Reconciling complex system models and fisheries advice: Practical examples and leads

Sigrid Lehuta1,a, Raphaël Girardin2, Stéphanie Mahévas1, Morgane Travers-Trolet2 and Youen Vermard1

1 IFREMER Ecologie et modèles pour l’halieutique, Rue de l’île d’Yeu, BP 2011, 44311 Nantes Cedex 03, France
2 IFREMER Halieutique Manche Mer du Nord, 150 Quai Gambetta, 62200 Boulogne-sur-Mer, France

Received 7 December 2015; Accepted 3 July 2016

Abstract – The move toward an ecosystem-based fisheries management (EBFM) requires new operational tools in order to support management decisions. Among them, ecosystem- and fisheries-based models are critical to quantitatively predict the consequences of future scenarios by integrating available knowledge about the ecosystem across different scales. Despite increasing development of these complex system models in the last decades, their operational use is still currently limited in Europe. Many guidelines are already available to help the development of complex system models for advice yet they are often ignored. We identified three main impediments to the use of complex system models for decision support: (1) their very complexity which is a source of uncertainty; (2) their lack of credibility, (3) and the challenge of communicating/transferring complex results to decision makers not accustomed to deal with multivariate uncertain results. In this paper, we illustrate these somehow theoretical “best practices” with tangible successful examples, which can help the transfer of complex system models from academic science to operational advice. We first focus on handling uncertainty by optimizing model complexity with regards to management objectives and technical issues. We then list up methods, such as transparent documentation and performance evaluation, to increase confidence in complex system models. Finally, we review how and where complex system models could fit within existing institutional and legal settings of the current European fisheries decision framework. We highlight where changes are required to allow for the operational use of complex system models. All methods and approaches proposed are illustrated with successful examples from fisheries science or other disciplines. This paper demonstrates that all relevant ingredients are readily available to make complex system models operational for advice.

Keywords: Ecosystem-based fisheries management / complex models / decision support / methodological solutions / participatory modeling / model sensitivity analysis / examples

1 Introduction

In 2012, the European Commission (EC) asked the International Council for the Exploration of the Sea (ICES) for an implementation plan for provision of ecosystem-based advice. ICES proposed a roadmap, defined by eco-regions, with successive milestones for the provision of support regarding Maximum Sustainable Yield (MSY) limits, data-limited stocks, mixed fisheries advice, multi-species, and wider ecosystem (ICES 2013a). It also plans that the influence on advice of other European initiatives such as the Marine Strategy Framework Directive (MSFD) and Marine Spatial Planning (MSP) be accounted for as they develop. Four years later, while the roadmap predicts that ICES should start providing “wider ecosystem” advice for most eco-regions, it is still uncertain how this kind of advice can be delivered.

Ecosystem-based fisheries management (EBFM) implies accounting for the interactions between ecosystem components, so direct and indirect impacts of management actions may be foreseen and planned. Integrated process-based models, which explicitly describe these interactions and evidence trade-offs between diverse societal objectives (FAO 2003), appear good candidates to support decision making in the scope of EBFM. A diversity of modeling approaches has arisen to support EBFM, from qualitative models to computer intensive quantitative models.

Qualitative models incorporating Bayesian belief networks (Melbourne-Thomas et al. 2013; Reum et al. 2015; Stafford et al. 2015) or time automata (Largouet et al. 2012) have been recently promoted for decision support (Zitek et al. 2009). They are based on a general understanding of the relationships that connect ecosystem variables (Dambacher et al. 2015) and can qualitatively predict system responses to external perturbations (Melbourne-Thomas et al. 2013). They also
are well-suited to data-limited systems (Largouet et al. 2012; Reum et al. 2015) and constitute an attractive approach for rapid model formulation and hypothesis testing regarding ecosystem structure and function (Melbourne-Thomas et al. 2013). If they can be useful to support long-term management strategies as tools to prioritize monitoring and interventions on critical ecosystems components (Dambacher et al. 2015), their utility for tactical advice in fisheries management (e.g. setting TAC, defining technical regulations, designing marine protected area (MPA) networks) is however limited. Indeed, they cannot produce numerical predictions and ambiguities frequently arise in the direction of the response, which resolution would require a nearly complete quantification of the model’s interaction strengths (Dambacher et al. 2015). This is why, many authors advocate for the use of quantitative modeling approaches as management decision tools (Hyder et al. 2015; Link et al. 2012; Plagányi et al. 2014; Rose 2012; Rose et al. 2010; Schmolke et al. 2010; STECF-SGMOS 2010; Steele et al. 2013; Thébault et al. 2014). These models, hereafter referred to as complex system models (CSM), are multi-specific, trophic-based, spatially explicit, and end-to-end models as well as models describing fleet dynamics and economic processes.

CSM, addressed in this paper are mechanistic and dynamic, with no analytical formulation, and therefore they rely on numerical simulation. Parameter number is often high, and parameterization is often a complex issue because integrated estimation is rarely possible. Therefore, parameterization is a pragmatic and iterative exercise, where parameter estimates are borrowed from literature, other models or regions, are newly estimated, or assigned a theoretical value. Finally developing and running these models is time-consuming. The development of CSM increased in the last 20 years from the first appearance of Ecopath models (Christensen and Pauly 1992) to end-to-end models such as Atlantis (Fulton 2010; Fulton et al. 2004) (ICES 2015a; Plagányi 2007; Prellezo et al. 2012). They have long been used primarily to improve scientific understanding of ecosystem functioning and in contrast with simpler approaches, there are few examples of their practical use for decision support, particularly in European fisheries (Hyder et al. 2015).

A reason is that complex system models needed for science are different from models needed for management. As Schmolke et al. (2010) accurately observed an important difference comes from the requirement for publication of integrated modeling framework being originality, while the requirement for their use as decision support tools is predictive capacity and suitability. Understandably, the level of demand for realism and accuracy is higher when ecological, economic and societal consequences are at stake (Dickey-Collas et al. 2014; Oreskes et al. 1994). More realism often leads to more complexity, which undeniably means more parameters that need to be estimated or measured. If uncontrolled, increased complexity can thus lead to increased prediction uncertainty. The relationship between model complexity and effectiveness, i.e. – model’s ability to provide understandable, testable and useful predictions regarding the question of interest, has thus been shown to be bell-shaped (Costanza and Sklar 1985; Grimm et al. 2005). In order to reach this theoretical optimum, model complexity should be carefully rationalized with regards to model objectives and available knowledge.

The absence of justification for increased complexity and appropriate supporting knowledge is often highlighted and advocated against the operational use of CSM (Dickey-Collas et al. 2014; Kraak et al. 2010; Link et al. 2010; Planque 2015; Rochet and Rice 2009; Schmolke et al. 2010). What these authors actually blame is rather the absence of a critical evaluation of model performance and the lack of rigor in the treatment of uncertainties. The difficult communication of the many assumptions and the lack of transparency of the development process often lead to a general feeling of mistrust and suspiciousness toward CSM, even when development is rigorously conducted. Even vivid supporters of complex models are reluctant to using them for tactical management, and advise to restrain from forecasting (Fulton et al. 2013; Rose et al. 2010). As Link et al. (2010) stated, “ecosystem models credibility will need to be established and the rigor of quality control/assurance and peer review will need to be comparable to what is currently conducted for data rich single and protected species stock assessments, to provide management advice in the U.S.”. This is probably the major obstacle to their operational use, even above considerations of cost, development and running time.

Nevertheless, complex models are increasingly used outside Europe to support management decisions; in Australia, Atlantis is used for management evaluation (Fulton et al. 2007) and has started to be used in the Californian current along the US West Coast in the context of an Integrated Ecosystem Assessment (IEA) and in a strategic environmental impact assessment (Kaplan et al. 2012; Kaplan and Marshall 2016). In Alaska, the results of food web models were used to revise the total allowable catch (TAC) for Pollock in 2006 (NPFMC 2006). In Europe, some initiatives of the use of complex models have to be mentioned: for instance, multi-species assessment models are used to provide single species assessment models with predation mortalities for a few stocks in the North Sea (ICES 2015b).

Outside fisheries management, CSM are also used to support decision making in important economic and environmental fields such as climate (e.g. Intergovernmental Panel on Climate Change), petroleum reserves (Mohaghegh 2014), radioactive waste disposal (Oreskes et al. 1994), weather forecasting (Hill et al. 2007), forest management (Pretzsch 2007) and chemical risk assessment (Grimm et al. 2009). So, why should the management of marine ecosystems in Europe be an exception?

Skern-Mauritzen et al. (2016) showed that ICES is well advanced compared to other international management bodies with regard to inclusion of environmental drivers into assessment models. However, it mainly consists of improvements of traditional models, and the format of outcomes and scientific advice remain unchanged. Products of CSM greatly differ from current ICES provided advice, so we need to think how multivariate outputs, scenarios and associated uncertainties could be efficiently delivered to managers. Moreover these new tools will need to find their place in the existing legal and institutional decision framework in place in Europe (e.g. MSFD, EC 2008a; EU 2013).
The understanding gained through CSM will only be used to support management if we can effectively address the impediments to their operational use. A variety of guidelines to ecosystem model development and use are available in the literature (Aumann 2007; Collie et al. 2016; FAO 2008; Fulton 2010; Link et al. 2012; Plagányi et al. 2014; Rose et al. 2010; Schmolke et al. 2010; STECF-SGMOS 2010; Steele et al. 2013; Thébaud et al. 2014; Townsend et al. 2008) and yet they are often ignored (Schmolke et al. 2010). It is not our intention to add a new one to the list. Rather, we intend to exemplify these “best practices” with readily available and applicable methods which can be routinely implemented towards operational use of CSMs.

In this paper, we present methodological tools and approaches to help (i) rationalize complexity with regards to management objectives and knowledge limitations, (ii) improve confidence in CSM through improved transparency regarding uncertainties and performance and (iii) transfer model results to advice in the current European legal and institutional framework. Given the large scope of our paper, the illustrations proposed here are not exhaustive at all and only rely on widely used, generic modeling platforms, leaving out models developed for more specific sites and questions, although potentially relevant.

2 Building “effective models” by rationalizing complexity

Although, all components of marine ecosystems are to some extent connected, it is foolish to model the totality of interactions in play at all scales, given our incomplete and imperfect knowledge and understanding of the functioning of marine ecosystems. In an advisory context, it would be oblivious. Priorities have to be set about which interactions and scales need to be modeled. Two aspects need to be considered to set these priorities: (1) the choice of the appropriate model (platform/model structure/processes to include/scales) regarding management objectives and (2) the need to account for increased uncertainty propagation and technical difficulties with increased complexity. The combination of both is referred to as “model effectiveness” by Costanza and Sklar (1985). In the following we first illustrate which complexity level needs to be used, and has actually been used, depending on the management context. Then we present technics to contain and even reduce complexity regarding available knowledge once the relevant model has been selected.

2.1 Management objectives drive complex system model structure

Several approaches are proposed to reach the “optimal” or “intermediate” level of complexity (Allen and Fulton 2010; Aumann 2007; Collie et al. 2016; Hannah et al. 2010; Hollowed et al. 2000; Hyde et al. 2015; Plagányi et al. 2014; Townsend et al. 2008; Weijerman et al. 2015). For instance, Townsend et al. (2008) provide best practices and guidance for the choice of features that need to be explicit in a model according to ecosystem characteristics. They list model features (species included, population stages, spatial and temporal scale, fleet description) and processes (recruitment, movement, fleet dynamics, trophic relationships, etc.) in a very systematic and efficient way. In addition to ecosystem characteristics, all authors report, that the question at stake, the management objective, should drive the choices of processes considered in the model. In its roadmap to integrated advice, ICES hierarchizes the type of advice that will progressively need to be delivered (single species and data poor, mixed fishery, multi-species, wider ecosystem, multi-species management plans, MSFD and MSP) (ICES 2013a). Using ICES chronology as a backbone, we detail the components/interactions that need to be considered depending on management questions and illustrate the guidance provided with examples of some complex system models presenting appropriate features (see Hollowed et al. (2000), Plagányi (2007), Thébaud et al. (2014) and Travers et al. (2007) for more exhaustive reviews and categorization of ecosystem and bio-economic models). Emphasis is put on model flexibility which allows a modular approach of model building and fine adjustment of complexity (Fig. 1).

2.1.1 Complex system models to improve single and datapoor stocks management

Currently ICES advice is based on well established procedures for single species management involving either stock assessment models or “data-poor” approaches (ICES 2015c). The increasing awareness of the risk of disregarding uncertainties in these simple management approaches has encouraged the development of Management Strategy Evaluation (MSE) loops (Kell et al. 2007; Smith 1999). It consists of dynamic coupling between a classical management model (a single species assessment model or a data-poor approach and a decision model such as Harvest Control Rules (HCR)) and an “operating model” of higher complexity, typically a CSM (Fig. 2). The operating model accounts explicitly for processes susceptible to bias single-species assessments (fishermen’s reactions, spatial heterogeneities, trophic relationship, and bias in data collection). The principle is to use the outputs of the operating model as inputs for the management procedures in order to (i) evidence where shortcuts in the system representation in traditional management tools (ignorance of spatial or seasonal process, trophic cascades etc) can jeopardize management success. It also serves (ii) quantifying the risk represented by uncertainties at every step of the decision process: from observation (collected data), advisory tools (assessment models, indicators, data-poor approaches) and implementation (fishermen’s behavior) (Butterworth 2007; Kell et al. 2007). In this context, CSM serve highlighting the need for a more complex management approach and possibly counter-balance known bias. Currently operating models used in Europe in MSE loops do not present a level of complexity much higher than the assessment models they aim to evaluate. For MSE purpose, CSM should include the processes that are thought to put simple approaches at risk. For instance, Marchal and Vermard (2013) plugged a simple assessment model (XSA, FLR package) to an ISIS-Fish model of the deep water fisheries in West-Scotland, because hyperstability (i.e. persistence of high catch per unit of effort (CPUE) at low stock level, due to density dependence in the spatial distribution of the stock
Fig. 1. Assessment of model complexity and skills against a range of management objectives. Grey boxes represent the minimum components and/or interactions required for a model to be competent to inform a given management objective. Bars represent example models that meet this requirement. This non-exhaustive list includes non-spatial models in red and spatially-explicit ones in blue. As some of these models are modular (i.e. can include or not a particular process or component depending on the application study), only their minimum level of complexity is represented here.

and fishermen’s behavior) was suspected to bias HCR based on CPUE. They evidenced the risks of using CPUE trends in this context as opposite to assessment model results. Similarly, Fay et al. (2011) used MSE loops to explore and optimize HCRs regarding uncertainties about spatial variability for a data-poor species (blue eye trevalla) in southeast Australia. MSE loops have long been considered in Australia (Punt et al. 2001) and are embedded in the modeling platforms Atlantis and InVitro (Fulton and Smith 2004; Gray et al. 2006). However the full loop is still rarely used or at least, we could not find published examples of specific insights gained from the dynamic loop (partly because of a lack of quantitative stock assessment methods (Punt et al. 2001)).

2.1.2 Models for managing mixed fisheries

It has been recognized for long that one pitfall of fisheries management advice is that technical interactions and fishermen’s behavior are not traditionally accounted for when giving stock-by-stock management advice (Fulton et al. 2011; Marchal and Vermard 2013). Technical interactions characterize the interactions occurring between stocks through fisheries. It happens when stocks are caught together either as part of the targeted set or as fish-catch (mixed-fishery). This phenomenon is of less importance in pelagic fisheries that can easily target a given schooling species, but it is frequent in demersal fisheries, where species live alongside each other. These technical interactions become a problem when all species are not sharing the same biological status and one can sustain a high fishing pressure whereas the other should be exploited cautiously. Ignoring the mixed-species aspect of a fishery leads to ignoring well-known fishing practices, which reduce management impact.

Particularly, in mixed fisheries, fishermen usually continue catching and discarding species for which the TAC is exhausted, as long as quota for other species caught simultaneously is available. The consideration of quotas as possibly ineffective constraint on fishing mortality started to be taken onboard by ICES in 2004 with the MTAC approach (Mixed Fisheries TAC, Vinther et al. 2004). It uses the single species assessment and estimates of catchabilities and effort per fleet to derive catches of each species corresponding to scenarios of fleet behavior (“stop fishing when first quota is reached”, “stop when cod quota is reached”, etc.). This approach was improved from 2006 on, using a simpler and more robust approach referred to as Fcube (Fleet and Fishery Forecast) (Ulrich et al. 2011). In 2012, mixed fisheries forecast based on Fcube were integrated for the first time into ICES stock advice for the main demersal stocks of the North Sea. This progress was made possible by the fact that (i) Fcube is specifically designed for advisory data collection and stock assessment methods in Europe (operational definition of métiers and fleets in the European Community’s Data Collection Framework (EC 2008b)); (ii) an ICES working group (WG MIXFISH) dedicated to update the model yearly (ICES 2015d) and (iii) demonstrated the interest of the approach to managers. Since then, mixed fisheries forecasts have been provided to managers every year to inform on single stock advice inconsistencies and potential “choke” species by fleet (a species
Fig. 2. Position of complex system models within the advisory and decision framework in the European context. Black arrows represent manager’s actions, starting from the formulation of questions and management objectives jointly with stakeholders. This drives the advisory process (white arrow) where blue elements represent authors’ recommendations for the inclusion of complex system models. Green arrows highlight where interactions in both ways with current approaches used to provide advice (grey boxes) would be required to better evaluate parameters or scenarios. Purple arrows emphasize the importance of promoting inclusion of stakeholders and managers, not only at the end but throughout the whole advisory process.

The model simplicity however shows its limits because it does not account for catchability at age (Ulrich et al. 2011), nor for fleet dynamics and reactions to management. In a context of increased constraints on fishing, and particularly tentative discard reduction due to the Landing Obligation (Council Regulation (EU) No. 1380/2013), it is very likely that fishermen will change their fishing practices, with consequences on catch profiles and catchabilities. More sophisticated models were therefore built using fisheries approach (i.e. effort distribution across fleets and métiers) among others ISIS-Fish (Lehuta et al. 2015; Mahévas et al. 2012; Marchal and Vermard 2013; Pelletier et al. 2009), DISPLACE (Bastardie et al. 2013), FLBEIA (Garcia et al. 2013) (see (Prellezo et al. 2012). They include explicit selectivity patterns and mesh size, population dynamics, spatial heterogeneities, and possibly fleet dynamics driven by economic factors. Such models only recently started to be used to assess multi-species management plans in Europe. For instance, in 2005, a multi-fleet, multi-stock simulation model (Poos et al. 2006) was used to answer a request from the Dutch ministry of Agriculture, Nature, Conservation and Food Quality on the likely effects of the recovery plan proposed by North Sea Regional Advisory Committee on sole and plaice stocks in the North Sea. While we may ignore more of these national initiatives to address public policy questions using fisheries approaches, the first call at the European scale for these models was launched by Scientific, Technical and Economic Committee for Fisheries (STECF, EC scientific committee) in 2014 for the evaluation of pluri-annual management plans, in application of the new Common Fishery Policy (CFP). The call explicitly requested models that include technical interactions, allow the split into fleet segments or gears, implement changes in selectivity (as a proxy to simulate area/time closures) and split catches into two categories, below and above Minimum Conservation Reference Size (STECF 2015). To answer this call, a large diversity of models were used (ISIS-Fish, Ecopath with Ecosim (EwE), FLBEIA, Fcube, IAM (Macher et al. 2008), Simfish (Bartelings et al. 2015) and Fishrent (Simons et al. 2014)).

Other efforts are conducted to account for fishermen behavior in traditional management tools. Within the European project MYFISH (EU FP7 289257), MSY targets have been re-estimated using yield as a function of increasing effort instead of yield as a function of F and accounting for fleet dynamics. It demonstrated that reference points are highly sensitive to hypotheses of fishermen behavior.

The shift to a fleet-based and effort-based approach also allows the consideration of fleet economic sustainability, which is an objective of the new CFP. Models that address fleet dynamics often describe explicitly the consequences of fishermen’s behavior on costs and revenues, and allow to derive profitability (Bastardie et al. 2010). Economic objectives such...
as Maximum Economic Yield can therefore be considered and evaluated accounting for exploitation patterns and economic structures (Guillen et al. 2013).

Although the focus of fleet based models is limited to commercial species, modeling fishing effort gives access to indirect effect of fishing. Particularly, the evaluation of effort distribution in space, allows assessment of fishing impact on habitats, on non-targeted species or/and on species for which fishing mortality is not assessed (ICES 2015e). Interactions between uses of the maritime domain can also be explored (Girardin et al. 2015).

2.1.3 Models for managing marine communities
(multi-species advice)

In these fleet-based models as well as in traditional assessment models, natural mortality parameter (M) is considered constant despite some early warning of the potential bias it could imply on model estimates (Daan 1987; Sims 1984). While such assumption might have limited effects on stock assessment when fishing mortality (F) was much higher than M, this is no longer the case. Indeed F has been greatly reduced in the recent years in order to preserve and even rebuilt stocks that were overfished or depleted. Thus, recent studies have re-addressed the implication of taking into account the variability of M – and more precisely its major component, the predation mortality (Hollowed et al. 2000; Johnson et al. 2015; Longo et al. 2015). In parallel, operational working groups have started to acknowledge the importance of species interactions, firstly by updating the mortality term in single-stock assessment models using time series of predation mortality (M2) derived from multi-species assessment (Gaichas et al. 2010; ICES 2015b), and secondly by directly using multi-species assessment models (e.g. Stochastic Multi-Species model (SMS), ICES 2014; Lewy and Vinther 2004) with the objective of deriving biological reference points (BRP) in a multi-species context (ICES 2013b).

A more qualitative use of trophic models can also be envisaged, such as the revision of the TAC for Walleye Pollack in Eastern Bering Sea in 2006. Faced with strong evidence for a threat on juvenile Pollack supported by a combination of ecosystem indicators and trophic models the Scientific and Statistical Committee agreed on a precautionary 10% decrease on the TAC provided by the single-species advice (North Pacific Fisheries Management Council’s Scientific and Statistical Committee minutes, December 2006).

Although the concept of MSY has been recently maintained as an operational target to be attained by 2020 (EU 2013), a caution exists regarding its applicability in a multi-species context (Larkin 1977). Multi-species models can thus be used to explore the potential yield of several species under various F. In their review, Tyrrell et al. (2011) pointed out the underestimation of M in traditional approaches especially for forage species and the shift for more conservative BRPs when predation is considered using multi-species models. For predator species, some simulations using SMS in the North Sea showed that the fishing mortality values leading to MSY (FMSY) derived from single-species model are smaller than FMSY computed in a multi-species context (ICES 2013b).

Finally, using a size-based multi-species model, Pinnegar et al. (2014) found that it was not possible to reach concomitantly the MSY values assessed independently for two competitor gadoids, cod and whiting, given the current exploitation patterns and technical interactions.

Beyond the usefulness of multi-species models for adapting traditional targets such as MSY to multi-species context, trophic models could also be used to explore the ecosystem effects of management actions, in order to avoid undesirable situations such as fishing-induced trophic cascades (Salomon et al. 2010). Such tools would thus allow defining a safe operating space (Scheffer et al. 2015) for the long-term exploitation of marine resources, associated with tipping points and time window of management effectiveness (Hughes et al. 2013; Marzlaff et al. 2016).

2.1.4 Models for managing under climate change

The impact of climate change on marine ecosystems and fisheries adds additional challenges to fisheries management, and forces fisheries managers to consider the resilience of both ecological and human system (Fluharty 2011). The subject has been widely explored during the last few decades (Brander 2007; Lehodey et al. 2006; Perry 2005). Complex ecosystem models allowed for instance investigating the combined effect of climate change and fishing on the ecosystem (Travers et al. 2007; Morgane Travers-Trolet et al. 2014). Atlantis enabled the evaluation of multiple management scenarios under climate change and highlighted the diversity of ecosystem response depending on the scenario and management objectives (Fulton et al. 2012; Kaplan et al. 2010).

On more operational aspects, these models can offer insights about the impact of accounting for climate change on operational indicators, such as MSY (Kaplan et al. 2010) and multi-species MSY (MMSY; Fulton 2011; Worm et al. 2009).

However few studies have focused on the impact on harvest strategy and management process (Haynie and Pfeiffer 2012). The development of complex models that encompass biophysical and human component of the ecosystem, such as Atlantis and InVitro, or the coupling of bio-physical model with ecosystem model such as Ecospace (Bulman et al. 2006; Gribble 2003) or size-spectrum models (Barange et al. 2014) allow simulating the propagation of climate change effects across the components of the ecosystem (Fulton 2011; Plagányi et al. 2011). The use of those models revealed for instance, the difference of fisherman’s responses to climate change depending on their flexibility, with for example an increase of incomes for large commercial fisheries while income decreased in smaller scale, less flexible fisheries (Fulton 2011).

2.1.5 Models for spatial management

Spatial management through the design and implementation of MPAs has been considered as effective tools for EBFM (Gewin 2004; Hall 1998; Roberts 2002). The implementation of coherent MPAs network by members’ state in Europe is a requirement of the MSFD (EC 2008a; Qiu and Jones 2013). The evaluation of the coherence of MPAs network and their efficiency remains challenging. Indeed, efficacy of protection
within the protected areas has generally been demonstrated (Colleter et al. 2012; Lester et al. 2009; Valls et al. 2012), however the benefit for adjacent fisheries is more complex to evaluate (Bastardie et al. 2014; Fock et al. 2011; Greenstreet et al. 2009).

Modeling approaches have been widely applied to assess the impact of spatial management using from single species model to end-to-end holistic model (Apostolaki et al. 2002; Brochier et al. 2013; Fulton et al. 2015; Holland 2000; Kaplan et al. 2012; Maury and Gascuel 1999; Pelletier and Mahevas 2005; Savina et al. 2013; Walters 2000). By considering explicitly the heterogeneities in the distribution of vessels, species and/or habitat, as well as competition for space and resources, spatial complex models avoid the bias of classical proxies such as reduction in fishing mortality or effort proportional to the closed surface. They also account for effort reallocation outside the MPA, a behavior which proved to diminish the expected benefits of MPAs.

Spatial modeling is also crucial to capture biological and life history trait such as migration, density dependence, connectivity between regions, both in terms of adult movement or pelagic larval dispersal between spatially-discrete regions or communities (D’Agostini et al. 2015; England et al. 2009; Savina et al. in press). Nevertheless, movement dynamics are rarely known and the implications of choices made regarding spatial resolution, are still active research themes.

Full ecosystem model such as Ecospace (Walters et al. 1999) and Atlantis (Fulton et al. 2007, 2004), as well as CSM focusing more particularly on fishermen’s behavior such as ISIS-Fish, and also lately DISPLACE have been applied to investigate the effects of MPAs around the world. These models have investigated essentially the impacts on biomass (Kraus et al. 2008; Valls et al. 2012), trophic networks functioning (Albouy et al. 2010; Colleter et al. 2012; Libralato et al. 2010), the impact outside the MPAs (Ainsworth et al. 2012; Guénette et al. 2014; Kaplan et al. 2012; Savina et al. 2013; Walters et al. 1999) and also their regional impact on fisheries’ economic performance and behavior, beyond the MPAs boundaries (Ainsworth et al. 2012; Bastardie et al. 2014; Kaplan et al. 2012; Kraus et al. 2008; Morzaria-Luna et al. 2012).

In these studies, the ecosystem modeling approach suggested that spatial management alone was not sufficient to reach EBFM objectives. They demonstrated that making space explicit in the model is necessary to design efficient management scenario.

2.1.6 Models for stakeholders and managers

Collie et al. (2016) suggested that expert panels are necessary for identifying the level of model complexity appropriate for specific management questions. The implementation of consultative processes involving stakeholders for the development of multi-species management plans (Regionalization, CFP) offers the opportunity to test the approach. Fulton et al. (2013) undertook such a participatory exercise for the implementation of an MSE in Australia’s North-West. They report the difficulties, which led to substitute the initially intended In-Vitro model (Gray et al. 2006) with a simplified model, developed in collaboration with stakeholders. Several other attempts are ongoing to try and use complex models within participatory approaches (EU project GAP2 and DiscardLess, ICES WAKIrish and WGAEWESS, Coselmar French regional project, ocean modeling forum) and the growing collaborations with social scientists should help overcome the difficulties to explain and interact around CSMs.

While the overall complexity of these models can be an impediment to stakeholders’ engagement, we believe that structured participatory approaches can facilitate model understanding and acceptance by non-modellers. We believe that the explicit description of processes in CSM is possibly closer to stakeholders representation of ecosystems than are abstract statistical models. For instance, the description of gear characteristics and fishing locations is probably more accessible and more susceptible to trigger feedback from fishermen, than partial fishing mortalities.

2.2 Methods to rationalize complexity according to knowledge and technical limitations

Rationalizing model complexity is also driven by the will to limit uncertainty and by costs and computational considerations. Managing uncertainty requires supporting model parameterization with validated knowledge and data. When knowledge is missing on key processes, hypotheses should be made with parsimony. However, data collection can be expensive. For instance, stomach content information could not be updated from 1991 to 2013 for the SMS model of the North sea (ICES 2014). Data availability is also especially problematic for spatial models, where implementation is data intensive and often implies refining temporal scale simultaneously to account for seasonal and even daily changes in spatial patterns. Data availability is often insufficient to support such fine-tuning. Indeed, while vessel monitoring system (VMS) give access to fine spatial and temporal resolution of effort distribution, fish population distribution are still often only known from surveys at a given time of year. Data availability to support model parameterization should thus be considered in order to limit foreseeable sources of large uncertainties. Running times also often increase with increased complexity which may impact the number of simulations that could later be done. In the following sections, we present two methods to try and adjust model complexity regarding these aspects.

2.2.1 Model coupling

An efficient way of ensuring adequate level of details for each layer (component) of the model is model coupling (Melbourne-Thomas et al. 2011; Shin et al. 2010). In particular, coupling allows adapting the scale to the level of available information and relevant processes: fine spatial and temporal scales for physical, physiological, individual processes (larval drift, prey encounter, plankton productivity) and larger scales for management (annual TAC, seasonal closures, MPA) or fleet dynamics. For instance (Travers-Trolet et al. 2014a) coupled a 3D hourly bio-physical model (ROMS-N2P2Z2D2) to a 2D bi-weekly multi-species fish model (Osmose) to explore trophic functioning under both environmental and fishing
pressure. Girardin (2015) coupled the Atlantis model of the Eastern English Channel, resolved at daily scale and based on irregular spatial polygons to represent the heterogeneity of habitats, with a fleet dynamics model at a monthly time step and whose areas are ICES statistical rectangles in order to account for fishermen’s behavior in the evaluation of spatial management strategies. With increasing questions raised regarding ocean evolution under global change (Cury et al. 2008), model coupling is more often used (e.g. Barange et al. 2014; Lefort et al. 2015; Pinsky and Fogarty 2012; Travers-Trolet et al. 2014b). This is also the approach adopted by two ICES working groups, WGSAM and WGMIXFISH, involved in advice for multi-species and mixed fisheries respectively, as they couple a module based on Fcube describing fleets with the SMS model of the North Sea (ICES 2015d, 2015e). Such couplings between already developed models could efficiently lead to end-to-end models encompassing models of marine socio-ecosystems from bio-physics to socio-economics as demonstrated by Melbourne-Thomas et al. (2011) for the Mexican state of Quintana Roo or by the currently under-development Symbioses framework for the Lofoten and Barents Sea (Caroll and Smit 2011). The increasing needs for model coupling also led to the development of coupling tools (e.g. Beecham et al. 2015). Coupling is also a way to bridge disciplines, make the best use of previous work and increase confidence since different model components have been developed by different and specialized teams (Thébaud et al. 2014; Travers et al. 2007).

2.2.2 Sensitivity analysis

Aside from common sense and expert groups, quantitative methods are also available to help guide model developments at a later stage. In particular, sensitivity analysis is a powerful tool to explore model behavior and response to parameter uncertainty. It consists of quantifying the relative influence of each input parameter on the outputs of the model through computing sensitivity indices (Saltelli et al. 2008). Sensitivity indices help (i) orient model development and choose an appropriate level of complexity by focusing on the most influential processes; (ii) gain insights in model functioning by highlighting strong interactions; (iii) orient data collection and estimation effort toward the most impacting sources of uncertainty; and (iv) focus uncertainty analysis on the most sensitive uncertain parameters. Numerous methods exist to carry out sensitivity analysis, from simple screening methods (e.g. Morris designs), factorial fractional or optimized designs and analyses of variance, or more simulation-intensive methods such Fourier Amplitude Sensitivity Test (FAST) and Sobol’s indices (Saltelli et al. 2008). All methods involve the combination of a statistical method to compute sensitivity indices and a parsimonious simulation design. These designs are optimized in terms of the number of simulations and values taken by parameters. The choice of a particular method depends, on the one hand, on the nature (continuous, discrete) and number of the inputs and on the second hand, on the number of simulations allowed. They offer trade-off between cost (number of simulations) and quality of the exploration of the parameter space. Many of these methods are model-free considering the model as a black box and therefore allow application to virtually every model (Faivre et al. 2013). We recommend to start using inexpensive and rough methods, such as the Morris method or group screening (Saltelli et al. 2008), and progressively refine the analysis as the number of parameters decreases (using for instance FAST or Sobol methods (Saltelli et al. 2008)). Indeed it has been empirically shown that regardless of the number of model parameters, there are often only a few (~10) parameters to which the outputs are highly sensitive (Faivre et al. 2013). Sensitivity analysis can be used to assess the robustness of conclusions to uncertainty in model structure. Alternative structures could be formulated in different ways: various formulations of a process, alternative discrete values of a parameter (e.g. alternative functional responses (Mackinson et al. 2003), alternative fishing behavior (Marchal et al. 2011). Simple sensitivity analysis approaches were used to optimize the number of functional groups in an Ecopath model of Port Cros National Park (Mediterranean, Prato et al 2014). Rochette et al. (2009) used sensitivity analyses to build an optimized semi-quantitative model of the Gironde Estuary trophic network. The identification of larval survival as the most sensitive process in the ISIS-Fish model of the Bay of Biscay anchovy fishery, led Lehuta et al. (2013b) to use outputs of a bio-physical larval drift model to refine the parameterization of larval survival.

3 Improving confidence in complex system models through increased transparency regarding uncertainties and performance

Modeling choices will always be somewhat subjective and different expert groups would probably end up with different models for the same management question and the same ecosystem. Being transparent about modeling uncertainties and confronting models to data are ways to learn about model performance and improve confidence in its predictions.

3.1 Improving communication and transparency around models

The co-existence of alternative complex system models to answer a given management question can be a strength (see Sect. 3.3). But, it can also create confusion among scientists and stakeholders, and raise the question of why consider several models to provide management recommendations and make decision? Answering these questions requires rigorous knowledge on each model regarding their utility and their quality, both in general and at the case study level. Reviews by Plagányi (2007) and Pretelzo et al. (2012), and the guidance and rigorous approaches presented in this paper, provide material to start answering. Initiatives have been taken within research projects and expert groups where different models were used simultaneously. For instance, the need for comparing hypotheses and outputs led to the development of different forms of documentation. Standardized “model matrices” or “summary sheets” with various levels of details, have been developed within ICES working groups (Working Group on Integrating Ecological and Economic Models (SGIMM), Working Group on Multi-species Assessment
Methods (WGSAM), within European networks (Eur-Oceans model shopping tool), within European projects (EU project SOCIOEC and DEVOTES) and by the Joint Research Centre in the context of the evaluation of pluri-annual management plans. The same need was identified within the EwE community and led to the creation of EcoBase, a repository of Ecopath applications worldwide with detailed information on each model goal and parameters (Colleter et al. 2015). Californian current Atlantis review panel similarly expressed concern about using a model for management applications unless full and formal documentation was freely available (NMFS 2014).

Beyond static documentation of a model, WGSAM developed “keyruns”, which are standard reference runs of a given model, accepted within the working group, and which are carried out every three years or after each substantial change to the model or data (ICES 2014). It was established because the group provides single-species assessment working groups with updated and documented predation mortality estimates from multi-species models. The requirement for a standardized documentation format when models are used for advice is further emphasized by Schmolke et al. (2010). They propose a format encompassing the entire modeling process (from parameterization to uncertainty analysis) and acknowledging its iterative nature (cyclic revisits of model formulation and parameterization). It thus allows easy access and understanding of model characteristics and performances. The “parameterization” section of a TRACE documentation could also benefit from Dankel et al. (2012) proposal to use pedigree matrices for quantitative scientific information. The “pedigree” completes the, usually provided, value, unit, and standard deviation of a parameter, with expert judgment of the reliability and a characterization of the origin and status of the information. We found only one publication reporting the use of pedigree matrices in the context of EBFM (Issaris et al. 2012). These could easily fit the current structure of international datasets under development within ICES (InterCatch, FishFrame Regional Database http://www.ices.dk/marine-data/data-portals/Pages/RDB-FishFrame.aspx) and supported by working groups involved with data collection (e.g. WGBIOP, PGDATA). It would probably also help progress on the characterization of uncertainty domains, a major difficulty in uncertainty analyses (Kraak et al. 2010; Rochet and Rice 2009).

These forms of documentation may however not be appropriate for stakeholders. The participatory approach conducted during the recent evaluation of the ecosystem impacts of Pacific sardine harvest showed that open dialogue could actually be sufficient. The impact assessment was conducted in the context of the ocean modeling forum (oceannmodelingforum.org) using five different complex system models. Although models were not exposed in details to stakeholders, model acceptance by managers and stakeholders was described as not being an issue, because of the frequent and transparent exposure of scientific advances to non-modelers (Tessa Francis Ocean Modeling Forum managing director, pers. comm.).

### 3.2 Confronting models to data

Pastoors et al. (2007) report that fishermen more easily related to model fit to past time series than to technical complexities of the modeling process to decide whether a model was trustworthy. Confronting models to data, quantitatively informs modelers, stakeholders and managers on the degree of confidence that can be placed on the model.

Beyond the philosophical considerations on “establishment of truth” and the appropriate terminology to use (“verification”, “validation”, “confirmation” (Oreskes et al. 1994; Sargent 2007; Sterman 1984)), the exercise of confronting model outputs to independent observations is not popular within ecosystem modeler community. However, following Sterman (1984) and Link et al. (2012), we believe that the reluctance to confront models to observations participates to the lack of credibility of these models and holds up their operational use. Moreover, this is modeler’s responsibility to be transparent about their model performance when public policy is at stake (Dickey-Collas et al. 2014; Oreskes et al. 1994).

Here we distinguish calibration from validation. Calibration is the adjustment of model parameter values to fit certain patterns or data (using optimization algorithms for instance). It is often used as a way to estimate parameters internally for relatively simple and quick models and is seldom applied to CSM. When used, it is often restricted to the estimation of few parameters (Ecopath: respiration; OSMOSE: larval mortalities; ISIS-Fish: accessibility parameters). Calibration results, although informative, do not constitute a strict validation. This is because these models are often over-parameterized with respect to available data (also called non identifiability) and there are possibly a large number of alternative parameterizations that lead to equally good reproduction of data (Aumann 2007; Cory et al. 2008; Oreskes et al. 1994; Wiegand et al. 2003). Mackinson (2014) also evidenced the sensitivity of model parameterization to the data used for calibration.

Validation assesses model ability to capture observed patterns of the studied system and implies confrontation of model results to observations which were not used for parameterization. This is a major problem for validation of CSM, because the parameterization phase often requires all available data. When possible, the issue should be anticipated and a subset of available observations kept aside for validation purposes. An alternative approach is to confront emergent properties of the model to observations, because parameterization is done at lower-level (e.g. individual level) and thus does not use these observations at system-level (e.g. population level).

In addition, because of the over-parameterization issues mentioned above, authors highlight that in order to be “structurally valid” a model should probably reproduce multiple time series of data at multiple scales. Pattern oriented modeling (POM) applies this principle (Aumann 2007; Cory et al. 2008; Grimm et al. 2005; Wiegand et al. 2003). It has been successfully applied to an end-to-end coupled model of the southern Benguela foodweb to valid individual observed diets, the seasonality of population size-structure and the structure of the foodweb (Travers-Trolet et al. 2014b). Marzloff et al. (2013) also used POM to valid the capacity of the model of seaweed-based rocky reef communities in Eastern Tasmania to demonstrate shifts in community composition. Model skill assessment is a powerful method to assess model performance thoroughly and is commonly carried out for bio-physical models (Allen and Somerfield 2009; Friedrichs et al. 2007;
Jolliff et al. 2009). It forces asking the relevant questions regarding model objectives such as the type of properties the model should verified (unbiased, in phase, following central limit tendencies, accurate, see Sterman 1984 for corresponding evaluation metrics), and regarding the outputs (including the timeframe) that should in priority match data (Lehuta et al. 2013b). Moreover multivariate graphical solutions were developed to allow easy display of trade-offs (Jolliff et al. 2009).

Nonetheless, validation should not be overvalued because, as stated by Lynch et al. (2009), and particularly true in marine systems, observations are possibly erroneous, and accuracy of data must be considered. It is also admitted that it exists no objective test of validity, so the step should rather be seen as a mean to evaluate model strengths and weaknesses and assess model appropriateness for a purpose. A model should not be thrown out as soon as a criterion of model performance is bad. As Link et al. (2012) stated, there is a lot from learn to failures of models to reproduce past events as illustrated by several studies (e.g. Aumann 2007; Lehuta et al. 2013b; Mackinson 2014; Romagnoni et al. 2015). More generally, confronting model to data (in calibration or validation), is an objective way to select among several alternative parameterizations of the same model (Lehuta et al. 2013b) or to weight alternative parameterizations or models in an ensemble approach (Stoltz 2010; Thorpe et al. 2015) (see Sect. 3.2).

Finally, Sterman (1984), reports considerations that validation is an entirely social process and a model is realistic to the extent that it is understood and accepted. Calibration and validation can be jointly performed in a more qualitative approach to lead to a common model acceptance by all stakeholders. Rosa et al. (2014) report how a participatory approach around an Ecopath model helped sort out conflicting hypotheses about the decrease in health of a tropical coastal lagoon. The impact of (i) bridge building proposed by fishermen and (ii) invasive species proposed by scientists were independently parameterized in Ecopath. Results showed that including impacts of bridge building represented better accordance with the underlying data.

3.3 Transparently reporting on uncertainty

3.3.1 Dealing with parameter uncertainty

The ways in which uncertainty is propagated in simulation models are often criticized (Kraak et al. 2010; Planque 2015; Rochet and Rice 2009). Although, there is an abundant literature about the treatment of uncertainty in fisheries models (Harwood and Stokes 2003; Hill et al. 2007; Link et al. 2012), little applies to CSM. Indeed, carrying out uncertainty analyses is a challenge with CSM, because of the number of parameters and long running times. We do not pretend to have solutions to these methodological and technical issues. Nevertheless, we propose to decrease the complexity of this analysis by performing a global sensitivity analysis as a first step of uncertainty analysis. We advise that sensitivity analysis be used as an objective means to select the most sensitive factors that effectively need to be considered in uncertainty analysis (Lehuta et al. 2010). Furthermore, experimental design theory that support many sensitivity analysis methods also offers alternatives to Monte-Carlo approaches, with optimized simulation designs (Kleijnen 2005) and the use of metamodels (see for instance Collie (2003), for an application in fisheries).

3.3.2 Dealing with structural uncertainty

Dealing with structural uncertainty has received more recent attention (Fulton 2010; Hill et al. 2007; Hyder et al. 2015; Link et al. 2012). The relatively subjective choices in the processes to include, together with the multiplication of models, and the possible multiple model formulations that can achieve the same level of fit to data (especially when trade-offs between scales and outputs have to be made) increased awareness of the necessity to account for model structural uncertainty (Hyder et al. 2015; Weijerman et al. 2015). The most popular approach to handle this type of uncertainty is called multi-model approach. Multi-model approaches have been used for a long time in climate research and whether forecast (Hill et al. 2007; Solomon et al. 2007) and have started developing for marine ecosystems (http://www.oceanmodellingforum.org, http://www.marine-ecosystems.org.uk, Fish-MIP initiative, Forrest et al. 2015; Gårdmark et al. 2012; Hobday 2010; Meier et al. 2014; Smith et al. 2015; Thorpe et al. 2015).

Multi-model approaches are considered to be more robust especially in changing conditions, because they protect against punctual single model failures and offers complementary visions (cross-disciplinary) of the same system. Moreover, while an agreement between models increases the confidence in the robustness of models outputs, a disagreement between models is still informative because it highlights sensitivity to key features of the system. Two paths emerge within multi-model approaches. The first one consists in showing the variability between models (model envelope), while the second one aims at selecting the best model or averaging available models (model averaging). Model envelope approach requires a common modelling protocol with an experimental setup using standardized data and scenarios as basis of multi-model intercomparison and assessment (Rosenzweig et al. 2013; Rötter et al. 2011). Even if model envelope based approach looks for common features of the response shown by most or all models, it offers the opportunity to analyse the origin of the discrepancies between models as a basis for model improvements (Haddeland et al. 2011; Schellnhuber et al. 2014). However, taking into account structural uncertainty in the forecasts may lead to several (sometimes highly contrasted) scenarios of the future of the system. The strength of model averaging or model selection is to deliver a single scenario. The method either selects the “best” model or creates an average model, by weighting each model of the collection with a performance criterion such as Akaiake Information Criterion (AIC; Burnham and Anderson 2004). This approach (also called ensemble approach) has been largely applied in all fields using statistical models. When statistical performance of goodness of fit or accuracy are lacking, like in most complex modelling exercises, each model is weighted equally or expert judgment is required (Link et al. 2012). We recommend not opposing these two multi-model approaches against each other: model envelope
is a relevant strategy to quantifying risk, while model averaging is useful to provide a single value and therefore is more suitable to decision-makers.

In combination with validation, uncertainty analysis provides the material for stakeholders to decide if the model is appropriate to address tactical questions (TAC setting, achievement of target values) or if it should restrict to strategic more qualitative questions (related to ecosystem trends, or relative comparisons between strategies).

4 Transferring model results to advice in the current European legal and institutional framework

Even the most appropriate and rigorous model is of no use, if its results cannot be transferred efficiently in the current management setting. Leslie et al. (2015) observed that major difficulties in implementing EBFM come from pre-existing management entities that may not be well-suited for an ecosystem approach. The decision framework in Europe so far relies on advice delivered by ICES and STECF, based on the Common Fishery Policy, and mainly restricted to quotas corresponding to precautionary approach and latter MSY approach. However, with the new CFP, it is currently transitioning to EBFM through the implementation of MSFD and the development of multi-species management plans at regional level. ICES also makes clear its will to broaden the type of advice it is delivering. If we want CSM to be used, we need to match them with the existing decision framework. In the following we examine how CSM development and results can find their place within the current institutional and legal settings and where arrangements will be needed to ensure an optimal use of their potential.

4.1 Delivering complex product

Communication of complex outputs is a challenge given that managers in Europe are used, and usually prefer, to deal with a unique number (advised TAC) or two (fishing mortality and spawning stock biomass). However, the beginning of an acceptation of more complex advice is demonstrated, for instance by the success of the inclusion of mixed fisheries advice in the advice sheets under the impulsion of the MIXFISH working group or by the intrinsically multivariate nature of MSFD indicators. The recent review of the Atlantis model of the Californian current also argues in the sense of a more important engagement of modelers with advisory bodies, first to improve model understanding by managers and second to improve the relevance of the tested scenarios and products, which were often considered too stylized (NMFS 2014).

It is also the responsibility of managers to more clearly specify objectives so appropriate evaluation criteria can be identified and produced. A key advantage of CSM is their ability to evidence trade-offs in management objectives, a crucial need for EBFM (Kaplan et al. 2012). Literature is developing to propose visualization methods that ease the communication of multivariate results from classic radar plots and traffic lights developed for indicator-based assessment, to stylized images (Collie et al. 2016; Levin et al. 2015; Link et al. 2010; Trenkel et al. 2007). Graphical solutions such as Taylor diagrams have also been proposed to display model skill assessment results for coupled bio-physical models (Jolliﬀ et al. 2009).

Communicating uncertainties is challenging but becomes critical when it comes to complex system model outputs (Hyder et al. 2015). Methods start to be proposed to visualize and make model results and uncertainty more easily accessible to managers (Booshehrian et al. 2012; Spiegelhalter et al. 2011). Moreover we need to teach managers to accept the use of uncertain predictions towards decision making (Kraak et al. 2010). Conversely, we need them (possibly with our help) to specify the kind of uncertainty framework they are willing to accept (precautionary, robust management, risk assessment) (Charles 1998; Dankel et al. 2012; Francis and Shotton 1997; Harwood and Stokes 2003; Hill et al. 2007).

4.2 Informing European initiatives: Indicator approaches and Marine Spatial Planning

At the European scale, Marine Strategy Framework Directive (MSFD, 2008/56/EC), is the legal framework developed to implement EBFM. Indicators and their target reference points are the tools to monitor ecosystem status and progress toward “Good Environmental State” (GES, that is nevertheless still not well defined) within this framework. Now that MSFD is about to enter its effective implementation phase, new tools are sought to evaluate the management measures and scenarios to restore GES, particularly modeling platforms (Joint Research Centre 2014). Moreover, MSE has long been proposed to evaluate the performance of indicator-based HCRs (Butterworth et al. 2010).

More recently CSM have been proposed to evaluate the properties of MSFD indicators (Piroddi et al. 2015). Indeed, in order to be informative indicators need to be sensitive and exclusive to the pressure of interest (often fishing), which means that they should be robust to other sources of variation (environment, uncertainties) (Roche and Trenkel 2003). However, it is generally impossible to disentangle causes of variations in empirically collected metrics and authors proposed to rely on simulation (Fulton et al. 2005; Lehuta et al. 2013a; Livingston et al. 2005; Pelletier et al. 2008; Travers et al. 2006). Fulton et al. (2005) used Atlantis to assess the ability of indicators to detect trends in variables of interest in the ecosystem. Travers et al. (2006) investigated the relative impact of fishing and environmental forcing demonstrating antagonistic and synergistic effects and highlighting the importance of the scale at which indicators are defined. Lehuta et al. (2013a) used sensitivity analyses to evaluate the relative impact of management measures, environmental scenarios and model uncertainties on indicators proposed to monitor the achievement management objectives, evidencing the low proportion of them actually primarily sensitive to management changes. Based on Ecopath, Ecosim and Ecotroph simulations across three ecosystems, Bourdau et al. (2016) compared the ability of five classic trophic indicators to reflect trends in ecosystem health. The question is further explored within the IndiSeas project (Coll et al. 2016), where several ecosystem models covering nine
ecosystems worldwide simulated the same fishing and environmental scenarios in order to derive generic conclusions of the abilities of ecosystem indicators to detect fishing impact.

Finally marine spatial planning initiatives are also developing under the EU commission impulse in order to manage multiple and conflicting uses of the marine space (EU 2014). Among other tools, spatial ecosystem models were put forward as good complements to classical systematic conservation planning tools. Metcalfe et al. (2015) and Reecht et al. (2015) coupled a conservation planning tool (Marxan with Zones) with CSM (resp. Ecopath with Ecosim and ISIS-Fish) for the design of a MPA network in the English Channel. The objective is to overcome the drawbacks of the classical conservation planning tools which ignore complex ecological processes and dynamic responses. Christensen et al. (2009) even replaced a conservation planning tool by using Ecospace in an optimization mode in order to optimize protected area placement according to ecological, social and economic criteria.

4.3 Fitting in institutional settings

As observed by Skern-Mauritzen et al. (2016), the progressive increase in use of complex models for advice and the rigorous evaluation of their performance probably needs to start as a bottom-up process, pushed by scientists. However it becomes evident that institutional support is necessary in order to support, organize efforts and standardize procedures. Recent changes in ICES strategic plan demonstrate the will to assist the transition to EBFM with the creation of a benchmark steering group (BSG) and of Integrated Ecosystem Assessment (IEA) groups at the regional level. Changes also need to be made in the way STECF addresses its calls and in the format of the related assessment workings groups if CSM are required.

ICES BSG has two main objectives: i) facilitate the transfer of science into advice and ii) advance the benchmark process in ICES and develop the concept of regional ecosystem benchmarks. Benchmarks are peer review process of assessment methodologies, where the data, hypotheses and results are reviewed and new proposals for improvement of the methodology are examined. BSG already oversees the benchmarking process for stock assessment models and intends to generalize good practices and quality assurance to any tool used to inform management, particularly in the scope of EBFM. While the structure is there, it is to anticipate that the benchmark process for a complex model will be an even longer and more difficult exercise than for an assessment model. The recent review of the Californian current Atlantis model proves the intensity of the task, but also its feasibility and sets the basis for a successful assessment (Kaplan and Marshall In press). In particular, they stress the need for terms of reference specific to CSM, which could be inspired from the methods proposed in Section 3. A few additional problems are however worth mentioning and solutions are still to be found (Link et al. 2012). The availability of relevant experts for these week-long workshops, already considered a problem for current single stock benchmarks, is likely to be amplified. Moreover the fastidiousness of the validation process (50% of the total time required for the entire modeling process, Aumann 2007) as well as the absence of an appropriate space to disseminate and leverage the results (neither to the scientific community nor to managers) discourage scientists to carry them out. Incentives have probably to be found to encourage modelers to involve in benchmarking and advice.

ICES promotes the development of Integrated ecosystem assessments (IEA) at regional level to coordinate international effort and develop tools for the provision of ecosystem advice (Walther and Möllmann 2013). IEA constitutes an overarching framework where the recommendations for ecosystem-based management (interdisciplinarity, stakeholder involvement, consideration of societal expectations, on top of more classical technical or scientific challenges) are made explicit and embedded in a process. Levin et al. (2009) proposed a five step loop starting with a scoping of management objectives, then going through development of ecosystem indicators and targets, risk analysis, assessment of ecosystem status relative to the objectives and management strategy evaluation. In this process, CMS have a crucial role to play, in particular to support the final MSE step so managers and citizen-stakeholders determine whether EBFM objectives have been achieved (deReynier et al. 2010). But, as demonstrated in this paper, they can also assist the choice of indicators (Sect. 4.2) and targets (Sect. 2.1) and help conduct risk analysis (Sect. 3) (Hamilton et al. 2015). Examples of successfully conducted IEAs start to appear in the US (Levin et al. 2015, 2014). Despite apparent structure and stewardship at ICES level, some raised the concern that the complex nature of IEAs and the lack of motivation for researcher to involve in the advisory process may make IEAs effective implementation difficult (Dickey-Collas 2014).

Even before full IEAs can be conducted in European fisheries, models have already been required to answer STECF calls for the evaluation of management plans for instance (STECF 2015). Typically, the evaluation meetings of STECF last about a week and it is frequent that expectations, scenarios and objectives are reviewed in course of the week. This format is not suited for complex simulation models that take several hours for a run, especially if a rigorous uncertainty analysis has to be carried out. STECF-SGMOS (2010) working group for the development of the EBFM in European seas recommended building reference ecosystem models for each reference ecosystem and to set up dedicate groups to apply and update the models thus ensuring operational readiness. While we support this proposal, we also stress the need to maintain a diversity of CSM (see Sect. 3.2.2) rather than one reference model, and this, despite obvious costs and time implications.

5 Conclusion

As Peck (2004) put it: “the world is complex and we need all the tools that we can muster to understand it”. Given complex EBFM objectives, CSMs have a key role to play within the new multidisciplinary framework that is starting to emerge in particular within IEA groups (Fig. 2). Furthermore the insights and quantitative predictions offered by CSM are necessary to support tactical, fisheries-specific management decisions.
Many complex system models have been developed in the last decades to inform ecosystem-based management (Hyder et al. 2015; Plagányi 2007; Pellezlo et al. 2012; Weijerman et al. 2015). While new developments might be suitable in certain contexts, the priority for the practical implementation of EBFM is rather to make a better use of available tools. Pragmatism was often the rule to select the models that were used to answer a specific management question (“because it is there”, “because developers are available for the meeting”, “because it is the only model for the area”). These considerations are obviously practical reasons that cannot be overlooked; however appropriateness in terms of complexity, performance, robustness and operationality, are better criteria for models selection and adaptation. The guidance provided here should help reach a compromise between availability and appropriateness: either because one or several models readily present(s) the relevant features according to the question; or because coupling or on the contrary simplifications can help reach an optimal level of complexity. We show that a wide range of methods required to adjust model complexity are available.

Additionally, the review carried out in this paper shows that successful examples of sensitivity analysis, coupling, validation and use for advice. Finally thank you to Isaac Kaplan and Marie Savina-Rolland for sharing experiences that enriched the content of the paper.

Acknowledgements. Authors are grateful to Martin Marzluff and an anonymous referee, whose thorough contributions improved a lot the quality of the manuscript. Authors would like to thank the French Association for fisheries science for the debate held during the 12th forum of the association (Montpellier, July 2015) which was a source of inspiration for this paper. Authors are also grateful to the members of the ICES working groups WGSAM and WGIPEM and to the Benchmark Steering Group for fruitful discussions on complex model validation and use for advice. Finally thank you to Isaac Kaplan and Marie Savina-Rolland for sharing experiences that enriched the content of the paper.

References


Fulton E.A., Johnson P., Gorton R., 2012, SE Fisheries and Climate Simulations: Barriers to Adaptation. 2010023 - Quantitative testing of fisheries management arrangements under climate change using Atlantis, Appendix 3, Australia, CSIRO.


Girardin R., 2015, Ecosystem and fisheries’ behaviour modelling: two crucial and interacting approaches to support Ecosystem Based Fisheries Management in the Eastern English Channel, Université Lille 1.


Mahévas S., Méhault S., Bertignac M., 2012, Consequences of change in selectivity versus the establishment Marine Protected Areas on the hake–nephrops fishery in the Bay of Biscay. Presented at the ICES Annual Science Conference.


