i,t+\Delta t}$), is modeled as a function of abundance at time *t* ($N_{i,t}$) and the following sources of mortality:

$$N_{i,t+\Delta t} = N_{i,t}\, e^{-\Delta t(F_s(t) + M_s(t) + M_{\epsilon_i}(t) + P_i(t))} \quad \text{(b.9)}$$

where F_s is the fishing mortality rate and M_s the natural mortality rate of species *s*. M_{ϵ_i} is the starvation mortality rate and P_i the predation mortality rate, both related to the fish school *i*.

The natural mortality rate (p_{19} in Table 2) and starvation mortality (already described in equation b.5) are part of the SA. The fishing mortality rate used in the NHCE OSMOSE model is provided by time (t) and species size (in the case of anchovy). We decided to include the fishing mortality in the SA by creating a fishing mortality multiplier f_m (p_{21} in Table 2). It is a positive value with a baseline value equal to 1. This fishing multiplier is used in the SA, by multiplying the fishing mortality rate (matrix of the fishing mortalities, time vs. species size).

Another source of mortality is the larval mortality rate ($M_{0,s}$). This parameter controls the number of eggs and larvae for species *s*, at age *0*, and at time $t + \Delta t$, ($N_{s,0,t+\Delta t}$), as follows:

$$N_{s,0,t+\Delta t} = N_{s,0,t}\, e^{-\Delta t M_{0,s}} \quad \text{(b.10)}$$

where $N_{s,0,t}$ represents the number of eggs and larvae in the system at time *t*, and $M_{0,s}$ is the larval mortality rate by species (p_{20} in Table 2). This parameter is a vector (one larval mortality rate by time step) in the NHCE OSMOSE model. The SA takes this parameter using the mean of the vector (mean over time).

Baseline values of natural mortality rate, larval mortality, and fishing mortality were estimated by a calibration process following a logarithmic scale. We chose the same scale to perform the experiments as part of the SA.

- Reproduction: any school whose length is greater than the size at sexual maturity (s_{mat}, p_{26} in Table 2) enters the reproduction process at the end of the time step. This allows the generation of a new school at the egg stage. For a given species:

$$N_{eggs} = Frac_{fem} \alpha \, season \, B_{mat} \quad \text{(b.11)}$$

where $Frac_{fem}$ is the fraction of females (also called sex ratio, p_{22} in Table 2), α is the relative fecundity. This parameter indicates the number of eggs per gram of mature female. The $season$ is a file that provides the spawning distribution within a year, and B_{mat} is the sum of the biomass of the schools that reached sexual maturity.

$Frac_{fem}$ is part of the SA. We chose the logit scale because this parameter only takes values from 0 to 1. The size at sexual maturity (s_{mat}) is also part of the SA. However, we re-parameterized it by s_{mat}/ L_{∞} in logit scale (with values from 0 to 1). It ensures that s_{mat} will always be smaller or equal to L_{∞}, but never bigger than L_{∞}.

Appendix C: Range information for the NHCE OSMOSE parameters

The calculation of parameter ranges based on data was performed as follows. For example, considering a parameter such as the asymptotic length of a fish species (in cm and linear scale), if the range estimated from the literature is between 10 cm (lower limit, L_l) and 60 cm (upper limit, L_u), with a baseline value of 40 cm (the value used in the model), we can express that the parameter has a lower bound of -75% of the baseline value and an upper bound of +50% of the baseline value. These bounds are to be provided for each parameter included in the SA (e.g., see Table 1 in Tranter et al., 2021). Using the parameter bounds, we can also express the parameter range as:

$$R_l = \left[\frac{L_l}{mean\,(L_l, L_u)} - 1 \right] 100 \qquad \text{(c.1)}$$

$$R_u = \left[\frac{L_u}{mean\,(L_l, L_u)} - 1 \right] 100 \qquad \text{(c.2)}$$

where R_l and R_u are the lower (l) and upper (u) ranges in percentage (%). These equations express the same parameters bounds found in the literature but representing the range around the mean of both limits (L_l and L_u). Using the numerical example above related to the asymptotic length, and equations c.1 and c.2, R_l equals -71.43% and R_u equals 71.43%. Applying the PR criterion, the range of variation (R) is the absolute value of R_l or R_u that we calculated for all the parameters for which data is available.

Table C.1. *Range values for the NHCE OSMOSE parameters. Using the same set of parameters presented in Table 2, we provide additional information related to the use of the PR criterion and range estimation. In this application, only four parameter sources were identified: HQD (high-quality data), MC (model calibration), ER (empirical relationships), and GS (guesstimated). When data were available, this information was used for range estimation (R in %). When data were not available, we followed the steps presented in Figure 3 for range estimation. This information was used to build Figure 5.*

Reference	Parameters used in the SA	Parameter source	Range information	Baseline value (x)	R (%)	R source
p_{27}	$c\ (g\ cm^{-3})$	HQD	[0.006, 0.007] (Ochoa et al., 2020)	0.0065	7.69	data
p_{25}	$L_\infty\ (cm)$	HQD	[19.16, 22.4] (Goicochea, data)	19.50	7.80	data

p_{23}	$L_{t=0}(cm)$	HQD	-	1.9682	10.79	protocol
p_{26}	$\frac{s_{mat}}{L_{\infty}}$	HQD	[10, 12.5] (Fishbase)	0.6154	11.11	data
p_{22}	$Frac_{fem}$	HQD	[0.498, 0.648] (Castillo, 2012)	0.50	13.09	data
p_{24}	$k\,(y^{-1})$	HQD	[0.6, 0.8] (Oliveros et al., 2010)	0.76	14.30	data
p_{21}	f_m	MC	-	1	30.05	protocol
p_{20}	M_0 $(month^{-1})$	MC	-	9.5821	30.05	protocol
p_{19}	$M\,(y^{-1})$	MC	-	0.9177	30.05	protocol
p_{18}	I_{max} $(g\ body\ g^{-1}y^{-1})$	ER	[1.56, 6.68] (Gislason and Helgason 1985)	3.50	62.14	data
p_{17}	ε_{crit}	ER	-	0.57	68.57	protocol
p_{16}	*egg size* (cm)	ER	[0.04, 0.28] (Cury and Pauly, 2000)	0.10	75	data
p_{15}	$a_{thr}\,(y)$	GS	-	0.35	85	protocol
p_{14}	$M_{\varepsilon_{max}}\,(y^{-1})$	GS	-	0.10	85	protocol
p_{13}	predator-prey size threshold ($\frac{s_{thr}}{L_{\infty}}$, cm)	GS	-	0.5128	85	protocol
$p_9 - p_{12}$	predator-prey size ratios of anchovy	GS	-	$\frac{\theta_{stage\,1}}{(\pi/2)} = 0.0008$ $\frac{\alpha_{stage\,1}}{(\pi/2)} = 0.0784$ $\frac{\theta_{stage\,2}}{(\pi/2)} = 0.0032$ $\frac{\alpha_{stage\,2}}{(\pi/2)} = 0.1020$	85	protocol
$p_1 - p_8$	predation accessibility of prey to predator	GS	-	$A(anchovy,\ hake) = 0.1$ $A(anchovy,\ sardine) = 0.9$ $A(anchovy,\ jack\ mackerel) = 0.9$ $A(anchovy,\ chub\ mackerel) = 0.9$	85	protocol

				$A(anchovy,\ mesopelagics) = 0.55$ $A(anchovy,\ munida) = 0.05$ $A(anchovy,\ humboldt\ squid) = 0.6$ $A(anchovy,\ euphausiids) = 0.6$		

Appendix D: Additional results

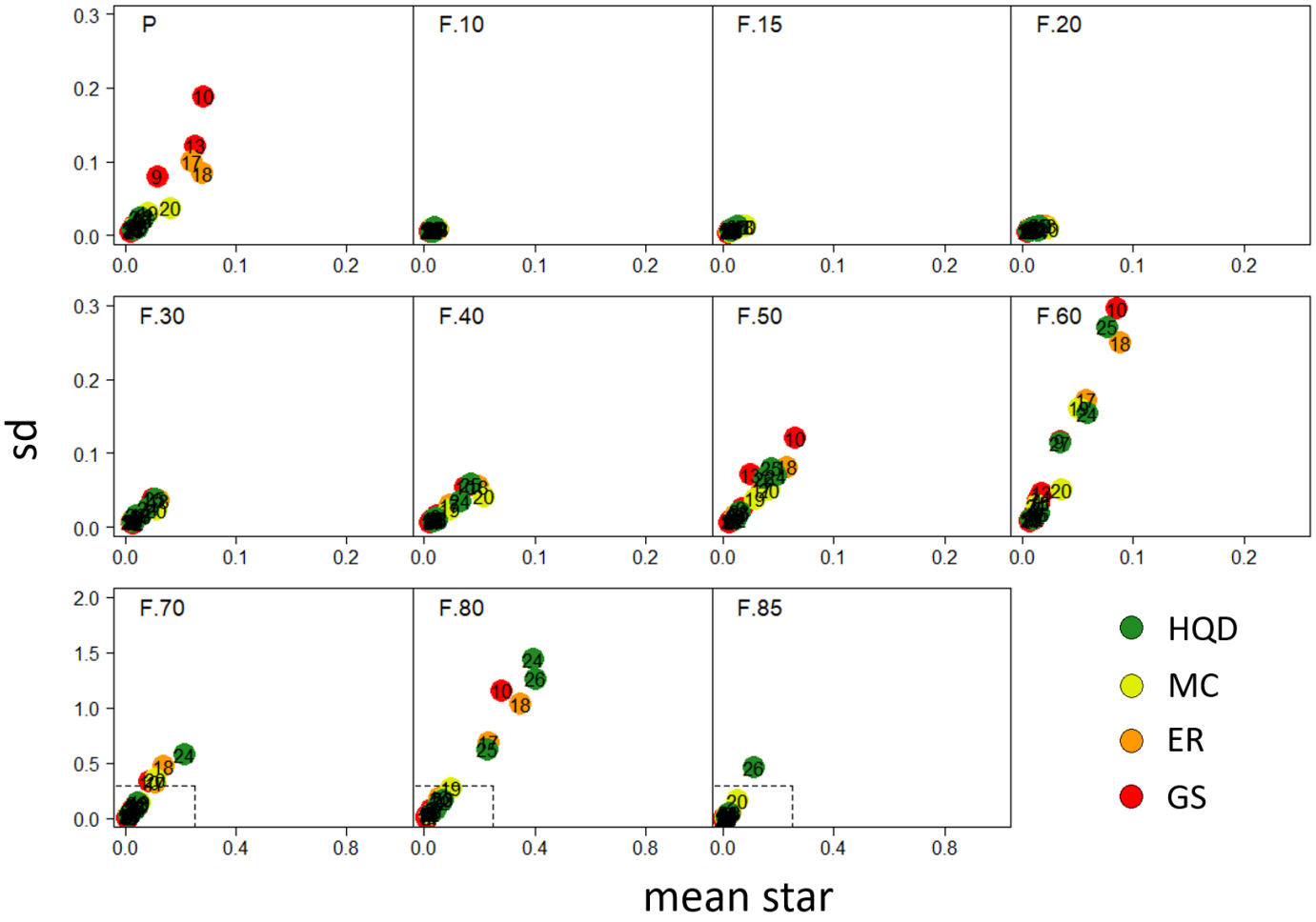

Figure D.1. *Results of Morris SA related to the P and F.x experiments (x {10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%}). The point colors represent the parameter source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)); and despite the use of fixed ranges (i.e., F experiments) we used the same colors for those experiments. Labels (inside colored dots) represent the reference number of the parameter (see Table 2). The axes of the first eight experiments are drawn as black dotted lines in the plots of the experiments F.70, F.80, and*

F.85.

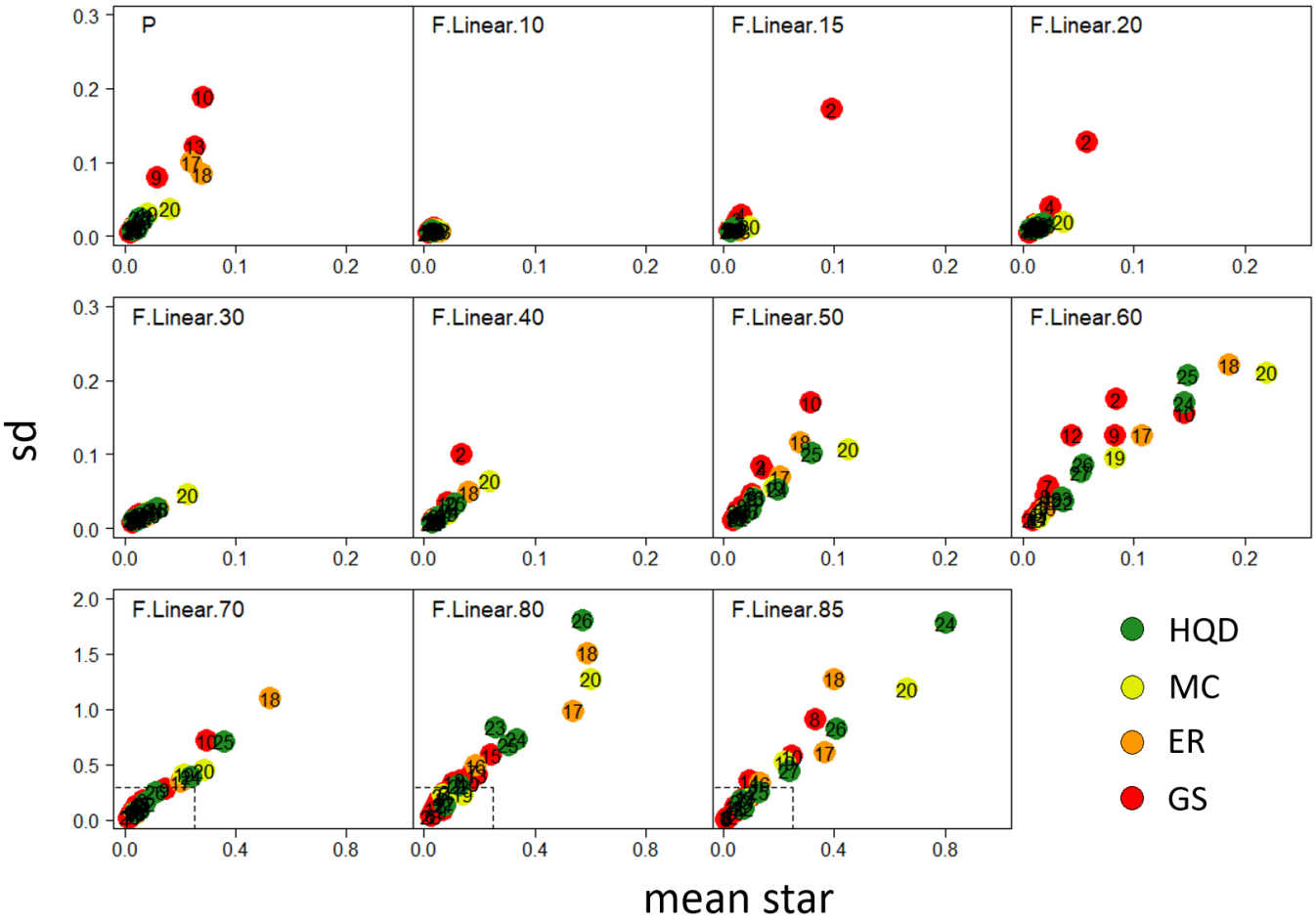

Figure D.2. *Results of Morris SA related to the P and F.Linear.x experiments (x {10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%}). The point colors represent the parameter source using the PR criterion (green for high-quality data (HQD), yellow for model calibration (MC), orange for empirical relationships (ER), and red for guesstimate parameters (GS)); and despite the use of fixed ranges (i.e., F experiments) we used the same colors for those experiments. Labels (inside colored dots) represent the reference number of the parameters (see Table 2). The axes of the first eight experiments are drawn as black dotted lines in the plots of the experiments F.70, F.80, and F.85.*

Appendix A

One Earth

Gold Standard for the Global Goals | 1 ton of CO_2 offset

CellPress

Review

Identifying uncertainties in scenarios and models of socio-ecological systems in support of decision-making

Mark D.A. Rounsevell,[1,2,*] Almut Arneth,[1] Calum Brown,[1] William W.L. Cheung,[3] Olivier Gimenez,[4] Ian Holman,[5] Paul Leadley,[6] Criscely Luján,[6,7] Stéphanie Mahevas,[8] Isabelle Maréchaux,[9] Raphaël Pélissier,[9] Peter H. Verburg,[10,11] Ghislain Vieilledent,[9] Brendan A. Wintle,[12] and Yunne-Jai Shin[7,13]

[1]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research/Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, and the Department of Geo-ecology (IFGG), Karlsruhe, Germany
[2]School of GeoSciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK
[3]Changing Ocean Research Unit, Institute for the Oceans and Fisheries, the University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[4]CEFE, CNRS, University Montpellier, EPHE, IRD, University Paul Valéry Montpellier 3, Montpellier, France
[5]Cranfield Water Science Institute, Cranfield University, Bedford MK43 0AL, UK
[6]Laboratoire d'Ecologie Systématique Evolution, Université Paris-Saclay, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France
[7]Institut de Recherche pour le Développement, MARBEC (Univ Montpellier, IRD, Ifremer, CNRS), Montpellier, France
[8]EMH, Ifremer, Centre Atlantique, Rue de l'Ile d'Yeu, BP 21105, 44311 Nantes Cedex 03, France
[9]AMAP Lab, University of Montpellier, CIRAD, CNRS, INRAE, IRD, 34000 Montpellier, France
[10]Institute for Environmental Studies, VU University Amsterdam, de Boelelaan 1087, 1081 HV Amsterdam, the Netherlands
[11]Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland
[12]School of Ecosystem and Forest Science, University of Melbourne, Melbourne, VIC 3010, Australia
[13]University of Cape Town, Cape Town, South Africa
*Correspondence: mark.rounsevell@kit.edu
https://doi.org/10.1016/j.oneear.2021.06.003

SUMMARY

There are many sources of uncertainty in scenarios and models of socio-ecological systems, and understanding these uncertainties is critical in supporting informed decision-making about the management of natural resources. Here, we review uncertainty across the steps needed to create socio-ecological scenarios, from narrative storylines to the representation of human and biological processes in models and the estimation of scenario and model parameters. We find that socio-ecological scenarios and models would benefit from moving away from "stylized" approaches that do not consider a wide range of direct drivers and their dependency on indirect drivers. Indeed, a greater focus on the social phenomena is fundamental in understanding the functioning of nature on a human-dominated planet. There is no panacea for dealing with uncertainty, but several approaches to evaluating uncertainty are still not routinely applied in scenario modeling, and this is becoming increasingly unacceptable. However, it is important to avoid uncertainties becoming an excuse for inaction in decision-making when facing environmental challenges.

INTRODUCTION

"The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people so full of doubts."[1] With this phrase, Bertrand Russell highlights the imperative of embracing uncertainty rather than fooling ourselves into thinking that it does not exist. This holds especially true for how we understand the natural world, including the increasingly important role of humans in socio-ecological systems. We know that socio-ecological systems are complex. They are non-linear, bifurcate, and have feedbacks and tipping points, all of which makes their future development inherently uncertain and difficult to predict. Indeed, the future is a place we can never know; we cannot observe it, and we cannot measure it. Yet, decision-makers are challenged with planning short- to long-term strategies for preserving biodiversity and the contributions of nature to people[2] and, so, we need to anticipate what the future may hold.

The scientific response to this challenge has been the development of scenarios to explore the uncertainty space of plausible, but unknown, futures.[3] Scenarios are not predictions, but are "a plausible and often simplified description of how the future may develop based on a coherent and internally consistent set of assumptions about key driving forces and relationships."[4] Scenarios are commonly underpinned by qualitative descriptions (narrative storylines) of the underlying direct and indirect drivers of change, including policy options,[3,5] which are often translated into impacts on biodiversity, ecosystem services, and complex socio-ecological systems using models in a storyline and simulation approach.[3] Hence, scenarios can be qualitative,

quantitative, or both. As such, scenarios and models are invaluable tools in guiding long-term, strategic policies that prompt management actions and increase public awareness of the future threats to nature.[6]

Due to the complexity of socio-ecological systems, but also to advances in knowledge and observation capacity, models are being developed with increasing complexity, involving many processes and feedbacks, and integrating multiple components of the ecosystem, from the physical environment to human societies. Examples include, land-use models,[7] agent-based models,[8] marine ecosystem models,[9,10] models of trophic levels,[11] dynamic vegetation models,[12,13] state and transition landscape models,[14] and niche-based models of species response to climate and land-use change.[15] There has been a strong focus on developing comprehensive modeling tools from empirical evidence,[16,17] but, until now, far less effort has been dedicated to exploring the uncertainties within these models, especially when used to quantify scenarios.

Identifying and quantifying future uncertainties may be key in achieving buy-in from stakeholders, to prompt evidence-based decision-making, and to shift mindsets on the perception of the future threats to biodiversity, ecosystems, and ecosystem services. To increase the influence of scenario and modeling analyses on policy and to trigger appropriate management responses, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) has strongly encouraged the use of scenarios and models, but warns that these "should be applied with care, taking into account uncertainties and unpredictability associated with model-based projections."[3] A critical challenge for improving scenarios and models of socio-ecological systems is to augment the scientific capacity in quantifying the uncertainty within and among model projections.[18]

Here, we review the current state of knowledge about the uncertainties associated with scenarios and models of socio-ecological systems within the context of decision-making, by which we mean the policy decisions made within private or public sector organizations. In doing so, we seek to address some of the key challenges raised by Elsawah et al.[19] that relate to uncertainty, such as the role of stakeholder engagement in the co-development of scenarios, linking scenarios across multiple geographical, sectoral, and temporal scales, improving the links between qualitative and quantitative scenarios, addressing surprises, addressing scenario consistency, communicating scenarios, and linking scenarios to decision-making. We do not aim to undertake an exhaustive evaluation of scenarios and model types. Instead, we use examples from a very wide range of scenarios and models to illustrate a comprehensive review of sources of uncertainty. A comprehensive review of sources of uncertainty in scenarios and models does not require a comprehensive review of scenarios and models. A wider ranging review can be found in the IPBES[3] assessment of scenarios and models.

We provide an overview of how uncertainty is treated within socio-ecological systems analysis and how understanding these uncertainties can enhance confidence in the creation of the next generation of scenarios and models. This is novel in both tackling a comprehensive review of sources of uncertainty in scenarios and models, exploring the implications of these uncertainties for decision-making and in setting out a number of potential solutions and recommendations for how to deal with these uncertainties.

TYPES OF UNCERTAINTIES

We focus on three categories of uncertainty: scenario uncertainty, model uncertainty, and decision-making uncertainty (see Table 1) across terrestrial and marine realms. We explore the whole chain of steps needed to create socio-ecological scenarios and models that are useful for decision-makers, from narrative storylines, the representation of human and biological processes in models, the estimation of model parameters, and model initialization and evaluation. Some of these sources of uncertainty relate to differences in worldviews, some to the limits of our current knowledge and others to our capacity to represent processes within models, including the reliability of model input data across spatial and temporal scales. Figure 1 shows the types of uncertainty (from Table 1) in the steps from observational data, model development, the construction of qualitative storylines and quantitative scenario projections that together provide input to decision-making.

SCENARIO UNCERTAINTY

Linguistic uncertainty

Linguistic uncertainty has been classified into five distinct types: vagueness, context dependence, ambiguity, indeterminacy of theoretical terms, and under-specificity.[20] Of these, ambiguity and vagueness arguably occur most commonly, largely because scenario terminology is often based on common language words. Indeed, the word "scenario" itself derives from the language of the theater. Yet, different communities can sometimes attribute different meanings to the same "precise" word, i.e., their use is ambiguous. For example, the word "pathways" is used as a synonym for "projections" or "trajectories" (as in the shared socio-economic pathways),[21] or alternatively it is used to describe a set of time-dependent actions that are required to achieve a future vision.[2] Using the term in one sense can lead to confusion if it is interpreted as being used in the other sense. Vagueness relates to statements with insufficient precision. For example, "population growth will increase strongly over the coming 50 years" tells us nothing about what a strong population growth actually looks like. Is it a doubling of population, or tripling, or something else? These different types of linguistic uncertainty commonly occur in narrative storylines, and they are especially important considerations when communicating the outcomes of scenario processes to decision-makers. Recent development of information technology provides a means to minimize linguistic uncertainty by building ontologies, i.e., an ensemble of formal definitions of concepts and their relationships within the domain of interest, and their synonyms or equivalents in closely related domains. While domain-specific ontologies exist in ecology that facilitate data mining and sharing,[22] to our knowledge, there is no widely accessible controlled vocabulary or thesaurus standardizing the meaning of the basic concepts used in scenarios of socio-ecological systems, as is the case with ontologies related to the Intergovernmental Panel on Climate Change (IPCC).[23]

Table 1. Sources of uncertainty and their description in scenarios and models of socio-ecological systems

Uncertainty sources	Description	Uncertainty types
Scenario uncertainty	The qualitative description of alternative worldviews and their development into the future and the quantification of model input parameters that are conditional on these descriptions.	*Linguistic uncertainty*. The use of similar terms to mean different things in different research communities, e.g., pathways, ensembles, boundary conditions. *Narratives storyline uncertainty*. The limits to imagining unknown futures (e.g., unknown unknowns). This can relate, for example, to alternative worldviews or the uncertainties associated with participatory processes arising from internal consistency and knowledge limitations. *Scenario parameter uncertainty*. The estimation of quantitative parameters from narrative storylines that are subsequently used in models. Scenario parameter uncertainty follows from the interpretation of quantitative values from qualitative narratives, e.g., the number of people in a "high population growth" scenario.
Model uncertainty	The representation of processes in models and how this is done.	*Structural (epistemic) uncertainty*. The uncertainties associated with the choice and the representation of processes in models. *Input data uncertainties*. The variability in baseline data conditions that are used to initialize a model, including thematic classification, i.e., how classes are defined in, for example, land-use maps. *Error propagation uncertainty*. The amplification (or dampening) of the transmission of errors across multiple coupled models. The role of meta-modeling and indirect effects (such as cross-sectoral interactions).
Decision uncertainty	Communicating and translating the results of scenario and modeling studies into decision-making.	*Data interpretation for decision-making*. Selective use of data or information from different sources and their interpretation. *Analyzing at relevant spatiotemporal scales*. The selection of spatiotemporal scales at which simulated data are analyzed, and the granularity of derived indicators (e.g., level of integration across biodiversity facets, merging subsets of ecosystem services). *Decision-making tools*. The variety of decision-supporting methods, e.g., multi-criteria decision analysis.

Narratives storyline uncertainty

The first step in the construction of scenarios is often the development of qualitative, narrative storylines.[5] These describe alternative trajectories in the key drivers of change (and their interactions) with a focus on socio-economic change. Socio-economic trajectories can also be associated with changes in physical conditions, such as climate change, where a change in climate is assumed to be internally consistent with drivers of, for example, societal consumption patterns and industrialization.[24] The uncertainties associated with the development of narrative storylines arise from how to create this internal consistency using mental models,[25] as well as the difficulty of imagining futures for which there are no historical analogs and representing a sufficient range of possible futures.[26,27] This affects the "plausibility" of narrative storylines in terms of whether the assumed causal relationships reflect real-world development, or the worldviews of the storyline developer. A particular case of this problem are "black swans," which reflect shocks or surprises to a system, i.e., events that are unexpected or assumed to have a low probability of occurring, but which have a high impact.[28] Black swans by their very nature can be difficult to anticipate or imagine, and are often unprecedented historically. The most appropriate way of dealing with uncertainties in storyline development is to clearly state and document the assumptions that underpin a narrative, and to communicate these assumptions when reporting a scenario study.[29]

Most narrative storylines focus on the supply side of natural resource systems (e.g., crop production or fish harvesting), and say little about the demand side (e.g., consumption patterns, such as dietary preferences) or the economic and institutional transformations that implicitly underlie the storylines. Although many "stylized" scenarios exist for diets, e.g., what would be the consequences for biodiversity of people becoming vegetarian or vegan,[30,31] these do not account for the transitions from where we are today to this assumed future situation.[32] Hence, the uncertainties associated with these transitions are not explicit.

Existing storylines of marine ecosystems largely focus on a narrow set of direct drivers, such as fishing or climate change,[33] or short-term policy interventions (such as protected areas or

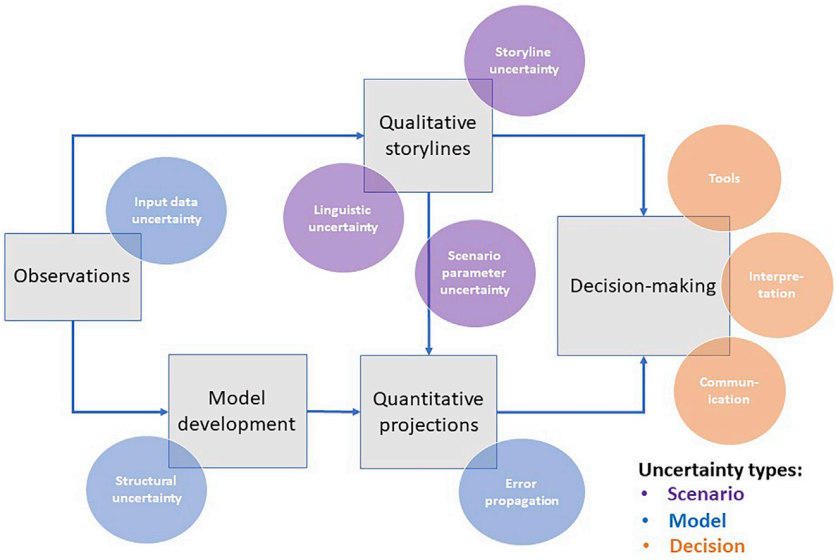

Figure 1. Sources of uncertainty in scenarios and models of socio-ecological systems within the context of decision-making

The circled sources of uncertainty are addressed in the main text: purple refers to scenario uncertainty, blue to model uncertainty, and orange to decision uncertainty.

management of fishing effort). Moreover, the consideration of indirect drivers, such as seafood demand from changes in population, consumption patterns or international trade, are not explicit in most marine storylines. Recent studies increasingly focus on expanding the scope of uncertainties by developing storylines that consider multiple drivers and policy interventions, in particular the interactions between climate change, fishing, and management.[34–36]

Terrestrial studies have a longer tradition of evaluating multiple, often cross-scale drivers in developing narrative storylines.[37] However, uncertainties arise from an overreliance on climate change as a driver, and not accounting for other drivers that are critical for socio-ecological systems, such as invasive alien species, trade in wild species, or air and water pollution.[2] Furthermore, uncertainties also arise from failure to account for indirect, cross-sectoral interactions.[37]

Participatory approaches, by which narrative storylines are co-created with stakeholders, add richness and diversity to storyline development, and strengthen the link between storylines and scenario quantification with models,[38] but are highly dependent on the selection of individual stakeholders and the extent of their explicit and tacit knowledge. Stakeholder mapping exercises[38] that seek to maximize stakeholder diversity are one way of resolving this problem. Participatory approaches are well developed in the marine realm, especially in fisheries management and marine spatial planning.[39,40]

Scenario parameter uncertainty

Simulation models can quantify the outcomes of narrative storylines for specific indicators. This requires the translation of the qualitative statements within a storyline into quantitative model inputs, which in itself has potential to introduce additional uncertainties.[5] We draw a distinction here between "scenario parameter uncertainty" and "model parameter uncertainty." Scenario parameter uncertainty derives from the translation of qualitative narratives into quantitative values, and so is dependent on the scenario itself, i.e., the quantitative values vary across scenarios. For example, a scenario parameter could be the number of people in a high, medium, or low population growth storyline. In general, scenario parameters relate to the socio-economic components of socio-ecological systems and may themselves be model inputs. Model parameter uncertainty refers to the estimation of parameters within the functions that represent modeled processes, e.g., a rate constant or capacity, and often, but not always, relate to the biophysical components of socio-ecological systems. Hence, model parameter uncertainty depends on the system and the model of that system, and is independent of a scenario. Scenario parameter quantification often uses best-guess estimates that sometimes draw on uncertain, historical analogs. However, the majority of these studies do not account for the uncertainties associated with the process of estimating scenario parameters themselves. A few exceptions to this have defined "credible" parameter ranges,[41] or have used conditional probabilistic futures methods.[42]

In the conditional probabilistic approach, probability distribution functions (PDFs) are created for the scenario parameters that are conditional on the assumptions within a scenario storyline, thus reflecting the uncertainty range in the estimation of a scenario parameter.[42–44] When combined with Monte Carlo sampling across the PDFs and multiple model simulations this approach is able to explore the range of scenario outcomes that are contingent on the uncertainties of scenario parameter inputs, although subjective assumptions and choices made in Monte Carlo sampling can introduce uncertainty in model outcomes.[45] Conditional probabilistic approaches have been used to explore whether scenario parameter uncertainty leads to divergent or (more commonly) convergent outcomes across scenarios.[43] Being computationally intensive, this method is less tractable for models with long run times, which constrains its application for many large-scale models. However, run times are also affected by the temporal and spatial resolution as well as the spatial extent of a model, and computational capacity is becoming increasingly less important.

Apart from these examples of scenario parameter uncertainty being quantified and communicated, there is little quantification of the uncertainties arising from different management and policy actions to achieve stylized scenarios,[2] e.g., assumptions of vegetarianism,[31] maximizing long-term fishing catches,[46] and the rate of change in fishing technologies that have been identified as key drivers of increasingly effective fishing effort that impacts marine biodiversity.[47] Management practices are especially important when representing adaptation processes within models in which responses are consistent with time-varying,

scenario-specific barriers and enablers, e.g., societal values and governance.[48] Overall, there are considerable gaps in current knowledge about scenario parameter uncertainty.

Model uncertainty

Structural uncertainty

Models simplify the representation of the real world in different ways and so produce different responses to the same scenario assumptions. These responses depend on how a model is structured and parameterized and on the timescale, all of which can lead to structural model uncertainty. Hence, modeling is the art of making choices in a given context, and structural uncertainties reveal the variety of these choices.[49,50] The more knowledge we try to formalize within models through process-based understanding, the more uncertainty we may potentially cause or reveal. One could argue that simple, parsimonious models are better than complex models for robust forecasting,[51,52] but there is no universal evidence of a relationship between model complexity and model robustness. Parsimonious models that are based on observed trends may lead to low uncertainty within the range of conditions for which they were calibrated, but can lead to high uncertainty when applied over longer timescales or in scenarios with large deviations from current trends.[53,54] However, focusing on parsimony misses the point about why we build models. We model to experiment with elements of the natural world to explore, explain, and understand how they work.[51]

Many models of climate, land use, and biodiversity are increasing in complexity by the addition of components, processes, and model coupling.[55–57] More complex models may, arguably, be better at representing system dynamics over longer time scales or under changing conditions than simpler models.[58] For example, oversimplifying biodiversity representation in vegetation models has long been an impediment to detailed understanding and robust projections of ecosystem dynamics and distribution.[59,60] This has motivated a finer representation of species or traits diversity,[61–64] which allows better exploration of the role of the interactions between diversity and ecosystem functioning in shaping the future of natural systems.[65,66] However, this does not necessarily lead to less uncertainty, since the representation of feedbacks and path dependency may lead to dramatic changes in system behavior, potentially increasing the range of possible responses and associated uncertainty. Furthermore, increasing model complexity may also lead to problems with the traceability of the origins of uncertainty and inconsistencies between different model components.[67] These problems may be further compounded within models that include stochastic process representations, leading to internal variability and multiple model outcomes. However, stochastic approaches based, for example, on Monte Carlo methods can be useful in representing uncertainty in model structure.[68]

Models can support improved understanding of how resource management can adapt to environmental change and thereby inform decision-making and policy processes. However, a better representation of adaptation processes is required in models in general. For example, substantial differences have been found between the extensive, available empirical knowledge about societal adaptation processes and their representation in models of land and water sectors.[69] Only a minority of models take account of the management choices that underpin adaptation measures or the constraints (financial, institutional, social, etc.) that may limit the uptake and effectiveness of adaptation;[70] factors that are likely to be influenced (positively or negatively) by the scenario setting. The pervasiveness of simplistic, over-optimistic approaches to simulate the role of adaptation in reducing impacts and vulnerabilities or in exploiting the benefits associated with climate and socio-economic changes means that studies may produce findings that cannot meaningfully inform decision-making about appropriate adaptation strategies.

Incremental model improvement aims to increase a model's ability to predict plausible responses to uncertain, environmental change conditions. The drawback of incremental improvements is that they can cause "lock-in" of an existing model structure or ways of doing things.[71] Moreover, even incremental changes in model structure require substantial investment in time and effort. The exploration of alternative structural specifications in models is often done for local- to regional-scale studies.[72–74] At the global level, the investment required to build new models may be substantially larger than maintaining existing models. Global-scale models often need long-term institutional funding, thus limiting the number of research groups that have the capacity for such effort. Hence, the diversity of model structures and modeling paradigms is low in global-scale modeling compared with regional-scale models.[75] For example, many global-scale economic models still use optimization approaches based on the assumptions of neoclassical economics that are known to be limited.[76]

Better understanding of structural uncertainty is often achieved by trying to learn from model inter-comparison exercises[7,77] (see Box 1) for the comparison of model results with observed data.[78,79] Model inter-comparisons and the closely related ensemble modeling approach have proven highly beneficial for improving the credibility of climate change projections, such as through the Coupled Model Inter-comparison Project (CMIP).[80] Similar multi-model efforts, in which different models that address a similar question are run using a standardized simulation protocol and the same input data, are only starting for impact models projecting future terrestrial[2,75] and marine biodiversity (Fish-MIP).[33,81]

The comparison of model outputs with observational data,[106] or benchmarking, can provide pointers toward the conditions under which a model performs better or worse, as well as revealing the sources of uncertainty. Diverse sets of observations are needed to assess both the magnitude and seasonal and interannual variability of modeled outputs.[82] Specialized experiments, such as free-air carbon enrichment studies, herbivore exclosures, or remotely sensed trait information[90–92] can also be used to test the realism of specific simulated processes. Taken together, these datasets can be used to test whether models correctly capture existing relationships between variables (or incorrectly assume existing relationships, which are not supported by observations). At least for vegetation models, studies have begun to systematically explore the use of scoring of model performance against a range of observations.[82] Two further common approaches to model improvement are: (1) the addition or re-specification of certain model components and (2) the simple calibration of model parameters to increase the model fit to data. Calibration may lead to either overfitting of the model or

Box 1. Model benchmarking, inter-comparison projects, and ensembles

Benchmarking is the repeated confrontation of models with a range of observations to establish a track record of model developments. Observational datasets in themselves are uncertain,[82,83] so benchmarking needs transparent information on which observations were used. Some global models already routinely undergo a systematic confrontation against data when new processes are added (e.g., for the terrestrial carbon cycle).[84–89] Recent approaches allow scoring of model performance against a wide range of observations for global vegetation models.[82] Observational data for benchmarking include multiple-site and remote-sensing products of, e.g., fraction of absorbed photosynthetically active radiation, gross primary productivity, net primary productivity, burnt area, river discharge, or atmospheric CO_2 concentration. Specialized experiments or datasets, such as free-air carbon enrichment studies, herbivore exclosures, or remotely sensed trait information[90–92] can also be used to test the realism of specific simulated processes. Diverse data are needed to assess both magnitude and seasonal and interannual variability of modeled processes.[82] These datasets can be used to test whether models correctly capture existing relationships between variables (or incorrectly assume existing relationships, which are not supported by observations). Physics, climate, and biogeochemistry observations are generally more numerous, systematically measured, and available on different spatiotemporal scales, whereas biodiversity data are more disparate and contain many gaps (e.g., the GOOS marine initiative),[93] so benchmarking is much more challenging for biodiversity models.

In models of climate, oceans, and ecosystem dynamics, stochastic sensitivity analyses (sometimes called "perturbed physics experiments") are applied (see also section parameter uncertainty) where model-internal parameter values are sampled across a parameter-space to explicitly and transparently test parameter-value uncertainty.[94] These analyses are computationally expensive and, so, have not been sufficiently exploited with coupled and integrated models. But, a number of studies have demonstrated their application both in offline models (e.g., related to vegetation or land-use change modeling) and in coupled models (e.g., related to carbon cycle-climate feedbacks).[42,44,95–98] Results help to identify those parameters to which a model is most sensitive, but can also inform sensitivity analysis of other models for those values. The outcomes aid the interpretation of, e.g., model ensembles as the magnitude of uncertainty seen in a single model's output from stochastic parameter sensitivity analysis can be compared with the spread in output within a model ensemble.

The currently most widely used approaches to quantify model uncertainty in climate change, land-use change, exploitation, and ecosystem modeling are inter-comparisons and model ensembles.[7,99–102] Ensemble modeling has proven highly beneficial for improving the credibility of climate change projections with international model inter-comparison efforts such as the Coupled Model Inter-comparison Project (CMIP).[80] It is only starting for impact models projecting future terrestrial[2,75] and marine biodiversity (Fish-MIP).[9,103] In model inter-comparisons, different models that address a similar question are run using a standardized simulation protocol and the same input data. Output comparison helps to identify whether models agree or disagree in the simulated time series or spatial patterns. In some cases, an ensemble mean is used based on the notion that the average across a range of models would "average-out" some of the structural and parameter-related uncertainties and yield more robust results.[15,94,104] However, the comparison between individual models and the "ensemble mean" might unintentionally also lead to the model being "re-tuned" to fit better to the average model response. Furthermore, "families" of similar models (or with similar development heritage) tend to bias the mean, as they are each given the same weight as a genuinely different model. So far, most ensemble studies do not identify and exclude (or give different weight to) models that fail to fulfill certain quality-assurance criteria (based on scores in a benchmarking exercise). This has started, however, to be the case for the terrestrial models used in the annual global carbon budget calculation.[105] In view of the often still untested model structural and parameter uncertainties, deriving probabilistic estimates of uncertainty from model ensembles must be viewed critically.[94]

to issues relating to equifinality. In overfitting, a calibrated model may represent a specific place and time very well, but it sacrifices generality when applied to other places and times. The comparison between individual models and an "ensemble mean" might unintentionally also lead to the model being "re-tuned" to fit better to the average model response.

Equifinality occurs when different functional or process representations in a model lead to the same outcome.[107–109] This reduces the range of the modeled outputs, but at the same time may conceal structural uncertainty, since it can be difficult to track which mechanisms within a model lead to the equifinal outcomes. The effect of equifinality can be evaluated by comparing the overall model outcomes against independent datasets,[58] but also by comparing different process representations within the model itself. This is important when assumptions are made, for example, in how to model the management choices that underpin land-use change.[110] While different approaches to representing management choices may, in the short term, lead to similar land-use outcomes, they may wrongly represent longer-term adaptation and behavior under resource constraints. In this case, empirical data on management choices may be more useful in validating the model process than validating the short-term model outcomes.

In a review of land-use models, little over half were validated independently, and many conflated calibration with validation.[70,111] Although this can be explained to some extent by the limited availability of consistent empirical datasets for different time periods, it still increases the risk of overfitting in many model applications. In other words, a model both trained and validated on historical data may not accurately project the full range of outcomes in a non-stationary future. However, calibration to improve model fit can, in part, compensate for the subjective decisions made by modelers concerning the selection of observed input datasets (e.g., which meteorological, economic,

or demographic variables), alternative process algorithms (e.g., reference evapotranspiration), and initial conditions (e.g., land-use classes and their distribution).[112,113] Nevertheless, the consequences of these choices may still be unclear when the model is perturbed beyond the historical conditions represented in the calibration data, leading to potentially large uncertainty in the magnitude and direction of impacts.[113]

Input data uncertainty

It is difficult to decouple model structural uncertainty from model input data uncertainty, since models with a different structure commonly use different input data.[7,104] Models of socio-ecological systems are data demanding for parameterization, calibration, and initialization of simulations, including large demands for baseline data. Uncertainties in the use of data can emerge from measurement errors, data scarcity, or a mismatch between the resolution and scope of the available data, and the needs of the model. These uncertainties are amplified when models include additional processes, represent processes at finer spatial scales, or expand the spatial and temporal scope of simulations. For example, data availability has been assessed for several mechanisms known to play a key role in mediating species responses to climate change, such as physiological processes, evolutionary potential, and species interactions.[114] Even for the best-studied species, data were at best incomplete if not entirely absent. In recent years, the scientific community has gone to great lengths to increase access to biodiversity data through the development of networks of high-quality monitoring systems (observation systems, instrumented sites, and remote-sensing sensors),[115–118] data repositories (e.g., GBIF.org; obis.org), or citizen science programmes.[119–121]

For correlative species distribution models,[122,123] the lack of accuracy and comprehensiveness of the species data and of the relevance and completeness of the predictors can critically impact the relevance of the fitted niche models and hence of the resulting outcomes.[124,125] Data deficiencies and biases in this specific approach include samples of species' occurrences that are too small or do not include absences, or have missing covariates; the latter being known to introduce significant spatial correlation in the errors of the analysis.[126–129]

Trait-based approaches have been developed to leverage limited data and allow model prediction for a broad range of species, including poorly studied ones. Traits are individual features that inform individual performance.[130] Both correlative and process-based models have used trait parameters to simulate higher-level processes. This includes population growth rate or range shifts in plant,[64,131–133] fish,[134–136] or reptile and amphibian communities.[137] Trait data availability is increasing rapidly (e.g., open digital repository;[138,139] www.fishbase.org), but it remains highly variable across taxonomic groups and geographic areas. It is also strongly correlated with the ease in measuring traits: so-called "soft" structural traits have been more often measured than "hard" physiological traits, although the latter often provide key information on species responses to non-present analog conditions, such as tolerance to drought or higher temperatures.[140–142] In addition, functional ecologists often report species mean trait values, resulting in a lack of assessment of intraspecific trait variability[142] despite increasing evidence for its role in species adaptation and coexistence.[143–146] These are both crucial in establishing biodiversity projections.[147]

Uncertainties related to initial conditions are less well studied in socio-ecological models,[148] although they have been identified as important in some studies. For example, variability in the data used to represent initial land-use conditions between different models of land-use change contributed a substantial part to the variation across future land-use projections[7] with distinct spatial differences in the level of uncertainty.[104] Differences in initial data can arise from different definitions of the same land cover type and different data acquisition approaches.[7,104] Similarly, errors in the initialization of forest structure in large-scale simulations of vegetation models can result from limited sampling and coarse resolution (for example, of large-scale, remote-sensing products), and have been found to propagate in subsequent model prediction uncertainty.[73,149,150]

Several methods are available to address input data uncertainties. Hierarchical modeling techniques and other statistical methods can address different sources of uncertainty explicitly in modeling frameworks.[146,151,152] Sensitivity and uncertainty analyses[43,153,154] can help identify and prioritize the need to reduce parameter uncertainty given limited time and resources and hence guide the empirical effort of data collection through iterative cycles of data-model fusion.[16,155,156] In stochastic sensitivity analyses (sometimes called "perturbed physics experiments"; see Box 1) model-internal parameter values are sampled across parameter-space to explicitly and transparently test parameter-value uncertainty.[94] These analyses are computationally expensive and, so, have not been sufficiently exploited with coupled and integrated models. But, a number of studies have demonstrated their application both in offline models (e.g., related to vegetation or land-use change modeling) and in coupled models (e.g., related to carbon cycle-climate feedbacks).[42,44,95–98] Results help to identify and rank those parameters to which a model output is most sensitive, but can also inform sensitivity analysis of other models for those values. The outcomes aid the interpretation of, e.g., model ensembles as the magnitude of uncertainty seen in a single model's output from stochastic parameter sensitivity analysis can be compared with the spread in output within a model ensemble.

Data assimilation techniques can bridge the gap between data availability and model requirements. In particular, inverse modeling, such as approximate Bayesian computation use a wide range of data to refine values of input parameters.[157–160] With these methods, parameter distributions provided by the available data (prior parameter estimate) are iteratively adjusted (posterior parameter estimate) by comparing simulation outputs with observed data at different scales, e.g., element fluxes derived from eddy-flux measurements,[161] tree size distribution derived from inventory data,[162] or remote-sensing products.[163]

A promising avenue in terms of data assimilation is the spectrometry imagery of functional diversity,[90,164] which, at least for terrestrial ecosystems, can help to bridge the gap between biodiversity data available from field surveys and the amount of data required to better control for uncertainty in continental- and global-scale models. This raises new technical challenges in terms of data standardization (corrections and inter-calibration of remote-sensing images) and methods for data extraction.[165] It also raises the issue that the input data themselves

often derive from modeled products. For example, in modeling the terrestrial C-cycle, the same level of uncertainty is possible for several DGVMs forced by the same climate scenario (based on a single emissions scenario and climate model), as for a single DGVM forced by inputs from several climate scenarios (with different emissions and climate models).[166]

Error propagation uncertainty

Uncertainties from error propagation arise in coupled model systems when the inputs to one model (e.g., a model of climate impacts on ecosystems) derive from the outputs of another model (e.g., a climate model). In some cases, several models are coupled together leading to serious error propagation especially at the end of the chain of coupled models.[167,168] Error propagation becomes even more important when there are dynamic feedbacks between models.

Coupled models are common in integrated assessment, which seeks to explore the interactions between, as well as within, different socio-ecological systems.[56] Integrated assessment models (IAMs) focus, for example, on the connections between the economy, the energy system, and land cover change[169] at global-scale levels. However, regional IAMs have also demonstrated the importance of adopting a cross-sectoral approach for impact assessments.[37] Indeed, the impacts of climate change as reported by the IPCC may be over- or underestimated because they fail to account for cross-sectoral interactions.[37] A source of uncertainty in coupled models is when simplified, meta-models replace complex models to facilitate data flows across systems.[37,153] However, these uncertainties may be acceptable since the indirect effects of one sector on another sector are often more important than the changes within a single sector itself.[37] Similar issues arise for models that do not consider cross-scale impacts, since one scale level is highly dependent on the boundary conditions defined by a higher-scale level.[76]

Different methods can evaluate the uncertainties arising from error propagation, with qualitative methods being of particular utility. Dunford et al.[168] combined formal numerical approaches, modeler interviews, and network analysis to provide a holistic uncertainty assessment of a regional integrated assessment model that considered both quantifiable and unquantifiable uncertainty. Maps of modeler confidence (the counterpart of uncertainty) were created from fuzzy-set methods and network analysis to show that validation statistics are not the only factor driving modeler confidence. Several other factors, such as the quality and availability of validation data, the meta-modeling process, trust between modelers, derivation methods, and pragmatic factors, such as time, resources, skills, and experience were also found to be important.[168]

For most simple models (e.g., linear Gaussian models), the variance of the prediction associated with error propagation can be computed analytically, paying attention to the dependence between variables and the associated covariance.[170] In the majority of cases, modeling involves complex models that are non-linear and non-Gaussian for which variance computation is analytically intractable. In such cases, error propagation can be evaluated through simulation using, for example, Monte Carlo methods.[171] A Monte Carlo-based approach to evaluate the propagation of uncertainties in a regional integrated assessment model, showed that, rather than the uncertainties "exploding" in importance, there was convergence across a range of contrasting scenarios.[43] This implies that if fully understood, uncertainties arising from error propagation can be managed successfully. However, the assessment of error propagation through simulation is computationally demanding and, in general, only applicable to models with rapid run times.

Model output-input chains and feedbacks can become complex and lead to unacceptable levels of uncertainty for decision-making.[168] Where possible, major sources of uncertainties (data, model, parameters) should be identified a priori to allow propagating errors with a minimum number of simulations. Comprehensive sensitivity analysis is also useful in identifying emergent uncertainties.[153] Structured sensitivity analysis (also referred to as scenario-neutral approaches and impact-response surfaces) is valuable in evaluating whether the emergent behavior in coupled models as a response to simple perturbations is consistent with understanding or influenced by error propagation, although sensitivity analysis as a method has been criticized.[172] Hierarchical Bayesian models can be useful tools to incorporate and propagate errors from multiple sources (data, parameters, models), through the computation of the predictive posterior distribution.[173]

UNCERTAINTIES IN DECISION-MAKING AND DECISION METHODS

Intrinsic uncertainties in decision-making

Uncertainty pertaining to environmental processes and ecological theory is interesting from an academic perspective, but it becomes a practical issue when it impinges on the ability of managers, planners, and policy makers to make relevant science-based decisions to achieve societal objectives.

Despite multiple uncertainties, decisions are still made about natural resource management. However, the decision-making process is itself messy and difficult to predict, depending as it does on the context, on the individuals involved (with their conscious and unconscious biases), on the breadth of values attributed to nature (including non-quantifiable ones), on the efficient exchange of knowledge between science and policy, and on time lags in policy implementation.[174] Decision-making is often disorganized and politicized, and has to deal with many trade-offs, as well as co-benefits, making it difficult to generalize about how uncertainty in scenarios and models affects decision-making processes. There is a significant body of work in decision theory and operations research on dealing with epistemic uncertainty in decision-making. However, further understanding is still needed on the relationship between science and the social and political processes of decision-making, and this is an important area of future research in environmental management.

What can be stated is that different degrees of uncertainties and levels of controllability may be more effectively managed by different strategies and approaches.[3] Controllability here refers to the degree of control that a decision-maker has over the system being managed. Controllability tends to be higher when decision horizons are shorter, when the decision-maker has direct and sole jurisdiction over the places and/or resources being managed, or when stakeholders do not vary widely in their aspirations for the outcomes of management. Controllability

covaries with uncertainties over temporal and spatial scales. It tends to be higher at local and national scales relative to regional and global scales.[175] When the system is highly controllable, and uncertainties about the future are low, it may be most effective to implement optimal control tactics. Optimal control tactics generally involve "predict-then-act," such as determining catch or fishing quotas.[176] In situations where controllability is low and uncertainty is high, robustness analysis[177] in support of scenario planning[178] may be favored.[179]

In this section, we further discuss how uncertainties in scenarios and models can contribute to decision-making uncertainty, as well as the tools that are available to address these uncertainties and their limitations.

How uncertainties are communicated to decision-makers

How uncertainties are accounted for in decision-making is strongly dependent on how these uncertainties are communicated to decision-makers. In international science-policy processes, such as IPCC or IPBES, formalized uncertainty language is used to communicate levels of confidence in the assessment of scientific evidence,[180] including results from scenarios and models. This approach is generally qualitative, although attempts have also been made to use quantitative probabilistic statements. Whether this approach is effective in communicating uncertainty to policy communities is debatable,[181] although some benefit to decision-makers is likely since government-approved assessment reports continue to use uncertainty language.

How uncertainties are accounted for in decision-making is also strongly dependent on how these uncertainties manifest into the different indicators that are provided to decision-makers, e.g., Living Planet Index,[182] species richness,[183] extinction risk,[184] and monetary value of ecosystem services.[185] Communicating alternative scenario outcomes thus requires appropriate indicators that are understandable and meaningful to decision-makers, and above all responsive to different drivers in an expected way, i.e., with low uncertainty. Within the same scenario or model, the way the output variables are transformed, integrated, and combined into indicators does not result in the same level of uncertainty,[186] or in the same strength of the signal-to-noise ratio.[185] The granularity of an indicator can be key (from population, to multispecies, to whole community level for example), as well as the choice of the spatial and temporal scales at which it is integrated. The portfolio statistical concept developed in economics and used by analogy in ecology, explains why dynamics may be extremely volatile at small scales (and high biodiversity granularity, e.g., population biomass), but less variable at more aggregated scales (and low biodiversity granularity, e.g., community biomass).[187] International initiatives, such as the Group on Earth Observations Biodiversity Observation Network (https://geobon.org), the Global Ocean Observing System (www.goosocean.org), and the Biodiversity Indicators Partnership (www.bipindicators.net), have proposed a number of indicators and essential biodiversity variables to characterize changes in biodiversity status under global change. However, the selection of indicators has been done mostly under the criteria of measurability and accessibility at the global scale,[116,118] but the performance of indicators in capturing changes and associated uncertainty have rarely been tested in a systematic way.[188,189]

It is not possible to say whether communicating to decision-makers the uncertainties in scenarios and models of socio-ecological systems actually changes decision-making in practice or not. There is no objective measure of the "success" of communicating uncertainties, nor is there a counterfactual to explain whether alternative decisions would have been made in the absence of knowledge about uncertainties.

How decision-making tools address uncertainties

A great number and variety of tools exist to support decision-makers in dealing with various kinds of uncertainty when making decisions.[6] A key role of decision support tools is to provide a framework that allows decision-makers and stakeholders to separate deliberations about what represents a desired outcome (competing objectives and preferences that arise from differing values) from deliberations about the facts of the matter; the probability that a particular course of action will result in a particular outcome. Therefore, it can be useful to think about different decision support tools in terms of how they deal with competing values and uncertainty (see Figure 2).

Decision support tools vary in terms of how they deal with spatial scale and extent, cultural and administrative complexity, multiple stakeholders, and competing values and uncertainty.[6] In Figure 2, we outline a small sample of the decision support approaches that deal with uncertainty to varying degrees with the aim of highlighting the breadth of opportunities for addressing competing values and models using existing decision support approaches, and these approaches are summarized in Table S1.

Despite the widespread development of decision support tools, the capacity of these tools to support objective decision-making may often be limited, especially where high levels of complexity and uncertainty make interpretability difficult. For example, when uncertain trade-offs between different ecosystem services are at stake, tools designed to support decisions are usually required to impose artificial boundaries or quantifications, and to limit and render comparable the broad, diverse range of services in question.[190–193] This implicitly involves the same value-based judgment under uncertainty that a decision-maker would be faced with in the absence of such a tool, but often obscures its subjective nature. More systematic biases also exist. Knowledge about socio-ecological systems is growing so rapidly and on so many fronts that it is very difficult to capture accurately. Social science knowledge in particular is consistently neglected, perhaps because most tool developers are natural scientists.[194,195] This also contributes to the neglect of cultural services, and their uncertainties, in ecosystem services assessments.[196] Even tools that sacrifice coverage are likely to prove to be too complex and uncertain to be used and understood by stakeholders as originally intended.[190]

Decision support tools therefore run the risk of obscuring uncertainty and subjectivity rather than helping to overcome it. This can be revealed, and to some extent overcome, where tools are used in participatory settings that allow for interrogation of assumptions, representation, and outcomes by a range of stakeholders.[197] Comprehensive uncertainty evaluation can

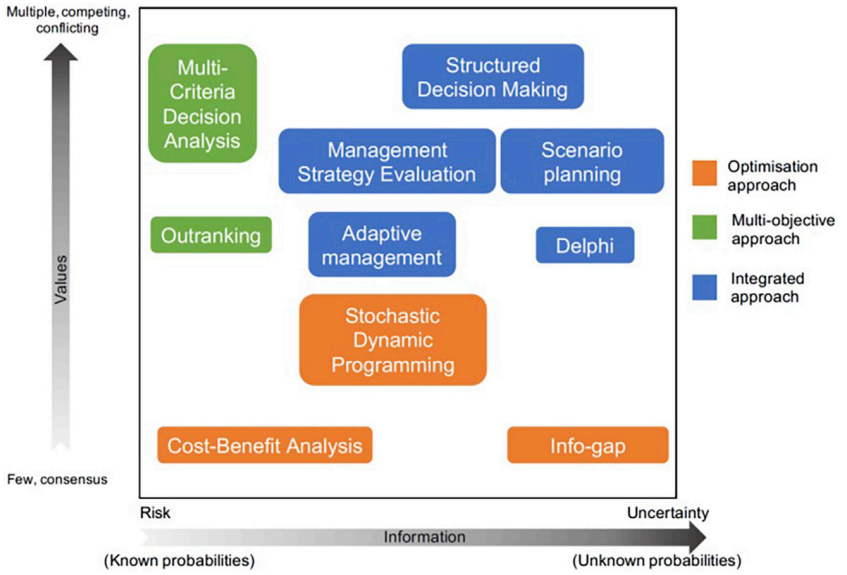

Figure 2. A sample of decision tools to support decision-making in the presence of competing values and uncertainty

See Table S1 for tool summaries and key references. Optimization approaches (orange) are a broad family of approaches that utilize either simple (cost benefit) or more sophisticated (info-gap) mathematical formulations that maximize an objective function. Multi-objective approaches (green) focus more on characterizing the competing values and preferences of decision stakeholders through more deliberative, or sometimes hybrid deliberative/quantitative processes. Integrated approaches (blue) tend to bring a suite of deliberative and quantitative tools together into a framework that seeks good decisions (e.g., Adaptive Management and Structured Decision-Making).

play an important role in this process,[198] but is not itself sufficient. Rather, improved and more comprehensive methods of accounting for subjectivity and uncertainty within nominally objective decision processes remain a priority.[199]

DISCUSSION: WAYS FORWARD

It is important to recognize the many sources of uncertainties that exist in scenarios and models of socio-ecological systems. It is also important to avoid these uncertainties becoming a disincentive for action when facing environmental challenges, within either the science or decision-making domains. Importantly, decision-makers should not use uncertainty as an excuse for inaction. There is no panacea for dealing with uncertainty, but a portfolio of approaches may provide an opportunity to better understand and cope with uncertainty. This portfolio might include a range of methods from Model Inter-comparison Projects (MIPs), validation against independent data, error propagation analysis, to learning from uncertainty to guide model improvement. Table 2 provides a summary of the approaches to addressing uncertainty that are discussed throughout this article. Figure 3 also provides a visual representation of these approaches with referencing to Table 2. Together, these provide a checklist of the types of actions that can be implemented when dealing with uncertainties of scenarios and models of socio-ecological systems within the context of supporting decision-making.

A number of ways of dealing with uncertainty are still not routinely applied in scenario modeling and this is becoming increasingly unacceptable. For instance, statistical parameter uncertainty analysis may not be possible for all parameters for all models, but it can be done at least for a subset of model parameters. Likewise, the confrontation of models with data is inadequately done. In many cases, there may be insufficient data to do this properly, but using this as an excuse to do nothing at all is simply wrong. In situations where data are lacking, one should start with qualitative "common sense" tests, such as by Turner et al.,[200] who identified future projected rates of change in bioenergy adoption to be three times faster than the historical precedent for the most rapidly changing land use.

Likewise, creating better scenarios of uncertain futures would benefit from consideration of a wider range of socio-economic and natural system drivers going beyond a focus on climate change alone.[2] This includes, for instance, drivers of biodiversity loss, such as biomass extraction, invasive alien species, and pollution.[2] Many scenarios are also weak at relating indirect drivers (i.e., the underlying socio-economic-political causes of change) to direct drivers. We need to move beyond the representation of stylized scenarios of, for example, consumption patterns, to scenarios and models that account for the role of human behavioral processes in affecting ecological change. This includes better representation of how policy and conservation initiatives affect people with the knock-on effects this has for ecosystems.[201] This is critical in better evaluating the considerable role of humans in causing ecological degradation, and in informing the decision processes that can do something about it through restoration and effective ecosystem management.[202]

Within this review, we have focused on models and scenarios of socio-ecological systems. However, it is clear from the literature that there is a bias toward the "ecological" aspects rather than the "social" aspects of such systems, such that many modeling approaches do not adequately capture the full range of interacting human and natural processes. We view this as a major research gap in current modeling and scenario exercises, and suggest that further development in this field would benefit from a greater focus on the social phenomena that are critical in understanding the functioning of nature on a human-dominated planet.

Uncertainty is often seen as the problem, while instead it could be interpreted as a "space" to manage socio-ecological systems in more desirable directions. Uncertainty also helps to target future effort in model development and to identify areas that lack understanding and, so, are priorities for future research. However, structural uncertainty needs to go beyond the improvement of model components and details, by re-evaluating the fundamental principles and assumptions of a model structure. Furthermore, part of the total uncertainty in the future of

Table 2. Potential solutions and recommendations to address uncertainty in models and scenarios of socio-ecological systems for different sources of uncertainty

	For these sources of uncertainty								
	Scenario uncertainty			Model uncertainty			Decision-making uncertainty		
Potential solutions and recommendations	Storyline	Linguistic	Parameter	Structural	Input	Error propagation	Tools	Communication	Interpretation
1. Stakeholder mapping exercises to address uncertainty in participatory processes	✓								
2. Explicitly state and document the assumptions that underpin a scenario narrative, and communicate these assumptions when reporting a scenario study	✓	✓							
3. Building ontologies		✓							
4. Defining credible scenario parameter ranges or using conditional probabilistic methods			✓						
5. Considering a wider range of socio-economic and natural system drivers that go beyond a focus on single drivers alone, e.g., climate change			✓						
6. Model inter-comparison exercises and model ensembles			✓	✓					
7. Developing coupled socio-ecological systems models that identify and represent important feedbacks to support the inclusion of feedbacks in scenarios			✓	✓					
8. Model benchmarking (see Box 1)				✓					
9. Validation against independent data, including the confrontation of models with empirical data				✓					
10. Going beyond the improvement of model components and details, by re-evaluating the fundamental principles and assumptions of a model structure				✓					
11. Developing scenarios and models that better account for the role of human behavioral processes in affecting ecological change				✓					
12. Learning from uncertainty to guide model improvement				✓					
13. Qualitative "common sense" tests, where independent validation data are lacking				✓	✓				

(*Continued on next page*)

Table 2. ***Continued***

Potential solutions and recommendations	For these sources of uncertainty								
	Scenario uncertainty			Model uncertainty			Decision-making uncertainty		
	Storyline	Linguistic	Parameter	Structural	Input	Error propagation	Tools	Communication	Interpretation
14. Hierarchical statistical modeling techniques and other methods, such as sensitivity and uncertainty analyses				✓	✓				
15. Increasing data access, e.g., developing high-quality monitoring systems (observations, instrumented sites, and remote-sensing sensors), data repositories, or citizen science					✓				
16. Data assimilation techniques, such as inverse modeling, e.g., approximate Bayesian computation					✓				
17. Error propagation analysis through, for example, qualitative methods, formal numerical approaches, modeler interviews, and network analysis						✓			
18. Simulation using, for example, Monte Carlo methods						✓			
19. Application of decision support tools to policy questions							✓	✓	
20. International initiatives to standardize indicators and make them available								✓	
21. Systematic testing of the performance of indicators in capturing socio-ecological changes and associated uncertainty								✓	
22. Defining appropriate indicators that are clear, concise, and responsive to different drivers								✓	
23. Improved and more comprehensive methods of accounting for subjectivity and uncertainty within nominally objective decision processes									✓
24. Co-creation and decision support in a participatory setting that allows for interrogation of assumptions, representation, and outcomes by a range of stakeholders	✓								✓

See Table 1 and the visual presentation in Figure 3. This list does not preclude other relationships between solutions and uncertainty sources that may be feasible.

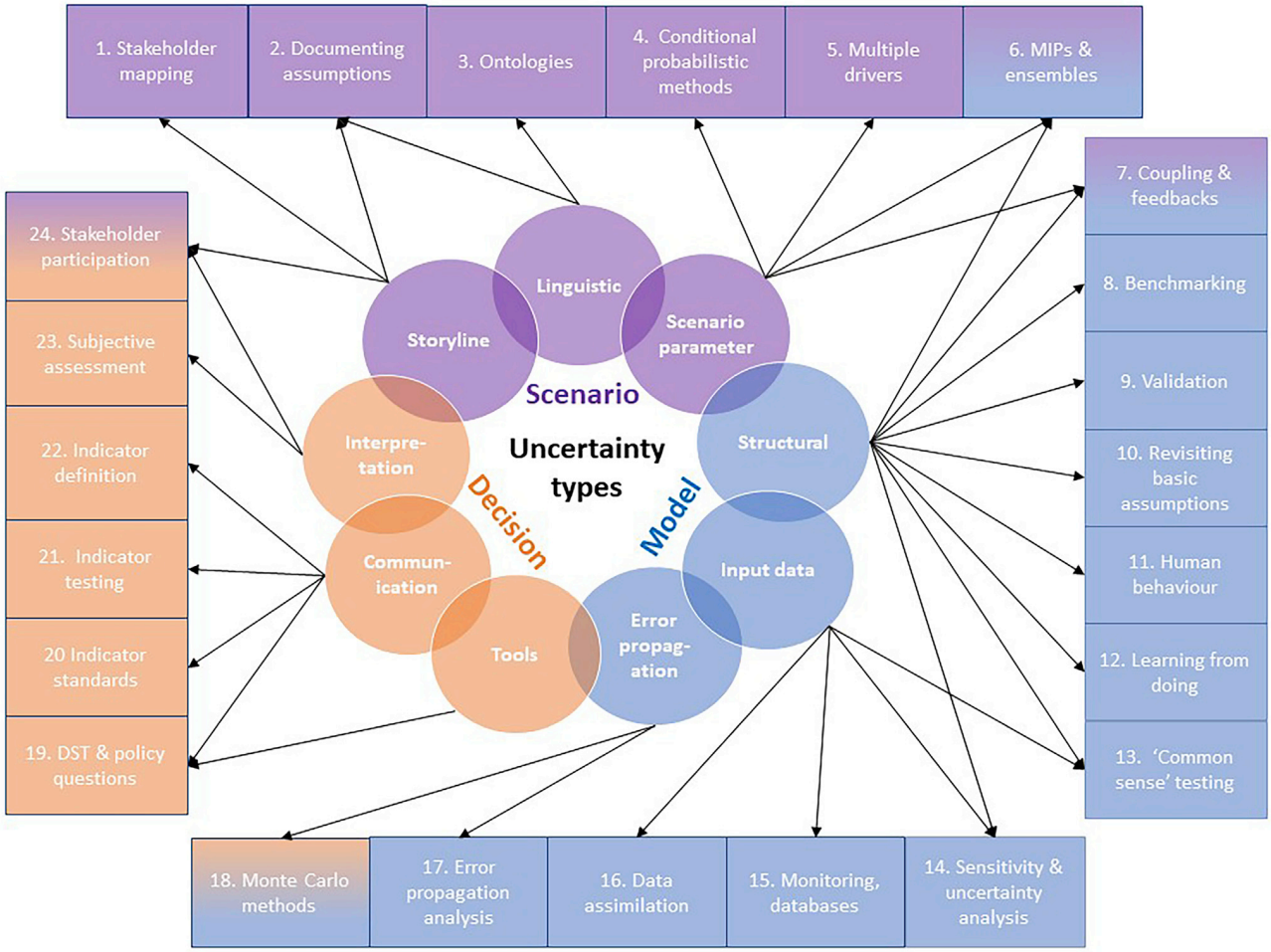

Figure 3. Visual summary of the types of uncertainties in scenarios and models of socio-ecological systems and ways of addressing them
The uncertainties are categorized as scenario, model, and decision uncertainties (see Table 1). More details about the numbered methods for addressing uncertainties are provided in Table S1. The color coding refers to the sources of uncertainty (see Table 1), with the gradient-shaded boxes indicating methods that apply to more than one uncertainty source.

socio-ecological systems actually derives from current and future decisions and, thus, from a decision-maker or citizen point of view, represents less of an "uncertainty" than our "societal leeway" or choices. Disentangling and documenting the different sources of uncertainties in socio-ecological systems is critical in allowing the design and initiation of informed and efficient actions. Many things about the future will always be uncertain, but we may wish to avoid the foolish and the fanatical by adopting the wisdom of doubt. Data and knowledge about socio-ecological systems are increasing rapidly, and knowledge improvement is often concomitant with awareness raising about system complexity. This leads to the paradox that, as technical knowledge increases, what we ignore is increasingly more important than what we know.

Uncertainty in science should not imply uncertainty in making decisions that respond to environmental problems.[203] Ironically, scientists see the quantification of uncertainty as underpinning scientific rigor, whereas others see it as a sign of weakness in the underlying science.[204] Too often, such a fallacy has become a flawed means of discouraging the endorsement of policies against environmental problems, such as climate change or biodiversity. Knowledge of uncertainty should inspire action rather than indifference and guide decision-making, rather than prevent it.[203]

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.oneear.2021.06.003.

ACKNOWLEDGMENTS

This work is part of the USBIO project funded by the LabEx CeMEB (ANR-10-LABX-04-01). M.D.A.R., A.A., and C.B. would like to acknowledge the funding of the Helmholtz Association. Y.-J.S. acknowledges funding support from the Biodiversa and Belmont Forum project SOMBEE (BiodivScen ERA-Net COFUND program, ANR contract no. ANR-18-EBI4-0003-01), the European Union's Horizon 2020 research and innovation program under grant agreements no. 869300 (FutureMARES) and no. 817578 (TRIATLAS), and the Pew Fellows Program in Marine Conservation at The Pew Charitable Trusts. R.P.

was partly funded by the FACCE Era-Gas project 3DForMod (ANR contract no. ANR-17-EGAS-0002-01). O.G. was partly funded by CNRS "Mission pour l'interdisciplinarité" through its "Osez l'interdisciplinarité" call and the French National Research Agency (grant ANR-16-CE02-0007). W.W.L.C. acknowledges funding support from NSERC Discovery Grant (RGPIN-2018-03864). B.A.W acknowledges funding from the Australian Research Council under ARC DP170104795. Natasha Cadenhead assisted in production of Figure 2.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptualization, writing, and editing of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Russell, B. (1945). History of Western Philosophy (Simon & Schuster).
2. IPBES (2018). The IPBES regional assessment report on biodiversity and ecosystem services for Europe and Central Asia. In) (IPBES).
3. IPBES (2016). The methodological assessment report on scenarios and models of biodiversity and ecosystem services (IPBES).
4. Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-Being: Synthesis (Island Press).
5. Rounsevell, M.D.A., and Metzger, M.J. (2010). Developing qualitative scenario storylines for environmental change assessment. Wiley Interdiscip. Rev. Clim. Change *1*, 606–619.
6. Acosta, L.A., Wintle, B.A., Benedek, Z., Chhetri, P.B., Heymans, S.J., Onur, A.C., Painter, R.L., Razafimpahanana, A., and Shoyama, K. (2016). Using scenarios and models to inform decision making in policy design and implementation. In IPBES, 2016: Methodological Assessment of Scenarios and Models of Biodiversity and Ecosystem Services, S. Ferrier, K.N. Ninan, P. Leadley, R. Alkemade, L.A. Acosta, H.R. Akçakaya, L. Brotons, W.W.L. Cheung, V. Christensen, and K.A. Harhash, et al., eds. (IPBES), pp. 35–81.
7. Alexander, P., Prestele, R., Verburg, P.H., Arneth, A., Baranzelli, C., Batista e Silva, F., Brown, C., Butler, A., Calvin, K., Dendoncker, N., et al. (2017). Assessing uncertainties in land cover projections. Glob. Change Biol. *23*, 767–781.
8. Brown, C., Seo, B., and Rounsevell, M. (2019). Societal breakdown as an emergent property of large-scale behavioural models of land use change. Earth Syst. Dyn. *10*, 809–845.
9. Tittensor, D.P., Eddy, T.D., Lotze, H.K., Galbraith, E.D., Cheung, W., Barange, M., Blanchard, J.L., Bopp, L., Bryndum-Buchholz, A., Büchner, M., et al. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: fish-MIP v1.0. Geoscientific Model. Dev. *11*, 1421–1442.
10. Travers, M., Shin, Y.J., Jennings, S., and Cury, P. (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. Prog. Oceanography *75*, 751–770.
11. Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V., Smith, M.J., Scharlemann, J.P.W., and Purves, D.W. (2014). Emergent global patterns of ecosystem structure and function from a mechanistic general ecosystem model. PLoS Biol. *12*, e1001841. https://doi.org/10.1371/journal.pbio.1001841.
12. Pavlick, R., Drewry, D.T., Bohn, K., Reu, B., and Kleidon, A. (2013). The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. Biogeosciences *10*, 4137–4177.
13. Prentice, I.C., Bondeau, A., Cramer, W., Harrison, S.P., Hickler, T., Lucht, W., Sitch, S., Smith, B., and Sykes, M.T. (2007). Dynamic global vegetation modeling: quantifying terrestrial ecosystem responses to large-scale environmental change. In Terrestrial Ecosystems in a Changing World (Springer Berlin Heidelberg)), pp. 175–192. https://doi.org/10.1007/978-3-540-32730-1_15.
14. Daniel, C.J., Frid, L., and Sleeter, B.M. (2016). State-and-transition simulation models: a framework for forecasting landscape change. Methods Ecol. Evol. *7*, 1413–1423. https://doi.org/10.1111/2041-210X.12597.
15. Buisson, L., Thuiller, W., Casajus, N., Lek, S., and Grenouillet, G. (2010). Uncertainty in ensemble forecasting of species distribution. Glob. Change Biol. *16*, 1145–1157. https://doi.org/10.1111/j.1365-2486.2009.02000.x.
16. Medlyn, B.E., De Kauwe, M.G., Zaehle, S., Walker, A.P., Duursma, R.A., Luus, K., Mishurov, M., Pak, B., Smith, B., Wang, Y.-P., et al. (2016). Using models to guide field experiments: *a priori* predictions for the CO_2 response of a nutrient- and water-limited native Eucalypt woodland. Glob. Change Biol. *22*, 2834–2851.
17. Cury, P.M., Shin, Y.J., Planque, B., Durant, J.M., Fromentin, J.M., Kramer-Schadt, S., Stenseth, N.C., Travers, M., and Grimm, V. (2008). Ecosystem oceanography for global change in fisheries. Trends Ecol. Evol. *23*, 338–346.
18. Pereira, H.M., Leadley, P.W., Proença, V., Alkemade, R., Scharlemann, J.P.W., Fernandez-Manjarrés, J.F., Araújo, M.B., Balvanera, P., Biggs, R., Cheung, W.W.L., et al. (2010). Scenarios for global biodiversity in the 21st century. Science *330*, 1496–1501.
19. Elsawah, S., Hamilton, S.H., Jakeman, A.J., Rothman, D., Schweizer, V., Trutnevyte, E., Carlsen, H., Drakes, C., Frame, B., Fu, B., et al. (2020). Scenario processes for socio-environmental systems analysis of futures: a review of recent efforts and a salient research agenda for supporting decision making. Sci. Total Environ. *729*, 138393. https://doi.org/10.1016/j.scitotenv.2020.138393.
20. Regan, H.M., Colyvan, M., and Burgman, M.A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecol. Appl. *12*, 618–628.
21. O'Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R., and van Vuuren, D.P. (2014). A new scenario framework for climate change research: the concept of shared socioeconomic pathways. Climatic Change *122*, 387–400.
22. Madin, J.S., Bowers, S., Schildhauer, M.P., and Jones, M.B. (2008). Advancing ecological research with ontologies. Trends Ecol. Evol. *23*, 159–168.
23. Sleeman, J., Finin, T., and Halem, M. (2018). Ontology-grounded topic modeling for climate science research. arXiv, arXiv:1807.10965v2.
24. van Vuuren, D.P., Kriegler, E., O'Neill, B.C., Ebi, K.L., Riahi, K., Carter, T.R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R., et al. (2014). A new scenario framework for climate change research: scenario matrix architecture. Climatic Change *122*, 373–386.
25. Metzger, M.J., Schröter, D., Leemans, R., and Cramer, W. (2008). A spatially explicit and quantitative vulnerability assessment of ecosystem service change in Europe. Reg. Environ. Change *8*, 91–107.
26. Maier, H.R., Guillaume, J.H.A., van Delden, H., Riddell, G.A., Haasnoot, M., and Kwakkel, J.H. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: how do they fit together? Environ. Model. Softw. *81*, 154–164.
27. Trutnevyte, E., Guivarch, C., Lempert, R., and Strachan, N. (2016). Reinvigorating the scenario technique to expand uncertainty consideration. Climatic Change *135*, 373–379.
28. Taleb, N.N. (2007). The Black Swan: The Impact of the Highly Improbable (Random house), p. 400.
29. Metzger, M.J., Rounsevell, M.D.A., den Heiligenberg, H.A.R.M., Pérez-Soba, M., and Hardiman, P.S. (2010). How personal judgment influences scenario development: an example for future rural development in Europe. Ecol. Soc. *15*, 5. http://www.ecologyandsociety.org/vol15/iss2/art5/.
30. Henry, R.C., Alexander, P., Rabin, S., Anthoni, P., Rounsevell, M.D.A., and Arneth, A. (2019). The role of global dietary transitions for safeguarding biodiversity. Glob. Environ. Change *58*, 101956.
31. Vuuren, D.P. Van, Stehfest, E., Gernaat, D.E.H.J., Van Den Berg, M., Bijl, D.L., Boer, H.S. De, Daioglou, V., Doelman, J.C., Edelenbosch, O.Y., Harmsen, M., et al. (2018). The need for negative emission technologies. Nat. Clim. Change *8*, 391–397. https://doi.org/10.1038/s41558-018-0119-8.
32. Brown, C., Alexander, P., Arneth, A., Holman, I., and Rounsevell, M. (2019). Achievement of Paris climate goals unlikely due to time lags in the land system. Nat. Clim. Change *9*, 203–208. https://doi.org/10.1038/s41558-019-0400-5.
33. Lotze, H.K., Tittensor, D.P., Bryndum-Buchholz, A., Eddy, T.D., Cheung, W.W.L., Galbraith, E.D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J.L., et al. (2019). Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. Proc. Natl. Acad. Sci. *116*, 12907–12912.
34. Gaines, S.D., Costello, C., Owashi, B., Mangin, T., Bone, J., Molinos, J.G., Burden, M., Dennis, H., Halpern, B.S., Kappel, C.V., et al. (2018). Improved fisheries management could offset many negative effects of climate change. Sci. Adv. *4*, 1–9.

35. Dueri, S., Guillotreau, P., Jiménez-Toribio, R., Oliveros Ramos, R., Bopp, L., and Maury, O. (2016). Food security, biomass conservation or economic profitability? Projecting the effects of climate and socio-economic changes on the global skipjack tuna fisheries under various management strategies. Glob. Environ. Change *41*, 1–12.

36. Maury, O., Campling, L., Arrizabalaga, H., Aumont, O., Bopp, L., Merino, G., Squires, D., Cheung, W., Goujon, M., Guivarch, C., et al. (2017). From shared socio-economic pathways (SSPs) to oceanic system pathways (OSPs): building policy-relevant scenarios for global oceanic ecosystems and fisheries. Glob. Environ. Change *45*, 203–216.

37. Harrison, P.A., Dunford, R.W., Holman, I.P., and Rounsevell, M.D.A. (2016). Climate change impact modelling needs to include cross-sectoral interactions. Nat. Clim. Change *6*, 885–890.

38. Kok, K., Bärlund, I., Flörke, M., Holman, I., Gramberger, M., Sendzimir, J., Stuch, B., and Zellmer, K. (2014). European participatory scenario development: strengthening the link between stories and models. Climatic Change *128*, 187–200.

39. Planque, B., Mullon, C., Arneberg, P., Eide, A., Fromentin, J.M., Heymans, J.J., Hoel, A.H., Niiranen, S., Ottersen, G., Sandø, A.B., et al. (2019). A participatory scenario method to explore the future of marine social-ecological systems. Fish Fish. *20*, 434–451.

40. Gopnik, M., Fieseler, C., Cantral, L., McClellan, K., Pendleton, L., and Crowder, L. (2012). Coming to the table: early stakeholder engagement in marine spatial planning. Mar. Pol. *36*, 1139–1149.

41. Pedde, S., Kok, K., Onigkeit, J., Brown, C., Holman, I., and Harrison, P.A. (2019). Bridging uncertainty concepts across narratives and simulations in environmental scenarios. Reg. Environ. Change *19*, 655–666.

42. Henry, R.C., Engström, K., Olin, S., Alexander, P., Arneth, A., and Rounsevell, M.D.A. (2018). Food supply and bioenergy production within the global cropland planetary boundary. PLoS ONE *13*, 1–17.

43. Brown, C., Brown, E., Murray-Rust, D., Cojocaru, G., Savin, C., and Rounsevell, M. (2014). Analysing uncertainties in climate change impact assessment across sectors and scenarios. Climatic Change *128*, 293–306.

44. Engström, K., Olin, S., Rounsevell, M.D.A., Brogaard, S., Van Vuuren, D.P., Alexander, P., Murray-Rust, D., and Arneth, A. (2016). Assessing uncertainties in global cropland futures using a conditional probabilistic modelling framework. Earth Syst. Dyn. *7*, 893–915.

45. Beulke, S., Brown, C.D., Dubus, I.G., Galicia, H., Jarvis, N., Schaefer, D., and Trevisan, M. (2006). User subjectivity in Monte Carlo modeling of pesticide exposure. Environ. Toxicol. Chem. Int. J. *25*, 2227–2236.

46. Costello, C., Ovando, D., Clavelle, T., Strauss, C.K., Hilborn, R., Melnychuk, M.C., Branch, T.A., Gaines, S.D., Szuwalski, C.S., Cabral, R.B., et al. (2016). Global fishery prospects under contrasting management regimes. Proc. Natl. Acad. Sci. *113*, 5125–5129.

47. Rousseau, Y., Watson, R.A., Blanchard, J.L., and Fulton, E.A. (2019). Evolution of global marine fishing fleets and the response of fished resources. Proc. Natl. Acad. Sci. *116*, 12238–12243.

48. Holman, I.P., Brown, C., Carter, T.R., Harrison, P.A., and Rounsevell, M. (2019). Improving the representation of adaptation in climate change impact models. Reg. Environ. Change *19*, 711–721.

49. Levins, R. (1966). The strategy of model building in population biology. Am. scientist *54*, 421–431.

50. Prentice, I.C., Liang, X., Medlyn, B.E., and Wang, Y.-P. (2015). Reliable, robust and realistic: the three R's of next-generation land-surface modelling. Atmos. Chem. Phys. *15*, 5987–6005.

51. Evans, M.R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C.M., Merz, M., O'Malley, M.A., Orzack, S.H., Weisberg, M., et al. (2013). Do simple models lead to generality in ecology? Trends Ecol. Evol. *28*, 578–583.

52. Yatat, V., Tchuinté, A., Dumont, Y., and Couteron, P. (2018). A tribute to the use of minimalistic spatially-implicit models of savanna vegetation dynamics to address broad spatial scales in spite of scarce data. Biomath *7*, 1812167.

53. Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S., Legay, M., Pagé, C., et al. (2012). Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. Ecol. Lett. *15*, 533–544.

54. Bugmann, H., Seidl, R., Hartig, F., Bohn, F., Brüna, J., Cailleret, M., François, L., Heinke, J., Henrot, A.J., Hickler, T., et al. (2019). Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. Ecosphere *10*. https://doi.org/10.1002/ecs2.2616.

55. De Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P., and Steppe, K. (2012). Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model. Geoscientific Model. Dev. *5*, 1091–1108.

56. Robinson, D.T., Di Vittorio, A., Alexander, P., Arneth, A., Michael Barton, C., Brown, D.G., Kettner, A., Lemmen, C., O'Neill, B.C., Janssen, M., et al. (2018). Modelling feedbacks between human and natural processes in the land system. Earth Syst. Dyn. *9*, 895–914.

57. Xu, X., Medvigy, D., Powers, J.S., Becknell, J.M., and Guan, K. (2016). Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. New Phytol. *212*, 80–95.

58. Fisher, R.A., Koven, C.D., Anderegg, W.R.L., Christoffersen, B.O., Dietze, M.C., Farrior, C.E., Holm, J.A., Hurtt, G.C., Knox, R.G., Lawrence, P.J., et al. (2018). Vegetation demographics in Earth System Models: a review of progress and priorities. Glob. Change Biol. *24*, 35–54.

59. Moorcroft, P.R. (2006). How close are we to a predictive science of the biosphere? Trends Ecol. Evol. *21*, 400–407.

60. Purves, D., and Pacala, S. (2008). Predictive models of forest dynamics. Science *320*, 1452–1453.

61. Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J., and Thonicke, K. (2015). Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. Glob. Change Biol. *21*, 2711–2725. https://doi.org/10.1111/gcb.12870.

62. Pavlick, R., Drewry, D.T., Bohn, K., Reu, B., and Kleidon, A. (2013). The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. Biogeosciences *10*, 4137–4177. https://doi.org/10.5194/bg-10-4137-2013.

63. Maréchaux, I., and Chave, J. (2017). An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. Ecol. Monogr. *87*, 632–664. https://doi.org/10.1002/ecm.1271.

64. Van Bodegom, P.M., Douma, J.C., Witte, J.P.M., Ordoñez, J.C., Bartholomeus, R.P., and Aerts, R. (2012). Going beyond limitations of plant functional types when predicting global ecosystem-atmosphere fluxes: exploring the merits of traits-based approaches. Glob. Ecol. Biogeogr. *21*, 625–636. https://doi.org/10.1111/j.1466-8238.2011.00717.x.

65. Mokany, K., Ferrier, S., Connolly, S.R., Dunstan, P.K., Fulton, E.A., Harfoot, M.B., Harwood, T.D., Richardson, A.J., Roxburgh, S.H., Scharlemann, J.P.W., et al. (2016). Integrating modelling of biodiversity composition and ecosystem function. Oikos *125*, 10–19.

66. Sakschewski, B., Von Bloh, W., Boit, A., Poorter, L., Peña-Claros, M., Heinke, J., Joshi, J., and Thonicke, K. (2016). Resilience of Amazon forests emerges from plant trait diversity. Nat. Clim. Change *6*, 1032–1036.

67. Voinov, A., and Shugart, H.H. (2013). Integronsters', integral and integrated modeling. Environ. Model. Softw. *39*, 149–158.

68. Jarnevich, C.S., Cullinane Thomas, C., Young, N.E., Backer, D., Cline, S., Frid, L., and Grissom, P. (2019). Developing an expert elicited simulation model to evaluate invasive species and fire management alternatives. Ecosphere *10* (5), e02730. https://doi.org/10.1002/ecs2.2730.

69. Holman, I.P., Brown, C., Carter, T.R., Harrison, P.A., and Rounsevell, M. (2018). Improving the representation of adaptation in climate change impact models. Reg. Environ. Change *19*, 711–721. https://doi.org/10.1007/s10113-018-1328-4.

70. Brown, C., Alexander, P., Holzhauer, S., and Rounsevell, M.D.A. (2017). Behavioral models of climate change adaptation and mitigation in land-based sectors. Wiley Interdiscip. Rev. Clim. Change *8*, e448. https://doi.org/10.1002/wcc.448.

71. Van Nes, E.H., and Scheffer, M. (2005). A strategy to improve the contribution of complex simulation models to ecological theory. Ecol. Model. *185*, 153–164.

72. De Weirdt, M., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P., and Steppe, K. (2012). Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model. Geoscientific Model. Dev. *5*, 1091–1108.

73. Joetzjer, E., Pillet, M., Ciais, P., Barbier, N., Chave, J., Schlund, M., Maignan, F., Barichivich, J., Luyssaert, S., Hérault, B., et al. (2017). Assimilating satellite-based canopy height within an ecosystem model to estimate aboveground forest biomass. Geophys. Res. Lett. *44*, 6823–6832. https://doi.org/10.1002/2017GL074150.

74. Naudts, K., Ryder, J., McGrath, M.J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., Campioli, M., et al. (2015). A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes. Geoscientific Model. Dev. *8*, 2035–2065.

75. Kim, H., Rosa, I.M.D., Alkemade, R., Leadley, P., Hurtt, G., Popp, A., Anthoni, P., Arneth, A., Baisero, D., Caton, E., et al. (2018). A protocol for an intercomparison of biodiversity and ecosystem services models using harmonized land-use and climate scenarios. Geosciences Model. Dev. *11*, 4537–4562. https://doi.org/10.5194/gmd-11-4537-2018.

76. Rounsevell, M.D.A., Arneth, A., Alexander, P., Brown, D.G., de Noblet-Ducoudré, N., Ellis, E., Finnigan, J., Galvin, K., Grigg, N., Harman, I., et al. (2014). Towards decision-based global land use models for improved understanding of the Earth system. Earth Syst. Dyn. *5*, 117–137.

77. Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S., Legay, M., Pagé, C., et al. (2012). Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. Ecol. Lett. *15*, 533–544.

78. Powell, T.L., Galbraith, D.R., Christoffersen, B.O., Harper, A., Imbuzeiro, H.M.A., Rowland, L., Almeida, S., Brando, P.M., da Costa, A.C.L., Costa, M.H., et al. (2013). Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought. New Phytol. *200*, 350–365.

79. Restrepo-Coupe, N., Levine, N.M., Christoffersen, B.O., Albert, L.P., Wu, J., Costa, M.H., Galbraith, D., Imbuzeiro, H., Martins, G., da Araujo, A.C., et al. (2017). Do dynamic global vegetation models capture the seasonality of carbon fluxes in the Amazon basin? A data-model intercomparison. Glob. Change Biol. *23*, 191–208.

80. Taylor, K.E., Stouffer, R.J., and Meehl, G.A. (2012). An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. *93*, 485–498.

81. Tittensor, D.P., Walpole, M., Hill, S.L.L., Boyce, D.G., Britten, G.L., Burgess, N.D., Butchart, S.H.M., Leadley, P.W., Regan, E.C., Alkemade, R., et al. (2014). A mid-term analysis of progress toward international biodiversity targets. Science *346*, 241–244.

82. Kelley, D.I., Prentice, I.C., Harrison, S.P., Wang, H., Simard, M., Fisher, J.B., and Willis, K.O. (2013). A comprehensive benchmarking system for evaluating global vegetation models. Biogeosciences *10*, 3313–3340.

83. Luo, Y.Q., Randerson, J.T., Abramowitz, C., Bacour, e., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J.B., Fisher, R., et al. (2012). A framework of benchmarking land models. Biogeosciences *10*, 3857–3874.

84. Arneth, A., Niinemets, Ü., Pressley, S., Bäck, J., Hari, P., Karl, T., Noe, S., Prentice, I.C., Serça, D., Hickler, T., et al. (2007). Process-based estimates of terrestrial ecosystem isoprene emissions: incorporating the effects of a direct CO_2-isoprene interaction. Atmos. Chem. Phys. *7*, 31–53.

85. De Kauwe, M.G., Medlyn, B.E., Zaehle, S., Walker, A.P., Dietze, M.C., Hickler, T., Jain, A.K., Luo, Y., Parton, W.J., Prentice, I.C., et al. (2013). Forest water use and water use efficiency at elevated CO_2: a model-data intercomparison at two contrasting temperate forest FACE sites. Glob. Change Biol. *19*, 1759–1779.

86. Hickler, T., Smith, B., Prentice, I.C., Mjöfors, K., Miller, P., Arneth, A., and Sykes, M.T. (2008). CO_2 fertilization in temperate FACE experiments not representative of boreal and tropical forests. Glob. Change Biol. *14*, 1531–1542.

87. Morales, P., Sykes, M.T., Prentice, I.C., Smith, P., Smith, B., Bugmann, H., Zierl, B., Friedlingstein, P., Viovy, N., Sabate, S., et al. (2005). Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. Glob. Change Biol. *11*, 2211–2233.

88. Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., Holmér, J., and Arneth, A. (2015). Modelling the response of yields and tissue C:N to changes in atmospheric CO_2 and N management in the main wheat regions of western Europe. Biogeosciences *12*, 2489–2515.

89. Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. Biogeosciences *11*, 2027–2054.

90. Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R., and Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. Science *355*, 385–389. https://doi.org/10.1126/science.aaj1987.

91. Tanentzap, A.J., and Coomes, D.A. (2012). Carbon storage in terrestrial ecosystems: do browsing and grazing herbivores matter? Biol. Rev. *87*, 72–94.

92. Walker, A.P., Zaehle, S., Medlyn, B.E., De Kauwe, M.G., Asao, S., Hickler, T., Parton, W., Ricciuto, D.M., Wang, Y.-P., Wårlind, D., et al. (2015). Predicting long-term carbon sequestration in response to CO_2 enrichment: how and why do current ecosystem models differ? Glob. Biogeochem. Cycles *29*, 476–495.

93. Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley, D.M., et al. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Change Biol. *24*, 2416–2433.

94. Parker, W.S. (2013). Ensemble modeling, uncertainty and robust predictions. Wiley Interdiscip. Reviews-Climate Change *4*, 213–223.

95. Booth, B.B.B., Jones, C.D., Collins, M., Totterdell, I.J., Cox, P.M., Sitch, S., Huntingford, C., Betts, R., Harris, G.R., and Lloyd, J. (2012). High sensitivity of future global warming to land carbon cycle uncertainties. Environ. Res. Lett. *7*. https://doi.org/10.1088/1748-9326/7/2/024002.

96. Lienert, S., and Joos, F. (2018). A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions. Biogeosciences *15*, 2909–2930.

97. Wramneby, A., Smith, B., Zaehle, S., and Sykes, M.T. (2008). Parameter uncertainties in the modelling of vegetation dynamics—effects on tree community structure and ecosystem functioning in European forest biomes. Ecol. Model. *216*, 277–290.

98. Zaehle, S., Sitch, S., Smith, B., and Hatterman, F. (2005). Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. Glob. Biogeochem. Cycles *19*, GB3020. https://doi.org/10.1029/2004GB002395.

99. Lawrence, D.M., Hurtt, G.C., Arneth, A., Brovkin, V., Calvin, K.V., Jones, A.D., Jones, C.D., Lawrence, P.J., de Noblet-Ducoudre, N., Pongratz, J., et al. (2016). The land use model intercomparison project (LUMIP) contribution to CMIP6: rationale and experimental design. Geoscientific Model. Dev. *9*, 2973–2998.

100. Le Quéré, C., Andrew, R.M., Canadell, J.G., Sitch, S., Korsbakken, J.I., Peters, G.P., Manning, A.C., Boden, T.A., Tans, P.P., Houghton, R.A., et al. (2016). Global carbon budget 2016. Earth Syst. Sci. Data *8*, 605–649.

101. Rabin, S.S., Melton, J.R., Lasslop, G., Bachelet, D., Forrest, M., Hantson, S., Kaplan, J.O., Li, F., Mangeon, S., Ward, D.S., et al. (2017). The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions. Geoscientific Model. Dev. *10*, 1175–1197.

102. Sitch, S., Friedlingstein, P., Gruber, N., Jones, S.D., Murray-Tortarolo, G., Ahlström, A., Doney, S.C., Graven, H., Heinze, C., Huntingford, C., et al. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. Biogeosciences *12*, 653–679.

103. Lotze, H.K., Tittensor, D.P., Bryndum-Buchholz, A., Eddy, T.D., Cheung, W.W., Galbraith, E.D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J.L., et al. (2018). Ensemble projections of global ocean animal biomass with climate change. bioRxiv, 467175.

104. Prestele, R., Alexander, P., Rounsevell, M.D.A., Arneth, A., Calvin, K., Doelman, J., Eitelberg, D.A., Engström, K., Fujimori, S., Hasegawa, T., et al. (2016). Hotspots of uncertainty in land-use and land-cover change projections: a global-scale model comparison. Glob. Change Biol. *22*, 3967–3983. https://doi.org/10.1111/gcb.13337.

105. Le Quéré, C., Andrew, R.M., Canadell, J.G., Sitch, S., Ivar Korsbakken, J., Peters, G.P., Manning, A.C., Boden, T.A., Tans, P.P., Houghton, R.A., et al. (2018). Global carbon budget 2018. Earth Syst. Sci. Data *10*, 2141–2194. https://doi.org/10.5194/essd-10-2141-2018.

106. Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., et al. (2015). Predictive ecology in a changing world. J. Appl. Ecol. *52*, 1293–1310. https://doi.org/10.1111/1365-2664.12482.

107. Beven, K., and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. *249*, 11–29.

108. Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., and Zhang, L. (2009). Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. Ecol. Appl. *19*, 571–574.

109. Rykiel, E.J. (1996). Testing ecological models: the meaning of validation. Ecol. Model. *90*, 229–244.

110. Medlyn, B.E., Robinson, A.P., Clement, R., and McMurtrie, R.E. (2005). On the validation of models of forest CO_2 exchange using eddy covariance data: some perils and pitfalls. Tree Physiol. *25*, 839–857.

111. van Vliet, J., Bregt, A.K., Brown, D.G., van Delden, H., Heckbert, S., and Verburg, P.H. (2016). A review of current calibration and validation practices in land-change modeling. Environ. Model. Softw. *82*, 174–182.

112. Remesan, R., and Holman, I.P. (2015). Effect of baseline meteorological data selection on hydrological modelling of climate change scenarios. J. Hydrol. *528*, 631–642.

113. Remesan, R., Begam, S., and Holman, I.P. (2019). Effect of baseline snowpack assumptions in the HySIM model in predicting future hydrological behaviour of a Himalayan catchment. Hydrol. Res. *50*, 691–708.

114. Urban, M.C., Bocedi, G., Hendry, A.P., Mihoub, J.B., Pe'er, G., Singer, A., Bridle, J.R., Crozier, L.G., De Meester, L., Godsoe, W., et al. (2016). Improving the forecast for biodiversity under climate change. Science *353*, aad8466. https://doi.org/10.1126/science.aad8466.

115. Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P., and Saatchi, S. (2008). Measuring and modelling biodiversity from space. Prog. Phys. Geogr. *32*, 203–221. https://doi.org/10.1177/0309133308093606.

116. Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., et al. (2013). Essential biodiversity variables. Science *339*, 277–278. https://doi.org/10.1126/science.1229931.

117. Muller-Karger, F.E., Miloslavich, P., Bax, N.J., Simmons, S., Costello, M.J., Sousa Pinto, I., Canonico, G., Turner, W., Gill, M., Montes, E., et al. (2018). Advancing marine biological observations and data requirements of the complementary essential ocean variables (EOVs) and essential biodiversity variables (EBVs) frameworks. Front. Mar. Sci. *5*, 1–15.

118. Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley, D.M., et al. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Change Biol. *24*, 2416–2433.

119. Affouard, A., Goëau, H., Bonnet, P., Lombardo, J.-C., Joly, A., and Goeau, H. (2017). Pl@ntNet App in the Era of Deep learning. ICLR: International Conference on Learning Representations (Toulon, France), hal-01629195.

120. Delbart, N., Beaubien, E., Kergoat, L., and Le Toan, T. (2015). Comparing land surface phenology with leafing and flowering observations from the PlantWatch citizen network. Remote Sensing Environ. *160*, 273–280. https://doi.org/10.1016/j.rse.2015.01.012.

121. Giraud, C., Calenge, C., Coron, C., and Julliard, R. (2016). Capitalizing on opportunistic data for monitoring relative abundances of species. Biometrics *72*, 649–658. 10.1111/biom.12431.

122. Elith, J., and Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. *40*, 677–697.

123. Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. Ecol. Lett. *8*, 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x.

124. Buisson, L., Thuiller, W., Casajus, N., Lek, S., and Grenouillet, G. (2010). Uncertainty in ensemble forecasting of species distribution. Glob. Change Biol. *16*, 1145–1157.

125. Jiménez-Valverde, A., Lobo, J.M., and Hortal, J. (2008). Not as good as they seem: the importance of concepts in species distribution modelling. Divers. Distributions *14*, 885–890. https://doi.org/10.1111/j.1472-4642.2008.00496.x.

126. Barry, S., and Elith, J. (2006). Error and uncertainty in habitat models. J. Appl. Ecol. *43*, 413–423. https://doi.org/10.1111/j.1365-2664.2006.01136.x.

127. Bean, W.T., Stafford, R., and Brashares, J.S. (2012). The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography *35*, 250–258. https://doi.org/10.1111/j.1600-0587.2011.06545.x.

128. Lobo, J.M., Jiménez-Valverde, A., and Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. Ecography *33*, 103–114. https://doi.org/10.1111/j.1600-0587.2009.06039.x.

129. Syfert, M.M., Smith, M.J., and Coomes, D.A. (2013). The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. PLoS ONE *8*, e55158. https://doi.org/10.1371/journal.pone.0055158.

130. Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., and Garnier, E. (2007). Let the concept of trait be functional!. Oikos *116*, 882–892.

131. Marechaux, I., and Chave, J. (2017). An individual-based forest model to jointly simulate carbon and tree diversity in Amazonia: description and applications. Ecol. Monogr. *87*, 632–664. https://doi.org/10.1002/ecm.1271.

132. Pollock, L.J., Morris, W.K., and Vesk, P.A. (2012). The role of functional traits in species distributions revealed through a hierarchical model. Ecography *35*, 716–725.

133. Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., Peñuelas, J., and Thonicke, K. (2015). Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. Glob. Change Biol. *21*, 2711–2725.

134. Guiet, J., Aumont, O., Poggiale, J.C., and Maury, O. (2016). Effects of lower trophic level biomass and water temperature on fish communities: a modelling study. Prog. Oceanography *146*, 22–37.

135. Blanchard, J.L., Jennings, S., Holmes, R., Harle, J., Merino, G., Allen, J.I., Holt, J., Dulvy, N.K., and Barange, M. (2012). Potential consequences of climate change for primary production and fish production in large marine ecosystems. Philosophical Trans. R. Soc. B *367*, 2979–2989.

136. Moullec, F., Velez, L., Verley, P., Barrier, N., Ulses, C., Carbonara, P., Esteban, A., Follesa, C., Gristina, M., Jadaud, A., et al. (2019). Capturing the big picture of Mediterranean marine biodiversity with an end- to-end model of climate and fishing impacts. Prog. Oceanography *178*, 102179.

137. Allen, W.L., Street, S.E., and Capellini, I. (2017). Fast life history traits promote invasion success in amphibians and reptiles. Ecol. Lett. *20*, 222–230.

138. Parr, C.S., Wilson, N., Leary, P., Schulz, K., Lans, K., Walley, L., Hammock, J., Goddard, A., Rice, J., Studer, M., et al. (2014). The Encyclopedia of Life v2: providing global access to knowledge about life on Earth. Biodiversity Data J. *2*, e1079.

139. Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., et al. (2011). Try—a global database of plant traits. Glob. Change Biol. *17*, 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x.

140. Griffin-Nolan, R.J., Bushey, J.A., Carroll, C.J.W., Challis, A., Chieppa, J., Garbowski, M., Hoffman, A.M., Post, A.K., Slette, I.J., Spitzer, D., et al. (2018). Trait selection and community weighting are key to understanding ecosystem responses to changing precipitation regimes. Funct. Ecol. *32*, 1746–1756. https://doi.org/10.1111/1365-2435.13135.

141. Paine, C.E.T., Deasey, A., and Duthie, A.B. (2018). Towards the general mechanistic prediction of community dynamics. Funct. Ecol. *32*, 1681–1692. https://doi.org/10.1111/1365-2435.13096.

142. Shipley, B., De Bello, F., Cornelissen, J.H.C., Laliberté, E., Laughlin, D.C., and Reich, P.B. (2016). Reinforcing loose foundation stones in trait-based plant ecology. Oecologia *180*, 923–931. https://doi.org/10.1007/s00442-016-3549-x.

143. Adler, P.B., Smull, D., Beard, K.H., Choi, R.T., Furniss, T., Kulmatiski, A., Meiners, J.M., Tredennick, A.T., and Veblen, K.E. (2018). Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. Ecol. Lett. *21*, 1319–1329. https://doi.org/10.1111/ele.13098.

144. Albert, C.H., Grassein, F., Schurr, F.M., Vieilledent, G., and Violle, C. (2011). When and how should intraspecific variability be considered in trait-based plant ecology? Perspect. Plant Ecol. Evol. Syst. *13*, 217–225.

145. Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., Rudolf, V.H.W., Schreiber, S.J., Urban, M.C., and Vasseur, D.A. (2011). Why intraspecific trait variation matters in community ecology. Trends Ecol. Evol. *26*, 183–192. https://doi.org/10.1016/j.tree.2011.01.009.

146. Clark, J.S., Dietze, M., Chakraborty, S., Agarwal, P.K., Ibanez, I., LaDeau, S., and Wolosin, M. (2007). Resolving the biodiversity paradox. Ecol. Lett. *10*, 647–659. https://doi.org/10.1111/j.1461-0248.2007.01041.x. PMID: 17594418.

147. Kissling, W.D., Walls, R., Bowser, A., Jones, M.O., Kattge, J., Agosti, D., Amengual, J., Basset, A., van Bodegom, P.M., Cornelissen, J.H.C., et al. (2018). Towards global data products of essential biodiversity variables on species traits. Nat. Ecol. Evol. *10*, 1531–1540. https://doi.org/10.1038/s41559-018-0667-3.

148. Payne, M.R., Barange, M., Cheung, W.W.L., MacKenzie, B.R., Batchelder, H.P., Cormon, X., Eddy, T.D., Fernandes, J.A., Hollowed, A.B., Jones, M.C., et al. (2015). Uncertainties in projecting climate-change impacts in marine ecosystems. ICES J. Mar. Sci. *73*, 1272–1282.

149. Hurtt, G.C., Fisk, J., Thomas, R.Q., Dubayah, R., Moorcroft, P.R., and Shugart, H.H. (2010). Linking models and data on vegetation structure. J. Geophys. Res. Biogeosciences *115*, G00E10. https://doi.org/10.1029/2009JG000937.

150. Rödig, E., Cuntz, M., Rammig, A., Fischer, R., Taubert, F., and Huth, A. (2018). The importance of forest structure for carbon fluxes of the Amazon rainforest. Environ. Res. Lett. *13*, 054013. https://doi.org/10.1088/1748-9326/aabc61.

151. Beale, C.M., and Lennon, J.J. (2012). Incorporating uncertainty in predictive species distribution modelling. Philos. Trans. R. Soc. B: Biol. Sci. *367*, 247–258. https://doi.org/10.1098/rstb.2011.0178.

152. Clark, J.S. (2003). Uncertainty and variability in demography and population growth: a hierarchical approach. Ecology *84*, 1370–1381.

153. Kebede, A.S., Dunford, R., Mokrech, M., Audsley, E., Harrison, P.A., Holman, I.P., Nicholls, R.J., Rickebusch, S., Rounsevell, M.D.A., Sabaté, S., et al. (2015). Direct and indirect impacts of climate and socio-economic change in Europe: a sensitivity analysis for key land- and water-based sectors. Climatic Change *128*, 261–277.

154. Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., and Wagener, T. (2016). Sensitivity analysis of environmental models: a systematic review with practical workflow. Environ. Model. Softw. *79*, 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008.

155. Urban, M.C., Bocedi, G., Hendry, A.P., Mihoub, J.B., Pe'er, G., Singer, A., Bridle, J.R., Crozier, L.G., De Meester, L., Godsoe, W., et al. (2016). Improving the forecast for biodiversity under climate change. Science *353*, aad8466. https://doi.org/10.1126/science.aad8466.

156. Williams, M., Richardson, A.D., Reichstein, M., Stoy, P.C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D.Y., Kattge, J., et al. (2009). Improving land surface models with FLUXNET data. Biogeosciences *6*, 1341–1359. https://doi.org/10.5194/bg-6-1341-2009.

157. Hartig, F., Dyke, J., Hickler, T., Higgins, S.I., O'Hara, R.B., Scheiter, S., and Huth, A. (2012). Connecting dynamic vegetation models to data—an inverse perspective. J. Biogeogr. *39*, 2240–2252. https://doi.org/10.1111/j.1365-2699.2012.02745.x.

158. Lagarrigues, G., Jabot, F., Lafond, V., and Courbaud, B. (2015). Approximate Bayesian computation to recalibrate individual-based models with population data: illustration with a forest simulation model. Ecol. Model. *306*, 278–286. https://doi.org/10.1016/j.ecolmodel.2014.09.023.

159. LeBauer, D.S., Wang, D., Richter, K.T., Davidson, C.C., and Dietze, M.C. (2013). Facilitating feedbacks between field measurements and ecosystem models. Ecol. Monogr. *83*, 133–154. https://doi.org/10.1890/12-0137.1.

160. Van Oijen, M., Rougier, J., and Smith, R. (2005). Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiol. *25*, 915–927.

161. Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., et al. (2001). FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. *82*. https://doi.org/10.1175/1520-0477.

162. Anderson-Teixeira, K.J., Davies, S.J., Bennett, A.C., Gonzalez-Akre, E.B., Muller-Landau, H.C., Joseph Wright, S., Abu Salim, K., Almeyda Zambrano, A.M., Alonso, A., Baltzer, J.L., et al. (2015). CTFS-ForestGEO: a worldwide network monitoring forests in an era of global change. Glob. Change Biol. *21*, 528–549.

163. Fischer, F.J., Maréchaux, I., and Chave, J. (2019). Improving plant allometry by fusing forest models and remote sensing. New Phytol. *223*, 1159–1165. https://doi.org/10.1111/nph.15810.

164. Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J.H.C., Reich, P., Bahn, M., et al. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. Remote Sensing Environ. *218*, 69–88. https://doi.org/10.1016/j.rse.2018.09.006.

165. Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K.Y., Forrester, D.I., Haeni, M., et al. (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Glob. Change Biol. *23*, 177–190. https://doi.org/10.1111/gcb.13388.

166. Ahlström, A., Schurgers, G., Arneth, A., and Smith, B. (2012). Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. Environ. Res. Lett. *7*, 44008.

167. Cheung, W.W.L., Frölicher, T.L., Asch, R.G., Jones, M.C., Pinsky, M.L., Reygondeau, G., Rodgers, K.B., Rykaczewski, R.R., Sarmiento, J.L., Stock, C., et al. (2016). Building confidence in projections of the responses of living marine resources to climate change. ICES J. Mar. Sci. *73*, 1283–1296.

168. Dunford, R., Harrison, P.A., and Rounsevell, M.D.A. (2015). Exploring scenario and model uncertainty in cross-sectoral integrated assessment approaches to climate change impacts. Climatic Change *132*, 417–432.

169. Patt, A.G., van Vuuren, D.P., Berkhout, F., Aaheim, A., Hof, A.F., Isaac, M., and Mechler, R. (2010). Adaptation in integrated assessment modeling: where do we stand? Climatic Change *99*, 383–402.

170. Lo, E. (2005). Gaussian error propagation applied to ecological data: post-ice-storm-downed woody biomass. Ecol. Monogr. *75*, 451–466. https://doi.org/10.1890/05-0030.

171. Hilborn, R., and Mangel, M. (1997). The Ecological Detective: Confronting Models with Data (Princeton University Press), p. 336.

172. Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., and Wu, Q. (2019). Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. Environ. Model. Softw. *114*, 29–39.

173. Clark, J.S. (2005). Why environmental scientists are becoming Bayesians. Ecol. Lett. *8*, 2–14. https://doi.org/10.1111/j.1461-0248.2004.00702.x.

174. Brown, C., Alexander, P., Arneth, A., Holman, I., and Rounsevell, M.D.A. (2019). Achievement of Paris climate goals unlikely due to time lags in the land system. Nat. Clim. Change *9*, 203–208. https://doi.org/10.1038/s41558-019-0400-5.

175. Low, S., and Schäfer, S. (2020). Is bio-energy carbon capture and storage (BECCS) feasible? The contested authority of integrated assessment modeling. Energy Res. Soc. Sci. *60*, 101326.

176. Holland, D.S. (2010). Management Strategy Evaluation and Management Procedures: Tools for Rebuilding and Sustaining Fisheries (OECD Publishing). OECD Food, Agriculture and Fisheries Working Papers, no. 25. https://doi.org/10.1787/5kmd77jhvkjf-en.

177. Regan, H.M., Ben-Haim, Y., Langford, B., Wilson, W.G., Lundberg, P., Andelman, S.J., and Burgman, M.A. (2005). Robust decision-making under severe uncertainty for conservation management. Ecol. Appl. *15*, 1471–1477.

178. Peterson, G.D., Cumming, G.S., and Carpenter, S.R. (2003). Scenario planning: a tool for conservation in an uncertain world. Conservation Biol. *17*, 358–366.

179. Allen, C.R., Angeler, D.G., Fontaine, J.J., Garmestani, A.S., Hart, N.M., Pope, K.L., and Twidwell, D. (2017). Adaptive management of rangeland systems. In Rangeland Systems: Processes, Management and Challenges, D.D. Briske, ed. (Springer), pp. 373–394.

180. Helgeson, C., Bradley, R., and Hill, B. (2018). Combining probability with qualitative degree-of-certainty metrics in assessment. Climatic Change *149*, 517–525.

181. Bradley, R., Helgeson, C., and Hill, B. (2017). Climate change assessments: confidence, probability, and decision. Philos. Sci. *84*, 500–522.

182. Visconti, P., Bakkenes, M., Baisero, D., Brooks, T., Butchart, S.H.M., Joppa, L., Alkemade, R., Di Marco, M., Santini, L., Hoffmann, M., et al. (2016). Projecting global biodiversity indicators under future development scenarios. Conservation Lett. *9*, 5–13.

183. Beckmann, M., Gerstner, K., Akin-Fajiye, M., Ceauşu, S., Kambach, S., Kinlock, N.L., Phillips, H.R.P., Verhagen, W., Gurevitch, J., Klotz, S., et al. (2019). Conventional land-use intensification reduces species richness and increases production: a global meta-analysis. Glob. Change Biol. *25*, 1941–1956.

184. Rounsevell, M.D.A., Harfoot, M., Harrison, P.A., Newbold, T., Gregory, R.D., and Mace, G.M. (2020). A biodiversity target based on species extinctions. Science *368*, 1193–1195.

185. Shin, Y.-J., Houle, J.E., Akoglu, E., Blanchard, J.L., Bundy, A., Coll, M., Demarcq, H., Fu, C., Fulton, E.A., Heymans, J.J., et al. (2018). The specificity of marine ecological indicators to fishing in the face of environmental change: a multi-model evaluation. Ecol. Indicators *89*, 317–326.

186. Lehuta Sigrid Mahevas Stephanie, L.F.P.P.P. (2013). A simulation-based approach to assess sensitivity and robustness of fisheries management indicators for the pelagic fishery in the Bay of Biscay. Can. J. Fish. Aquat. Sci. *70*, 1741–1756.

187. Schindler, D.E., Armstrong, J.B., and Reed, T.E. (2015). The portfolio concept in ecology and evolution. Front. Ecol. Environ. *13*, 257–263.

188. Shin, Y.J., Bundy, A., Shannon, L.J., Blanchard, J.L., Chuenpagdee, R., Coll, M., Knight, B., Lynam, C., Piet, G., and Richardson, A.J. (2012). Global in scope and regionally rich: an IndiSeas workshop helps shape the future of marine ecosystem indicators. Rev. Fish Biol. Fish. *22*, 835–845.

189. Fu, C., Xu, Y., Bundy, A., Grüss, A., Coll, M., Heymans, J.J., Fulton, E.A., Shannon, L., Halouani, G., Velez, L., et al. (2019). Making ecological indicators management ready: assessing the specificity, sensitivity, and threshold response of ecological indicators *105*, 16–28.

190. Bagstad, K.J., Semmens, D.J., Waage, S., and Winthrop, R. (2013). A comparative assessment of decision-support tools for ecosystem services quantification and valuation. Ecosystem Serv. *5*, 27–39.

191. Schulp, C.J.E., Burkhard, B., Maes, J., Van Vliet, J., and Verburg, P.H. (2014). Uncertainties in ecosystem service maps: a comparison on the European scale. PloS one *9*. https://doi.org/10.1371/journal.pone.0109643.

192. Watkiss, P., Hunt, A., Blyth, W., and Dyszynski, J. (2015). The use of new economic decision support tools for adaptation assessment: a review of

methods and applications, towards guidance on applicability. Climatic Change *132*, 401–416.

193. Wegner, G., and Pascual, U. (2011). Cost-benefit analysis in the context of ecosystem services for human well-being: a multidisciplinary critique. Glob. Environ. Change *21*, 492–504.
194. Minx, J.C., Callaghan, M., Lamb, W.F., Garard, J., and Edenhofer, O. (2017). Learning about climate change solutions in the IPCC and beyond. Environ. Sci. Pol. *77*, 252–259.
195. Watts, D.J. (2017). Should social science be more solution-oriented? Nat. Hum. Behav. *1*, 0015. https://doi.org/10.1038/s41562-016-0015.
196. Grêt-Regamey, A., Sirén, E., Brunner, S.H., and Weibel, B. (2017). Review of decision support tools to operationalize the ecosystem services concept. Ecosystem Serv. *26*, 306–315.
197. Daw, T.M., Coulthard, S., Cheung, W.W.L., Brown, K., Abunge, C., Galafassi, D., Peterson, G.D., McClanahan, T.R., Omukoto, J.O., and Munyi, L. (2015). Evaluating taboo trade-offs in ecosystems services and human well-being. Proc. Natl. Acad. Sci. *112*, 6949–6954.
198. Uusitalo, L., Lehikoinen, A., Helle, I., and Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. Environ. Model. Softw. *63*, 24–31.
199. Estévez, R.A., and Gelcich, S. (2015). Participative multi-criteria decision analysis in marine management and conservation: research progress and the challenge of integrating value judgments and uncertainty. Mar. Pol. *61*, 1–7.
200. Turner, P.A., Field, C.B., Lobell, D.B., Sanchez, D.L., and Mach, K.J. (2018). Unprecedented rates of land-use transformation in modelled climate change mitigation pathways. Nat. Sustainability *1* (5). https://doi.org/10.1038/s41893-018-0063-7.
201. Larrosa, C., Carrasco, L.R., and Milner-Gulland, E.J. (2016). Unintended feedbacks: challenges and opportunities for improving conservation effectiveness. Conservation Lett. *9*, 316–326.
202. Arneth, A., Olsson, L., Annette, A., Erb, K.-H., Hurlbert, M., Kurz, W.A., Mirzabaev, A., and Rounsevell, M.D.A. (2021). Restoring degraded lands. Annu. Rev. Environ. Resour. *46*. https://doi.org/10.1146/annurev-environ-012320-054809.
203. Lewandowsky, S., Risbey, J.S., Smithson, M., Newell, B.R., and Hunter, J. (2014). Scientific uncertainty and climate change: Part I. Uncertainty and unabated emissions. Climatic Change *124*, 21–37.
204. Howe, L.C., Macinnis, B., Krosnick, J.A., Markowitz, E.M., and Socolow, R. (2019). Acknowledging uncertainty impacts public acceptance of climate scientists' predictions. Nat. Clim. Change *9*, 863–867. 10.1038/s41558-019-0587-5.

Appendix B

Automatizing the analysis of computer simulations in R: some limitations and recommendations

Criscely Luján[1,2]*, Ricardo Oliveros-Ramos[3], Wencheng Lau-Medrano[4], Nicolas Barrier[5], Laure Velez[1], Philippe Verley[6], Yunne-Jai Shin[1]

[1] MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Montpellier. France

[2] Laboratoire d'Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91400, Orsay. France

[3] Inter-American Tropical Tuna Commission 8901 La Jolla Shores Dr, La Jolla, CA 92037, USA. United States of America

[4] Instituto del Mar del Perú. Esquina Gamarra y General Valle s/n, Chucuito, Callao. Perú

[5] MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Sète. France

[6] IRD, CIRAD, UMR AMAP. Boulevard de la Lironde, 34398, Montpellier Cedex 5. France

* criscelylujan@gmail.com (CL)

Abstract

Computer simulations are programs that numerically solve a model of a system or a process. Simulation tools with diverse levels of complexity and sophistication have been widely used and developed by computational biologists. With the increasing number of R users, there are many packages already implemented to analyze computer simulations in biology, from running the model to the analysis of model outputs. However, the way these tools are built can sometimes limit their use. Based on a review of R packages found in CRAN we identify four main limitations in packages developed for the analysis of computer simulations: i) the inclusion of too many functions, with ii) non-standard R naming conventions, resulting in iii) non-customizable plots of simulation outputs, which are iv) produced as side effects of long

calculations. The aim of this paper is to propose how to make the best use of the S3 object-oriented approach in order to avoid these issues. We present an illustrative example by using the marine ecosystem model OSMOSE and its associated R package **osmose**. This analysis provides some guidance for R programmers, software developers and scientists in general aiming to improve future development of packages in computer science.

Author summary:

A growing number of computer simulation tools are being developed to help a wide user community to solve, run and analyse models in different scientific disciplines. We undertook a review of such tools implemented in the R language which allows to show their main limitations from a user perspective. We then suggest ways to improve the performance of these simulation tools in the R language. We specifically illustrate some ready-to-use solutions based on the S3 Object-oriented programming system, and using an existing R package of a marine ecosystem model as a case study. Our analysis provides some guidance for R programmers, software developers and scientists in general aiming to improve future development of R packages in computer science.

1. Introduction

Computer simulations are defined as programs that contain a model of a system (natural or artificial) or a system process (De Jong and Van Joolingen, 1998). The development of computer simulations has been rising for the last fifty years in many fields of science (Winsberg, 2010). Currently their applications are spread in a wide range of science subjects such as physics, chemistry, biology, engineering, computer science, among others (Rutten et al., 2012).

R (Ihaka and Gentleman, 1996) is a free software and open source project that runs on almost any standard computing platforms and operating systems. R was developed with the aim to

perform statistical computing and one of its key advantages over many other softwares is its sophisticated graphics capabilities (Peng, 2016). Nowadays R shows substantial increases in popularity as it ranks 13th in the TIOBE index (https://www.tiobe.com/tiobe-index/, updated at May 2021), a measure of popularity of programming languages. The increasing demand of computer simulations together with the huge number of R users within the scientific community has pushed the development of R packages dedicated to the use of models. The Comprehensive R Archive Network (CRAN: https://cran.r-project.org/) is the official package repository that hosts thousands of tools including those which are dedicated to the use of models. This repository allows access to the code and documentation of packages.

With the objective to make a state of the art of the tools dedicated to computer simulations, we performed a systematic review of the packages hosted in the CRAN using four keywords ("modelling", "modeling", "computer", and "simul"; see **Appendix A** for detailed methodology). The search resulted in 292 packages related to computer simulations currently available in CRAN. We consider this is a lower estimate of the number of R packages as some packages known by the authors were not included in this search due to the mismatch of keywords (e.g. `CHNOSZ` (Dick, 2019), `photobiology` (Aphalo, 2015), `phytools` (Revell, 2012), `r4ss` (Taylor et al., 2019) and `soilphysics` (da Silva and de Lima, 2015)) or because they are not available in CRAN yet and are hosted on other platforms (e.g. GitHub, GitLab).

A more detailed review of a sample of 73 packages (see **Table A.1** in **Appendix A**) which were published between 2015 and 2020 showed the use of computer simulation packages across 15 science fields including: ecology, geophysics, seismology, environmental sciences, social sciences, hydrological sciences, biology, among others. These tools are not only diverse in terms of thematics of application but also in terms of complexity and code extension. The use of these tools allows the execution of models, but also the analysis and visualization of model inputs and outputs.

The development of R packages as interfaces between the core programs of a model and the handling of input/output intends to ease the work with computer simulations, especially for users with more expertise in the theoretical aspects behind the model than in the computing and programming fields. However, the development of an R package involves a set of unique challenges, and as a lesson from the packages' review we found that standard programming requirements are not always well satisfied.

In this article we focus on four issues that we identified as principal limitations in the development and use of packages dedicated to computer simulations, providing some recommendations that are consistent with R standards:

- **Too many functions**: R packages that are developed for performing simulation analyses often include a large amount of functions. Some packages that are part of our review, e.g., `biomod2` (Thuiller et al., 2020), `BacArena` (Bauer et al., 2017) and `EpiModel` (Jenness et al., 2018), are characterized by the export of more than 100 functions in total, while `surveillance` (Meyer et al., 2017) has more than 400 functions. Another package that was not part of our search in CRAN is `photobiology` (Aphalo, 2015) which exports more than 900 functions. The inclusion of too many functions in the use of a package is a limitation for the users who are bound to memorize many function names, in addition to understanding how the different functions work and are linked, and to finally learn how to interpret the function outputs. Therefore, the inclusion of generic functions through all package operations can be useful. Generic functions - for example a function called `summary` - are commonly and intuitively interpreted by R users as a function providing the summary of a specific R object.
- **Non-standardized names**: The selection of appropriate names (for functions and function arguments) is fundamental in the development of an R package. Correctly naming functions helps users to understand how a package works. Moreover, the same

arguments' names can also be used in different functions within the package, reducing the number of names to be learnt by the users. Potential naming problems arise when packages are developed without following any type of naming conventions (Bååth, 2012), or in other cases mixing them. And even when a package follows a naming convention, sometimes the chosen name does not provide any information about the use of the function or the nature of the argument. Within our package review set, only 39.7% of the packages present standardized function names, indicating potential issues in the way the code is developed for these tools.

- **Non-customizable plots**: Packages working with models' inputs and outputs usually include visualization and plotting functions. However the importance is not only to provide code for graphs, but also to provide some options to customize graphs. Only this way a user can quickly analyze and understand the outcome of model simulations, considering that the needs and the objectives of the simulations will differ from one user to another. Providing basic graphs can be a limitation for users who need more elaborate graphs but do not have enough programming skills or time to code the graphs themselves. Providing non customizable graphs is a limitation for users in general, even for those having high experience and programming skills, since the plot functions included in the package do not allow changes on the plots appearance to get more sophisticated and customized plots (e.g. changes in legends, text size, colors). In our package review, 72.6% of the tools provided a function to make graphs, but only 18.9% of them allowed the customization of graphs. The objective of a package should be focused in providing users well-elaborated graphs that are as flexible as possible to offer users a large range of options.
- **Plots as side effects of complex calculations**: Computer simulations involve many calculations that are usually included in the packages along a three tiers structure: i)

loading of the input data, ii) performing calculations and, iii) presenting model outputs (usually by means of graphs). Sometimes all three stages are performed by a single function. Nevertheless, this all-in-one approach can limit the users in their understanding of the steps in-between the different processes, and can affect time optimization as well. It is therefore highly recommended to work with independent functions, each of them being aligned with some specific process (reading data, running intermediate analyses and displaying results). That way the users can improve their understanding as well as integrate additional levels of analyses, such as merging data, including exploratory analyses in-between and also customizing plots.

In the present work we propose ways to deal with the four limitations described above when developing R packages. We specifically provide some guidance on how to make the best use of the S3 object-oriented programming (OOP) system to analyze computer simulations. We illustrate our purpose using the marine ecosystem model OSMOSE (Shin and Cury, 2004) and its associated R package **`osmose`** (Shin et al., 2020). Through this example, we aim to help the future development of packages and tools across different scientific fields, and for a variety of programmers, software developers and scientists developing R packages to perform computer simulations.

2. The S3 object-oriented programming system

Object-oriented programming (OOP) is a programming paradigm. In an OOP system, programmers define the data type of a data structure, and also the types of operations that can be applied to the data structure. In this way, the data structure becomes an *object* , and a large software can be considered as a society of objects (Kak, 2003). Then, all objects that possess the same attributes and exhibit the same behavior are grouped into collections, called *classes*

(Garrido, 2003). The operations between objects are determined by functions known as *methods* (Craig, 2007).

2.1 OOP systems in R

R has four OOP implementations: S3, S4, RC (reference classes) and R6. The "base types" system is not object-oriented, however it provides the building blocks for the OOP systems in R (Wickham, 2014). The differences between these OOP systems depend on how classes and methods are defined. Most OOP programming languages (like Java, C++ and C#) implement a style of OOP called *message passing object-oriented*. This style is based on sending messages between objects, and based on these messages the objects determine which method to call. S3 works differently and implements a style called *generic function object-oriented*. Instead of passing messages, it uses a function called generic function that decides which method to call (Wickham, 2014). S3 is the simplest OOP system in R, the most commonly used system in CRAN packages and the only OOP system used in the base and stats packages (Wickham, 2014).

2.2 Visualization and analysis using S3 in R

Working with the S3 system in R, if we create an atomic vector using the `sample` function, this object will be of the class `integer`. And then if we encode that object as a factor using the `as.factor` function, this object will be of the class called `factor`:

```
# Object of 'integer' class
x0 = sample(x = c(1:10), size = 20, replace = TRUE)
class(x0)
#> [1] "integer"
# Object of 'factor' class
```

```
x1 = as.factor(x0)
class(x1)
#> [1] "factor"
```

After the creation of the objects x0 and x1 we can use generic functions that already exist in R, and produce different results according to the class of the objects:

```
plot(x0)
plot(x1)
```

Although they were created using the same generic function (`plot`), the plots produced for the two objects are different (**Figure 1**). The plot of x0 shows a scatter plot while the plot of x1 a bar plot. This is due to the use of the S3 OOP system. Using this system, the generic function `plot` is calling different functions (methods) to create a plot depending on the class of the objects. Then for the object of integer class (x0) the method used is the `plot.default`. While for the object of factor class the method used is the `plot.factor`. However, users should never call these specific methods directly, but instead rely on the generic function (`plot`) to call the corresponding methods.

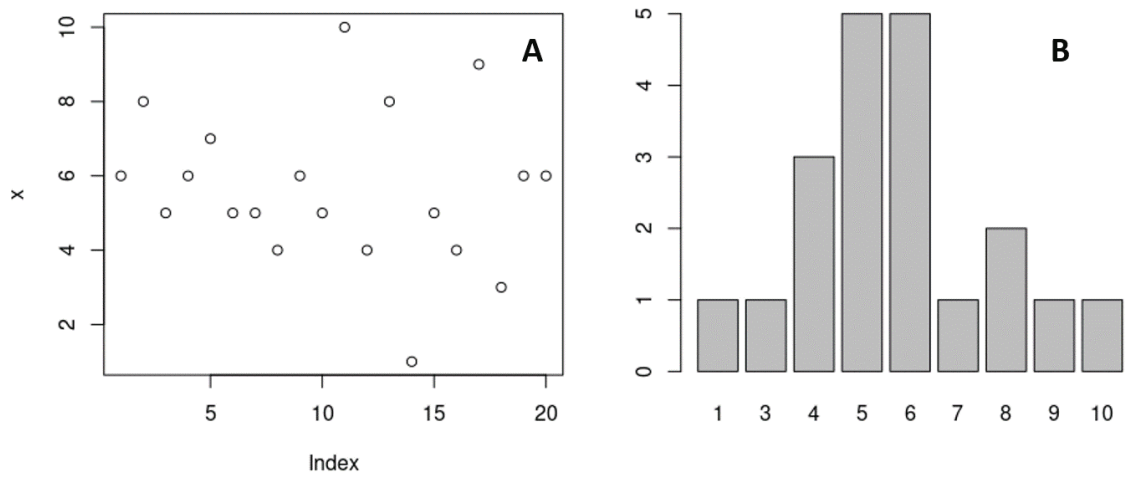

Figure 1: The use of `plot` generic function for different objects. The graph A (on the left) is a scatter plot of the object x0 (`plot(x0)`) of the class integer. The graph B (on the right) is a bar plot of the object x1 (`plot(x1)`) of the class factor . Differences in the plots' appearance are due to the class of the objects. The generic function plot calls the `plot.default` function when the object is of the class integer and the `plot.factor` function when the object is of the class factor.

The principal use of S3 OOP in R is for the `print`, `summary`, and `plot` methods (Wickham, 2019). These methods allow one generic function to display the objects differently depending on their class. To check the list of all the available methods for a S3 generic function use the `methods` function (for example: `methods(plot)`).

3. Illustrating the S3 approach using the `osmose` R package

3.1 The `osmose` R package

osmose is the R package created for the OSMOSE (Object-oriented Simulator of Marine ecOSystEms) model (Shin and Cury, 2001, 2004). OSMOSE is a spatial, age and size-structured individual-based model (IBM) of the dynamics of marine fish and macro-invertebrate communities. OSMOSE represents the major processes of the fish life cycle (growth, predation, natural and starvation mortalities, reproduction and migration) to simulate the functioning of marine food webs under climate change and fisheries pressure.

The programming language used for coding the core processes in OSMOSE is Java (Gosling, 2000) but the model can be run within the R environment with the **osmose** R package. The implementation of this package in R makes the use of the model easier for the many scientists who are very familiar with the R programming language.

Information about the OSMOSE model can be found in its official webpage (http://www.osmose-model.org) and also in its user guide (https://documentation.osmose-model.org/). Some OSMOSE applications can be found in the scientific literature, e.g., Travers et al. (2006); Marzloff et al. (2009); Travers et al. (2010); Travers-Trolet et al. (2014); Grüss et al. (2015); Halouani et al. (2016); Oliveros-Ramos et al. (2017); Fu et al. (2017); Moullec et al. (2019).

osmose is available at CRAN (https://cran.r-project.org/web/packages/osmose/index.html) and provides tools to build and run simulations using the OSMOSE model. It also allows the processing of model configuration and outputs producing graphics and reports. The package code and its new implementations can also be found in its GitHub repository (https://github.com/osmose-model/osmose).

As part of the OSMOSE model configuration, a large amount of data is needed to run the model but also to analyze the model outputs. There are two important functions in the package: `run_osmose` and `read_osmose`. The first function is focused on running the model, and the

second is focused on reading the model configuration and the model outputs. A schematic representation of these functions is presented in **Figure 2**.

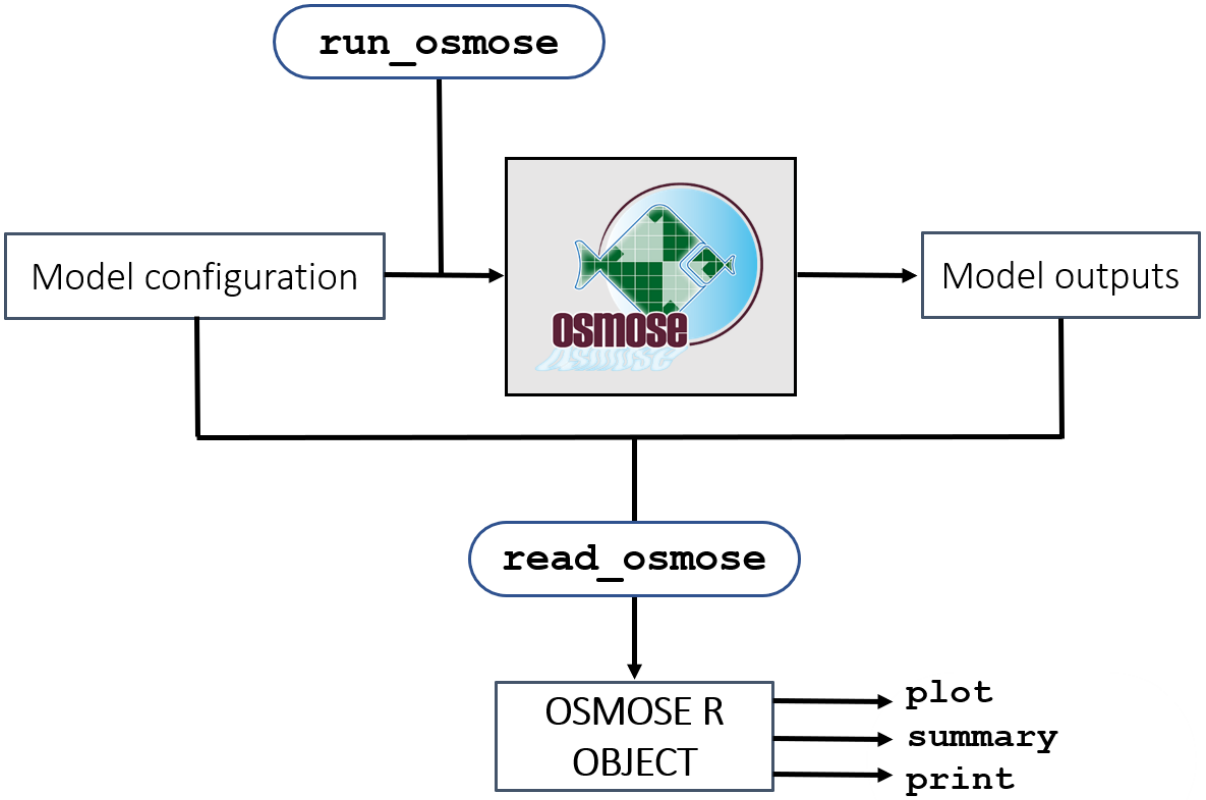

Figure 2: Diagram representing the functioning of the **osmose** package. To use this package we need the input information that constitutes the configuration of the model (parameterization, forcing variables). This information can be read by the use of the `read_osmose` function. This function can also be called to read all the outputs after running the model by the use of the `run_osmose` function. The object we get can be used by generic functions such as `plot`, `summary` or `print`.

3.2 Using the `osmose` R package

The use of S3 OOP in this package allows an easy manipulation of objects by using classes, methods and generic functions. A summary of principal classes and methods in **osmose** is presented in **Table 1**.

Using **osmose** we explain how we dealt with the four limitations previously described. The code used in the following example is presented in the **Appendix B** section, including the creation of the two objects: `outputs` and `configuration`. Objects of the classes `osmose` and `osmose.config` respectively. Using these two objects and a generic function, like `plot`, we can produce graphs of the model configuration and model outputs. See **Figure 3**, **Figure 4** and the following code example:

```
# Figure 3: Plot for model outputs
plot(outputs, what = "biomass")

# Figure 4: Plot for model inputs (model configuration)
plot(configuration, what = "species")
```

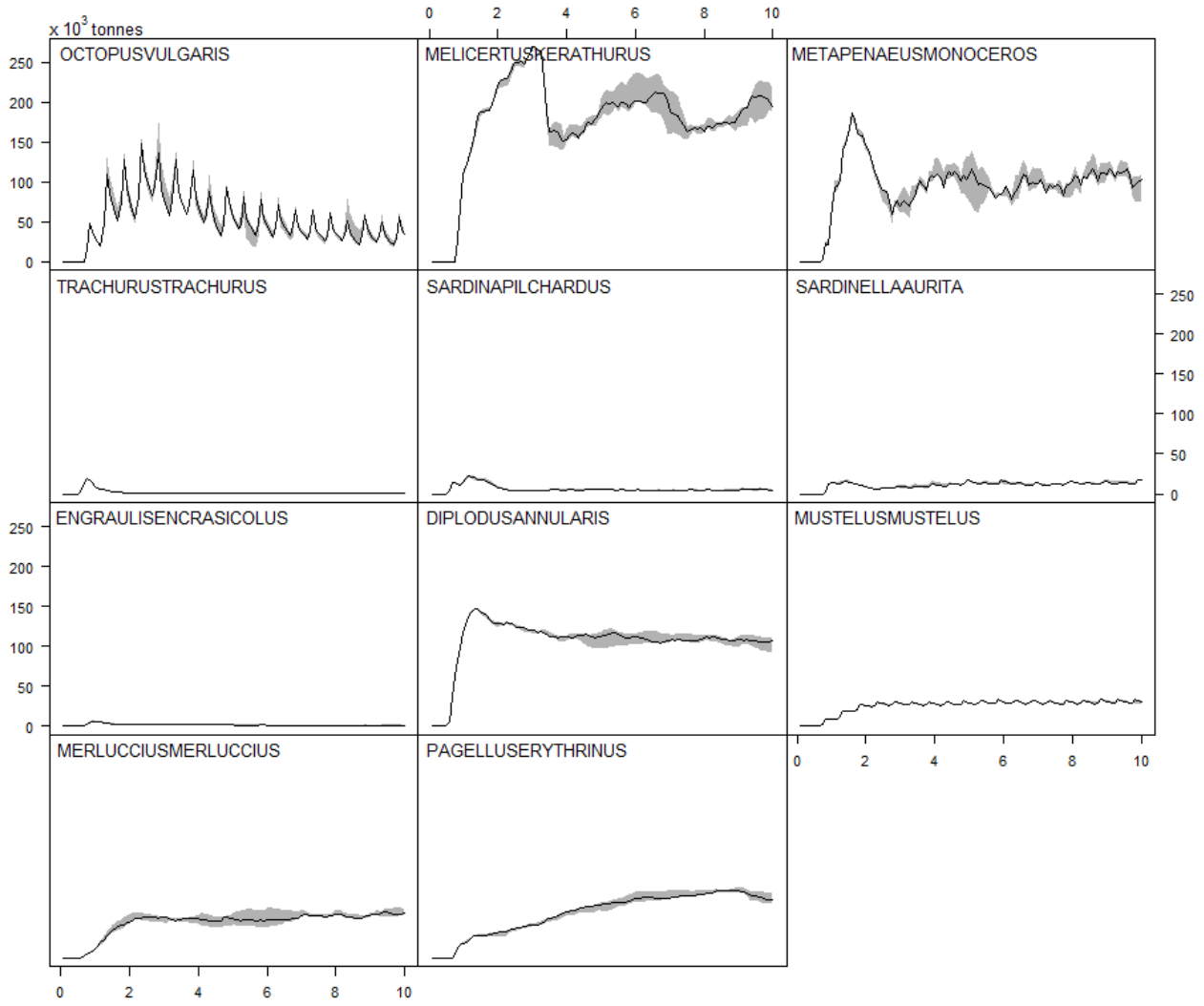

Figure 3: Plot for model outputs. Using an object called `outputs` (created in the section **Appendix B**) and the generic function `plot` we created a graph for the simulated biomass of the species included in the model. We need to specify the variable that is plotted by the use of the argument `what`.

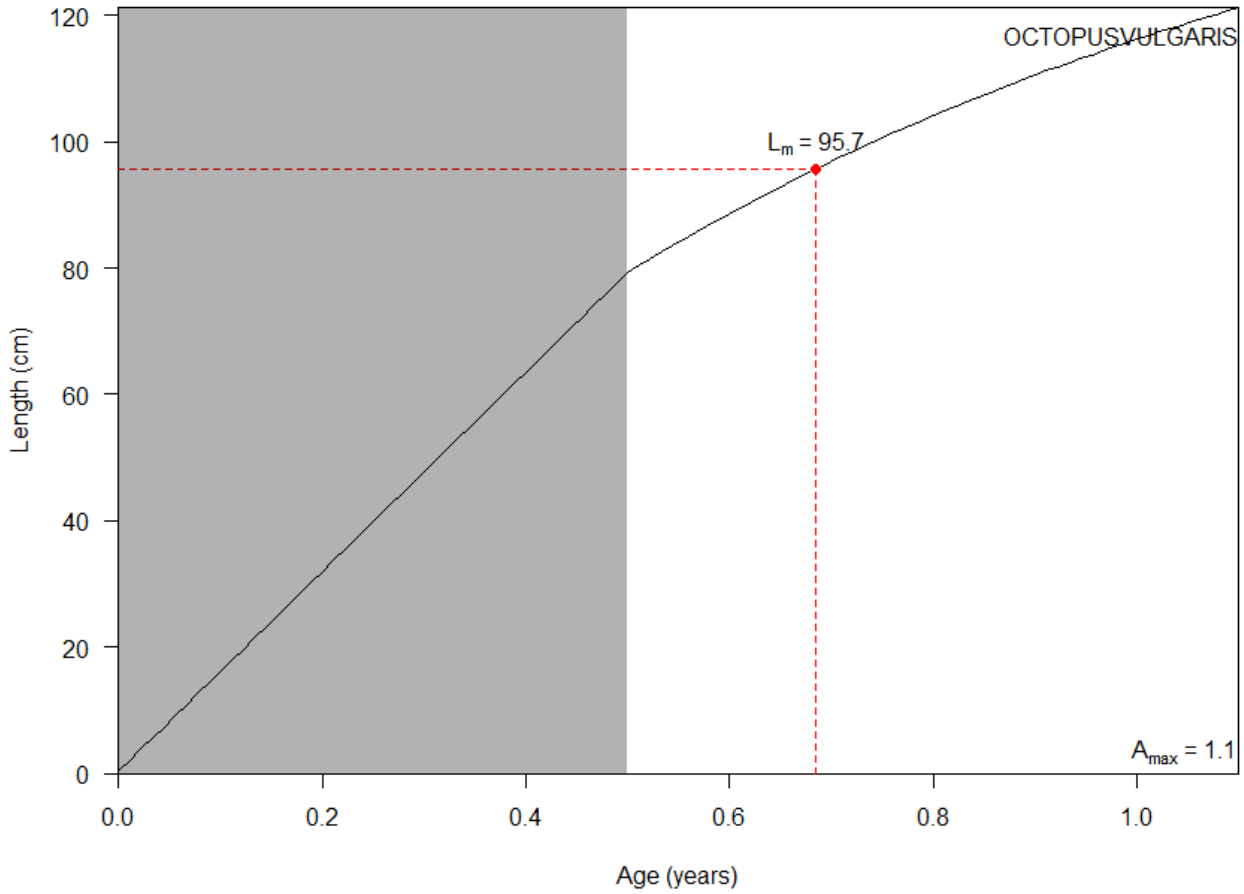

Figure 4: Plot for model inputs. Using an object called `configuration` (created in the section **Appendix B**) and the generic function `plot` we created a graph of the growth parameters of the first species included in the model. We need to specify the variable that is plotted by the use of the argument `what`.

This example shows that with the use of a reduced number of functions and a few lines of code, users can generate graphs in an easy way. Here we are only using one generic function (`plot`) and based on the class of the objects the corresponding method is called without user intervention (e.g. `plot.osmose`). The generic function only needs to be provided with an object of the `osmose`'s classes and the variable name to be plotted by the use of the argument

"what". This argument specifies the process (as part of the model configuration) or the variable (as part of the model output) to be plotted.

osmose also includes plot types for OSMOSE R objects. The selection of the plot types is done through the use of the argument called "type" (as in the plot generic function of the graphics R package). If this parameter is not specified in argument, it always uses the default value (type = 1). For more details about all plot types in **osmose**, see its internal code in GitHub (https://github.com/osmose-model/osmose).

Graphical differences between default plots and customized plots are presented in the next example of code. We used an object of class osmose to make a graph using the variable called biomass (what = "biomass") and the plot type number 2 (type = 2). The default plot (**Figure 5-A**) can be produced based on three arguments (x, what and type). These arguments can be used through all plot methods in **osmose** allowing the standardization in the arguments names of functions.

The plot can be customized using additional arguments such as: col, lty, lwd, cex, cex.axis and ylim (**Figure 5-B**). These arguments' names are the same as in the plot generic function in R, and therefore comply with standardization requirements. Moreover, the customization of plots is not only about the arguments that change the plot appearance, but also about the arguments that change the information presented in the plot. An example of that feature is in the use of the arguments called: species, speciesNames, and initialYear. The first argument is about the selection of components to plot (indicating which species will appear on the graph), while the second one specifies which names will be assigned to each species (e.g, using scientific names). Since this is a graph of time series, this plot also allows us to manipulate the year dimension in the x-axis by using the argument called initialYear. These examples show that users can produce well-elaborated graphs in an easy and intuitive

way but the code of the package is also flexible enough to allow users to make changes in the appearance and content of the graphs.

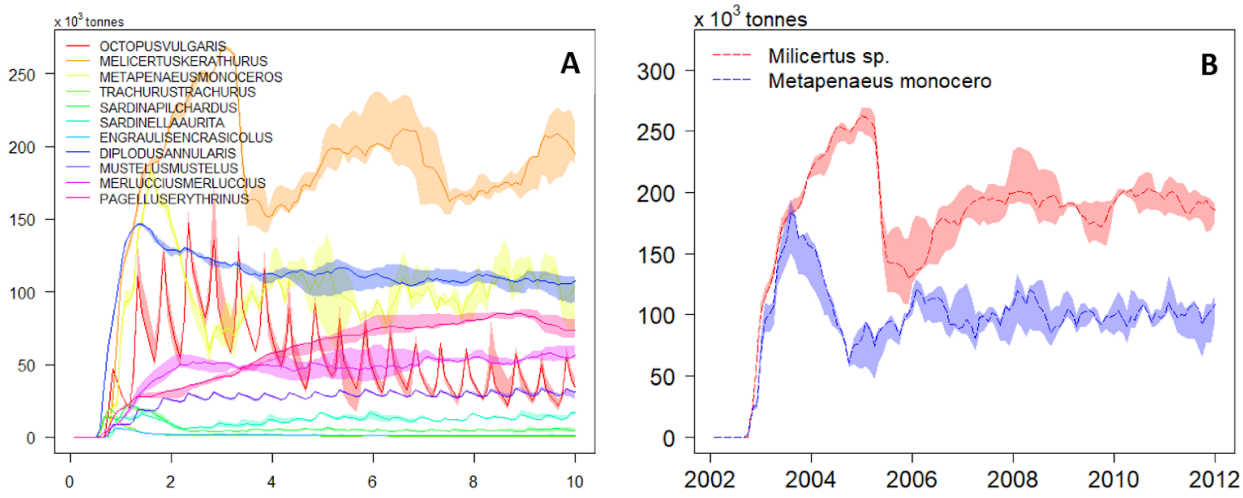

Figure 5: Customization of plots. Calling the same object of class `osmose` two plots were created. The graph **A** (on the left) is a default plot of the variable called biomass using the plot type number 2 (`type = 2`). The graph **B** (on the right) is a customized plot using the same variable and plot type (`biomass` and `type = 2`) but including changes about the appearance of the plot (`col`, `lty`, `lwd`, `cex`, `cex.axis`, and `ylim`) as well as the content of the plot: selection of species (`species = c(1, 2)`), species names (`speciesNames = c("Milicertus sp.", "Metapenaeus monocero")`), and the initial year in the x-axis (`initialYear = 2002`).

```
# Default plot
plot(outputs, what = "biomass", type = 2)

# Customized plot
```

```
plot(outputs, what = "biomass", type = 2, col = c("red", "blue"), lty =
5,lwd = 1.5, cex = 1.3, cex.axis = 1.5, ylim = c(0, 320), species = c(1,
2), speciesNames = c("Milicertus sp.", "Metapenaeus monocero"),
initialYear = 2002)
```

The above code example shows that subsequent to the simulation run, the creation of the "outputs" object allows re-use of the same object to create different plots and analyses, without needing to read and run the model again. This is due to the structure of this package which works in three tiers: loading the inputs, performing calculations and plotting the results. Everything needed for plotting (in this example, the model outputs) are assigned to an object ("outputs") that is used to display the results and possibly run statistical analyses. At the level of the creation of plots, graphical changes can be done using the same object. In case users require very specific analyses, more complex analyses or plots that are not included in the package yet, users can extract the variable of interest by using the supplied `get_var` function to then make additional calculations and produce graphs outside the package.

Table 1: Examples of classes and methods found in **osmose**

Class	Method	Description
`osmose`	`print, summary, plot, report, get_var`	Main `osmose` class produced with `read_osmose` function when model outputs are read
`osmose.config`	`plot, get_var`	Class produced with `read_osmose` function when model configuration are read
`osmose.biomass`	`plot`	Exclusive class of `osmose` outputs and the variable biomass
`osmose.abundance`	`plot`	Exclusive class of `osmose` outputs and the variable abundance

`osmose.yield`	`plot`	Exclusive class of `osmose` outputs and the variable yield

3.3 Implementing the `osmose` R package

To overcome the main issues encountered when developing packages for computer simulations (see introduction), we adopted a range of decisions for the implementation of **`osmose`**:

Keeping the number of functions as low as possible: There are only two principal functions to start using **`osmose`**: `run_osmose` to run the computer simulation and `read_osmose` to read the model configurations and model outputs. As mentioned above, the use of generic functions and classes and the creation of methods (three important components of the S3 OOP system) allowed to reduce the number of functions used in the package. For methods, **`osmose`** uses the same generic function names (for example: `plot` and `get_var`) across different classes such as `osmose` and `osmose.config`. This way the users only need to recall that there are two functions (called `plot` and `get_var`) which allow them to make graphs and obtain variables, respectively. However, in terms of code (internally in the package), each class uses different methods (functions) which results will depend on the handled object: i) `plot.osmose.config` and `get_var.osmose.config` for objects of the class `osmose.config`, and `plot.osmose` and `get_var.osmose` for objects of the class `osmose`.

Five methods were developed for the `osmose` class. Three of them use common generic functions in R: `print`, `summary`, and `plot`. For the other two methods, generic functions were created: `report` and `get_var`. These methods allow to:

- `print.osmose`: print information about the outputs of an `osmose` object.
- `summary.osmose`: summarize information about the outputs of an `osmose` object.

- `plot.osmose`: produce different plots from the `osmose` object.
- `report.osmose`: export a report of an `osmose` object.
- `get_var.osmose`: extract a variable that is part of the `osmose` outputs.

In the `osmose.config class`, two methods were created:

- `plot.osmose.config`: produce different plots for an `osmose.config` object.
- `get_var.osmose.config`: extract a variable that is part of the `osmose` configuration.

Standardizing names: The two main **osmose**'s functions, `run_osmose` and `read_osmose`, follow the underscore_separated standard naming convention (Bååth (2012)). Depending on the use of `read_osmose` the class names can be: i) `osmose` class when reading the model outputs, or ii) `osmose.config` class when reading the model configuration. Both types of objects are related to the OSMOSE model, therefore we kept the name "osmose" in both class names. However, we separate the second class name by using "." for the names of the classes (for example: `osmose.config`). This avoids confusion with the functions names that use the underscore_separated naming convention.

Concerning generic functions, the naming was carefully done along the standards adopted for the names of common generic functions in R. For example, the function `plot` is a generic function commonly used for plotting R objects. We chose the same function name for both classes of **osmose**. Other generic functions developed for `osmose` class are also common generic functions in R (e.g., `print` and `summary`) while functions such as `get_var` and `report` are specific to **osmose** and were assigned descriptive names that are the same in both classes of **osmose**. Both `get_var` and `report` are important generic functions for **osmose** users, they allow to extract a variable to perform additional analyses and to generate reports, respectively.

The importance of standardizing names is also taken into account in the naming of functions' arguments. For example, as arguments of the generic function `plot` and the methods `plot.osmose` and `plot.osmose.config`, we reuse the arguments' names of the `plot.default` function in R, such as: `lty`, `lwd`, `cex`, `cex.axis`, `col`, among others. In this way, users who are already familiarized with the use of the `plot.default` function in R, can readily use the same arguments in the plot methods of **osmose**. We also created other arguments for plot methods such as `what`, `type`, `species` and `initialYear`, which are all the same across the methods.

Customizing plots on demand: Allowing users to customize plots was a real challenge in the development of **osmose** due to the high number of variables with different dimensions in the model configuration and model outputs. However that was managed by developing plot types, maintaining flexibility in the use of plot arguments as much as possible. For example, choosing an object of class `osmose` the use of the argument type looks as simple as the following code:

```
# Using the 'outputs' object created in the section Appendix B
# Plot of biomass
plot(outputs, what = "biomass", type = 1)
```

What is internally happening in the package is that due to the class of the object (`osmose`), the generic function (`plot`) calls the method `plot.osmose`. Then this function uses the generic function (`get_var`) and the corresponding method (`get_var.osmose`) to obtain an object of the class `osmose.biomass`:

```
plot.osmose = function(x, what = "biomass", ...) {
x = get_var(x, what = what, ...) # using here: 'get_var.osmose' method
```

```
# now the class of 'x' is osmose.biomass
plot(x, ...) # using here: 'plot.osmose.biomass' method
return(invisible())
}
```

Due to the class of the "x" object (`osmose.biomass`), the plot method called is `plot.osmose.biomass`. The structure of this method is the following:

```
plot.osmose.biomass = function(x, type = 1, ...) {
osmosePlots2D(x, type = type, ...)
return(invisible())
}
```

`osmosePlots2D` is an internal function in **osmose** that was created for variables that have 2 dimensions. In this example, the variable (`what = biomass`) has information by species and time. Then the way the biomass variable is plotted is precisely determined by the type argument passed to the function `osmosePlots2D`. By default it is using the type number 1 (`type = 1`).

Internal plot functions such as `osmosePlots2D` process information according to the dimensions of the variables to be plotted, regardless of the content of the information. The function `osmosePlots2D` can thus be reused for other variables that have 2 dimensions like for example the OSMOSE variables called "abundance" and "yield". Using the same idea we created other functions called `osmosePlots3D` and `osmosePlots4D`, for other variables that are in three or four dimensions, respectively.

One important argument of the internal plot functions is the `type`. This argument determines how the variable is plotted. Then, many types of plots can be included in each of the internal plot functions `osmosePlots2D`, `osmosePlots3D`, `osmosePlots4D` (**Figure 6**).

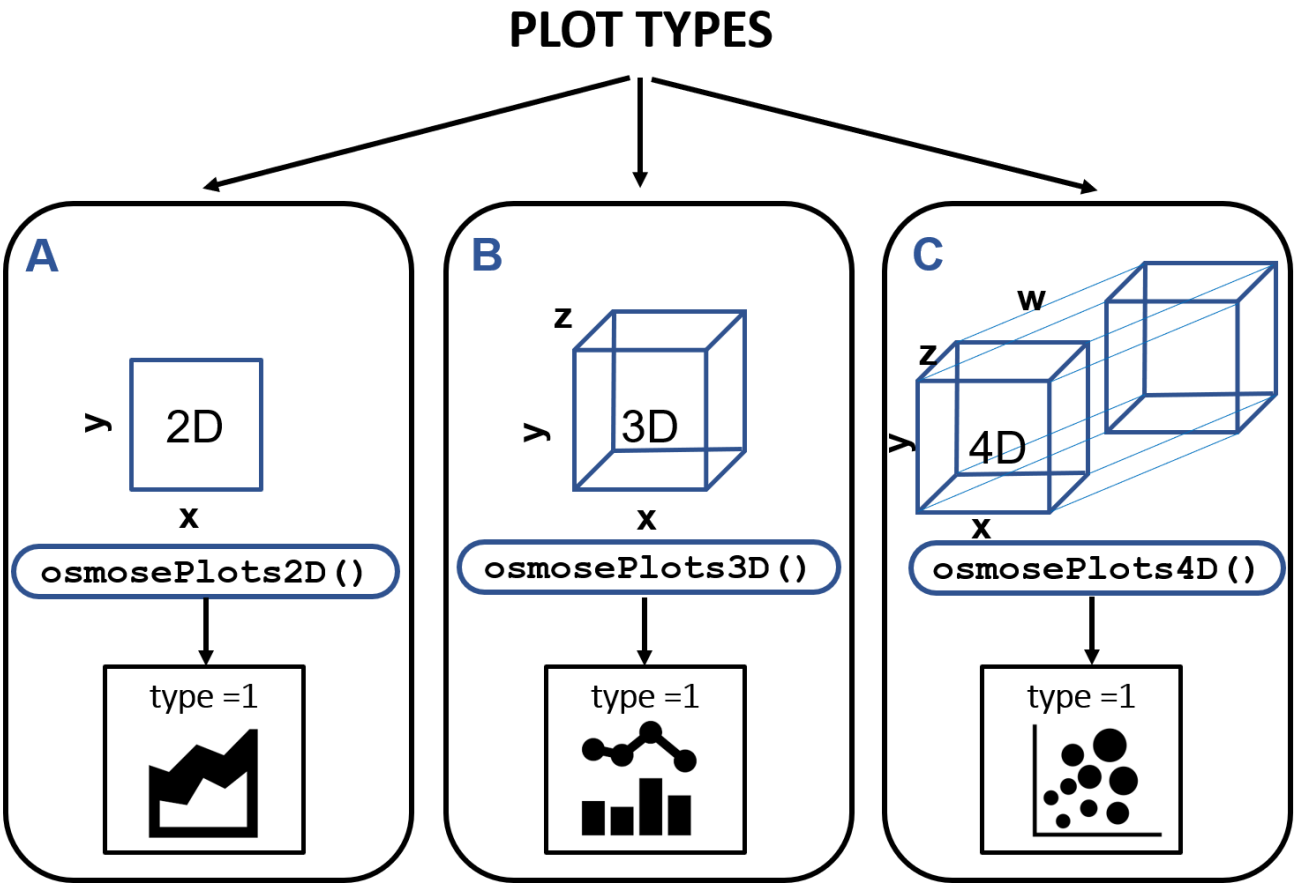

Figure 6: Schematic representation of the use of internal functions (`osmosePlots2D`, `osmosePlots3D`, `osmosePlots4D`) for information processing. In this example, we want to plot a variable of an OSMOSE object (configuration or outputs) containing information in several dimensions that need to be processed before plotting. The internal plot functions process the information, using `osmosePlots2D` for variables that have 2 dimensions, `osmosePlots3D` for 3 dimensions, and `osmosePlots4D` for 4 dimensions. The different ways to plot the variables are included in these internal plot functions by the use of the

argument called "type" (`type = 1`). In this representation we included only one plot type for each internal plot function.

The way the code of the plot types is structured allows to create as many plot types as possible for each internal plot function. Therefore, future plot types that can be created for OSMOSE variables (as part of model configuration or model outputs) will not disrupt the functions that were already developed but will expand the code of the package by creating more types. On the other hand, the code can be extended to plot other OSMOSE variables. This can be done including more functions with the same structure than `plot.osmose.biomass`, for example by changing the variable "`biomass`" for "`mortality`".

Dissociating the simulation and plots sequences: The way we developed **`osmose`**, having functions organized in three steps, avoided the creation of plots as side effects of complex calculations. The first step includes functions to run the model (`run_osmose`). Then after running the model, if nothing is changed about the configuration, we do not need to go through this step again. This first step also involves reading the model configuration and model outputs (`read_osmose`). The second step includes functions that perform intermediate analyses, such as extracting and transforming a variable that is part of the outputs (e.g. biomass). Finally, the third step consists in plotting the results of the model runs. This step reuses the work done before and the object created from running the model.

4. Conclusions and perspectives

There are many R packages already available to facilitate the use of computer simulations. The review of the CRAN packages we performed showed that the way these tools are implemented can sometimes limit their use. We identified four main limitations: i) inclusion of too many

functions, with ii) non-standard R naming conventions, to produce iii) non-customizable plots, which are iv) produced as side effects of long calculations. Our review also revealed that only 47.9% of the analyzed packages dedicated to computer simulations used the OOP system, and among them 82.9% adopted the S3 OOP and 34.3% the S4 OOP (so 17.2% using a mix of both S3 and S4 OOP). This is consistent with the literature (Wickham, 2014) which reports that S3 is one of the most common OOP systems in CRAN packages. However, the limitations raised in this study are regularly encountered, even if an OOP system was used, whether S3 or S4 OOP.

The aim of this study was to explain how to make the best use of the S3 OOP system, describing the benefits of using this system for running and analyzing computer simulations. We selected this particular S3 system due to its popularity in R. However the coding decisions adopted for **osmose** and that we explained through this study are applicable to any other package that uses any other OOP system. In the case of the **osmose** package, S3 enables users to handle complex OSMOSE model applications in an intuitive way and with a fast learning curve. From the perspective of a developer, the use of the S3 OOP system allows us to work with a tool that can evolve easily. The **osmose** code can be used to create new structures in terms of analyses (new objects, classes and methods). Analyses such as the calibration, the validation or the sensitivity analysis of the OSMOSE model are essential, and more generally for any type of applied model. These analyses can be easily included in the development of **osmose** by for example the creation of new classes and their corresponding generic functions and methods.

The code to produce graphics is an important component in **osmose**. Even though many plot types are already implemented in this package, the core of the OSMOSE model continues to evolve which requires the creation of new graphics and new types of analyses. This is all facilitated by an appropriate use of the S3 approach. It also allows users to call specific

packages to customize the graphs according to their needs and preferences (e.g. by using `Tidyverse` or `lattice` packages).

The reflection of this work goes beyond the **`osmose`** case study, opening the discussion on how existing techniques (like the use of OOP) can provide important benefits in the development of tools that are highly used in science. We focused on S3 but similar solutions can also be implemented by using other OOP systems in R and even in another programming language.

5. Acknowledgement

Criscely Luján was supported by an Peruvian individual doctoral research grant from the "Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica" (contract n° 134-2016-FONDECYT). Yunne-Jai Shin, Nicolas Barrier, Ricardo Oliveros-Ramos and Criscely Luján have been partially funded by the Biodiversa and Belmont Forum project SOMBEE (BiodivScen programme, ANR contract n°ANR-18-EBI4-0003-01), and the European Union's Horizon 2020 research and innovation programme under grant agreement N◦ 817578 (TRIATLAS project).

6. References

1. P. J. Aphalo. The r4photobiology suite. UV4Plants Bulletin, 2015(1):21–29, 2015. URL https://doi.org/10.19232/uv4pb.2015.1.14.

2. R. Bååth. The State of Naming Conventions in R. The R Journal, 4(2):74–75, 2012. URL https://doi.org/10.32614/RJ-2012-018.

3. E. Bauer, J. Zimmermann, F. Baldini, I. Thiele, and C. Kaleta. Bacarena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. PLoS Computational Biology, 13(5): e1005544, 2017. URL https://doi.org/10.1371/journal.pcbi.1005544.

4. I. D. Craig. Object-oriented programming languages: Interpretation. 8, 2007. URL https://doi.org/10.1007/978-1-84628-774-9.

5. A. R. da Silva and R. P. de Lima. Soilphysics: An r package to determine soil preconsolidation pressure. Computers & geosciences, 84:54–60, 2015. URL https://doi.org/10.1016/j.cageo.2015.08.008.

6. T. De Jong and W. R. Van Joolingen. Scientific discovery learning with computer simulations of conceptual domains. Review of educational research, 68(2):179–201, 1998. URL https://doi.org/10.3102/00346543068002179.

7. J. M. Dick. Chnosz: Thermodynamic calculations and diagrams for geochemistry. Frontiers in Earth Science, 7:180, 2019. URL https://doi.org/10.3389/feart.2019.00180.

8. C. Fu, N. Olsen, N. Taylor, A. Grüss, S. Batten, H. Liu, P. Verley, and Y.-J. Shin. Spatial and temporal dynamics of predator-prey species interactions off western canada. ICES Journal of Marine Science, 74(8):2107–2119, 2017. URL https://doi.org/10.1093/icesjms/fsx056.

9. J. M. Garrido. Object-oriented programming: from problem solving to Java. Charles River Media, 1 edition, 2003.

10. Gosling, J., 2000. The Java Language Specification. Addison-Wesley Professional, Boston, MA.

11. A. Grüss, M. J. Schirripa, D. Chagaris, M. Drexler, J. Simons, P. Verley, Y.-J. Shin, M. Karnauskas, R. Oliveros-Ramos, and C. H. Ainsworth. Evaluation of the trophic structure of the west florida shelf in the 2000s using the ecosystem model osmose. Journal of Marine Systems, 144:30–47, 2015. URL https://doi.org/10.1016/j.jmarsys.2014.11.004.

12. G. Halouani, F. B. R. Lasram, Y.-J. Shin, L. Velez, P. Verley, T. Hattab, R. Oliveros-Ramos, F. Diaz, F. Ménard, M. Baklouti, et al. Modelling food web structure using an end-to-end approach in the coastal ecosystem of the gulf of gabes (tunisia). Ecological modelling, 339:45–57, 2016. URL https://doi.org/10.1016/j.ecolmodel.2016.08.008.

13. R. Ihaka and R. Gentleman. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics, 5(3):299–314, 1996. URL https://doi.org/10.1080/10618600.1996.10474713.

14. S. M. Jenness, S. M. Goodreau, and M. Morris. EpiModel: An R package for mathematical modeling of infectious disease over networks. Journal of Statistical Software, 84(8):1–47, 2018. URL https://doi.org/10.18637/jss.v084.i08.

15. A. C. Kak. Programming with Objects: A Comparative Presentation of Object-Oriented Programming With C++ and Java. John Wiley and Sons, Inc, 2003. URL https://doi.org/10.1109/9780470547144.

16. M. Marzloff, Y.-J. Shin, J. Tam, M. Travers, and A. Bertrand. Trophic structure of the peruvian marine ecosystem in 2000–2006: insights on the effects of management scenarios for the hake fishery using the ibm trophic model osmose. Journal of Marine Systems, 75(1-2):290–304, 2009. URL https://doi.org/10.1016/j.jmarsys.2008.10.009.

17. S. Meyer, L. Held, and M. Höhle. Spatio-temporal analysis of epidemic phenomena using the r package surveillance. Journal of Statistical Software, 77(11), 2017. URL https://doi.org/10.18637/jss.v077.i11.

18. F. Moullec, L. Velez, P. Verley, N. Barrier, C. Ulses, P. Carbonara, A. Esteban, C. Follesa, M. Gristina, A. Jadaud, et al. Capturing the big picture of mediterranean marine biodiversity with an end-to-end model of climate and fishing impacts. Progress in Oceanography, 178:102179, 2019. URL https://doi.org/10.1016/j.pocean.2019.102179.

19. R. Oliveros-Ramos, P. Verley, V. Echevin, and Y.-J. Shin. A sequential approach to calibrate ecosystem models with multiple time series data. Progress in oceanography, 151:227–244, 2017. URL https://doi.org/10.1016/j.pocean.2017.01.002.

20. Peng, R. D. (2016). R programming for data science (pp. 86-181). Leanpub.

21. L. J. Revell. phytools: an r package for phylogenetic comparative biology (and other things). Methods in ecology and evolution, 3(2):217–223, 2012. URL https://doi.org/10.1111/j.2041-210X.2011.00169.x.

22. N. Rutten, W. R. Van Joolingen, and J. T. Van Der Veen. The learning effects of computer simulations in science education. Computers & Education, 58(1):136–153, 2012. URL https://doi.org/10.1016/j.compedu.2011.07.017.

564 23. Y.-J. Shin and P. Cury. Exploring fish community dynamics through size-dependent trophic
565 interactions using a spatialized individual-based model. Aquatic Living Resources, 14(2):65–80, 2001.
566 24. Y.-J. Shin and P. Cury. Using an individual-based model of fish assemblages to study the response
567 of size spectra to changes in fishing. Canadian Journal of Fisheries and Aquatic Sciences, 61(3):414–431,
568 2004. URL https://doi.org/10.1139/f03-154 .
569 25. Y.-J. Shin, T. Morgane, V. Philippe, R. Oliveros-Ramos, L. Velez, N. Barrier, C. Lujan, M. Hurtado,
570 and W. Lau-Medrano. osmose: Object Oriented Simulator of Marine Ecosystems, 2020. URL
571 https://CRAN.R-project.org/package=osmose . R package version 3.3.4.
572 26. I. G. Taylor, I. J. Stewart, A. C. Hicks, T. M. Garrison, A. E. Punt, J. R. Wallace, C. R. Wetzel, J. T.
573 Thorson, Y. Takeuchi, K. Ono, C. C. Monnahan, C. C. Stawitz, Z. T. A'mar, A. R. Whitten, K. F. Johnson, R. L.
574 Emmet, S. C. Anderson, G. I. Lambert, M. M. Stachura, A. B. Cooper, A. Stephens, N. L. Klaer, C. R.
575 McGilliard, I. Mosqueira, W. M. Iwasaki, K. Doering, and A. M. Havron. r4ss: R Code for Stock Synthesis,
576 2019. URL https://CRAN.R-project.org/package=r4ss . R package version 1.36.1.
577 27. W. Thuiller, D. Georges, R. Engler, and F. Breiner. biomod2: Ensemble Platform for Species
578 Distribution Modeling, 2020. URL https://CRAN.R-project.org/package=biomod2 . R package version
579 3.4.6.
580 28. M. Travers, Y.-J. Shin, L. Shannon, and P. Cury. Simulating and testing the sensitivity of
581 ecosystem-based indicators to fishing in the southern benguela ecosystem. Canadian Journal of
582 Fisheries and Aquatic Sciences, 63(4):943–956, 2006. URL https://doi.org/10.1139/f06-003 .
583 29. M. Travers, K. Watermeyer, L. Shannon, and Y.-J. Shin. Changes in food web structure under
584 scenarios of overfishing in the southern benguela: comparison of the ecosim and osmose modelling
585 approaches. Journal of Marine Systems, 79(1-2):101–111, 2010. URL
586 https://doi.org/10.1016/j.jmarsys.2009.07.005 .
587 30. M. Travers-Trolet, Y. Shin, and J. Field. An end-to-end coupled model roms-n2p2z2d2-osmose of
588 the southern benguela foodweb: parameterisation, calibration and pattern-oriented validation. African
589 Journal of Marine Science, 36(1):11–29, 2014. URL https://doi.org/10.2989/1814232X.2014.883326 .

31. H. Wickham. Advanced R. Chapman and Hall/CRC, first edition edition, 2014. ISBN13: 978-1466586963.

32. H. Wickham. Advanced R. Chapman and Hall/CRC, second edition edition, 2019. ISBN13: 978-0815384571.

33. E. Winsberg. Science in the age of computer simulation. University of Chicago Press, 2010. ISBN10: 0226902048.

Appendix C

run_up: Running Uncertainty aNalysis UP complex models

Automated uncertainty analysis for complex models in R. This package has been created especially to assist in complex model work where the uncertainty estimation of model inputs is required. It is a generic tool that can be used for any type of model. It provides tools for the creation of ranges of variation and the estimation of sensitivity indices, initially using Morris analysis and the calculation of elementary effects.

See < https://github.com/CriscelyLP/run_up > for more details.

Version:	0.1.0
Depends:	R (≥ 2.15)
Created:	2020-10-05
Author:	Criscely Luján [aut, cre], Ricardo Oliveros-Ramos [ctb]
Maintainer:	Criscely Luján < criscelylujan at gmail.com >
BugReports:	https://github.com/CriscelyLP/run_up/issues
License:	GPL-2
NeedsCompilation:	no

Concluding remarks and perspectives

The study of uncertainty in complex ecosystem models: what have we learned?

The importance of the study and quantification of uncertainty is highlighted in many research works (IPBES 2016; Payne et al., 2016; Steenbeek et al., 2021) and has been the focus of many dedicated scientific conferences (e.g., SAMO, the scientific analysis of model outputs; Saltelli et al., 2021) and summer schools organized around the world. The study of uncertainty brings a mixture of theoretical and practical challenges, for which there is no predefined recipes of steps to follow (Rounsevell et al., 2021). However, the presentation of a portfolio of approaches with practical applications on various case studies brings the opportunity to illustrate how to deal with the uncertainty properly. In this research work, we focus on two frequently used techniques to study uncertainty in models: sensitivity analyses and uncertainty analyses. Both approaches bring two different views to the study of uncertainty (Saltelli et al., 2019), for which, according to our literature review, numerous methods and computational tools have been developed to facilitate their implementation.

A difficulty in studying uncertainty in models, especially models characterized by a large number of parameters, is when the quantification of uncertainty in their input parameters is lacking. In these cases, a recurrent practice is to use arbitrary ranges of variation for the input parameters, but this can generate biased results. In this thesis, we propose the parameter reliability criterion, which describes parameter uncertainty based on its reliability as a function of the source of information used to estimate the parameter value of a model. The presentation of this criterion is a practical way to address uncertainty of complex models, but technically this criterion has no limitation to be used in any other mathematical model.

There are some strong computational limitations for running uncertainty analyses of complex ecosystem models that have a large number of input and output parameters, and can be stochastic. UA of these models would typically require the execution of numerous simulations, for which the use of very sophisticated computing machines is required, and which can lead to the generation of large volumes of model outputs. In this work that focused on the NHCE OSMOSE marine ecosystem model, we were faced with these technical complexities. To deal with this, the use of a supercomputer was essential (i.e., the DATARMOR super calculator, https://wwz.ifremer.fr/pcdm/Equipement), as well as the development of computer programs to automatize the uncertainty simulations and the analyses of its many outcomes. We thus co-developed an R package (the OSMOSE R package, Shin et al., 2020) that facilitated the execution of the simulations implemented for this thesis and helped a lot in processing the results. With this we want to emphasize that with the use of complex models, it is also indispensable to have tools (i.e., packages implemented in different programming languages), which make use of rigorous standards in programming, with standardized names, customizable graphics and most importantly, which provide independent functions that allow working on specific tasks (e.g., reading data, executing intermediate analyses and calculating model results). Promoting the development of such tools is a big step towards facilitating the work with complex models; besides this will provide good conditions for running and analyzing a large number of simulations for the study of model uncertainty.

The parameter reliability criterion: what is its contribution ?

The reliability criterion presented in the thesis introduces the basic assumption that parameters derived from data or information with higher reliability have lower uncertainty than parameters derived from data with lower reliability. We use this criterion to draw uncertainty around the input parameters of a model, where the reliability criterion assigns smaller ranges of variation for parameters with higher reliability and wider ranges for parameters with lower reliability. Above all, we use the reliability criterion to deal with situations where there is an absence of probability distributions for model parameters or ranges of variability that are supported by observational or experimental data. This criterion is accompanied

by a set of rules, and framed in a protocol to implement sensitivity analyses. However, this criterion can be technically used to implement uncertainty analyses as well. Although they do not study uncertainty in the same way, both types of analyses require a quantification of the uncertainty in model parameters. The protocol is flexible and generic enough to be adapted to various models and improved. For example, the creation of a new parameter ranking category would just require to follow the same generic recommendations of the protocol such as the adoption of a monotonic function between parameter reliability and uncertainty, and common rules for the calculation of parameter ranges.

As part of the protocol, we also open the discussion to other important technical elements in the study of uncertainty in complex models. We discuss the importance of complying with the mathematical constraints of the parameters included in the analysis and the importance of the re-parameterization that this may require. Complex models can have parameters that do not only include a single numerical value, but may also accept, e.g., vectors, time series, or matrices, and that are distributed on different scales (e.g., linear, logarithmic, logit). All this must be considered when describing the uncertainty around the model parameters, and during the implementation of the selected method to study the uncertainty in the model outputs. These considerations have been explained throughout this thesis, and we have also implemented them in the freely available code we developed to accompany the simulations run in this research (runup R package).

Uncertainty in the NHCE OSMOSE model: what did this analysis reveal?

This thesis addresses the study of uncertainty in the OSMOSE model applied to the northern Humboldt Current Ecosystem (NHCE). In this context, the study of uncertainty in this model allowed us the exploration of the degree of confidence that we can place in its predictions, but it has also been an opportunity to obtain results that could guide future improvements of the NHCE OSMOSE model.

As part of the results obtained from the parametric uncertainty analysis, our work revealed that, as a result of the complex interactions of the NHCE OSMOSE

model, the uncertainty from one species' parameters can impact the rest of the food web. The parameters of two species in particular (i.e., euphausiids and Humboldt squid) had the greatest impact on the modelled system. As these two species are not the most well monitored species of all the species modelled, these results highlight the need for more observations and knowledge about them. In general, the quality of the data for all species needs to be improved, also taking into account the time scale since the behaviour of uncertainty also changes over time, which needs further investigations. Moreover, the conducted uncertainty analysis calls for improving the representation of its components by evaluating the assumptions in its structure. Using the NHCE OSMOSE model, the propagation of species uncertainty affected very little the Peruvian anchoveta (Engraulis ringens), a species considered key in the ecosystem. This is explained by the weak predator-prey relationships between the anchoveta and the rest of the species in the model, indicating an under-representation of the interactions of this species in the NHCE OSMOSE model. For this, future versions of the NHCE OSMOSE model should include the effect of some of its main predators explicitly (such as seabirds, Bertrand et al., 2012) whereas they only appear as a mortality parameter in the current configuration of the NHCE OSMOSE model.

The uncertainty of marine ecosystem models is most often analysed on model output variables such as biomass and abundance of modelled species (e.g., Bracis, et al., 2020; Hansen et al., 2019). However, these models also provide several ecological indicators in output (e.g., Halouani et al., 2019; Moullec, 2019), but little is known about their behaviour under uncertainty. Our work revealed that the response of indicators to uncertainty in the model inputs was not homogeneous; some indicators dampened the level of uncertainty more (e.g., the slope of the size spectrum, the marine trophic index, and the trophic level of fish communities), but others amplified it (e.g., large fish indicators, LFIs). Given that these results were obtained using only two uncertainty scenarios, we strongly recommend further analyses, including the NHCE OSMOSE model as the model under study, but also through an ensemble of ecosystem models that allow for a more rigorous assessment of the behaviour of indicators under uncertainty. Computational costs must always be considered in this type of models, as this was one of the limitations we had to deal with.

Our uncertainty study included the execution of uncertainty analyses and sensitivity analyses. For the sensitivity analyses which entailed a huge number of sim-

ulations to run, we selected only the parameters related to Peruvian anchovy to illustrate the protocol's application and the use of the parameter reliability criterion. Because this was a restrictive analysis, the most sensitive parameters identified are not necessarily the most sensitive parameters of the NHCE OSMOSE model. However, we showed that among the set of anchovy parameters and on average for all experiments run (including the protocol), the four parameters with the strongest main effects are the critical threshold of predation efficiency, the maximum rate of ingestion, natural mortality rate, and larval mortality rate.

Perspectives of the research

Much remains to be done regarding the study of uncertainty in the NHCE OSMOSE model. In this thesis, we have quantified the effect of parametric uncertainty (using the explicitly modelled species parameters) on the model outputs under study, calling for improvements in the quality of the data and the representation of the NHCE species. Furthermore, these data should be improved on both temporal and spatial scales. The outputs of the model under study are also spatially distributed. The characterization of uncertainty in a spatial manner has not yet been studied and needs to be addressed. On the other hand, the OSMOSE modelling platform uses spatial distribution maps as forcing. These maps are built using statistical models (e.g., generalised statistical models, generalised linear models, etc.). The effect of the uncertainty of these forcing has not yet been quantified in the OSMOSE NHCE model, nor in any other case study of the OSMOSE modelling platform. Another very relevant source of uncertainty in the NHCE OSMOSE model, and more generally important in complex ecosystem models, is the uncertainty due to initial conditions.

References

Bertrand, S., Joo, R., Arbulu Smet, C., Tremblay, Y., Barbraud, C., Weimerskirch, H. (2012). Local depletion by a fishery can affect seabird foraging. Journal of Applied Ecology, 49(5), 1168-1177.

Bracis, Chloe, Sigrid Lehuta, Marie Savina-Rolland, Morgane Travers-Trolet, and Raphaël Girardin. 2020. "Improving Confidence in Complex Ecosystem Models: The Sensitivity Analysis of an Atlantis Ecosystem Model." Ecological Modelling

431: 109133.

Halouani, G., Le Loc'h, F., Shin, Y. J., Velez, L., Hattab, T., Romdhane, M. S., Lasram, F. B. R. (2019). An end-to-end model to evaluate the sensitivity of ecosystem indicators to track fishing impacts. Ecological Indicators, 98, 121-130.

Hansen, Cecilie, Kenneth F Drinkwater, Anne Jähkel, Elizabeth A Fulton, Rebecca Gorton, and Mette Skern-Mauritzen. 2019. "Sensitivity of the Norwegian and Barents Sea Atlantis End-to-End Ecosystem Model to Parameter Perturbations of Key Species." PloS One 14 (2): e0210419. IPBES (2016). The methodological assessment report on scenarios and models of biodiversity and ecosystem services (IPBES).

Moullec, F., Velez, L., Verley, P., Barrier, N., Ulses, C., Carbonara, P., ... Shin, Y. J. (2019a). Capturing the big picture of Mediterranean marine biodiversity with an end-to-end model of climate and fishing impacts. Progress in Oceanography, 178, 102179.

Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., et al. (2016). Uncertainties in projecting climate-change impacts in marine ecosystems. ICES Journal of Marine Science, 73(5), 1272-1282.

Rounsevell, M. D., Arneth, A., Brown, C., Cheung, W. W., Gimenez, O., Holman, I., et al. (2021). Identifying uncertainties in scenarios and models of socio-ecological systems in support of decision-making. One Earth, 4(7), 967-985.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., ... Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environmental modelling software, 114, 29-39.

Saltelli, A., Jakeman, A., Razavi, S., Wu, Q. (2021). Sensitivity analysis: A discipline coming of age. Environmental Modelling Software, 146, 105226.

Steenbeek, J., Buszowski, J., Chagaris, D., Christensen, V., Coll, M., Fulton, E. A., et al. (2021). Making spatial-temporal marine ecosystem modelling better–A perspective. Environmental Modelling Software, 145, 105209.

Yunne-Jai Shin, Travers Morgane, Verley Philippe, Ricardo Oliveros-Ramos and Laure Velez. osmose: Object Oriented Simulator of Marine Ecosystems. (2020). R package version 4.3.2. http://www.osmose-model.org/.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences du végétal : du gène à l'écosystème (SEVE)

Titre: Faire face à l'incertitude dans les modèles complexes : une application au modèle écosystémique OSMOSE dans l'écosystème nord du courant de Humboldt

Mots clés: modèles complexes, analyse d'incertitude, analyse de sensibilité, modélisation des écosystèmes, OSMOSE, écosystème nord du courant de Humboldt.

Résumé: L'augmentation de la complexité des modèles permet d'obtenir une représentation plus réaliste des systèmes naturels. Cela peut également conduire à la création d'outils très complexes, dont il est nécessaire d'étudier les sources d'incertitudes, ainsi que le degré de confiance que nous pouvons accorder à leurs prédictions. Cette thèse porte sur l'étude de l'incertitude dans le modèle d'écosystème marin OSMOSE appliqué à l'écosystème nord du courant de Humboldt. Ce travail explore les différentes méthodes et outils disponibles pour les études d'incertitude. Nous avons mis en œuvre une analyse d'incertitude, dont les principaux résultats ont montré que l'incertitude d'un modèle complexe dépend de l'incertitude de quelques paramètres d'entrée, qu'elle se propage dans le temps et que, selon les sorties du modèle sélectionnées, celles-ci peuvent atténuer ou amplifier le niveau d'incertitude de ses résultats. Nous présentons également un protocole basé sur le critère de fiabilité des paramètres. Cela nous permet de classer les paramètres du modèle en fonction de la source d'information utilisée pour estimer leurs valeurs et d'attribuer un niveau d'incertitude aux paramètres du modèle. Enfin, nous proposons une série de recommandations pour les futures études d'incertitude utilisant des modèles complexes.

Title: Dealing with uncertainty in complex models: an application to the OSMOSE ecosystem model of the northern Humboldt current system

Keywords: complex models, uncertainty analysis, sensitivity analysis, ecosystem modelling, OSMOSE, northern Humboldt current ecosystem.

Abstract: Models of increasing complexity help achieve a more realistic representation of natural systems. This can also lead to the creation of very complex tools, which sources of uncertainty must be studied, as well as the degree of confidence we can place in their predictions. This thesis focuses on studying the uncertainty of the OSMOSE marine ecosystem model applied to the northern Humboldt current ecosystem. This work explores several methods and tools available for uncertainty studies. We implemented an uncertainty analysis, where our main results show that the uncertainty in a complex model depends on the uncertainty in a few inputs, which propagates in time, and depending on the model outputs selected, these may dampen or amplify the level of uncertainty in model results. We also propose a protocol based on the parameter reliability criterion. This allows classifying model parameters according to the source of information used to estimate their values and assigning a level of uncertainty to model parameters. Finally, we provide a series of recommendations for future uncertainty studies using complex models.