



A computational approach to managing coupled human–environmental systems: the POSEIDON model of ocean fisheries

Richard M. Bailey¹  · Ernesto Carrella¹ · Robert Axtell² · Matthew G. Burgess³ · Reniel B. Cabral³ · Michael Drexler⁴ · Chris Dorsett⁴ · Jens Koed Madsen¹ · Andreas Merkl⁴ · Steven Saul⁵

Received: 12 July 2017 / Accepted: 12 May 2018 / Published online: 9 June 2018
© The Author(s) 2018

Abstract

Sustainable management of complex human–environment systems, and the essential services they provide, remains a major challenge, felt from local to global scales. These systems are typically highly dynamic and hard to predict, particularly in the context of rapid environmental change, where novel sets of conditions drive coupled socio-economic–environmental responses. Faced with these challenges, our tools for policy development, while informed by the past experience, must not be unduly constrained; they must allow equally for both the fine-tuning of successful existing approaches and the generation of novel ones in unbiased ways. We study ocean fisheries as an example class of complex human–environmental systems, and present a new model (POSEIDON) and computational approach to policy design. The model includes an adaptive agent-based representation of a fishing fleet, coupled to a simplified ocean ecology model. The agents (fishing boats) do not have programmed responses based on empirical data, but respond adaptively, as a group, to their environment (including policy constraints). This conceptual model captures qualitatively a wide range of empirically observed fleet behaviour, in response to a broad set of policies. Within this framework, we define policy objectives (of arbitrary complexity) and use Bayesian optimization over multiple model runs to find policy parameters that best meet the goals. The trade-offs inherent in this approach are explored explicitly. Taking this further, optimization is used to generate novel hybrid policies. We illustrate this approach using simulated examples, in which policy prescriptions generated by our computational methods are counterintuitive and thus unlikely to be identified by conventional frameworks.

Keywords Simulation · Policy · Agent-based modelling · Optimization · Fisheries · Socio-economic · Decision-support systems

Handled by Barbara Neumann, Institute for Advanced Sustainability Studies eV, Germany.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11625-018-0579-9>) contains supplementary material, which is available to authorized users.

✉ Richard M. Bailey
richard.bailey@ouce.ox.ac.uk

¹ School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

² Department of Computational and Data Sciences, George Mason University, Fairfax, VA 22030, USA

Introduction

As the ‘great acceleration’ continues, achieving sustainable use of resources from coupled human–environmental systems (‘CHE’ systems) remains one of the key challenges facing humanity (Kotchen and Young 2007; Schlueter et al. 2012). A broad range of ‘ecosystem services’ (Daily 1997) is provided by links between ecosystems (including linked

³ Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA

⁴ Ocean Conservancy, 1300 19th Street NW, Washington, DC 20036, USA

⁵ Arizona State University, Phoenix, AZ, USA

physical and biogeochemical processes) and complex human systems with embedded social, political, economic, and cultural components (Boyd and Banzhaf 2007; Carpenter et al. 2009; Crossman et al. 2013). These large systems are intrinsically difficult to manage, in part because of competition over resource allocation ('common action problems', e.g., Harding 1968; Ostrom 1999), but also because by their nature, they are among other things multi-scaled, spatially heterogeneous, time-varying, highly path-dependent, adaptive, and affected by both internal and external influences (An 2012; Liu et al. 2007). Predicting the effect of any intervention in the face of such complexity presents difficulties, and it is the associated inherent uncertainty that drives the current need for innovative models and novel analytical approaches.

Ocean fisheries illustrate aspects of many CHE systems, namely, a dynamic environmental–ecological system, strong human presence, and a complex regulatory landscape (Fulton et al. 2011; Hunt et al. 2013). As a result, management of marine living resources is a well-known difficult problem, and is far from resolved (Glaser et al. 2014). In 2016 the FAO (2016) reported on the state of fisheries globally, finding 31.4% were fished at biologically unsustainable rates. Christensen et al. (2014) described a 65% global decline in predator fish biomass over the preceding 40 years, interpreted as due to overfishing. While some (mostly developed-world) fisheries are sustainably managed or in strong recovery, defining paths to recovery for the remainder, while maintaining acceptable levels of service provision over a spectrum of developed–developing-world contexts, is a key challenge (Costello et al. 2016). In the present work, we focus on marine fisheries as an example CHE system, and describe both a new model (POSEIDON) and a new computational approach to goal-orientated policy generation. However, our approach is generalizable to other types of CHE systems.

Goal-orientated approaches—where policies are explicitly mapped on to specified desired outcomes (optimal sustainable harvest, biological conservation, social equity, for example), informed by expert knowledge and by simulation of sets of candidate policies—have proved effective in many contexts (e.g., Marine Spatial Planning; see also Cabral et al. 2016a). Indeed, part of the solution for improved sustainable management of natural resources is better implementation of known successful policies (e.g., Costello et al. 2016). However, tailoring existing approaches is not necessarily feasible when, for example, novel combinations of conditions occur (e.g., through anthropogenically induced combinations of environmental stressors), or because lack of management capacity precludes application of whole classes of otherwise successful management approaches.

In any governance context, there exists an array of individual factors that can in principle be varied by regulators to

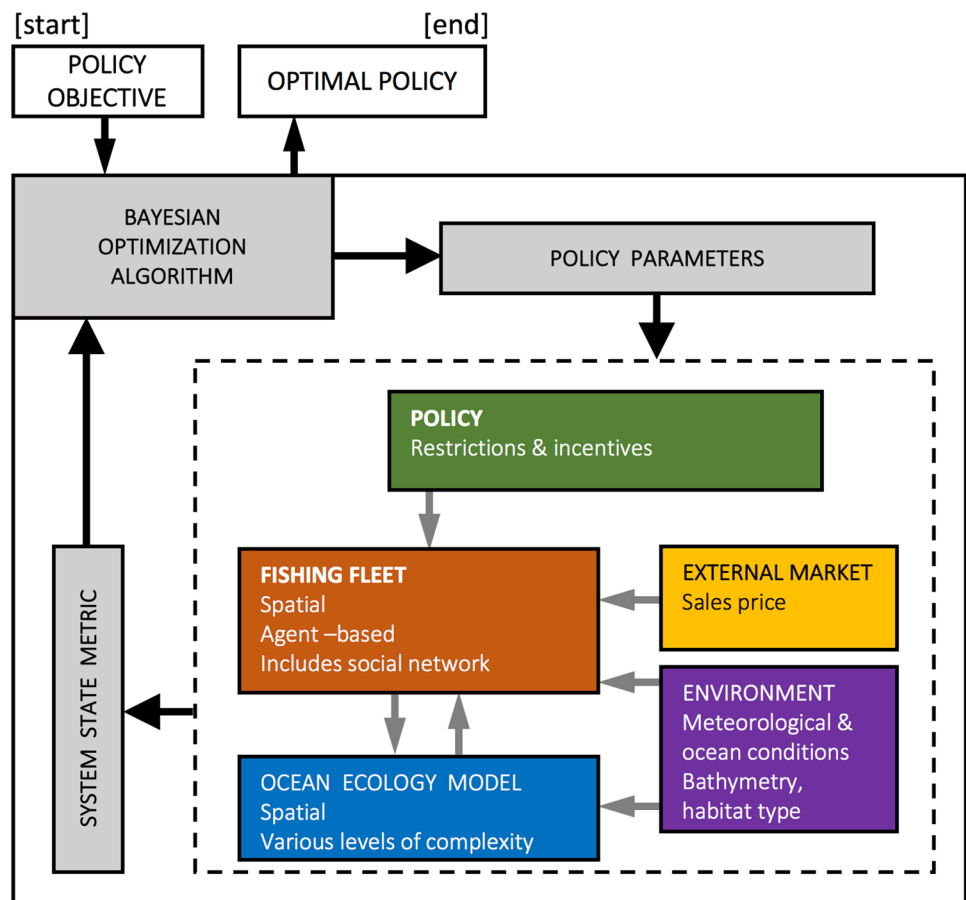
attempt to influence behaviour, either directly (e.g., exclusion) or indirectly (e.g., market-based incentives or technological constraints). Where there are relatively few factors, systematic exploration of possible combinations of measures may be feasible using simulations. In fisheries, this type of simulation approach is referred to as 'management strategy evaluation' (see Punt et al. 2016, for review). Alternatively, the system response to individual policies may be estimated using controlled experiments (including adoption of adaptive management in real time; Walters 1986) or analysis of historical empirical data (Porch et al. 2007). However, extensive experimentation in real human–environmental systems is seldom, if ever, achievable. In part this is due to the overwhelmingly large number of experimental factors involved ('the curse of dimensionality'), both in policy choice and in the number of states in which the system may exist. There are also obvious practical and ethical problems associated with experimentation on real human–environmental systems. This lack of opportunity for either empirical experimentation or 'brute-force' theoretical assessment of 'all' possible policy outcomes, means a different approach is required for screening policy choices. In this paper, we introduce the POSEIDON model for ocean fisheries (see Fig. 1), and use it to explore policy choice and novel methods of policy generation. We explore the behaviour of the model, and its capabilities in generating appropriate behavioural responses, at a conceptual level. We simplify as many aspects of the model as possible, to remove extraneous influences and provide a focus on the core behaviour of the model in the absence of additional complication.

Computation and optimization as a solution

We present here a computational approach to policy development. In essence, the process starts with a decision regarding the desired system state/outcome—the management objective; following this, an automated computational process is initiated which uses simulations of the CHE system to find policies that most closely achieve the desired outcome. This process is in a sense the reverse of policy evaluation strategies, where simulations are used to rank the outcomes of pre-defined policies. In later sections of this paper we use this approach to both optimize existing policies and to generate new hybrid policies.

If policies can be encoded as parameter sets, such that varying the parameters varies the policies, then automated numerical optimization is in principle possible. Multiple policy variants give rise to multiple parameter sets, which can be considered together as a large combinatorial parameter space, representing a set of possible policies. Any given policy or set of policies is then represented (defined) by a point or region in this space and in principle a CHE system model can be used to evaluate their success in relation to the

Fig. 1 POSEIDON model structure, including optimization routine. Elements within the dashed line are the core POSEIDON model (with titles of each module and brief descriptive key words). Outside of this is the optimization routine, which iteratively adjusts policy parameters, based on system state, to achieve specified policy objectives. In the present conceptual version of the model, the environment and market modules were not activated, and both environmental conditions and sales prices were kept constant throughout all simulations. Further details are provided in the ESM



policy objective(s). This allows a goal-orientated approach to policy development in which the starting point is the definition of the desired outcome (the emergent system state), defined by a ‘scoring function’ (SF); the optimization process then efficiently searches the parameter space to identify parameter combinations (policies) that best achieve the desired outcome, maximizing the value of the SF, subject to constraints. The emphasis in this process is, therefore, on defining the desired state for the system (and relevant metrics), and removes the need to attempt any design of policies to achieve the identified goals.

This approach requires sophisticated models to predict policy outcomes, which include the adaptive ‘counter-measures’ potentially employed by agents within the system in response to changing policies. This requirement amplifies the dimensionality problem, as regulations are now input parameters of a complicated, non-linear computational model. To be a feasible proposition, the search over policy parameters that maximizes the SF must be a highly efficient routine (Lee et al. 2015; see discussion of the ‘QBME’ method, Stonedahl and Wilensky 2010). An inherent trade-off emerges between the needs of optimization (which works better with fewer parameters and faster models) and realism (which involves more parameters and slower execution

time). This trade-off can be tamed somewhat by efficient optimization. Below we describe the use of a Bayesian meta-model to guide the optimization process, and also exploit the parallelizable nature of agent-based models. While the optimization approach is promising, it is not without possible drawbacks and the potential for ‘brittleness’ is one such weakness, and is discussed below.

The need for simulating adaptive responses

The set of subcomponents necessary for adequately realistic simulation of human–environmental systems, and the balance between parsimonious representation of necessary components and the need for realism and adaptability, is specific to each application, but we mention here what we consider to be relevant principles.

Fundamentally, the model must be driven by mechanistic processes to perform well under novel conditions. It must have sufficient granularity to capture relevant heterogeneity, both spatially and temporally, in all model components. This applies equally to the human components. For example, representing human agents as aggregated, homogeneous, rational, optimal social-economic entities is a poor assumption in many contexts (Conlisk 1996). Models of

social behaviour in natural environments should capture the known heterogeneity of the actors (their motives, preferences, etc.), the effects of their interactions through social networks (e.g., Barnes et al. 2016), and some characteristics of their adaptive behaviour. After all, those who are being regulated typically adapt in some way to new policies, and in potentially unexpected and undesirable ways (Kydland and Prescott 1977). This triggers the need to first be able to predict actors' behaviour and then to use this ability to fine-tune regulation parameters (as part of the optimization process). The problem of trying to optimally manage adaptive agents is well-suited to computational approaches, because many of the conventional tools of control theory break down in such settings (Kydland and Prescott 1977).

Agent-based modelling approaches are uniquely suited in this regard, as the actions and behaviours of individual actors are explicitly represented (An 2012; Filatova et al. 2013). There is a risk when modelling humans of 'hard-wiring' behaviour and adaptation paths in to the model (Railsback and Grimm 2011). This can happen directly ('if x then y ', for example, informed by simplified decision trees derived from social surveys) or indirectly (e.g., by focusing only on certain behavioural parameters in a dynamic optimization). Hard-wiring generates fragility, in the sense that when the model is applied to new situations (that the modeller ignored or could not predict), the quality of the results is compromised (Grimm and Railsback 2012). One approach (which is adopted in our fisheries model described below) is instead to build agents with generic exploration routines that automatically adapt to changes in their environment (Berry et al. 2002; Tesfatsion 2003). That is, agents make choices with the same algorithm in all contexts (e.g., any imposed policy), and this same process incorporates the different incentives generated by each rule to produce different final behaviour. This generic approach can be made specific using relevant empirical data, either directly through model fitting or indirectly through interviews and focused group discussions, as a way to tune the adaptation hyperparameters to local behaviour (subsuming hard to quantify factors like appetite for risk).

Methods

The model

In this paper, we introduce concepts and methods for computationally augmented policy development, using fisheries as an example. To avoid introducing the added complexities of parameterizing, simulating and interpreting models of specific real-world fisheries, the model presented here is a conceptual version of the more general POSEIDON model (Fig. 1). It nonetheless includes many of the components

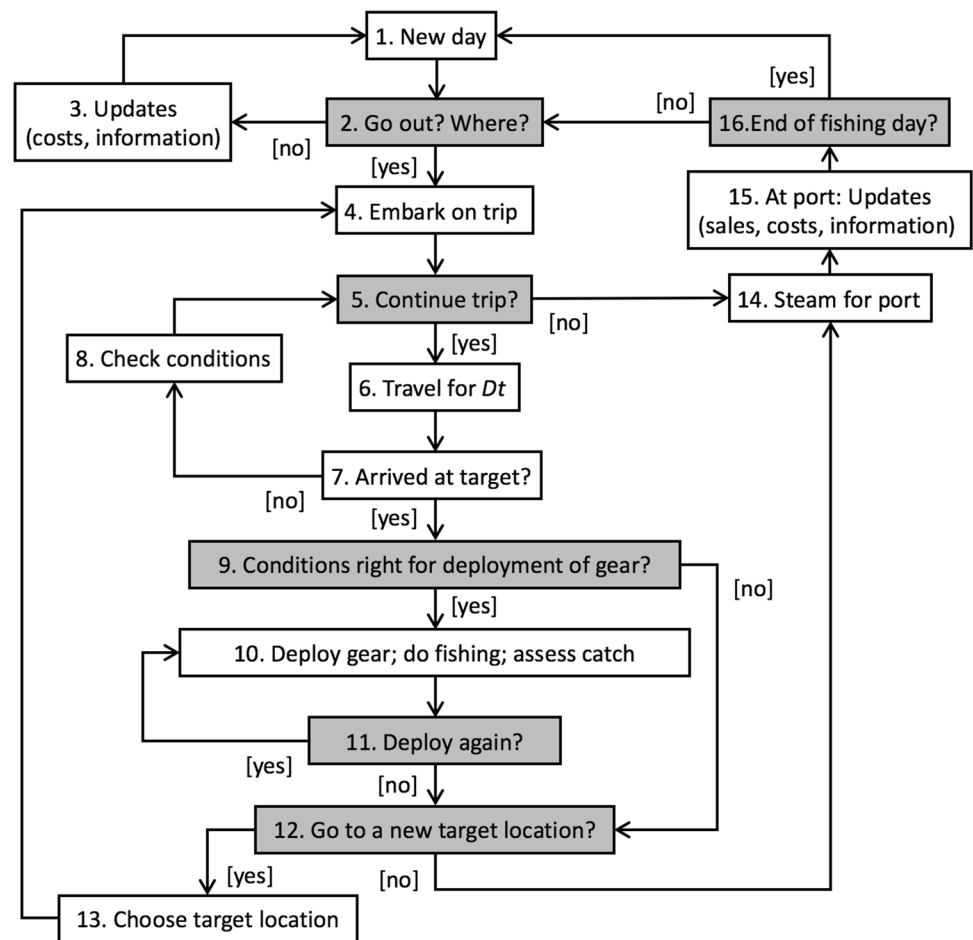
necessary for modelling real-world contexts, and real-world site-specific applications will be the focus of subsequent work.

The domain of the model is a near-shore capture fishery, represented spatially by a shoreline (which houses a single port), and an ocean. The ocean contains spatially distributed fish biomass, and a fishing fleet that can traverse the ocean catching fish. The framework includes a rudimentary representation of internal markets (for tradeable quotas—see below) plus external market signalling through pricing (in the present conceptual version external pricing remains fixed). Policies can be imposed on the fleet using various restrictions and financial incentives. A full technical description of the model is given in the Electronic Supplementary Material (ESM). The ocean component is modular and all ocean module options are spatially explicit with biomass that responds to (is depleted by) fishing pressure. In the simplest case we have small numbers of non-interacting fish species, with population density that grows locally (per spatial cell) according to a simple logistic model, and diffuses spatially according to the local gradient (following Soulié and Thébaud 2006; Cabral et al. 2010). A significantly more sophisticated option for the biology is the OSMOSE model (Shin and Cury 2001, 2004; Grüss et al. 2015) a computational model of fish dynamics simulated at school level. All options are described fully in the ESM, and in examples shown in the main paper the logistic model is used throughout.

While considerable effort has already been expended in developing models of marine ecological systems [e.g., OSMOSE (ibid.), ATLANTIS (Fulton 2010; Fulton et al. 2004), ECOSIM (Christensen and Walters 2004)], the human components of ocean system models have received less attention, and for this reason much of our focus has been on fleet behaviour (Fulton et al. 2011; van Putten et al. 2012). Rather than treating the fleet as a homogenous population, we simulate individual boats (including crew) as autonomous agents. Behaviour is determined by the multiple decisions the agents make each day, such as whether to go fishing, where to fish, which gear to use (Fig. 2). These decisions must be made without precise and full knowledge of the simulated world, but rather on what can be gathered from agent's past experience and from members of their social network (see also Little et al. 2004; Little and McDonald 2007 for discussion of foraging strategies and influence of social networks). The value of this information then decays relatively rapidly, as conditions change (e.g., biomass distribution, market conditions, profit opportunities), and as resources are simultaneously exploited by competitor agents.

We model fishers' decisions as so-called 'bandit problems' (Katehakis and Veinott 1987), in which the fishers' goal is to allocate resources (e.g., time spent fishing) among competing options (e.g., fishing locations), with information

Fig. 2 Daily routine of the fishing agents. Decision points are shown as grey boxes



that is initially limited but which increases as subsequent choices are made. The ‘multi-armed bandit’ problem provides a framework for studying the *exploration–exploitation* trade-off faced when repeatedly choosing among a finite set of options (i.e., the relative benefits of continuing to *exploit* an existing choice versus *exploring* another option) (Bubeck and Cesa-Bianchi 2012; Kuleshov and Precup 2014) (pseudo-code is available; see ESM §1.1). Within our implementation of this scheme, agents initially make a random choice from the available options (in any given context), evaluate the outcome of this choice and compare this to the success of others in their social network. In subsequent choices, and with some probability, they choose to continue with their present choice, copy the choice of a more successful agent, or randomly explore another option. We refer to this as explore–exploit–imitate (EEI). As this continues over time, and in parallel across the fleet, it generates an evolved group response that is successful in terms of whatever the agents’ value (which, in this relatively simple abstract model, is profit, but could be defined with arbitrary levels of sophistication). The ‘rules’ of the policies, therefore, do not change the behaviour algorithms, but do affect the agent (fisher) behaviour indirectly both at individual

and group levels, through incentives in relation to profit. That is, choices that result in greater profit (or other benefit) will be reinforced and adopted more widely. A benefit of this approach is that fishers’ choice behaviour is dynamic, and general enough that it works regardless of the biological model it exploits, or the policies it operates under. This results in a hallmark of agent-based models, which is the emergence of high-level dynamic patterns from low-level rules/incentives.

In this conceptual version of the model, environmental conditions (e.g., conditions that affect fishery productivity, such as ocean temperature) are unchanging and have no effect on agent behaviour or ecology. The nature of the imposed policies is described below, and as a simplification, we enforce full compliance in all agents in the present version. However, adding dynamic environmental conditions or options for fisher compliance with regulations are natural extensions of the model, planned for subsequent work.

Optimization methods

Simulated policies are defined here by parameter sets, such that varying the policy parameter values changes the

policies. A simple example would be a seasonal fishery closure policy, defined by two numbers (t_1, t_2) representing the start and end day of the annual season, operating over some period of years. It is then in principle possible to search the full parameter space (combinations of t_1, t_2), evaluating model performance for each combination against some desired outcome at the system level, calculated using a relevant scoring function. The goal might be to maximize total catch, c , by varying both t_1 and t_2 , and the model would be used to provide values of c , over the two dimensions t_1, t_2 . The parameter values (t_1, t_2) that elicit the “best” response (largest value of c) could then in principle be found. In this sense we can treat the model as a ‘black-box’ function, where the input is the policy parameter set and the output is a score based on the simulation outcome; finding the “best” policy is a function maximization problem. In the present case we use Bayesian optimization (Shahriari et al. 2016) to achieve this outcome. Bayesian optimization works by creating a meta-model of the simulation outcomes, iteratively simulating new policies and using the outcomes to update the meta-model. In computational models (including agent-based models) the search for optimality is tied to the question of the number of simulation runs necessary for a given level of confidence that the output found is the best outcome (the global optimum), versus a relatively good outcome (a local optimum). The advantage of Bayesian optimization is that it answers both questions at once, and with great efficiency. The posterior distribution (over all dimensions) generated by the Bayesian optimizer provides not only the average expected value of running a new simulation under any parameter combination, but also its uncertainty. The precise formulation of the scoring function is of course important in this process. In the above example of a seasonal fishery closure, whether the catch (c) being maximized was defined as that at the end of some interval or as the sum over that interval, may have a significant effect on the optimized values of t_1 and t_2 .

There are secondary benefits of tying together optimisers and computer simulations. First is the ability to calibrate the model. To achieve this, the scoring function is changed to represent some distance of the model output from empirical data (Hartig et al. 2011; Grazzini and Richiardi 2015). The optimizer then proceeds to tune model parameters to achieve the best fit (with the usual caveats of calibration, in particular the need to cross validate the results to have a fair estimate of the fit quality). A second benefit is the ability to assess the ‘brittleness’ of solutions, which is a response to a core difficulty: computational models are somewhat opaque (source code for many models is long and difficult to read, for non-experts particularly), parameters interact non-linearly and the input space is so large that most of it lays unexplored. In developing and testing models there is a risk associated with focusing on model configurations and parameter choices that

produce expected behaviour. The more the realistic behaviour of a model depends on fine-tuning of parameters, the greater its ‘brittleness’. A model system that is brittle tends to operate acceptably within a relatively small volume of the control parameter space, but performance degrades sharply otherwise (Bush et al. 1999). Fine-tuning of parameters may lead to improved performance, but risks increasing the brittleness of the modelled behaviour. A trade-off potentially exists between finely tuned high performance (e.g., better model/data comparison, highly sensitive to parameter choice) and robust poorer performance (with lower sensitivity to change in parameters). Testing for robustness both of the model behaviour, and of the parameters defining the policies generated by our approach, is, therefore, essential. Here, we use the Automated Non-linear Tests (ANTs) approach (Miller 1998), in which we objectively define the degree to which parameters must be changed to force the model in to ‘incorrect’ behaviour. We do this by devising a scoring function to favour unwanted/unrealistic model behaviour and searching within given bounds (using Bayesian optimization) for parameters that generate the unwanted behaviour. Any investigated model behaviour can then be described as robust under the identified parameter variation if it cannot be ‘denatured’ through this approach. This is conceptually allied to the ‘pattern-orientated’ approach of Jakoby et al. (2014), in that both approaches guard against inclusion of unrealistic parameter values, and those parameter sets that lead to unwanted behaviour. The ANTs approach does have limitations and these are outlined in ESM §2.1.3, along with ANTs results for each of the model behaviours shown.

Quantifying trade-offs

Implicit in the discussion of optimization and scoring functions so far, has been the notion that there exists a single outcome targeted by the policy and that this was the sole focus of the optimization, or that if a combination of outcomes existed then an appropriately weighted summation of these outcomes had been defined. While for many such problems it is possible to combine and explicitly weight multiple variables, it is not necessarily desirable to define a priori a single numerical definition of policy success. An alternative is to assign multiple objectives and present a set of policies that represents the best possible trade-offs between them. This is multi-objective optimisation (Deb 2001; Luke 2009) the output of which is a Pareto front: the set of efficient (highest scoring) choices over the full range of policy variable combinations (see Jacobsen et al. 2017, for a fisheries example). One way to understand the Pareto front is to consider it as the ‘budget constraint’ for the policy maker, in that it expresses what must be given up in one objective to improve another. We produce Pareto fronts in the present work using the NSGA-II

algorithm (Deb et al. 2002; see ESM §4.3). Examples of simulated trade-offs between ‘competing’ policy objectives are discussed below.

Approach to model assessment

As the model presented here is a conceptual model, full validation through quantitative comparison to empirical data is not a relevant means of assessment. While a range of biological representations can be used within the POSEIDON model (described above), for the present simulations, the highly simplified ‘logistic growth with diffusion’ model was used. This provides a sufficiently responsive spatially distributed stock for the present model experiments, but is not adequately realistic for data/model comparisons (an externally validated model such as OSMOSE (ESM §2.5) would be used in this case). The stronger focus in the present work on fleet dynamics does, however, require evidence that our formulation of vessel behaviour, and in particular the simple bandit algorithms that determine individual agent choices, are an adequate way to model fishing fleets in broad terms. Evidence is found in the nature of emergent behaviours generated in response to: (1) imposed policies, including marine protected areas (MPA’s), seasonal fishery closures, fishing gear (technology) regulations, and the use of tradeable and non-tradeable catch quotas and (2) non-policy-related

factors, including fuel and fish sale prices—the fleet should respond in qualitatively realistic ways under such changes.

Optimization experiments

To explore the use of our policy optimization methods in both fine-tuning individual policies and in generating novel policy combinations, we present three cases: (1) the setting of tradeable and non-tradeable quotas in the context of geographically mixed and geographically separated target/bycatch species; (2) the optimal placement of an MPA when trading-off catch against conservation in a mixed (two boat types) fishery; and (3) the use of the optimizer to generate novel (hybrid) policies in the context of a trade-off between catch and conservation.

To improve readability in the following section, we present sequentially for each experiment, the background, methods and results, after first presenting the model assessment results.

Experiments and results

Model assessment results

Model assessment results are summarised in Table 1 (with references to relevant ESM sections for full details). We focus first on the fleet response to two relatively direct

Table 1 Summary list of model behaviour in response to simulated policy and changes in boundary conditions

Conditions	Results	Comments
Marine protected area [ESM §2.4]	Fishing the line	Murawski et al. (2005), Parnell et al. (2007), Kellner et al. (2007) and Cabral et al. (2016b)
Gear restrictions on habitat [ESM §2.4]	Fishing the habitat line	Hannah (2003) and Bellman et al. (2005)
Two co-located species with different biomasses [ESM §2.7]	Target switching until biomasses equalize	Sutton and Ditton (2004), Béné and Tewfik (2001) and Katsukawa and Matsuda (2003)
Fuel price increase [ESM §2.3 and §2.8]	Fishing closer to port and/or switch to gear with better efficiency	Poos et al. (2013), Abernethy et al. (2010) and Basurko et al. 2013
ITQ versus TAC: heterogeneous catchability [ESM §3.1]	Positive correlation between catchability and total catches for both policies	A ‘common sense’ finding
ITQ versus TAC: heterogeneous efficiency [ESM §3.1]	Positive correlation between efficiency and catches for ITQ but not TAC	A ‘common sense’ finding
ITQ–TAC–open access: fish prices negatively proportional to landings [ESM §3.2]	ITQ solves race to fish and maximizes profits	Birkenbach et al. (2017)
ITQ versus TAC: choke species geographically separated [ESM §3.3]	ITQ incentivizes fishing away from choke species; TAC does not	Miller and Deacon (2017)
ITQ versus TAC: choke species co-located with target [ESM §3.4]	ITQ incentivizes selection of gear that avoids bycatch; TAC incentivizes higher catchability overall	Miller and Deacon (2017)

The first column provides the simulation context and reference to the section in the ESM, where full details can be found. ‘Results’ summarises the model behaviour, and ‘Comment’ provides the basis on which the model behaviour is considered reasonable

effects: changes in the total biomass and distribution of fish species, and changes in fuel price. To summarise these results, agents react to local biomass depletion by moving to fish in areas of the ocean with higher biomass, and when biomass fluctuates, agents naturally target more abundant species, all without explicit knowledge of stock levels and locations. With regard to fuel price, agents incur a financial cost in buying fuel, and respond to changes in fuel prices by changing fishing location, fishing at distances from port that better balance the trade-off between fuel costs and catch (for any given distribution of biomass in the ocean). Furthermore, when we allow for adoption of more fuel efficient gear, higher fuel prices drive faster uptake of this gear by the agents. While both sets of responses are relatively easy (for knowledgeable humans) to predict, and would be expected of any real fleet, we note that none of these behaviours are programmed into the agents. For example, agents have no built-in concept of distance, and all related changes in fishing location emerge as a consequence of their decision-making process. Other responses, to policies, are perhaps less obvious a priori, but nonetheless expected in hindsight.

Following imposition of a (no-take) marine protected area (MPA), which agents can traverse but not fish, agents react by ‘fishing the line’ (McClanahan and Kaunda-Arara 1996; Kellner et al. 2007) to benefit from ‘spill-over’ effects of fish leaving the protected area, in spite of the concepts ‘border’ and ‘line’ not being part of their programming. We can impose global quotas (total allowable catch—TAC—per species per season over the whole fishery), in which the season for a species is closed once its TAC is reached. Here, the agents naturally learn as a group to race (competitively) to catch the full quota as quickly as possible. This is a classic commons problem, as there is no incentive not to fish as effectively as possible. Implementation of a simulated individual tradeable quota (ITQ) system resolves this “race to fish”, and results in greater overall earnings (as explained by Branch et al. 2006). Furthermore, mismatches between quotas and availability incentivise agents to target species whose quota is cheaper either by changing to more species-selective gear or by changing fishing location if the species occupy different regions of the ocean. Where we introduce variation in the efficiency of gear or fuel use, the more efficient agents prosper further under the ITQ system by buying quota from less efficient agents, therefore, eventually leading to consolidation of the fleet (if we allow boats to exit the fishery once earnings drop below some prescribed level; also shown in the model of Little et al. 2009). We assessed the robustness of these seemingly realistic behaviours using ANTs (described above), and found all behaviours were persistent under at least $\pm 20\%$ variation in relevant parameter values.

In summary, relatively simple (but adaptive) agents used in POSEIDON are able to reproduce a broad range of fleet

behaviours observed in real fisheries. For example, without being explicitly coded in to the model, under open access a race-to-fish emerges endogenously from the simulated fleet, which then responds appropriately to policies aimed at solving this problem (ESM §3.3). Similarly, agents learn to avoid bycatch through location choice (ESM §3.4) and gear selection (ESM §3.5), in the presence of complex policy scenarios. This flexibility suggests great potential for POSEIDON to answer a wide range of policy-related questions. The following sections provide use-case examples of policy optimization, both for fine-tuning prescribed policies, and for generating policy hybrids.

Experiment 1: tradeable versus non-tradable quotas

As described above, a well-used output control in fishery management is the imposition of fishery-wide quotas of total allowable catches (TAC, in units of mass), enforced at species (or species group) level. Here, the fishery season remains open for that species until total catches across the fleet reach the TAC. An unintended consequence of this approach is the ‘race to fish’ it incentivizes between fishers who aim to maximize, competitively, their individual catch (hence profit) (see review by Branch et al. 2006). By distributing amongst the fishers, before the season starts, permits/quotas to catch a given amount of fish within the season, and then allowing them to trade these quotas (an individual tradeable quota [ITQ] system, see Costello et al. 2008), the race-to-fish can be eliminated, as the incentives to race no longer exist. According to Costello et al. (2008), 121 fisheries worldwide were using this approach by 2003, distributing the TAC amongst some proportion of the extant fishers. We describe two experiments to explore the use of TACs (Expt. 1a) and ITQs (Expt. 1b) using the present model.

Experiment 1a (TACs): methods

As in model runs described above, we use two versions of a highly simplified simulated world to demonstrate optimization of both TACs and ITQs. Two moderately mobile species of fish inhabit a rectangular ocean: in the first case, species ‘A’ (‘red fish’) live in the upper half, species ‘B’ (‘blue fish’) live in the lower half (both can be caught and sold), and populations regrow logistically (with a rate constant of 0.7/year; ESM §2.1.2); in the second case, the populations of red and blue fish are well-geographically mixed, occurring homogeneously in all parts of the ocean in an arbitrary proportion (the conclusions we present are not sensitive to this ratio). The chosen policy goal, through adjusting quota levels, is to maximize the total red fish catch over 20 years ($\sum_{t=1}^{20} C_t^A$) while preserving the maximum blue fish biomass ($M_{t=20}^B$) (here, index t is model year, C is total annual catch, and superscript A and B refer to species). The scoring function

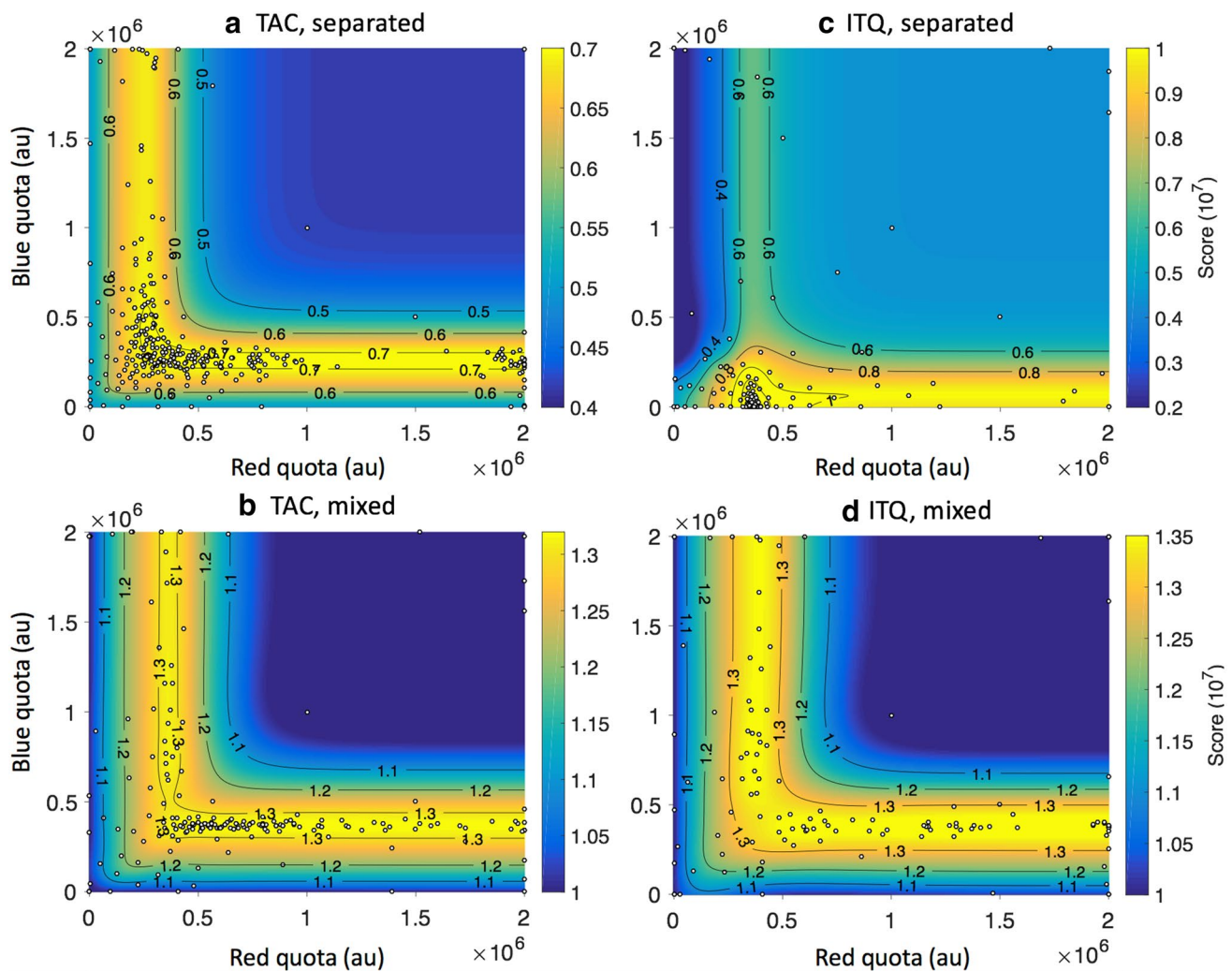


Fig. 3 Results from the Bayesian optimization process for TAC and ITQ quota allocation. Each black dot ($n=200$) represents the outcome of a 20-year model run with unique Q_{red} and Q_{blue} combinations (red/blue fish quota values), from which scores are calculated. Plots **a, b** relate to the geographically separated north/south red/blue fish

modelled world, while **c, d** are for geographically mixed populations. Plots **a, c** show the estimate of the scoring function (over the red/blue quota space) for the TAC policy; plots **b, d** show equivalent data for the ITQ simulation. Colours represent estimated policy scores (S), with values shown in the side colour bars

S (necessary for the optimization procedure), is defined $S = M_{t=20}^B + \sum_{t=1}^{20} C_t^A$ (also see caption to Fig. 3). For both the ITQ and the TAC there are two parameters to set: annual quotas for species A and species B, Q_{red} and Q_{blue} , respectively (which in the case of the ITQ are distributed equally amongst the $N=100$ agents; we did not investigate how variation in the initial allocation of quota impacts our results, and this will be explored in future work).

Experiment 1a (TACs): results

Figure 3 shows the estimated score distributions generated during the optimization of quotas for TAC (Fig. 3a, b) and ITQ (Fig. 3c, d) policies. The position of these points is decided iteratively by the Bayesian Optimizer. As each result

(value of S) is calculated, following each successive run of the model, the optimizer updates the estimate of the distribution of scores (and associated uncertainty) across the $[Q_{\text{red}}, Q_{\text{blue}}]$ parameter space, and provides the ‘coordinates’ for the next model run. The contours plotted in Fig. 3 represent this meta-model of the score distribution at the end of 200 simulation runs (equivalent plots of the uncertainty in the estimated value of the meta-model are shown in ESM §4.2).

For the ‘separated’ and ‘mixed’ cases (Fig. 3a, b), the scores produce an L-shaped pattern. For ‘separated’ species, with the two limbs intersect at $Q_{\text{red}} \approx Q_{\text{blue}} \approx 2.6e5$, while for ‘mixed’ the intersection happens at $Q_{\text{red}} \approx Q_{\text{blue}} \approx 3.75e5$ (arbitrary units). For the agents, whose goal is to maximize profit, the fleet-wide TACs provide no individual incentive to target either species, since both can be sold for profit, and

agents tend to catch whatever is close to port (to reduce fuel costs). This results in well-mixed yearly landings. Since the policy score is a simple (non-weighted) sum, catching reds contributes to the score by the same amount as preserving blues; hence, the L-shaped results are shown in Fig. 3a, b. That is, to avoid a low policy score by year 20, total catch must be constrained (to preserve blue and ensure red fish are not caught at a rate that depletes the stock too rapidly) but not constrained to the extent that catches are too heavily limited. This is achieved by constraining either the red or the blue quota, and once this constraint is present for one species, the results (policy score) are insensitive to the quota for the other species so long as that quota is large enough avoid further constraining the catch. This explanation applies to both the ‘separated’ and the ‘mixed’ cases, with the optimizer ‘discovering’ what is effectively the maximum profit yield for red species in both cases (over the 20 year period), which is larger for the ‘mixed’ case as there is greater total biomass.

Experiment 1b (ITQs): methods

We now turn to the ITQ under otherwise identical model conditions (‘mixed’ and ‘separated’). Quotas for each species are allocated to each agent at the start of each simulated season. Throughout the season, agents participate in an open auction market for quotas at the end of each day. Each agent needs sufficient quotas to cover their catch and based on their earnings needs and catch rate, they decide whether to buy/sell quotas from/to other fishers (see ESM §3.1 for details of quota valuation and trading). If they catch more than the quotas they have, they sell only the portion of the catch covered by the quota (the remaining fish are discarded) and they are disqualified from fishing until the end of the season (when the quota allowance refreshes—see ESM §3 for full description). In this experiment the boat holds are relatively small and the effect of discards is insignificant.

Experiment 1b (ITQs): results

In this experiment, the optimizer returns a TAC for each species (Q_{red} and Q_{blue}) which is then distributed equally amongst the fishers. Results are shown in Fig. 3c, d: in the case of mixed populations, optimal quotas are $Q_{\text{red}} \approx Q_{\text{blue}} \approx 3.75e5$; for geographically separated populations, the results are $Q_{\text{red}} \approx 3.75e5$ and $Q_{\text{blue}} \approx 0$.

Comparing Fig. 3a and c shows, for the case of geographically separated fish populations, the optimal quota allocation under TAC and ITQ to be significantly different. The optimal blue quota under the ITQ is found to be zero. Agents operating under the simulated ITQ are in general incentivized to avoid species whose quotas are rare, and in this is an extreme instance of this tendency. Agents who by

experimentation land in the south of the map and catch blue fish are unable to fish again for the entire season and those within their social network learn not to imitate them. The Bayesian optimizer exploits this ability of the fleet to adapt, and converges, quite logically, to a scheme, where red quota is at the discovered MSY and blue fish are conserved by setting their quota to zero. To demonstrate this difference is caused by agents’ reaction to policy, we run the same optimizations on the well-geographically mixed distribution of red and blue fish (see Fig. 1c, d). Agents in this case are unable to modulate the ratio of blue to red catches by fishing location choice, and the optimal ITQ and TAC quotas are largely indistinguishable (yielding the familiar L-shaped optimum, since blue/red landings are always correlated). A somewhat counterintuitive result from our model is that the optimal total quota allocated under an ITQ is not necessarily equal to the optimal (fishery-wide) TAC. In the case outlined above, this is because ITQs, unlike TACs, incentivise changes in fishing location choices (further discussion in ESM).

Experiment 2: optimizing marine protected area (MPA) placement

Experiment 2 (MPA placement): methods

In this experiment, we investigate the optimal placement of an MPA in a simulated world, where there is only one species of fish (with biomass density that increases linearly with distance from shore), and two types of fishing agents [real-world decisions, e.g., Watts et al. (2009), rest on larger bodies of information, but we maintain relative simplicity here, to ease interpretation]. In our simulation, the first agent type has large boats with large travel range, large holds and efficient gear (high probability of catching fish per unit effort); the second has small boats, with small holds, inefficient gear and limited travel range (see ESM §4 for details). The policy goal is to protect the catch of small-scale fishers, in the face of competition from the larger boats while maximizing the total fishery catch. The total catch, which is most likely to be dominated by the larger boats, is expressed in terms of summed catches from all boats over a 20-year period. The policy choice is the size and location of a limited-entry MPA in which small-scale fishers can fish but larger boats cannot (traversing the MPA is allowed for all). This situation is akin to a developing country setting (e.g., Philippines), where commercial vessels are allowed to fish only beyond a prescribed distance from shore, while small-scale fishers face no such restrictions. In Experiment 1, there was an implicit assumption that the two terms in the score function (S) were equally weighted, allowing for a straightforward optimization of each policy. However, in this second example

we wish to explore the possible trade-off between the two objectives (total catch versus small-boat catch), and we explore the associated Pareto front.

Experiment 2 (MPA placement): results

Results from Experiment 2 are shown graphically as a Pareto front in Fig. 4a, where the nature of the trade-off between the two policy objectives can be seen. Three example protected areas, generated by the Bayesian optimizer, are also shown in Fig. 4b–d. Figure 4b shows the MPA that prioritizes small-boat catches. The optimizer creates a large MPA around the port within which small boats fish, without competition from larger boats. Figure 4c shows the MPA that maximizes total catch, at the expense of small boat catch. When prioritizing total catches the optimizer no longer protects the small boats, which then must compete against the larger boats close to port. It does, however, still create an MPA further out at sea. This region of the ocean is too far from port for small boats to reach and the effect of the MPA is to prevent overfishing by the large boats. The optimizer ‘discovers’ this protected area is necessary to maximize long-term catches and prevent early depletion, and the (large boat) agents duly learn to fish the line to maximize their catch. Finally, Fig. 4c represents the numerically optimal trade-off between the extremes (the equally weighted case).

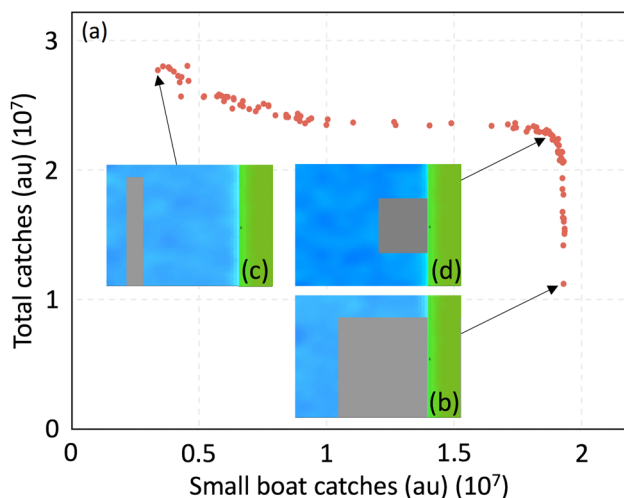


Fig. 4 Exploration of trade-offs associated with MPA placement. **a** Pareto front showing the range of outcomes due to size and placement of the MPA. A strong trade-off exists between small-scale fishers’ income and total catches. Parts **b–d** show the simulation map together with example MPAs generated by the optimizer at points shown along the front. Green areas are land and the port is located (vertically) in the centre of the land

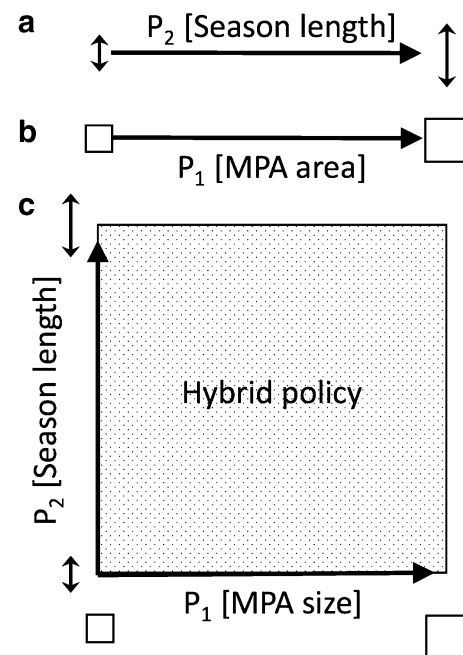


Fig. 5 Generation of hybrid policies. In this example, two simplified policy tools (Season length and MPA area) are available. Each policy is defined by a single parameter (depicted graphically in parts **(a)** and **(b)**). The set of possible combinations of policies is described by the 2D-space shown in part **(c)**, with points in this space relating to the two parameters. Scoring the outcome of these choices results in a surface, with maximum value indicating the most effective parameter choice, and this can be extended to a many dimensional volume representing aspects of many individual policies, over which optimization can be performed

Experiment 3: generating hybrid policies

A natural development from the first two examples is to move away from the need to decide a priori the particular policy to be optimized in any given context. In real world situations this choice may be constrained by practicalities, but in principle (and in the modelling context) we can make such unconstrained choices. To achieve this, we expand the ‘policyspace’ by allowing the optimizer to take control simultaneously of all policy parameters across all applicable policies (in this case, policies for MPAs, TAC and ITQ allocations, and season start- and end-dates). This provides opportunity for the optimizer to ‘create’ higher-scoring hybrid policies by combining aspects of different individual policies (see Fig. 5 for an explanatory example). We describe two examples of the use of this approach that echo the previous experiments: Experiment 3a in which target/bycatch species are geographically mixed, and Experiment 3b in which they are separated. Computationally, optimization of hybrid policies is more intensive (the search space is larger), and while it is also possible that optimal hybrid policies may

retain the simplicity of individual existing policies (i.e. hybrids are no better than the best individual policy), there is also a risk of the process creating impractical levels of complication in the highest scoring policy. Ways of avoiding these potential difficulties are also discussed below.

Experiment 3a (geographically mixed species): methods

The model configuration for Experiment 3a is identical to Experiment 1a (two species of geographically mixed fish populations, red [A] and blue [B]), and we use the same scoring function, with higher scores for greater catch of red fish and greater conservation of blue fish: $S = M_{t=20}^B + \sum_{i=1}^{20} C_{t=i}^A$. As a baseline case, we optimize a fisher-wide TAC and the optimizer returns two policy parameters ($Q_{\text{red}}, Q_{\text{blue}}$), and a score for the policy outcome (S). We do the same for other individual policies: (1) a permanent MPA (4 parameters defining MPA location); (2) a temporary MPA (4 location parameters plus 1 parameter defining the number of days the MPA is active, counting from day 1); (3) season length (1 parameter, number of days the season is open, counting from day 1); and (4) ITQ (2 parameters controlling initial quota for each species per fisher). For the generation of hybrid policies, we combine the parameters of the TAC, temporal MPA and season length policies (8 parameters in total) and allow the optimizer to search over this enlarged policy space.

Experiment 3a (geographically mixed species): results

The optimization routine yields a policy score, as defined above, for the TAC (denoted by the relevant subscript) of $S_{\text{TAC}} = 1.3e7$, which we normalized to 1 and use as a baseline for comparison to the other policies. For each of the different policy options, the optimal policy version was found (using the Bayesian optimizer); a further 100 simulations were then run, with the optimal policy imposed for the duration of the 20 year simulation, to provide a mean and standard deviation for each score. The resultant normalized scores (S) and standard deviations (σ) are: permanent MPA, $S_{\text{pMPA}} = 1.09$ ($\sigma_{\text{pMPA}} = 0.04$); temporary MPA, $S_{\text{tMPA}} = 1.48$ ($\sigma_{\text{tMPA}} = 0.008$); season length, $S_{\text{SL}} = 0.99$ ($\sigma_{\text{SL}} = 0.005$); ITQ, $S_{\text{ITQ}} = 1.03$ ($\sigma_{\text{pMPA}} = 0.006$).

For this model setup, the hybrid policy produced results that were indistinguishable from the top scoring individual policy (the temporary MPA; $S_{\text{Hybrid}} = 1.45$, $\sigma_{\text{Hybrid}} = 0.0158$). While the expectation might be that more degrees of freedom (policy parameters) would routinely produce higher scores, in this case the optimizer effectively ‘switches off’ both of the TACs (by setting red, blue TACs to $1.2e6$ and $2.0e6$, respectively, which are so large as to be non-binding) and the season length restriction (by setting season length to 366 days). The resultant policy is a relatively large MPA adjacent to the port (see ESM, §4.5.1), which is imposed for

part of the year (283 days). The optimizer (without being coded to do so, and working only to maximize the score, S) exploits the logistic growth characteristics of the individual fish populations. It sets a policy that avoids local biomass being depleted to the point, where growth rates become drastically reduced. Without any regulation, the fleet would naturally generate radial ‘fishing fronts’, first depleting cells closest to port and working outwards. Were the effort more dispersed, fewer cells would be emptied and the recruitment rates would be higher (for the same global biomass). A temporary MPA close to port is a way to achieve a more dispersed effort. When the MPA region is open, it is exploited by the agents (it is close to port and more profitable), but since the MPA is only open for a portion of the year, the effect of fishing is insufficient to cause major depletion. For the rest of the year agents ‘fish the line’ around the MPA (see ESM §2.4 for equivalent examples), which causes some local depletion but this is tempered by the days agents spend fishing within the protected area. No other policy approach is able to improve on this situation.

Experiment 3b (geographically separated species): methods

In the second hybrid optimization, we re-run the same scenario as for Experiment 3a, with the exception that red/blue fish are geographically split: red in northern half, blue in southern half of the ocean (as for Experiment 1).

Experiment 3b (geographically separated species): results

In the case of geographically separated species, the highest scoring individual policy is the ITQ, with individual blue and red quotas being $Q_{\text{red}} \approx 3.75e5$ and $Q_{\text{blue}} \approx 0$, with a score of $\sim 1e7$. As for Experiment 3a, normalized mean scores and standard deviations for 100 model runs of each optimized individual policy were calculated, here yielding: permanent MPA, $S_{\text{pMPA}} = 1.54$ ($\sigma_{\text{pMPA}} = 0.03$); temporary MPA, $S_{\text{tMPA}} = 1.54$ ($\sigma_{\text{tMPA}} = 0.03$); season length, $S_{\text{SL}} = 1.00$ ($\sigma_{\text{SL}} = 0.02$); ITQ, $S_{\text{ITQ}} = 1.66$ ($\sigma_{\text{pMPA}} = 0.02$).

Unlike Experiment 3a, a higher score can in this case be achieved using the hybrid approach. The optimizer finds a mixture of an MPA plus a restricted season to be superior to the ITQ (the normalized score is $S_{\text{Hybrid}} = 1.70$, $\sigma_{\text{pMPA}} = 0.01$). The hybrid policy includes a year-round MPA covering the entire area, where blue fish live and a short season length (118 days) to fish the remaining (red fish) areas; the quotas (TACs) are set at levels so high as to be un-restrictive ($1.1e6$ for red fish, and $1.2e6$ for blue, although the blue TAC is inactive in this particular case as blue fish are in a protected area, but is calculated as this is not necessarily the case for other hybrid policies). This

hybrid solution combines the obvious effect of protecting blue fish using an MPA and discovers that effort control through season length returns a better long-term yield than the use of quotas. The improvement of the hybrid policy over the ITQ, while statistically significant ($p < 1e-10$) is not dramatic in terms of score, but demonstrates that hybrid policies can indeed produce better outcomes than individual policies in some cases.

Discussion

The conceptual version of POSEIDON presented here is highly simplified compared to real fisheries, and the sensitivity of the observed model outcomes to the various simplifications imposed requires careful investigation before meaningful lessons can be learnt for specific fisheries. For example, increasing the heterogeneity of the agents, the initial allocation of quotas in the simulated ITQ, and the level of fidelity with which physical/environmental/economic conditions are represented, are all likely to affect the model outcomes. Nonetheless, the flexibility of POSEIDON in simulating known behavioural characteristics of fishing fleets in response to a broad range of management policies suggests considerable potential for real-world application, once additional empirically informed model complexity is included.

The use of numerical optimization to achieve policy goals, both by tuning existing policies and by generating policy hybrids, has been demonstrated in the previous sections. A necessary part of the approach advocated is the use of models that include human actors who behave adaptively under the influence of management policies. Rather than driving our simulated agents with statistical models of historical/empirically informed decision outcomes, we provided the agents with a means of making decisions. The use of an explore–exploit–imitate (EEI) algorithm represents what is perhaps the simplest approach to developing this fully adaptive group (fleet) response. While individual human behaviour is undoubtedly more complex than this algorithm, demonstration that this approach is indeed reasonable for fleet-level simulation is the appearance of higher level behaviour (emergent in modelled time and/or space) which is unexpected or hard to predict—for example, the emergence of ‘fishing the line’ around an MPA. The features listed in Table 1 provide some confidence that EEI agents collectively react realistically (i.e., produce behavioural patterns documented empirically) across a range of relevant contexts, and provide means for gaining insights in to the outcomes of real-world management decisions. For example, results of Experiment 1 show that optimal ITQ allocation is not achieved necessarily by dividing the optimal TAC amongst the fishers. This is by no means an obvious result and is due to the adaptation of the fleet to

financial constraints/opportunities and the spatial distribution of biomass.

However, while the use of EEI agents is potentially powerful, it is also not without limitations. It is not necessarily expected that the transient response of the EEI fleet would closely match that of a real fishing fleet. This is because, under the present scheme, improvements in behaviour are incremental, requiring both ‘bad’ choices by some agents, which are then avoided by others, and randomly found ‘good’ choices which are then actively reinforced. More cognitively sophisticated agents (and real fishers) would be expected, through a variety of potential mechanisms, such as forward planning, to discount predictably bad choices or risky options and, therefore, adapt their behaviour more rapidly. Such additional behavioural features would require further (empirical) calibration, which adds complication and potential biases, and these were avoided in the present conceptual work, at the potential cost of less-realistic dynamic responses.

An unsurprising feature of the optimization approach is that the outcome (the optimal policy) is sensitive to the specification of the scoring function. For example, if the management goal is to maximize whole fishery earnings over a short interval, the optimizer may find policies that lead to full collapse of the stock, if this is the best way to maximize catch and earnings over that particular time interval. Specification of the scoring function, therefore, requires some effort, and in real-world contexts, some expert knowledge. We do not see this as a drawback, but rather an appropriate allocation of human effort. Users of this approach must spend time deciding what is required from the system (the preferred system state), and how best to quantify it, rather than attempting the difficult task of deciding which specific policies might achieve that state, and which may then be tested using simulation or other means.

The ability to optimize across a range of competing objectives is a necessary inclusion within this approach, and we present tools which allow policy trade-offs to be explored. In real-world applications, decisions over trade-offs would likely entail factors such as management costs, and a variety of risks and rewards (e.g., Little et al. 2016). These have not been included in the present model, but are a natural extension, and while likely complex in nature, would be conceptually straightforward to include in relevant policy scoring functions.

Results from Experiment 3 show it is possible to use optimization to generate hybrid policies, by presenting a broad collection of policy options to the optimizer. Two interesting observations came of Experiment 3 in this regard. In the first experiment, the optimiser effectively ‘turned off’ some policy options to achieve the best score (e.g., ignoring season length by setting it to > 365 days). In doing so, it effectively

reverted to a single policy option of a temporal MPA. In the second, it produced a hybrid policy with better outcomes than the best of the single policies. The improvement of the hybrid policy over this single policy (an ITQ) was statistically significant ($p < 1e-10$), but was not dramatic in terms of score. It demonstrates, nonetheless, that improvements are possible.

While the examples described for Experiment 3 showed the optimiser turning off some policies to achieve the best score, or producing relatively simple hybrids, the optimiser could in principle create a complex mix of policy components. Such a mix might be unworkable in terms of real-world implementation, and in some contexts it may be necessary to set an arbitrary level for maximum policy complexity. A response to this might be to adjust the scoring function, such as directly “penalizing” for policy complexity or by including the associated implementation costs. Alternatively, the optimization could be run as a multi-objective optimisation, with policy complexity or implementation cost as one of the dimensions. In the case of Experiment 3b (and imagining a real-world application), the minor increase in score (as defined) may also entail a subsidiary advantage in terms of ease of implementation (i.e., it may be cheaper and easier to implement a season closure and an MPA than to run an ITQ), and this could naturally be included in the score, such that the difference between the single and hybrid options may be more stark.

The approach of using computational methods to evaluate candidate policies, or to generate hybrids of existing policies, has a number of advantages over more traditional assessments in which policy candidates are created a priori and scored using simulations. Two classes of advantages are: (1) avoiding constraining the policy type and (2) optimization in light of unforeseen (or poorly resolved) risks and opportunities.

On point (1), it is notoriously difficult to confidently predict what type of policy (typically referenced against those already in use) is likely to be successful (let alone optimal in some sense) in achieving specified goals over extended time periods. In some well-resourced fisheries, strict top-down (TD) control may be successful (e.g., enforced protected areas or observers enforcing no-discard policies); in resource-poor contexts, policies that facilitate self-organizing/-policing bottom-up (BU) responses (e.g., TURF reserves; Christy 1982; Afflerbach et al. 2014) may promote better outcomes (as compared to a sub-optimal implementation of TD controls). In the optimization scheme we describe, decisions on whether to opt for TD or BU approaches, or mixtures of both, need not be made early in the process, thereby closing off potentially (and unexpectedly) useful options. The optimization process naturally ‘experiments’ with a wide range of policies that incorporate specified TD controls and naturally include BU responses

due to the adaptive nature of the group (fleet) response. It is the score of the outcome that is the focus of the optimization, and unless factors associated with penalizing TD or BU solutions are included in the scoring function, all possibilities will be explored.

Point (2) is largely about unintended consequences, be they detrimental or serendipitous, that may operate over a variety of temporal and spatial scales. An immediate and well-studied consequence of incorporating spatial effort controls, for example, may be the displacement of effort (ABPmer 2017). Dynamic responses may be less obvious. For example, a suite of policy goals may include the long-term stability of fish catch, defined perhaps by scoring highly for reduced catch variance (or discounting for variability in catch; Mangel 2000). There is evidence in some systems that policies intended to smooth-out shorter term variation may increase the risk of pushing the system towards malign critical transitions in the longer term. Actively suppressing higher frequency variability risks losing valuable (perturbation response) information on the underlying system state, but also may condition the system towards resilience to a limited spectrum of disturbances (Carpenter et al. 2015). In other words, managing for short-term variability (which may be politically favourable) may entail higher long-term risk to both economic and biological sustainability—a somewhat counterintuitive possibility. The optimization procedures we describe could in principal accommodate such phenomena, requiring no special vigilance or even knowledge of such effects by the user. Policies that promote strong short-term stability would be down-weighted once the associated longer term risks were realised in the model results (so long as the model was run for sufficient time), and it may not be obvious why these policies were avoided. If such concerns are placed also within the context of non-analogue present/future environmental conditions, and the recognition that many complex systems become increasingly fragile as levels of stress are increased (Scheffer et al. 2009), the need to use a freely searching, process-based, coupled human–environmental model becomes even more compelling.

Conclusions

We argue there is a strong need for new approaches to generating policies in the domain of complex human–environmental systems. Models which couple environmental processes to complex (adaptive) human components are essential for exploring system-level responses to changes in policy. We have employed agent-based modelling in our ocean fisheries model (POSEIDON), which captures qualitatively a wide range of empirically observed fishing behaviour and fisheries responses. The fleet responds adaptively to novel changes in policy and/or other conditions and this frees the model

user to explore policies and behaviours beyond those which have been observed empirically in the past. As such, we can define policy objectives (of arbitrary complexity) and use Bayesian optimization (over multiple model runs) to find policies that best meet the management goals. In addition, we can generate hybrid policies which outperform single policies by running the optimization over all (individual policy) parameters simultaneously. For future work, we identify three key areas: (1) a focus on *adaptive management* strategies, in which there is optimization of management algorithms rather than the parameters of statically imposed policies; (2) incorporation of agents with greater cognitive sophistication than the explore–exploit–imitate agents currently used, to explore transient aspects of policy response; and (3) ‘concrete models’ of specific real-world fisheries, to explore the effects on behaviour and policy choice of heterogeneity in both the ecology and the fishing fleet.

Acknowledgements This work was funded by the Oxford Martin School (within the Sustainable Oceans Programme; PI Bailey) and also by Ocean Conservancy, supported by grants from Packard Foundation, Walton Family Foundation, Gordon and Betty Moore Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abernethy KE, Trebilcock P, Kebede B, Allison EH, Dulvy NK (2010) Fuelling the decline in UK fishing communities? *ICES J Mar Sci* 67(5):1076–1085
- ABPmer (2017) Displacement of fishing effort from marine protected areas. ABPmer Report No. R.2790. Commissioned Reports, Number 241. York. <http://publications.naturalengland.org.uk/publication/5674265573064704>
- Afflerbach JC, Lester SE, Dougherty DT, Poon SE (2014) A global survey of “TURF-reserves”. *Territorial Use Rights for Fisheries coupled with marine reserves*. *Glob Ecol Conserv* 2:97–106
- An L (2012) Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecol Model* 229(24):25–36
- Barnes ML, Lynham J, Kalberg K, Leung P (2016) Social networks and environmental outcomes. *Proc Natl Acad Sci* 113(23):6466–6471
- Basurko OC, Gabiña G, Uriondo Z (2013) Energy performance of fishing vessels and potential savings. *J Clean Prod* 54:30–40
- Bellman MA, Heppell SA, Goldfinger C (2005) Evaluation of a US west coast groundfish habitat conservation regulation via analysis of spatial and temporal patterns of trawl fishing effort Canadian. *J Fish Aquat Sci* 62(12):2886–2900
- Béné C, Tewfik A (2001) Fishing effort allocation and fishermen’s decision making process in a multi-species small-scale fishery: analysis of the conch and lobster fishery in Turks and Caicos Islands. *Hum Ecol* 29(2):157–186
- Berry BJ, Kiel LD, Elliott E (2002) Adaptive agents intelligence and emergent human organization: capturing complexity through agent-based modeling. *Proc Natl Acad Sci* 99(3):7187–7188
- Birkenbach AM, Kaczan DJ, Smith MD (2017) Catch shares slow the race to fish. *Nature* 544(7649):223–226
- Boyd J, Banzhaf S (2007) What are ecosystem services? The need for standardized environmental accounting units. *Ecol Econ* 63:616–626
- Branch TA, Hilborn R, Haynie AC, Fay G, Flynn L, Griffiths J, Marshall KN, Randall JK, Scheuerell JM, Ward EJ, Young M (2006) Fleet dynamics and fishermen behavior: lessons for fisheries managers Canadian. *J Fish Aquat Sci* 63:1647–1668
- Bubeck S, Cesa-Bianchi N (2012) Regret analysis of stochastic and nonstochastic multi-armed bandit problems. arXiv preprint. <http://arxiv.org/abs/1204.5721>
- Bush SF, Hershey J, Vosburgh K (1999) Brittle system analysis. arXiv preprint cs/9904016. <http://arxiv.org/abs/cs/9904016> [csNI]
- Cabral RB et al (2010) Effect of variable fishing strategy on fisheries under changing effort and pressure: an agent-based model application. *Ecol Model* 2212:362–369
- Cabral RB, Halpern BS, Costello C, Gaines SD (2016a) Unexpected management choices when accounting for uncertainty in ecosystem service tradeoff analyses. *Conserv Lett*. <https://doi.org/10.1111/conl.12303>
- Cabral RB, Gaines SD, Johnson B, Bell TW, White C (2016b) Drivers of redistribution of fishing and non-fishing effort after the implementation of a marine protected area network. *Ecol Appl*. <https://doi.org/10.1002/eap.1446>
- Carpenter SR, Mooney HA, Agard J, Capistrano D, DeFries RS, Díaz S, Dietz T, Duraipappah AK, Oteng-Yeboah A, Pereira HM, Perrings C (2009) Science for managing ecosystem services: beyond the Millennium Ecosystem Assessment. *Proc Natl Acad Sci* 106(5):1305–1312
- Carpenter RS, Brock WA, Folke C, van Nes EH, Scheffer M (2015) Allowing variance may enlarge the safe operating space for exploited ecosystems. *PNAS* 112(46):14384–14389
- Christensen V, Walters C (2004) Ecopath with ecosim: methods, capabilities and limitations. *Ecol Model* 172:109–139
- Christensen V, Coll M, Piroddi C, Steenbeek J, Buszowski J, Pauly D (2014) A century of fish biomass decline in the ocean Marine. *Ecol Prog Ser* 512:155–166
- Christy FT (1982) Territorial use rights in marine fisheries: definitions and conditions. *FAO Fisheries Technical Paper* 227
- Conlisk J (1996) Why bounded rationality? *J Econ Lit* 34(2):669–700
- Costello C, Gaines SD, Lynham J (2008) Can catch shares prevent fisheries collapse? *Science* 321(5896):1678–1681
- Costello C, Ovando D, Clavelle T, Strauss CK, Hilborn R, Melnychuk MC, Branch TA, Gaines SD, Szuwalski CS, Cabral RB, Rader DN (2016) Global fishery prospects under contrasting management regimes. *Proc Natl Acad Sci* 113:5125–5129
- Crossman ND, Burkhard B, Nedkov S, Willemsen L, Petz K, Palomo I, Drakou EG, Martín-Lopez B, McPhearson T, Boyanova K, Alkemade T, Ego B, Dunbar MB, Maes J (2013) A blueprint for mapping and modelling ecosystem services. *Ecosyst Serv* 4:4–14
- Daily G (1997) *Nature’s services: societal dependence on natural ecosystems*. Island Press, Washington
- Deb K (2001) Multi-objective optimization using evolutionary algorithms. *Wiley-Interscience series in systems and optimization*. Wiley, Chichester
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6(2):182–197. <https://doi.org/10.1109/4235.996017>
- FAO (2016) *The state of world fisheries and aquaculture*. FAO, Rome

- Filatova T, Verburg PH, Parker DC, Stannard CA (2013) Spatial agent-based models for socio-ecological systems: challenges and prospects. *Environ Model Softw* 45:1–7
- Fulton EA (2010) Approaches to end-to-end ecosystem models. *J Mar Syst* 81(1–2):171–183
- Fulton EA, Parslow JS, Smith ADM, Johnson CR (2004) Biogeochemical marine ecosystem models II: the effect of physiological detail on model performance. *Ecol Model* 173:371–406
- Fulton EA, Smith AD, Smith DC, van Putten IE (2011) Human behaviour: the key source of uncertainty in fisheries management. *Fish Fish* 12(1):2–17
- Glaser SM, Fogarty MJ, Liu H, Altman I, Hsieh CH, Kaufman L, MacCall AD, Rosenberg AA, Ye H, Sugihara G (2014) Complex dynamics may limit prediction in marine fisheries. *Fish Fish* 15(4):616–633
- Grazzini J, Richiardi M (2015) Estimation of ergodic agent-based models by simulated minimum distance. *J Econ Dyn Control* 51:148–165
- Grimm V, Railsback SF (2012) Designing formulating and communicating agent-based models. Agent-based models of geographical systems. Springer, Dordrecht, pp 361–377
- Grüss A, Schirripa MJ, Chagaris D, Drexler M, Simons J, Verley P, Shin YJ, Karnauskas M, Oliveros-Ramos R, Ainsworth CH (2015) Evaluation of the Trophic structure of the West Florida Shelf in the 2000s using the ecosystem model OSMOSE. *J Mar Syst* 144:30–47. <https://doi.org/10.1016/j.jmarsys.2014.11.004>
- Hannah RW (2003) Spatial changes in trawl fishing effort in response to footrope diameter restrictions in the US west coast bottom trawl fishery. *N Am J Fish Manag* 23(3):693–702
- Harding G (1968) The tragedy of the commons. *Science* 162(3859):1243–1248
- Hartig F, Calabrese JM, Reineking B, Wiegand T, Huth A (2011) Statistical inference for stochastic simulation models—theory and application. *Ecol Lett* 14:816–827
- Hunt LM, Sutton SG, Arlinghaus R (2013) Illustrating the critical role of human dimensions research for understanding and managing recreational fisheries within a social-ecological system framework. *Fish Manag Ecol* 20(2–3):111–124
- Jacobsen NS, Burgess MG, Andersen KH (2017) Efficiency of fisheries is increasing at the ecosystem level. *Fish Fish* 18:199–211
- Jakoby O, Grimm V, Frank K (2014) Pattern-oriented parameterization of general models for ecological application: towards realistic evaluations of management approaches. *Ecol Model* 275:78–88
- Katehakis MN, Veinott AF (1987) The multi-armed bandit problem: decomposition and computation. *Math Oper Res* 12(2):262–268. <https://doi.org/10.1287/moor.12.2.262>
- Katsukawa T, Matsuda H (2003) Simulated effects of target switching on yield and sustainability of fish stocks. *Fish Res* 60(2):515–525
- Kellner JB, Tetreault I, Gaines SD, Nisbet RM (2007) Fishing the line near marine reserves in single and multispecies fisheries. *Ecol Appl* 17(4):1039–1054. <https://doi.org/10.1890/05-1845>
- Kotchen MJ, Young OR (2007) Meeting the challenges of the anthropocene: towards a science of coupled human–biophysical systems *Glob Environ Change* 17(2):149–151
- Kuleshov V, Precup D (2014) Algorithms for multi-armed bandit problems. arXiv preprint. <http://arxiv.org/abs/1402.6028>
- Kydland F, Prescott E (1977) Rules rather than discretion: the inconsistency of optimal plans. *J Polit Econ* 85(3):473–491
- Lee J-S, Filatova T, Ligmann-Zielinska A, Hassani-Mahmooei B, Stonedahl F, Lorscheid I, Voinov A, Polhill G, Sun Z, Parker DC (2015) The complexities of agent-based modeling output analysis. *J Artif Soc Soc Simul* 18(4):4
- Little LR, McDonald AD (2007) Simulations of agents in social networks harvesting a resource. *Ecol Model* 204:379–386
- Little LR, Kuikka K, Punt AE, Pantus F, Davies CR, Mapstone BD (2004) Information flow among fishing vessels modelled using a Bayesian network. *Environ Model Softw* 19:27–34
- Little LR, Punt AE, Mapstone BD, Begg GA, Goldman B, Williams AJ (2009) An agent-based model for simulating trading of multi-species fisheries quota. *Ecol Model* 220:3404–3412
- Little LR, Punt AE, Dichmont CM, Dowling N, Smith DC, Fulton E, Sporic M, Gorton RJ (2016) Decision trade-offs for cost-constrained fisheries management. *ICES J Mar Sci* 73:494–502
- Liu J, Dietz T, Carpenter SR, Alberti M, Folke C, Moran E, Pell AN, Deadman P, Kratz T, Lubchenco J, Ostrom E (2007) Complexity of coupled human and natural systems. *Science* 317(5844):1513–1516
- Luke S (2009) Essentials of metaheuristics: a set of undergraduate lecture notes. Lulu, Morrisville. <https://cs.gmu.edu/~sean/book/metaheuristics/Essentials.pdf>
- Mangel M (2000) Irreducible uncertainties sustainable fisheries and marine reserves. *Evol Ecol Res* 2(4):547–557
- McClanahan TR, Kaunda-Arara B (1996) Fishery recovery in a coral-reef marine park and its effect on the adjacent fishery. *Conserv Biol* 10:1187–1199
- Miller JH (1998) Active nonlinear tests (ANTs) of complex simulation models. *Manag Sci* 44(6):820–830
- Miller SJ, Deacon RT (2017) Protecting marine ecosystems: regulation versus market incentives. *Mar Resour Econ* 32(1):83–107. <https://doi.org/10.1086/689214>
- Murawski SA, Wigley SE, Fogarty MJ, Rago PJ, Mountain DG (2005) Effort distribution and catch patterns adjacent to temperate MPAs. *ICES J Mar Sci Journal du Conseil* 62(6):1150–1167
- Ostrom E (1999) Coping with tragedies of the commons. *Annu Rev Polit Sci* 2:493–535
- Parnell PE, Dayton PK, Margiotta F (2007) Spatial and temporal patterns of lobster trap fishing: a survey of fishing effort and habitat structure. *Bull South Calif Acad Sci* 106(1):27–37
- Poos JJ, Turenhout MN, van Oostenbrugge AE, Rijnsdorp AD (2013) Adaptive response of beam trawl fishers to rising fuel cost. *ICES J Mar Sci* 70(3):675–684
- Porch CE, Turner SC, Schirripa MJ (2007) Reconstructing the commercial landings of red snapper in the Gulf of Mexico from 1872 to 1963. In: American Fisheries Society symposium, vol 60. American Fisheries Society Bethesda
- Punt AE, Butterworth DS, Moor CL, De Oliveira JA, Haddon M (2016) Management strategy evaluation: best practices. *Fish Fish* 17:303–334
- Railsback SF, Grimm V (2011) Agent-based and individual-based modeling: a practical introduction. Princeton University Press, Princeton
- Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, Held H, van Nes EH, Rietkerk M, Sugihara G (2009) Early-warning signals for critical transitions. *Nature* 461:53–59
- Schlueter M, McAllister RRR, Arlinghaus R, Bunnefeld N, Eisenack K, Hoelker F, Milner-Gulland EJ, Müller B, Nicholson E, Quaas M, Stöven M (2012) New horizons for managing the environment: a review of coupled social-ecological systems modeling. *Nat Resour Model* 25(1):219–272
- Shahriari B, Swersky K, Wang Z, Adams RP, Freitas N (2016) Taking the human out of the loop: a review of bayesian optimization. *Proc IEEE* 104(1):148–175. <https://doi.org/10.1109/jproc.2015.2494218>
- Shin Y-J, Cury P (2001) Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. *Aquat Living Resour* 14(2):65–80. [https://doi.org/10.1016/S0990-7440\(01\)01106-8](https://doi.org/10.1016/S0990-7440(01)01106-8)
- 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. *Can J Fish Aquat Sci* 61(3):414–431. <https://doi.org/10.1139/f03-154>
- Soulié J, Thébaud O (2006) Modeling fleet response in regulated fisheries: an agent-based approach. *Math Comput Model* 44:553–564
- Stonedahl F, Wilensky U (2010) Finding forms of flocking: evolutionary search in ABM parameter-spaces. In: Bosse T, Geller A, Jonker CM (eds) *Multi-agent-based simulation XI. Lecture notes in computer science*, vol 6532. Springer, Berlin, pp 61–75. https://doi.org/10.1007/978-3-642-18345-4_5
- Sutton SG, Ditton RB (2004) The substitutability of one type of fishing for another. *N Am J Fish Manag* 2:536–546
- Tesfatsion L (2003) Agent-based computational economics: modeling economies as complex adaptive systems. *Inf Sci* 149(4):262–268
- van Putten IE et al (2012) Theories and behavioural drivers underlying fleet dynamics models. *Fish Fish* 13(2):216–235
- Walters CJ (1986) *Adaptive management of renewable resources*. Macmillan Publishers Ltd, Basingstoke
- Watts ME, Ball IR, Stewart RS, Klein CJ, Wilson K, Steinback C, Lourival R, Kircher L, Possingham HP (2009) Marxan with zones: software for optimal conservation based land- and sea-use zoning. *Environ Model Softw* 24:1513–1521