F

FACILITATION

MICHAEL W. McCOY, CHRISTINE HOLDREDGE, AND BRIAN R. SILLIMAN
University of Florida, Gainesville

ANDREW H. ALTIERI
Brown University, Providence, Rhode Island

MADS S. THOMSEN
University of Aarhus, Roskilde, Denmark

Facilitation includes direct or indirect interactions between biological entities (i.e., cells, individuals, species, communities, or ecosystems) that benefit at least one participant in the interaction and cause harm to none. Research on ecological facilitation has steadily increased over the past three decades and is now appreciated as a fundamental process in ecology. Facilitation also has many important implications for problems in applied ecology and conservation.

HISTORICAL CONTEXT

Understanding the processes governing species coexistence and community structure is a central goal of community ecology. Historically, most ecological research has focused on the negative effects of abiotic or biotic interactions as the primary drivers of species occurrence and the organization of ecological communities. Consequently, negative ecological interactions, such as competition and predation, contribute disproportionately to the conceptual foundation upon which most ecological theory is built. However, over the past three decades there

has been a growing body of literature highlighting the important role of facilitative interactions for population- and community-level processes.

Before discussing recent conceptual advancements leading from facilitation research, it is useful to briefly reflect on the history of ecology and facilitation research. The focus of ecologists on antagonistic species interactions can be traced back even to the publication of *The Origin of Species* in 1859, in which Charles Darwin metaphorically described species as wedges. In this metaphor, ecological space (the resource pie) is divided into a series of wedges that represent species' population or range sizes (i.e., proportion of the pie occupied). The addition of a new species to the resource pie or an increase in the size of the wedge of an existing species must necessarily come at the expense of other species. Indeed, the footprint of this metaphor is imprinted on the conceptual foundations of ecology and lies at the heart of many core ecological concepts, such as the competitive exclusion principle, the niche concept, island biogeography, community assembly, and community invasibility. Interestingly, in addition to setting the stage for research on negative interactions, Darwin also laid the groundwork for studying facilitation by recognizing that reciprocally positive interactions could arise in nature as a result of organisms acting with purely selfish interests. However, his insights did not permeate ecological thought until the mid-twentieth century. Indeed, it is now widely recognized that consideration of facilitation fundamentally changes our understanding of many core ecological concepts and greatly enhances the generality and depth of ecological theory. The stage has been set for rapid development of ecological knowledge as positive interactions, negative interactions, and neutral processes

are incorporated into a more sophisticated archetype for ecology.

INTRASPECIFIC FACILITATION, MUTUALISM, AND COMMENSALISM

Although ecologists traditionally invoke facilitation to describe interspecific interactions, facilitation between individuals of the same species (intraspecific facilitation) also plays a key role in driving population and community dynamics. Under some circumstances, organisms can experience positive density dependence whereby individuals living in aggregations have higher growth rates, survivorship, and reproductive output. These benefits arise via a wide variety of mechanisms ranging from the reductions in per capita risk that occur as predator consumption rates saturate at high prey densities (e.g., dilution effects) to the buffering effects of neighbors against harsh physical stressors (e.g., in rocky intertidal zones). Positive density dependence also occurs at low densities via Allee affects where a species' population growth rate rises with increasing density via increased fertilization success and propagule survival.

Facilitative interactions among populations and species are generally categorized as mutualism or commensalism. Mutualism is a specific form of facilitation in which interactions are reciprocally positive for all participants (Fig. 1). Mutualism includes highly transient reciprocally positive interactions as well as interactions that have emerged as a result of long coevolutionary histories between participants (such as plant-pollinator and plant-disperser networks). Commensalism, in contrast, is reserved for cases of facilitation where at least one participant in an interaction is positively affected while others are neither positively nor negatively impacted. Although

FIGURE 1 Example of mutualism. In this interaction, coral (*Pocillopora sp.*) provides habitat, shelter, and foods for crabs (*Trapezia rufopunctata.*), which, in turn, provide the coral protection by repelling coral predators. Photograph by Adrian Stier.

the literature is replete with examples of mutualisms, there is a relative scarcity of examples of commensalism in nature. Consequently, some ecologists debate the relevance of this term, arguing that species interactions are more likely asymmetrical in strength (i.e., one species exhibits a strong positive response to the other, while the other exhibits a weak positive [or negative] response to the first), rather than being truly commensal in nature.

THEORETICAL PERSPECTIVES

While there has been much recent interest in facilitation, there is still much to learn about how positive interactions influence population and community dynamics. Theoretical developments on positive interactions have only recently begun to move beyond phenomenological descriptions to identify general mechanisms that drive ecological dynamics. Indeed, a number of recent advancements have illustrated the essential role of positive interactions for key ecological phenomena such as ecological community assembly, determining the geographical distributions of species, maintaining species coexistence, and influencing the diversity and stability of ecological communities.

Simple two-species models have most commonly been used to investigate the effects of positive interactions on species coexistence and to examine the environmental conditions where positive interactions are expected. For example, as early as 1935, Gause and Witt examined two-species Lotka–Volterra models and showed that positive interactions can be destabilizing when they are strong because they create positive feedbacks between species (e.g., mutualisms) that can lead to infinite population growth. When interaction strengths are weak or strongly asymmetrical (e.g., commensalisms), however, positive interactions can have stabilizing effects, especially when they occur in conjunction with external mortality sources such as predation, disturbance, and stressful environments. Recent investigations have shown that incorporating nonlinearities via density dependence in cost–benefit functional responses (i.e., positive effects saturate with increasing population density) has a stabilizing effect in these models. In fact, mutualistic communities characterized by nonlinear functional responses have positive complexity–stability relationships that suggest positive interactions may be important drivers of community resilience as a whole.

Phenomenological models (e.g., simulations and agent-based models) have extended the findings of a large body of empirical research to generate important insight into the conditions where positive interactions are expected to play a key role in species coexistence and persistence. In general, positive interactions are important for promoting species

persistence in severe environments (e.g., arctic, salt marsh, and desert ecosystems) and extending the geographic distributions of species along range boundaries. Specifically, positive interactions can expand species range limits by enabling the expansion of the realized niche into more severe environments than would be possible without positive interactions. In these cases, one species makes local environments favorable for colonization by a second species by directly or indirectly enhancing access to resources, dispersal rate, or provision of refuge from competitors, predators, or abiotic stress. Facilitative interactions, however, are often context dependent. Most interactions between species have positive and negative components, and the relative strength of positive and negative effects is often determined by the environmental context of the interaction. For example, mutualism can change to competition along a stress gradient, and so species can exist as facilitators and competitors in different zones of a landscape.

Although models examining the effects of positive interactions on the dynamics of species pairs have provided progressive insights into the conditions where facilitation promotes species coexistence, population stability, and range limits, species rarely interact and coexist in isolation from other species or habitats. Indeed, most natural communities are characterized by complex webs of interacting species that are spatially linked to other communities via dispersal or species movements. It is likely that the effects of facilitation will not be intuitive extensions of two-species models in these multispecies and multipatch assemblages.

Several recent studies have, in fact, begun to examine spatially explicit multispecies metacommunity models that include diverse interspecific interactions and environmental stress gradients. These models reveal some general expectations for the role of facilitation in community assembly and ecosystem diversity and function. For example, poor habitat quality and low spatial connectedness are expected to favor the emergence of highly stable local communities that are strongly facilitative but characterized by low diversity and low productivity. In contrast, high habitat quality and high connectedness among metacommunity patches are expected to promote local communities that are less facilitative and less stable but characterized by high diversity and productivity (Fig. 2).

EMPIRICAL PERSPECTIVES

Foundation Species and Community Facilitation

The structure and dynamics of many widely recognized ecological communities are influenced by facilitation. In fact, many communities are identified by a foundation species that provides or creates the physical structure,
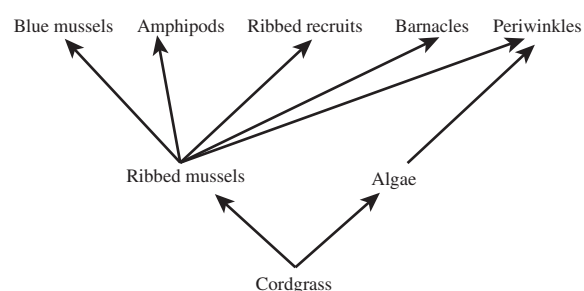

FIGURE 2 Example of a facilitation cascade in cordgrass/mussel bed communities. The establishment of cordgrass intiates a facilitation cascade whereby the establishment of ribbed mussels is facilitated by cordgrass. The synergistic effects of the cordgrass and mussels, in turn, facilitate the establishment of a variety of other taxa, including other species of mussels, barnacles, amphipods, and snails. Photograph by Andrew Altieri.

conditions, and boundaries of a community (e.g., kelp beds, coral reefs, hardwood forests, mangrove stands, salt marshes, phytotelmata, and so on), which directly or indirectly facilitates a diverse array of species. Though a range of interactions including faciliation, competition, and predation may occur among species in the community, the overall persistence of the community is facilitated by the foundation species. The exact mechanism by which a foundation species facilitates a community varies among habitats. In physically stressful environments, foundation species typically ameliorate environmental stress, whereas in more benign environments they more typically provide refuge from predation. By creating patches of suitable habitat, foundation species influence community structure on a landscape scale and can be important drivers for both local and regional patterns of diversity.

Although community facilitation is often attributed to a single foundation species, it can also be driven by the synergistic interactions between two or more species that in concert provide the foundation for communities. Facilitation of communities via synergistic interactions among

foundation species has been explicitly identified in only a few habitats, such as cobble beaches and coral reefs, but could be widespread in habitats that are defined by mixtures of species, such as sea grass meadows. However, like other forms of facilitation, community facilitation is likely context dependent, with the interaction between foundation species changing from synergistic to antagonistic along an environmental gradient. At a landscape scale, this context dependency can lead to abrupt changes in species composition if foundation species exist as faciliators and competitors in different zones of the landscape.

Ecosystem Facilitation

Facilitation is also an important process operating at the highest level of ecological organization—ecosystems. Ecologists have long recognized the importance of fluxes of energy, matter, and organisms for driving ecosystem processes. When this spatial coupling benefits foundation species and their associated species assemblages in a way that increases ecosystem services (e.g., increased stability, productivity, resilience, and so on), then the spatial coupling serves as a form of "ecosystem facilitation." Ecosystem facilitation differs from community facilitation in that it typically occurs along spatial gradients and refers to positive interactions across ecosystem boundaries, whereas community facilitation typically occurs along stress gradients and refers to species facilitation within a single community or ecosystem type. For example, in aquatic systems ecosystem facilitation occurs when plankton produced in the pelagic zone sinks and provides food and energy to support and maintain the benthic ecosystem (benthic–pelagic coupling). Ecosystem facilitation is particularly common between aquatic and terrestrial ecosystems and between productive and unproductive systems. For example, offshore sea grass beds and coral reefs can attenuate storm waves to protect intertidal mangroves and salt marshes that then provide additional energetic protection for inland terrestrial ecosystems. Mangroves and salt marshes can also reciprocally facilitate offshore sea grasses and coral reefs by catching and retaining suspended sediments and nutrients that can drive epiphytic growth and shading that is harmful to seagrass and coral ecosystems.

In these and most other examples of ecosystem facilitation, the positive interactions among ecosystems occur via provision of limiting resources (spatial subsidies of nutrients, organic matter, shelter, habitat to live in) or by reducing environmental stressors (e.g., sediments, toxins, flooding, storm disturbances, abrasion). However, spatial coupling among ecosystems is not always facilitative and can become detrimental in some contexts. For example, although plankton deposition is important for benthic ecosystems, excessive depositions may ultimately disrupt the system by choking filter feeders and stimulating bacterial blooms.

Finally, it is important to note that research on ecosystem facilitation, in contrast to population and community facilitation, is dominated by "correlative studies." Natural ecosystems are typically too large and complicated to manipulate, so the mechanistic basis of ecosystem facilitation is often difficult to identify. Thus, where manipulative experiments cannot be conducted, a weight-of-evidence approach is needed that combines rigorous site selection and data collection criteria, statistical modeling, and natural history.

CONSERVATION APPLICATIONS

Conservation biology is a goal-focused science whose primary objective is to study, protect, and preserve preidentified targets. Historically, conservation targets were specific species of concern, and conservation strategies focused on minimizing negative species and environmental interactions. Stressor-reducing approaches, however, fail to incorporate positive interactions into their designs and are being replaced by strategies that identify and harness positive interactions. Specifically, recent advancements in conservation research have instigated the expansion of conservation targets to include entire ecosystems (i.e., restoration of foundation species) and the functions and services they provide.

The common omission of positive interactions in conservation and restoration approaches has likely resulted in missed opportunities to enhance conservation projects at no increased cost, as many natural synergisms do not emerge from restoration and preservation designs focused on minimizing negative interactions. For example, when restoring coastal marshes and mangroves, traditional planting designs have been plantation style, with all plugs planted at far enough and equal distances from each other to ensure no competition. Experimental work in these systems over the past 20 years, however, has shown that mangrove and marsh plants, when growing in stressful mudflats, grow better in larger clumps and when these clumps are placed closer together. The improved growth stems from the benefits plants receive from the aeration of soils by other nearby plants. These studies also clearly show that these cooperative benefits far outweigh the negative impacts of competition for nutrients between plugs. Thus, by not updating its designs and theoretical approaches to incorporate recent findings that highlight the importance of positive interactions in the success of species under harsh physical conditions, restoration

ecology is failing to take advantage of naturally occurring synergisms among species and individuals.

While this wetland restoration example is focused on the importance of incorporating positive interactions at the population level, such synergisms can also occur at the ecosystem level. For example, some approaches to protecting shoreline habitats have already begun to incorporate positive interactions among species, ecosystems, and man-made structures. The overall goal is to maximize positive interactions that surrounding or overlapping ecosystems would naturally provide each other but were lost under the old paradigm of coastal protection (remove all buffers in favor of stronger man-made structures). The combined use of hard structures to fend off flooding and erosion and wetland plant ecosystem restoration can be effective if we identify compatible and complementary aspects of engineering and vegetation adaptation measures. An excellent example comes from Dutch engineering and conservation efforts, where coastal engineers have tried to "build with nature" to increase the resilience of their man-made structures to oceanic disturbance. Levees built to prevent flooding during storms are maintained with a thick grass cover to increase their integrity. In addition, more recent efforts have focused on placing willow trees and marsh grasses just ahead of man-made levees to reduce wave action on and water levels around the protective barrier. The benefits of this type of positive-interaction engineering go well beyond natural ecosystems enhancing the integrity and efficacy of man-made structures, as the services of the planted ecosystems are not limited to this one interaction. For instance, planted marshes likely increase fishery production in surrounding areas, and willows increase carbon sequestration and habitat for local songbirds. Human utilization of the shorelines also increases, as the natural ecosystems planted on top of and around man-made structure provide recreational opportunities.

Despite advances in positive interactions and facilitation theory in ecological research over the past 20 years, the concepts have failed to make a large impact on conservation and restoration ecology. As highlighted in the examples above, incorporating positive interactions into conservation and restoration practices can occur at the organismal, population, community, and ecosystem levels and can reap substantial benefits with little additional investment in resources. Conservation and restoration plans simply need to be modified to explicitly integrate positive interactions. The old paradigm of applying terrestrial forestry and wildlife theory (i.e., minimizing competition and predation on target species) to modern-day conservation and restoration efforts needs to be updated with current ecological theory revealing that positive interactions among species under harsh conditions (i.e., those of stressed targets) are paramount to those species continued existence.

SEE ALSO THE FOLLOWING ARTICLES

Conservation Biology / Metacommunities / Resilience and Stability / Restoration Ecology / Spatial Ecology / Stress and Species Interactions / Succession / Two-Species Competition

FURTHER READING

Altieri, A., B. Silliman, and M. Bertness. 2007. Hierarchical organization via a facilitation cascade in intertidal cordgrass bed communities. *The American Naturalist* 169: 195–207.

Bertness, M., and R. Callaway. 1994. Positive interactions in communities. *Trends in Ecology & Evolution* 9: 191–193.

Bruno, J., J. Stachowicz, and M. Bertness. 2003. Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution* 18: 119–125.

Stachowicz, J. J. 2001. Mutualism, facilitation, and the structure of ecological communities. *Bioscience* 51: 235–246.

Thomsen, M. S., T. Wernberg, A. H. Altieri, F. Tuya, D. Gulbransen, K. J. McGlathery, M. Holmer, and B. R. Silliman. 2010. Habitat cascades: the conceptual context and global relevance of facilitation cascades via habitat formation and modification. *Integrative and Comparative Biology* 50: 158–175.

FISHERIES ECOLOGY

ELLIOTT LEE HAZEN

Joint Institute for Marine and Atmospheric Research, University of Hawaii

LARRY B. CROWDER

Stanford University, California

Fisheries ecology is the integration of applied and fundamental ecological principles relative to fished species or affected nontarget species (e.g., bycatch). Fish ecology focuses on understanding how fish interact with their environment, but fisheries ecology extends this understanding to interactions with fishers, fishery communities, and the institutions that influence or manage fisher behaviors. Traditional fisheries science has focused on single-species stock assessments and management with the goal of understanding population dynamics and variability. But in the past few decades, scientists and managers have analyzed the effects of fisheries on target and nontarget species and on supporting habitats and food webs and have attempted to quantify ecological linkages

(e.g., predator–prey, competition, disturbance) on food web dynamics. Fisheries ecology requires understanding how population variability is influenced by species interactions, environmental fluctuations, and anthropogenic factors.

THE DISCIPLINE

While some researchers view fisheries ecology as basic ecology applied to fisheries management, the complex interactions among individual organisms, trophic levels, and a changing environment have spanned a number of ecological fields. The field of fisheries ecology was described by Magnuson in 1991 as the intersection of ichthyology, fisheries science, and ecology, and also including theory from a variety of ecological disciplines including physiological ecology, behavioral ecology, population ecology, community ecology, ecosystem ecology, and landscape ecology (Fig. 1). More recently, the fields of historical ecology and social-ecological systems have aimed to better define a baseline for change and to re-integrate human dimensions into ecology. These diverse fields highlight the interdisciplinary nature of fisheries ecology but also the interactions at various scales and across individuals, species, communities, and ultimately ecosystems.

Physiological Ecology

The field of physiological ecology initially examined the effects of physical features such as dissolved oxygen (DO), temperature, and salinity on mortality of individual organisms. Similar studies have extended data from individual organisms to identify potential habitat for an entire species using thermal, salinity, and DO preferences and support the concept of environmentally mediated niches.

Behavioral Ecology

Behavioral ecology includes sensory perception and decision making in individual organisms. One particularly applicable example is Pacific salmon that live in saltwater and spawn in freshwater, homing to the same stream where they hatched. By following chemical cues from river outflow, they are able to return to their natal stream to spawn. Decision-making choices such as migration from breeding to feeding grounds and foraging theory such as when to leave one feeding ground for another have been aided by the development of archival tags
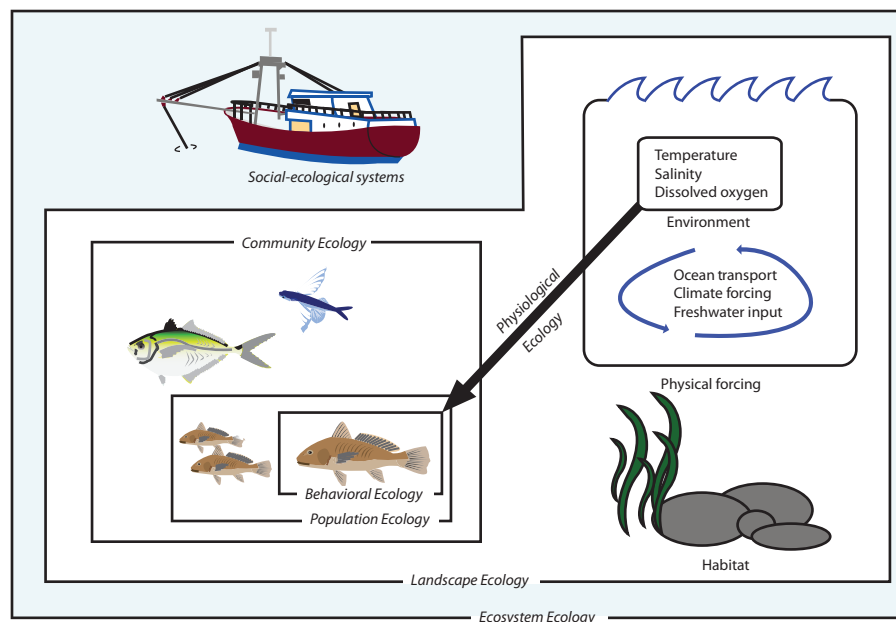

FIGURE 1 Fisheries ecology consists of multiple ecological disciplines (shown in italics) and can focus on the behavior of a single species to the ecosystem and on the economic effects of fishing. *Physiological ecology* focuses on the effects of physical variables in the environment on cell biology and mortality; *behavioral ecology* focuses on the response of individuals or schools of fish to sensory cues; *population ecology* focuses on measuring growth rates, biomass, and age structure of a population; *community ecology* focuses on the interactions among species and how that affects food web structure and population dynamics; *landscape ecology* focuses on how physical features including habitat affect distribution and abundance of one or many species; *ecosystem ecology* examines trophic interactions, physical forcing, and the resiliency of ecosystem states; *social-ecological systems* focuses on the human dimensions of how fishermen affect and are affected by fisheries and the environment; and *historical ecology* focuses on long, often prehistoric time series (not shown).

(e.g., satellite) that can measure movement patterns in otherwise hard-to-study species.

Population Ecology

Population ecology has served as the primary tool of fisheries management since the mid-twentieth century. Since the passing of the Magnuson–Stevens Act in 1976, fishery targets have been set by modeling the maximum sustainable yield (MSY) that keeps a population at an optimal spawning stock biomass (SSB). In unfished populations, biomass would plateau as competition for prey and habitat result in a maximum carrying capacity. The MSY is the peak of the curve at which fisheries stocks are kept at a maximum population growth rate by fishery harvest, although interaction effects make this target difficult to achieve. Life history theory has also been incorporated into management such that individual life stages, particularly those that are vulnerable, can be managed independently. Transition matrices can be used to model probabilities of transition from one life stage to another (e.g., juvenile to spawning adult) as a function of fishery and natural mortality.

Community Ecology

Community ecology focuses on multiple species and their interaction strengths through competition or predation. As predation is a driver of mortality for the prey and fitness for the predator, it has become the dominant model for understanding and modeling species interactions. A classic example of how predation and competition can structure ecosystems is the overfishing of cod (*Gadus morhua*) in the northwest Atlantic, where human predators drove cod to functional extinction. Even with fisheries closures, cod have not recovered as would have been predicted by a single-species model because of competition with other predators in the ecosystem such as dogfish. Because cod were a key interactor in the northwest Atlantic food web, their reduction allowed other species to thrive and occupy their niche, further hampering their recovery. This example highlights both the need to consider multispecies interactions in addition to cod's top predator, human fishers.

Landscape Ecology

While landscape ecology may seem a misnomer for a subdiscipline of fisheries ecology, the focus on patterns in spatial structure and scale have become an important consideration in understanding both species distributions and interactions. Traditional landscape ecology examines two-dimensional distribution in the terrestrial environment. In marine systems, depth becomes an important dimension. Ecological processes are inherently scale dependent both in their interactions and the way they are measured. For example, small zooplankton react to microturbulence in the water at the scales of centimeters, while large fish interact with fronts and eddies at the scale of 100–1000 meters. Moreover, because prey and predator often aggregate at different scales, their interactions can alternate from in and out of phase, dependent on the scale of observation. Understanding these scales of interaction and applying appropriate measurements and models is necessary to move toward spatially explicit ecology and management.

Ecosystem Ecology

Ecosystem ecology has focused on connecting multiple species interactions with physical forcing mechanisms with the goal of understanding how species affect and are affected by the environment. Ecosystem models have used energy flow among trophic levels as a currency to understand ecosystem change. Examples include the effects of direct fishery removal and answering scenario-based questions such as how management policies or environmental change could affect ecosystems of interest. Examples include Ecopath and Atlantis, both of which allow scenarios to propagate through the food web.

Social-Ecological Systems

In the past decade, human dimensions have been recognized as an often-overlooked component of ecosystems, both as drivers of change and as users of ecosystem services. Fishers have been modeled as economic actors, a more complex predator that seeks out prey (fishery landings) with an informed knowledge and cost structure (e.g., license and fuel costs). Economic-based behavior models can directly inform management decisions, creating a more holistic view of fisheries management as managing fishers and the fishery in concert. The future of fisheries ecology will require merging the theoretical with the applied using relevant spatial and temporal frameworks to create interdisciplinary fisheries management tools that focus on the ecosystem from physical processes to human behavior.

CHALLENGES IN FISHERIES ECOLOGY

Direct Effects of Fishing

Fishing has been one of the earliest forcing effects and has had the greatest impact on marine ecosystems. Between 50–90% of top predators have been removed from the world's oceans, including species that are not targeted by the fishery, such as a number of shark species. Many top predators such as elasmobranchs and marine mammals are long lived and have slow reproductive rates, making sustainable harvest near impossible. The longer it takes

an individual to become reproductive, the likelihood increases that individuals will be caught before they have a chance to reproduce. As top predator biomass declines, fishing pressure often shifts to lower trophic levels to maintain harvest rates, a process that Daniel Pauly has termed "fishing down marine food webs." Many of the mid-trophic species such as sardines and anchovies are important forage fish that serve as primary prey resources in pelagic food webs. Fisheries targeting mid-trophic-level species can result in user conflicts between fisheries and protected species needs. For example, Steller sea lions (*Eumetopias jubatus*) in the western Aleutian Islands are listed as endangered because they have declined over 40% between 2000 and 2008 following severe stock reductions from fur harvest. In 1999, federal managers closed pollock (*Theragra chalcogramma*) fishing in Steller sea lion critical habitat, and in 2010 have suggested closing Pacific cod (*Gadus macrocephalus*) and Atka mackerel (*Pleurogrammus monopterygius*) because they are a primary prey item of juvenile and adult sea lions.

Fishing pressure leads to direct biomass reduction of target species but can have other direct effects. Because most fisheries have size limits, catching only large fish can have population-level effects. In 2007, Law described the resulting intense selection pressure that can alter the genetic composition of fished populations. Populations mature earlier and at smaller body size in response to years of size-selective fishing. As only large individuals are caught in trawl nets, slow-growing and smaller fish that can escape the mesh have increased reproductive success. In the northwest Atlantic, Atlantic cod mature earlier and at smaller body size in response to years of fishing for larger individuals. Changes in population genetics may not be easily reversible and will require an evolutionary perspective to help long-term protection of species and maximization of yields.

Indirect Effects of Fishing

When fisheries target species that are strong interactors, extreme fishing pressure can result in trophic cascades, regime shifts, or in worst cases, ecosystem collapse. In diverse systems, competitors can proliferate and replace the niche that was made available by species removal. Reduction in apex predators can relieve both competition and predation pressure from mid-level predators resulting in an abundance of mid-trophic species, a process that has been termed mesopredator release. Although mid-tropic-level species are usually less valuable from a fisheries-removal perspective, their release can minimize cascading effects through the food web. In extreme cases, fisheries can trigger regime shifts, particularly in systems
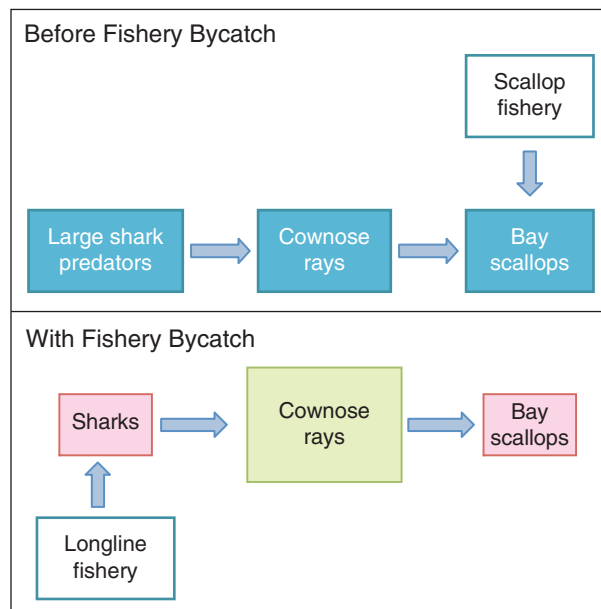

FIGURE 2 Representation of a trophic cascade from reduction of sharks in coastal North Carolina. Each box represents a species or a functional group, with arrows representing predation pressure. As longline fishery effort increased, declines in large sharks resulted in increases in small sharks, skates, and rays. With increased mesopredator abundance, bay scallop abundance crashed, resulting in the closure of the long-standing fishery.

with low functional diversity, when a strong competitor or keystone species is removed from the food web.

Trophic cascades can restructure entire ecosystems and can have unforeseen effects on fisheries landings (Fig. 2). In 2007, Myers and colleagues observed a trophic cascade following the removal of a complex of top predators in the western Atlantic. High bycatch rates in coastal fisheries led to population declines in seven species of coastal sharks of 87–99% from 1970 to 2005. With the functional removal of these top predators, predation pressure was released on smaller elasmobranchs (skates, rays, small sharks) in the mid-Atlantic. During the same period of large shark decline, the smaller elasmobranchs increased at a rate of 1.2 to 23% per year. The cownose ray (*Rhinoptera bonasus*) undergoes a mass migration southward to Florida wintering grounds each autumn and feeds primarily on bivalves in shallow water. With mid-trophic-predator increases such as the cownose ray, bay scallop (*Argopecten irradians*) population numbers crashed, leading to the closure of a century-old fishery in 2004. Predator exclusion experiments confirmed the cownose ray migration increased predation pressure on the bay scallop to the point of reproductive failure and fishery closure.

In addition, when ecosystem engineers are lost, such as algal grazers, food webs can shift from a top-down to

bottom-up control. In 2006, at sites where herbivorous grazers were low in abundance, Newman and colleagues found that there was an increase in fleshy algae growth that can cover and smother coral reefs. Instead of top-down control where predators and grazers maintain a coral-dominated ecosystem, the system can become bottom up where algal density is nutrient driven. Although complex food webs are more resilient to change, reduction in diversity leaves them vulnerable to collapse and they can be difficult if not impossible to recover.

In addition to harvesting top predators, overharvesting of important prey resources such as forage fish can decrease prey availability to top predators. If there is no excess production in forage fish, top predators and fisheries can compete for resources. In the highly productive Peruvian upwelling system, sardine (*Sardinops sagax*) and anchovies (*Engraulis ringens*) cycle based on the strength of upwelling, driving the distribution of top predators. During El Niño years, a lack of offshore transport results in warm nutrient-poor water dominating the coastal environment and leading to reduced productivity and stock biomass of forage fish. This reduction of an important forage base echoes through the food web and results in a decline in seabirds and marine mammal predators. Fishery landings also are proportional to anchovy and sardine populations, but the competition between fisheries and top predators can lead to further declines in top predator numbers. These interactions between fishery needs and marine food webs echo the need to manage the ecosystem as a whole.

Bycatch

Bycatch includes all target and nontarget species that are inadvertently killed, injured, or otherwise incapacitated but not retained. If fishers retain nontarget species, they are considered catch. For some marine species such as loggerhead turtles (*Caretta caretta*), bycatch mortality encompasses nearly all of their fishing-related mortality. From a population dynamics perspective, bycatch-associated mortality is indistinct from direct harvest mortality. For bycatch often more than target catch, it is challenging to determine the impacts on marine ecosystems because less information is available regarding the magnitude of removals. Life history characteristics (growth, age at reproduction) can be better predictors of fishing impact than trophic level, and behavioral characteristics of the species may play a major role in their vulnerability, such as birds that are attracted to discarded bycatch or baited longline hooks. In examples where bycatch species are slower growing and later maturing than target catch, their removals can exceed mortality limits and trigger management action before target populations meet quota. This effect can place bycatch taxa at risk of inadvertent overexploitation in even well-managed fisheries.

Seabirds, sea turtles, marine mammals, and chondrichthyans are among the most bycatch sensitive of the long-lived taxa. Although few fisheries target chondricthyans (sharks, rays, and chimaeras), they are common in bycatch and are increasingly retained in many fisheries. Global reported landings have increased steadily since 1984, and numerous studies have found declines in abundance exceeding 95%. For example, shark species are commonly caught as bycatch in the northwest Atlantic longline fishery. In 2003, Baum and colleagues used logbook data to examine rates of population decline in sharks from 1986 to 2000. All of the species caught including coastal and oceanic species showed extreme population declines (60–89%), including localized extinctions such as white sharks (*Carcharodon carcharias*) off the coast of Newfoundland. In some fisheries, juveniles of target species are considered bycatch because when caught below the legal limit, they will be discarded with many injured or already dead. From a fisheries management perspective, the main difference between bycatch and target catch is simply a difference in available data.

Recreational and Artisanal Fisheries

Recreational and commercial fisheries are fundamentally different and present distinct management challenges. Unlike most commercial fisheries, recreational fisheries remain open access with little data available to managers. While most U.S. states issue saltwater fishing licenses, there is generally no mechanism to limit total effort, making it difficult to regulate or reduce total harvest. In many regions, recreational fisheries lack consistent data collection on effort, landings, discards, and expenditures. In the United States, the National Marine Fisheries Service currently uses telephone and intercept surveys to gather data. But the high volume and diffuse nature of angler effort presents sampling challenges, making it difficult for managers to monitor trends in real time. Recreational anglers are also driven by entirely different motivations than their commercial counterparts. Managers often suffer from the mistaken assumption that recreational fisheries are self-regulating and that anglers will exit the fishery if the fishing experience somehow declines in quality. But if anglers are willing to tolerate low catch rates, recreational fishers may be more likely to exploit fisheries to the point of collapse.

The impact of large numbers of small-scale fishers can be rapid and similar in magnitude to commercial fisheries. While artisanal fisheries often operate in a smaller geographic area than their commercial counterparts, the near-shore environment can still suffer from local depletion of fish stocks, habitat damage, and bycatch. In 2007, Peckham and fellow researchers in Baja California Sur, Mexico, found a high degree of overlap between loggerhead sea turtle (*Caretta caretta*) hotspots and small-scale fishery effort. After assessing bycatch rates in bottom-set gillnets and pelagic longlines, two small-scale fisheries had loggerhead turtle bycatch rates (individuals per hook) an order of magnitude greater than industrial longline fisheries. Yearly catch was on the order of 1000 loggerheads, a comparable number to the 1300 loggerheads taken per year in the North Pacific industrial longline fleet. While artisanal fisheries typically do not operate crushing or habitat-altering gear such as trawl nets or dredges, they can damage key structure-forming organisms when walking on coral reefs or using spears or nets. In order to manage global fisheries and the cumulative effects from anthropogenic impacts, better data on artisanal fishing effort, catch rates, and bycatch numbers are needed.

The heterogeneity and complexity of small-scale fisheries provide challenges to assessing their impacts and contrast directly with large commercial-fishing operations. Consequently, mitigation of the impacts of artisanal fisheries proves challenging and may be region and/or gear specific. There is strong evidence for impacts on community structure and trophic relationships, but the general direction and processes do not differ from other sectors of the fishing community. Artisanal fisheries may be somewhat unique with respect to (a) predominance in the inshore zones, (b) multispecies catch and low discard ratios, and consequently (c) low nutrient input from discards to other marine fauna. Better management of post-harvest losses could reduce the demand and subsequent rate of removals. Where artisanal fisheries overlap with intrinsically vulnerable taxa such as sea turtles or marine mammals, a large number of small-scale fishers can rival the impact of commercial fisheries and generate rapid declines.

However, some of the best examples of fisheries management also come from small-scale fisheries where communities manage their marine and terrestrial systems in concert and as a renewable resource rather than a rush to catch fish. These management principles have been incorporated in commercial fisheries where fishers are given "ownership" of the resource on a yearly basis. This approach works both by preventing overcapacity in the fishery and by providing security to have fishing rights in future years and providing a sense of stewardship.

Fishery Recovery

Unfortunately, few examples exist of successful recovery of heavily exploited fishes. A 2004 study by Caddy and Agnew showed that the North Atlantic swordfish recovery was successful because depletion was not excessive (70% of the maximum sustainable yield), recovery time was relatively short (10 years), and the fishing fleets cooperated in management due to their similarities and shared incentives. Otherwise, evidence for recovery of long-lived, slowly maturing top predators is limited and dependent on many factors, both biological and political. Scientists are beginning to agree that the escalating crisis in marine ecosystems is in large part a failure of governance rather than a failure in fishers. Management structures such as catch shares provide a degree of ownership to the fishery and incentives toward preserving future catch rates. Techniques that optimize fishery involvement and consequently incentives toward future sustainability may be a large part of the solution. Many recent assessments and government objectives have called for a transition from managing sectoral activities, including fisheries, toward ecosystem-based management. Environmentalists have sought to implement marine reserves to maintain the structure and function of marine ecosystems. But this too is a sectoral approach. Traditional single-species management has a clearer recovery goal, specifically a certain spawning stock biomass to support future fishing efforts. But it is more difficult to define recovery goals in an ecosystem framework.

THE ROLE OF MANAGEMENT IN FISHERIES ECOLOGY

Traditional Fisheries Management

Single-species management relies on finding the inflection point where fish stock growth slows due to competition for food and habitat at high population sizes. Fisheries quotas are set and divided among commercial license holders for the duration of the fishing season. Once the quota is reached or the season ends, the fishery is closed. Even with appropriate precautionary quotas, overcapacity in fleets and a race to reach fisheries quotas can lead to overfishing. Changes in governance that limit access could both reduce overall fishing pressure and increase catch per unit effort for the remaining fishers.

Although almost 1/3 of the world's stocks have been estimated as overfished, there are examples of successfully managed and sustainably certified fisheries, such as halibut and numerous salmon species in Alaska.

These fisheries have limited bycatch, and the freshwater migration in salmon life history improves data access for stock assessments. However, natural variability in population dynamics, compounding stressors, and economic considerations can add complexity and can hamper successful fisheries management. The lack of sufficient data is still a large problem resulting in many poorly understood and poorly managed fisheries. Fish stocks often span management and governmental boundaries requiring multijurisdictional and international cooperation to adequately manage migratory stocks. Management techniques such as ecosystem-based management and integrated ecosystem assessments aim to incorporate multiple-species interactions and are not limited by economic boundaries.

Catch Shares

Catch shares are a form of single-species management that provides additional incentives to fishers to harvest sustainably. Rather than set an individual quota per license, each fisher or cooperative of fishers purchases a share of the fishery, essentially providing ownership of their quota. Catch shares have been implemented in various forms, including individual transferable quotas (ITQs), fishery cooperatives, and spatial management rights. Catch shares can be successful because they serve as a rights-based management strategy providing fishers with a dependable asset. It is important to note that catch shares are an implementation of management targets and would be compatible with ecosystem approaches.

Ecosystem-Based Management

Ecosystem-based management (EBM) focuses on the entire ecosystem across multiple trophic levels and includes humans. The basic tenet of EBM is to manage in concert how interannual and decadal environmental variability and fishery pressure will affect other species in the ecosystem, adjacent ecosystems when appropriate, and fisheries economics and behavior. In addition to managing biomass, it is important to conserve biodiversity, food web structure, and ecosystem function. The entire ecosystem needs to be managed in an economic context as well, to understand how ecosystem effects will translate to fishers' livelihoods. Because there are many users of the ecosystem, a successful ecosystem approach must include all relevant sectors of society and to be truly interdisciplinary.

Measuring the ecosystem effects of fishing is a daunting task, as many abiotic, ecological, and anthropogenic factors act in synergy. The development of marine food web models (e.g., Ecopath/Ecosim, Atlantis) has allowed fisheries ecologists to assess the anthropogenic impacts on marine ecosystems at temporal and spatial scales that are too large and complex for experimental studies. Because variability in climate forcing can propagate up food webs to affect higher trophic levels, management needs to set targets that incorporate predicted biological responses to climate forcing even though predictions can be difficult. Widespread use of single-species harvesting policies based on maximum sustainable yield can lead to removal of a suite of top predators and extensive loss of ecosystem function. Despite advances in ecosystem approaches, applying robust models to improve ecosystem-based management of commercial fisheries still requires investment in the collection of sufficient biological and environmental data in addition to extensive model-validation exercises.

Marine Spatial Planning

Place-based management and marine spatial planning (MSP) can provide a far more promising approach to implementing ecosystem-based management. Rather than individual sectoral agencies managing their activities everywhere, responsible authorities could collaborate to manage all the human activities in a focused place. These places might align with ecosystem boundaries, socioeconomical boundaries, and/or jurisdictional boundaries. In practice, management always occurs in a delimited space, with processes that cross management boundaries.

The biophysical component of marine ecosystems provides the basic template on which all human activities, including fisheries, occur and that various forms of governance regulate. Approaches to MSP and ocean zoning consider basic ecological concepts so that human activities can be conducted in ways that maintain ecosystem functioning, provide sustainable ecosystem services on which people depend, and maintain resilient ecosystems that can respond to environmental change.

Place-based management of marine ecosystems requires a hierarchy of management practices starting at the most general level with the concept of ecosystem-based management and moving toward the development of an integrated approach that accords priority to the maintenance of healthy, biologically diverse, productive, and resilient ecosystems. The key to success in place-based management of marine ecosystems is to design governance systems that align the incentives of stakeholders, in this case fishermen, with the objectives of management. MSP that fully incorporates the underlying ecosystem template and explicitly integrates the socioeconomic and governance overlays

can form the basis for adequate protection of marine ecosystems and the sound use of marine resources, including fisheries.

FISHERIES ECOLOGY MOVING FORWARD

As fisheries management continues the transition to ecosystem-based management spanning physical variability to human dimensions, fisheries ecology will need to become more holistic and interdisciplinary. Recent National Ocean Policy (NOP) reports have tasked managers with developing ecosystem approaches to management that incorporate species interactions and translate to ecosystem services. New tools and research approaches are required that can evaluate management scenarios and their downstream effects on ecosystem function and services.

Fisheries ecology research includes theoretical developments, field-based experiments, and meta-analyses of fisheries data. Research on food web dynamics, particularly under changing climate regimes, will be critical in informing ecosystem models. Field-based experiments that can be built into long time series also allow comparisons of species interactions, abundance, and distribution before and after management decisions or climate regime shifts. These indicators will become critical in detecting ecosystem-wide perturbations, predicting future ecosystem dynamics, and parameterizing mass balance models to predict ecosystem changes. Mass balance models (e.g., Ecopath/Ecosim, Atlantis) are our primary tool for measuring trophic transfer for an entire food web and forecasting change under varied management and climate scenarios. Spatially explicit distribution models (e.g., linear models, generalized additive models, neural networks) can be used to identify biotic and abiotic drivers of distribution and can identify critical spatial scales of distribution and ecological processes. Spatially explicit approaches become necessary for identifying ecological and anthropogenic patterns at relevant scales to inform marine spatial planning.

Ultimately, coastal MSP will require tradeoff-based interaction models incorporating physical forcing up to economics and human behavior. Ecosystem-based management has evolved from an idea into a management framework, yet tools are still being developed to manage entire ecosystems in addition to individual stocks. In 2009, Levin and colleagues proposed integrated ecosystem assessments (IEAs) as a framework to aggregate scientific findings and to inform EBM at various scales and across sectors, although few have been developed to date. More specifically, IEAs provide a five-step pathway to quantitatively assess physical, biological, and socioeconomic factors in concert with defined ecosystem management objectives. This five-step transition from (1) scoping, which involves scientists, stakeholders, and managers identifying key EBM goals, (2) indicator development, which serve as representative proxies for the overall ecosystem state, (3) risk analyses, which assess the risk human activities and natural processes have on indicators, (4) management evaluation, which uses ecosystem models to evaluate outcomes and tradeoffs from proposed management strategies, and (5) ecosystem monitoring and evaluation, which measures the ongoing effectiveness of implemented management at appropriate scales and includes interactions among indicators.

Stakeholders, economists, and social scientists need to play a role in each step of the IEA to ensure the effect of management decisions on fishers and communities remain central goals of ecosystem-based management. More research is desperately needed on recreational and artisanal fisheries, as both target catch and bycatch from these fisheries are often absent in EBM. The field of fisheries ecology is in transition from primarily reductionist single-species approaches to integrated multiple-species responses forming a broader, interconnected framework that crosses multiple disciplines, management sectors, and jurisdictional boundaries to ensure sustainable fisheries for future generations.

SEE ALSO THE FOLLOWING ARTICLES

Beverton–Holt Model / Ecological Economics / Ecosystem Services / Food Webs / Marine Reserves and Ecosystem-Based Management / Population Ecology / Ricker Model

FURTHER READING

Caddy, J. F., and D. J. Agnew. 2004. An overview of recent global experience with recovery plans for depleted marine resources and suggested guidelines for recovery planning. *Reviews in Fish Biology and Fisheries* 14: 43–112.

Hilborn, R. 2007. Reinterpreting the state of fisheries and their management. *Ecosystems* 10: 1362–1369.

Law, R. 2007. Fisheries-induced evolution: present status and future directions. *Marine Ecology Progress Series* 335: 271–277.

Levin, P. S., M. J. Fogarty, S. A. Murawski, and D. Fluharty. 2009. Integrated Ecosystem Assessments: developing the scientific basis for ecosystem-based management of the ocean. *PLoS Biology* 7: 23–28.

Lubchenco, J., and N. Sutley. 2010. Proposed U.S. policy for ocean, coast, and great lakes stewardship. *Science* 328: 1485–1486.

Magnuson, J. T. 1991. Fish and fisheries ecology. *Ecological Applications* 1: 13–26.

Myers, R. A., J. K. Baum, T. D. Shepherd, S. P. Powers, and C. H. Peterson. 2007. Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* 315: 1846–1850.

Pauly, D., V. Christensen, J. Dalsgaard, R. Froese, and F. Torres. 1998. Fishing down marine food webs. *Science* 279: 860–863.

Walters, C. J., V. Christensen, S. J. Martell, and J. F. Kitchell. 2005. Possible ecosystem impacts of applying MSY policies from single-species assessment. *ICES Journal of Marine Science* 62: 558–568.

Worm, B., et al. 2009. Rebuilding global fisheries. *Science* 325: 578–585.

FOOD CHAINS AND FOOD WEB MODULES

KEVIN McCANN AND GABRIEL GELLNER

University of Guelph, Ontario, Canada

Ecological systems are enormously complex entities. In order to study them, ecologists have tended to greatly simplify and deconstruct food webs. A result of this simplification has been the development of modular theory—the study of isolated subsystems. Figure 1 shows a common three-species subsystem (the food chain) and a common four-species subsystem (the diamond module) extracted from a whole food web. This entry briefly reviews modular theory with a focus on these two ubiquitous modules. It ends by discussing how the results from this modular theory hope to ultimately contribute to a theory for whole systems.

MODULES AND MOTIFS: TOWARD A THEORY FOR WHOLE SYSTEMS

The term module has been around for some time, although it appears to have had subtly different meanings. Robert Paine first used the term module to describe a subsystem of interacting species; he considered a module as a consumer and its resources that "behave as a functional unit." However, the term did not initially seem to catch on, and it was later reintroduced to ecology in slightly different form by Robert Holt in order to facilitate theoretical development beyond the well-developed theory of single-species and pairwise interactions. Holt defined modules specifically as communities of intermediate complexity beyond the well-studied pairwise interactions but below the diversity found in most natural systems (e.g., combination of 3–6 interacting species).

In a sense, the modular approach discussed by Holt sought to ask if we can extend predator–prey and competition theory in a coherent fashion to small subsystems. Similarly, network theorists—looking for properties in all kinds of nature's complex networks—have simultaneously considered the idea of underlying modular subnetworks that form the architecture for whole webs. Their terminology for this subnetwork topology is motif. Generally, these motifs are defined for two-species interactions (e.g., consumer–resource interactions), three-species interactions (e.g., food chains, exploitative competition, appar-

ent competition, and the like), and beyond. Each motif is an i-species class characterized by the set of all possible interactions within that group. In this case, the single node is of little interest in quantifying network structure since it is everywhere by definition. Nonetheless, ecologists recognize that single nodes (i.e., populations) are dynamically important (e.g., cohort cycles). Further, there is a well-developed dynamic theory for populations in ecology.

In what follows, we will use the most general definition of modules to represent all possible subsystem connections, including the one-node/species case through to the n-node/species cases. In an attempt to delineate the different terminology, Robert Holt has argued that he sees modules "as motifs with muscles." This is reasonable since Holt's modular theory has always sought to understand the implications of the strength of the interactions on the dynamics and persistence of these units. The term module here, therefore, will be used to mean all motifs that include interaction strength.

Network and graph theorists have added to our ability to categorize the structure of real food webs. Network theorists have developed techniques that allow us to rigorously consider which motifs are common—or in the language of network theory, which motifs are over- or underrepresented in nature. Here, as with much of ecology, one compares nature to some underlying null model, and so one must remain cautious about what overrepresentation actually means. Nonetheless, this technique allows researchers to quantify the relative presence of motifs and focus the development of modular theory on common natural subsystems (i.e., overrepresented modules of real webs).

It is natural for ecological theory to move beyond consumer–resource interaction in order to explore common subsystems of food webs. In fact, Robert May, who championed the classical many-species whole food web matrix approach, argued that models of intermediate complexity may be a more direct path to interpreting how food web structure influences population dynamics and stability than matrices.

This entry first briefly reviews consumer–resource theory (hereafter *C–R* theory) before discussing two common higher-order modules, the three-species food chain and the four-species diamond module. The dynamics of the food chain module, which appears to be the most often overrepresented of all three-species modules in food webs, is reviewed, followed by a consideration of the common four-species diamond module, which is, in a sense, simply two food chains
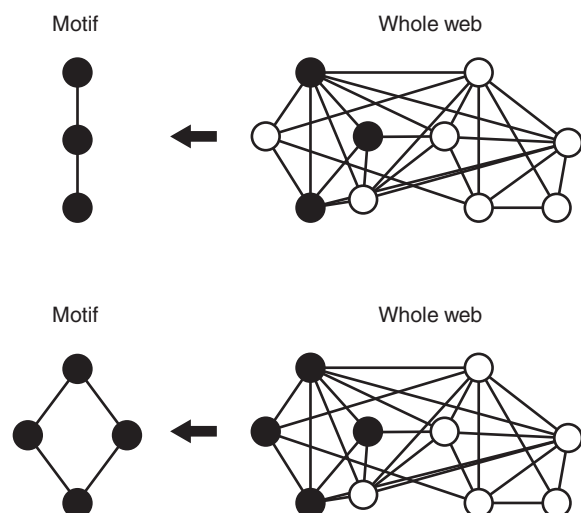
FIGURE 1 Food webs can be deconstructed into subsystems or motifs. Here, two common motifs are shown. A "common" structure means that these motifs appear with greater frequency than expected by pure chance. Different types of networks have different types of ubiquitous structure. It has been found, for example, that information processing networks are different in structure than energy processing–based networks.
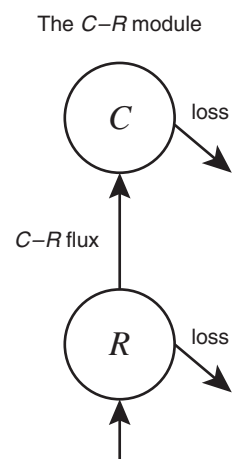

FIGURE 2 The *C–R* interaction module one of the basic building blocks. The basic fluxes are shown: the *C–R* flux (consumption) and the loss rates. Ultimately, flux in and flux out are related to any reasonable metric of direct interaction strengths.

concatenated (Fig. 1). This module also appears common in natural systems.

All the while, attention is paid to existing *C–R* theory in order to interpret the dynamic consequences of these modules. Food web modules have the tendency of displaying subsystem signatures such that oscillations can often be attributed to particular underlying *C–R* interactions. Given this, *C–R* theory forms a framework that allows the examination food web module dynamics from an interaction strength perspective.

A FUNDAMENTAL MODULE: CONSUMER–RESOURCE INTERACTIONS

Figure 2 shows the material fluxes that accompany a consumer–resource interaction in nature. There exist fluxes between the consumer and its resource (the interaction itself), there are fluxes into the resource (nutrient or biomass uptake), and there are also fluxes out of the consumer and the resource (e.g., mortality). These *C–R* fluxes occur repeatedly in all the underlying *C–R* modules that ultimately make up whole food webs (e.g., Fig. 1). These fluxes show that consumer–resource theory yields a very general and powerful result when considered from a flux-based interaction strength perspective.

Specifically, there is a tendency for all *C–R* models to produce a destabilizing response to increased production, increased interaction rates, or decreases in mortality. Figure 3 shows the dynamic response across a gradient in interaction strength. The gradient follows the ratio of *C–R* flux relative to the consumer loss term. *C–R* models that have high *C–R* flux rates and low mortality loss rates tend to be less stable. "Less stable" warrants some clarification since stability is often assessed in a multitude of ways. Figure 3A expresses this loss of stability as the change from equilibrium dynamics (for low *C–R* flux:loss ratios) to wildly oscillating dynamics at high *C–R* flux:loss ratios. As an example, this result of increased variability with increased flux-based interaction strengths generally occurs for the well-known Rosenzweig–MacArthur model.

Figure 3B, on the other hand, shows a different but very related reduction in stability. Here, solutions remain in equilibrium dynamics, but the return time back to the equilibrium increases as the *C–R* flux rate:loss ratio increases. Part of this increase in return time is due to the fact that the solutions take on oscillatory decays (i.e., eigenvalues become complex). This increased return time also can be seen in that the negative real part of the eigenvalues of the stable equilibria move closer to zero (i.e., weaker attraction to equilibrium). This increased oscillatory decay, and weakened attractor, drive a greater return time to equilibrium and so reduces stability. This destabilization result is qualitatively similar the Rosenzweig–MacArthur model result of Figure 3A. Note that these oscillatory decays can quite readily turn into cycles, or quasi-cycles, in stochastic settings.

There is a large food web literature that employs matrices to examine dynamics. It is of interest to ask whether
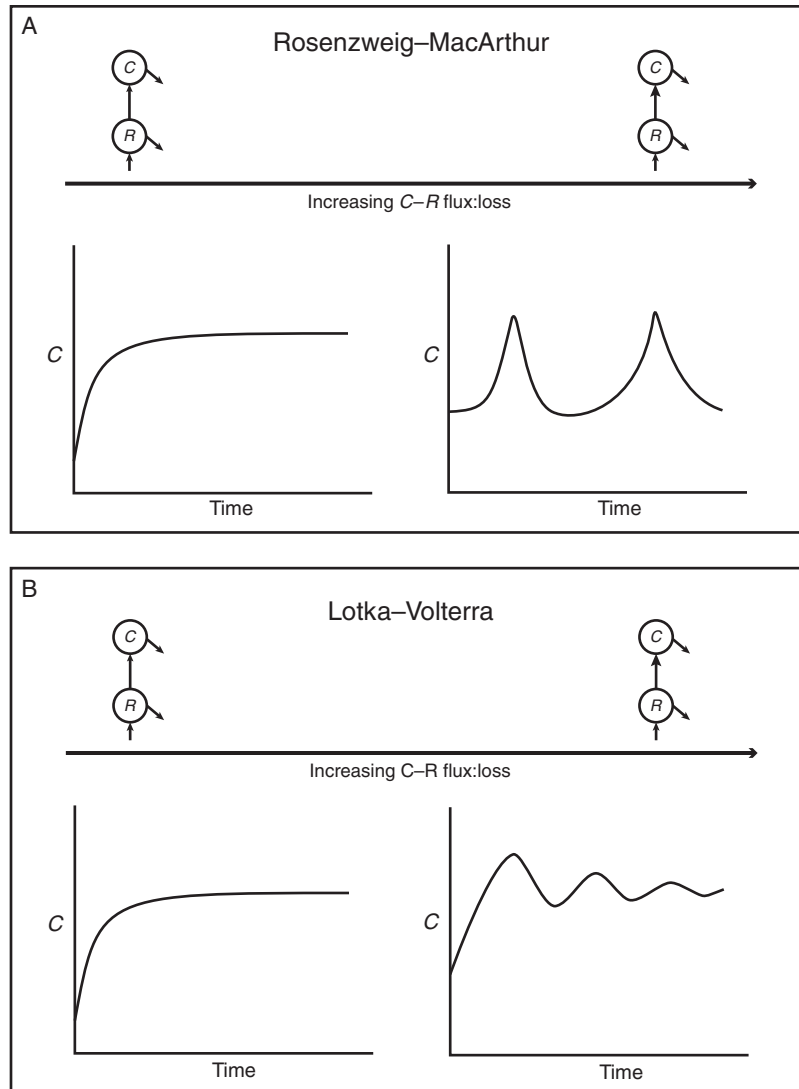
FIGURE 3 The influence of changing interaction strength (here defined as *C–R* flux: *C* loss ratio) on the dynamics of the module. Increasing *C–R* flux:loss ratio destabilizes the interaction by (A) yielding wild oscillations as in the Rosenzweig–MacArthur model, or (B) yielding oscillatory decay as in the Lotka–Volterra model.

the *C–R* matrix approach gives similar answers to the *C–R* theory discussed above. Here, a compatible result is also found since the elements of the matrix can be decomposed in a similar manner (i.e., interaction terms and loss terms). In the *C–R* matrix, indeed for all matrices, the diagonal elements of the matrix are measures of the loss rate, while the off-diagonal entries are measures of the fluxes through the consumer–resource interactions. Analogous to the previous results, increasing the off-diagonal interaction elements relative to the diagonal loss terms tends to decrease the linear stability of the system up to some maximum. This is satisfying, as it means that different mathematical approaches yield similar results.

The above results, taken together, can be stated biologically as follows:

THE PRINCIPLE OF INTERACTION STRENGTH (1)

Any biological mechanism that increases the strength of the flux through a consumer–resource interaction relative to the mortality rate of the consumer ultimately tends to destabilize the interaction.

The degree of this destabilization will necessarily depend on the model formulation, but all models tend to follow this general pattern. This result can be stated differently, and to some advantage for the remaining sections of this entry, as follows:

Any biological mechanism that decreases the strength of the interaction relative to the mortality tends to stabilize that interaction.

The next section turns to modular theory to point out how *C–R* theory explains the dynamics of more speciose food web modules.

FOOD CHAINS: COUPLED *C–R* INTERACTIONS

There exists a lot of theory describing the dynamics of food chains. Researchers, for example, have shown that it is relatively easy to find complex or chaotic dynamics in a common food chain model. Specifically, these researchers have found chaos readily occurrs in the Rosenzweig–MacArthur food chain model when the attack rates were high relative to the mortality rates. This result is strongly related to the interaction strength principle (1) of *C–R* theory stated in the last section. Note that in such a case the underlying interactions that comprise the food chain (i.e., both predator–consumer (*P*–C and *C–R* interactions) would tend to produce cycles if isolated (Fig. 3A).

It is important to realize that chaos often emerges when multiple oscillators interact. Multiple oscillators, once coupled, tend to produce a complex mixture of their underlying frequencies (e.g., Fig. 4A). In the case of the *P–C–R* food chain, there are two underlying oscillators, a predator–consumer oscillator (*P–C*) and the consumer–resource oscillator (*C–R*), which mix to drive complex dynamics (Fig. 4A). The signatures of these underlying oscillators in chaotic attractors are present and visible in the power spectrum of their time series. These results resonate with the interaction strength principle (1) of *C–R* theory stated above.

It becomes interesting to consider what happens when a weaker interaction (i.e., low *C–R* flux:loss ratio) is coupled with a stronger interaction (i.e., high *C–R* flux:loss rate). Figure 4B depicts such a case. Then there is a potentially oscillatory interaction (the *C–R* interaction) and an interaction that would not oscillate in isolation (the *P–C* interaction). This case introduces an example of weak interaction theory. The *C–R* interaction strength corollary (2) reviewed in the previous section allows the interpretation of the results that follow.

Note that in Figure 4B, the stable *P–C* interaction, in effect, adds an additional mortality loss to the consumer (*C*). Thus, the *C* in the *P–C–R* chain experiences more loss than such a consumer would experience in isolation. All else being equal, from the interaction strength corollary (2) it would be expected that any such increase in loss to the consumer would tend to stabilize the underlying
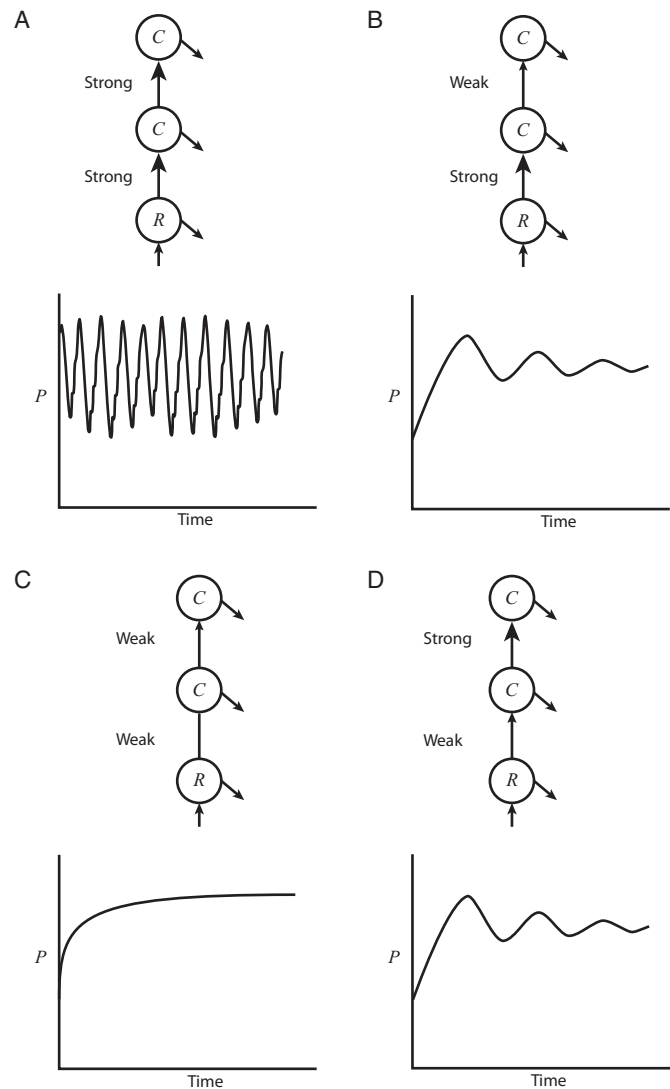

FIGURE 4 The influence of changing interaction strength on the dynamics of the food chain module. (A) Two strong *C–R* interactions couple to produce wild oscillations as in the Rosenzweig–MacArthur model. (B) The weak *P–C* interaction mutes (increases loss) without adding another oscillator and so stabilize the *C–R* interaction. The resultant *P–C–R* food chain dynamic, in this depicted case, is a stable equilibrium, although an oscillatory approach to equilibrium occurs. (C) All interactions are weak and the dynamics are stable. (D) Weak *C–R* interaction nullifies potentially strong *P–C* interaction, yielding stable dynamics.

C–R interaction (Fig. 4B). Indeed, this is what generally happens when a weakly consuming top predator, *P*, taps into a strong *C–R* interaction. The interaction gets less oscillatory—it mutes the *C–R* oscillator—and can even produce stable equilibrium dynamics.

It is worth pointing out that a range of food chain dynamics can be thought of from this very simple, but powerful, perspective. For example, two stable underlying interactions, the *P–C* and the *C–R* interaction, tend to yield stable equilibrium food chain dynamics (Fig. 4C).

Further, and similar to the case discussed in Figure 4B, a weak interaction in the basal C–R can mute a potentially strong P–C interaction (Fig. 4D).

To understand this last result, consider a thought experiment starting with the scenario in which both underlying interactions in the food chain are strong. In other words, imagine starting with the case depicted in Figure 4A. Now, if the strength of the basal C–R interaction is experimentally reduced, the result becomes like the case depicted in Figure 4D. As the lower level C–R interaction is reduced, the amount of energy that reaches the P–C interaction begins to be reduced. As the corollary (2) states, any mechanism that reduces the flux:loss ratio ought to stabilize the underlying interaction. Therefore, one expects the potentially oscillatory P–C interaction to be muted by the reduced C–R interaction strength, and it is (compare Fig. 4A to Fig. 4D).

In summary, a weak C–R interaction means that little actual production reaches the top predator. As such, not enough biomass energy is transmitted up the food chain to drive the potential oscillations in the P–C interaction. In such a case, the weak basal interaction effectively mutes the potentially unstable higher-order food chain interaction. This is again the "interaction strength corollary" operating—a biological mechanism that weakens flux stabilizes a potentially unstable interaction. Properly placed weak interactions therefore can readily inhibit oscillatory dynamics.

Numerous fairly hefty mathematical papers that look at the bifurcation structure of this commonly explored model are consistent with the simple ideas presented here. These papers, however, catalog in more detail dynamical outcomes such as multiple basins of attraction, quasi-periodicity, and chaotic transients. Additionally, researchers are beginning to explore the role of stage structure on food chains, which appear to quite readily produce multiple basins of attraction.

HIGHER-ORDER MODULES: COUPLED FOOD CHAINS AS AN EXAMPLE

The results of the last section appear to work for simple food web modules. That is, multiple strong interactions coupled together readily beget complex dynamics. In addition, weak interactions properly placed can mute potentially oscillatory interactions. However, in more complex food web modules another stabilizing mechanism arises.

As an example, this other stabilizing mechanism arises in cases where predators, P, feed on consumers (C_1 and C_2) competing for a common basal resource (Fig. 5). This module is ubiquitous in real food webs and has become referred to as the diamond module. There is evidence
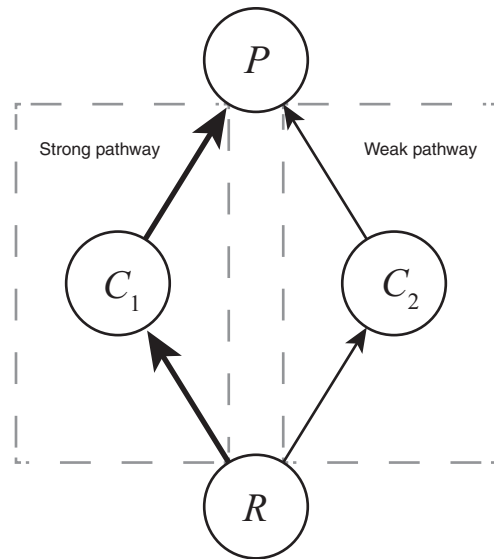

FIGURE 5 The diamond four-species food web module. One pathway is composed of strong interactions and the other, weak interactions due to very general growth–defense tradeoffs. In this food web module, the fast-growing species tends to be highly competitive but also very vulnerable to predation. This tradeoff produces relatively stable modules compared to just the strong food chain model.

that organismal tradeoffs play a major role in governing the diamond module since tradeoffs frequently manifest around parameters involved with consumption and growth rates. As an example, high tolerance to predation tends to correlate with a lower growth-rate life history.

This tradeoff occurs because organisms that have adapted to grow rapidly allocate energy to somatic and reproductive effort, with little to no energy allocated into costly physical structures that impede a predators' ability to consume them. Such a rapidly growing consumer is expected to be an extraordinary competitor in situations without a predator. Nonetheless, their ability to produce biomass also makes them vulnerable to predators. Thus, this same consumer species is both productive and potentially strongly consumed by a predator (i.e., the strong chain of Fig. 5)—it is the precise recipe for nonequilibrium food chain dynamics.

On the other hand, an organism that puts much energy into the development of defense structures (e.g., porcupine) may also be expected to grow more slowly and be less competitive when not in the presence of predators. Since predators have less impact on this organism, this slow-growing species will become an excellent competitor when predator densities are high. This species forms the middle node in the weak chain of Figure 5.

This tradeoff, therefore, creates a combination of strong interaction and weak interaction pathways (Fig. 5). In isolation, the strong pathway chain (P–C_1–R in Fig. 5)

tends toward instability, while the weak interaction pathway (P–C_2–R in Fig. 5) tends toward heightened stability (Fig. 4C). Together, they blend into a system that has far more stability than the strong interaction pathway alone.

This reduction of the whole system to coupled subsystems allows one to take a simplifying view of this otherwise staggering mathematical problem. If one can find a mechanism that tends to stabilize all the underlying oscillators, then this ought to eliminate the occurrence of oscillatory dynamics in the full system. Similarly, such a mechanism that reduces the amplitude of the underlying oscillators also ought to reduce the amplitude of the dynamics of the full system. Previous work suggest that this type of result may occur frequently.

Toward this effect, one can perform theoretical experiments on a strong focal chain by adding new C–R interactions, one at a time, until the diamond module is created. Figure 6B highlights the numerical experiment of adding exploitative competition (C_2) to a food chain model undergoing chaotic dynamics (Fig. 6A). In this case, C_2, is competitively inferior to C_1, so its ability to persist is mediated by the selective predation of the top predator, P, on C_1. Here, one expects exploitative competition to inhibit the oscillating C_1–R subsystem by deflecting energy away from the strong, potentially excited interaction. It does exactly this, producing oscillatory but muted or well-bounded limit cycles (Fig. 6B). In this particular example, the system still does not reach an equilibrium solution over this range, as the muting potential of the added competitor simply is not capable of deflecting enough energy to cause period-doubling reversals all the way to an equilibrium value.

Now, if one completes the diamond module and allows the top predator to also be a generalist and feed on C_2 weakly, then Figure 6C shows that this final addition has now produced a stable equilibrium dynamic. This final result is not immediately obvious and so a comment is in order. There is another somewhat hidden stabilizing mechanism in this last result. The differential-strength pathways tend to readily produce asynchronous consumer dynamics that the consumer can average over.

As an example, if one perturbs this system by adding additional basal resources, this first tends to increase both of the consumers. They increase synchronously, however; soon this C increase drives a decrease in total resource densities, R, and an increase in total predator densities, P. The system is suddenly quite tricky for consumers to eke out a living in, since there are simultaneously low-resource (high competition) and high-predation conditions.

In this precarious situation, the differential strength pathways produce asynchronous C dynamics. This occurs because as P grows, it starts to consume a lot of C_1, freeing C_2 from the suppressive grips of the superior competitor, C_1. Once freed from competition and yet not strongly consumed by the top predator, P, the weak-pathway consumer, C_2, starts to grow. Thus, one consumer is decreasing and the other consumer is increasing.
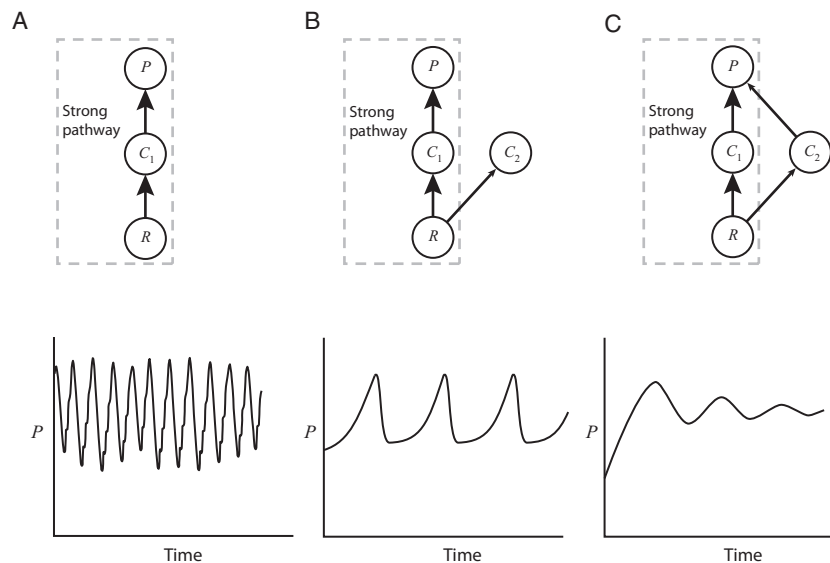
FIGURE 6 Time series for three different food web configurations that ultimately end up as the diamond module with weak and strong pathways. (A) Food chain composed of just strong pathways with complex, highly variable dynamics. (B) One weak competitive interaction added to the food chain mutes the oscillations, leaving a smaller, less complex oscillation. (C) Another weak interaction, completing the diamond module, changes the dynamics to stable equilibrium dynamics.

This out-of-phase consumer dynamic, in a sense, can sum together to give the top predator a relatively stable resource base. That is, when one consumer is low, the other is high, and so the top predator is buffered against the low resource densities of any one prey item. Clearly, this is a much more stabilizing situation than when both consumers are simultaneously low, in which case, the top predator tends to precipitously decline.

In summary, despite the complexity of this system, which includes multiple attractors and numerous bifurcations, the qualitative result remains: relatively weak links, properly placed, simplify and bound the dynamics of food webs. Here, "properly placed" is a result of the underlying biologically expected tradeoffs. On the other hand, strong interactions coupled together are the recipe for chaos and/or species elimination.

DATA AND THEORY

There are accumulating empirical and experimental results that appear consistent with this simple food web module theory. First, it is clear that *C–R* interactions, food chains, and simple food web modules can produce complex oscillatory dynamics in experimental settings. Further, it appears clear that increasing flux through these interactions by increasing nutrients tends to drive more complex dynamics. Thus, strong coupled interactions do indeed drive complex dynamical phenomena in highly controlled experiments.

Experiments have also begun to test whether weak interactions can mute such wild oscillations. Two recent controlled experiments have found that weak interactions in a simple food web module can decrease the variability of the population dynamics. One of these experiments found that some of this stability may also be due to the asynchronous resource dynamics discussed above. Along these lines, an extensive study based on field data from Lake Constance found that edible (strong) and less edible (weak) phytoplankton varied out of phase shortly after a synchronized pulse in zooplankton following spring turnover. Both controlled experiments and field data show signs that weak–strong pathways may produce asynchronous out-of-phase dynamic responses to increases in predator densities.

Finally, a recent analysis of interaction strengths in a Caribbean food web found the cooccurrence of two strong interactions on consecutive levels of food chains occurred less frequently than expected by chance. Thus, real food webs may not commonly produce strongly coupled food chains. Further, where they did find strongly coupled links in food chains there was an overrepresentation of omnivory. While not discussed here, this latter result may imply that omnivory acts in a stabilizing fashion in this case.

SEE ALSO THE FOLLOWING ARTICLES

Bifurcations / Bottom-Up Control / Chaos / Food Webs / Ordinary Differential Equations / Predator–Prey Models / Stability Analysis / Top-Down Control

FURTHER READING

Fussmann, G. F., S. P. Ellner, K. W. Shertzer, and N. G. Hairston. 2000. Crossing the Hopf bifurcation in a live predator–prey system. *Science* 290: 1358–1360.

Hastings, A., and T. Powell. 1991. Chaos in a 3-species food-chain. *Ecology* 72(3): 896–903.

Holt, R. D. 2002. Community modules. In A. C. Gange and V. K. Brown, eds. *Multitrophic interactions in terrestrial ecosystems*. The 36th Symposium of the British Ecological Society. London: Blackwell Science.

McCann, K., and P. Yodzis. 1995. Bifurcation structure of a 3-species food-chain model. *Theoretical Population Biology* 48: 93–125.

McCann, K. S., A. M. Hastings, and G. Huxel. 1998. Weak trophic interactions and the balance of nature. *Nature* 395: 794–798.

Muratori, S., and S. Rinaldi. 1992. Low- and high-frequency oscillations in three dimensional food chain systems. *SIAM Journal on Applied Mathematics* 52: 1688.

Rip, J., K. McCann, D. Lynn, and S. Fawcett. 2010. An experimental test of a fundamental food web motif. *Proceedings of the Royal Society: London B: Biological Sciences* 277: 1743–1749.

FOOD WEBS

AXEL G. ROSSBERG

Queen's University Belfast, United Kingdom

Food webs are the networks formed by the trophic (feeding) interactions between species in ecological communities. Their relevance for population ecology follows directly from the importance of trophic interactions for maintenance and regulation of populations. Food webs are crucial for community ecology whenever trophic interactions control competition, and hence community composition and species richness. Theoretical representations of food webs range from simple directed graphs to complex dynamic models integrating much of theoretical ecological.

STATUS OF THE THEORY

The study of food webs developed from a science that simply records data, through a phase of cataloging and identifying patterns in the data, and then moved toward interpreting data and patterns, first in terms of phenomenological models and later in terms of general ecological mechanisms. A stage in which the theoretical foundations would be settled and their implications studied in depth has not been reached, yet.

By now, a large number of food web models have been developed that combine elements from a certain set of recurring model components with specific original ideas. As these models are becoming more complex, work to systematically evaluate and compare their predictive power is limited by computational issues, differences in model focus, and deficits in the available data. There is currently no standard model for food web structure and/or dynamics. Understanding the roles played by different model elements in generating the patterns found in data remains a major theoretical challenge.

Fisheries management is perhaps the only area where progress has been made in applying complex food web models to real-world problems. While long-term forecasts of the dynamics of interacting, harvested fish populations are limited by high model sensitivity and the problem of food web stability (see below), food web models of fish communities help understanding medium-term changes in abundances, mortalities, and population growth rates.

BASIC ELEMENTS OF FOOD WEB MODELS

Topological and Quantitative Food Webs

Food webs are often represented as in Figure 1, by directed graphs in which species form the nodes and feeding interactions are represented by arrows (trophic links) pointing from resources to consumers, i.e., in the direction of energy flow. By convention, consumers are drawn above their resources whenever possible, so that energy flows upward and top predators are on the top.

Some empirical food webs restrict the set of species included to those that are directly or indirectly eaten by certain consumers (*sink webs*) or those directly or indirectly feeding on certain resources (*source webs*). Theory nearly exclusively considers *community food webs*, implying the idealizations of sharply delineated ecological communities and well-defined sets of member species.
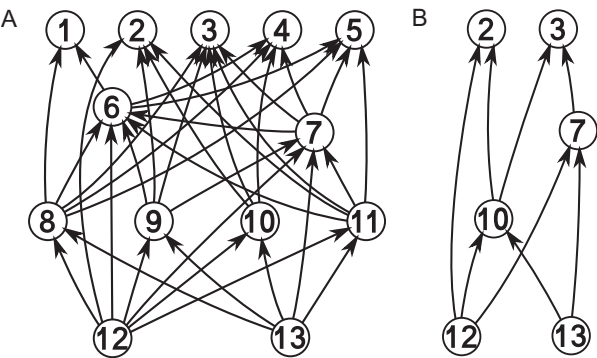

FIGURE 1 The food web of (A) Morton Creek (after G. Wayne Minshall, *Ecology* 48: 139–149, Fig. 3) and (B) the community that persists in the model described in Box 1.

A food web described exclusively by a species list and an adjacency matrices $A_{ij} = 0, 1$ or a corresponding graph is called a *binary food web*. This distinguishes *topological food webs* from *quantitative food webs* that are described by a link-strength matrix C_{ij} and, potentially, species abundances.

Link-Strength Functions and Trophic Niche Space

The trophic link strength C_{ij} between a resource i and a consumer j is often modeled as a function of characteristics of these two species. That is, one assumes that each species i is characterized by a set of vulnerability traits $\mathbf{v}_i$ and a set of foraging traits $\mathbf{f}_i$ such that, with an appropriate choice of the link-strength function $c(\cdot, \cdot)$, trophic link strength is given by $C_{ij} = c(\mathbf{v}_i, \mathbf{f}_j)$. Vulnerability and foraging traits may not be independent. Body size or preferred habitat, for example, affect species in both their roles as resources and as consumers.

The function $c(\cdot, \cdot)$ is often modeled such that for any biologically possible (or technically allowed) combination

TABLE 1
The sets of species lumped into each network node in Figure 1A

Node	Name	Species
1	*Phagocata gracilis*	
2	Decapoda	*Orconectes rusticus rusticus*
		Cambarus tenebrosus
3	Plecoptera	*Isoperla clio*
		Isogenus decisus
4	Megaloptera	*Nigronia fasciata*
		Sialis joppa
5	Pisces	*Semotilus atromaculatus*
		Rhinichthys atratulus
6	*Gammarus minus*	
7	Trichoptera	*Diplectrona modesta*
		Rhyacophila parantra
8	Asellus	*A. brevicaudus*
		A. intermedius
9	Ephemeroptera	*Baetis amplus*
		Baetis herodes
		Baetis phoebus
		Epeorus pleuralis
		Centroptilum rufostrigatum
		Pseudocloeon carolina
		Paraleptophlebia moerens
10	Trichoptera	*Neophylax autumnus*
		Glossoma intermedium
11	Diptera	Tendipedidae
		Simulium sp.
		Tipulidae
		Pericoma sp.
		Dixa sp.
		Other
12	Detritus	
13	Diatoms	(Unresolved)

SOURCE: After G. Wayne Minshall, *Ecology*, 48: 139–149.

of vulnerability traits $\mathbf{v}$ there is a specific combination of foraging traits $\mathbf{f}$ that maximizes the interaction strength $c(\mathbf{v}, \mathbf{f})$. Disregarding those $\mathbf{f}$ that do not match any $\mathbf{v}$, one can then identify the set of biologically possible $\mathbf{v}$ with the set of maximizing $\mathbf{f}$. This unified set of trophic traits is called the *trophic niche space*. Under some mild conditions on the convexity of $c(\cdot, \cdot)$ (i.e., that $c(\mathbf{v}, \mathbf{f})$ decreases the more $\mathbf{f}$ differs from the value matching $\mathbf{v}$), satisfied in most models, the interaction strength between two species i and j is then the stronger the closer $\mathbf{f}_j$ is to $\mathbf{v}_i$ in trophic niche space. By fixing some threshold value c_t, one can assign a foraging niche to each consumer j as the set of all $\mathbf{v}$ for which $c(\mathbf{v}, \mathbf{f}_j) > c_t$. The *foraging niche* of a consumer j surrounds $\mathbf{f}_j$ in trophic niche space. Many food web models are defined by combining a link-strength function with rules for assigning trophic traits to species.

An example for a link-strengths function is

$$c(\mathbf{v}, \mathbf{f}) = c_0 \exp\left[v^{(0)} + f^{(0)} - \frac{1}{2}\sum_{i=1}^{D} \sigma_i \left(v^{(i)} - f^{(i)}\right)^2\right]. \quad (1)$$

In this case, the trophic traits $\mathbf{v}$ and $\mathbf{f}$ are $(D + 1)$ dimensional real-valued vectors with components $v^{(0)}, \ldots, v^{(D)}$ and $f^{(0)}, \ldots, f^{(D)}$, respectively, σ_i equals either $+1$ or -1, and c_0 is a scale constant. Many link-strength functions found in the literature can be brought into this or similar forms. Although mixed sign structures are conceivable as well, theorists generally assume $\sigma_i = +1$ for all i. For a given $\mathbf{v}$ and fixed $f^{(0)}$, trophic link strength is then maximized by matching $f^{(1)}, \ldots, f^{(D)}$ with $v^{(1)}, \ldots, v^{(D)}$. Thus, trophic niche space is here a D-dimensional vector space. The *trophic baseline traits* $v^{(0)}$ and $f^{(0)}$ determine the overall vulnerability of a resource and the overall aggressivity of consumers, respectively.

The idea that trophic link strength depends only on the consumer and the resource is a simplifying idealization (think of prey hiding in bushes). A modification of trophic link strength by a third species has been called a *rheagogy* (based on the Greek *rheô*, "flow," and *agôgeô*, "influence"). Rheagogies can complicate the interpretation of food webs and, when sufficiently strong, qualitatively affect community structure.

Dynamic Food Web Models

Population-dynamical food web models describe how the abundances of all species in a community change over time as a result of trophic interactions. They build on the premise that trophic interactions dominate population dynamics, so that nontrophic effects can be modeled in highly simplified form. Historically, population sizes have been quantified by the number of individuals; recent models generally operate with biomass or bio-energy instead.

Minimal ingredients for population-dynamical food web models are appropriately parametrized sub models for (1) functional and numerical responses of consumers to varying resource abundances, (2) nontrophic losses (death and/or metabolic losses), and (3) the population dynamics of basal species. Box 1 provides an example.

BOX 1. A SIMPLE POPULATION-DYNAMICAL FOOD WEB MODEL

This model describes trophic interactions between 13 (trophic) species. The population dynamics of consumer species (i = 1–11) is modeled by a generalized Lotka–Volterra model of the form

$$\frac{dB_i}{dt} = -RB_i + \overbrace{\epsilon \sum_j C_{ji}B_j B_i}^{\text{numerical response}} - \sum_j \overbrace{C_{ij}B_i}^{\text{functional response}} B_j \quad (2)$$

and that of the basal species (i = 12, 13) as

$$\frac{dB_i}{dt} = \overbrace{r(1 - B_i/K)B_i}^{\substack{\text{basal} \\ \text{dynamics}}} - \sum_j \overbrace{C_{ij}B_i}^{\substack{\text{functional} \\ \text{response}}} B_j. \quad (3)$$

With A_{ij} denoting the adjacency matrix corresponding to Morton Creek (Fig. 1A), trophic link strengths are set to pseudo-random values $C_{ij} = A_{ij} \cos(13i + j)^2$, the efficiency for conversion of resource biomass to consumer biomass to ϵ = 0.5, the effective consumer respiration rate to R = 0.2, the growth (or replenishment) rate of basal species to r = 1, and their carrying capacities to K = 1. Simulations are initiated by setting all B_i = 1 and continued over 1000 unit times. Species with $B_i < 10^{-10}$ are removed as extinct. The remaining community, shown in Figure 1B, approaches a stable, feasible fixed point.

Food web assembly models describe how the structure of food webs changes as a result of a sequence of invasions and extinctions. Invading species are either chosen from a predetermined species pool (representing a metacommunity) or generated at random. Assembly models in which trophic traits of invading species are determined by randomly modifying traits of resident species are called *evolutionary food web models*. This is often described as "speciations" of resident species in large "mutation" steps. The scheme is perhaps more realistic than this interpretation. The actual speciations can have occurred allopatrically or at distant locations or times. In the focal community, such processes would simply be reflected by invasions of species that are more or less similar to existing ones.

Examples of evolutionary models in the literature cover cases with small and large mutation steps, as well as cases where the decision which species to invade or to

extinguish is made depending on population dynamics (populations falling below a given threshold are removed as extinct), depending on other properties of the food web or purely at random (neutral evolution).

PATTERNS AND MECHANISMS

Models and Data

Generic models of complex systems such as food webs are unlikely to reproduce all aspects of empirical data. One therefore has to distinguish between patterns in empirical data, on one hand, and models capable of reproducing and explaining specific sets of patterns, on the other hand. This distinction is blurred when data are characterized in terms of models fitted to them. Yet models fitted to data are useful even on purely theoretical grounds—for example, for identifying the mechanisms by which specific combinations of model elements reproduce particular empirical patterns, or to understand in how far different mechanisms interact under realistic conditions. The discussion here will emphasize mechanisms that relate model elements to patterns, since understanding these should allow us to recombine model elements to balance model complexity against desired descriptive or predictive capabilities.

The sheer amount of data that accurate quantitative descriptions of complete natural food webs would require dashes any hopes that such descriptions will become available in the foreseeable future. Each food web data set is a compromise between accuracy and completeness, and different empiricists will set different priorities. To make food webs manageable, trophic links are often inferred rather than measured, *trophic species* are introduced for special compartments such as "detritus" or to summarize groups of similar species, especially at lower trophic levels, and interactions with neighboring communities are not represented (see, e.g., Table 1, Fig. 1A). Theory interpreting such data has to take these limitations into account.

Size Selectivity

Big fish eat small fish. Patterns of body-size selectivity are evident in food web data, especially for aquatic communities. Models capture this phenomenon by assigning a trophic trait "size" to each species and constraining the relative sizes of resources and consumers in feeding interactions. For models and data that focus on predatory interactions rather than parasites, pathogens, or grazing, the pattern "large eats small" dominates.

Because individuals of many species grow substantially before they mature, the size range of resources fed on by a species as a whole can be large. Pairs of species eating each other's offspring lead to loops in food webs, and

species eating their own offspring lead to *cannibalism*. As a simplified representation of these complications, food web models sometimes allow consumers to choose resources among all smaller species, admitting a few exceptions where small eats large. When no exceptions are allowed, food web adjacency matrices can be brought into lower-triangular form by assigning indices to species according to size. One of the earliest topological food web models, the *cascade model* conceived by Cohen, Briand, and Newman in 1989, generates adjacency matrices simply by randomly setting a fraction of matrix entries below the diagonal to 1 and leaving all others 0. The cascade model was shown to generate topologies more similar to empirical webs than other, comparably simple models.

Phylogenetic Constraints

Evolutionarily related species are similar, and similar species tend to have similar resources and consumers. For example, a bird eating a seed is likely to eat other kinds of seeds; and a bird eating an insect is more likely to eat other kinds of insects than to eat seeds because insects are more similar to each other than to seeds. Empiricists are taking these phylogenetic constraints on food web topology for granted when defining trophic species taxonomically (e.g., Table 1). Phylogenetic analyses of food webs confirm this pattern. They also show that consumer sets of species are inherited more strongly than resource sets. For example, granivorous and insectivorous birds may differ in their diets and yet share bird-eating raptors as common consumers.

These phylogenetic constraints can be described by evolutionary food web models in which both vulnerability and foraging traits are inherited, the former stronger than the latter. Such models generate characteristic patterns in food web topologies, that is, patterns evident even without knowledge of the underlying phylogeny. Before considering other mechanisms to explain patterns in food webs, one should therefore always ask first whether these patterns are simply consequences of phylogenetic and size constraints, two empirically well-established facts. Most patterns discussed below are of this kind.

Link-Strength Distributions

Food webs contain many more weak links than strong links, that is, distributions of link strengths are highly skewed toward weak links. This pattern is found independent of whether link strength is measured in terms of absolute biomass or energy flows, or normalized to predator and/or prey abundances. While early studies found link strengths to be exponentially distributed,

more recent work points toward power-law or log-normal distributions (a quantity is log-normally distributed if its logarithm is normally distributed).

Log-normal and power-law distributions may be empirically indistinguishable. Assuming log-normal link strength distributions, very strong links are unlikely to be observed because corresponding consumer–resource pairs are rare, and very weak links are unlikely to be observed because they are hard to detect. Thus, only a finite range in link-strength magnitude is empirically accessible, and it is well known that the upper tail of a log-normal distribution can, over a range small compared to the width of the distribution, be approximated by a power law.

Log-normal distributions of trophic link strength arise naturally in models where the link-strength function is a product of many factors,

$$c(\mathbf{v}, \mathbf{f}) = \prod_k c_k(\mathbf{v}, \mathbf{f}), \qquad (4)$$

and the factors $c_k(\mathbf{v}, \mathbf{f})$ are all positive and depend differently on consumer and/or resource traits. An example for such a link-strength function is Equation 1 with large D. To the extent that the trophic traits of member species are distributed randomly within a community, $\log c(\mathbf{v}, \mathbf{f})$ is then the sum of many random numbers and can be approximated by a normal distribution. Hence, $c(\mathbf{v}, \mathbf{f})$ is log-normally distributed.

Two empirical patterns—(a) that link strengths are similarly distributed independent of their precise definition, and (b) that resource abundances are bad predictors of diets—are both reproduced by models in which the variance of logarithmic link strengths is large compared to that of logarithmic abundances.

Very weak trophic links are empirically indistinguishable from absent trophic links. Quantitative food web models will therefore in general be more parsimonious if all possible link are modeled as present, but logarithmic link strengths are allowed to vary broadly, such that, as observed, the fraction of links sufficiently strong to be detected (*connectance*) is small. In this view, adjacency matrices A_{ij} derive from link-strength matrices C_{ij}, e.g., by thresholding link strengths. Empirical food web topologies might therefore best be understood by combining a theory for link strengths C_{ij} with an observation model.

Degree Distributions

For directed graphs, one defines the *in-degree* of a node as the number of incoming links, and the *out-degree* as the number of outgoing links. In food web theory, in-degree (number of resources) is also called *generality* of a species, and out-degree (number of consumers) is called its *vulnerability* (not to be confused with vulnerability traits).

Both, in- and out-degrees are distributed much more broadly in food webs than would be expected for links assigned with equal probability to each species pair (which yields binomial distributions). Out-degrees often have an approximately even distribution between zero and some upper limit. This pattern is reproduced by models implementing the "large eats small" rule mentioned above, but otherwise assigning links from consumers to a given resource independently and with equal probability.

Most empirical distributions of in-degrees (and sometimes out-degrees) are approximately exponential (geometric). At least, values near zero tend to be the most frequent and the distributions have long upper tails. This skewed structure can be reproduce by evolutionary models with high heredity of vulnerability traits. The skewed distribution of clade sizes characteristic of phylogenetic trees then leads to the observed skewed distribution of in-degrees, provided consumers tend to forage on a single resource clade. Near-exponential degree distributions are sometimes also found in models without phylogenetic constraints, but the underlying mechanism in this case has not been identified, yet.

For better comparison across food webs, in- and out-degrees can be normalized by the mean out-degree (= mean in-degree = link density). Distributions of normalized degrees are quite similar among food webs and have been hypothesized to follow universal functional forms.

Degree distributions and link-strength distribution are related through the *multivariate distribution* of the strengths of all links from and toward one species. For instance, the hypothesized universality of normalized in-degree distributions implies dependencies among link strengths: a link from one resource to a consumer makes the occurrence of a link of similar strength from another resource to the same consumer more likely (Rossberg, Yanagi, Amemiya, and Itoh, *Journal of Theoretical Biology* 243: 261–272).

Network Motifs

Food webs contain certain types of small connected subgraphs with three or four species more often (others less often) than expected by chance. Using smart randomization algorithms, it can be shown that the prevalence of these *network motifs* is not explained by the degree distributions particular to food webs. Yet it remains unclear how far network motifs are simply consequences of

phylogenetic and size constraints, or whether some motifs call for independent explanations.

Intervality

For small food webs, it is often possible to re-index species such that, in each column of the adjacency matrix, the 1s form a contiguous block (consecutive ones property). For larger food webs, this is usually not possible, but the number of 1s one would need to add in the matrix to archive this is much smaller than expected by chance. This phenomenon is called *food web intervality*, with reference to a related concept in graph theory.

Evolutionary food web models with high heredity of vulnerability traits naturally reproduce this phenomenon (see Fig. 2). Its historical interpretation, however, was different. Influenced by an ongoing discourse on ecological niches, Joel E. Cohen, when discovering intervality in 1977, interpreted it as the signature of an approximately one-dimensional trophic niche space. In fact, food webs defined by thresholding link-strength function Equation 1 with $D = 1$ and fixed $v^{(0)}$ will always be perfectly interval. A subsequent search for the trophic trait corresponding to this single dimension, however, remained unsuccessful. Body size, a likely candidate, does not to constrain topology sufficiently to explain the pattern. By now we know that phylogenetic constraints offer a more parsimonious explanation. Yet models constructed around one-dimensional

Phylogenetic tree   Consumers

Palaeontological time

FIGURE 2 The phylogenetic explanation of intervality. Put a food web's member species in an order in which they could appear on the tips of a phylogenetic tree. If consumers feed on groups of species that are similar because they are related, then the 1s in each column of the adjacency matrix (shown in part) will automatically form contiguous blocks (gray boxes). That is, the food web is interval. In practice, food web intervality in large communities is rarely perfect, neither in data nor in models implementing this mechanism *via* phylogenetically correlated vulnerability traits.

niche spaces are still popular as simple tools to generate interval food webs.

Block Structure

By visual inspection, one easily recognizes large rectangular *blocks* of high connectance in empirical adjacency matrices. More precisely, there are pairs of large species sets (P_r, P_c) such that trophic links between resources from P_r and consumers from P_c are much more frequent than trophic links on average. If P_r and P_c are identical, one speaks of a *compartment* structure. Compartments are rare in food web topologies, but they can result from joining food webs of separated habitats—and be modeled as such.

Blocks with nonoverlapping or partially overlapping P_r and P_c arise in models with phylogenetic constraints. If, for example, all members of P_r are closely related and heredity of vulnerability traits is high, then all members of P_r have similar vulnerability traits. Members of P_c are those species with foraging traits located near the vulnerability traits of the members of P_r in niche space. The members of P_c are not necessarily related.

Covariation Patterns

Early food web theory worked with large collections of food web datasets that were each too small to exhibit much structure on their own. Theory therefore focused on the question how simple quantitative properties of topological food webs, such as the number of species, the fraction of top species (species without consumers), the mean length of resource–consumer chains, and so on, covary among datasets. Stochastic food web models reproducing these covariation patterns were sought.

Later, this idea was adapted to smaller collections of larger, high-quality food webs. The problem then amounts to defining a vector of food web properties $\mathbf{x}$ and finding models such that, for each empirical data set i, its properties $\mathbf{x}_i$ are likely to co-occur in the output of a stochastic model for an appropriate set of model parameters $\mathbf{p}_i$. Model-selection methods such as the Akaike Information Criterion have been used to guard against model overparametrization. Applicability of this analysis is limited by its computational cost. In the most advanced application so far, involving 14 food web properties and models with up to 6 free parameters, a million independent model samples had to be generated to fit each of 17 empirical datasets. The model most faithfully reproducing covariation patterns was the *matching model,* which generates food web topologies by
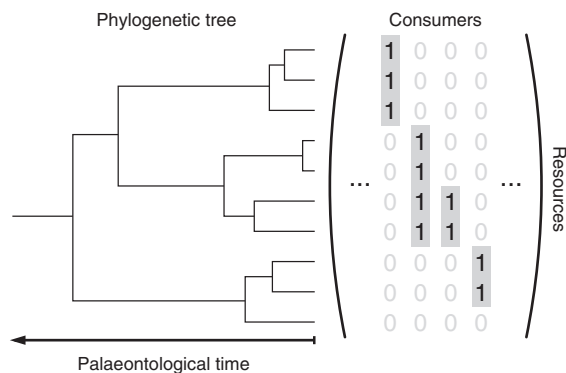
combining a large-eats-small rule with phylogenetically correlated trophic traits.

THE PROBLEM OF FOOD WEB STABILITY

What Is the Problem?

Many practical questions that food web theory strives to address relate directly or indirectly to population dynamics. A plausible approach to answering such questions is to construct a quantitative food web for the community in question, to set up equations modeling the population dynamics of the community, and to simulate this system. However, in the majority of cases, one will find this approach to fail: as simulation time proceeds, many species in the community go extinct, despite their observed coexistence in reality.

An example is shown in Figure 1. The empirical food web of Morton Creek is equipped with population dynamics as described in Box 1. When simulating the system, 7 of the 13 trophic species in this food web go extinct. Variations of the parametrization suggested in Box 1 lead to similarly devastating results.

The mathematical concept most adequate for the study of this phenomenon is community permanence. A community is permanent if no population ever drops below an arbitrary but fixed positive threshold, except for initial transients, and all populations remain bounded from above, independent of initial conditions. However, for the sake of simplicity, theorists often consider instead community feasibility (existence of a fixed point where all populations are positive) or the linear stability of feasible fixed points. The three concepts are related: linearly stable communities are typically permanent, and all permanent communities have a feasible fixed point.

What follows is a critical discussion of possible explanations for these difficulties in building realistic, stable population-dynamical food web models and of suggestions that have been made to overcome these. Efforts have been made to cover the most important ideas, yet the list is not complete.

Self-Limitation of Populations

Trophic interactions are not the only factors limiting population growth. For basal species (species not feeding on others), this is obvious. But even consumer population dynamics will be affected by nontrophic factors, such as the spread of diseases or competition for suitable breeding space. In models, such effects are generally captured by self-limiting, density-dependent contributions to the population growth rate (intraspecific competition). In the Jacobian matrix computed at a community fixed point (controlling linear stability; see Stability Analysis), these self-limiting terms lead to negative contributions to the diagonal, which stabilize the system. The magnitude of these diagonal contributions required to achieve linear stability has been used as a measure of system stability (*diagonal strength*).

One finds that substantial intraspecific competition is required to stabilize large food web models by this mechanism alone. The question whether in nature self-limitation of sufficient magnitude regularly occurs appears to be open. In agriculture, population densities much higher than naturally observed can be reached simply by providing sufficient food and protection from natural enemies. Besides, one needs to ask for any mechanism of nontrophic self-limitation whether the same mechanism could also mediate competitive interactions with other, similar species, in which case this mechanism would equally contribute to destabilizing a community as it helps stabilizing it.

Adaptive Foraging

In laboratory studies, consumers that are let to feed on two kinds of resources often feed disproportionally more on the more abundant resource (*adaptive foraging, prey switching*). The reasons can be active behavioral adaptation by the consumer, or passive mechanisms such as shifts in the consumer's feeding grounds. Evolutionary adaptation can have similar effects, too. Independent of the mechanism, the phenomenon leads to a release of rare species from predation pressure. Species that would otherwise go extinct can survive at low abundances. Adaptive foraging therefore stabilizes communities.

However, foraging adaptation appears difficult to measure in the field, leaving the question open how much of this effect should realistically be allowed in models. Besides, current representations of adaptive foraging in food web models often do not allow for switching to become weaker for pairs of very similar resources, an unlikely feature, which effectively disables competitive exclusion through indirect competition.

Sparse Food Webs

In food web models with self-limitation, stability breaks down when the link density becomes too large, unless other, stabilizing model elements are included. With respect to linear stability, this effect was mathematically explained in 1972 by Robert May. Later simulation studies demonstrated a similar constraint for community persistence. It appears as yet unclear if this phenomenon prevails in food web models without self-limitation. Empirically, the question if there is a natural limit on link density is unsettled. A large number of studies indicate that link density increases with food web size, but these

studies have been criticised for insufficiently taking account of known biases in empirical data.

Slow Consumers

The larger the typical body mass M of a species, the slower its physiological and ecological dynamics are. Ecologically relevant rate constants of dimension 1/Time, such as consumption rates, respiration rates, or maximal Malthusian growth rates, are known to scale approximately as $M^{-1/4}$ (allometric scaling).

Consumers are often larger than their resources. As a result, ecological processes are slower for consumers than for their resources. When defined appropriately, trophic link strength therefore decreases toward higher trophic levels. This pattern has been shown to stabilize community dynamics and to enhance the likelihood of feasible communities. Randomization of link strengths in model food webs exhibiting this pattern, leaving only the overall distribution of link strengths intact, drastically destabilizes these food webs. Further, it has been shown that for consumer–resource body-mass ratios above 10–100, typical for the observed range of values, this stabilization is particularly efficient.

Stable Trophic Modules

When analyzing certain three-species sub-graphs found in food webs in isolation, i.e., ignoring their interactions with the rest of the web, they are found to be stable (or feasible) more often than expected by chance. It is unlikely that this pattern is simply a consequence of the "slow consumers" pattern described above. Thus, stable trophic modules appear to be an independent phenomenon contributing to food-web stability. This is confirmed by studies directly relating the relative frequency of modules to the stability of random food webs.

Weak Links

Relatively weak trophic links can damp destabilizing oscillations in food webs, that is, when removing these links oscillations become stronger. While this effect is important, it is often overinterpreted to the extent that any weak trophic link would be stabilizing. Simulations by McCann, Hastings, and Huxel (*Nature*, 395: 794–798), for instance, show that a minimum link strength is required. The broader question if specific link-strength distributions contribute to stabilization or destabilization of food webs or emerge from stability constraints requires more research.

Assembly and Evolution

Population-dynamical food webs emerging in assembly models with or without an evolutionary component tend to be substantially larger and more complex than model food webs of the same type obtained by first fixing links and abundances according to some (partially) random algorithm and then simulating dynamics until a persistent community remains. Apparently, assembly selects a set of particularly stable food webs among the set of all possible webs. The particular structural features selected for remain unidentified. Food webs tend to *saturate* in assembly processes; that is, a state is reached in which the perturbations of the community resulting from the invasion of one species lead, on average, to the extinction of one species. There are indications that such saturated states have particular dynamical properties combining fast and slow relaxation processes.

It is plausible to conjecture that natural communities are in such saturated states, too. This would explain the persistent difficulties to reproduce food web stability in models, without offering an immediate solution. Because in saturated communities populations coexist in a delicate ecological balance, models of such communities, unless reproducing population dynamics to very high fidelity, would not be persistent or, after adding stabilizing model elements, be more stable than in reality. Recent observations do indeed indicate that natural communities saturate due to limits to coexistence. However, conclusive evidence linking these observations to community saturation in food web assembly models appears to be missing.

SEE ALSO THE FOLLOWING ARTICLES

Allometry and Growth / Assembly Processes / Food Chains and Food Web Modules / Foraging Behavior / Networks, Ecological / Predator–Prey Models / Stability Analysis / Two-Species Competition

FURTHER READING

Berlow, E. L., A.-M. Neutel, J. E. Cohen, P. C. de Ruiter, B. Ebenman, M. Emmerson, J. W. Fox, V. A. A. Jansen, J. I. Jones, G. D. Kokkoris, D. O. Logofet, A. J. McKane, J. M. Montoya, and O. Petchey. 2004. Interaction strengths in food webs: issues and opportunities. *Journal of Animal Ecology* 73: 585–598.

Bersier, L.-F. 2007. A history of the study of ecological networks. In F. Képès, ed. *Biological networks*. Hackensack, NJ: World Scientific.

Cohen, J. E., F. Briand, and C. M. Newman. 1990. *Community food webs: data and theory.* Berlin: Springer.

de Ruiter, P. C., V. Wolters, J. C. Moore, and K. O. Winemiller. 2005. Food web ecology: playing jenga and beyond. *Science* 309: 68–71.

Drossel, B. 2001. Biological evolution and statistical physics. *Advances in Physics* 50: 209–295.

Loeuille, N. 2010. Consequences of adaptive foraging in diverse communities. *Functional Ecology* 24: 18–27.

May, R. M. 1973. *Stability and complexity in model ecosystems.* Princeton: Princeton University Press.

McKane, A. J. 2004. Evolving complex food webs. *European Physical Journal B:* 38: 287–295.

Montoya, J., S. Pimm, and R. Solé. 2006. Ecological networks and their fragility. *Nature* 442: 259–264.

Pimm, S. 2002. *Food webs.* Chicago: University of Chicago Press. Reprinted with a new foreword.

Williams, R. 2008. Effects of network and dynamical model structure on species persistence in large model food webs. *Theoretical Ecology* 1: 141–151.

Woodward, G., B. Ebenman, M. Emmerson, J. Montoya, J. Olesen, A. Valido, and P. Warren. 2005. Body size in ecological networks. *Trends in Ecology & Evolution* 20: 402–409.

FORAGING BEHAVIOR

THOMAS CARACO

State University of New York, Albany

An animal's survival and reproduction demand that it consume energy and nutrients produced by other organisms. Some animals acquire essential resources in a comparatively simple manner; consider an aquatic filter feeder extracting organic matter from flowing water. Other animals capture resources in a manner requiring a complex series of actions, sometimes involving social relationships; consider a group of lions ambushing, capturing, and then competing for a gazelle. The study of foraging behavior spans diverse questions concerning the mechanisms, evolution, and ecological consequences of animals' food consumption. Foraging theory, more specifically, assumes that natural selection can shape behaviors that directly govern an animal's energy acquisition. Foraging theory has helped advance understanding of the remarkable diversity observed among different species' feeding behavior. Furthermore, models of individual or social foraging can be linked with models for population dynamics to predict stability and complexity at the level of ecological communities.

NUMERICAL AND FUNCTIONAL RESPONSES

A consumer population's birth and death rates, as well as migration rates between populations, may depend on temporal and spatial variation in food availability. Change in a consumer population's density, driven by change in food abundance, is termed the consumer's numerical response to resource density.

Consumers often affect the population dynamics of the biotic resources they exploit. The rate at which consumers deplete food abundance depends on both consumer density and the amount of the food resource eaten per unit time by each consumer. The latter quantity is termed the consumer's functional response to resource density. Properties of the consumer's foraging behavior can directly impact the functional response.

OPTIMAL FORAGING THEORY

Optimal foraging theory (OFT) asks how behaviors governing the acquisition and consumption of resources contribute to survival and reproductive success. OFT offers an understanding of prominent behaviors by evaluating their potential adaptive significance. Characterizing foraging as "optimal" follows from the premise that variation in a behavior influencing an individual's survival or reproduction can be subject to optimizing (i.e., stabilizing) selection. Given this premise, foraging theory commonly invokes mathematical optimization as a metaphor for natural selection. OFT has drawn criticism, largely because the models seldom include intrinsic limitations (e.g., genetic constraints) on phenotypes and their evolution. Foraging theorists reply that they seek general principles linking environments to behavior, predictions independent of any particular organism's mechanistic constraints. Furthermore, any OFT model appreciates that constraints on the forager–environment interaction (constraints often defining the problem) limit the choices or options available to the consumer. OFT has produced quantitative hypotheses about behavior—predictions subject to rejection through experimentation or observation—and the theory has advanced understanding of why certain animals forage as they do.

Behavioral ecologists generally define OFT as the study of solitary, independent foragers and refer to models of interacting foragers as social foraging theory (SFT). For clarity, this distinction is adopted here; some questions concerning social foragers do not apply to solitaries, and methods ordinarily used to solve the two types of models differ.

Model Structure

Models in OFT first identify feasible phenotypes. The set may be discrete (e.g., prey types a predator encounters) or continuous (e.g., the length of time an ambush predator remains at one location). Second, the model specifies limitations intrinsic to the organism (e.g., inability to distinguish prey types) or extrinsic (e.g., time available to feed). As indicated above, OFT stresses the latter constraints. Finally, the model's objective function specifies a quantitative relationship between feasible behavioral phenotypes and a "currency of fitness." That is, the model formalizes the hypothesis that lifetime reproductive success (Darwinian fitness) correlates with a measure of foraging performance, the currency. Maximizing the objective function (or minimizing cost) identifies optimal behavior; predictions are deduced from the model's solution. Testing the

predictions asks if the behavior of interest has the functional significance proposed in the model. OFT does not suggest that every trait of a forager is an adaptation.

Diet Breadth

Specialist consumers exploit a narrow range of resources; generalists are less selective. OFT addresses this distinction in a series of models for the prey types (different foods) included in a forager's diet. A basic version, called the contingency model, answers the following question. Given encounter with an item of a recognizable prey type, should the forager consume the item or reject it and search for a more rewarding prey type? The number and identity of the prey types a forager accepts specify its diet breadth.

The contingency model (Fig. 1) assumes that a forager can search for different prey types simultaneously, since prey are intermingled. But the forager discovers only one item per encounter. When an item is accepted, the forager must stop searching and handle the food to extract

FIGURE 1 The contingency model's optimal diet. Prey types are ranked from highest to lowest profitability, the ratio of net energy yield per item to handling time per item. Profitabilities are indicated by red ✳ symbols. Blue ☆ symbols indicate mean long-term rate of energy gain for a diet composed of the k highest ranked prey types ($k = 1, 2, . . . , 10$). If the profitability of prey type ($k + 1$) exceeds the rate of gain for a diet composed of the first k prey types, the optimal (rate-maximizing) diet includes prey type ($k + 1$). If the rate of gain for a diet of k prey types exceeds the profitability of prey type ($k + 1$), the optimal diet cannot include prey type ($k + 1$) or any lower-ranked prey. (A) Optimal diet includes first three prey types. Profitability of fourth ranked prey is less than optimal diet's rate of gain. (B) Same profitabilities as in Part A, but encounter rates will all prey are increased. Rates of energy gain increase, and optimal diet includes only two highest ranked prey types. (C) Profitabilities decline less rapidly with rank when compared to Part A. Optimal diet includes the six highest ranked prey types.

energy. Prey types can differ in density (and hence encounter rate during search), net energy yield per item, and handling time. The model hypothesizes that a forager's fitness should increase with its average long-term rate of gaining energy. Hence, an optimal diet breadth maximizes the rate of energy gain.

To find the optimal diet, the model evaluates each prey type by the ratio of its net energy yield per item to handling time per item. This ratio is termed the type's profitability. The model's solution requires that the most profitable type be included in the diet. Adding the prey type ranked second implies that the forager will encounter food it accepts more often while searching. But taking the second type will decrease the mean energy yield per item accepted and/or increase the mean handling time per item. The forager faces a tradeoff between faster prey encounter while searching and reduced mean profitability per item eaten. The model's solution yields a simple rule: the forager should expand its diet if its long-term rate of gaining energy when specializing on the most profitable type is less than the profitability of the second type. If this is true, the expanded diet increases the rate of energy gain. Proceeding from the highest rank in descending order, the profitability of each prey type is compared to the long-term rate of gain for the diet including all types of higher rank, and no others. The first set of prey types where the rate of energy gain exceeds the profitability of the next lower ranked type is the optimal diet. The decision to accept or reject a prey type does not depend on its density, but does depend on the densities of all types of higher profitability. The model predicts that a given prey type is either always accepted or always rejected, and it predicts more specialized diets when profitability decreases steeply across ranks or when densities of the most profitable prey types are increased. Later versions examine diet breadth when prey types are encountered simultaneously, when discriminating prey types imposes a cost, and when profitability of each prey type varies randomly.
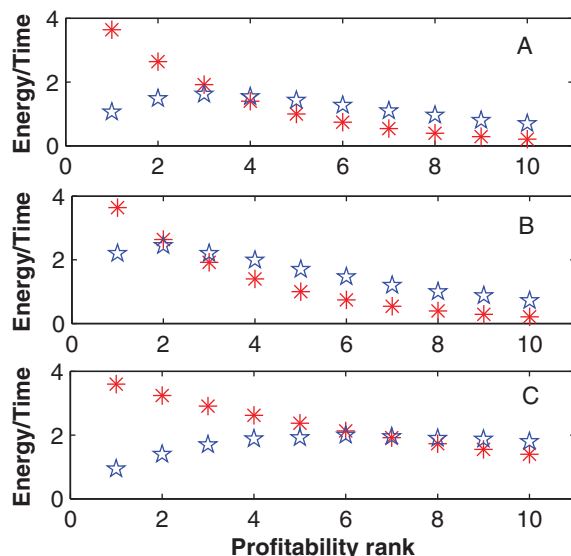
Herbivory and Dietary Constraints

The contingency model's assumptions apply to many carnivores, insectivores, and granivores, since search and handling ordinarily are exclusive activities. Furthermore, food consumption by these foragers simultaneously yields both energy and other required nutrients. However, understanding the diverse dietary ecologies of herbivores, animals that consume only (or mostly) green plant material, demands modified approaches.

A generalist mammalian herbivore sees palatable food everywhere in its environment; its diet breadth may have no relationship to searching effort. Variation in the availability of different plants may contribute to the complexity of herbivore diets, but foraging theory for generalist herbivores emphasizes that plant species can vary in digestibility, nutrient profiles, and toxins. Some herbivores include less digestible material in the diet to slow down the rate at which more digestible material passes through the gut. Only digestible material is absorbed and converted to stored energy; too great a rate of passage may reduce the energy extracted from higher-quality food. Some herbivores balance their intake of energy and essential nutrients (crude protein or sodium) by consuming combinations of different plants. Other herbivores' mixed diets may expose them to different anti-herbivore compounds, while averting too great a consumption of any single plant toxin.

OFT models generalist herbivore diets by subjecting fitness maximization to constraints assuring that nutritional, physiological, or ecological criteria are satisfied simultaneously. Some models identify "strategies" of energy maximization or time minimization. Among feasible diets, one may provide the most energy; another diet may be energetically and nutritionally feasible while minimizing foraging time, and so reducing the herbivore's hazard of predation.

Patch Residence Time

For many animals, foraging consists of repeated cycles of travel between food patches and resource extraction within these patches. One of OFT's most enduring results concerns the length of time a forager should remain in a patch, under the hypothesis that selection favors increases in the long-term rate of energy gain. The solution to the patch-residence problem, called the marginal value theorem, has been applied to a number of seemingly different questions in evolutionary ecology.

The patch-residence model assumes an environment containing one or more patch types. The forager knows the mean travel time between patches and recognizes each patch type upon entry. Patch types differ in resource availability; more productive patches yield a greater net energetic gain for fixed residence time. In any type of patch, the forager's energetic gain decelerates as residence time increases. The rate of energy gain is maximal as the forager begins to exploit a patch, and it declines continuously with residence time, due to resource depression. That is, depletion of food (or evasive action by the forager's prey) lowers the rate at which the forager gains energy. Given a reduced rate of gain as residence time increases, the marginal value theorem asks when an optimal forager should leave and travel to the next patch.

For an environment with only one patch type, the model predicts that increased travel time between patches increases the optimal residence time. In an environment with many patch types, a forager that maximizes its long-term rate of energy gain will leave each patch at the same rate of increase in its energy gain within the patch, and that rate equals the long-term gain rate. Energy gain and residence time may differ among patch types, but the *derivative* of energy gain within the patch (the marginal value), with respect to residence time for that type, is identical across patch types for the optimal forager. This result generated remarkable interest among ecologists, and a number of related models followed. Some relax the assumption of an "omniscient" forager. For example, experience within a patch might help the forager discriminate better from worse patches. Other models compare simpler rules for departure; a forager might leave every patch after a fixed residence time elapses, after capturing a fixed number of prey, or as soon as the time since the forager last found food exceeds a critical "giving up" time.

Risk-Sensitivity

In winter, a forager may have only the daylight hours to consume energy fulfilling its 24-hour metabolic demands. During breeding, an individual might have to capture enough prey each day to meet its needs and those of rapidly developing offspring. For these foragers, failure to consume a required amount of energy during a limited period imperils survival or reproduction. If we further assume that energy intake varies randomly among foraging periods, as must often be true, then models for risk-sensitive behavior apply.

Risk-sensitive behavior implies that an individual's preferences respond not only to average benefits but also to the variance in benefits associated with different actions. To demonstrate the idea, consider the "small bird in winter." A forager has T time units available. Total energy intake by time T must exceed the individual's physiological requirement R, or its chance of surviving the nonforaging period is reduced significantly. For simplicity, let the forager choose between two habitats to search for food. Within a habitat, the animal discovers food clumps as a random process. When the forager discovers a clump, the amount of energy available within the clump varies randomly.

Foraging ends at time T; total energy intake is the sum of the amount consumed within each clump discovered.

Under reasonable assumptions, the distribution of energy intake follows a bell-shaped curve. The expected total intake is simply the product of the mean number of clumps discovered and the mean energy available per clump. The variance of the total energy intake increases with the variance of the number of clumps discovered and with the variance in the energy available per clump. The animal behaves as if it knows the mean and variance of energy intake for each habitat.

A plausible currency of fitness is the probability that energy intake fails to exceed the requirement. A risk-sensitive forager should choose its habitat to minimize the probability that its intake is less than or equal to R. If the intake variance is equal for the two habitats, the forager should choose the habitat with the greater mean. If the habitats offer the same mean intake but different variances, the choice is not so simple. If the expected intake exceeds the requirement R (food is plentiful), the forager should choose the lower-variance habitat. However, when the mean intake does not exceed R (so that survival is jeopardized), the forager should choose the higher-variance habitat; the animal should "gamble" when losing energy. When both mean intake and its variance differ between habitats, they interactively govern probabilities of energetic failure and so combine to predict foraging preference.

State-Variable Models

The preceding examples of models in OFT make static predictions. That is, the expression (or choice) of a behavioral phenotype maps directly to a fitness score. More generally, an action may contribute directly to survival and reproduction, or may contribute indirectly by changing the animal's state. The new state and the advance of time together can affect the animal's next action. Feedback between behavior and state continue until a final time (e.g., end of the day) is reached and fitness is scored. State-variable models predict sequences of actions between initial and final times, as a function of state. The models are termed dynamic, rather than static.

Definition of state depends on the question of interest. In foraging theory, state usually refers to the individual's level of energetic reserves. To demonstrate, recall the diet-choice problem, but in a dynamic context. As the foraging period commences, the animal might accept or reject a prey type based on its initial energy reserve. As its reserve grows or decays, and as the time remaining to forage declines, the animal might expand or contract its diet. That is, the predicted diet breadth can vary with state, even if prey densities and profitabilities remain constant. At the end of the period, a hypothesized "terminal

reward" function maps the final energy reserve to survival and reproduction. Dynamic state-variable models take the expected value of the terminal reward as currency of fitness; optimal behavior, for given reserve and time remaining to forage, maximizes this expectation.

State-variable models ordinarily require computational solution, so that general predictions are not always apparent. Some interesting applications of state-variable models concern costs of suboptimal behavior. Suppose a forager makes a prey-choice "error" in the middle of the day and then follows the optimal policy until the final time. If the error has little effect on the value of the terminal reward, selective pressure is likely weak. However, if the error induces a significant fitness cost, selective pressure on the state-time combination where the suboptimal choice occurred may be strong.

OFT: Final Comment

The methods used in OFT parallel models in bioeconomics—theory developed to manage ecological resources optimally. Predictions of foraging theory have been applied in anthropology, microeconomics, and psychology. Some ethologists, students of behavioral mechanisms, suggest that OFT offers complex models for simple environments and that animals may use simpler rules that deal efficiently with complex environments. As knowledge concerning the neural bases of decision making increases, a combined functional and mechanistic understanding of foraging may emerge.

SOCIAL FORAGING THEORY

Social foraging implies that the functional consequence of an individual's actions depends on both the individual's behavior and the behavior of other foragers, often competitors for the same resource. Social foraging theory (SFT) models generally rely on methods of game theory. Models for dietary choice or patch departure when individuals forage in groups usually make predictions that differ, at least in detail, from the corresponding model in OFT. To emphasize the distinction between solitary and social foraging, this section reviews some issues that concern social foragers only.

Group Size

Groups ordinarily encounter prey (or food patches) more often than do solitaries, and groups can capture larger prey. Group membership may provide the opportunity to learn locations of food or to acquire a foraging skill via social learning. A group member may be safer from predation than is a solitary forager. But as group size

increases, most (or all) group members experience greater competition for food, often leading to aggressive interaction. These benefits and costs, along with mechanisms regulating recruitment/expulsion of group members, govern a foraging group's equilibrium size.

IDEAL FREE DISTRIBUTION

Suppose that a population of identical individuals exploits food occurring only in a small number of patches. Suitability of a patch is given by a constant, representing food density; constancy implies that consumption does not reduce food availability. Increasing the number of consumers occupying a patch decreases the feeding rate of each individual in that patch; consumers interact only through scramble competition.

The ideal free distribution (IFD) predicts consumer density in each patch. "Ideal" implies that an individual knows and chooses that patch where its rate of food consumption is maximal. "Free" implies that a consumer can move between patches without energetic cost or behavioral interference. Foragers move until nothing can be gained by moving elsewhere; the sizes of consumer groups equilibrate when each individual has the same resource-consumption rate. The IFD predicts input matching, where the fraction of consumers in a patch equals the fraction of the total resource available in that patch. That is, the distribution of consumers matches the distribution of resources. The IFD is stable in that an individual switching from one patch to another will reduce its resource consumption, as long as all other consumers do not move.

The IFD has prompted a number of further models. Less than ideal foragers fail to discriminate resource-consumption rates or may learn a patch's suitability only after sampling. Consumers may not be free; travel between patches can be costly. Consumers will not be identical if some forage more efficiently, and interference among individuals will affect the impact of local density on resource-consumption rates. Each of these altered assumptions can predict a different patch-occupation pattern.

AGGREGATION ECONOMIES

In an aggregation economy, benefits of group foraging outweigh costs when groups are small. But as groups become large, competitive interactions eventually increase costs beyond any attainable benefits of foraging socially. Therefore, aggregation-economy models assume that the individual's currency of fitness, as a function of the size of its group, has a single peak. The associated group size is termed, perhaps inappropriately, the optimal group size.

Predicted group size depends on how groups form and dissolve, and can depend on genetic relatedness among group members.

If solitaries can freely enter any group where membership increases their fitness, equilibrium group size will likely exceed the optimal size. However, solitaries may hesitate to enter groups of close relatives, if doing so reduces each relative's fitness. If group members collectively accept or repel a solitary trying to join the group, the equilibrium group in the absence of relatedness will be the optimal size. If, however, the solitary is a relative of current group members, the solitary could be admitted.

Social Parasitism

Different members of the same group may choose alternate methods to obtain food. Consider the producer–scrounger distinction, an example of social parasitism. Producers expend effort finding and capturing prey; a producer gets a meal only when it generates a feeding opportunity. Scroungers avoid costs of producing and attempt to exploit every feeding opportunity provided by the group's producers. If all individuals have chosen to produce, the first individual switching to scrounging will have more chances to feed than any other group member. When the scrounger phenotype is rare, its fitness should exceed a producer's fitness. If all individuals have chosen to scrounge, no food is discovered. As long as a producer can obtain a greater-than-average portion of the food it discovers, the first producer will have a fitness exceeding that of a scrounger. When these conditions hold, the frequency of scrounging will equilibrate where each phenotype has the same fitness. The predicted equilibrium will depend on environmental attributes (e.g., prey density) and the model's fitness currency, but in each case the equilibrium frequency of scrounging will qualify as stable.

Scrounging can appear because individuals seek to increase their own food consumption or to reduce their foraging costs. For a given group size, more frequent scrounging (i.e., reduction in the number of producers) reduces total food consumption across group members. Each individual's pursuit of its own advantage means that every group member obtains less resource, a consequence of social parasitism.

FORAGING BEHAVIOR TO POPULATION DYNAMICS

Models of foraging behavior can be written into the growth equations of consumer–resource systems, integrating individual-level processes with the analysis of

ecological interactions. Some combined models evaluate consequences of particular foraging preferences or functional responses. Other models assume that foragers respond optimally to varying prey density, to predict effects of adaptive behavior on community stability. The body of results is complex; this section lists only a few prominent lessons.

Suppose that an individual forager's effect on the prey population's growth declines as prey density increases. The consequent decelerating functional response does not tend to reduce density fluctuations in a consumer–resource interaction. However, a sigmoid functional response accelerates at intermediate prey densities, so that the prey mortality imposed by each forager increases with the density of prey. Hence, at some prey densities a sigmoid functional response can stabilize population dynamics. When a consumer population preys on two species, a sigmoid functional response can arise if foragers switch between resources and so concentrate predation on the more common prey. Predator switching can, therefore, stabilize the three-species interaction. When a switching predator prevents one prey species from excluding another competitively, the predator's impact is termed a keystone effect.

Dynamical consequences of foraging preference, and its impact on details of the functional response, have been deduced in analyses of three-species food chains. A resource is exploited by a consumer that, in turn, is exploited by a third species. The third species might be an omnivore (exploiting both the resource and the consumer) or a top predator specializing on the consumer; omnivory should exert the greater stabilizing influence on density fluctuations.

Parasitoids often exploit a host population with a highly clumped spatial distribution; many patches contain few hosts, and some patches contain many hosts. An inefficient forager fails to respond to host spatial heterogeneity, while an optimal forager searches patches with the greatest host density. In models of this interaction, optimal patch use by the parasitoid tends to stabilize the densities of the two species. Finally, consider a predator with access to two prey species of differing profitabilities. Suppose that the contingency model's average rate of energy gain enters the dynamics as a component of both prey mortality rates and the predator's birth rate. The predator always includes the prey of higher profitability in its diet. It adds or drops the second prey as the density of the preferred prey changes, according to the optimal diet's choice criterion. The resulting pattern of prey consumption does not tend to stabilize the dynamics,

and it can be destabilizing. In general, adaptive foraging may or may not promote stable ecological interaction; predictions—not surprisingly—depend on model details.

SEE ALSO THE FOLLOWING ARTICLES

Behavioral Ecology / Energy Budgets / Evolutionarily Stable Strategies / Predator–Prey Models

FURTHER READING

Clark, C. W., and M. Mangel. 2000. *Dynamic state variable models in ecology: methods and applications*. Oxford: Oxford University Press.

Giraldeau, L.-A., and T. Caraco. 2000. *Social foraging theory*. Princeton: Princeton University Press.

Houston, A. I., and J. M. McNamara. 1999. *Models of adaptive behaviour: an approach based on state*. Cambridge, UK: Cambridge University Press.

Stephens, D. W., J. S. Brown, and R. C. Ydenberg, eds. 2007. *Foraging: behavior and ecology*. Chicago: University of Chicago Press.

Stephens, D. W., and A. S. Dunlap. 2008. Foraging. In R. Menzel, ed. *Learning theory and behavior*. Oxford: Elsevier.

Stephens, D. W., and J. R. Krebs. 1986. *Foraging theory*. Princeton: Princeton University Press.

FOREST SIMULATORS

MICHAEL C. DIETZE
University of Illinois, Urbana–Champaign

ANDREW M. LATIMER
University of California, Davis

Forest simulators are computer models used to predict the state and dynamics of a forest. As such, forest simulators are on the more complex end of ecological models, both because of the inherent complexity of forest communities and because these models are typically focused on predicting real assemblages of trees, not abstract "forest vegetation." Diverse motivations have driven the development of forest simulators, but the objectives fall into two general classes: (1) to test and extend ecological theory and (2) to predict responses to management action and environmental change. Increasing scientific concern with climate change and the role of forests in global C and N cycles, together with advances in computational power and modeling, are increasing the importance of forest simulators as predictors of forest responses.

OBJECTIVES OF FOREST SIMULATORS

Forest simulators serve to synthesize our reductionist information about how forests work into a coherent,

quantitative framework that can predict mechanistically based on first principles and permit us to verify that inclusion of all the "parts" we study in detail allows us to reconstruct the "whole." In this regard, forest simulation can drive theory by forcing us to codify our assumptions, allowing data–model mismatch to identify false assumptions or understudied processes. A related goal has been to test theoretical predictions about forest dynamics with data from specific systems. Examples include investigating different theories of species coexistence and the roles of disturbance and site history in forest dynamics.

Beyond theory, forest simulators also play an important role in management and policy. A number of applied forest simulators are routinely used to predict growth and yields, such as the U.S. Forest Vegetation Simulator (FVS) and the Canadian Tree and Stand Simulator (TASS). These tend to be far ahead of most ecological models in terms of the diversity of factors they include that impact forest growth, but they also suffer the problem of overparameterization, which leads to high forecast uncertainty. The incredibly high data demands for fully calibrating such models means that they are regularly used with default parameters that may not be appropriate for a given site or situation. In the last few decades, there has also been an explosion of forest simulation research focused on global change issues. The goal here is to make projections that help clarify the potential impacts of global change on forests, such as the change in ecosystem services or the loss of biodiversity, and equally importantly to characterize feedbacks from forests to the climate system via energy, water, and carbon fluxes. These global change applications share the goals of informing policy and management and prioritizing directions for further research. Finally, a more recent application of forest simulators has been in data assimilation, where the goal is to estimate the current state of the forest, rather than some future state, given the constraint of incomplete data. For example, a forest simulator might be used to estimate the structure of a forest that would be compatible with an observed lidar profile and then to make inferences about the likely range of values for other stand properties.

CLASSES OF FOREST SIMULATORS

Forest simulators encompass a wide range of models dealing with different ecological processes and operating across a large range of spatial and temporal scales. While there are exceptions, most forest simulators can be divided into two groups, one that is focused on community ecology and the other on ecosystem ecology.

Within the first group, forest simulation is dominated by a class of models generally referred to as gap models because of their origin in simulating forest gap dynamics, the dominant disturbance for many forest types. Gap models originated in the early 1970s with patch-based models such as JABOWA and FORET that accounted for the height-based competition for light among trees of different sizes and species. These models generally predict dynamics driven by growth rate and shade tolerance, with fast-growing but shade-intolerant early successional species giving way over time to slower-growing but shade-tolerant late successional species. The 1990s saw the development of truly spatially explicit, individual-based forest simulators such as SORTIE. In these models, the crowns of individual trees interact with each other in three dimensions and the understory light environment is more heterogeneous, driven by the overlap of the shadows cast by each individual tree. Similarly, in these spatial individual-level models, seed dispersal becomes an explicit two-dimensional process, with models differing as to whether they treat dispersal from a Lagrangian (individual seed, e.g., SORTIE) vs. Eulerian (seed density, e.g., SLIP) viewpoint. In addition to the spatially explicit IBMs, there are also a number of landscape patch models, such as LANDIS, that take a simpler representation of each individual patch but which represent the broader scale interactions of vegetation with the abiotic environment and which are often focused on broad-scale spatial-pattern and disturbance feedbacks.

In contrast with community-focused gap models are ecosystem-focused forest simulation models. These models are focused primarily on fluxes and pools of carbon but may represent other biologically important cycles as well, most commonly water and nitrogen. Forest ecosystem models tend to be much simpler in terms of their representation of interactions among individuals but more complicated in their representation of physiological processes, such as photosynthesis, carbon allocation, and respiration. These models are also more likely to represent belowground processes such as rooting, soil moisture, and soil biogeochemical cycles. There is a much wider range of spatial and temporal scales represented in forest ecosystem models than in forest community models, from individual trees up to the globe and from near instantaneous in time to millennial. That said, the biological processes involved tend to have particular scales they operate at, and thus models are generally built around specific spatial and temporal scales. Indeed, some of the major remaining challenges in forest modeling—both conceptually and computationally—revolve around scaling.

SPATIAL SCALES

Individual Scale

The spatial scales represented explicitly by forest models range from <1 m to global (Fig. 1). At the finest spatial scales are the spatially explicit individual-based models, such as SLIP and SORTIE, that represent the exact location of individual trees and the spatial interactions between trees. The primary focus of these models is competition for light, which is the limiting resource in most forests and which drives interspecific and intraspecific interactions, tree growth patterns, and demography. Fine-scale processes include the 3D representation of light based on ray-tracing algorithms, which are particularly important for capturing the high degree of heterogeneity in the light environment of forest gaps (Fig. 2). Also occurring at a fine scale is crown competition, 2D seed dispersal, and density-dependent interactions in the youngest life history stages

(e.g., seed bank, seedlings). The focus of these models has thus far been on autogenic fine-scale heterogeneity, rather than fine-scale exogenous heterogeneity in soils or topography, as a mechanism for promoting coexistence.

Patch Scale

The next spatial scale represented by forest models is the "patch" scale, which is on the order of 10–30m in diameter depending on the model, thus encompassing several to dozens of individuals in what is assumed to be a locally homogeneous, common environment. Patch-based models average over the fine-scale variability of spatially explicit models, and the size of patches are set assuming that every individual within a patch is able to compete with every other and that the mortality of the dominant canopy tree is sufficient to convert a patch to a forest gap. Light within a patch-based model is usually represented by a vertical

A Global

Dynamic global vegetation models
Vegetation layers in GCMs

Biomass
low
high

B Regional

Dynamic regional vegetation models
Grid-based models with spatially
implicit subgrid processes

C Landscape

Ecohydrological models
Spatially explicit community
mosaic models
Disturbance and fire models

D Stand and tree levels

Gap models
Patch-based models
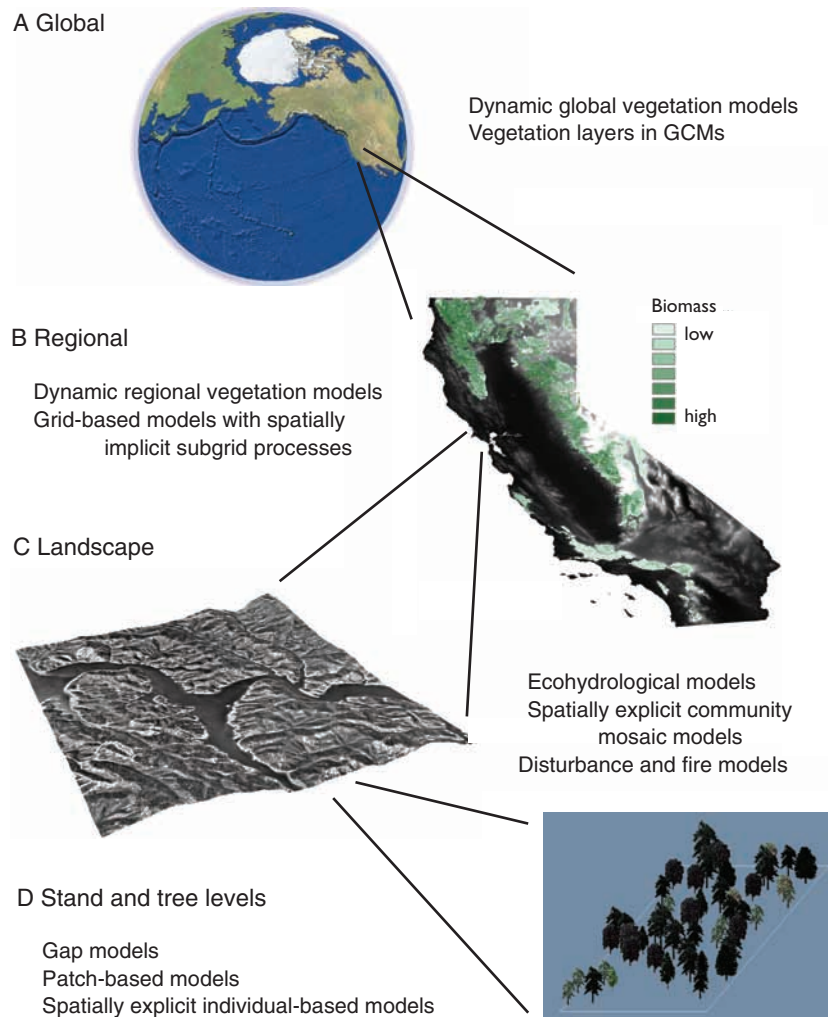Spatially explicit individual-based models

FIGURE 1 Spatial scales addressed by different classes of forest simulators: (A) global vegetation; (B) regional vegetation or forest; (C) landscapes; and (D) forest stands and individual trees.
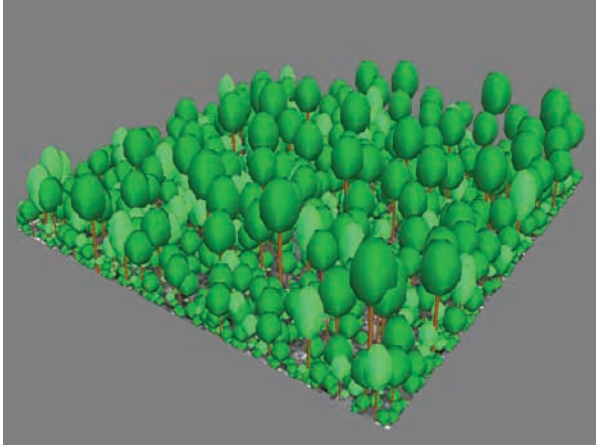
FIGURE 2 Visual representation of forest dynamics in the spatially explicit forest simulator SLIP.

gradient, in which case these models tend to overestimate light levels in gaps, though some models do consider the shadow cast by each patch onto neighboring patches.

The patch is the fundamental scale for a large fraction of models from both the community and ecosystem perspectives, with one key difference: community models are still individual based and thus include multiple trees of multiple sizes and species within a single patch, whereas ecosystem models are based on aggregate carbon pools of a single plant functional type. One ramification of this difference is that while community models will have multiple canopy layers, ecosystem models typically have either a single layer of foliage, often referred to as the "big leaf," or two layers of foliage, representing the functional difference in leaf structural and photosynthetic properties between sun-leaves and shade-leaves. Another key difference in the two modeling approaches at this scale is that because community models are individual based they are focused on the demography of individuals. This means that the fundamental dynamics are conceived in terms of individual demographic responses: the growth rate of a tree based on its size, species, and light environment, the fecundity of individual trees as a function of size and sometimes growth, and the mortality of whole individual trees as a function of growth rate. The inclusion of individual mortality means that almost all forest community models are stochastic, while ecosystem models are almost all deterministic. A consequence of this is that community modelers usually analyze models based on runs with large numbers of patches to average over stochastic dynamics, whereas ecosystem models usually have just one patch in which mortality is simply a deterministic "coarse litter" flux term.

Landscape Scale

The next scale up above patches is the landscape scale (Table 1). The spatial extent of landscapes can vary considerably, from hundreds of meters to tens of kilometers or more. The critical feature of landscape-scale models is not their absolute geographical extent but rather the fact that they account for environmental heterogeneity among patches and aim to provide insight into the effects of such heterogeneity on community or ecosystem dynamics. This heterogeneity can be in terms of the physical template of the landscape itself (e.g., topography, soils, hydrology, microclimate), anthropogenic heterogeneity in the landscape due to land use and fragmentation, or autogenic heterogeneity generated by large-scale disturbances. There is a greater emphasis in landscape modeling on real landscapes rather than on conceptual ones, which are common in gap models that are often focusing on more theoretical questions about the process of succession and community assembly. With this focus on real landscapes also comes a greater emphasis on applied problems and management. A frequent "natural" extent for landscape models is the watershed. Landscape-scale models are most often community focused (e.g., LANDIS, MetaFor), though there are also a number of landscape-scale ecosystem models (e.g., RHESSYS, ForClim), the majority of which are coupled to watershed hydrology models to address ecohydrological questions. Another common feature of landscape models is that there is greater emphasis on spatially contagious processes such as disturbance and dispersal. The most studied of these processes is fire; there are many forest landscape models coupled to fire models that range in complexity from simple "contagious" process models to very detailed mechanistic models of fire spread and intensity (e.g., BEHAVE, FIRE-BGC). While the fundamental unit in landscape models is the patch, the representation of processes within each patch is often simplified compared with patch-scale models. Landscape-scale models are often operating at a spatial scale that encompasses thousands or more patches and necessarily focuses on the distribution of vegetation types and stand ages across patches, rather than the states and dynamics of individual patches. In applying landscape models, users typically assume that they are large enough that the states of constituent patches reach a steady-state distribution (i.e., Watt's patch mosaic) despite the fact that individual patches are far from equilibrium.

Regional to Global Scale

Above the landscape scale are models that take a regional to global perspective on forest dynamics. The questions driving research at this scale primarily surround climate

Model	Spatial Scale	Temporal Scale	Phenom. or Mechanistic	Descriptive or Predictive	Deterministic or Stochastic	Point or Areal
SLIP (Scaleable Landscape Inference and Prediction)	Individual	Annual	Phenom.	Proscr.	Stochastic	Area
SORTIE	Individual	>Annual	Phenom.	Proscr.	Stochastic	Area
TASS (Tree and Stand Simulator)	Individual	Annual	Phenom.	Proscr.	Stochastic	Area
JABOWA (concatenation of authors Janak, Botkin, Wallis)	Patch	Annual	Phenom.	Proscr.	Stochastic	Point
FORET (Forests of Eastern Tennessee)	Patch	Annual	Phenom.	Proscr.	Stochastic	Area
FVS (Forest Vegetation Simulator)	Patch	>Annual	Phenom.	Proscr.	Stochastic	Point
LANDIS (Forest Landscape Disturbance and Succession)	Landscape	>Annual	Phenom.	Proscr.	Stochastic	Area
MetaFor (Forest Meta-model)	Landscape	Annual	Phenom.	Proscr.	Stochastic	Area
RHESSYS (Regional Hydro-Ecologic Simulation System)	Landscape	Daily	Mech.	Proscr.	Determ.	Area
ForClim (Forests in a changing Climate)	Landscape	Monthly	Phenom.	Proscr.	Stochastic	Point
LPJ-GUESS (Lund-Postdam-Jena General Ecosystem Simulator)	Globe	Daily	Mixed	Proscr.	Stochastic	Avg. point
Hybrid	Globe	Daily	Mech.	Proscr.	Stochastic	Avg. point
ED (Ecosystem Demography)	Globe	Subdaily	Mech.	Proscr.	Determ.	Area
CLM (Community Land Model)	Globe	Subdaily	Mech.	Proscr.	Determ.	Wt. point
Sheffield DGVM (Dynamic Global Vegetation Model)	Globe	Daily	Mech.	Proscr.	Determ.	Wt. point
Orchidee (Organizing C & Hydrology in Dynamic Ecosys.)	Globe	Subdaily	Mech.	Proscr.	Determ.	Wt. point
LPJ (Lund-Potsdam-Jena)	Globe	Daily	Mech.	Proscr.	Determ.	Wt. point
Biome-BGC (Biome BioGeochemical Cycles)	Globe	Daily	Mech.	Proscr.	Determ.	Point
CASA (Carnegie-Ames-Stanford-Approach)	Globe	Monthly	Mech.	Descr.	Determ.	Point

NOTE: Many additional excellent models exist in every category. Spatial scale generally refers to the broadest spatial extent the model is designed to run at, though the individual-based models (IBM) typically function at the scale between patch and landscape. Temporal scale refers to the time step of the model. For point vs. area, wt. point refers to models that have multiple points within a grid cell that are weighted by their proportional area while avg. point refers to models that have multiple stochastic replicates within each grid cell that are averaged. See the text section "Process Representation in Forest Simulators" for discussion of other groupings.

change impacts on the carbon cycle and to a lesser extent on biogeographic/biodiversity issues, though these are usually resolved only to the level of biome or plant functional type rather than to species (e.g., Community Land Model, Sheffield DGVM, Orchidee, LPJ). These models all have an ecosystem component, and only a small subset considers community processes (e.g., ED, LPJ-GUESS). However, there is a growing recognition that disturbance history and successional processes can strongly influence the carbon cycle. These models are typically run on a grid where the grid cells are often much larger in extent than the landscapes in the landscape models. When these models include processes at individual through landscape scales, they must represent them as spatially implicit subgrid processes. For example, in the Ecosystem Demography (ED) model forest stands of different ages are not given spatial locations but are represented by the proportion of the landscape that is in each age class.

Since most global models are based on deterministic ecosystem models, the dynamics of these grid cells are essentially identical to that of a single patch or a weighted average of noninteracting patches representing different plant functional types. Models at this scale include the dynamic global vegetation models (DGVMs) that represent the terrestrial ecosystems in general circulation models (GCMs). While these global models are no longer strictly forest models, almost all originated as forest models (e.g., Forest–BGC evolved into Biome–BGC) and were later modified to incorporate other vegetation types. Because of the emphasis on global change within this research community, there have been a much larger number of model intercomparison projects focused on these models than on other classes of forest models. These include early efforts such as VEMAP (Vegetation/Ecosystem Modeling and Analysis Project) and VEMAP2 focused on the continental United States as well as more recent intercomparisons such as the global scale C4MIP (Coupled Carbon Cycle Climate Model Intercomparison Project), the LBA (Large Scale Biosphere Atmosphere) focused on Amazonia, and the two NACP (North American Carbon Program) intercomparison projects, one focused on the continental scale and the other on site-level comparisons to the Ameriflux network.

TEMPORAL SCALES

Forest models resolve processes that range in temporal scale from the near instantaneous to the millennial. Because the

processes involved in community models are essentially demographic, they tend to focus on a narrower range of time scales, from annual to centennial. In contrast, all ecosystem models resolve intra-annual dynamics and some resolve subdaily processes down to a very fine scale. There are two reasons for forest models to resolve processes at subdaily scale. The first reason is to capture the diurnal cycle of photosynthesis using mechanistic photosynthesis models that are driven by instantaneous values of light, temperature, humidity, CO_2, and wind speed. Since these mechanistic models are nonlinear, photosynthesis models operating at a coarser time step either have to make approximations based on an "average" day or use more empirical relationships. The second reason for subdaily modeling is to explicitly resolve the mass and energy budgets of the land surface. These budgets are calculated using a class of process models referred to as land surface models that include a large number of environmental processes beyond the strictly ecological (e.g., boundary layer mixing, snow physics, hydrology, and so on). The primary motivation for including a land surface submodel within a forest ecosystem model is to be able to couple the ecosystem model with an atmospheric model, which requires a lower boundary condition for the land surface. By operating at fine temporal scale and by including atmospheric, vegetation, and hydrological processes, land surface models aim to capture the turbulent mixing and other energy flows that mediate feedbacks among the soil, vegetation, and the atmosphere that are vital to climate projections and to understanding the role of forests in global climate.

At daily to monthly time scales, the processes resolved by forest models are ecophysiological in nature, such as photosynthesis, respiration, carbon allocation, phenology, decomposition, and biogeochemical cycling. Models that have a daily or monthly time scale as their smallest time step typically resolve an explicit mass balance but assume that the energy budget is controlled by some external meteorological driver. At annual to multiannual time steps, forest models typically resolve growth, mortality, reproduction, and disturbance. For most ecosystem models these processes are not resolved explicitly, while for most forest community models this represents the fundamental time step and these processes are the basis for their dynamics. Since most community models ignore intra-annual processes, their calculations for demography are typically based on data-driven empirical relationships rather than physiology. As such, community models are often more constrained to field data, especially with respect to long-term dynamics, but because they generally rely on correlations rather than well-defined mechanisms,

they are less suitable for extrapolating responses to novel changes in the environment drivers or novel combinations of environment variables. In contrast, mechanistic ecosystem models are more robust to extrapolation to different conditions, but they often fail to represent long-term dynamics both because they do not include the successional processes that dominate long-term dynamics and because they are often only calibrated to short-term data.

PROCESS REPRESENTATION IN FOREST SIMULATORS

Beyond space and time, forest models can also be classified by how they represent different processes. Below are presented four important contrasts in model dynamics: phenomenological vs. mechanistic, descriptive vs. predictive, stochastic vs. deterministic, and point-based vs. area-based.

Phenomenological vs. Mechanistic

As alluded to above, the phenomenological/statistical versus mechanistic/physiological dichotomy in many ways reflects the community/ecosystem distinction, but it is more useful to view this as a continuum because at some scale of biological organization all our ecological models are phenomenological and within ecosystem models there is a good bit of variability in how different processes are represented. However, the crux of the distinction lies in whether tree growth is based on correlations with environmental variables or on mechanistic representations of NPP/photosynthesis because this distinction largely determines our degree of belief in extrapolating to novel conditions. In principle, other demographic transitions might also be modeled mechanistically, but in fact mechanistic models for mortality simply do not exist, and those for fecundity are rare and difficult to parameterize. The link between growth and productivity is largely one of mass balance—a given amount of net carbon uptake translates into a given amount of growth, and the only real issue is allocation. Mortality, on the other hand, is a complex and multifaceted phenomenon that is often gradual, with many drivers, feedbacks, and lags. Typically, forest gap models assume that mortality is a function of growth rate and disturbance, while in ecosystem models mortality can be as simple as assuming some constant background rate. Beyond mortality and growth, fecundity can be either phenomenological or mechanistic (usually some fixed fraction of NPP), but in either case it is usually poorly constrained to data.

In theory, dispersal can be either phenomenological or mechanistic, though in practice we are unaware of a

forest model that has been coupled to a mechanistic dispersal model, but this is bound to happen soon due to their increasing popularity. Mechanistic dispersal models are of varying complexity, but all are fundamentally based on wind speed and seed drag or on movement patterns of animal dispersers. Phenomenological dispersal models, on the other hand, are all based on dispersal kernels, which are probability density functions that give the probability a seed will travel a given radial distance from the parent. Either way, data and theory suggest long-distance dispersal (LDD) is a highly stochastic and inherently unpredictable process. One reason for the use of mechanistic dispersal is that LDD is almost impossible to determine from seed trap data. While the role of LDD in community dynamics is well recognized, its importance for ecosystem responses is less well understood—most large-scale models lack explicit dispersal but instead assume one of two extreme cases that define the endpoints in LDD: (i) new seed is available at all places at all times and thus dispersal is not limiting or (ii) all seed rain is local.

Descriptive vs. Predictive

Another important dichotomy is between models that are descriptive versus predictive. Predictive models attempt to predict biotic responses given a set of initial conditions and meteorological drivers and thus can be run into the future conditioned on meteorological scenarios. Descriptive models, on the other hand, typically require other biotic variables to be specified as drivers. Most commonly these are remotely sensed data, such as LAI, fAPAR, albedo, and the like. Because these models are more constrained by data, they are expected to do a better job of diagnosing unobserved biotic variables. For example, atmospheric inversion models such as the CarbonTracker typically base their continental-scale ecosystem carbon fluxes on descriptive models such as CASA. The tradeoff is that such models cannot be run into the future, and thus climate change forecasts are all based on predictive models.

Stochastic vs. Deterministic

A third contrast is between stochastic and deterministic models. As mentioned above, most ecosystem models are deterministic and most community models are stochastic. This difference is due to mortality and the spatial scales of the models. In fine-scale models, the death of an individual tree is an all-or-nothing event and has a large impact on the microenvironment, and thus these deaths are represented stochastically. In broad-scale models, in contrast, mortality is often modeled as a carbon flux term. Since the fine-scale dynamics of individual tree mortality and gap dynamics are thought to play a large role in overall forest structure and composition, the failure to represent these gaps is one of the main limitations of deterministic ecosystem models at long time scales. The approaches to accommodate this scaling problem can be divided into two categories. First, there are ecosystem models, such as LPJ-GUESS and Hybrid, that are coupled with stochastic gap models and scale up by sampling (i.e., running a large number of replicate stochastic patches). Second, there is the Ecosystem Demography model (ED) and models derived from the ED that treat mortality as a deterministic process and accommodate this by explicitly modeling the distribution of stand ages across the landscape. In essence, mortality is thought of as affecting some fraction of each patch in each year, which is reset to a stand age of 0, while the remainder of the patch does not experience mortality. Simulations with a stochastic version of ED show that the deterministic approach accurately captures the mean of the stochastic version and also is more efficient and tractable, as demonstrated in work by Moorcroft and collaborators (2001).

Point-Based vs. Area-Based

The final contrast considered here is between point-based and area-based models and has to do with how models represent space. Most regional/global models are actually point or patch models that are 0D or 1D (vertically structured) and are simply run on a grid (models represent the nodes on the grid). A few contain spatially implicit subgrid processes, where different patches within a grid cell represent different fractional areas, and thus could be considered to be quasi-area-based. Stand level models are area based in terms of a grid of patches (where each patch truly fills the area allocated to it) or are IBM that are spatially explicit and represent area in 2D or 3D. Landscape models fall in between in that they are explicitly based on a map of polygons or grid cells but these grid cells can start to get too big to represent every tree in them or to safely assume all trees within a cell are interacting. Understanding how a model represents space affects how processes scale in the models, what data can be used to calibrate or test the model, and how we interpret model parameters and model dynamics. For example, a leaf property such as maximum photosynthetic rate means very different things if it is referring to an individual leaf on a tree, the whole forest canopy within a patch, or the aggregate carbon uptake across a 1×1 degree lat/lon grid cell.

DISTURBANCE AND STEADY STATES

A number of disturbances have been included in forest models, the most common being gap phase disturbance, from which gap models derive their name, and fire. Gap phase disturbance can either be autogenic, driven by the mortality of a large adult tree, or externally generated by windthrow or ice storms. As mentioned above, fire models vary enormously in their complexity from simple contagious processes to complex simulations. A number of other disturbances are also included sporadically in different models, such as land-use/land-change, droughts, insects, and pathogens, but overall these have received far less attention than fire and gaps.

One of the reasons the representation of disturbance is so critically important to forest models is that they have such a large impact on if and when an ecosystem reaches steady state. Most community-focused models do not assume that the system is at equilibrium at the start of a run since they are interested in the transient dynamics. Community models often start from bare ground or (less often) from some observed or "typical" composition/structure. That said, community models are often run out to some steady state with a lot of emphasis placed on what that steady state is (despite the fact that there's very little information to judge if the steady state is correct). This is in part a reflection of their conception around questions of long-term coexistence. As the spatial scale increases, more and more models use "steady state" as the initial condition for the computer experiments. This is done even when there is widespread recognition that a particular system is not in steady state (and open debate as to whether any ecosystem ever is in steady state). There are two interconnected reasons for this. First, at broad spatial extents datasets do not exist to serve as the initial conditions. There may be partial information from inventories or remote sensing, but many state variables are unconstrained, especially soil properties such as carbon and nitrogen content. The second reason for a steady-state assumption at a broad scale is that models at these scales generally do not explicitly represent successional dynamics and subgrid (landscape, patch) heterogeneity. Current research into ecological data assimilation is in its infancy, one of its goals is to get around the equilibrium assumption at these scales and to acknowledge the impact of this uncertainty on model predictions. Given what we know about the importance and prevalence of disturbance and transient dynamics in forest community and ecosystem dynamics, this is a vital area of research.

CHALLENGES AND CONCLUSIONS

Forest simulators are likely to continue to play a large role in ecological research for the foreseeable future. Many important basic and applied questions about forest models remain unanswered, and important challenges face model developers. This final section highlights issues believed by the authors to be the most important. In a nutshell, the major challenges for forest simulators are that they are very data intensive, hard to initialize correctly, computationally expensive, lack clear analytical solutions, and face a number of scaling issues, particularly when it comes to bridging the community/ecosystem dichotomy. One unifying characteristic of forest models, whether they are ecosystem- or community-oriented, is that because they are generally aimed at predicting real ecosystems they include a lot of processes and require a large number of parameters. Work by Pacala and colleagues in the 1990s on the SORTIE model was a key turning point in the shift from parameterization of models from "the literature" to being much more data driven and connected to experiments designed with model parameterization as an explicit goal. This is still an ongoing change in perspective, though there is a growing recognition of the importance of formalizing data–model fusion and the propagation of uncertainty through models.

Data for Parameterization and Generalization

One important remaining challenge is to better understand to what extent parameters at one site can be applied to another site. In general, gap model parameters are considered site specific, and for larger models the impact of ecotypic variation is largely unknown. Site-to-site variability is not just a "nuance" parameter for models but has large impacts on our conceptual understanding of how forests work and in testing how general our theories of forest dynamics are. Beyond parameterization, forest models are also data intensive when it comes to initialization and drivers. As discussed in the last section, moving away from simple initial conditions that are either "bare ground" or "steady state" to ones that are based on the current state of specific forests requires large amounts of information. Both community and ecosystem models have so many internal state variables that it is virtually impossible to initialize a model precisely for even a single patch, let alone at broader scales, especially once one acknowledges that empirical measurement error is often nontrivial for many ecological processes (especially belowground dynamics). As ecology moves into a "data rich era" thanks to modern observational technologies (e.g., remote sensing, eddy covariance) and research networks (e.g., NEON, FLUXNET, LTER), these challenges will move from the insurmountable toward the routine as ecologists

become more adept at data assimilation and informatics. This is not to say that we won't always be data limited, but that we will be more sophisticated at dealing with the uncertainties. We expect an emerging focus for forest model research will be on determining the quantity, quality, and type of data required to represent and forecast forest dynamics.

Computation

Beyond data, one of the persistent challenges in forest modeling has been computation. While forest simulators have come a long way since the early days of punch cards, the complexity of our models and the scales that we wish to run them on seems likely to continue to outpace Moore's law. In general, forest models are among the most computationally intensive models in ecology. At fine spatial scales, the inclusion of spatially explicit processes can dominate computation (e.g., light and dispersal can be >95% of the computation) and the algorithms involved get disproportionately slower as the spatial scale increases. For broad-scale models, the sheer size of the simulation is usually daunting. For fast time-scale models, such as coupled ecosystem/atmosphere models that include land surface models, the closure of the surface energy budget is computationally expensive and can necessitate complex dynamic numerical integration routines. In all cases, what underlies this computational demand is the fact that forest models lack an analytical solution and thus need to be understood using numerical experiments. The combination of model complexity and lack of a closed-form solution can make forest models difficult to interpret and hampers the ability to reach broad general conclusions. Progress has been made in finding analytical approximations to forest models, and this is an important area of future research and is also closely related to the issues of scaling and crossing the community/ecosystem dichotomy. Partnerships among ecologists, modelers, and mathematicians will be as important as increasing computer power in making these models more useful and interpretable.

Scaling Issues

The frequent dichotomies in the function of forest models (community/ecosystem, annual/diurnal, fine scale/large scale) arise because the processes that affect overall forest dynamics span such a wide range of scales. For computational reasons, it is often impossible to explicitly represent processes important to one class of dynamics (e.g., the emergence of successional dynamics from tree-to-tree competitive interactions) at broader spatial scales. Indeed, individual-based models seem to be limited to a scale of a few km due to the nonlinear scaling of the

computation involved as much as the sheer number of trees that need to be tracked. Given that upscaling individual-based models, or even patch-based models, to regional and global scales will effectively never be computationally possible, an important unresolved question is in what ways do the broad-scale ecosystem models lose representative and predictive power by excluding finer-scale processes. These processes are, especially, (a) neighborhood competition and gap dynamics, (b) the importance and persistence of nonequilibrium dynamics, and (c) the landscape-scale effects of interactions with the abiotic environment. We cannot solve this problem by brute-force computation, so it is essential to understand, by extensive model comparison, analytical insight, and large-scale field campaigns, what is lost in scaling and to devise new scaling approaches. As mentioned above, there are already a small number of models (LPJ-GUESS, Hybrid, and ED) that explicitly attempt to integrate ecosystem and community perspectives and to bring together processes operating across a large range of spatial and temporal scales, but these are just the start and many opportunities for innovation remain.

Species and Functional Types

Another challenge in bridging the community/ecosystem dichotomy is that most community models are parameterized around individual species, whereas most broad-scale ecosystem models are built around plant functional types (PFTs). While the use of PFTs is in part driven by the computational demands of representing diversity, it is more often a reflection of the availability of data to accurately parameterize models. This data limitation is only in part a reflection of what trees have been studied but is also a function of what data are available to modelers. Although there are a number of plant trait database initiatives in progress, these databases need to be made more public and there needs to be a greater incentive for field researchers to archive and document data and to deposit it in such databases. Only with such data can modelers and functional ecologists assess how best to summarize species to the level of functional type and whether important dynamics are lost in doing so. There also needs to be more concerted effort on gap-filling research to constrain the processes that drive model uncertainties, such as belowground dynamics.

Prospects for Forest Simulators

Forests structure the ecological dynamics of many ecosystems, influence regional-scale weather patterns, and dominate carbon fluxes from terrestrial vegetation. They are also economically important globally and locally,

presenting difficult land management challenges such as those arising from logging and fire policy and enforcement. For these reasons, forest simulators will play an increasingly central role both in forecasting global change and in assessing its impacts on existing forests and management practices. Yet current models, despite their increasing sophistication and power, remain highly data dependent and often make predictions without a robust accounting of uncertainty. Because of this, it remains very difficult to do model intercomparisons and to assess model performance confidently. One of most pressing challenges, accordingly, is the availability and integration of data. There is likely to be rapid progress on this front as large new data sources and computational methods become available and widespread. A second set of challenges lies at the intersection of community and ecosystem models: understanding how competitive spatial dynamics and nonequilibrium successional processes influence ecosystem processes and broad diversity patterns, determining how to scale these processes efficiently, and assessing the adequacy of functional types to bridge between species-level dynamics and ecosystem function. Finally, richer information about the belowground components of ecosystem function—including soil microbial ecology and the role of mycorrhizae in flows of energy and nutrients—are fertile areas of investigation, and belowground dynamics are becoming an important frontier of forest modeling. While these are all very active areas of research, there are no clear answers yet, and progress will depend on collaboration among mathematicians, modelers, and field ecologists.

SEE ALSO THE FOLLOWING ARTICLES

Computational Ecology / Dispersal, Plant / Environmental Heterogeneity and Plants / Gap Analysis and Presence/Absence Models / Gas and Energy Fluxes across Landscapes / Integrated Whole Organism Physiology / Landscape Ecology / Plant Competition and Canopy Interactions / Stoichiometry, Ecological

FURTHER READING

Bugmann, H. 2001. A review of forest gap models. *Climatic Change* 51: 259–305.

Gratzer, G., C. D. Canham, U. Dieckmann, A. Fischer, Y. Iwasa, R. Law, M. J. Lexer, H. Sandmann, T. A. Spies, B. E. Splechtna, and J. Szwagrzyk. 2004. Spatio-temporal development of forests—current trends in field methods and models. *Oikos* 107: 3–15.

Larocque, G., J. Bhatti, R. Boutin, and O. Chertov. 2008. Uncertainty analysis in carbon cycle models of forest ecosystems: research needs and development of a theoretical framework to estimate error propagation. *Ecological Modelling* 219: 400–412.

McMahon, S. M., M. C. Dietze, M. H. Hersh, E. V. Moran, and J. S. Clark. 2009. A predictive framework to understand forest responses to global change. *Annals of the New York Academy of Sciences* 1162: 221–236.

Moorcroft, P. R., G. C. Hurtt, and S. W. Pacala. 2001. A method for scaling vegetation dynamics: the Ecosystem Demography model (ED). *Ecological Monographs* 71: 557–586.

Pacala, S. W., C. D. Canham, J. Saponara, J. A. Silander, Jr., R. K. Kobe, and E. Ribbens. 1993. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* 66: 1–43.

Perry, G. L. W., and J. D. A. Millington. 2008. Spatial modelling of succession-disturbance dynamics in forest ecosystems: concepts and examples. *Perspectives in Plant Ecology, Evolution and Systematics* 9: 191–210.

Pretzsch H., R. Grote, B. Reineking, T. Rötzer, and S. Seifert. 2008. Models for forest ecosystem management: a European perspective. *Annals of Botany* 101: 1065–1087.

Scheller, R. M., and D. J. Mladenoff. 2007. An ecological classification of forest landscape simulation models: tools and strategies for understanding broad-scale forested ecosystems. *Landscape Ecology* 22: 491–505.

Shugart, H. H., and T. M. Smith. 1996. A review of forest patch models and their application. *Climatic Change* 34: 131–153.

FREQUENTIST STATISTICS

N. THOMPSON HOBBS
Colorado State University, Fort Collins

Frequentist statistics provide a formal way to evaluate ecological theory using observations. Frequentist inference is based on determining the probability of observing particular values of data given a model that describes how the data arise. This probability provides a basis for discarding models that make predictions inconsistent with observations. The probability of the data conditional on a model also forms the foundation for maximum likelihood estimation, which has been the method of choice for estimating the values of parameters in ecological models.

AIMS AND BACKGROUND

Purpose

Ecological theory seeks general explanations for specific phenomena in populations, communities, and ecosystems. Virtually all scientific theory achieves generality by abstraction, by portraying relationships in nature as mathematical models. Models are abstractions that make predictions. Statistical analysis provides a process for evaluating the predictions of models relative to observations, and in so doing provides a way to test ecological theory. Frequentist statistics, also known as classical statistics, have been the prevailing system for statistical inference in ecology for decades.

Textbooks that introduce frequentist statistics usually emphasize methods—how to estimate a parameter, conduct a test, find confidence limits, estimate power, and so on. Because these texts give only brief treatment of

the statistical principles behind these methods, students introduced to classical statistics often fail to understand what they are doing when they apply the procedures they learned. They have difficulty extending familiar procedures to unfamiliar problems. They may fail to understand fundamental concepts, for example, the meaning of *P* values or confidence intervals. This entry will depart from the customary introductory material by describing the principles that underpin frequentist methods. The purpose of the chapter is to provide a conceptual foundation for understanding classical methods and for appreciating their relationship to other inferential approaches.

Models, Hypotheses, and Parameters

A brief treatment of terminology and notation is needed for this foundation. Assume that a hypothesis expressed verbally (abbreviated as *H*) must be translated into a parameter or parameters (abbreviated as θ) to allow the hypothesis to be evaluated with observations. Thus, a simple hypothesis *H = The mean rate of capture of prey by a predator is 11.5 prey per day* is translated into $\theta = 11.5 \text{ d}^{-1}$. A more detailed hypothesis uses a model to explain variation in the parameter in terms of independent variables. In this way, models are mathematical expressions of verbal hypotheses. So, a more detailed, explanatory hypotheses might be expressed as *H = The average number of prey captured by a single predator per day increases asymptotically as prey density increases because searching and handling are mutually exclusive processes.* This verbal hypothesis is expressed mathematically as

$$\mu = \frac{aV}{1 + ahV}, \tag{1}$$

where μ is the average capture rate of prey per predator per day (time^{-1}), *a* is the search rate (area/time), *h* is the handling time (time), and *V* is the density of prey (area^{-1}). Note that is this case, θ is a vector composed of two parameters, *a* and *h*.

In the discussions that follow, the terms hypothesis and model will be used interchangeably. Moreover, because a model is composed of parameters, θ will also be used to abbreviate models.

HISTORY

The frequentist approach to statistical inference can be traced to the influential work of J. Newman, E. S. Pearson, and R. A. Fisher during the mid-twentieth century. The seminal work of Newman and Pearson developed statistical procedures for making decisions about two alternative actions based on observations. Their work focused on identifying procedures that had the best operational characteristics for separating these alternatives. Two operational characteristics were important. A test should have a low probability of rejecting a hypothesis that is true and a low probability of failing to reject a null hypothesis that is false. The common scientific practice of rejecting a null hypothesis and accepting an alternative has its roots in these ideas. In contrast, R. A. Fisher was more concerned with the use of observations as evidence for one hypothesis over another. In particular, he developed the use of likelihood as a way to quantify the relative support in data for alternative values of parameters or models. The unifying idea between the work of Newman and Pearson on the one hand and Fisher on the other is that probability is defined in terms of the relative frequency of observations and that probability is objectively verifiable and, as a result, can be used to objectively evaluate scientific hypotheses.

FIRST PRINCIPLES

Probability from Frequency

The term frequentist comes from the definition of probability as the relative frequency of observations chosen randomly from a defined population. Imagine that we have an unknown quantity that represents any possible outcome of an experiment or a sample; we call this quantity a random variable, or *Y*. Examples of random variables relevant to ecological theory might include the average height of a tree in a forest patch, the number of offspring produced by a bird, the number of species in a habitat, or the biomass of plants produced on agronomic-style plots. The shorthand **y** will be used to describe a set of specific observations on the random variable, and the shorthand y_i to represent a single observation. So, if *Y* is the possible average height of trees, then $y_i = 3.3$ meters is an example of the observed average height. For occasional cases where we can speak of a single observation or a set in the same context, I will use *y* without a subscript.

The probability that we would observe a particular value, $Y = y_i$, is based on the frequency of that value relative to other values given many repetitions of an experiment or a sample. Simply put, the probability of y_i is the number of times that we observe y_i divided by the total number of observations. The estimated value of the probability of $Y = y_i$ asymptotically approaches the true value as the number of observations approaches infinity. The important message here is that frequentist statistics is based on a specific definition of probability: the relative frequency of observations in an infinite number of repetitions of an experiment or a sample.

Probability Distributions

We treat ecological quantities of interest as random variables because we expect variation when we observe them, variation that arises from many sources—genetic and phenotypic differences among individuals, differences in environmental conditions among sites, errors that arise in our observations, and so on. It follows that if a model predicts no difference in a random variable of interest and if we observe differences in the data, we need to know if the differences we observe should be taken as evidence refuting the model or if these differences would be reasonably expected to arise from natural variation. We use probability distributions to portray variation. Because the idea of a probability distribution and its mathematical description in a density function are central to all statistical analysis, some intuition for probability distributions is developed here.

First, consider the case where data are discrete, which means that integers are the only possible values that we can observe—when we count things, we obtain discrete data. Assume we take all of the y_i in the population and sort them into bins according to their value (Fig. 1). In this case, the bins for sorting the values of y_i are defined by a range of integers (Fig. 1). The relative frequency of values of y_i can be summarized in a histogram, where the heights of the
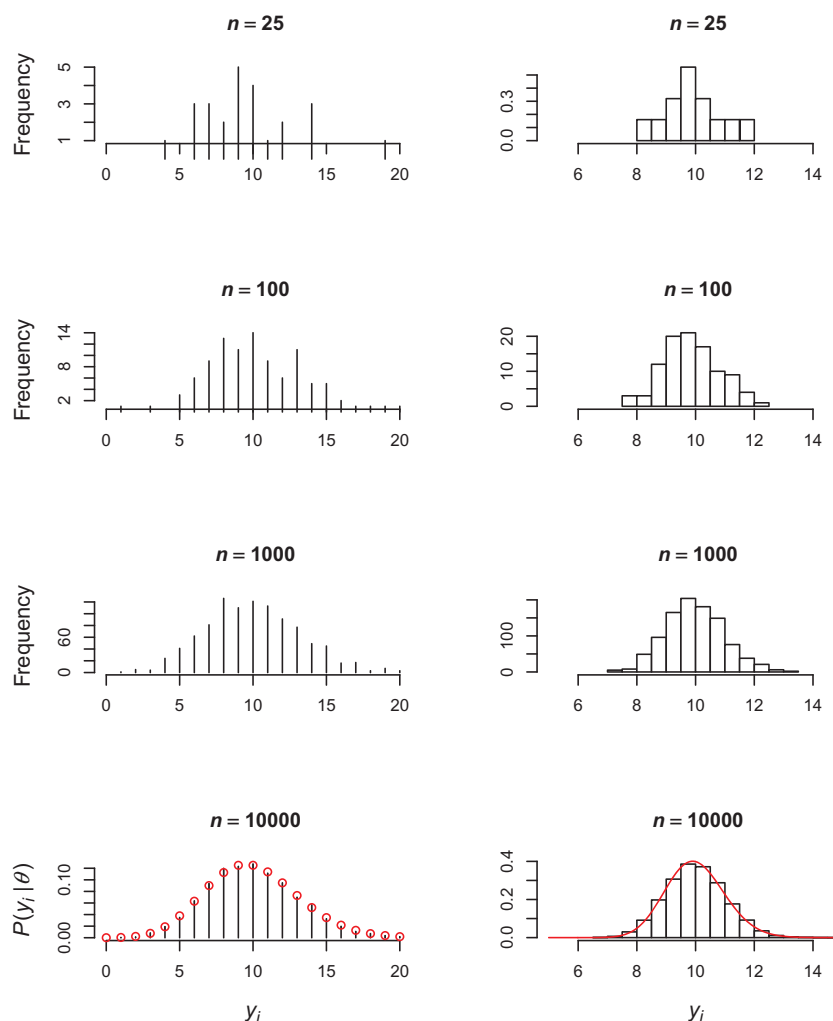

FIGURE 1 Illustration of the relationship between the relative frequency of observations (y_i) drawn from a population with parameters θ and their representation in a probability distribution for discrete (left column) and continuous data (right column). Each panel represents the assignment of n randomly chosen observations from a population to "bins" according to the values observations. The central tendency and shape of the distribution becomes better defined as the number of observations increases, which can be seen in the progression of histograms from the top to the bottom of the figure. As the number of observations approaches infinity, the frequency of each observations divided by the number of observations defines the probability of an observation conditional on the population parameter(s) (bottom row). A discrete density function calculates this probability for discrete data (red circles in lower left panel). A probability density function (red line in lower right panel) calculates the probability density for continuous data. (red solid line in lower right panel).

bars above each integer bin gives the number of observations that we assign to it. If we rescale the height of the bars in the histogram by dividing the number the y_i in each bin by the total number of the observations, then the heights of all of the bars sum to 1 and the height of an individual bar gives the probability the of ith observation, $\Pr(Y = y_i)$. As the number of observations gets large (strictly speaking, approaches infinity), then our rescaled histogram defines the probability distribution of the data (Fig. 1).

When data are continuous, our interpretation is somewhat different. Continuous data (for example mass, length, time, temperature) cannot be accurately represented as integers; they must be expressed as real numbers. With continuous data, our binning example works only as the width of the bins becomes infinitely narrow (Fig. 1), causing an important mathematical distinction between probability distributions for discrete and continuous data. When data are discrete, we can talk about the probability that an individual observation, y_i, takes on a given value because the total area of our rescaled histogram (Fig. 1, lower left panel) can be represented as a sum of the heights of the bars multiplied by their width. This sum must = 1. However, when we make our bins "infinitely narrow," we can no longer sum them to find their area. In this case, we must use definite integration to find the area under curve (Fig. 1, lower right panel), which also equals 1. Because we use definite integration to find probability, we cannot talk about the probability of an observation (that is a single value of y_i), because integration is defined only for intervals, not for points. Thus, for continuous data, we can only talk about probability of ranges of values. As a tangible example, we can estimate the probability that the depth of a stream is between 50 and 70 cm [$\Pr(50 \leq y_i \leq 70)$] using the definite integral of the probability distribution between 50 and 70, but we cannot estimate the probability that a stream is 50 cm deep. Instead, for a single observation of continuous data we must talk about probability density, rather than probability, a distinction that will be clarified in the subsequent section.

Density Functions

With an infinite number of observations of our random variable, the relative frequencies of the values of observations define a probability distribution. This distribution can be portrayed mathematically using an equation called a discrete density function for discrete data (also called a probability mass function and a probability function) and a probability density function for continuous data (Fig. 1). These functions can be represented in a general way as

$$\Pr(Y = y_i \mid \theta) = f(y_i, \theta), \qquad (2)$$

where the left-hand side of Equation 2 is the probability that we would observe y_i conditional on the value of the parameter θ. Avoiding statistical formalism, "conditional on" means that we fix the value of θ and we seek a value for the probability of a variable y_i, given the specific, fixed value of θ. The right-hand side of Equation 2 is the discrete density function or the probability density function. Equation 2 returns a probability for discrete data and a probability density for continuous data. Probabilities, of course, must be between 0 and 1, but a probability density can take on any value such that the integral of the density function over all values of $y_i = 1$. The probability (or probability density) returned by the function is determined by the value of the observation (y_i) and the parameter(s) (θ).

To illustrate these concepts, consider the Poisson discrete density function, which is often used to describe the probability that we would count a number of objects or events per unit of time or space, given that the average number of counts $= \theta$:

$$\Pr(y_i \mid \theta) = \frac{\theta^{y_i} e^{-\theta}}{y_i!}. \qquad (3)$$

Illustrating this function using the simple capture rate example, above, we hypothesize that the mean capture rate $\theta = 11.5$. We then ask, what is the probability that a predator would capture 15 prey items in a day if the mean capture rate $= 11.5 \text{ d}^{-1}$? Thus,

$$\Pr(y_i = 15 \mid \theta = 11.5) = \frac{11.5^{15} e^{-11.5}}{15!} = .063. \quad (4)$$

We conclude that if the mean of the distribution is 11.5 d^{-1}, then we expect to observe 15 prey captured in a day only 6% of the time given many observations.

There are many different density functions that are used to represent the probability distributions of different kinds of observations—the choice of which one to use in a frequentist analysis depends on how the data arise. The most familiar one is the normal distribution, but there are many others. Particularly useful to ecologists are the binomial, multinomial, negative binomial, Student's t, F, chi-squared, uniform, and gamma distributions.

Parameters and Moments

All discrete density functions and probability density functions have one or more parameters (θ). These parameters determine the specific relationship between the inputs and the outputs of the function. All probability distributions also have moments that describe the central tendency and shape of the distribution. For some density functions, notably the normal and the the Poisson, the parameters are

identical with the first two moments (the mean and the variance). However, for all other distributions there is an algebraic relationship between shape parameters and moments, but they are not identical. Translating moments into shape parameters and shape parameters into moments can be accomplished by a method called moment matching that exploits the algebraic relationship between them to solve two equations in two unknowns.

FREQUENTIST INFERENCE BASED ON THE PROBABILITY OF THE DATA

Hypothesis Testing

Statistical hypothesis testing seeks to evaluate scientific hypotheses, which are most often stated as models. Frequentist inference depends on answering the following question: if our hypothesis is true, how probable are the data that we observe? If the observed data have a very low probability under the hypotheses (or model), then the hypothesis can be discarded as false. Notice that this approach to inference sees the hypothesis as fixed while the data are variable.

We gain insight about the hypothesis as follows. We choose a value of θ to represent the hypothesis and choose an appropriate discrete density function or probability density function specifying the probability that a random variable Y takes on a specific value, y_i, conditional on θ. We take an observation (or observations) and use the chosen density function to determine the probability that the random variable would take on values greater than or equal to the observed value. If the data are improbable if the hypothesis were true, we conclude it must be false.

As a simple example, consider a study of survival in a population of birds. We wish to test the hypothesis that the probability that an adult survives from one year to the next is .45, and so we define $\theta = .45$. We will test this hypothesis by observing 100 birds and determining the number that survive a year later. An appropriate probability distribution for this hypothesis is the binomial, which gives the probability of observing a specific number of successes (k = observed number of birds that survive) on a given number of trials (n = number of birds observed) and an overall mean probability of a success (θ = hypothesized survival probability). Thus,

$$\Pr\left(Y = y_i \mid \theta\right) = \binom{n}{k} \theta^k (1 - \theta)^{n-k}. \tag{5}$$

We observe that 57 birds survive. Thus, we ask, what is the probability that we would observe 57 birds alive if the true value of the annual survival probability $= .45$? The probability of obtaining 57 survivals from a sample of 100 assuming that the true survival probability $= .45$ is

$$\Pr\left(Y = 57 \mid \theta = .45\right) = \binom{100}{57} \theta^{57}(1 - \theta)^{100-57} \tag{6}$$
$$= .004,$$

and the probability of observing 57 *or more* survivors is

$$\sum_{n=57}^{100} \binom{100}{n} \theta^n (1 - \theta)^{100-n} = .01. \tag{7}$$

If the true value of the survival probability is .45, and we were to repeat our sample many, many times, we would expect that only 1% of those samples would include 57
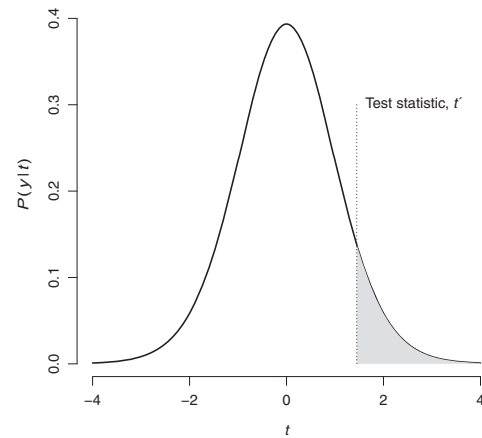

FIGURE 2 Illustration of a test statistic calculated as a function of the data. In this example, we have two data sets, **y** = [11, 16, 18, 24, 16, 8, 32, 8, 6, 13] and **z** = [18, 14, 21, 26, 19, 25, 37, 10, 11, 23]. We want to know if the mean of the distribution from which **z** was drawn exceeds the mean of the distribution of **y**, or, alternatively, if the variation in the distributions is sufficiently great that the observed difference in means could be expected to arise from chance alone. We compose a null hypothesis: "The mean of the population from which **z** was drawn is not greater than the mean of the population from which **y** was drawn." To test this hypothesis, we choose a function of the data, (i.e., a test statistic) that has a known probability distribution under the null hypothesis. Because we have equal sample sizes and the variance of the two samples is the same, we use $t = \frac{\bar{z} - \bar{y}}{\sqrt{(s_y^2 + s_z^2)\frac{1}{n}}}$ where s^2 symbolizes the variance of each sample and n is the sample size. Student's t distribution gives the probability density for given values of t conditional on the degrees of freedom = $2n - 2$. Using the data at hand, the value of t is 1.45. The probability of obtaining a value $\geq$1.45 if the null hypothesis is true is $P = .06$ as indicated by the shaded area in the figure. If we used a fixed significance level of $\alpha = .05$, we would fail to reject the null hypothesis. In this case there would be two possible interpretations: (1) there is no difference between the means or (2) our sample size was not large enough to detect the difference. Alternatively, we could use the value of P as "evidence" against the null hypothesis and could conclude that there is reasonably strong evidence against it—we would expect a value of $t \geq$ 1.45 only 6% of the time given many repetitions of the sample if the mean of the distribution of **z** were not greater than the mean of the distribution of **y**.

or more survivors. This result forms evidence that a survival probability of .45 is improbable, allowing us to reject the hypothesized value.

The example above uses the observations directly to test a hypothesis, but more often, we calculate a test statistic as a function of the observations. Probability tells us that any quantity that is a function of a random variable is also a random variable. Exploiting this fact, we let $t = g(\mathbf{y})$ be a function of the observations and let $T = g(Y)$ be the corresponding random variable. We call T a test statistic if we know the probability distribution of t when the hypothesis is true (H) and if larger values of t provide stronger evidence against the hypothesis. We can use a test statistic to evaluate the hypothesis by taking observations on the random variable and calculating a significance level, P, as

$$P = \Pr[T \geq g(\mathbf{y}) \mid H]. \tag{8}$$

Assuming many repeated experiments or samples, the significance level gives the probability that we would observe a test statistic more extreme than the one we observed given that the hypothesis is true. If that probability is low, then we can conclude that the hypothesis is false. An example of using a test statistic is given in Figure 2.

Null Hypothesis Testing

The null hypotheses forms a central concept in frequentist statistics, a concept that extends from the argument of logicians that propositions can only be determined to be false; they cannot be proven to be true. "True" statements are those that withstand repeated attempts to show they are false. Thus, if a researcher seeks to determine if there is a difference between two means, he or she seeks to falsify the null hypothesis of no difference. In the Newman–Pearson school, the P value provides a basis for choosing between two actions, which has been translated into the idea of rejecting the null hypothesis and accepting the alternative. Acceptance or rejection of the null hypothesis is accomplished using a critical value, which is the value of T for a set significance level $= \alpha$. If the observed value of T exceeds the critical value, then we reject the hypothesis; if it fails to exceed it, then we fail to reject. The specific value of P matters only in the context of choosing whether to reject the hypothesis. Newman and Pearson viewed α as an operational characteristic of the statistical test, specifically, the probability of making a type I error, rejecting a hypothesis that is true. This level was fixed, and any value of a test statistic associated with $P > \alpha$ provided a bias for falsifying the hypothesis. In contrast, R. A. Fisher's writing about P also included the idea that

significance levels provide evidence against the hypothesis and that the smaller the value, the stronger the evidence. Moreover, Fisher also developed the idea of likelihood, which diverged markedly from the approach of Newman and Pearson. (Likelihood is discussed below.)

Confidence Intervals

Thus far, we have discussed the role of observations in testing hypotheses. Observations are also useful for estimating parameters—for example, means, variances, proportions, and rates. Because of the inherent variability among individual observations, we need a way to express uncertainty in parameter estimates, which is accomplished with confidence intervals. The frequentist view of a confidence interval starts with the idea that there is a fixed value of a parameter of interest (θ). Presume we take many samples or run many experiments to estimate θ. We estimate a $1 - \alpha$ confidence interval as a range of values of the random variable of interest $[l(Y), u(Y)]$ such that the range would fail to bracket the true, fixed value of the parameter $(1 - \alpha) \times 100\%$ of the time. More formally,

$$\Pr\left[l(Y) < \theta < u(Y)\right] = 1 - \alpha. \tag{9}$$

FREQUENTIST INFERENCE BASED ON THE LIKELIHOOD OF THE PARAMETER

Frequentist inference discussed thus far relies on estimating the probability of the data conditional on the parameter. In this framework, the parameter is considered fixed and the data are variable (Fig. 3). Likelihood reverses this relationship by considering the parameter to be variable and the data fixed. In the likelihood framework, we ask what is the likelihood of the parameter given that we have an observation with a particular value?

Definition of Likelihood

Likelihood is defined by the relationship,

$$L(\theta \mid y_i) = c \Pr(y_i \mid \theta), \tag{10}$$

which simply says that the likelihood of the parameter conditional on the data is proportional to the probability of the data conditional on the parameter. When the data are continuous and the right-hand side of Equation 10 is a probability density function, then likelihood is proportional to probability density. As a result of the relationship in Equation 10, $\Pr(y_i \mid \theta)$ is often referred to as a likelihood function.

Remembering that we express hypotheses as different values for parameters, the difference between probability-and

Probability of the data
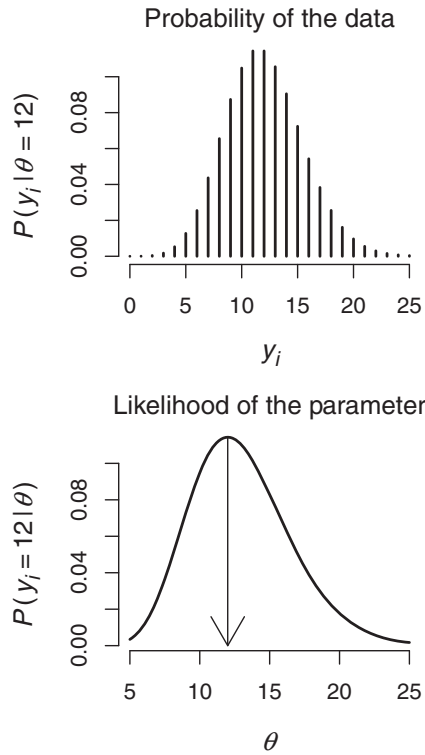

Likelihood of the parameter

FIGURE 3 Illustration of the difference between a probability distribution (top graph) and a likelihood profile (bottom graph) using a Poisson discrete density function, $\Pr(y_i \mid \theta) = \frac{\theta^{y_i} e^{-\theta}}{y_i!}$. In the probability distribution, we fix the mean ($\theta = 12$) and we vary the value of the data to determine how probable the data are a given value of the mean of the distribution. In the likelihood profile, we fix the value of the data ($y_i = 12$) and we vary the value of the parameter to determine the likelihood of the parameter for a given observed value of the data. The arrow shows the maximum likelihood estimate of the parameter.

likelihood-based inference is as follows. In the probability framework, we assume a value for the parameter and evaluate the probability that the data would arise if the fixed parameter value is correct. In the likelihood framework, we assume we have the data in hand, i.e., they are fixed, and we want to evaluate alternative hypotheses represented as different values of the parameter. Because the purpose of likelihood is to compare the evidence in data for alternative values of parameters (read alternative models, hypotheses; see the section "Strength of Evidence," below), the value of the constant c does not matter; for the purpose of model comparison, we can assume $c = 1$. This assumption is fundamental to likelihood analysis. It means that a single value for likelihood cannot be interpreted; rather, interpretation depends on comparing likelihoods for different hypotheses. Likelihood cannot be used to evaluate a single hypothesis, but rather can only be used to evaluate support in data for one hypothesis relative to another, a core property of likelihood theory called "relativity of evidence." This property is explored further below.

Equation 10 describes the likelihood for a single observation (y_i); how do we estimate the likelihood of sets of observations? The likelihood of n independent observations conditional on the value of the parameter(s) is simply the product of the individual likelihoods,

$$L(\theta \mid \mathbf{y}) = c \prod_{i=1}^{n} \Pr(y_i \mid \theta). \qquad (11)$$

The log likelihood of the parameter given multiple observations is obtained using the sum of the individual log likelihoods,

$$\ln L(\theta \mid \mathbf{y}) = ln(c) + \sum_{i=1}^{n} \ln \Pr(y_i \mid \theta). \qquad (12)$$

Log likelihoods are often used to estimate parameters because of computational advantages and because of their relationship to other statistical quantities.

Probability Distributions Compared to Likelihood Profiles

The proportionality in Equation 10 might suggest that frequentist inferences based on probability and inferences based on likelihood are similar, but there are fundamental differences between the two approaches. The basis for these differences is revealed by comparing a probability distribution and a likelihood profile (Fig. 3). A probability distribution has values of the data (i.e., $Y = y_i$) on the x-axis, while a likelihood profile has values of the parameter (i.e., θ) on the x-axis. For both approaches, the value on the y-axis is calculated using the function, $f(y_i, \theta)$ (Eq. 2). However, for probability distributions the parameter is held constant in this function and the value of data varies, while in the case of likelihood profiles the data are held constant and we vary the value of the parameter. It is also important to note that the probability distribution can be discrete or continuous, but the likelihood profile is always continuous. A critical difference between the probability distributions and likelihood profiles is that the area under the likelihood profile $\neq 1$, while the area under the probability distribution $= 1$.

Maximum Likelihood

The likelihood profile helps us to understand the concept of maximum likelihood (Fig. 3). The maximum likelihood estimate of the parameter in a model is defined as the value of θ that maximizes $L(\theta \mid \mathbf{y})$, which can be seen graphically as the value of θ at the peak of the likelihood profile (Fig. 3). For simple models, maximum likelihoods can be found using calculus, but more complex models require numerical methods. Maximum likelihood

is the prevailing approach to estimation of parameters in the frequentist framework.

Strength of Evidence

To evaluate one model relative to another, we use the ratio of the likelihoods for each model evaluated at the maximum likelihood values of the parameters $(\hat{\theta}_1, \hat{\theta}_2)$:

$$R = \frac{L(\hat{\theta}_1 \mid y)}{L(\hat{\theta}_2 \mid y)}. \tag{13}$$

The quantity R measures the strength of evidence for one model over another. So, presuming that the numerator contains the greater of the two likelihoods, we can say that evidence for model 1 is R times stronger than the evidence for model 2. Although this statement provides a perfectly clear summary of the relative strength of evidence for the two models, it is also possible to calculate a P value (Eq. 8) for the difference between the models with a likelihood ratio test. The likelihood ratio test uses $-2 \ln R$ as a test statistic with the chi-square distribution as a basis for the level of significance. The log of the maximum likelihoods is also used to compare evidence for models using information theoretics, for example, Akaike's information criterion (AIC).

RELATIONSHIPS BETWEEN FREQUENTIST AND BAYESIAN INFERENCE

This section briefly discusses key similarities and differences between the two inferential approaches used to evaluate ecological theory with data, frequentist and Bayesian statistics.

Bayesian inference estimates the posterior distribution of model parameters, i.e., $\Pr(\theta \mid \mathbf{y})$. Bayes' law provides the foundation for this estimation:

$$\Pr(\theta \mid \mathbf{y}) = \frac{\Pr(\mathbf{y} \mid \theta)\Pr(\theta)}{\Pr(\mathbf{y})}. \tag{14}$$

The left-hand side of Equation 14 gives the probability of the parameter conditional on the data, which resembles the likelihood profile (Fig. 3), except that the area under the curve now equals 1. The probability of the parameter (i.e., $\Pr(\theta)$) summarizes the information about the parameter that we knew before conducting an experiment or taking a sample—hence, it is called the prior distribution or prior. The probability of the data $[\Pr(\mathbf{y})]$ serves as a normalizing constant, which assures the left-hand-side is a true probability. Because the probability of the data is a constant for any given data set, we can simplify Equation 14 as

$$\Pr(\theta \mid \mathbf{y}) \propto \Pr(\mathbf{y} \mid \theta)\Pr(\theta). \tag{15}$$

There are obvious similarities between Equation 15, which defines Bayesian inference, and Equation 10, which defines likelihood. Although it is often said that the difference between frequentist and Bayesian inference is the use of prior knowledge, this statement is not accurate, because it is perfectly feasible to calculate maximum likelihood estimates that include $\Pr(\theta)$. Moreover, Bayesians and frequentists agree that all of the information in the data about the parameter is transmitted through the likelihood function, $\Pr(\mathbf{y} \mid \theta)$.

So, what is the difference between the two approaches? A major divergence is the frequentist and Bayesian view of the parameters and the data. In the frequentist framework, the data are seen as random variables and the parameter is a fixed but unknown constant. In the Bayesian framework, the parameter is viewed as a random variable, which allows calculation of its probability distribution. The practical effect of including the probability of the data in the denominator of Equation 14 is to normalize the likelihood profile, such that the area under the curve integrates to 1. So, computationally, the primary difference between likelihood and Bayesian methods is that likelihood uses maximization to find the parameter that yields the highest probability of the data, while Bayesian methods integrate the likelihood function (or sum for discrete data) over all values of the parameter to obtain a probability distribution of the parameter. Estimates of means of parameters will be identical using the two approaches for simple analyses when priors are uninformative or when identical informative priors are included in maximum likelihood and Bayesian analysis.

A second divergence is the way the two frameworks define probability. As described above, frequentists hold that probability can only be defined in terms of the frequency of an observation relative to a population of potential observations. In contrast, Bayesians view probability as a measure of the state of knowledge or degree of belief about the value of a parameter, a definition that is consistent with the view of parameters as random variables. Historically, this philosophical difference was a cause for heated argument between frequentists and Bayesians. However, it is important to remember that both inferential approaches are abstractions. The idea that the probability of a parameter represents a state of knowledge is no more or less abstract than the idea of probability based on an infinite number of experiments or trials that we never observe.

The contemporary view of the two approaches focuses less on philosophy and more on pragmatism. Modern computational algorithms, particularly Markov chain

Monte Carlo methods, and fast computers have made it feasible to use Bayesian methods to evaluate complex models that heretofore defied analysis using maximum likelihood. If prior, objective information exists, then it can be included in the Bayesian analysis; if it does not, then priors can be made uninformative (with the caveat that there are frequentists who maintain uninformative priors do not exist). When priors are uninformative, Bayesian analysis resembles "normalized likelihood." Many contemporary ecologists use both approaches for evaluating models with data, their choice being guided by the practical requirements of the problem rather than a dogmatic commitment to an inferential philosophy. However, the freedom to match problems to analyses requires a familiarity with fundamental statistical principles. This entry has introduced those principals for frequentist inference.

SEE ALSO THE FOLLOWING ARTICLES

Bayesian Statistics / Information Criteria in Ecology / Markov Chains / Statistics in Ecology

FURTHER READING

Bolker, B. 2008. *Ecological models and data.* In *R.* Princeton: Princeton University Press.
Edwards, E. W. F. 1992. *Likelihood.* Baltimore, MD: Johns Hopkins University Press.
Royall, R. 1997. *Statistical evidence: a likelihood paradigm.* Boca Raton, FL: Chapman and Hall/CRC.

FUNCTIONAL TRAITS OF SPECIES AND INDIVIDUALS

DUNCAN J. IRSCHICK AND CHI-YUN KUO

University of Massachusetts, Amherst

The notion of animal function centers on the idea of biological roles for certain physical traits. Researchers study animal function in several ways, including mechanistic studies of how muscles, soft tissues, and nerves integrate to drive movement and how structural elements (bones, tendons, muscles, etc.) interact to bolster or inhibit animal motion. For example, one basic function of the vertebrate hind limb is movement to facilitate locomotion, such as during walking or running. An essential element for the notion of function is the ability to perform dynamic tasks,

such as biting, running, walking, flying, and so on. Such ecological tasks, otherwise known as whole-organism performance capacities, are often essential to the survival of animals, as it enables them to occupy environments, escape predators, and capture prey. Examples of whole-organism performance traits include maximum speed or maximum acceleration incurred during short sprints of locomotion, endurance capacity, or bite force (the maximum amount of force that an animal can produce during biting). In other words, function is a broad term used to describe what the phenotypic trait (e.g., hand, limb) is generally selected and well suited for, whereas performance describes how well an organism can accomplish a certain functional task that is ecologically relevant.

HISTORICAL BACKGROUND

The importance of animal function within the broader field of ecology has emerged slowly over the last 40 or so years. Attempts to integrate animal function into ecology first began when ecologists were attempting to understand why and how species could occupy different niches within ecological communities. Early attempts to address this phenomenon were largely devoid of animal function and relied nearly exclusively on measurements and comparisons of morphological traits (e.g., body size, limb elements). However, during the 1970s and 1980s, the field of ecomorphology arose in an attempt to link the functional traits of animals (originally only morphology, but later changing to measurements of performance) with variation in habitat use. The basic premise of ecomorphology is that variation in morphology and function should be tightly correlated (Fig. 1) with variation in habitat use among species that occupy distinct niches. In some cases, researchers have shown that interspecific morphological and functional differences could enable species to access different resources, supporting the basic tenants of niche theory that such differences have evolved to reduce competition. Therefore, ecomorphology is a paradigm that allows researchers to study how different species have become adapted, or not, to their environments. Over the last 40 years, this field has matured

A   *Interspecific*   **Morphology → Performance → Habitat use**

B   *Interspecific*   **Morphology → Performance → Fitness**

FIGURE 1 A heuristic diagram showing relationships between morphology, performance, and either habitat use (panel (A), intraspecific), or fitness (panel (B), interspecific). The arrows do not imply causality directly, but generally mean that variation at each level (e.g., morphology) influences the next level up (e.g., performance) more than the other way around.

and changed, especially in regards to researchers adding in direct measurements of animal function that enable a more complete mapping of morphology → performance → resource use. This trend has been driven by the emergence of field-portable technologies for measuring animal performance, such as portable racetracks, external monitors for measuring speed or metabolic rate, and high-speed video cameras. While the field of ecomorphology has focused primarily on the concept of adaptation among well-defined species, many of the same ideas and approaches have been applied to intraspecific variation to test ideas about adaptation in the context of natural and sexual selection. The idea here is to examine linked variation among individuals between three levels: morphology, performance, and fitness (with fitness substituting for resource use at the intraspecific level; Fig. 1). This practice both tests whether individual variation in morphology is predictive of variation in performance and also whether this morphological variation is adaptive by being under the influence of natural selection. In practice, researchers have not always examined all three components simultaneously, and have mixed and matched various elements (e.g., morphology → performance, performance → fitness, etc.), but this approach is complementary to the interspecific studies. The intraspecific approach relating morphology, performance, and fitness is fundamentally important for ecology because it addresses how environmental and intrinsic factors influence demographic processes of life, death, and reproduction. This explosion of papers examining links between ecology and function at both interspecific and intraspecific levels has allowed evolutionary ecologists to address issues such as how communities are structured, whether members of ecological communities differ in just one aspect (e.g., morphology), or more (morphology, function) aspects, and whether animals use different functional strategies that enable them to occupy different habitats.

INTERSPECIFIC VARIATION: EVOLUTIONARY RADIATIONS

Studies that have linked animal function/performance to ecology among species have done so at various scales. Some have investigated sympatric, or nearly so, groups of closely related species that co-occur within the same community, whereas others have examined broader radiations of species that occur across a wide range of habitats (and therefore could not reasonably be called communities). No studies, to our knowledge, have comprehensively examined multiple taxonomic groups within the same community in terms of examining relationships between function and ecology.

One of the unresolved questions in ecology concerns whether animal species can play different roles within communities. Are species closely bound to the use of a single resource, or can they use many different resources, and therefore become generalists? Functional tests are especially valuable for addressing these questions because they are often used directly to acquire resources (such as in the case of feeding), whereas phenotypic variation alone may be deceiving. The classic (and somewhat dated) view of animal communities is that each species occupies separate niches and uses nonoverlapping resources. While this view has been updated and modified with the addition of new data over the least several decades, recent functional studies have revealed new twists for both how communities and broader radiations perform tasks, and how those tasks can be mapped onto phenotypic variation. Even in the absence of empirical data, however, the expectation of one-to-one (i.e., each species with a distinct phenotype and function occupying a distinct niche) matching is probably unrealistic in most cases for two reasons. First, most phenotypic traits are multifunctional (one-to-many), meaning they can perform several different tasks (consider the human hand, which can perform many different tasks, including gripping objects, throwing, or punching). Second, because of this generality, many different phenotypes might also perform the same function (many-to-one).

Studies with fish and bats both show that there are either weak links between animal function and resource use (bats), or some evidence of many-to-one mapping between jaw morphology and function (fish). Fish jaws present great potential for multifunctionality; they are composed of a complex series of levers (jaw bones, muscles, and soft tissue) that can be moved and used in a variety of ways (e.g., by projecting certain mouthparts, suction, etc.). Fish also feed in myriad ways, ranging from suction feeding (in which no real biting or teeth are required), to durophagy (the crushing of hard prey, such as snails, which typically requires a robust jaw and teeth). Research by Peter Wainwright and his colleagues shows that fish jaws from large and complex coral reef communities, and from larger fish evolutionary radiations, seem to exhibit both many-to-one and one-to-many mapping of fish jaw lever systems (a proxy for functional data on fish jaws) and the kind of prey they consume (e.g., suction feeding, durophagy). In other words, species that consume similar prey can have different jaw morphologies with similar functional capacities. On the other hand, morphologically similar species can differentiate in diet due to functional versatility. The potential implication of this phenomenon is that the most common force driving

evolutionary divergence within communities, namely, interspecific competition, is unlikely to result in a one phenotype–one function mapping. This also shows the importance of examining functional data, as simple examination of phenotypic data would lead to the erroneous conclusion that communities are highly ordered and structured, whereas the inclusion of functional data shows more clearly the complex many-to-one or one-to-many patterns. Moreover, these patterns also suggest that morphological differences may have arisen in some cases for reasons other than resource use, although the alternative causes remain opaque. Other studies with bats show that there are also weak linkages between diet and bite force performance, except for highly specialized species, such as those that consume nectar or lap blood. This general lack of connection between function and diet observed in bats and fish is reminiscent of classic ecomorphological studies that showed little correspondence between habitat use and morphological form, such as for some Midwestern bird communities studied by John Wiens and his colleagues. These early studies were instrumental in broadening the approaches of ecomorphologists beyond the well-studied ideas of classic niche segregation theory.

At the other extreme, there seem to exist some systems that show links between morphology, function, and resource use, most notably within relatively simple island systems that likely arose via adaptive radiation. *Anolis* lizards of the Caribbean are notable for their high level of diversity (Fig. 2), both in terms of species numbers (over 400 species in the Caribbean, Central and South America)

FIGURE 2 Images of five different *Anolis* lizard ecomorphs from the Dominican Republic showing the diversity of form in this genus. Carribean anoline lizards appear show a clear division of matching between morphology, habitat use, and function, suggesting a relatively rare example of "one-to-one" matching. Photographs by Duncan J. Irschick.

as well as behavior and morphology. Caribbean islands are composed of assemblages of anole lizard species of varying numbers (from over 50 in Cuba and Hispaniola to 7 in Jamaica), and each possesses several ecomorphs, which are groups of species that occupy distinct habitats (e.g., tree trunks, crowns of trees) and have corresponding differences in limb, body, and tail dimensions, as well as performance capacities, such as maximum sprint speed on flat surfaces and on rods of varying diameters. This matching of morphology, habitat use, and performance capacities is repeated across islands with remarkable consistency, reinforcing the view that ecological communities can show a one-to-one pattern. Similarly, Darwin's finches (*Geospiza*), of which only 14–15 species exist, show clear relationships between beak morphology (namely, the width and depth of the beak), maximum bite force, and the average size of the prey they consume (e.g., either hard or softer prey), leading to both strong relationships between resource use and function, and also a division of species that access largely nonoverlapping sets of resources. The only caveat to this latter example is that bird beaks are also used for song, a vital behavior employed primarily by males to attract females, and there is evidence that the morphological and kinematic features that promote bird song do not always coincide with features that might be useful for other tasks, such as feeding.

The ecological factors that underlie the formation of communities or that have influenced the broader radiations of animals may explain some of these different outcomes. Two key factors may be the rapidity of the radiation and the simplicity of the ecological context in which species evolve. In cases where evolutionary radiations are rapid and where interspecific competition is intense, likely because of limited resources, then extreme specialization for different roles and a corresponding one-to-one matching can evolve and persist, a scenario that closely fits the profile of Caribbean anoles. By contrast, when ecological conditions are more complex, and the number of potential species interactions increases greatly, and/or if the radiation or community has evolved over a much longer time period, there may be greater evolutionary pressures for species to be opportunistic to survive, therefore creating selection pressures for animals to utilize resources in a myriad of ways and potentially setting the stage for one-to-many or many-to-one mapping.

INTRASPECIFIC VARIATION: METHODOLOGY

Within species, the notion of variation in animal function has often been regarded as a nuisance to be ignored or minimized. Researchers often minimize variation by examining certain sets of animals (e.g., particular age classes) or by gathering large sample sizes such that the standard error around a mean is minimized. One obvious exception to this methodology is the field of behavioral ecology, which seeks to understand the behavior of individuals in relation to intrinsic or extrinsic factors. Whereas reducing variation may be desirable in certain situations, there is an increasing appreciation that intraspecific variation in morphology is often closely tied to variation in function, behavior, and resource use, often in a complex fashion that can affect ecological dynamics. The amount of intraspecific variation in functional traits can be surprisingly large, and many performance traits show a skewed distribution with a few exceptional performers, and a relatively large number of average athletes, although the reasons for this pattern are poorly understood.

Determining whether intraspecific functional variation is "real" (see below) as opposed to random measurement error or simply meaningless variation, is challenging, but two issues are paramount: First, is the trait repeatable? That is, does an animal always run quickly or bite hard if given repeated opportunities to do so? Second, can intraspecific variation in functional capacities be reliably mapped onto morphology, or is there a disconnect among individuals among these and other traits? For example, among individual animals, thicker muscle fibers are usually correlated with the production of greater force, and this means, in general, that variation among individuals in the overall size of their muscles has a direct consequence for an important ability—strength. If no such relationship exists, then variation within each level might carry less value. A third, more rarely applied idea is that variation in both the phenotype and function should have a genetic basis (i.e., be heritable). It is this last ingredient that facilitates evolution via natural selection and therefore sets the table for microevolutionary change.

ANIMAL PERFORMANCE AND FITNESS

Perhaps the most fundamental role of intraspecific functional variation in regards to ecology concerns how it is linked with fitness. Because performance traits are believed to more directly affect fitness compared to morphology per se, there is a prevailing view that natural and sexual selection should be stronger on performance than on morphology, although the available data indicate that the intensity of selection on these two kinds of traits are similar, perhaps because of interrelationships between them.

Do high levels of performance capacity increase the odds for survival and reproductive success? Of 23 studies reviewed by Duncan J. Irschick and Jean-Francois

Le Galliard in 2008, about half showed positive directional selection on performance in relation to survival. In other words, about half the time, high levels of performance increase the odds of survival, whereas the other half of the time, there was no relationship between survival and performance. This supports the view that performance traits play an important role in the demographic dynamics of life and death within ecological communities. Interestingly, stabilizing selection appears to be rare, only being documented in a few isolated cases, a pattern that is similar for morphological traits. Given how frequently environmental conditions can change, thereby altering densities of predators and prey, and therefore the relevant selection pressures, this finding suggests that directional selection may occur in some years and not others, perhaps leading to an on and off "blinking" among years, with a broader pattern of stabilizing selection over microevolutionary time periods. However, there is little data on how variation in performance capacity influences lifetime fitness, or even reproductive success. The data to date suggest that survival selection may not translate into higher reproductive success, perhaps because performance traits appear important primarily in the context of eluding predators or perhaps capturing prey, with few documented links to female choice.

PERFORMANCE TO HABITAT USE

While the correspondence between morphology, performance, and resource use among individuals (intraspecific variation) has been examined less than for among well-defined species or populations, such relationships, at least in theory, should hold. Ecologists have long recognized that individuals of the same species often differ substantially in morphology and can even occupy distinct niches. Some of the various ways that intraspecific variation can occur within a species include ecological dimorphism, ontogenetic shifts, and the presence of behaviorally distinct morphs, often among males.

Many studies have demonstrated both ecological dimorphism and variation among ontogenetic classes in ecology, but few have integrated functional data. In some cases, functional differences among the sexes enables males and females to access different resources, such as in common collared lizards (*Crotaphytus collaris*), in which males have both larger body sizes and relatively larger heads and therefore can bite harder than females. The larger gapes and higher bite forces of male collared lizards enable them to consume larger and bulkier prey than females. Similarly, ontogenetic shifts in morphology and function can enable different size classes to access different resources, such as in sheepshead fish (*Archosargus probatocephalus*), in which oral jaw-crushing force increases with body size across nine ontogenetic classes, enabling larger size classes to consume harder prey, such as bivalves and crabs.

Only a few studies have examined whether morphs within species differ in functional capacities and whether such variation has an impact on their social status, behavior, or ecology. *Uta* lizards form distinct male morphs marked by different throat colors, which correspond to different social types (territorial, mate-guarder, roaming). These morphs differ in circulating levels of testosterone (T), a hormone known to influence vertebrate muscle and certain kinds of fast-twitch performance capacity, a hormonal difference that may enable dominant morphs to more effectively patrol their territories and evict intruders during male conflicts. Other studies with lizard male morphs (*Urosaurus*) similarly show few differences in body shape or performance, suggesting that social behavior, perhaps driven in part by hormones, may be a more important determinant of social status. Not all intraspecific morphs exist as male social classes, and in some cases they may represent the nascent stages of speciation. In African seedcracker finches (*Pyrenestes*) that occur in Cameroon, two morphs exist; the large morphs have larger bills and wider lower mandibles, whereas the small morphs have small bills and narrower lower mandibles. Both feed on sedge seeds, but the large morph feeds more efficiently on harder seeds and the small morph feeds more efficiently on softer seeds, a dietary difference that is reflected in the apparent ability to crush prey and that has arisen and been maintained through disruptive selection.

FUNCTIONAL STRATEGIES

One of the advantages of examining functional traits is the potential for different strategies among individuals within species. Such strategies are analogous to human athletes employing different strategies to win despite some physical constraints. For example, the diminutive (relative to her competitors) Janet Evans (5'6", 119 lbs) won swimming races by rapidly flapping her arms against the water (the "windmill" method), while much larger female swimmers used the more traditional method of using their long arms to take long and slower strokes. The potential for such strategies in animals is great, but the study of such strategies is in its infancy conceptually and methodologically, and only a few tantalizing examples exist. When bats are forced to fly with an extra load (saline water, harmlessly voided by urination) they do not uniformly respond with the same set of compensatory

kinematic (movement) mechanisms, suggesting individual strategies to cope with an ardurous mechanical demand. A recent published study suggested that different strategies can develop in animals exposed to certain environmental conditions. In recent work by Dror Hawlena and colleagues, grasshoppers raised in an environment with predators (spiders) altered their jumping kinematics, which resulted in faster and longer jumps than those raised in a predator-free environment. The difference was solely due to alterations in jumping kinematics without conspicuous morphological change, in contrast to many studies showing morphological modifications that facilitate certain kinds of performance (e.g., burst speed) when animals develop with predators, such as tadpoles in the presence of dragonflies. Animals can also exhibit developmental plasticity in terms of how performance capacity interacts with environmental factors such as food. Research on common lacertid lizards (*Lacerta vivipara*) shows that when juvenile lizards are fed *ad libitum*, those with low endurance catch up and those with high endurance decline. Dietary restriction, paradoxically, allows juvenile lizards with high endurance to retain their locomotor advtange as they mature. In other words, there may be genetic variation both in overall performance proficiency and in how animals respond to environmental cues as they mature.

One promising yet largely unexplored area concerns variation in motivation to perform maximally. Such individual difference in willingness, if proven, would be analogous to "behavioral syndromes" documented in the field of behavioral ecology. This phenomenon refers to consistent individual correlative patterns among multiple behaviors, as when some individuals are bolder than others across all contexts (e.g., being more aggressive toward conspecific individuals, less flighty toward predators, and so on). Similarly, such "performance syndromes" might exist within species, as ontogenetic classes of animals (e.g., juveniles, adults) often show variation in their propensity to run to the limits of their maximum capacities, perhaps because of the differential threat they perceive from predators as a result of vulnerability at different life stages. Interestingly, if such variation in motivation or other manifestations of performance syndromes existed, this might be one reason for a general lack of correspondence between morphology and performance within species. These compelling examples provide suggestive hints as to how animals functionally cope with different environments, but far more data are needed.

SEE ALSO THE FOLLOWING ARTICLES

Behavioral Ecology / Integrated Whole Organism Physiology / Movement: From Individuals to Populations / Phenotypic Plasticity

FURTHER READING

Alexander, R. M. 2003. *Principles of animal locomotion*. Princeton: Princeton University Press.
Bennett, A. F., and R. B. Huey. 1990. Studying the evolution of physiological performance. *Oxford Survey of Evolutionary Biology* 7: 251–284.
Biewener, A. A. 2003. *Animal locomotion*. Oxford: Oxford University Press.
Hawlena, D., K. Holger, E. R. Dufresne, O. J. Schmitz. 2011. Grasshoppers alter jumping biomechanics to enhance escape performance under chronic risk of spider predation. *Functional Ecology* 25: 279–288.
Irschick, D., J. Meyers, J. Husak, and J. Le Galliard. 2008. How does selection operate on whole-organism functional performance capacities? A review and synthesis. *Evolutionary Ecology Research* 10: 177–196.
Podos, J., and S. Nowicki. 2004. Beaks, adaptation, and vocal evolution in Darwin's finches. *BioScience* 54: 501–510.
Vogel, S. A. 1988. Life's devices: the physical world of animals and plants. Princeton: Princeton University Press.
Wainwright, P. C., and S. M. Reilly, eds. 1994. *Ecological morphology: integrative organismal biology*. Chicago: University of Chicago Press.
Wiens, J. A., and J. T. Rotenberry. 1980. Patterns of morphology and ecology in grassland and shrubsteppe bird populations. *Ecological Monographs* 50: 287–308.