

Multispecies models for population dynamics: Progress, challenges and future directions

Jonatan Marquez¹, Stefan Vriend², Emily G. Simmonds², Marie V. Henriksen³, Lisa Sandal², Marlène Gamelon⁴, Christophe Coste⁵, Knut Hovstad⁶, and Aline Magdalena Lee¹

¹Norges teknisk-naturvitenskapelige universitet

²Norwegian University of Science and Technology

³Norwegian Institute of Bioeconomy Research

⁴Centre for Biodiversity Dynamics

⁵NTNU

⁶Norwegian Biodiversity Information Centre

April 20, 2022

Abstract

Understanding how population dynamics are influenced by species interactions and the surrounding community is crucial for addressing many ecological questions, but requires modelling of complex systems involving direct, indirect and often asymmetric species interactions. Progress in developing multispecies models that can tackle this task is being made in multiple subfields of ecology, often with varying approaches and end goals but also facing shared challenges. We review some of the main challenges and the ways in which they are being addressed, highlighting a wide variety of methods that can support the development of multispecies models for understanding population dynamics. The main challenges that we examine are estimation of species interactions from limited data, the necessity of simplifications, and handling uncertainty in complex, multispecies models. In addition to reviewing a wide variety of approaches and methods for dealing with these challenges, we discuss future directions and make suggestions for how we believe the development of multispecies models for understanding population dynamics can move forward more efficiently.

Multispecies models for population dynamics: Progress, challenges and future directions

Running title: Multispecies models for population dynamics

Jonatan F. Marquez^{1*}, Stefan J.G. Vriend¹, Emily G. Simmonds^{1,2}, Marie V. Henriksen³, Lisa Sandal¹, Marlène Gamelon^{1,4}, Christophe F.D. Coste¹, Knut Anders Hovstad^{1,5}, Aline M. Lee¹

¹ Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway.

² Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway.

³ Department of Landscape and Biodiversity, Norwegian Institute of Bioeconomy Research, 7031 Trondheim, Norway.

⁴ Université Lyon 1, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 69622 Villeurbanne, France

⁵The Norwegian Biodiversity Information Centre, 7446 Trondheim, Norway

Jonatan F. Marquez: jonatan.fredricson@gmail.com

Stefan J.G. Vriend: svriend@gmail.com

Emily G. Simmonds: emilygsimmonds@gmail.com

Marie V. Henriksen: marie.henriksen@nibio.no

Lisa Sandal: lisa.sandal@ntnu.no

Marlène Gamelon : marlene.gamelon@univ-lyon1.fr

Christophe F.D. Coste: christophe.f.d.coste@ntnu.no

Knut Anders Hovstad: knut.hovstad@artsdatabanken.no

Aline M. Lee: lee@alumni.ntnu.no

Keywords: Community, data integration, interspecific interaction, functional group, hybrid model, latent variable, model complexity, model uncertainty, population ecology, species associations.

Article type: Synthesis

Statement of authorship : AML conceived the research idea. All authors (JFM, SJGV, EGS, MVH, LS, MG, CFDC, KAH, AML) contributed to the literature search. AML and JFM reviewed and structured the findings. JFM led the writing of the manuscript with contributions from all other authors.

Data accessibility statement: No data was used.

Number of words in the main text: 7311

Number of words in the abstract: 164

Number of words in boxes: 264 and 261

Number of figures: 3

Number of boxes: 2

Number of references: 176

Corresponding author (*) : Jonatan F. Marquez, Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway; Tel.: +47 93040892; e-mail address: jonatan.fredricson@gmail.com

Abstract

Understanding how population dynamics are influenced by species interactions and the surrounding community is crucial for addressing many ecological questions, but requires modelling of complex systems involving direct, indirect and often asymmetric species interactions. Progress in developing multispecies models that can tackle this task is being made in multiple subfields of ecology, often with varying approaches and end goals but also facing shared challenges. We review some of the main challenges and the ways in which they are being addressed, highlighting a wide variety of methods that can support the development of multispecies models for understanding population dynamics. The main challenges that we examine are estimation of species interactions from limited data, the necessity of simplifications, and handling uncertainty in complex, multispecies models. In addition to reviewing a wide variety of approaches and methods for dealing with

these challenges, we discuss future directions and make suggestions for how we believe the development of multispecies models for understanding population dynamics can move forward more efficiently.

Introduction

Understanding how population dynamics are influenced by species interactions and surrounding communities is crucial for our understanding of the workings of whole communities and for our ability to better predict the dynamics of individual species (Pimm 1982; Marzloff *et al.* 2016). Over recent years, it has become apparent that single-species population models are often not sufficient to predict population dynamics in multispecies systems, and that management decisions based on such models can have detrimental consequences (Kinze & Punt 2009; Legović & Geček 2010; Engelhardt *et al.* 2020). Single-species models fail to adequately capture dynamics in many real systems because communities are composed of complex networks of interactions with continuous feedback effects. These feedbacks are ignored in single-species models, even when abundances of interacting species are included as covariates, limiting the realism and predictive potential of the models (Kissling *et al.* 2012). These limitations of single-species models are prompting an increased interest among ecologists and managers in developing multispecies models that improve understanding of the functioning of multispecies dynamics (Fulton *et al.* 2019), and that provide comprehensive information for effective ecosystem management (Daan & Sissenwine 1990; Plagányi *et al.* 2014).

There are two main approaches to developing multispecies models. The first approach builds up from the field of population ecology by joining single-species models into multispecies frameworks, incorporating species interactions. Conceptually, this approach can be traced back to the classical deterministic **Lotka-Volterra models** (Note: Bold terms in the text are explained in the Glossary and Multispecies model types glossary used to describe the population dynamics of pairs of predator-prey or competing species (Lotka 1925; Volterra 1928). These early models have since provided a basis for more complex models of food webs and competitors, incorporating several species and more realistic characteristics such as spatial dynamics, environmental variability and population structure (e.g. Roughgarden 1975; Holt & Lawton 1994; Amarasekare 2008; Gamelon *et al.* 2019; Rüger *et al.* 2019; Lee *et al.* 2020). In addition, several types of single-species population models aimed at quantifying population abundance and understanding drivers of population dynamics have recently been extended to multispecies versions to improve the understanding of the roles of interacting species (e.g. **integrated population models** (Péron & Koons 2012; Barraquand & Gimenez 2019), **integral projection models** (Adler *et al.* 2010; Kayal *et al.* 2018)), while the use of simulation models such as **individual based models** (Breckling *et al.* 2005; DeAngelis & Grimm 2014; Grimm *et al.* 2017) has also been shifted more towards multispecies modelling. Because models built up from single-species population models handle individual populations and their interactions explicitly, they tend to require large amounts of data and ecological knowledge, generally limiting their use to small subsets of species within a community.

The second major class of multispecies models focuses on understanding systems at the community level, including attributes such as community structure, biomass, energy flow, species richness and stability (Ulanowicz 1972; Pimm 1982; Tarnecki *et al.* 2016). Because these models focus on the dynamics of the community as a whole, they have historically tended to treat species as interchangeable and have not been concerned with the fate of specific species. Adding more species- or population-level detail to these models is a second approach to producing new multispecies models to understand population dynamics. Currently, most community-level models are unsuitable as predictors of population dynamics because they oversimplify population processes and make broad assumptions about the systems under study (Hollowed *et al.* 2000). However, they play an important role in identifying knowledge gaps and interactions (Plagányi 2007; Travers *et al.* 2007; Collie *et al.* 2016), and are often essential tools for addressing macroecological issues, such as consequences of harvesting, climate or habitat change (Pacifi *et al.* 2017) on community dynamics. By adding more population-level detail and mechanisms to such models, they can also be used to understand the dynamics of individual populations within communities (García-Callejas *et al.* 2018).

Thus, we are seeing a shift towards multispecies modelling of population dynamics, both by expanding single-

species models to incorporate more species and by adding more species-specific dynamics to community-level models. One can imagine an ultimate goal of these two approaches meeting in the middle, producing multispecies models that can both describe the dynamics of individual species and capture the complexity of whole communities. Recently, we have started to see **hybrid models** that could provide a first step in such a development by embedding a detailed population dynamic model within a community-level model allowing information to flow back and forth between the two (e.g. Breckling *et al.* 2005; Makler-Pick *et al.* 2011; Schmolke *et al.* 2019). Such hybrid models have shown potential for exploring ecological questions such as how ecosystem regime shifts affect the dynamics of a particular species (Gray & Wotherspoon 2015; Fulton *et al.* 2019), but they currently remain limited to simpler systems due to their complexity and insufficient species-specific data (Makler-Pick *et al.* 2011). Better collaboration between different modelling disciplines could speed up progress on this front and unlock more of the innate potential of hybrid modelling (Mokany *et al.* 2016).

In parallel, some complex multispecies frameworks and software have been developed for exploring specific questions, such as the effects of alternative harvesting strategies and protected areas, mainly in a fisheries context (Tjelmeland & Bogstad 1998; Pauly *et al.* 2000; Begley & Howell 2004). These frameworks typically use a combination of data, previously established parameter estimates and broadscale assumptions about processes, and are presented as pre-packaged software for exploring specific systems and questions. While such pre-packaged fisheries model frameworks have great value for their use, their context sets them somewhat apart from the more general modelling challenges discussed in this paper. We will therefore not focus on these types of model frameworks here, but refer our readers to the many reviews of these that already exist in the literature (Hollowed *et al.* 2000; Plagányi 2007; Collie *et al.* 2016).

In general, development of multispecies models incorporating population dynamics and species interactions faces numerous challenges. Natural systems are complex, containing many direct and indirect, often asymmetric and environmentally sensitive, interactions (Montoya *et al.* 2006; Morin 2011; Ovaskainen *et al.* 2017). This complexity, and the resultant high data demands, represents a major obstacle to achieving more realistic multispecies models capable of providing accurate mechanistic understanding of ecological processes. However, estimating the dynamics of a subset of species within a community carries the risk of omitting important processes, introducing biases, and limiting the usefulness of models for understanding population dynamics and projecting them into the future (Fath *et al.* 2007). In addition, uncertainty at different levels is propagated in complex models, causing a trade-off between biological realism and parameter certainty (Collie *et al.* 2016). It is therefore essential to find ways to simplify natural systems in models while capturing the relevant biological processes and minimizing the introduction of biases, increased uncertainty and inaccurate conclusions (Essington 2004; Berlow *et al.* 2009).

The current shift towards multispecies modelling is taking place across subfields of ecology simultaneously but often independently (Mooij *et al.* 2010). Consequently, many of the challenges inherent in multispecies modelling, such as estimation of species interactions, dealing with high model complexity or identification and measuring of model uncertainty, are being addressed in different ways, with little coordination (Mooij *et al.* 2010). In this paper, we review the ways in which the major challenges of developing multispecies models and estimating their parameters have been addressed in the literature. We hope that by bringing together different methods and approaches across fields, we can learn from each other and make faster progress towards more realistic multispecies models for population dynamics. We discuss constructive strategies to promote multispecies model development in light of promising techniques, as well as challenges and limitations.

Estimating interactions

Accurate understanding and representation of **interspecific interactions** is key for the development of robust multispecies models but challenging because of the sheer number of **potential interactions** in most natural systems and the many factors that can influence them. The number of potentially interacting species pairs in a system with S species is $S(S-1)/2$. However, many potential interactions never occur due

to differences in biological traits, such as morphology, size and phenology (i.e., **forbidden links**; Olesen *et al.* 2011; González-Varo & Traveset 2016). For example, nocturnal and diurnal species do not coincide in time and some pollinator's morphology prevents them from reaching the nectar of certain flowers (Olesen *et al.* 2011), reducing the number of interactions to be estimated for a given system. In contrast, other interactions are age-, sex- or phenotype dependent, and special consideration of population structures and complex life cycles are needed in the models, resulting in additional data requirements and more interaction terms (Lafferty *et al.* 2008; Strauss *et al.* 2017; Gamelon *et al.* 2019; Torres-Campos *et al.* 2020). Developing models that consider variation in phenotypes, behaviours, and demography is important for understanding changing interactions, and improving model projections and predictions (González-Varo & Traveset 2016).

The number of **realized interactions** is further influenced by local conditions affecting aspects of foraging biology and resource utilization, such as relative abundances, species assemblages and environmental conditions (Beckerman *et al.* 2006; Vázquez *et al.* 2009; Spiesman & Gratton 2016; Delmas *et al.* 2019). For example, studies of simple arthropod food webs have shown that more prey species do not necessarily increase the number of trophic interactions because predators tend to focus on preferred prey while ignoring the other (Torres-Campos *et al.* 2020). Similarly, increasing the number of potentially competing pollinator species has been shown to cause resource partitioning among pollinators, which can lead to divergence of floral traits that benefit the main pollinating species while hindering access to others (Temeles *et al.* 2016). The dynamic nature of species interactions presents a challenge for the development of realistic models aiming to capture large portions of communities or extrapolate from one time and place to another (Chamberlain *et al.* 2014; Gray & Wotherspoon 2015).

Several methods have been developed to identify and quantify species interactions, with variable suitability depending on the species and study system. Some methods assess species interactions directly, while others infer interactions indirectly based on **species associations** (i.e., inferred from species co-occurrences or correlated density dynamics). Direct observations in the field can provide valuable information about how species behave and interact under natural environmental conditions that correlational and comparative studies cannot. This can, however, be infeasible in species-rich systems and when species and their interactions are rare, weak and/or elusive (e.g. nocturnal, deep-sea, small or cryptic species) because the chance that both species and their interaction occur when the observer is present is low (Jordano 1987; Trijoulet *et al.* 2019). Manipulation experiments (e.g. enclosure, exclosure, population augmentation treatments) represent powerful tools to test hypotheses and better understand the mechanistic relationships between species but they can be costly and logistically challenging, especially when studying large or highly mobile species (Schmitz 2004; Wood *et al.* 2019). Here, we present additional alternative methods, discussing their potential and challenges.

Diet analysis

Feeding plays a central role in many species' interactions. Therefore, methods to understand what and how organisms eat are important tools for assessing interactions. Instead of observing an animal eat, information about the diet of a species can be obtained by analysing ingested or excreted material through stomach content and faecal analyses (Nielsen *et al.* 2018). Historically, these types of analyses have mostly been done visually, providing valuable information about the prey, such as size or life stage. However, visual examinations can also produce biases, for example between hard and soft body prey, since soft tissues dissolve faster (Nielsen *et al.* 2018). DNA analyses of the stomach content and faeces can greatly improve the identification of prey to species level, and are particularly useful when studying smaller organisms, such as insects (Titulaer *et al.* 2017; Horswill *et al.* 2018; Curtsdotter *et al.* 2019), but are less useful for identifying subgroups or traits within the prey species. The main limitations of stomach content and faecal analyses are that they provide diet information from relatively short periods (e.g. last foraging day), and might provide little information if the animal has not fed recently or if the prey's DNA degrades during digestion (Russell *et al.* 1992). Analyses of stable isotope ratios, mostly of nitrogen and carbon, and other biomarkers provide information about assimilated materials, thereby providing long-term dietary information from an individual (Nielsen *et al.* 2018). Stable isotope ratio analyses can identify a species' trophic position and

preferred foraging areas, but are often unable to quantify prey to the species level (e.g. Vander Zanden *et al.* 1999; Blois *et al.* 2013). Because no diet tracing method is bias free, there is a growing interest in developing methods that combine different techniques to benefit from the advantages and minimize the shortcomings (Nielsen *et al.* 2018). For example, using information from visual stomach content analyses as priors in Bayesian isotope mixing models allows accurate quantitative diet estimates (Chiaradia *et al.* 2014). Similarly, combining stable isotope, DNA and morphological analyses provides good estimates of prey diversity and subtle changes in trophic levels, while minimizing invasiveness and frequency of sampling (Horswill *et al.* 2018).

Time series correlations

A variety of methods have been developed to infer interactions indirectly from other types of commonly available data, such as counts, presence-absence data or species traits. When several species are monitored at the same site, the changes in one species' abundance or demographic rates (e.g. survival, reproduction) can be related to the population dynamics of the other species (Certain *et al.* 2018). In some cases, the predation or competition pressure of a species on others can be quantified by removing them in laboratory or field experiments and measuring the changes in the dynamics of others (Pacala & Silander 1990; Wilson & Tilman 1991; Wootton 2001). However, removing a species from an ecosystem is seldom feasible. Therefore, other methods classically used in demography can be used to infer interactions indirectly through long-term demographic data collected at the individual or population level, such as counts and capture-mark-recapture (CMR) data. For example, by relating age-specific breeding success and survival of chamois *Rupicapra rupicapra* to annual population counts of cohabiting red deer *Cervus elaphus*, competition for food resources was found to negatively influence chamois breeding success of primiparous and senescent females (Gamelon *et al.* 2020). Thus, individual-based long-term monitoring programs running on several interacting species at the same location can provide data to estimate species interactions, even in the absence of detailed direct observations of the interaction.

Joint species distribution models

Analogously, spatial patterns in species abundances can reveal associations between species. **Species distribution models** (SDMs) were originally developed to infer species habitat preferences from spatial abundance/occurrence data and environmental data (Kearney & Porter 2009). More recently, SDMs have been developed that model multiple species jointly (**joint species distribution models**; JSDMs). JSDMs assume that biotic interactions create non-random spatial patterns in occurrence or abundance. Therefore, by first accounting for the (dis)similarities in species' responses to spatial environmental patterns through an environmental covariance matrix, these models can reveal species' spatial associations from the residual covariance matrix (Ovaskainen & Abrego 2020). How well these species associations represent the true underlying species interactions in a JSDM depends on whether the important environmental covariates have been included in the model (Pollock *et al.* 2014; Dorazio *et al.* 2015; Ovaskainen *et al.* 2017; Zurell *et al.* 2018). In addition, because species associations can be influenced by common habitat preferences or migration patterns, additional information about the species, such as traits and phylogeny, are often used in combination with spatial data to estimate the probability of an interaction occurring (Morales-Castilla *et al.* 2015). Dormann *et al.* (2018) present a useful checklist to facilitate the interpretation of such estimates and avoid major pitfalls.

As long-term spatial datasets become increasingly available, approaches are also being developed that jointly account for spatial and temporal dynamics to estimate species interactions (Schliep *et al.* 2018). For example, using a multispecies competitive community dynamics framework, **time-series JSDM**s are being used to infer species interactions from species associations as functions of local species abundances in previous years and local environmental conditions (Mutshinda *et al.* 2011; Ovaskainen *et al.* 2017). By combining temporal and spatial information, time-series JSDMs have a high potential to provide more accurate estimates of underlying species interactions, since observed covariations are based on multiple points in time and space, rather than representing only a snapshot or a summary of all the dynamics in a large region (Ives *et al.*

2003; Ovaskainen *et al.* 2016, 2017; Thorson *et al.* 2016). However, spatiotemporal models are also inherently more complex, which makes them computationally challenging and less user-friendly (Norberg *et al.* 2019). Moreover, as with any inference method, model outputs are influenced by the amount, quality and spatial structure of the data. It is therefore important to always evaluate the results in light of ecological knowledge.

Trait-based approaches

Trait-based approaches have emerged in recent decades as useful methods to study community dynamics by characterizing individuals or species by key traits rather than as species with prescribed interactions (McGill *et al.* 2006; Degen *et al.* 2018; Kiørboe *et al.* 2018). Trait-based models have the potential to describe more sophisticated communities since they can reduce the number of parameters to be estimated (Kiørboe *et al.* 2018; Curtsdotter *et al.* 2019). Well-defined biological traits comparable across species, such as body size, mobility, or defense strategy, may provide useful information about an individual's mortality, growth, metabolism, and trophic role in a community. For example, larger competitors often exert a dominant competitive pressure over smaller ones (Kohyama 1992), predators preferentially feed on prey of a specific body mass relative to their own (Brose 2010; Kalinkat *et al.* 2011) and active-searching predators are more likely to encounter prey but likely to attract predators, compared to passive ambush predators (Kiørboe *et al.* 2018). These generalizations allow initial parameterization of interaction networks, even without direct data on species interactions. Currently, the majority of trait-based approaches are size-based (Kiørboe *et al.* 2018). Although the relationship between body mass and ecological interaction is well supported across many taxa (Pope *et al.* 2006; Hartvig *et al.* 2011; Boitet *et al.* 2012; Schneider *et al.* 2012; Curtsdotter *et al.* 2019), body size is not sufficient to describe the complex interactions of many systems (Jonsson *et al.* 2018; Curtsdotter *et al.* 2019; Keppeler *et al.* 2020). Further examination of key biological traits and how they interact with each other will therefore help improve the development of trait-based approaches. These general relationships seem unlikely to be accurate or common enough to replace more targeted estimation of interaction strength. Nonetheless, they can produce biologically plausible models (Brose 2010), and capture major patterns of population dynamics in some systems, and can be a useful starting point to help us fill in gaps in otherwise well-parameterized models.

Functional responses

Mathematically formulating the effects of realized interactions in a model of multispecies population dynamics entails estimating the influence of one species' population density on the population growth of the other (i.e., **functional response**). This function of population density is for simplicity often assumed to be linear. While linearity may adequately capture interspecific competition interactions, they may be less suited for modelling interactions among trophic levels, e.g. predator-prey dynamics (Certain *et al.* 2018). For example, functional responses can take different shapes depending on how a species searches, handles and processes prey, with potentially large effects on population dynamics (Spalinger & Hobbs 1992; Koen-Alonso 2007; Castillo-Alvino & Marvá 2020). Additionally, functional responses can vary between habitats and life stages, where a prey species might itself be an important competitor or even predator on young individuals of the predator species (Essington 2004), highlighting the importance of considering population structure. While assuming simple functional responses is a useful first step for multispecies models, it has been argued that relying strongly on them could hinder advances towards better mechanistic understanding of multispecies dynamics and limit their projectability into the future (Hunsicker *et al.* 2011; Kalinkat *et al.* 2011; Rosenbaum & Rall 2018). Assessing the effects of several types of functional responses on model outputs is one suggested solution, particularly because several functional response types can sometimes result in similarly well-fitting models of empirical data (Butterworth & Plagányi 2004; Koen-Alonso 2007; Kinzey & Punt 2009).

Data collection and utilization

Sampling design and technology

The development of multispecies models relies heavily on access to high-quality data, but how can data collection be improved to help estimation of species interactions and validation of multispecies models? Ideally, long-term multispecies data sets for a wide number of cohabiting species and environmental variables should be collected to ensure that the population dynamics and interactions are well covered under a wide range of conditions. In practice, the high costs and logistic challenges of acquiring such data sets have limited their availability to economically important species (e.g. harvested communities; ICES 2019) or simpler and mostly self-contained communities (e.g. islands, lakes; Christensen *et al.* 2013). However, studies on the effectiveness of different sampling designs have highlighted strategies through which data collection can be improved to the benefit of multispecies modelling (Lahoz-Monfort *et al.* 2014; Trijoulet *et al.* 2019; Zhang *et al.* 2020). In particular, there is evidence that sampling a greater number of sites at low intensity gives more representative system-wide estimates of interactions compared to sampling fewer sites at high intensity (Bogstad *et al.* 1995; Latour *et al.* 2003). Sampling a single site more rigorously easily results in biases towards the interactions occurring in that location, while sampling a greater range of sites gives a more representative picture of species interactions across its range.

In addition to finding ways to improve the efficiency of sampling strategies, recent technological developments have increased the quantity and quality of data for studying species interactions. Technologies that automate data collection, such as drones, GPS trackers, movement sensors, video/audio recorders and image recognition software (Weinstein 2015; Marvin *et al.* 2016), can reduce time and costs, thereby allowing greater sampling coverages (e.g. number of species, geographical area, higher spatial and temporal resolutions) and increasing the probability of recording an interaction. Increased environmental interest by the public has also led to the development of citizen science and crowdsourcing initiatives that can provide unprecedented amounts of data (Chandler *et al.* 2017; Devarajan *et al.* 2020). An example of this is the Global Biodiversity Information Facility (GBIF) which as of April 2022 has more than 1.9 billion species occurrence observations publicly available. However, citizen science data face limitations related to inconsistencies in sampling effort, sampling biases, and errors (Zipkin & Saunders 2018). Various statistical techniques are being developed and used to account for these sampling issues by, for example, modelling random effect and hierarchical structures (Kelling *et al.* 2015), but the majority of such data remains unused or limited to broader macroecological studies (Theobald *et al.* 2015; Heberling *et al.* 2021).

Data integration

As multispecies models tend to be data demanding, modelling methods that can simultaneously include and take full advantage of a variety of data sources are valuable. Such approaches have the potential to shorten the time series required to provide good estimates, improve the cost-effectiveness of monitoring programs, and improve the modelling of data-poor species. For instance, Barraquand & Gimenez (2019) found that combining data on capture-recapture, counts and reproduction to estimate dynamics of interacting multi-stage populations using **integrated community models** could provide accurate estimates of interactions, while also requiring shorter time series than studies using only count data. They evaluated the benefits of the different types of data to the model results and the costs of collecting these, and concluded that collecting reproduction data instead of capture-recapture data was a more cost-effective strategy, especially for abundant species (Barraquand & Gimenez 2019).

Integrating similar data types collected in different ways, such as abundance data through camera traps and transects, or citizen science and scientific surveys is a useful strategy in single-species population modelling (Besbeas *et al.* 2002; Lee *et al.* 2015; Zipkin & Saunders 2018; Isaac *et al.* 2020). Recent studies indicate that the advantages inherent in data integration methods in single-species models are also present in multispecies frameworks (Péron & Koons 2012; Fithian *et al.* 2015; Lahoz-Monfort *et al.* 2017; Barraquand & Gimenez 2019; Miller *et al.* 2019). Data integration methods allow models to maximize the information extracted from each dataset, while considering the weaknesses and strengths of each one (Miller *et al.*

2019). Similarly, combining data on similar species can improve the estimates of each species individually (Lahoz-Monfort *et al.* 2017), and of data-poor species in particular (Fithian *et al.* 2015). For instance, using SDM, Fithian *et al.* (2015) found that when faced with presence-only data for a species, using presence-only and presence-absence data from other species facilitated information sharing across species, which improved parameterization for the data-poor species by leveraging information from closely connected species. Such data sharing is an additional advantage of multispecies models over single species ones in many systems (Kindsvater *et al.* 2018).

Model structure and simplifications

We have discussed methods that are helpful for estimating or inferring large numbers of species interactions for multispecies models, as well as ways through which ecological data are becoming more detailed and increasingly available. However, translating the high complexity of most natural systems to models often leads to increased uncertainty, and difficulty parameterizing and interpreting the results. Therefore, even the most comprehensive multispecies models require some simplifications. Dynamic multispecies population models have historically started as simplified versions of the dynamics of a small subset of species (e.g., Lotka-Volterra model), onto which complexity was added in the form of, for example, life stages, spatial dynamics, or environmental influence. In contrast, community or network models aim to describe or understand entire or large parts of an ecosystem and therefore need to simplify the description of these communities. They tend to do this by finding ways to reduce the number of interactions that need to be estimated separately in the model without reducing model performance (Morin 2011; Collie *et al.* 2016).

One way to reduce the number of interactions to be estimated is by reducing the number of **nodes** in the model, i.e., the number of community components (Fig. 1). Aggregating species into groups based on taxonomic, trophic or/and ecological similarity (i.e., **trophospecies** or **functional groups**), can help the development of simplified community models that cover a large proportion of the **community** and yet maintain key properties of more complex models (Hood *et al.* 2006; Ulanowicz *et al.* 2014; Olivier & Planque 2017). This strategy has the added benefit of helping to understand and make better predictions of rare and data-poor species because it allows one to “borrow” information from common, closely related species, or species with similar traits that are likely to respond similarly to the environment, thereby increasing the sample size used to estimate the parameters of the node. Similarly, sampling error and stochasticity can have a smaller negative effect on the model predictability when species are grouped (Agarwal *et al.* 2021). However, since this approach regards various species as equal, information about individual species and their dynamics is lost (Simmons *et al.* 2019). Thus, the model outputs become sensitive to the criteria used to classify species, which is largely dependent on the research question (Fath *et al.* 2007; Pacifici *et al.* 2014). For example, a species classification based on taxonomical or ecological similarities might be better suited for addressing impacts of habitat change, while trophic similarities might be better suited for modelling harvesting impacts and energy flows. It is also important to assess the sensitivity of a model to different species classification criteria because differences in classification methods (e.g. cluster analyses, expert knowledge, model-based) can yield contrasting species groupings and model results (Picard *et al.* 2012; Olivier & Planque 2017).

Large communities can also be divided into subgroups or **modules** based on substructures within interaction networks (Olesen *et al.* 2007; Dormann & Strauss 2014; Fig. 1). Modules represent recurring non-random groupings of species within the community that interact more with each other than with species from other modules (Olesen *et al.* 2007). Identifying modules within ecosystems is therefore a good strategy to find subsets of species that can be modelled independently from the rest of the community (Allesina *et al.* 2005). Different modules within a system can then be used as nodes of a coarser community model. As a result of fewer interactions, the task of modelling large communities becomes more manageable. Similarly, the number of interactions can in some cases be reduced by identifying species with weak interactions with the rest of the community and omitting them from the model. For example, rare species are sometimes assumed to exert such weak competition or predation pressure in relation to common or dominant species that their influence is ignored (Canard *et al.* 2012). However, extensive research of the system may be required before making

such assumptions (Terry & Lewis 2020). Weak interactions can increase in importance over time (Terry et al. 2017) and, even if they remain weak, can still be important for maintaining the structure and stability of complex systems (McCann et al. 1998). Removal of weak interactions may therefore not always result in realistic model predictions.

Instead of grouping or omitting species to reduce the number of nodes the same functions can sometimes be used to describe different interactions and processes (e.g. growth, foraging, dispersal, reproduction, competitiveness) while only varying their parameterization to best represent each species (McDermot & Rose 2000; Reuter 2005; Buchmann *et al.* 2011; Grimm & Berger 2016). For example, some forest models use the same function to describe competitive interactions (e.g. based on vertical leaf area distribution) and same growth function, but adjust the growth parameters to each species (Kohyama 1992). Similarly, some models of fish communities assume trophic interactions between fish to be size-dependent and species-independent, thus, the same predation function can be used across species (Giacomini *et al.* 2013; González-Varo & Traveset 2016). This type of simplification is often used in **agent-based modelling**, also known as individual-based modelling (IBM) among ecologists when the agents represent individuals. IBMs can assign the same biological “behaviour” (i.e., growth model, interaction model, dispersal model, etc.) to individuals from different species or groups of species and efficiently simulate the complex dynamics of some interacting species within a community (DeAngelis & Grimm 2014). It has also been argued that, analogous to how modellers tend to use just a few well-established functional responses, there could be a small subset of well-established functions to describe other types of species behaviours that influence interspecific dynamics, like foraging or home range, with well-understood properties and requirements, thereby facilitating model development and communication (Grimm & Berger 2016).

Latent variable approaches can also be a useful way to reduce the dimensionality of multispecies models. Latent variables are unobserved variables that can be used to represent the main axis of (co)variation among species (Warton *et al.* 2015). For instance, in JSDMs with latent variable structures, all pairs of species associations or co-occurrences are modelled jointly by searching for the leading axes of variation unaccounted for by the environmental effects. This creates linear combinations of several variables, limiting the dimensionality of the multispecies data (Thorson *et al.* 2015; Ovaskainen & Abrego 2020; van der Veen *et al.* 2021). **Instructural equation modelling** (SEM), latent variables typically represent a theorized environmental effect measured by one or more indicator variables (Grace *et al.* 2010). In Bayesian network analysis, latent variables are used to group nodes with similar roles in the network and can thereby reduce the complexity of the modelled system (Kim *et al.* 2018). Latent variables can also be used to estimate interaction probabilities where nodes with similar latent positions in the network structure are assumed to be more likely to interact (Rohr *et al.* 2016; Kim *et al.* 2018).

While simplifications are a useful and necessary part of modelling, oversimplifications can lead to poorer model performance and loss of predictive power, especially under changing conditions (Raick *et al.* 2006; Berlow *et al.* 2009). Ideally, one would always compare the simplified models to more complete and complex models to assess their effectiveness and accuracy, as well as to identify the trade-offs of the simplifications (Raick *et al.* 2006). However, that would entail having abundant data to develop the complex models first, which is usually not an option. In practice, decisions on model structure and simplifications are often based on data availability instead of robust knowledge about ecological functionality (see e.g. Lafferty *et al.* 2008; Dunn *et al.* 2017). Because such decisions will continue to be necessary, especially in data-poor studies, it is important that the simplification methods, (e.g. criteria used to aggregate species or standardize links) are systematically documented to facilitate comparative studies that help highlight the strengths and weaknesses of each approach (Olivier & Planque 2017). We must not overlook the importance of having a robust understanding of the individual building blocks of natural ecosystems even if we aim to model whole systems (Koen-Alonso 2007). In the long run, this will promote the development of more encompassing and realistic models, while enabling us to limit their complexity and data requirements through ecologically grounded simplifications.

Dealing with uncertainty

Uncertainty is a feature of all statistical and mathematical models that result from simplifying natural processes using imperfect data to estimate unknown processes (Berlow *et al.* 2009). Epistemic or systematic uncertainty (Regan *et al.* 2002) enters the modelling process because of (1) errors in data or insufficient data, (2) random and non-random variation in nature, and (3) assumptions and simplifications about the parameters and model structure (Regan *et al.* 2002; Walker *et al.* 2003; Koo *et al.* 2017; Fig. 2). While many of these sources of uncertainty are shared with single-species models, the greater complexity in multispecies models complicates the task of quantifying their effects, increases the number of pathways through which uncertainty can propagate and increases their potential influence on the overall model output (Zhang *et al.* 2015). It is therefore crucial that sources of uncertainty are identified, quantified and reported in any multispecies model.

Uncertainty in population-specific data can be accounted for through techniques derived from single-species frameworks, such as observation models that estimate sampling error. However, as multispecies models require more diverse data, accounting for measurement errors and/or systematic biases associated with all the data sources becomes more challenging (Regan *et al.* 2002). For instance, estimation of species interactions often requires sampling multiple sources (i.e., several species) and types of data (e.g., interaction types, frequencies) simultaneously, each with some degree of sampling error. Also, species interactions are expected to be influenced by environmental variables, which are themselves estimated with some degree of uncertainty (Koo *et al.* 2017). Similarly, because species interactions are sometimes estimated indirectly within the models (e.g., Ovaskainen *et al.* 2016), interaction estimates become model outputs and subject to additional uncertainty. Expanding datasets and improvements in measuring and identification techniques have great potential to reduce the degree of uncertainty in the data and make models less sensitive to prior assumptions (Cressie *et al.* 2009). However, obtaining more and better data is still limited by logistical challenges (Zhang *et al.* 2015). This is another reason why diversifying the types of independent data collected can be beneficial, as different model parameters can be informed by multiple data sources (e.g., fecundity, census, mortality-at-age) simultaneously in a single framework (Kindsvater *et al.* 2018). Identification and propagation of data uncertainty through the modelling process, and critical assessments of the conditions under which the models are useful are important to minimize errors, biases and misleading projections (Wells & O'Hara 2013; Certain *et al.* 2018; Engelhardt *et al.* 2020).

As mentioned in previous sections, multispecies models often have to rely on inferred interactions or researchers' assumptions about the processes giving rise to the observed data (Milner-Gulland & Shea 2017). However, different assumptions lead to different results. This source of uncertainty can be particularly difficult to quantify because the resulting measurements of uncertainty associated with the likelihood function of the model do not inform about the correctness of the model, but about the certainty in the parameters, already assuming the model structure is true (Kinze & Punt 2009). Instead, this structural uncertainty can be accounted for, quantified or reduced through model comparison, model averaging, or validation of predictions (Regan *et al.* 2002; Koo *et al.* 2017).

Uncertainty cannot be reduced to zero (Milner-Gulland & Shea 2017), but we can explore ways to minimize uncertainty and report it transparently. Recognizing and quantifying all sources of uncertainty is essential for evaluating model usefulness and identifying model weaknesses for future research (Zhang *et al.* 2015). Acknowledging uncertainty can also improve models directly. For example, including prior knowledge about ecological preferences in multispecies SDMs as uncertain, instead of fixed, has been shown to improve both predictability and accuracy (Vermeiren *et al.* 2020). However, recognition and analysis of uncertainty in multispecies models has received relatively little attention, a lack that has been argued to be a major hindrance for the use of multispecies models in management contexts (Thorpe *et al.* 2015). Development of methods to consistently quantify and reduce uncertainty in multispecies models is therefore important going forward and should happen simultaneously with the development of the models themselves. This will ensure that we have the necessary toolkit to achieve usable model outputs and give biologically meaningful insights that can be used in management and conservation contexts. This will also help us to identify parts of the

community or sampling designs associated with higher uncertainty, providing a simple way to improve data collection.

DISCUSSION

Multispecies models have huge potential as tools for understanding and predicting the dynamics of interacting species, and helping to disentangle the effects of these interactions from other processes affecting population dynamics, such as climate change, habitat changes and harvesting (Kinzey & Punt 2009; Péron & Koons 2012; Swallow *et al.* 2016). Rapid progress is made towards fulfilling this potential, but important challenges remain. Currently, the greatest limitations to multispecies model development are data availability and model complexity, leading to difficulties in parameterization and large uncertainty.

Data availability is improving (McCallen *et al.* 2019), but data collection will always be burdened by logistical limitations. It is therefore important that the right types of data are collected to maximize their use for multispecies modelling, for instance by ensuring that data collection gives attention to species interactions and data are collected for multiple cohabiting species simultaneously (Trijoulet *et al.* 2019). Still, collecting data on all components and processes is virtually impossible in most ecosystems. Therefore, it is important to consider grounded ecological knowledge about a system before making decisions regarding the community components that are sampled or omitted during data collection. This may be achieved by, for example, identifying community modules and impossible or missing links (Dormann & Strauss 2014; Terry & Lewis 2020). Additionally, prior assumptions about our knowledge of community functioning must be challenged, given that ecosystems are ever-changing and discrepancies between experts' knowledge can have important impacts on the models (Picard *et al.* 2012; Terry & Lewis 2020). Evaluations of ongoing multispecies monitoring programs can help identify weaknesses in the data collection (e.g., poor sampling of a particular species or region), which, when addressed, could greatly benefit multispecies model performance (Carvalho *et al.* 2016; Zhang *et al.* 2020). In addition, it would be extremely beneficial to foster collaborations among researchers and improve organization of data collection. By increasing collaborations, researchers working on different species could coordinate sampling efforts in the same region, thereby producing data more useful to multispecies modelling without incurring additional costs. Finally, development of methods that promote more efficient use of limited data, such as data integration, have an important role to play in multispecies modelling development.

Model complexity is another major challenge for the development of multispecies models. Attempting to capture the inherent complexity of natural ecosystems in mathematical models causes issues related to computational and data requirements, model parameterization, number of potential error sources, and interpretability and transferability of model outputs. Assumptions and simplifications will continue to be necessary. For example, multispecies population dynamics models need to rely on the assumption that the subset of species and processes included in the models are sufficient to describe the dynamics of interest, and simplified community models must rely on the criteria chosen to group certain species or simplify species interactions. It is crucial that such simplifications are grounded on robust ecological knowledge, rather than on data limitations (Lafferty *et al.* 2008), and that they are transparently reported so that future studies can compare and assess the effects of different simplifying strategies on model outputs (Fulton *et al.* 2003).

The lack of adequate quantification and reporting of uncertainty currently represents a major challenge for the development of multispecies models. Uncertainty is inevitable in ecological modelling, and it is often accentuated in increasingly complex models. In multispecies models, uncertainty comes associated with many sources, which makes the quantification and propagation of individual sources of uncertainty more difficult. There is a great need for methods to assess uncertainty consistently so that models can be compared and evaluated. Although uncertainty is often viewed as a negative model characteristic with regard to practical applications (Pappenberger & Beven 2006), a model with unknown uncertainty is certainly less useful than a model with high, but reported, uncertainty (Keenan *et al.* 2011). In some cases, high uncertainty in

multispecies models can be a positive outcome if it means that the model accounts for important external processes affecting the populations rather than regarding these as random variation, as single-species models would do (Hollowed *et al.* 2000; Kinzey & Punt 2009).

Ultimately, a promising way forward that maximizes data use is the combination of modelling approaches with contrasting strengths and weaknesses (Strauss *et al.* 2017). For example, Schmolke *et al.* (2019) built a hybrid model consisting of a food web simulation model for an aquatic system coupled with an IBM for a single species of interest (Fig. 3). For each time step, the food web model provided biomass estimates for different parts of the food web, based on environmental variables, estimates of vital rates, and expected interactions. Estimates of prey biomass from this model were then fed into the IBM, which modelled in detail the transfer of prey biomass to individuals of the focal species, and the resulting biomass estimate for the focal species was fed back into the food web model. While this model did not include stochastic effects and focused on a single species, the general approach of coupling more detailed population models with a system-wide community model could represent a promising method for incorporating species interaction effects into predictive models of population dynamics (Ernest *et al.* 2011). Other types of hybrid modelling frameworks have also been proposed, for example, by linking a spatially explicit dynamic model of a plant community to an IBM describing the behaviour of frugivorous birds to account for the effect of the birds on seed dispersal (Vincenot *et al.* 2011). As hybridizing models is a relatively new approach within ecology, and their development requires a diverse range of expertise, their potential and benefits remain largely unexplored, representing an exciting path moving forward (Kim *et al.* 2019).

In this review, we have identified common themes in relation to the challenges faced by different subfields of ecology aiming to model multispecies dynamics. Working together on shared challenges across subfields should promote faster and more fruitful progress towards a better understanding of multispecies dynamics. It has become customary for modelling studies to compile their developed models as readily available open-source software and code, which will undoubtedly promote development (Powers & Hampton 2019; Alston & Rick 2021). Moving forward, we emphasize that it is important that multispecies studies clearly outline the model objectives and assumptions. We also encourage increased communication across multispecies modelling disciplines, and hope that this leads to the development of new hybrid frameworks that successfully combine models with contrasting strengths and weaknesses, representing cost-effective ways to advance our understanding of multispecies dynamics (Gray & Wotherspoon 2015; Strauss *et al.* 2017).

Acknowledgements

This study was funded by the Research Council of Norway through the Centre of Excellence Centre for Biodiversity Dynamics (project 223257) and research project SUSTAIN (244647). MVH was supported by the Norwegian Research Council (280715).

Bibliography

- Adler, P.B., Ellner, S.P. & Levine, J.M. (2010). Coexistence of perennial plants: An embarrassment of niches. *Ecol. Lett.* , 13, 1019–1029.
- Agarwal, V., James, C.C., Widdicombe, C.E. & Barton, A.D. (2021). Intraseasonal predictability of natural phytoplankton population dynamics. *Ecol. Evol.* , 11, 15720–15739.
- Allesina, S., Bodini, A. & Bondavalli, C. (2005). Ecological subsystems via graph theory: The role of strongly connected components. *Oikos* , 110, 164–176.
- Alston, J.M. & Rick, J.A. (2021). A beginner’s guide to Conducting Reproducible Research. *Bull. Ecol. Soc. Am.* , 102, 1–14.
- Amarasekare, P. (2008). Spatial dynamics of foodwebs. *Annu. Rev. Ecol. Evol. Syst.* , 39, 479–500.

- Barraquand, F. & Gimenez, O. (2019). Integrating multiple data sources to fit matrix population models for interacting species. *Ecol. Modell.* , 108713.
- Beckerman, A.P., Petchey, O.L. & Warren, P.H. (2006). Foraging biology predicts food web complexity. *Proc. Natl. Acad. Sci. U. S. A.* , 103, 13745–13749.
- Begley, J. & Howell, D. (2004). An overview of Gadget, the globally applicable area-disaggregated general ecosystem toolbox. *Ices C.* , 1–15.
- Berlow, E.L., Dunne, J.A., Martinez, N.D., Stark, P.B., Williams, R.J. & Brose, U. (2009). Simple prediction of interaction strengths in complex food webs. *Proc. Natl. Acad. Sci. U. S. A.* , 106, 187–191.
- Besbeas, P., Freeman, S.N., Morgan, B.J.T. & Catchpole, E.A. (2002). Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters. *Biometrics* , 58, 540–547.
- Blois, J.L., Zarnetske, P.L., Fitzpatrick, M.C. & Finnegan, S. (2013). Climate change and the past, present, and future of biotic interactions. *Science (80-.)* , 341, 499–504.
- Bogstad, B., Pennington, M. & Vølstad, J.H. (1995). Cost-efficient survey designs for estimating food consumption by fish. *Fish. Res.* , 23, 37–46.
- Boit, A., Martinez, N.D., Williams, R.J. & Gaedke, U. (2012). Mechanistic theory and modelling of complex food-web dynamics in Lake Constance. *Ecol. Lett.* , 15, 594–602.
- Breckling, B., Müller, F., Reuter, H., Hölker, F. & Fränzle, O. (2005). Emergent properties in individual-based ecological models - Introducing case studies in an ecosystem research context. *Ecol. Modell.* , 186, 376–388.
- Brose, U. (2010). Body-mass constraints on foraging behaviour determine population and food-web dynamics. *Funct. Ecol.* , 24, 28–34.
- Buchmann, C.M., Schurr, F.M., Nathan, R. & Jeltsch, F. (2011). An allometric model of home range formation explains the structuring of animal communities exploiting heterogeneous resources. *Oikos* , 120, 106–118.
- Butterworth, D.S. & Plagányi, É.E. (2004). A brief introduction to some approaches to multispecies/ecosystem modelling in the context of their possible application in the management of South African fisheries. *African J. Mar. Sci.* , 26, 53–61.
- Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of Structural Patterns in Neutral Trophic Networks. *PLoS One* , 7, e38295.
- Carvalho, S.B., Gonçalves, J., Guisan, A. & Honrado, J.P. (2016). Systematic site selection for multispecies monitoring networks. *J. Appl. Ecol.* , 53, 1305–1316.
- Castillo-Alvino, H. & Marvá, M. (2020). The competition model with Holling type II competitive response to interfering time. *J. Biol. Dyn.* , 14, 222–244.
- Certain, G., Barraquand, F. & Gårdmark, A. (2018). How do MAR(1) models cope with hidden nonlinearities in ecological dynamics? *Methods Ecol. Evol.* , 9, 1975–1995.
- Chamberlain, S.A., Bronstein, J.L. & Rudgers, J.A. (2014). How context dependent are species interactions? *Ecol. Lett.* , 17, 881–890.
- Chandler, M., See, L., Copas, K., Bonde, A.M.Z., López, B.C., Danielsen, F., *et al.* (2017). Contribution of citizen science towards international biodiversity monitoring. *Biol. Conserv.* , 213, 280–294.
- Chiaradia, A., Forero, M.G., McInnes, J.C. & Ramírez, F. (2014). Searching for the true diet of marine predators: Incorporating Bayesian priors into stable isotope mixing models. *PLoS One* , 9, e92665.

- Christensen, T., Payne, J., Doyle, M., Ibarguchi, G., Taylor, J., Schmidt, N.M., *et al.* (2013). *Arctic Terrestrial Biodiversity Monitoring Plan. CAFF Monitoring Series Report Nr. 7*.
- Collie, J.S., Botsford, L.W., Hastings, A., Kaplan, I.C., Largier, J.L., Livingston, P.A., *et al.* (2016). Ecosystem models for fisheries management: Finding the sweet spot. *Fish Fish.*, 17, 101–125.
- Coulson, T. (2012). Integral projections models, their construction and use in posing hypotheses in ecology. *Oikos*, 121, 1337–1350.
- Cressie, N., Calder, C.A., Clark, J.S., Ver Hoef, J.M. & Wikle, C.K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecol. Appl.*, 19, 553–570.
- Curtsdotter, A., Banks, H.T., Banks, J.E., Jonsson, M., Jonsson, T., Laubmeier, A.N., *et al.* (2019). Ecosystem function in predator–prey food webs—confronting dynamic models with empirical data. *J. Anim. Ecol.*, 88, 196–210.
- Daan, N. & Sissenwine, M. (1990). Multispecies models relevant to management of living resources. *Environmentalist*, 10, 229–230.
- DeAngelis, D.L. & Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000Prime Rep.*, 6.
- Degen, R., Aune, M., Bluhm, B.A., Cassidy, C., Kędra, M., Kraan, C., *et al.* (2018). Trait-based approaches in rapidly changing ecosystems: A roadmap to the future polar oceans. *Ecol. Indic.*, 91, 722–736.
- Delmas, E., Besson, M., Brice, M.H., Burkle, L.A., Dalla Riva, G. V., Fortin, M.J., *et al.* (2019). Analysing ecological networks of species interactions. *Biol. Rev.*, 94, 16–36.
- Devarajan, K., Morelli, T.L. & Tenan, S. (2020). Multi-species occupancy models: review, roadmap, and recommendations. *Ecography (Cop.)*, 43, 1612–1624.
- Díaz, S., Purvis, A., Cornelissen, J.H.C., Mace, G.M., Donoghue, M.J., Ewers, R.M., *et al.* (2013). Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecol. Evol.*, 3, 2958–2975.
- Dorazio, R.M., Connor, E.F. & Askins, R.A. (2015). Estimating the effects of habitat and biological interactions in an avian community. *PLoS One*, 10, 1–16.
- Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F., *et al.* (2012). Correlation and process in species distribution models: Bridging a dichotomy. *J. Biogeogr.*, 39, 2119–2131.
- Dormann, C.F. & Strauss, R. (2014). A method for detecting modules in quantitative bipartite networks. *Methods Ecol. Evol.*, 5, 90–98.
- Dunn, R.P., Baskett, M.L. & Hovel, K.A. (2017). Interactive effects of predator and prey harvest on ecological resilience of rocky reefs. *Ecol. Appl.*, 27, 1718–1730.
- Engelhardt, E.K., Neuschulz, E.L. & Hof, C. (2020). Ignoring biotic interactions overestimates climate change effects: The potential response of the spotted nutcracker to changes in climate and resource plants. *J. Biogeogr.*, 47, 143–154.
- Ernest, C., Giannino, F., Rietkerk, M., Moriya, K. & Mazzoleni, S. (2011). Theoretical considerations on the combined use of System Dynamics and individual-based modeling in ecology. *Ecol. Modell.*, 222, 210–218.
- Essington, T.E. (2004). Getting the right answer from the wrong model: Evaluating the sensitivity of multispecies fisheries advice to uncertain species interactions. *Bull. Mar. Sci.*, 74, 563–581.
- Fath, B.D., Scharler, U.M., Ulanowicz, R.E. & Hannon, B. (2007). Ecological network analysis: network construction. *Ecol. Modell.*, 208, 49–55.

- Fithian, W., Elith, J., Hastie, T. & Keith, D.A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods Ecol. Evol.* , 6, 424–438.
- Fulton, E.A., Blanchard, J.L., Melbourne-Thomas, J., Plagányi, É.E. & Tulloch, V.J.D. (2019). Where the ecological gaps remain, a modelers' perspective. *Front. Ecol. Evol.* , 7, 1–22.
- Fulton, E.A., Smith, A.D.M. & Johnson, C.R. (2003). Effect of complexity on marine ecosystem models. *Mar. Ecol. Prog. Ser.* , 253, 1–16.
- Gamelon, M., Filli, F., Sæther, B.E. & Herfindal, I. (2020). Multi-event capture-recapture analysis in Alpine chamois reveals contrasting responses to interspecific competition, within and between populations. *J. Anim. Ecol.* , 89, 2279–2289.
- Gamelon, M., Vriend, S.J.G., Engen, S., Adriaensen, F., Dhondt, A.A., Evans, S.R., *et al.* (2019). Accounting for interspecific competition and age structure in demographic analyses of density dependence improves predictions of fluctuations in population size. *Ecol. Lett.* , 22, 797–806.
- García-Callejas, D., Molowny-Horas, R. & Araújo, M.B. (2018). Multiple interactions networks: towards more realistic descriptions of the web of life. *Oikos* , 127, 5–22.
- Giacomini, H.C., DeAngelis, D.L., Trexler, J.C. & Petrere, M. (2013). Trait contributions to fish community assembly emerge from trophic interactions in an individual-based model. *Ecol. Modell.* , 251, 32–43.
- González-Varo, J.P. & Traveset, A. (2016). The labile limits of forbidden interactions. *Trends Ecol. Evol.* , 31, 700–710.
- Grace, J.B., Anderson, T.M., Olf, H. & Scheiner, S.M. (2010). On the specification of structural equation models for ecological systems. *Ecol. Monogr.* , 80, 67–87.
- Gray, R. & Wotherspoon, S. (2015). Adaptive submodel selection in hybrid models. *Front. Environ. Sci.* , 3, 1–11.
- Grimm, V., Ayllón, D. & Railsback, S.F. (2017). Next-generation individual-based models integrate biodiversity and ecosystems: Yes we can, and yes we must. *Ecosystems* , 20, 229–236.
- Grimm, V. & Berger, U. (2016). Structural realism, emergence, and predictions in next-generation ecological modelling: Synthesis from a special issue. *Ecol. Modell.* , 326, 177–187.
- Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., *et al.* (2006). A standard protocol for describing individual-based and agent-based models. *Ecol. Modell.* , 198, 115–126.
- Hartvig, M., Andersen, K.H. & Beyer, J.E. (2011). Food web framework for size-structured populations. *J. Theor. Biol.* , 272, 113–122.
- Heberling, J.M., Miller, J.T., Noesgaard, D., Weingart, S.B. & Schigel, D. (2021). Data integration enables global biodiversity synthesis. *Proc. Natl. Acad. Sci.* , 118, e2018093118.
- Hollowed, A.B., Bax, N., Beamish, R., Collie, J., Fogarty, M., Livingston, P., *et al.* (2000). Are multispecies models an improvement on single-species models for measuring fishing impacts on marine ecosystems? *ICES J. Mar. Sci. J. du Cons.* , 57, 707–719.
- Holt, R.D. & Lawton, J.H. (1994). The ecological consequences of shared natural enemies. *Annu. Rev. Ecol. Syst.* , 25, 495–520.
- Hood, R.R., Laws, E.A., Armstrong, R.A., Bates, N.R., Brown, C.W., Carlson, C.A., *et al.* (2006). Pelagic functional group modeling: Progress, challenges and prospects. *Deep. Res. Part II Top. Stud. Oceanogr.* , 53, 459–512.
- Horswill, C., Jackson, J.A., Medeiros, R., Nowell, R.W., Trathan, P.N. & O'Connell, T.C. (2018). Minimising the limitations of using dietary analysis to assess foodweb changes by combining multiple techniques. *Ecol.*

Indic. , 94, 218–225.

Hunsicker, M.E., Ciannelli, L., Bailey, K.M., Buckel, J.A., Wilson White, J., Link, J.S., *et al.* (2011). Functional responses and scaling in predator-prey interactions of marine fishes: Contemporary issues and emerging concepts. *Ecol. Lett.* , 14, 1288–1299.

ICES. (2019). Working Group on Multispecies Assessment Methods (WGSAM). *ICES Sci. Reports* , 1, 320.

Isaac, N.J.B., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., *et al.* (2020). Data integration for large-scale models of species distributions. *Trends Ecol. Evol.* , 35, 56–67.

Ives, A.R., Dennis, B., Cottingham, K.L. & Carpenter, S.R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecol. Monogr.*

Jonsson, T., Kaartinen, R., Jonsson, M. & Bommarco, R. (2018). Predictive power of food web models based on body size decreases with trophic complexity. *Ecol. Lett.* , 21, 702–712.

Jordano, P. (1987). Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution. *Am. Nat.* , 129, 657–677.

Kalinkat, G., Rall, B.C., Vucic-Pestic, O. & Brose, U. (2011). The allometry of prey preferences. *PLoS One* , 6.

Kayal, M., Lenihan, H.S., Brooks, A.J., Holbrook, S.J., Schmitt, R.J. & Kendall, B.E. (2018). Predicting coral community recovery using multi-species population dynamics models. *Ecol. Lett.* , 21, 1790–1799.

Kearney, M. & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecol. Lett.* , 12, 334–350.

Keenan, T.F., Carbone, M.S., Reichstein, M. & Richardson, A.D. (2011). The model-data fusion pitfall: Assuming certainty in an uncertain world. *Oecologia* , 167, 587–597.

Kelling, S., Fink, D., La Sorte, F.A., Johnston, A., Bruns, N.E. & Hochachka, W.M. (2015). Taking a 'Big Data' approach to data quality in a citizen science project. *Ambio* , 44, 601–611.

Keppeler, F.W., Montaña, C.G. & Winemiller, K.O. (2020). The relationship between trophic level and body size in fishes depends on functional traits. *Ecol. Monogr.* , 0, 1–19.

Kim, B., Lee, K.H., Xue, L. & Niu, X. (2018). A review of dynamic network models with latent variables. *Stat. Surv.* , 12, 105–135.

Kim, Y., Son, J., Lee, Y.S., Lee, M., Hong, J. & Cho, K. (2019). Integration of an individual-oriented model into a system dynamics model: An application to a multi-species system. *Environ. Model. Softw.* , 112, 23–35.

Kindsvater, H.K., Dulvy, N.K., Horswill, C., Juan-Jordá, M.J., Mangel, M. & Matthiopoulos, J. (2018). Overcoming the data crisis in biodiversity conservation. *Trends Ecol. Evol.* , 33, 676–688.

Kinzey, D. & Punt, A.E. (2009). Multispecies and single-species models of fish population dynamics: Comparing parameter estimates. *Nat. Resour. Model.* , 22, 67–104.

Kjørboe, T., Visser, A., Andersen, K.H. & Browman, H. (2018). A trait-based approach to ocean ecology. *ICES J. Mar. Sci.* , 75, 1849–1863.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., Mcinerny, G.J., *et al.* (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *J. Biogeogr.* , 39, 2163–2178.

Koen-Alonso, M. (2007). A process-oriented approach to the multispecies functional response. *From Energ. to Ecosyst. Dyn. Struct. Ecol. Syst.* , 1–36.

Kohyama, T. (1992). Size-structured multi-species model of rain forest trees. *Funct. Ecol.* , 6, 206.

- Koo, K.A., Park, S.U., Kong, W.S., Hong, S., Jang, I. & Seo, C. (2017). Potential climate change effects on tree distributions in the Korean Peninsula: Understanding model & climate uncertainties. *Ecol. Modell.* , 353, 17–27.
- Lafferty, K.D., Allesina, S., Arim, M., Briggs, C.J., De Leo, G., Dobson, A.P., *et al.* (2008). Parasites in food webs: The ultimate missing links. *Ecol. Lett.* , 11, 533–546.
- Lahoz-Monfort, J.J., Harris, M.P., Morgan, B.J.T., Freeman, S.N. & Wanless, S. (2014). Exploring the consequences of reducing survey effort for detecting individual and temporal variability in survival. *J. Appl. Ecol.* , 51, 534–543.
- Lahoz-Monfort, J.J., Harris, M.P., Wanless, S., Freeman, S.N. & Morgan, B.J.T. (2017). Bringing it all together: Multi-species integrated population modelling of a breeding community. *J. Agric. Biol. Environ. Stat.* , 22, 140–160.
- Latour, R.J., Brush, M.J. & Bonzek, C.F. (2003). Toward ecosystem-based fisheries management: strategies for multispecies modeling and associated data requirements. *Fisheries* , 28, 10–23.
- Lee, A.M., Bjørkvoll, E.M., Hansen, B.B., Albon, S.D., Stien, A., Sæther, B.-E., *et al.* (2015). An integrated population model for a long-lived ungulate: more efficient data use with Bayesian methods. *Oikos* , 124, 806–816.
- Lee, A.M., Sæther, B.-E. & Engen, S. (2020). Spatial covariation of competing species in a fluctuating environment. *Ecology* , 101, 1–9.
- Legović, T. & Geček, S. (2010). Impact of maximum sustainable yield on independent populations. *Ecol. Modell.* , 221, 2108–2111.
- Lotka, J.A. (1925). Elements of physical biology. *Nature* , 116, 461–461.
- Makler-Pick, V., Gal, G., Shapiro, J. & Hipsey, M.R. (2011). Exploring the role of fish in a lake ecosystem (Lake Kinneret, Israel) by coupling an individual-based fish population model to a dynamic ecosystem model. *Can. J. Fish. Aquat. Sci.* , 68, 1265–1284.
- Marvin, D.C., Koh, L.P., Lynam, A.J., Wich, S., Davies, A.B., Krishnamurthy, R., *et al.* (2016). Integrating technologies for scalable ecology and conservation. *Glob. Ecol. Conserv.* , 7, 262–275.
- Marzloff, M.P., Melbourne-Thomas, J., Hamon, K.G., Hoshino, E., Jennings, S., van Putten, I.E., *et al.* (2016). Modelling marine community responses to climate-driven species redistribution to guide monitoring and adaptive ecosystem-based management. *Glob. Chang. Biol.* , 22, 2462–2474.
- McCallen, E., Knott, J., Nunez-Mir, G., Taylor, B., Jo, I. & Fei, S. (2019). Trends in ecology: shifts in ecological research themes over the past four decades. *Front. Ecol. Environ.* , 17, 109–116.
- Mccann, K., Hastings, A. & Huxel, G.R. (1998). Weak trophic interactions and the balance of nature. *Nature* , 395, 794–798.
- McDermot, D. & Rose, K.A. (2000). An individual-based model of lake fish communities: application to piscivore stocking in Lake Mendota. *Ecol. Modell.* , 125, 67–102.
- McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* , 21, 178–185.
- Miller, D.A.W., Pacifici, K., Sanderlin, J.S. & Reich, B.J. (2019). The recent past and promising future for data integration methods to estimate species' distributions. *Methods Ecol. Evol.* , 10, 22–37.
- Milner-Gulland, E.J. & Shea, K. (2017). Embracing uncertainty in applied ecology. *J. Appl. Ecol.* , 54, 2063–2068.

- Mokany, K., Ferrier, S., Connolly, S.R., Dunstan, P.K., Fulton, E.A., Harfoot, M.B., *et al.* (2016). Integrating modelling of biodiversity composition and ecosystem function. *Oikos* , 125, 10–19.
- Montoya, J.M., Pimm, S.L. & Solé, R. V. (2006). Ecological networks and their fragility. *Nature* , 442, 259–264.
- Mooij, W.M., Trolle, D., Jeppesen, E., Arhonditsis, G., Belolipetsky, P. V., Chitamwebwa, D.B.R., *et al.* (2010). Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquat. Ecol.* , 44, 633–667.
- Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015). Inferring biotic interactions from proxies. *Trends Ecol. Evol.* , 30, 347–356.
- Morin, P.J. (2011). *Community Ecology* . *Community Ecol.* 2nd edn. John Wiley & Sons, Ltd, Chichester, UK.
- Mutshinda, C.M., O'Hara, R.B. & Woiwod, I.P. (2011). A multispecies perspective on ecological impacts of climatic forcing. *J. Anim. Ecol.* , 80, 101–107.
- Nielsen, J.M., Clare, E.L., Hayden, B., Brett, M.T. & Kratina, P. (2018). Diet tracing in ecology: Method comparison and selection. *Methods Ecol. Evol.* , 9, 278–291.
- Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., *et al.* (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecol. Monogr.* , 89, 1–24.
- Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2011). Missing and forbidden links in mutualistic networks. *Proc. R. Soc. B Biol. Sci.* , 278, 725–732.
- Olesen, J.M., Bascompte, J., Dupont, Y.L. & Jordano, P. (2007). The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* , 104, 19891–19896.
- Olivier, P. & Planque, B. (2017). Complexity and structural properties of food webs in the Barents Sea. *Oikos* , 1339–1346.
- Ovaskainen, O. & Abrego, N. (2020). *Joint Species Distribution Modelling* . Cambridge University Press.
- Ovaskainen, O., Roy, D.B., Fox, R. & Anderson, B.J. (2016). Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods Ecol. Evol.* , 7, 428–436.
- Ovaskainen, O., Tikhonov, G., Dunson, D., Grøtan, V., Engen, S., Sæther, B.-E., *et al.* (2017). How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proc. R. Soc. B Biol. Sci.* , 284, 20170768.
- Pacala, S.W. & Silander, J.A. (1990). Field tests of neighborhood population dynamic models of two annual weed species. *Ecol. Monogr.* , 60, 113–134.
- Pacifici, K., Reich, B.J., Miller, D.A.W., Gardner, B., Stauffer, G., Singh, S., *et al.* (2017). Integrating multiple data sources in species distribution modeling: A framework for data fusion. *Ecology* , 98, 840–850.
- Pacifici, K., Zipkin, E.F., Collazo, J.A., Irizarry, J.I. & Dewan, A. (2014). Guidelines for a priori grouping of species in hierarchical community models. *Ecol. Evol.* , 4, 877–888.
- Pappenberger, F. & Beven, K.J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resour. Res.* , 42, 1–8.
- Pauly, D., Christensen, V. & Walters, C. (2000). Ecopath, Ecosim, and Ecospace as tools for evaluating ecosystem impact of fisheries. *ICES J. Mar. Sci.* , 57, 697–706.
- Péron, G. & Koons, D.N. (2012). Integrated modeling of communities: Parasitism, competition, and demographic synchrony in sympatric ducks. *Ecology* , 93, 2456–2464.

- Picard, N., Köhler, P., Mortier, F. & Gourlet-Fleury, S. (2012). A comparison of five classifications of species into functional groups in tropical forests of French Guiana. *Ecol. Complex.* , 11, 75–83.
- Pimm, S.L. (1982). Food webs. In: *Food Webs* (ed. Intergovernmental Panel on Climate Change). Springer Netherlands, Dordrecht, pp. 1–11.
- Plagányi, É.E. (2007). Models for an ecosystem approach to fisheries. *FAO Fish. Tech. Pap.* , 477, 108.
- Plagányi, É.E., Punt, A.E., Hillary, R., Morello, E.B., Thébaud, O., Hutton, T., *et al.* (2014). Multispecies fisheries management and conservation: Tactical applications using models of intermediate complexity. *Fish Fish.* , 15, 1–22.
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., *et al.* (2014). Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods Ecol. Evol.* , 5, 397–406.
- Pope, J.G., Rice, J.C., Daan, N., Jennings, S. & Gislason, H. (2006). Modelling an exploited marine fish community with 15 parameters - results from a simple size-based model. *ICES J. Mar. Sci.* , 63, 1029–1044.
- Powers, S.M. & Hampton, S.E. (2019). Open science, reproducibility, and transparency in ecology. *Ecol. Appl.* , 29, 1–8.
- Raick, C., Soetaert, K. & Grégoire, M. (2006). Model complexity and performance: How far can we simplify? *Prog. Oceanogr.* , 70, 27–57.
- Regan, H.M., Colyvan, M. & Burgman, M.A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol. Appl.* , 12, 618–628.
- Reuter, H. (2005). Community processes as emergent properties: Modelling multilevel interaction in small mammals communities. *Ecol. Modell.* , 186, 427–446.
- Rohr, R.P., Naisbit, R.E., Mazza, C. & Bersier, L.Fé. (2016). Matching-centrality decomposition and the forecasting of new links in networks. *Proc. R. Soc. B Biol. Sci.* , 283.
- Rosenbaum, B. & Rall, B.C. (2018). Fitting functional responses: Direct parameter estimation by simulating differential equations. *Methods Ecol. Evol.* , 9, 2076–2090.
- Roughgarden, J. (1975). Population dynamics in a stochastic environment: Spectral theory for the linearized N-species Lotka-Volterra competition equations. *Theor. Popul. Biol.* , 7, 1–12.
- Rüger, N., Condit, R., Dent, D.H., DeWalt, S.J., Hubbell, S.P., Lichstein, J.W., *et al.* (2019). Demographic tradeoffs predict tropical forest dynamics. *bioRxiv* , 168, 165–168.
- Russell, R.W., Hunt, G.L., Coyle, K.O. & Cooney, R.T. (1992). Foraging in a fractal environment: Spatial patterns in a marine predator-prey system. *Landsc. Ecol.* , 7, 195–209.
- Schaub, M. & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *J. Ornithol.* , 152, S227–S237.
- Schliep, E.M., Lany, N.K., Zarnetske, P.L., Schaeffer, R.N., Orians, C.M., Orwig, D.A., *et al.* (2018). Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. *Glob. Ecol. Biogeogr.* , 27, 142–155.
- Schmitz, O.J. (2004). From mesocosms to the field: the role and value of cage experiments in understanding top-down effects in ecosystems. In: *Insects and ecosystem function* (eds. Weisser, W.W. & Siemann, E.). Springer, Berlin-Heidelberg, pp. 277–302.
- Schmolke, A., Bartell, S.M., Roy, C., Green, N., Galic, N. & Brain, R. (2019). Species-specific population dynamics and their link to an aquatic food web : A hybrid modeling approach. *Ecol. Modell.* , 405, 1–14.

- Schneider, F.D., Scheu, S. & Brose, U. (2012). Body mass constraints on feeding rates determine the consequences of predator loss. *Ecol. Lett.* , 15, 436–443.
- Simmons, B.I., Cirtwill, A.R., Baker, N.J., Wauchope, H.S., Dicks, L. V, Stouffer, D.B., *et al.* (2019). Motifs in bipartite ecological networks: uncovering indirect interactions. *Oikos* , 128, 154–170.
- Spalinger, D.E. & Hobbs, N.T. (1992). Mechanisms of foraging in mammalian herbivores: New models of functional response. *Am. Nat.* , 140, 325–348.
- Spiesman, B.J. & Gratton, C. (2016). Flexible foraging shapes the topology of plant-pollinator interaction Networks. *Ecology* , 97, 1431–1441.
- Strauss, T., Gabsi, F., Hammers-Wirtz, M., Thorbek, P. & Preuss, T.G. (2017). The power of hybrid modelling: An example from aquatic ecosystems. *Ecol. Modell.* , 364, 77–88.
- Stroud, J.T., Bush, M.R., Ladd, M.C., Nowicki, R.J., Shantz, A.A. & Sweatman, J. (2015). Is a community still a community? Reviewing definitions of key terms in community ecology. *Ecol. Evol.* , 5, 4757–4765.
- Swallow, B., King, R., Buckland, S.T. & Toms, M.P. (2016). Identifying multispecies synchrony in response to environmental covariates. *Ecol. Evol.* , 6, 8515–8525.
- Tarnecki, J.H., Wallace, A.A., Simons, J.D. & Ainsworth, C.H. (2016). Progression of a Gulf of Mexico food web supporting Atlantis ecosystem model development. *Fish. Res.* , 179, 237–250.
- Temeles, E.J., Newman, J.T., Newman, J.H., Cho, S.Y., Mazzotta, A.R. & Kress, W.J. (2016). Pollinator competition as a driver of floral divergence: An experimental test. *PLoS One* , 11, 1–17.
- Terry, J.C.D. & Lewis, O.T. (2020). Finding missing links in interaction networks. *Ecology* , 101, 1–13.
- Terry, J.C.D., Morris, R.J. & Bonsall, M.B. (2017). Trophic interaction modifications: an empirical and theoretical framework. *Ecol. Lett.* , 20, 1219–1230.
- Theobald, E.J., Ettinger, A.K., Burgess, H.K., DeBey, L.B., Schmidt, N.R., Froehlich, H.E., *et al.* (2015). Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biol. Conserv.* , 181, 236–244.
- Thorpe, R.B., Le Quesne, W.J.F., Luxford, F., Collie, J.S. & Jennings, S. (2015). Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing. *Methods Ecol. Evol.* , 6, 49–58.
- Thorson, J.T., Ianelli, J.N., Larsen, E.A., Ries, L., Scheuerell, M.D., Szuwalski, C., *et al.* (2016). Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. *Glob. Ecol. Biogeogr.* , 25, 1144–1158.
- Thorson, J.T., Scheuerell, M.D., Shelton, A.O., See, K.E., Skaug, H.J. & Kristensen, K. (2015). Spatial factor analysis: A new tool for estimating joint species distributions and correlations in species range. *Methods Ecol. Evol.* , 6, 627–637.
- Titulaer, M., Melgoza-Castillo, A., Panjabi, A.O., Sanchez-Flores, A., Martinez-Guerrero, J.H., Macias-Duarte, A., *et al.* (2017). Molecular analysis of stomach contents reveals important grass seeds in the winter diet of Baird's and Grasshopper sparrows, two declining grassland bird species. *PLoS One* , 12, 1–17.
- Tjelmeland, S. & Bogstad, B. (1998). MULTSPEC - A review of a multispecies modelling project for the Barents Sea. *Fish. Res.* , 37, 127–142.
- Torres-Campos, I., Magalhaes, S., Moya-Larano, J. & Montserrat, M. (2020). The return of the trophic chain: Fundamental vs. realized interactions in a simple arthropod food web. *Funct. Ecol.* , 34, 521–533.
- Travers, M., Shin, Y.J., Jennings, S. & Cury, P. (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. *Prog. Oceanogr.* , 75, 751–770.

- Trijoulet, V., Fay, G., Curti, K., Smith, B. & Miller, T. (2019). Performance of multispecies assessment models: insights on the influence of diet data. *ICES J. Mar. Sci.* , 76, 1464–1476.
- Ulanowicz, R.E. (1972). Mass and energy flow in closed ecosystems. *J. Theor. Biol.* , 34, 239–253.
- Ulanowicz, R.E., Holt, R.D. & Barfield, M. (2014). Limits on ecosystem trophic complexity: Insights from ecological network analysis. *Ecol. Lett.* , 17, 127–136.
- Vazquez, D.P., Bluthgen, N., Cagnolo, L. & Chacoff, N.P. (2009). Uniting pattern and process in plant-animal mutualistic networks: A review. *Ann. Bot.* , 103, 1445–1457.
- van der Veen, B., Hui, F.K.C., Hovstad, K.A., Solbu, E.B. & O’Hara, R.B. (2021). Model-based ordination for species with unequal niche widths. *Methods Ecol. Evol.* , 2021, 1288–1300.
- Vermeiren, P., Reichert, P. & Schuwirth, N. (2020). Integrating uncertain prior knowledge regarding ecological preferences into multi-species distribution models: Effects of model complexity on predictive performance. *Ecol. Modell.* , 420, 108956.
- Vincenot, C.E., Giannino, F., Rietkerk, M., Moriya, K. & Mazzoleni, S. (2011). Theoretical considerations on the combined use of System Dynamics and individual-based modeling in ecology. *Ecol. Modell.* , 222, 210–218.
- Volterra, V. (1928). Variations and fluctuations of the number of individuals in animal species living together. *ICES J. Mar. Sci.* , 3, 3–51.
- Walker, W.E., Harremoes, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P., *et al.* (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integr. Assess.* , 4, 5–17.
- Warton, D.I., Blanchet, F.G., O’Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., *et al.* (2015). So many variables: Joint modeling in community ecology. *Trends Ecol. Evol.* , 30, 766–779.
- Weinstein, B.G. (2015). MotionMeerkat: Integrating motion video detection and ecological monitoring. *Methods Ecol. Evol.* , 6, 357–362.
- Wells, K. & O’Hara, R.B. (2013). Species interactions: Estimating per-individual interaction strength and covariates before simplifying data into per-species ecological networks. *Methods Ecol. Evol.* , 4, 1–8.
- Wilson, S.D. & Tilman, D. (1991). Component of plant competition along an experimental gradient of nitrogen availability. *Ecology* , 72, 1050–1065.
- Wood, C.L., Summerside, M. & Johnson, P.T.J. (2019). An effective method for ecosystem-scale manipulation of bird abundance and species richness. *Ecol. Evol.* , 9, 9748–9758.
- Wootton, J.T. (2001). Prediction in complex communities: Analysis of empirically derived Markov models. *Ecology* , 82, 580.
- Yodzis, P. & Winemiller, K.O. (1999). In search of operational trophospecies in a tropical aquatic food web. *Oikos* , 87, 327.
- Vander Zanden, M.J., Casselman, J.M. & Rasmussen, J.B. (1999). Stable isotope evidence for the food web consequences of species invasions in lakes. *Nature* , 401, 464–467.
- Zhang, C., Chen, Y. & Ren, Y. (2015). Assessing uncertainty of a multispecies size-spectrum model resulting from process and observation errors. *ICES J. Mar. Sci.* , 72, 2223–2233.
- Zhang, C., Xu, B., Xue, Y. & Ren, Y. (2020). Evaluating multispecies survey designs using a joint species distribution model. *Aquac. Fish.* , 5, 156–162.
- Zipkin, E.F. & Saunders, S.P. (2018). Synthesizing multiple data types for biological conservation using integrated population models. *Biol. Conserv.* , 217, 240–250.

Captions

Figure 1. Diagram illustrating four methods for reducing complexity of natural systems. Centre figure shows a hypothetical system with the true interactions and variables. Functional groups/trophospecies illustrates the idea of grouping species by trophic/functional similarity and treating each group as one component. Modules represent sub-groups of mostly mutually interacting species which can be modelled independently. Modules can also be linked to each other. General functions can define interactions between individuals based on traits (e.g., size), rather than between species. Latent variables are used to represent the variability caused by unknown factors (e.g., species or environmental variables).

Figure 2. Representation of potential sources of uncertainty. The thick/grey arrows on top illustrate the propagation of uncertainty through the modelling process. The red symbols highlight example causes for uncertainty, such as: (a) interactions taking place at night (represented by the moon symbol) or in difficult to observe/monitor places, (b) incorrect assumptions about the presence or absence of species interactions and, (c) uncertainty in estimated parameters. Lastly, model uncertainty can be reduced through increasing transparency.

Figure 3. Example of hybrid model that combines a food web model (brown) and individual based model (IBM; blue) to represent a focal species (S_3) in more detail. For each model iteration, values of the food web model components influence responses of individuals in the IBM, like movement, metabolism, excretions, fecundity, growth or behaviour. In turn, total changes in the population of the focal species (S_3) produced by the IBM feed back into the food web model, affecting the species that interact with the focal species in the following iteration (i.e., S_1 and S_5).

Figures

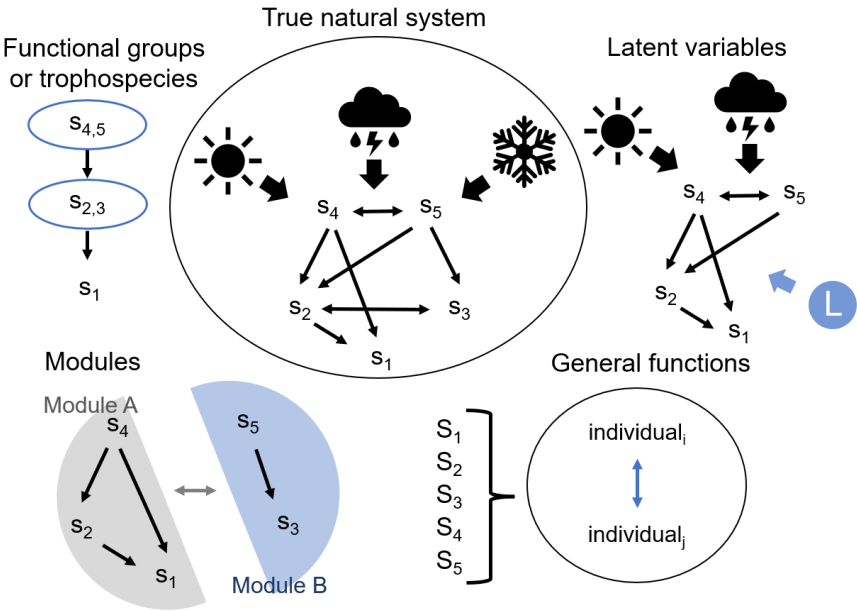


Figure 1. Diagram illustrating four methods for reducing complexity of natural systems. Centre figure shows a hypothetical system with the true interactions and variables. Functional groups/trophospecies illustrates

the idea of grouping species by trophic/functional similarity and treating each group as one component. Modules represent sub-groups of mostly mutually interacting species which can be modelled independently. Modules can also be linked to each other. General functions can define interactions between individuals based on traits (e.g., size), rather than between species. Latent variables are used to represent the variability caused by unknown factors (e.g., species or environmental variables).

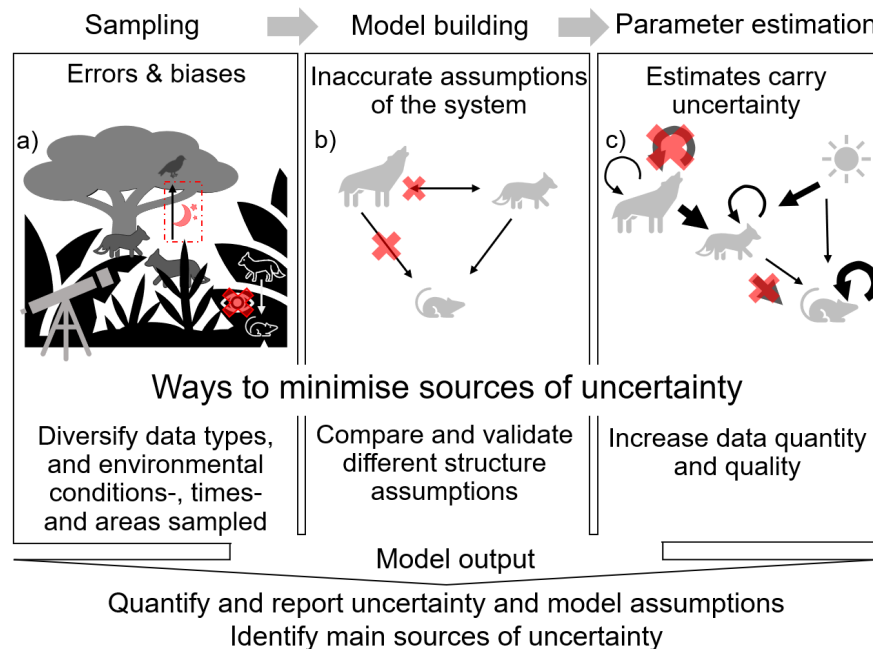


Figure 2. Representation of potential sources of uncertainty. The thick/grey arrows on top illustrate the propagation of uncertainty through the modelling process. The red symbols highlight example causes for uncertainty, such as: (a) interactions taking place at night (represented by the moon symbol) or in difficult to observe/monitor places, (b) incorrect assumptions about the presence or absence of species interactions and, (c) uncertainty in estimated parameters. Lastly, model uncertainty can be reduced through increasing transparency.

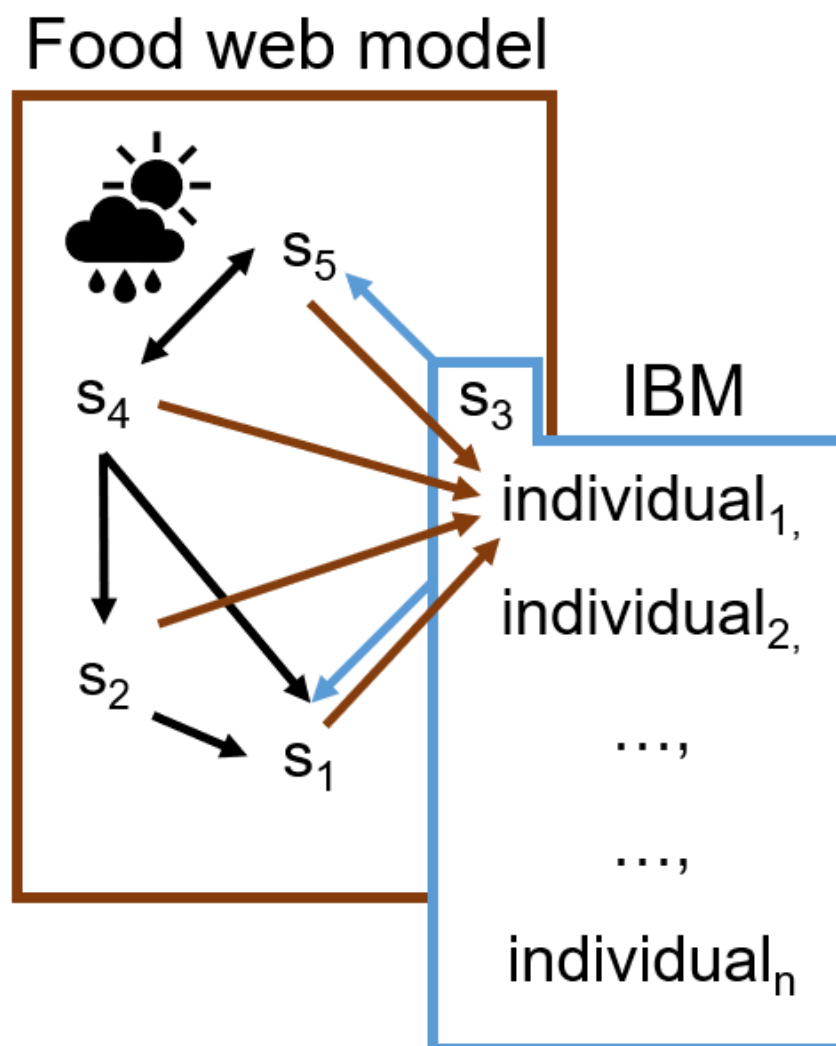


Figure 3. Example of hybrid model that combines a food web model (brown) and individual based model (IBM; blue) to represent a focal species (S_3) in more detail. For each model iteration, values of the food web model components influence responses of individuals in the IBM, like movement, metabolism, excretions, fecundity, growth or behaviour. In turn, total changes in the population of the focal species (S_3) produced by the IBM feed back into the food web model, affecting the species that interact with the focal species in the following iteration (i.e., S_1 and S_5).

Boxes