

Calibrating ecosystem models to support ecosystem-based management of marine systems

Jacob W. Bentley^{1,*}, David Chagaris², Marta Coll^{3,4}, Johanna J. Heymans^{5,6}, Natalia Serpetti⁷, Carl J. Walters⁸, Villy Christensen^{8,4}

¹Natural England, Nobel House, 17 Smith Square, London SW1P 3JR, United Kingdom

²IFAS Nature Coast Biological Station, University of Florida, Cedar Key, FL 32625, United States

³Institute of Marine Sciences (ICM-CSIC), Passeig Marítim de la Barceloneta 27–49, 08003 Barcelona, Spain

⁴Ecopath International Initiative, Barcelona 08193, Spain

⁵European Marine Board, Jacobsenstraat 1, Oostende 8400, Belgium

⁶Scottish Association for Marine Science, Scottish Marine Institute, Oban PA37 1QA, United Kingdom

⁷European Commission, Joint Research Centre, 1, 21027 Ispra VA, Italy

⁸Institute for the Oceans and Fisheries, University of British Columbia, Vancouver, BC V6T1Z4, Canada

*Corresponding author. Natural England, Nobel House, 17 Smith Square, London SW1P 3JR, United Kingdom. E-mail: Jacob.Bentley@naturalengland.org.uk

Abstract

Ecosystem models, such as Ecopath with Ecosim (EwE), provide a platform to simulate intricate policy scenarios where multiple species, pressures, and ecosystem services interact. Complex questions often return complex answers, necessitating evidence and advice to be communicated in terms of trade-offs, risks, and uncertainty. Calibration procedures for EwE, which can act as a source of uncertainty and bias in model results, have yet to be explored in a comprehensive way that communicates how sensitive model outputs are to different calibration approaches. As the EwE community has grown, multiple divergent approaches have been applied to calibrate models through the estimation of vulnerability multipliers: parameters that augment the consumption rate limits of predators. Here we explore the underlying principles of vulnerability multipliers as well as existing calibration approaches and their justification. Two case studies are presented: the first explores how vulnerability multipliers emerge based on the chosen calibration approach using simulated data, while the second takes two operational EwE models (Irish Sea and Northwest Atlantic Continental Shelf) and compares their outputs when calibrated following alternate calibration approaches. We show how calibration approaches can impact model-derived advice and provide a list of best practice recommendations for EwE calibration.

Keywords: ecopath with ecosim; ecosystem modeling; temporal dynamics; calibration; fisheries; indicators

Introduction

There is a growing demand for ecosystem models that can support decision-makers in their commitments to implement ecosystem-based management of fisheries and other marine resources (Craig and Link 2023, Karp et al. 2023, Rodriguez Perez et al. 2023). Ecosystem models facilitate simulations of multiple species, interspecific interactions, top-down and bottom-up processes, and how they are impacted by the prevailing environmental and anthropogenic pressures. Although ecosystem models have been around for more than four decades (Andersen and Ursin 1978; Laevasto and Larkin 1981), only recently, and on rare occasions, has their management utility been realized in real-world settings (Howell et al. 2021, Craig and Link 2023).

The complexity and associated uncertainty of ecosystem models have historically been seen to reduce their suitability for tactical advice (Hilborn and Walters 1992, Plagányi et al. 2014, Townsend et al. 2008), while the structural inertia of traditional management systems and regulatory frameworks has hampered the operational rollout of ecosystem-based management (EBM) and ecosystem models (Christie 2005, Marshak et al. 2017, Craig and Link 2023, Karp et al. 2023). This has changed, however, as the importance of going beyond tactical management and towards a more strate-

gic and trade-off-orientated approach has been made clear through initiatives such as the EU Landing Obligation (e.g. Celić et al. 2018) and the development of alternative energy with the associated consequences for spatial planning (e.g. Serpetti et al. 2021).

The development and application of ecosystem models are constrained by technical challenges such as data gaps, imprecise observations, and uncertainty in input parameters (Steenbeek et al. 2021, Karp et al. 2023). Parameter estimation in ecosystem models can be subject to bias and limited transparency due to the lack of representative data, large parameter space, long run-times, and the absence of a standardized calibration approach. Without appropriate guidance, model calibration, through which simulations are brought into alignment with observations, can introduce considerable uncertainty and bias into model predictions and stimulate the development of divergent calibration strategies that vary across model frameworks and even individual modelers (Pethybridge et al. 2019).

Decision-makers are increasingly exploring how ecosystem models can inform management and policy, particularly in the context of cumulative pressure assessment in a rapidly changing environment and multispecies trade-offs. Justifiable and reproducible approaches for calibrating ecosystem models are

therefore needed to ensure models are capable of reproducing retrospective trends and have predictive power. The calibration step is essential for establishing the form predator impacts on prey take over a range of predator density. The form of these relationships sets the scope for trophic interaction-dependent changes that have occurred and could potentially occur as a result of management decisions and environmental changes. Regardless of the modeling platform, careful consideration is needed at this step.

Calibration procedures for ecosystem models vary depending on model complexity and how they represent different components and processes of the ecosystem. Existing calibration procedures involve manual and iterative tuning approaches to fit to historical observations (Pethybridge et al. 2019) as well as the implementation of automatic calibration routines (e.g. Scott et al. 2016). As models increase in complexity, with the addition of more species or biophysical processes, the number of parameters that could be tuned generally grows, increasing the potential for overfitting with associated poor predictive power.

Ecopath with Ecosim (EwE; Christensen and Walters 2004) is the most widely used ecosystem modeling framework, used to explore fishery management options in the context of their ecosystem impacts (e.g. Mackinson et al. 2018), and increasingly to simulate the often cumulative impacts of climate change (e.g. Serpetti et al. 2017, Stock et al. 2023), Marine Protected Areas (MPAs; e.g. Püts et al. 2023), bioaccumulation of contaminants (e.g. Tierney et al. 2018), and infrastructure development such as offshore wind (e.g. Serpetti et al. 2021). EwE is being increasingly applied to inform intergovernmental advice, particularly within the International Council for the Exploration of the Sea (ICES). EwE has been used by ICES to enhance single species catch advice with ecosystem information (ICES 2020, Bentley et al. 2021), add qualitative food web descriptions to ecosystem overviews (e.g. ICES 2022), and explore options to incorporate ecosystem information into Integrated Ecosystem Assessments (IEA; ICES 2018). Continuing progress is also being made through ICES using EwE to quantify the impacts of pressures on ecosystem services (ICES 2023), explore the flow and sequestration of carbon, and provide quantitative food web indicators for ICES Ecosystem Overviews (ICES 2019a).

EwE includes a static mass-balance model of the marine system, Ecopath, that serves as the initial conditions for time-dynamic simulations in Ecosim (Walters et al. 1997). Ecosim predictions are sensitive to the Ecopath input parameters (usually biomass, production and consumption rates, diet, and fishery removals) as well as the predator–prey “vulnerability multipliers,” which can be manually set or estimated in Ecosim during model calibration to determine the functional form of predator consumption rates as predator biomass changes. The impacts of vulnerability multipliers are conditional on Ecopath inputs, which set the baseline predation mortalities. The concept underpinning vulnerability multipliers is derived from foraging arena theory (Ahrens et al. 2012). In Ecosim, vulnerability multipliers have implications for stock-recruit dynamics, density dependence and compensation, carrying capacity, stock resiliency, interspecific interactions, and ecosystem energy flow (Walters and Martell 2004). However, the effect of different calibration strategies on the estimation of vulnerability multipliers in Ecosim and emergent stock productivity estimates has not yet been demonstrated, nor do we under-

stand in a comprehensive way how sensitive model outputs are to different approaches and how this may influence derived advice.

The goals of this paper are to shed additional light on the Ecosim vulnerability multipliers and the ecological theory behind them, describe different Ecosim calibration approaches, and demonstrate the sensitivity of model outputs and management advice to those approaches. We provide an updated overview of foraging arena theory and its implementation in Ecosim, followed by a discussion of the life history and ecological basis for initializing vulnerabilities and the existing approaches to estimation in Ecosim. Case studies are provided to demonstrate the effect of model fitting strategies on a conceptual model as well as two ecosystem models with drastically different structures that have each been used in real-world management applications. It is the overarching aim of this paper to increase awareness and transparency in the EwE calibration approach and provide recommendations for the estimation of vulnerability multipliers to support the growing application of EwE as a decision support tool.

Ecosim vulnerability multipliers and their estimation

Behavioral, morphological, and physical mechanisms limit the rates at which prey encounter predators (e.g. Walters and Juanes 1993, Green and Côté 2014). Behaviors such as vertical migration, schooling, and hiding reduce the individual’s probability of capture (Gilwicz 1986, Jolles et al. 2022), while physical processes such as mixing or sinking in the water column can limit the spatial overlap of predator and prey (Morison et al. 2019). Simple mass-action predictors (such as the Lotka-Volterra model; Wangersky 1978) do not account for the effect of risk-sensitive foraging behavior on prey availability and assume that encounters between predators and prey are entirely random and linear. In such models, biomass and encounter rates are the only factors limiting consumption.

Ecosim predictions of consumption based on a simple mass-action model have been modified to account for spatial and temporal constraints as a result of risk-sensitive foraging behavior (Walters et al. 1997, 2000, Ahrens et al. 2012). Prey biomass pools in Ecosim are dynamically divided into vulnerable and invulnerable components, which imply behavioral or physical mechanisms that limit the rate at which prey become vulnerable to predation (Fig. 1; Walters et al. 1997). The transfer rates between these components determine the amount of prey available to a predator and thus the degree to which a change in predator biomass will impact predation mortality and prey biomass.

These transfer rates, or “vulnerabilities,” (v_{ij}) as they are more widely known, influence predator and prey biomasses by regulating the consumption rates (Q_{ij}) of a predator j on its prey group i , as shown in Equation (1):

$$Q_{ij} = \frac{v_{ij}a_{ij}B_iB_jT_iT_jS_{ij}M_{ij}/D_j}{v_{ij} + v_{ij}T_iM_{ij} + a_{ij}M_{ij}B_iS_{ij}T_j/D_j/A} \times f(Env_t). \quad (1)$$

Equation (1) has grown over time in response to user requests for new functionality being added to impact consumption. Here, a_{ij} is the effective search rate of predator j feeding on prey i , which directly influences the time spent feeding and, thus, the risk of predation. The terms T_i and T_j are the relative feeding times of prey i and predator j , which can be adjusted via a feeding time adjustment rate: larger values represent a

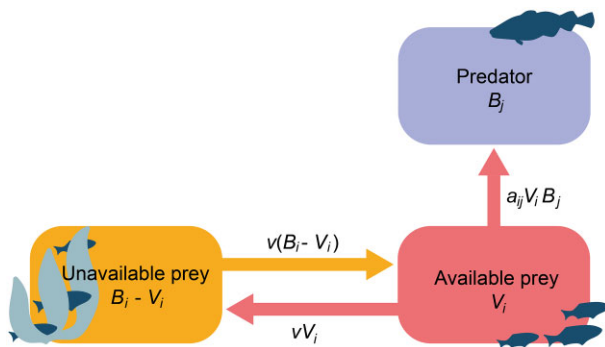


Figure 1 Simulation of flow between available (V_i) and unavailable ($B_i - V_i$) prey biomass in Ecosim. a_{ij} is the search rate of prey i by predator j , and v is the exchange rate between the vulnerable and invulnerable states. Fast equilibrium between the two prey states implies $V_i = vB_i/(2v + aB_j)$. Based on Walters et al. (1997).

more rapid adjustment of foraging time and reduced predation risk. S_{jj} are the seasonal and long-term forcing effects, which can be defined by the user to modify consumption in direct relation to environmental conditions. M_{ij} are the mediation forcing effects, which users can construct to account for changes in consumption that may have been indirectly caused by the behaviors of other groups or external factors. For example, small pelagic fish may become more vulnerable to predation by seabirds if they are driven to the surface by mediators such as large pelagic predators (Dill et al. 2003). Handling time, D_j , limits the rate of consumption and represents the notion that predators have limited time available for foraging due to the time spent in pursuit of prey or manipulating/ingesting captured prey (Holling 1959, Schoener 1971). The size of the foraging arena is denoted as A . Finally, $f(Env_t)$ represents an environmental response function that restricts the size of the foraging arena to account for external drivers, which may change over time (Christensen et al. 2014). This function has principally been used to make consumption rates dynamic with time and space in response to temperature and habitat (e.g. Christensen et al. 2015, Bentley et al. 2017, Serpetti et al. 2017, Corrales et al. 2018).

Users cannot adjust vulnerability exchange rates (v_{ij}) directly. Instead, users interact with these rates via vulnerability multipliers (k_{ij}), which can be more easily interpreted as the maximum increase in predation mortality rate that a predator can exert on a prey if the predator were to grow to its carrying capacity. The increase is relative to baseline Ecopath predation mortality rates ($M2$, where $M2 = Q_{ji}/B_i$). The vulnerability exchange rates (v_{ij}) are then set to the vulnerability multiplier (k_{ij}) multiplied by the baseline predation mortality ($M2$), i.e. $v_{ij} = k_{ij}M2$. Multipliers can range from one to infinity with two as the default value. When v_{ij} results in a low rate of exchange the relationship between predator density and mortality on prey saturates quickly. As k_{ij} is increased, resulting in an increase in v_{ij} , this relationship straightens.

The vulnerability multipliers within the Ecopath baseline establish the shape of the predator–prey relationship and this shape does not change when running Ecosim across years: Ecosim dynamically handles the consequences of changes in predator and prey abundance based on the baseline situation, including changes in carrying capacity over time. The default value of two assumes that the predation mortality rate can double at the highest, while a value of one means that the

predator is at its carrying capacity, which by definition means that it fully utilizes its prey, so it cannot increase the predation mortality its causing on the prey.

High vulnerability multipliers imply strong top-down control and weak bottom-up control. With top-down control, a doubling of predator abundance may result in close to a doubling in the predation mortality they are causing, which can occur when a predator is far from its carrying capacity. Of course, predator carrying capacity is also dependent on prey biomass, but the key aspect of top-down control is that it is the predator abundance that determines the rate of prey consumption. With low vulnerability multipliers where a predator is at its carrying capacity, any increase in consumption has to be linked to changes in prey productivity—i.e. in bottom-up factors (Christensen and Walters 2023).

For exploited species, it is thus extremely important to recognize that vulnerability multipliers do not only reflect the ecological limits caused by prey and predator behavior, but also how depleted the exploited species is in the base Ecopath model relative to the natural level (i.e. carrying capacity) that might be achieved if fishing were stopped. As such, for overexploited species to recover following reduced fishing, vulnerability multipliers need to be set relatively high so that predators can consume far more prey than in the initial Ecopath snapshot. Higher vulnerability multipliers tend to make groups more sensitive and responsive to changes in fishing mortality. It is also important to consider not only the consequences of risk-sensitive foraging but also the ecological consequences of predator depletion. As a predator population increases it may expand its distribution and access additional foraging areas, albeit this may be restricted by species ecology and habitat availability.

Over time, EwE users have adopted and developed multiple approaches to parameterize the vulnerability multipliers (Fig. 2). While some approaches derive estimates from *a priori* knowledge and ecological observations or hypotheses in data-poor situations (e.g. Rehren et al. 2022), users are more frequently turning to formal statistical estimation using calibration time series and a tuning process when time series are available (e.g. Scott et al. 2016). Statistical fitting routines estimate vulnerability multipliers that bring simulations closer in line with observations. However, users should be cautious, as statistically optimized multipliers may stray away from values that might be considered ecologically realistic. A thorough sense-check is always recommended. The following sections explore these different approaches in more detail.

Using ecology and history to estimate or initialize vulnerability multipliers

Vulnerability multipliers are perhaps easier to comprehend when it is assumed that they reflect how far an exploited predator is from their carrying capacity (e.g. interpreted as unfished state), and that vulnerability multipliers should facilitate consumption rates that enable a species to recover from their Ecopath biomass to their unfished biomass if fishing ceases. EwE can use the ratio between a group's unfished biomass and its Ecopath biomass to estimate vulnerability multipliers for exploited groups (see Box 1 in [Supplementary material](#)). This does, however, require information on the unfished biomasses of exploited species, which could either come from quantitative historical estimates (e.g. Rosenberg et al.



Figure 2 Pathways for estimating vulnerability multipliers (k_{ij}) in Ecosim. All pathways end with a reference to peer-reviewed examples.

2005) or a proxy method (e.g. the CPUE approach used by Rehren et al. 2022).

An added corollary is that the unfished state may be associated with high abundance of top predators and low abundance of their prey due to high predation mortality. If such top predator populations are fished down, predator release may cause the prey to increase. For those prey, the vulnerability multipliers should thus be set to a high value, even though the baseline model represents the unfished state. Mid-trophic-level organisms are likely to also be fished; therefore, their vulnerability multipliers could also be set to reflect related depletion.

It can indeed be difficult to specify reasonable vulnerability multipliers for non-exploited species. Here, vulnerability multipliers need to be considered in the context of the foraging arena: the fine-scale spatial structure of the trophic interactions and what proportion of prey may be vulnerable to predation at any moment (Fig. 3). The activity, spatial restrictions, and distributions of species provide insight into the likely vulnerability of prey to predation. This in turn provides a starting point from where it is possible to assign vulnerability multipliers. The distribution of predators could be restricted by limited mobility, habitat requirements, or the predation risk they face themselves, whereas prey vulnerability may be influ-

enced by the time they spend in and out of refuge. This can be related to the availability of shelter, such as macroalgae (e.g. Wilson et al. 1990), or specific ontogenetic life stages; e.g. juvenile fish may allocate less time to foraging, or be more restricted spatially (and thus unable to access vulnerable pools of prey) than their adult counterparts (Nunn et al. 2012). Different behaviors, such as dispersal behaviors (e.g. moving to spawning sites), aggressive behaviors, or evolutionary behaviors (e.g. changes in shoaling dynamics) may also influence vulnerability to predation (e.g. Ladich 2022, Sbragaglia et al. 2021).

Trophic levels have also been used to approximate vulnerability multipliers (e.g. Cheung et al. 2002, Kluger et al. 2016). This approach has been used in situations where time series data were unavailable and assumes that the vulnerability multipliers are proportional to the trophic level of the predator. If we refer to the relationship between vulnerability multipliers and unfished biomass, this approach assumes that higher trophic levels are further removed from their unfished biomass than lower trophic levels. This may seem like a reasonable assumption considering how global fisheries have historically targeted and depleted higher trophic level fish stocks (Shannon et al. 2014), but conflicts with the concept of using *a priori* knowledge to parameterize vulnerability multipliers

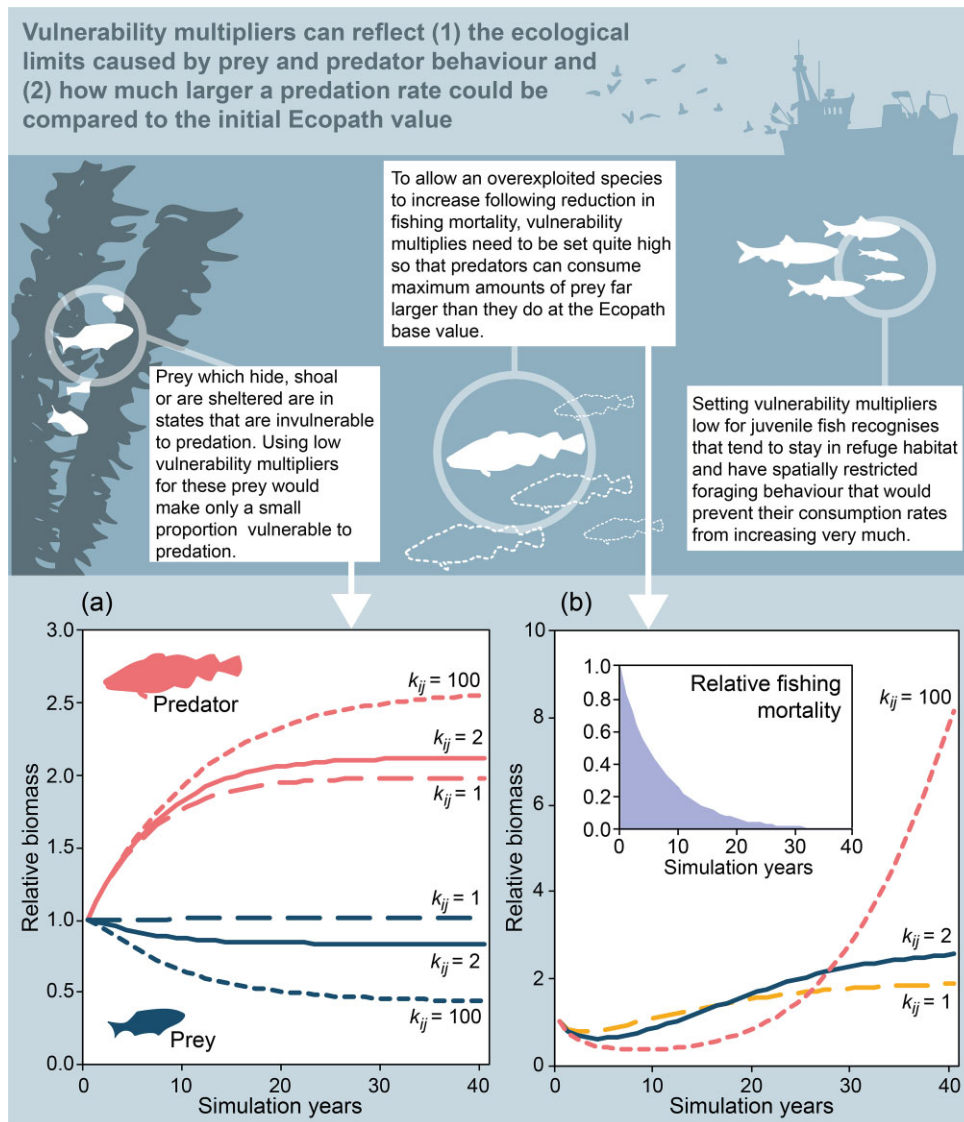


Figure 3 Using history and ecology to estimate vulnerability multipliers (k_{ij}) in Ecosim. The illustration at the top of the figure provides examples where k_{ij} estimates can be inferred from functional group ecology and life history. Model simulations demonstrate how (a) prey vulnerability influences predator–prey biomass trajectories when predator biomass increases and (b) how k_{ij} estimates impact the rate of functional group recovery following reduction in fishing (inset figure).

based on region-specific trends in historical exploitation or ecology.

Finally, an approach to setting vulnerability multipliers was recently applied by Chagaris *et al.* (2020) to constrain how much predation mortality by a given predator could increase relative to a prey's total natural mortality:

$$k_{ij} = \frac{M2_{cap} \times M_i}{M2_{base,ij}}, \quad (2)$$

where $M2_{cap}$ defines the percentage of the natural mortality of a prey that a predator can be responsible for (between 0 and 1), M_i is the natural mortality of prey i , and $M2_{base,ij}$ is the base predation mortality by predator j on prey i . There may be ecological reasons, or reasons derived from data, to prevent a single predator from being accountable for large proportions of a prey's natural mortality. Using these k_{ij} instead of default or model estimated values (which are often higher) may also be driven by ambitions for model perfor-

mance. Chagaris *et al.* (2020) found that extremely high k_{ij} estimated by Ecosim led to dynamic instability at high fishing mortality rates when evaluating equilibrium yield curves, and using $M2_{cap}$ values between 0.75 and 1.0 led to more reasonable estimates for F_{MSY} (the fishing mortality at maximum sustainable yield) while also constraining theoretical maximum predation mortality rates to values that were compatible with prey natural mortality rates.

Statistical approaches for estimating vulnerability multipliers

Evaluating and weighting time series

It is becoming more common for Ecosim vulnerability multipliers to be estimated using statistical routines, with Heymans *et al.* (2016) demonstrating that it is best practice to estimate vulnerability multipliers by calibrating model simulations to time-series reference data. Longer time series are preferable as

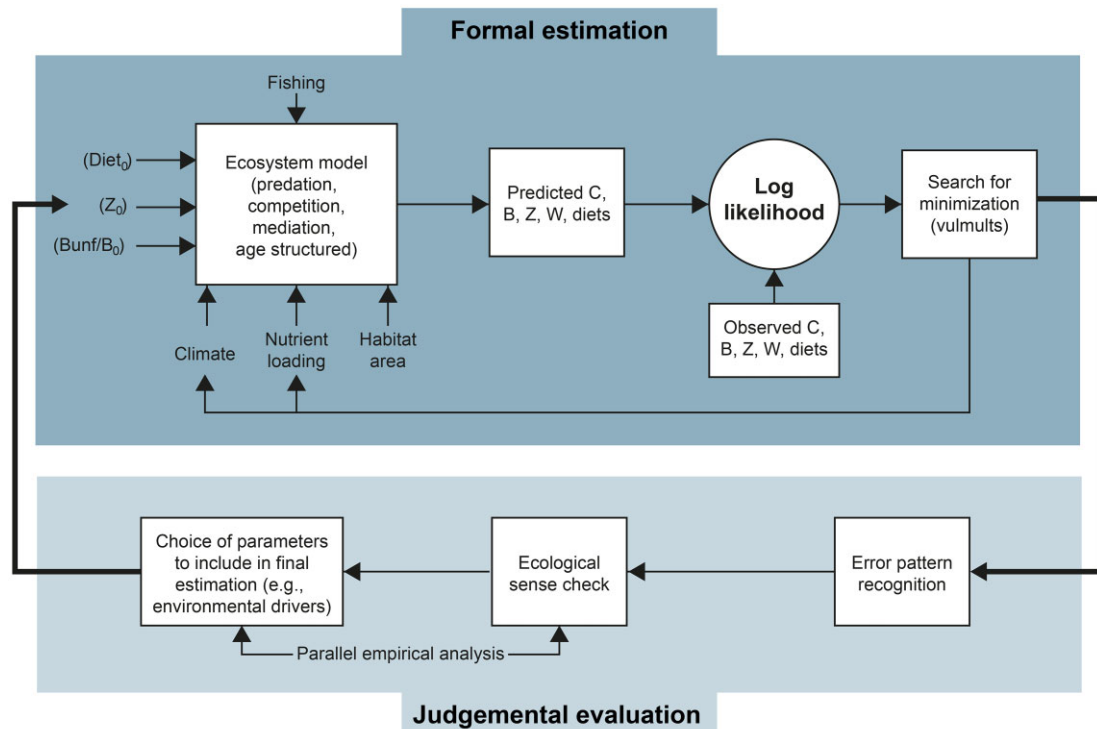


Figure 4 Overview of the Ecopath and Ecosim modeling processes. Using log likelihood criteria, the input parameters or anomalies (e.g. climate) may be estimated based on a non-linear search routine and vulnerability multiplier (vulmult) estimation. Prediction (fitting) failures after each estimation trial then inform judgmental changes in model structure and parameters. B is biomass, Z is total mortality, C is catch, and W is average weight. Subscript 0 refers to the Ecopath model base year, and CC to carrying capacity. Adapted from Christensen and Walters (2011).

they provide an opportunity to explore important drivers of change and tend to have strong contrast in the data, which improves the model's ability to estimate parameters, leading to more confidence in our assessment of ecosystem dynamics.

The quality of time series data is important, especially if the fitted model is to be used for management purposes, as vulnerability multipliers, and thus predation rates, will be augmented for simulations to fit against the data. In Ecosim, users can weight time-series data to represent how reliable or variable time series are compared to the other reference time series. Low weights imply that the data either has high variance or is unreliable (e.g. underestimated or uncertain catches). Weightings impact the contribution of time series to the assessment of model performance, where a weight of 0 indicates that the time series will not be used in the calculation of “goodness of fit.” Weightings can be assigned based on a qualitative assessment of data pedigree (e.g. based on data origin), or by using more quantitative information, such as confidence intervals from survey estimates, the retrospective analyses of stock assessment models, or signal-to-noise ratio assessments (Heymans et al. 2016).

The procedure of estimating vulnerability multipliers that improve the fit of model simulation to calibration data is based on the estimation of both the sum of squares (SS) and the Akaike Information Index (AIC) (Akaike 1998, Cavanaugh and Neath 2019). The SS is simulated using log likelihood criteria (Fig. 4): this is a weighted SS of log simulations from log observations, scaled in the case of relative observations (y) by the maximum likelihood estimate of the relative simulation scaling factor (q) in the equation $y = q \times n$, where n is the absolute observation. When generating a set of models under different fitting hypotheses, AIC is used to iden-

tify the model of best fit. AIC is a tool for model selection that penalizes for fitting too many parameters relative to the time series available for estimating the SS and is calculated as:

$$AIC = n \times \log \left(\frac{\min SS}{n} \right) + 2K, \quad (3)$$

where n is the total number of observations, or time-series values, from the loaded calibration time series and K is the number of parameters estimated. When sample size is small, there is a large probability that AIC will select models with too many estimated parameters (i.e. overfit models). AICc is used to address this potential overfitting by including a correction for small sample sizes:

$$AICc = AIC + 2K \times \frac{K - 1}{n - K - 1}. \quad (4)$$

As a rule of thumb, Burnham and Anderson (2004) state that AICc should be used unless $\frac{n}{K} >$ about 40. In other words, unless the number of estimated parameters equates to a minimum 2–3% of the amount of data, use AICc. AICc should therefore be used when assessing EwE model performance.

A word of caution, the AIC calculations assume that the observations are independent, whereas time-series data, such as typically used for ecosystem modeling have high autocorrelation. For this reason, it is advisable to test the impact on assumptions about n on model selection.

Multiple approaches have been developed to statistically estimate vulnerability multipliers (refer to Fig. 2). They can be estimated for predators, providing a single multiplier limit to all of a given predator's base predation rates, and they can be estimated for individual predator–prey relationships, which assumes that the multiplier limits are heterogeneous across

prey. Vulnerability multipliers are estimated by predator or predator–prey, and not just prey, as the terms impact the consumption rates and carrying capacities of predators and therefore have more influence on predator biomass dynamics when estimated by predator. Estimating by prey may also increase the risk of overfitting, which is discussed in more detail in the case studies. This choice between a predator or predator–prey approach tends to be associated with user preference, ecological justification, or determined based on the approach that produces the best fit model. Whether estimating predator or predator–prey vulnerability multipliers, there are a few ways to select which vulnerability multipliers should be estimated: (i) manual selection based on *a priori* knowledge or species priority; (ii) select vulnerability multipliers for groups with calibration time series; or (iii) select the most sensitive vulnerability multipliers (i.e. those that when changed have the largest impact on *SS*). Manually selecting vulnerability multipliers allows for an early integration of ecological information but may lead to a suboptimal model fit if the *SS* is not sensitive to the selected multipliers. Conversely, the sensitivity search may optimize model fit but it is purely statistical and fits will not necessarily make sense ecologically. Only estimating vulnerability multipliers for groups with time series acknowledges that, to some degree, the parameter should be constrained by the available time series. The level of group connectedness within the food web (e.g. group consumption and predation) may also constrain the parameter search if changes in vulnerability multipliers impact the *SS* contribution of other groups. Groups that do not have informative time series or, have low connectedness in the food web, have widely variable estimated vulnerability multipliers.

Choosing how many vulnerability multipliers to estimate, without overfitting is another point of confusion and discussion. The number of vulnerability multipliers that can be potentially estimated is often significantly more than the data available to constrain simulations. EwE best practices suggest that a conservative number of degrees of freedom and, therefore, parameters to estimate is one less than the number of calibration time series available ($K - 1$; Heymans *et al.* 2016). This approach recognizes that values within time series are highly autocorrelated, viewing each time series as an “independent observation,” but the $K - 1$ approach could be overly conservative, especially if long time series are available.

Both manual and automated statistical calibration routines are available in Ecosim to search for vulnerability multipliers. The manual approach can be arduous when testing multiple fitting hypotheses (e.g. with or without fishing effort or primary production anomalies), as the number of plausible fitting combinations can easily reach the hundreds, if not thousands, increasing the likelihood of user error. In the past, users have overcome this issue by only testing the n th fitting scenario (e.g. 5, 10, and 15 vulnerability multipliers, etc.; e.g. Alexander *et al.* 2015), however, this approach risks overlooking the vulnerability multiplier combination, which produces the best statistical fit. The stepwise fitting procedure developed by Scott *et al.* (2016) automates this process, allowing for a broad exploration of the parameter space, which accelerates the process and removes the problem of user error. Recent improvements to the automated approach have increased the computational speed by enabling multiple fitting scenarios to be tested simultaneously using computers multithreading capabilities (J. Steenbeek personal communication, September 20, 2023).

Novel approaches to estimate vulnerability multipliers using the manual and automated fitting routines have also been developed for two EwE models, which are being used operationally to inform fisheries catch advice (Chagaris *et al.* 2020, Bentley *et al.* 2021). Bentley *et al.* (2021) employed an approach that combined searches for predator vulnerability multipliers and predator–prey vulnerability multipliers, whereas the approach developed by Chagaris *et al.* (2020) uses the manual fitting tool in Ecosim to iteratively estimate the $K - 1$ most sensitive predator–prey vulnerability multipliers over multiple sequential (repeated) tuning iterations. Full methodologies for these approaches are provided in the [Supplementary material](#).

It is worth reiterating that statistical estimation of vulnerability multipliers does not necessarily have any bearing on ecology. While it is possible to exclude vulnerability multipliers from the search routine, there is currently no mechanism to include prior information or ecologically sensible bounds to constrain the limits for vulnerability multipliers included in the search routine. A judgment evaluation following the formal estimation of vulnerability multipliers should be applied to: (i) reflect on the ecological assumptions attached to estimated vulnerability multipliers; (ii) assess how realistic functional group simulations are (in hindcast and future); and (iii) understand and fix issues with model structure and parameterization (Fig. 4) (e.g. Corrales *et al.* 2017). It is possible to view the fit of each functional group to calibration time series and its contribution to the overall *SS* in Ecosim via the “*SS* group plots.” This is often used to screen issues with model simulations, such as contradicting trends or misalignment in initial time steps, and direct fixes.

However, what is often not accounted for when estimating vulnerability multipliers is their impacts on the advice products, such as F_{MSY} or food web indicators. The focus is often only on the goodness of fit of the model, but the impacts of estimated vulnerability multipliers on predictions and reference points should be evaluated (e.g. Rehren *et al.* 2022). We provide two case studies to explore how alternate fitting approaches impact the emergence of vulnerability multipliers and how vulnerability multipliers impact model outputs.

Case study 1: How fitting approaches and calibration data influence which vulnerability multipliers emerge

A hypothetical EwE model (Anchovy Bay: often used to test EwE scenarios; see Christensen and Walters 2023) was used to investigate how vulnerability multipliers emerge (and whether they re-emerge through fitting) and how this process is influenced by: (i) noise in the calibration data and (ii) the chosen approach for estimating vulnerability multipliers: “predator” vulnerability multipliers or “predator–prey” vulnerability multipliers. We investigated the impact of emerging vulnerability multipliers on biomass and catch simulations and estimates of F_{MSY} .

Building a base Ecosim model

Fictitious Ecosim simulations for Anchovy Bay were created by adding temporal trends to fishing effort and adjusting vulnerability multipliers. Simulated fishing effort trends (Supplementary Fig. 1) reflected trends often seen in reality: (i) sealers fishing effort followed an exponential decline as may be ex-

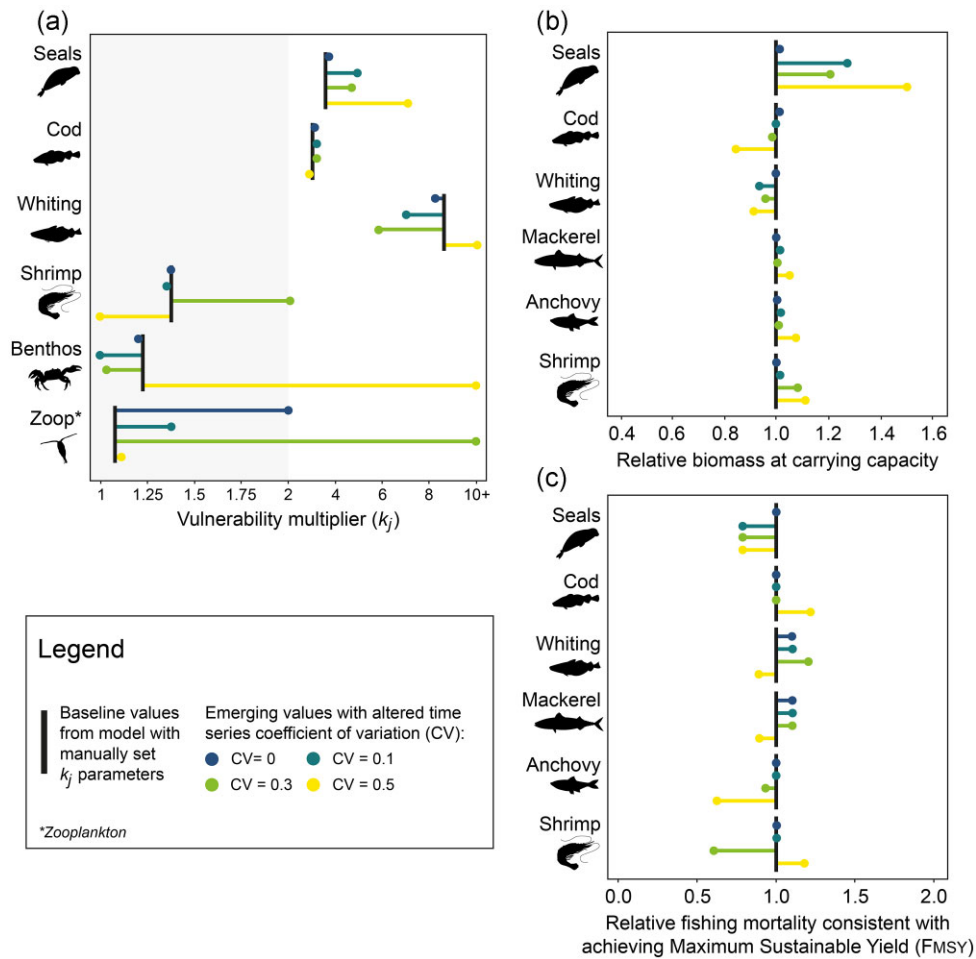


Figure 5 Estimation and impact of predator vulnerability multipliers (k_{ij}). The Anchovy Bay ecosystem model was calibrated against generated time series with incremental CV to identify the impact of time series quality on (a) k_{ij} re-emergence and how k_{ij} estimates impacted; (b) functional group carrying capacities in the absence of fishing; and (c) estimates of relative fishing mortality consistent with achieving maximum sustainable yield (F_{MSY}).

pected in response to conservation efforts/policy; (ii) trawlers fishing effort followed an exponential decline under the assumption that whitefish (cod and whiting) stocks have been overexploited, leading to reductions in effort to encourage stock recovery; (iii) seiners and bait boat effort followed a slight linear increase in response to growing demand; and (iv) shrimpers effort increased assuming fishers shifted their target species to shrimp following reduced opportunities to catch white fish. Vulnerability multipliers (k_{ij}) were adjusted following ecological assumptions and assumptions linked to the fishing effort trajectories (Supplementary Table 1). To distinguish between scenarios more easily, predator vulnerability multipliers will hereafter be denoted as k_j , while predator–prey vulnerability will remain as k_{ij} . For predator k_j estimates, a mix of high, low, and default k_j estimates were applied. For groups that were assumed to be overexploited, k_j values were estimated using the “Estimate Vulnerabilities” interface (as described in the Supplementary material). For the predator–prey k_{ij} estimates, the Ecosim sensitivity search was used to identify the 10 most sensitive predator/prey k_{ij} parameters, which were then adjusted to ensure a range of high and low k_{ij} estimates were included (Supplementary Table 2).

For the purpose of this exercise, these two simulations (one with predator k_j and one with predator–prey k_{ij}) were viewed as perfect representations of their ecosystems, i.e. the biomass

and catch simulations were “real observations” driven by the “true” vulnerability multipliers. The aim of the following exercise was to test whether, when using these “real observations” as calibration time series, the “true” vulnerability multipliers would re-emerge, and whether the addition of noise to the “real observations” had any impact on the emerging vulnerability multipliers. Biomass and catch simulations were extracted from Ecosim and four scenarios for observation data quality were prepared: noise [random noise, normally distributed around the mean (true) biomass trend to represent observation error] was added to the calibration time series with coefficients of variation (CV) of 0 (no noise), 0.1, 0.3, and 0.5 (Supplementary Fig. 2).

Predator vulnerability multipliers

Vulnerability multipliers were reset to the default value of two; fishing dynamics were not changed from those used to produce the “real observations.” The exported biomass and catch time series were used as calibration time series to estimate predator vulnerability multipliers for the functional groups seals, cod, whiting, shrimp, benthos, and zooplankton using the manual stepwise fitting interface. k_j values for groups that had values of two in the initial model were not altered.

Fig. 5 shows how k_j parameters emerged after model calibration, and how this altered functional group carrying

capacities in the absence of fishing and F_{MSY} estimates. k_j values, which emerged when estimated using the calibration time series with no noise were similar to the “true” k_j parameters (Fig. 5a). Adding noise to the calibration time series led to divergence between the estimated k_j values and the “true” parameters, highlighting the impact data quality can have on the fitting procedure and thus stressing the importance of evaluating the suitability of time series before using them to drive model calibration. While not tested in the simulations presented, it is important to reiterate that time series can be weighted to represent the reliability of the data relative to established confidence intervals or signal-to-noise assessments (Heymans et al. 2016). The variability in k_j re-emergence under the four data quality scenarios was also unique to specific functional groups; e.g. k_j estimates for cod showed greater re-emergence accuracy (or consistency) when compared to other functional groups. Cod is highly connected within the food web (i.e. cod is an opportunistic predator i.e. also preyed upon by higher trophic levels); therefore, vulnerability multipliers that improve the model fit tend to be more constrained due to their potential to have large cascading impacts on the wider food web. In addition, cod also experienced a period of collapse followed by recovery, which provides much-needed contrast for the model to reliably estimate the vulnerability multipliers.

Functional group carrying capacities and estimates of F_{MSY} were impacted by emerging k_j values (Fig. 5b and c). Carrying capacities from scenarios with k_j parameters calibrated against data with no noise were most similar to those achieved with the “true” k_j parameters (Fig. 5b), with dissimilarity generally increasing with the addition of noise to the calibration data. The importance of acknowledging the impact of k_j estimates beyond model fit is demonstrated with the resulting F_{MSY} estimates: relative changes to F_{MSY} estimates mirrored the deviations of estimated k_j values relative to the “true” k_j values (Fig. 5c). Increases in k_j values led to decreases in F_{MSY} , while decreases in k_j values led to increases in F_{MSY} . This is because higher k_j values enable groups to recover faster with the cessation of fishing and reach a higher carrying capacity, but they also decrease stock resilience to increases in F (functional groups decline faster and more severely if you increase their k_j). It is worth noting that where differences between true and estimated F_{MSY} occurred, they were not proportional to the difference in true and estimated vulnerability multipliers (i.e. large changes in k_{ij} do not result in equally large changes to F_{MSY}), as F_{MSY} is also constrained by other components, which limit production, including other mortality and the biomass of available prey as well as prey dynamics in the foraging arena.

Predator–prey vulnerability multipliers

Similar to the predator scenario, vulnerability multipliers were reset to the default of two, and the exported biomass and catch time series (generated with “true” predator–prey vulnerability multipliers) were used as calibration time series to estimate predator–prey vulnerability multipliers. Predator–prey k_{ij} values for the 10 most sensitive predator/prey k_{ij} parameters were estimated using the manual stepwise fitting interface. Fig. 6 shows how k_{ij} parameters emerged and how this altered functional group carrying capacities and F_{MSY} estimates.

In comparison to the emergence of predator vulnerabilities, the emergence of predator–prey vulnerabilities was

less constrained with poor k_{ij} re-emergence accuracy across all calibration data scenarios (Fig. 6a). Functional group carrying capacities showed higher dissimilarity from their baseline when compared to predator simulations and their baseline (Fig. 6b). Carrying capacity dissimilarity increased with the addition of noise to the calibration data; however, simulations with no/low noise were notably more dissimilar when estimating predator–prey vulnerabilities as opposed to predator vulnerabilities (Fig. 6b), which is due to the greater differences in k_{ij} estimates.

Relative F_{MSY} estimates, influenced by predator–prey k_{ij} values, showed higher dissimilarity from their baseline (Fig. 6c) when compared to F_{MSY} estimates influenced by predator k_{ij} values (Fig. 5c). The links between predator–prey k_{ij} values and F_{MSY} are less obvious than the links between predator k_j values and F_{MSY} due to the more complex interaction-specific consumption limits. This is particularly true for groups with mixed diets where diet composition is relatively homogenous across multiple prey (e.g. cod, whiting, seals, and mackerel), while links between predator–prey k_{ij} values and the F_{MSY} estimates for groups that are heavily dependent on a single prey group were observed for anchovy (F_{MSY} mirrors the anchovy/zooplankton k_{ij} estimates) and shrimp (F_{MSY} mirrors the shrimp/benthos k_{ij} estimates).

Case study 2: How fitting approaches impact the evidence derived from operationally used models

The Irish Sea EwE model and Northwest Atlantic Continental Shelf EwE model (hereafter called NWACS-MICE) have both been used to inform fisheries advice for their respective regions using ecological/ecosystem reference points (Chagaris et al. 2020, Bentley et al. 2021, Howell et al. 2021). Both models were designed to focus on commercial fisheries; however, they have very different structures in terms of model complexity (Table 1). In this case study, the two models were used to demonstrate the outcomes and management implications of vulnerability multiplier (k_j or k_{ij}) estimation and compared estimates of F_{MSY} and ecosystem indicators. Ecosystem indicators selected for this analysis included total system biomass, commercial biomass, total catch, system diversity (Kempton's Q; Kempton and Taylor 1976), the trophic level of the catch, and the trophic level of the community. Estimates of F_{MSY} and ecosystem indicators were compared across the following nine fitting approaches:

- (1) Predator k_j values, where the number of parameters estimated is one less than the number of available calibration time series ($K - 1$).
- (2) Predator k_j values estimated for all functional groups with time series.
- (3) Predator k_j values using the automated stepwise fitting approach, where the applied k_j values are taken from the model with the lowest $AICc$, and the vulnerabilities are reset to the default Equation (2) at each fitting iteration.
- (4) Predator k_j values using the automated stepwise fitting approach, where the applied k_j values are taken from the model with the lowest $AICc$, and the vulnerabilities are retained from previous fitting iterations.

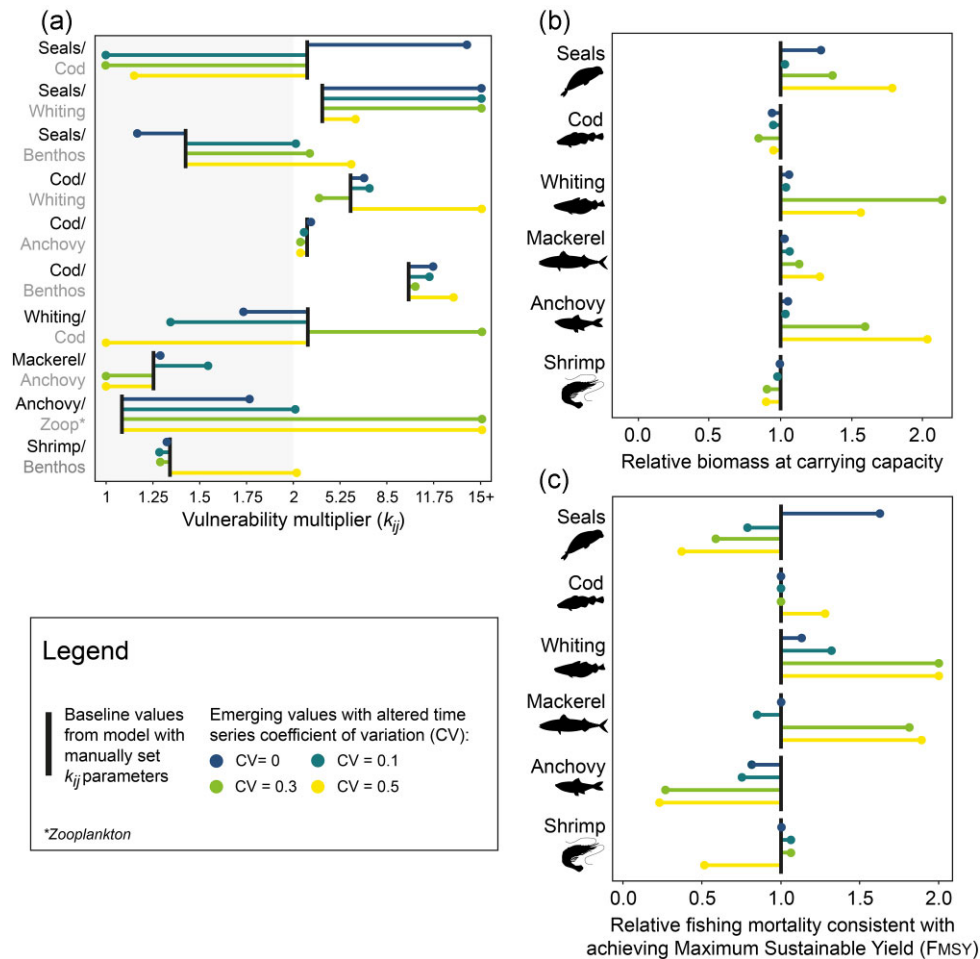


Figure 6 Estimation and impact of predator–prey vulnerability multipliers (k_{ij}). The Anchovy Bay ecosystem model was calibrated against generated time series with incremental CV to identify the impact of time-series quality on (a) k_{ij} re-emergence and how k_{ij} estimates impacted; (b) functional group carrying capacities in the absence of fishing; and (c) estimates of relative F_{MSY} .

Table 1. Comparison of key characteristics between static (Ecopath) and time-dynamic (Ecosim) models of the Irish Sea and Northwest Atlantic Continental Shelf (NWACS-MICE).

Static model	Irish Sea EwE model	NWACS-MICE
Total number of functional groups	41	17
Number of multi-stanza groups	4	5
Fleets	8 (gear specific)	8 (species specific)
Time dynamic model	Irish Sea EwE model	NWACS-MICE
Time-period	1973–2016	1985–2017
Calibration time series	52 (28 biomass, 24 catch)	28 (18 biomass, 10 catch)
Fitting approach	Predator and predator–prey search	Repeated predator–prey search

- (5) Predator–prey k_{ij} values, where the number of parameters estimated is one less than the number of available calibration time series ($K - 1$).
- (6) Predator–prey k_{ij} values using the automated stepwise fitting approach, where the applied k_{ij} values are taken from the model with the lowest $AICc$, and the vulnerabilities are reset to the default Equation (2) at each fitting iteration.

- (7) Predator–prey k_{ij} values using the automated stepwise fitting approach, where the applied k_{ij} values are taken from the model with the lowest $AICc$, and the vulnerabilities are retained from previous fitting iterations.
- (8) Predator–prey k_{ij} values using a repeated manual stepwise fitting approach, where the estimated k_{ij} ($K - 1$) are retained from one iteration to the next (with a total of 5 iterations) and the final configuration is that with the lowest $AICc$, as was done in Chagaris et al. (2020).
- (9) A combination of predator k_j and predator–prey k_{ij} values using the methods outlined by Bentley et al. (2020). Predator k_j values were estimated using the automated stepwise fitting approach in #3. Predator–prey k_{ij} values were estimated using a manual stepwise fitting approach and the remaining degrees of freedom. The number of additional predator–prey k_{ij} 's was determined by their $AICc$ score (note this approach was only carried out for the Irish Sea model in this study).

Alternate fitting approaches led to the emergence of different vulnerability multipliers in the corresponding models of best fit (as determined by sum of squared deviations and $AICc$) for the Irish Sea (Fig. 7) and NWACS-MICE (Fig. 8). Different fitting approaches impacted estimates of F_{MSY} in both models due to changes in species sensitivity to F with alternate vulnerability multipliers. Despite the increased complexity of the

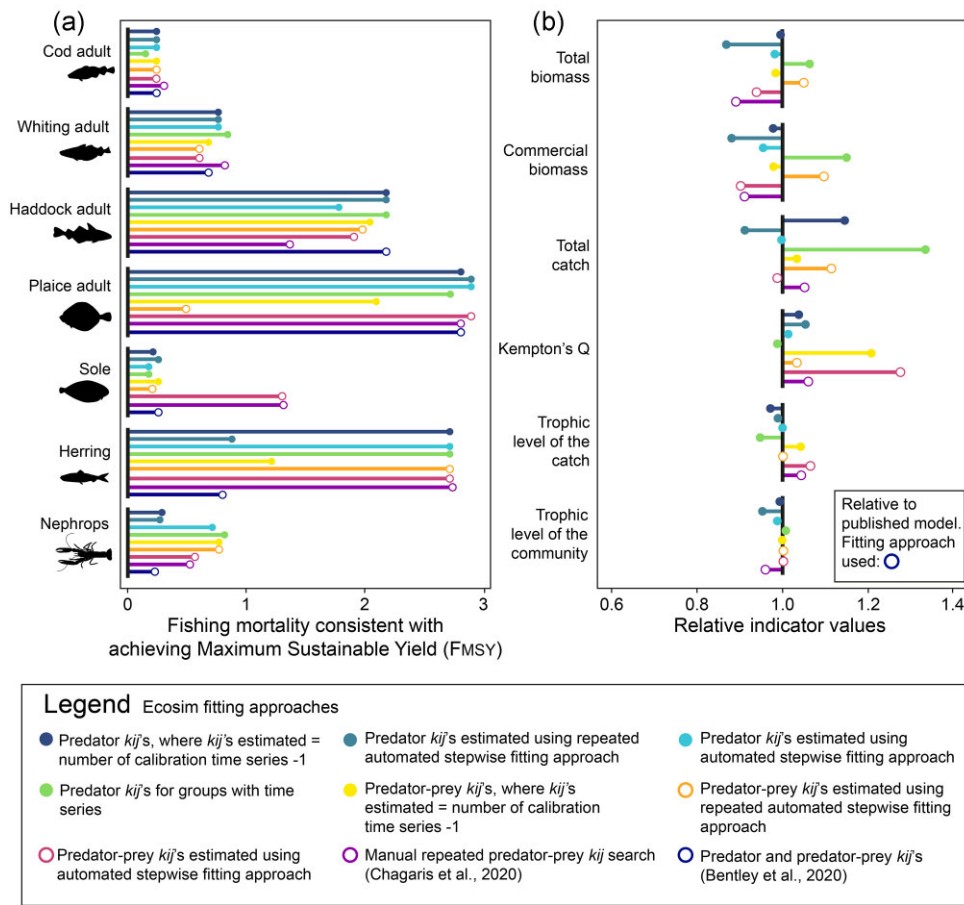


Figure 7 Irish Sea EwE outputs under alternate fitting approaches. Vulnerability multipliers for the Irish Sea Ecosim model were estimated following seven alternate fitting approaches. The impacts of alternate fitting approaches and vulnerability multiplier estimates are shown for (a) estimates of F_{MSY} and (b) indicators of ecosystem structure and function. The impacts of vulnerability multiplier estimates on indicator simulations are illustrated by comparing new simulations against the simulations from the published model. The published Irish Sea model has vulnerability multiplier values estimated using the predator and predator-prey approaches (Bentley et al. 2020).

Irish Sea model, the patterns in F_{MSY} variability are comparable between models, with certain species having consistent F_{MSY} estimates across the nine approaches for vulnerability multiplier estimation. This includes cod and whiting for the Irish Sea and striped bass for the NWACS. As demonstrated in Case Study 1, these species are opportunistic feeders, which are also predated on by higher trophic levels, giving them a relatively high degree of connectivity within the food web models, which may constrict the emergence of vulnerability multipliers. Additionally, these groups have experienced a period of collapse and, in some cases, recovery, which provides contrast for the model to estimate the vulnerability multipliers. Fitting approaches with similar properties resulted in more closely related F_{MSY} estimates (Supplementary Fig. 3). For example, F_{MSY} estimates generated by approaches that searched for vulnerability multipliers by “predator” tended to be more similar to each other when compared to those generated by approaches, which searched by “predator-prey,” and vice versa. This emergent trend is perhaps most clearly seen in the F_{MSY} estimates for menhaden and bluefish adults from the NWACS-MICE model (Fig. 8).

We investigated how the approach used to estimate vulnerability multipliers impacted derived ecosystem indicators (Figs

7b and 8b). These impacts were relatively small, with most deviations being within the range of 5–10% when compared to indicators from the published models. Trophic level indicators were particularly robust across estimation approaches, despite often larger differences being observed in diversity (Kempton's Q) catch and commercial biomass indicators. Balanced reconfiguration within the ecosystem models (i.e. increases in some species and decreases in others with similar trophic levels) enabled the trophic indicators to remain similar across approaches. However, the dissimilarity in trophic level of the catch in the NWACS-MICE model was generally higher across scenarios where vulnerability multipliers were searched by “predator.” This reflects the higher F_{MSY} reference points for adult weakfish and bluefish and lower F_{MSY} reference points for menhaden and herring produced under the same fitting approaches. Overall, the Irish Sea EwE model showed greater dissimilarity in derived indicators than the NWACS-MICE model. This outcome is likely linked to the increased complexity of the Irish Sea model, and how a repeated search provides the opportunity to adjust more predator-prey vulnerability multipliers. This may be less of a concern for low-complexity models as the parameter space is smaller, increasing the likelihood that the same vulnerability multipliers will be adjusted.

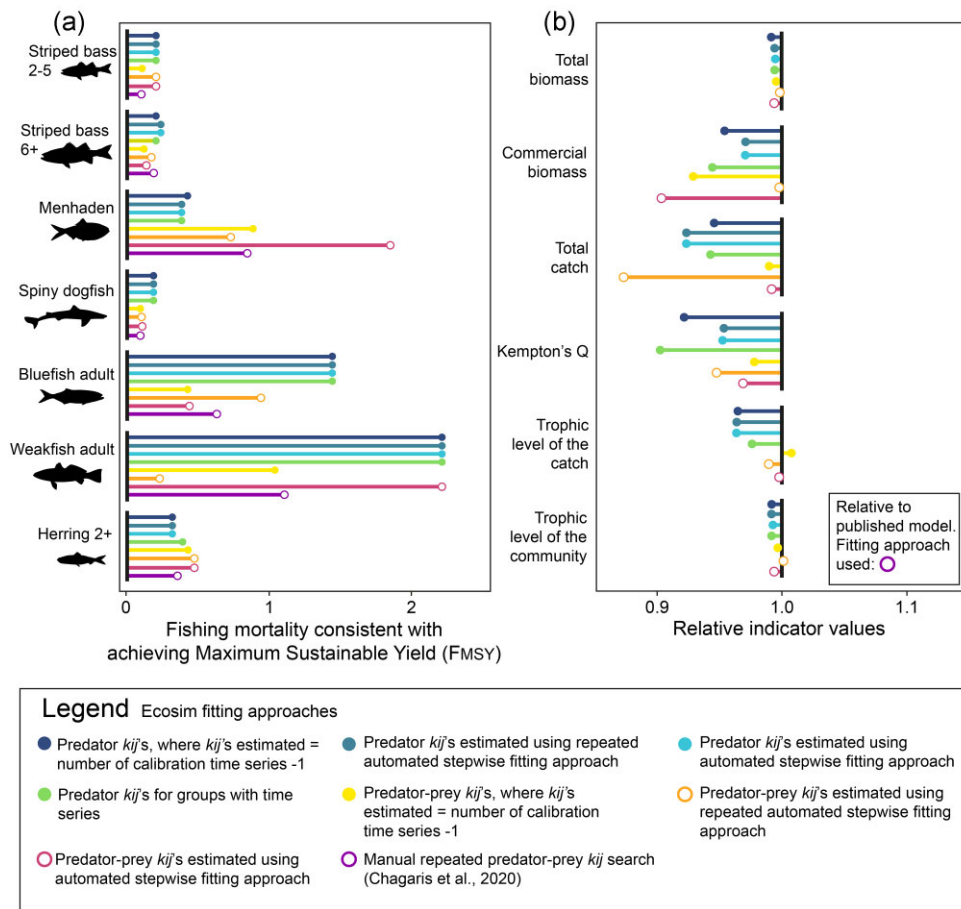


Figure 8 NWACS-MICE EwE outputs under alternate fitting approaches. Vulnerability multipliers for the NWACS-MICE Ecosim model were estimated following seven alternate fitting approaches. The impacts of alternate fitting approaches and vulnerability multiplier estimates are shown for (a) estimates of F_{MSY} and (b) indicators of ecosystem structure and function. The impacts of vulnerability multiplier estimates on indicator simulations are illustrated by comparing new simulations against the simulations from the published model. The published NWACS-MICE model has vulnerability multiplier values estimated using the manual repeated predator–prey vulnerability multiplier search approach (Chagaris et al. 2020).

Discussion

Limitations and future development

By describing the estimation of vulnerability multipliers and visualizing their impact on model simulations, this paper attempts to demystify the use of vulnerability multipliers, advance EwE uncertainty analyses, and contribute to our understanding of the importance of such parameters for model simulations and projections (Payne et al. 2016). Understanding and acknowledging the underlying assumptions of vulnerability multipliers, and the risk of overfitting, is a crucial step in the pursuit of operational and defensible EwE models that can be used to inform marine management decisions and objectives. Currently, this is an aspect of EwE modeling which is not always appropriately scrutinized, risking overconfidence in simulations and missed opportunities to convey model uncertainty and direct future research. We expect that one reason for the limited assessment of vulnerability multipliers is the general complexity of the concept, and how this has impacted the ways in which their use, purpose, and meaning have been diversely interpreted by the global community.

Challenges associated with parameter estimation are not unique to EwE or other ecosystem models. Complex single-species stock assessment models that integrate multiple data types under a statistical framework also struggle to estimate

key parameters that determine stock productivity, such as those describing the stock–recruit curve (Conn et al. 2010, Lee et al. 2012, Mangel et al. 2013). Yet these models are routinely used for managing fish stocks around the world. We have shown that, given the calibration approach used (i.e. by estimating predator vulnerability multipliers), EwE can return reasonable parameters. The complexity of ecosystem models should thus not limit their utility to management, considering that they are used appropriately in recognition of their tactical limitations, they can fit the data reasonably well, common sense checks are performed, and uncertainty is appropriately communicated.

It is not enough to estimate vulnerability multipliers and assume those that produce the best statistical hindcast fit are appropriate. Ecological reasoning and hypothesis testing must support statistical inference as it should when balancing Ecopath models (Link 2010), estimating primary production anomalies (e.g. Serpetti et al. 2017), and incorporating environmental drivers (e.g. Mackinson 2014). Part of this process should include a critical evaluation of calibration time series, as these, and their associated uncertainty, drive the statistical estimation of vulnerability multipliers. Using data with inherent inconsistencies will lead to variable and potentially biased estimates. Equally important is the lack of missing reference time series. Time series produce constraints, and when esti-

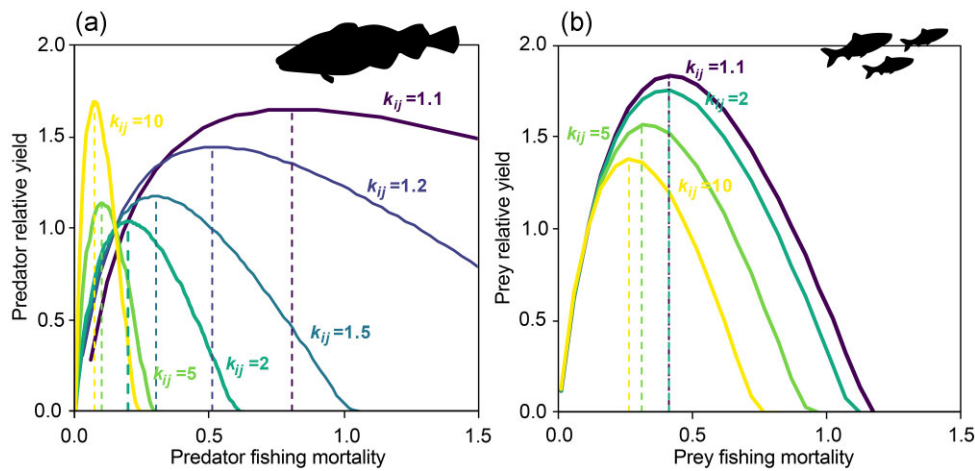


Figure 9 Effects of vulnerability multipliers on derived sustainable fishing advice. Estimations of F_{MSY} (Walters et al. 2005) are influenced by the value of vulnerability multipliers (k_{ij}). Changing predator k_{ij} values influences the F_{MSY} reference points for (a) the predator and (b) its prey. Dashed lines indicated the estimated F_{MSY} reference points for each k_{ij} scenario.

mating vulnerability multipliers for groups without time series, the lack of constraints allows the fitting procedure to explore a broad parameter space to let such groups indirectly impact other groups with time series. All of the above can impact model-derived management advice.

As shown in our case studies, F_{MSY} estimates can change in response to estimated vulnerability multipliers and their impacts on predator consumption rates, albeit with most changes being relatively conservative (Supplementary Fig. 3). With high vulnerability multiplier values, species are more sensitive to changes in F as deviations from the baseline Ecospace mortality rate have a greater impact on production rates. Biomass therefore declines at an increased rate as fishing mortality increases (and vice versa) when vulnerability multipliers are high, meaning maximum sustainable yields are achieved at lower fishing mortalities (Fig. 9a). As predator consumption rates increase with higher vulnerability multiplier values, prey experience higher predation rates, reducing the yield that can be obtained by fishing at F_{MSY} (Fig. 9b). Unreliable vulnerability multipliers are not easily apparent when reviewing model hindcast simulations against calibration time series data. Comparing Ecosim F_{MSY} to other estimates, or proxies (e.g. natural mortality), is one approach to assess vulnerability multipliers and has been demonstrated for the ICES key-run models of the North Sea (ICES 2016) and Irish Sea (ICES 2019b). Simulating models beyond observations, under alternate fishing or environmental scenarios, can also highlight issues with vulnerability multipliers by exploring group sensitivities and whether simulated responses to change fall outside of what might be considered ecologically reasonable. Future developments should also consider dependencies between vulnerability multipliers, whether correlation exists between vulnerability multipliers, and how this may impact the ability of a search routine to find stable solutions.

For EwE models to be of operational use, it should be possible to explain why estimated vulnerability multipliers are realistic. This could be based on knowledge of species ecology, carrying capacity, or natural mortality. We envisage that the future development of Ecosim will encourage users to think more critically when calibrating models by building options to restrict the statistical vulnerability optimization routine. The objective of this would be to enable users to con-

strain vulnerability multiplier estimation using *a priori* knowledge where, importantly, data is available to justify doing so. Increased control over the search for vulnerability multipliers could be used to set upper and lower parameter limits, or limits determined by carrying capacity, and penalize parameter combinations that operate outside of predefined limits. Such constraints would also have important implications for Ecospace, the spatial-temporal component of EwE. High vulnerability multipliers, and the large increases in predation mortality that they enable, can have disproportionately large impacts in Ecospace when prey are restricted to small areas (as predators are able to deplete them rapidly). Vulnerability multipliers in Ecospace require further consideration given how spatial heterogeneity may impact species physiology, habitat carrying capacity, and predator–prey interaction rates. Spatial considerations are implicit within the vulnerability concept and enable spatial considerations to be integrated indirectly into Ecosim. The necessity for vulnerability multipliers, or at least those in Ecosim, which go some way to indirectly recognizing spatial heterogeneity, may be negated by the explicit consideration of spatial heterogeneity in Ecospace. Alternate vulnerability multiplier combinations may be needed depending on the priority use of Ecosim or Ecospace and the different mechanistic role vulnerability multipliers may play across the two components.

Recommendations

Calibration methods for EwE are not prescriptive. Any one of the methods included in Fig. 2, or even new methods, may be suitable for use if they can be justified. That said, Case Study 1 showed that predator vulnerability multipliers are more likely to re-emerge than predator–prey vulnerability multipliers, and that re-emergence is impacted by data quality. Below, we have provided best practice recommendations to evaluate the appropriateness of vulnerability multipliers and their impact on model uncertainty:

- **Recommendation 1: Limit the number of vulnerability multipliers to be estimated.** The most efficient way to limit the number of parameters is to estimate by predator, add individual predator–prey combinations, if you have specific arguments for why this is necessary. Avoid

estimating vulnerabilities for groups without time series as the lack of constraints can lead to unrealistic estimates.

- **Recommendation 2: Explain vulnerability multipliers.** Provide justifications for setting initial vulnerability multipliers (or keeping the default). If estimating vulnerability multipliers using a statistical routine, check if they make sense relative to the exploitation and ecology of the predator and the ecology of the predator–prey interaction.
- **Recommendation 3: Sense check carrying capacities.** Vulnerability multipliers augment the upper limit for predator consumption rates, which dictates how predators respond to changes in mortality rates (e.g. release from fishing pressure or predation) or in prey biomass. It is important to review how predator biomass responds to such changes and critically evaluate whether the changes are plausible and whether the limits of estimates should be constrained (i.e. setting upper and lower limits).
- **Recommendation 4: Look beyond goodness of fit when evaluating model performance.** Combinations of vulnerability multipliers that achieve the best statistical fit (i.e. SS and AICc) do not necessarily produce the “best” model, if other model outputs, such as indicators, F_{MSY} reference points, and forward projections, are unlikely. Assessment of wider model performance should be undertaken to review vulnerability multipliers.
- **Recommendation 5: Perform vulnerability multiplier sensitivity analyses.** It is best practice to acknowledge and communicate model uncertainty. Calibrating Ecosim models, and thereby choosing one of multiple approaches to estimate vulnerability multipliers, introduces additional uncertainty into the process. Exploring model performance under alternate calibration approaches tests the sensitivity of model outputs to changes in vulnerability multipliers and identifies which vulnerability multipliers consistently emerge.

Summary and conclusions

Decision makers are being increasingly challenged to find solutions to complex problems, which require the assessment of risk, trade-offs, and cumulative pressures in the marine environment (Punt 2017, Piet et al. 2020, Pedreschi et al. 2021, Rodriguez Perez et al. 2023). EwE is actively being used to support both strategic and tactical management decisions, which address the cumulative impacts of multiple pressures and needs of various stakeholders (Craig and Link 2023, Karp et al. 2023). Given the growing operational use of EwE and other process-based models, it is of mounting importance that best practices are regularly revisited and developed to build our understanding of bias within models and uncertainty in model outputs. This is key to the wider acceptance and pragmatic use of EwE and other ecosystem modeling approaches. Calibration procedures for complex process-based models require particular attention and scrutiny given the sensitivity of model outputs and advice to alternate sets of calibration parameters. The way in which calibration procedures are communicated and understood can act as a barrier to their effective implementation and dissemination. It was our intention here to highlight the existing routes for parameter estimation and frame them in the context of the technical and

ecological mechanisms and theories that can guide their justified use. Through the case studies presented in this paper, we have demonstrated that assessing model calibration based on similarity between model simulations and calibration data alone is not adequate, and that statistical optimization routines should be complemented with a judgmental evaluation grounded in ecological theory and lessons, which can be learnt from available data and knowledge of the study region.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflict of interest: The authors have no conflict of interest to declare.

Author contributions

All authors contributed to the design of the manuscript. J.W.B. and D.C. carried out the modeling work and analyzed the data. All authors contributed to the writing, reviewing, and editing of the manuscript.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Funding

J.B. would like to acknowledge the EU Horizon project OceanICU (grant agreement no. 101083922). J.J.H. and V.C. would like to acknowledge the EU Horizon 2020 project EcoScope: ecocentric management for sustainable fisheries and healthy marine ecosystems project (grant agreement no. 101000302). M.C. would like to acknowledge EU Horizon project Ges4Seas (grant agreement no. 101059877), the Spanish project ProOceans (Plan Estatal de Investigación Científica, Técnica y de Innovación 2020, PID2020-118097RB-I00), and the “Severo Ochoa Centre of Excellence” accreditation (CEX2019-000928-S). V.C. also acknowledges NSERC Discovery Grant RGPIN-2019-04901.

References

- Ahrens RNM, Walters CJ, Christensen V. Foraging arena theory. *Fish Fisher* 2012;13:41–59. <https://doi.org/10.1111/j.1467-2979.2011.0432.x>
- Akaike H. Information theory and an extension of the maximum likelihood principle. In: *Selected papers of Hirotugu Akaike*. New York, NY: Springer, 1998, 199–213.
- Alexander KA, Heymans JJ, Magill S *et al.* Investigating the recent decline in gadoid stocks in the west of Scotland shelf ecosystem using a foodweb model. *ICES J Mar Sci* 2015;72:436–49. <https://doi.org/10.1093/icesjms/fsu149>
- Andersen KP, Ursin E. A multispecies extension to the Beverton and Holt theory of fishing, with accounts of phosphorus circulation and primary production. *Meddelelser fra Danmarks Fiskeri-og Havundersogelser* 1978;7:319–435.
- Bentley JW, Lundy MG, Howell D *et al.* Refining fisheries advice with stock-specific ecosystem information. *Fron Mar Sci* 2021;8:602072. <https://doi.org/10.3389/fmars.2021.602072>
- Bentley JW, Serpetti N, Fox CJ *et al.* Retrospective analysis of the influence of environmental drivers on commercial stocks and fishing

- opportunities in the Irish Sea. *Fisheries Oceanogr* 2020;29:415–35. <https://doi.org/10.1111/fog.12486>
- Bentley JW, Serpetti N, Heymans JJ. Investigating the potential impacts of ocean warming on the Norwegian and Barents Seas ecosystem using a time-dynamic food-web model. *Ecol Modell* 2017;360:94–107. <https://doi.org/10.1016/j.ecolmodel.2017.07.002>
- Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res* 2004;33:261–304.
- Cavanaugh JE, Neath AA. The Akaike information criterion: background, derivation, properties, application, interpretation, and refinements. *WIREs Comput Stat* 2019;11:e1460. <https://doi.org/10.1002/wics.1460>
- Celić I, Libralato S, Scarcella G et al. Ecological and economic effects of the landing obligation evaluated using a quantitative ecosystem approach: a Mediterranean case study. *ICES J Mar Sci* 2018;75:1992–2003. <https://doi.org/10.1093/icesjms/fsy069>
- Chagaris D, Drew K, Schueller A et al. Ecological reference points for Atlantic menhaden established using an ecosystem model of intermediate complexity. *Fron Mar Sci* 2020;7:606417. <https://doi.org/10.3389/fmars.2020.606417>
- Cheung WL, Watson R, Pitcher TJ. Policy simulation on the fisheries of Hong Kong marine ecosystem. *Fish Cent Res Rep* 2002;10, 46–53.
- Christensen V, Coll M, Buszowski J et al. The global ocean is an ecosystem: simulating marine life and fisheries. *Global Ecol Biogeogr* 2015;24:507–17. <https://doi.org/10.1111/geb.12281>
- Christensen V, Coll M, Steenbeek J et al. Representing variable habitat quality in a spatial food web model. *Ecosystems* 2014;17:1397–412. <https://doi.org/10.1007/s10021-014-9803-3>
- Christensen V, Walters CJ. Ecosim with Ecosim: methods, capabilities and limitations. *Ecol Modell* 2004;172:109–39. <https://doi.org/10.1016/j.ecolmodel.2003.09.003>
- Christensen V, Walters CJ. Progress in the use of ecosystem modelling for fisheries management. In: V Christensen, JL Maclean (Eds.) *Ecosystem Approaches to Fisheries: A Global Perspective*. Cambridge: Cambridge University Press, 2011, 189–205.
- Christensen V, Walters CJ. 2023. *Ecosystem Modelling with EwE*. Victoria, BC: Bccampus. <https://pressbooks.bccampus.ca/ewemodel/>
- Christie DR. Implementing an ecosystem approach to ocean management: an assessment of current regional governance models. Ocean ecosystem management: challenges and opportunities for regional ocean Governance. *Duke Environ Law Policy Forum* 2005;16:117–42. <https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1087&context=delpf>
- Conn PB, Williams EH, Shertzer KW. When can we reliably estimate the productivity of fish stocks?. *Can J Fish Aquat Sci* 2010;67:511–23. <https://doi.org/10.1139/F09-194>
- Corrales X, Coll M, Ofir E et al. Hindcasting the dynamics of an eastern Mediterranean marine ecosystem under the impacts of multiple stressors. *Mar Ecol Prog Ser* 2017;580:17–36. <https://doi.org/10.3354/meps12271>
- Corrales X, Coll M, Ofir E et al. Future scenarios of marine resources and ecosystem conditions in the eastern Mediterranean under the impacts of fishing, alien species and sea warming. *Sci Rep* 2018;8:14284. <https://doi.org/10.1038/s41598-018-32666-x>
- Craig JK, Link JS. It is past time to use ecosystem models tactically to support ecosystem-based fisheries management: case studies using Ecosim with Ecosim in an operational management context. *Fish Fish* 2023;24:381–406. <https://doi.org/10.1111/faf.12733>
- Dill LM, Heithaus MR, Walters CJ. Behaviorally mediated indirect interactions in marine communities and their conservation implications. *Ecology* 2003;84:1151–7. <https://www.jstor.org/stable/3107923>
- Gliwicz MZ. Predation and the evolution of vertical migration in zooplankton. *Nature* 1986;320:746–8. <https://doi.org/10.1038/320746a0>
- Green SJ, Côté IM. Trait-based diet selection: prey behaviour and morphology predict vulnerability to predation in reef fish communities. *J Anim Ecol* 2014;83:1451–60. <https://doi.org/10.1111/1365-2656.12250>
- Heymans JJ, Coll M, Link JS et al. Best practice in Ecosim with Ecosim food-web models for ecosystem-based management. *Ecol Modell* 2016;331:173–84. <https://doi.org/10.1016/j.ecolmodel.2015.12.007>
- Hilborn R, Walters CJ. Role of stock assessment in fisheries management. In: *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Boston, MA: Springer US, 1992, 3–21.
- Holling CS. The components of predation as revealed by a study of small-mammal predation of the European Pine Sawfly1. *Can Entomol* 1959;91:293–320. <https://doi.org/10.4039/Ent91293-5>
- Howell D, Schueller AM, Bentley JW et al. Combining ecosystem and single-species modelling to provide ecosystem-based fisheries management advice within current management systems. *Fron Mar Sci* 2021;7:607831. <https://doi.org/10.3389/fmars.2020.607831>
- ICES. *Report of the Working Group on Multispecies Assessment Methods (WGSAM)*, 9–13 November 2015, Woods Hole, MA: ICES CM 2015/SSGEPI, 2016, 20, 206pp.
- ICES. *Report of the Workshop on operational EwE models to inform IEAs (WKEW-IEA)*. 26–30 November 2018, Barcelona: ICES CM 2018/IEASG, 2018, 16, 32pp.
- ICES. Workshop on the design and scope of the 3rd generation of ICES Ecosystem Overviews (WKEO3). *ICES Sci Rep* 2019a;1:46. <https://doi.org/10.17895/ices.pub.5445>
- ICES. Working group on multispecies assessment methods (WGSAM). *ICES Sci Rep* 2019b;1:320. <https://doi.org/10.17895/ices.pub.5758>
- ICES. Workshop on an Ecosystem Based Approach to Fishery Management for the Irish Sea (WKIrish6; outputs from 2019 meeting). *ICES Sci Rep* 2020;2:32. <https://doi.org/10.17895/ices.pub.5551>
- ICES. 2022. Celtic Seas Ecoregion—Ecosystem overview. In Report of the ICES Advisory Committee, 2022. ICES Advice 2022, Section 7.1, <https://doi.org/10.17895/ices.advice.21731615>
- ICES. Workshop on Assessing CAPacity to supply Ecosystem Services (WKASCAPES; outputs from 2022 meeting). *ICES Sci Rep* 2023;5:25. <https://doi.org/10.17895/ices.pub.22842767>
- Jolles JW, Sosna MM, Mazué GP et al. Both prey and predator features predict the individual predation risk and survival of schooling prey. *eLife* 2022;11:e76344. <https://doi.org/10.7554/eLife.76344>
- Karp MA, Link JS, Grezlik M et al. Increasing the uptake of multispecies models in fisheries management. *ICES J Mar Sci* 2023;80:243–57. <https://doi.org/10.1093/icesjms/fsad001>
- Kempton RA, Taylor LR. Models and statistics for species diversity. *Nature* 1976;262:818–20. <https://doi.org/10.1038/262818a0>
- Kluger LC, Taylor MH, Mendo J et al. Carrying capacity simulations as a tool for ecosystem-based management of a scallop aquaculture system. *Ecol Modell* 2016;331:44–55. <https://doi.org/10.1016/j.ecolmodel.2015.09.002>
- Ladich F. Shut up or shout loudly: predation threat and sound production in fishes. *Fish Fish* 2022;23:227–38. <https://doi.org/10.1111/faf.12612>
- Laevastu T, Larkins HA. *Marine Fisheries Ecosystems: its Quantitative Evaluation and Management*. Farnham: Fishing News Books Ltd., 1981.
- Lee HH, Maunder MN, Piner KR et al. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models?. *Fish Res* 2012;125–126:254–61. <https://doi.org/10.1016/j.fishres.2012.03.001>
- Link JS. Adding rigor to ecological network models by evaluating a set of pre-balance diagnostics: a plea for PREBAL. *Ecol Modell* 2010;221:1580–91. <https://doi.org/10.1016/j.ecolmodel.2010.03.012>
- Mackinson S. Combined analyses reveal environmentally driven changes in the North Sea ecosystem and raise questions regarding what makes an ecosystem model's performance credible? *Can J Fish Aquat Sci* 2014;71:31–46. <https://doi.org/10.1139/cjfas-2013-0173>
- Mackinson S, Platts M, Garcia C et al. Evaluating the fishery and ecological consequences of the proposed North Sea multi-annual plan. *PLoS One* 2018;13:e0190015. <https://doi.org/10.1371/journal.pone.0190015>

- Mangel M, MacCall AD, Brodziak J *et al.* A perspective on steepness, reference points, and stock assessment. *Can J Fish Aquat Sci* 2013;70:930–40. <https://doi.org/10.1139/cjfas-2012-0372>
- Marshak AR, Link JS, Shuford R *et al.* International perceptions of an integrated, multi-sectoral, ecosystem approach to management. *ICES J Mar Sci* 2017;74:414–20. <https://doi.org/10.1093/icesjms/fsw214>
- Morison F, Harvey E, Franzè G *et al.* Storm-induced predator–prey decoupling promotes springtime accumulation of North Atlantic phytoplankton. *Fron Mar Sci* 2019;6:608. <https://doi.org/10.3389/fmars.2019.00608>
- Nunn AD, Tewson LH, Cowx IG. The foraging ecology of larval and juvenile fishes. *Rev Fish Biol Fish* 2012;22:377–408. <https://doi.org/10.1007/s11160-011-9240-8>
- Payne MR, Barange M, Cheung WW *et al.* Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES J Mar Sci* 2016;73:1272–82. <https://doi.org/10.1093/icesjms/fsv231>
- Pedreschi D, Vigier A, Höffle H *et al.* Innovation through consultation: stakeholder perceptions of a novel fisheries management system reveal flexible approach to solving fisheries challenges. *Mar Policy* 2021;124:104337. <https://doi.org/10.1016/j.marpol.2020.104337>
- Pethybridge HR, Weijerman M, Perryman H *et al.* Calibrating process-based marine ecosystem models: an example case using Atlantis. *Ecol Modell* 2019;412:108822. <https://doi.org/10.1016/j.ecolmodel.2019.108822>
- Piet G, Delacámara G, Kraan M *et al.* Advancing aquatic ecosystem-based management with full consideration of the social-ecological system. In: *Ecosystem-Based Management, Ecosystem Services and Aquatic Biodiversity: Theory, Tools and Applications*. New York, NY: Springer, 2020, 17–37. https://doi.org/10.1007/978-3-030-45843-0_2
- Plagányi ÉE, Punt AE, Hillary R *et al.* Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. *Fish and Fisheries* 2014;15(1):1–22. <https://doi.org/10.1111/j.1467-2979.2012.00488.x>
- Punt AE. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. *ICES J Mar Sci* 2017;74:499–510. <https://doi.org/10.1093/icesjms/fsv193>
- Püts M, Kempf A, Möllmann C *et al.* Trade-offs between fisheries, offshore wind farms and marine protected areas in the southern North Sea—winners, losers and effective spatial management. *Mar Policy* 2023;152:105574. <https://doi.org/10.1016/j.marpol.2023.105574>
- Rehren J, Coll M, Jiddawi N *et al.* Evaluating ecosystem impacts of gear regulations in a data-limited fishery—comparing approaches to estimate predator–prey interactions in Ecopath with Ecosim. *ICES J Mar Sci* 2022;79:1624–36. <https://doi.org/10.1093/icesjms/fsac077>
- Rodriguez Perez A, Tsikliras AC, Gal G *et al.* Using ecosystem models to inform ecosystem-based fisheries management in Europe: a review of the policy landscape and related stakeholder needs. *Fron Mar Sci* 2023;10:1196329. <https://doi.org/10.3389/fmars.2023.1196329>
- Rosenberg AA, Bolster WJ, Alexander KE *et al.* The history of ocean resources: modeling cod biomass using historical records. *Fron Ecol Environ* 2005;3:78–84. [https://www.jstor.org/stable/3868514\(?PMU?\)](https://www.jstor.org/stable/3868514(?PMU?))
- Sbragaglia V, Jolles WJ, Coll M *et al.* Fisheries-induced changes of shoaling behaviour: mechanisms and consequences. *Trends Ecol Evol* 2021;36:885–8.
- Schoener TW. Theory of feeding strategies. *Annu Rev Ecol Syst* 1971;2:369–404. <https://doi.org/10.1146/annurev.es.02.110171.02101>
- Scott E, Serpetti N, Steenbeek J *et al.* A stepwise fitting procedure for automated fitting of Ecopath with Ecosim models. *SoftwareX* 2016;5:25–30. <https://doi.org/10.1016/j.softx.2016.02.002>
- Serpetti N, Baudron AR, Burrows M *et al.* Impact of ocean warming on sustainable fisheries management informs the Ecosystem Approach to Fisheries. *Sci Rep* 2017;7:13438. <https://doi.org/10.1038/s41598-017-13220-7>
- Serpetti N, Benjamins S, Brain S *et al.* Modeling small scale impacts of multi-purpose platforms: an ecosystem approach. *Fron Mar Sci* 2021;8:694013. <https://doi.org/10.3389/fmars.2021.694013>
- Shannon L, Coll M, Bundy A *et al.* Trophic level-based indicators to track fishing impacts across marine ecosystems. *Mar Ecol Prog Ser* 2014;512:115–40. <https://doi.org/10.3354/meps10821>
- Steenbeek J, Buszowski J, Chagaris D *et al.* Making spatial-temporal marine ecosystem modelling better—a perspective. *Environ Model Softw* 2021;145:105209.
- Stock A, Murray CC, Gregr EJ *et al.* Exploring multiple stressor effects with Ecopath, Ecosim, and Ecospace: research designs, modeling techniques, and future directions. *Sci Total Environ* 2023;869:161719. <https://doi.org/10.1016/j.scitotenv.2023.161719>
- Tierney KM, Heymans JJ, Muir GK *et al.* Modelling marine trophic transfer of radiocarbon (¹⁴C) from a nuclear facility. *Environmental Model Softw* 2018;102:138–54.
- Townsend HM, Link JS, Osgood KE *et al.* *Report of the National Ecosystem Modeling Workshop (NEMoW)*. Silver Spring, MD: NOAA, 2008.
- Walters C, Christensen V, Pauly D. Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments. *Rev Fish Biol Fish* 1997;7:139–72. <https://doi.org/10.1023/A:1018479526149>
- Walters C, Pauly D, Christensen V *et al.* Representing density dependent consequences of life history strategies in aquatic ecosystems: ecoSim II. *Ecosystems* 2000;3:70–83. <https://doi.org/10.1007/s100210000011>
- Walters CJ, Christensen V, Martell SJ *et al.* Possible ecosystem impacts of applying MSY policies from single-species assessment. *ICES J Mar Sci* 2005;62:558–68. <https://doi.org/10.1016/j.icesjms.2004.12.005>
- Walters CJ, Juanes F. Recruitment limitation as a consequence of natural selection for use of restricted feeding habitats and predation risk taking by juvenile fishes. *CJFAS* 1993;50:2058–70.
- Walters CJ, Martell SJ. *Fisheries Ecology and Management*, Princeton, NJ: Princeton University Press, 2004.
- Wangersky PJ. Lotka–Volterra population models. *Annu Rev Ecol Syst* 1978;9:189–218. <https://doi.org/10.1146/annurev.es.09.110178.001201>
- Wilson KA, Able KW, Heck KL. Predation rates on juvenile blue crabs in estuarine nursery habitats: evidence for the importance of macroalgae (*Ulva lactuca*). *Mar Ecol Prog Ser* 1990;58:243–51.

Handling Editor: Mary Hunsicker