Processes and predictions in ecological models: logic and causality

Christian Damgaard

Department of Ecoscience, Aarhus University

Abstract

To make credible ecological predictions for terrestrial ecosystems in a changing environment and increase our understanding of ecological processes, we need plant ecological models that can be fitted to spatial and temporal ecological data. Such models need to be based on sufficient understanding of ecological processes to make credible predictions and account for the different sources of uncertainty. Here, I argue (1) for the use of structural equation models in a hierarchical framework with latent variables and (2) to specify whether our current knowledge of relationships among state variables may be categorized primarily as logical (empirical) or causal. Such models will help us to make continuous progress in our understanding of and ability to predict the dynamics of terrestrial ecosystems and provide us with local predictions with a known degree of uncertainty that are useful for generating adaptive management plans. The hierarchical structural equation models I recommend are analogous to current general epistemological models of how knowledge is obtained.

Keywords: structural equation model; hierarchical model; ecological prediction;

Introduction

The climate and environment are changing, and the effects of these changes on ecosystems and biodiversity are well documented (IPBES 2019). It is becoming increasingly important to develop credible ecological predictions, but the limits of such a predictive approach are an open question. To what extent is it possible to predict the future of ecosystems in a changing environment? What ecological processes are most likely to dominate future ecosystem changes? How fast will ecosystem changes occur? Here, I focus on the vegetation component of the ecosystem as a dynamic system that is influenced by climate, interactions among organisms and disturbances as an example of attempts to develop predictive ecological models in general.

The conceptual framework and study methods of plant ecology, and therefore of plant ecological models, are embedded in concepts of time and space. One of the few generalizations that we can make about vegetation is that it is spatially heterogeneous, and the spatial patterns of the vegetation depend mainly on the environment and historical legacies (Greig-Smith 1979; Ricklefs 2004; Svenning and Skov 2004; von Humboldt and Bonpland 1805). Even in a homogenous environment, plant species are aggregated at large and small spatial scales. At a larger scale, plant species are aggregated among sites due to random extinction events and limited possibilities for colonization (Cordonnier et al. 2006; Leibold et al. 2004; MacArthur and Wilson 1967; Rees et al. 2001). Within sites, plants may be aggregated due to clonal growth, limited seed dispersal and size-asymmetric competition (Herben et al. 2000; Pacala and Levin 1997; Stoll and Weiner 2000). Dominant plant species may affect the soil and lead to important feedback mechanisms, so the vegetation and soil may be described as a dynamic system in time and space, where different phases of the vegetation simultaneously may co-occur in different patches (Watt 1947) or the ecosystem can develop in a successional pattern.

Moreover, it is a truism that ecological processes occur in time, and they can only be understood when temporal dynamics are taken into account (Damgaard 2019a; Damgaard and Weiner 2017; Kratz et al. 2003; Pickett 1989). Generally, environments are non-stationary, and spatial and temporal variability in ecological systems are of fundamental importance to ecosystem dynamics (Chesson 2000b; Cushman 2010; Huston 1979). For example, the successional changes in tree-dominated ecosystems are expected to be highly erratic and unpredictable at the local scale due to variable factors, such as herbivore abundance (Schippers et al. 2014), while often being highly predictable at a larger scale. Temporal changes in the water level of Atlantic wetlands have been shown to affect both the nature of interspecific competitive relationships and the overall importance of competition for regulating population growth (Merlin et al. 2015).

When studying ecological processes, the ecosystem is a useful conceptual entity that focuses attention on important system properties (Tansley 1935). For example, Lenton et al. (2021) argue that self-perpetuating feedback cycles involving biotic as well as abiotic components may be critical for the stability and long-time success of ecosystems, and Tansley (1935) proposed that stable ecosystems will outlast unstable ecosystems. Furthermore, when ecosystems are situated near other ecosystems in a landscape, another level of emergent properties may arise among the ecosystems at the landscape level (Leibold et al. 2004). For example, environmental variation within a landscape may lead to a storage effect with coexistence of species that would not otherwise coexist (Chesson 2000a; Chesson and Warner 1981).

The goal of much ecological research is to understand the underlying causes of observed patterns, which enables us to make credible ecological predictions. However, an observed ecological pattern may be caused by several different processes and involve several important contingencies (Damgaard 2019a; Simberloff 2004; Vellend 2010), and the relationship between process and pattern may be influenced by time lags of unknown duration (Svenning and Sandel 2013). Consequently, although there is some level of generality, the quantitative importance of different ecological processes is generally unknown and expected to differ among sites. The resulting variation in vegetation dynamics among seemingly comparable sites leads to a level of complexity that limits our ability to generalize our understanding of the ecological mechanisms underlying observed patterns. While historical contingencies and variation among apparently similar sites preclude the generalization of ecological findings, careful local studies of the mechanisms conducted at the relevant spatial and temporal scales do provide causal information of the factors that determine important community dynamic features (Simberloff 2004; Weiner 1995).

There is increasing awareness of the limitations of classical community ecological models, e.g. Lotka-Volterra-type competition models, to predict ecosystem effects of a changing environment (Clark et al. 2020; Damgaard and Weiner 2021; Eigentler 2021; Martyn et al. 2020; Mayfield and Stouffer 2017; Neill et al. 2009; Vellend 2010). For example, it is widely appreciated that frequency-dependent species interactions, in which relatively rare species are favored over more common species, may play an important role in plant species co-existence and community dynamics in plant communities (Chisholm and Fung 2020; Connell et al. 1984). Species-specific soil-plant interactions have received increasing interest as a potentially important and general mechanism for regulating plant populations by hindering local establishment and growth of conspecific plant species in the next generation (Aldorfová et al. 2022; Heinen et al. 2020; Mazzoleni et al. 2015a; Mazzoleni et al. 2015b; van der Putten et al. 2013). In an attempt to cope with the ecological complexity, Vellend (2010) suggested partitioning all ecological processes that result in species abundance changes into a logically complete set of disjunct categories: selection, drift, dispersal and speciation (the nomenclature of these categories may be discussed in an ecological context, but the important characteristic is that they are logically disjunct and complete). Historically, such a classification scheme has been successful in evolutionary biology, and Vellend (2010) argued that the reason this approach has not been used more in ecology is due to universal familiarity with ecological patterns. For example, since prehistoric time it has been common knowledge that plant growth is reduced when plants grow close together, and it therefore seems natural to focus on the effects of competitive interactions rather than introducing a logically complete set of possible processes in community ecology.

The quantitative scientific exploration of the underlying mechanisms that lead to observed ecological patterns among ecological state variables in time and space relies on statistical analyses, but often the investigation meets the obstacle summarized in the well-known phrase “correlation does not imply causality”. More precisely and in the language of Jaynes (2003), the outcome of a statistical analysis is the quantification of logic dependencies (see box 1).

BOX 1

Probability theory: the logic of science

In an inspiring book published after his death, Jaynes (2003) argues that probability theory in a Bayesian setting may be thought of as a quantitative extension of Aristotelian logic, which enables us to engage meaningfully in plausible reasoning and provides a method for quantifying our belief in different hypothesis. From a number of verbal statements of the necessary requirements to a quantitative system of plausible reasoning, Jaynes develops the foundation of probability theory, and show that Bayesian statistics is a mathematical consistent tool for quantifying and communicating our uncertainties of a studied phenomenon (Cox 1946; Jaynes 2003).

Moreover, just as a true statement in the language of Aristotelian logic does not imply a causal mechanism, logical dependencies expressed in the extended quantitative logic of probability theory do not imply causal dependencies (Jaynes 2003). Logic is the general science of all possible relationships between objects, whereas causality is a specific type of ordered relationship that we hypothesize for the world to make sense.

The notion of an extended quantitative logic *sensu* Jaynes (2003) provides important insights into the nature of plausible reasoning and how we do science; scientific discoveries are most often made by generalization from an observed pattern using the induction principle, and extended quantitative logic shows why this is indeed possible and works in practice. By observing a number of white swans we may say something intelligent about the probability that the next swan we observe is white.

Therefore, I have chosen here to use the term “logical dependency” instead of “(statistical) association” or “correlation”, which are more commonly used in the relevant literature. More practically, the use of the concept of extended quantitative logic and Bayesian statistics impose us to estimate the uncertainties of our predictions.

The distinction between logical and causal dependencies is subtle and not always clear, but important nonetheless. Logical dependencies are investigated using probability theory by calculating the plausibility of different hypotheses when knowledge is limited, and even though it is assumed that the calculated probabilities are determined by causal dependencies, the calculation of such probabilities may be decoupled from causal dependencies. For example, if we want to know if a plant experienced competition from a neighboring plant at day t, when we know that the plant experienced competition at a later day $t + \tau$, then we may calculate the probability P (competition at day t | competition at day $t + \tau$). Such calculations allow us to make ecological predictions, even though the estimated relationship is not a causal dependence, since the outcome at a later day cannot influence the outcome at an earlier day. It can also be shown that the accumulation of new causally-independent data may lead to logical dependencies when investigating the plausibility of multiple hypotheses (Jaynes 2003).

Logical and causal dependencies are not distinct entities, but form a continuum from almost pure logical (statistical) relationships to primarily causal relationships, and often the development of a credible causal hypothesis, i.e. scientific knowledge, goes through a process of collecting relevant data and investigating data by increasingly refined steps of inductive reasoning (Jaynes 2003). The distinction is further complicated by unmeasured confounding factors that affect observed independent variables. These confounding factors may lead to "spurious correlations" or omitted variable bias (Gelman and Hill 2007). The effect of the confounding factors could potentially be teased out in an analysis of instrument variables or using structural equation models (Grace 2021; Rinella et al. 2020), and causal independence may be inferred from randomized experimental manipulations or controlled manipulations using the logical do-operator introduced by Pearl (Lindley 2002; Pearl 2009). However, many important ecological processes operate at such large spatial and temporal scales that experimental manipulations are impossible (Shipley 2016), although it may be possible in some cases to combine the information from manipulated experiments and large-scale observational studies (Benedetti-Cecchi et al. 2018).

Structural equation models

The combined effect of both logical and causal dependencies among state variables at the ecosystem level can be effectively modelled using structural equation models, where the observed logical dependencies are modeled using hypothesized causal relationships that often are specified as a directed acyclic graph (Grace et al. 2010; Grace et al. 2012; Shipley 2016; Wright 1921). The hypothesized causal relationships summarize the domain knowledge of the relevant ecological process and may be obtained using differential equation models or general assumptions of mass conservation. However, it is important to note that typically the causal relationships in structural equation models are assumed to be linear functions (but see Grace et al. 2012). This means that structural equation models normally do not compare the likelihood of different functional forms of the dependencies among the observed state variables, i.e. the nature of the ecological processes that link the observed variables is not investigated. For example, in a typical structural equation model of the possible dependencies between the observed growths of two plant species, it is not addressed whether the growth data may be adequately explained by Lotka-Volterra type of interspecific competition or whether it is relevant to assume some additional form of frequency-dependency (e.g. Clark et al. 2020; Damgaard and Weiner 2021). This insensitivity to the underlying processes that link the measured ecological state variables in the hypothesized causal network may be critical for the ecological predictions that are made using a fitted structural equation model. For example, if frequency-dependency plays a role in interspecific competitive interactions, then more species can coexist (Chisholm and Fung 2020).

Consequently, when specifying structural equation models it is important to model the underlying ecological mechanisms in more detail than is usually done and to apply hypothesized mechanistic functional relationships among state variables, e.g. by assuming that plant growth decreases with the relative frequency of the plant species in a non-linear way (Damgaard and Weiner 2021), or by replacing linear relationships with sigmoid growth functions, which are known to capture population ecological mechanisms at a sufficiently high level of aggregation to allow statistical treatment (Damgaard 2005; Damgaard and Weiner 2008). In this way, logical dependencies in structural equation models may be replaced by dependencies of a more causal nature, and in my opinion, the adopted terminology of logical and causal dependencies used here highlights this overall ecological research agenda. Clark et al. (2020) warn against the use of process models of relatively high complexity due to the risk of overfitting, but if frequency-dependency or feedback cycles are important ecosystem features, then such a level of model complexity is needed in order to make credible ecological predictions.

Grace et al. (2010), Shipley (2016) and others have made similar arguments concerning the use of structural equation models for making stronger inferences and better predictions in ecology. Here, I add to these points (1) the separation of logical from causal relationships in the analysis, thus making (2) the link to Jaynes's epistemological framework more direct.

Usually, there is sizeable variation in the measured independent variables due to measurement and sampling errors in ecological studies (Muff et al. 2015; Yanai et al. 2018). Such measurement errors are typically not taken into account in ecological and environmental modelling, where measured independent variables are treated as if they are constant entities. This is unfortunate, since it has been shown that even unbiased and normally distributed sampling and measurement errors may lead to important model and prediction bias, a phenomenon known as "regression dilution" (Carroll et al. 2006; Damgaard 2020b; Damgaard and Weiner 2021; Detto et al. 2019). Another typical problem in the analysis of ecological experiments is the use of the language of the experimental design, e.g. “control” and “treatment”, as orthogonal fixed independent factors, although there is often large variation and covariation among the variables that are manipulated in the experiment, e.g. soil humidity and temperature, which are the real independent variables of interest in the study and the ones on which inferences and conclusions are based (Damgaard et al. 2018).

How can we best improve the empirical modelling of plant ecological processes in space and time in order to make credible ecological predictions? Following the suggestion by Clark (2005; 2007) and several others, I argue for combining structural equation models (Grace et al. 2010; Pearl 2009; Shipley 2016) with the flexibility of hierarchical models with latent variables in a Bayesian setting for investigating logical and causal plant ecological dependencies in a coherent way and, at the same time, model the different sources of uncertainty (Clark 2007; Damgaard 2019b).

Hierarchical models

The use of hierarchical models with latent variables allows us to partition and model the different sources of uncertainty (Clark 2005; 2007; Wikle 2003). For example, in an empirical study of competitive plant growth there may be i) sampling and measurement errors when determining the size of plants at different times using non-destructive measuring methods, ii) structural uncertainty in the form of competitive interactions, e.g. whether frequency-dependent competitive effects play a role or not and iii) the possible confounding effects of unmeasured variables on individual plant performance at both an early and later growth stage due to spatial variation in soil nutrients, pathogens and other factors (Fig. 1).

The effect of sampling and measurement uncertainty is accounted for in hierarchical models by introducing latent variables that model the true, but unknown, value of different state variables (Clark 2007; Gelman and Hill 2007). The logical link between the latent variables and the corresponding observed data depends on the type of data and the measurement method. For example, in a study where plant biomass is measured destructively and it is known that the variance increases with the mean, then it may be appropriate to model the observed biomass by a gamma distribution with latent variables as mean parameters and a common scale parameter expressing the relationship between the mean and the variance. In a multi-species study of pin-point plant cover data, it may be appropriate to model the observed number of pin-point hits by a reparametrized Dirichlet-multinomial distribution, in which the latent variables are the relative mean cover of the different species and where the expected inter-specific spatial aggregation at the plot level is modelled by a parameter (Damgaard 2015; 2018).

When fitting fine-grained ecological data, it is relevant to examine how well different process models are supported by the data (e.g. Clark et al. 2020; Damgaard and Weiner 2021) by estimating the expected out-of-sample prediction error of the different functional models using the Watanabe-Akaike or other information criteria (Gelman et al. 2014), although the predictive performance of a model may be a poor measure of whether the correct causal relationships has been specified (Arif and MacNeil 2022).

In the absence of prior knowledge and when modelling spatial and temporal vegetation data on relatively large scales, it may be a useful strategy to keep it simple and model changes using logically complete sets of ecological process categories such as those suggested by Vellend (2010), of which selection at the community level is the most important category. Alternatively, it may be attractive to model some of the processes that are not of immediate interest by flexible non-parametric models if there is a real concern that some of the hypothesized functional relationships may be misspecified (Walker 2013).

In more complicated ecological models, it is important to specify whether the hypothesized dependencies among the state variables may be categorized as primarily logical or causal. In any case, when using relatively simple models it becomes increasingly important to quantify the uncertainties and restrain from making predictions outside the spatial and temporal domain of the data. For example, in a spatial and temporal comprehensive, but coarse-grained, study of vegetation changes in wet heathlands and how these changes were affected by the environment, changes in cover of selected species were modelled by simple linear models, thus ignoring any higher-order community dynamic effects (Damgaard 2019b). Instead, care was taken to partition spatial and temporal effects and different sources of uncertainty in a hierarchical model, which enabled short-term ecological predictions that may be used for generating local adaptive management plans (Damgaard 2020a). In this particular case, spatial and temporal effects were separated by assuming that the state of plant communities in the first year of the sampled time series was controlled by spatial processes, which included the effects of the environment, as well as historical legacies, contingencies and succession history at the site level prior to initiating sampling. The temporal processes were then modelled in a state-space model based on the observed yearly changes in vegetation cover during the sampling period (Fig. 2). In ecological systems with significant among-site dispersal, it may be relevant to model the effect of the dispersal in spatial and temporal diffusion models (e.g. Wikle 2003).

Ideally, the possible confounding effects of unmeasured variables should be investigated using an instrument variable analysis (Gelman and Hill 2007; Rinella et al. 2020). However, this is often not possible and, instead, we may account for possible confounding effects of unmeasured variables by modeling the unexplained covariance between the observed variables, such as the unexplained covariance between early and later growth stages (Fig. 1).

It is noteworthy that the treatment of measurement errors and causal dependencies in structural equation models in a hierarchical setting mimics general epistemological models (see Box 2) and, consequently, provides a general recipe for how to link models with data and causal hypotheses into a mathematical framework.

BOX 2

General epistemological models

From a more general epistemological point of view, hierarchical models with latent variables mirror well-known theoretical philosophical models of how knowledge of the world is obtained. Kant (1781) suggested that there is a fundamental division between a world of phenomena ("Das ding an sich") and the observer. When the observer senses a certain phenomenon (the object), a representation of the object transcends into the mind of the observer and starts to make sense (*a posteriori* knowledge). In this process, the representation of the object is merged with several sources of relevant *a priori* knowledge, such as the concepts of causality, space and time.

In a hierarchical model, data corresponds to Kant's phenomena, and the latent variables and their relationships correspond to the transcended representation of the objects in the mind of the observer. Similarly, Wittgenstein (1922) described how we make mental pictures of objects and how these pictures are models of reality. He suggested that the relationship among the different picture elements corresponds to the relationship among the objects and, thus, provides a representation of the structure of the reality. For example, in proposition 2.15 he writes: “That the elements of the picture are combined with one another in a definite way, represents that the things are so combined with one another. This connection of the elements of the picture is called its structure, and the possibility of this structure is called the form of representation of the picture.” (Wittgenstein 1922). This description is almost a definition of hierarchical structural models where the “pictures” corresponds to quantitative latent variables, although the latter is a more restricted concept.

Discussion

It is important to partition and model the different types of observed variance, and hierarchical models are often the best tool for this purpose. For example, in a relatively fine-grained study of plant competitive growth of *Festuca ovina* and *Agrostis capillaris* it was concluded that the inclusion of frequency-dependence and measurement error improved model performance greatly, but taking possible unmeasured variables into account did not. Furthermore, when sampling and measurement errors were taken into account, the resulting ecological prediction differed qualitatively from the corresponding non-hierarchical models without latent variables (Damgaard and Weiner 2021). In a study of tree fecundity Clark et al. (2004) found that random variation among individual trees and among years dominated the variation in tree fecundity. If this random variation was omitted from the analysis, parameter estimates would have been biased and led to qualitatively erroneous conclusions. They also found that estimated random variation could partly explain the observed pattern of coexistence among the tree species.

The current poor performance of ecological predictions is probably due to the relatively low quantity and quality of spatial and temporal ecological data. The increasing use of advanced technology in ecological monitoring, such as drones and satellites, may allow us to make better predictions. As fine-grained plant growth data at increasing spatial scales become available, it may be feasible to fit structural equation models, in which different hypothesized species interaction processes are statistically compared. By increasing the resolution of spatial and temporal ecological data, it should become possible to increase both prediction precision and accuracy. At the limit of better and better local data and increasingly detailed understanding of the ecological processes, the increase in predictive performance may reach a boundary. If the intrinsic boundary of ecological predictive performance is low due to historical or other contingencies, it may be necessary to rethink our current strategy of adaptive management of ecosystems based on ecological predictions in nature conservation.

For some ecosystems that operate on relatively small temporal and spatial scales, such as microbial decomposer systems, it may be possible to perform manipulative experiments that allow us to test and validate complete causal hypotheses of the system dynamics. However, the relevant temporal and spatial scales of most ecosystems are so large that a purely experimental approach is not feasible, so in our modelling of ecosystem processes we must recognize that our current ecosystem knowledge is incomplete and not be overconfident in hypothesized causal mechanisms. Consequently, ecological models with both causal and logical dependencies will be needed in the attempt to make ecological predictions. The goal of ecology as a science, however, is to increase our mechanistic understanding of the different processes that determine the abundances and distributions of species, and thereby enable us to convert assumed logical dependencies into more causal dependencies among state variables in our structural equation models. The current state of structural equation models in plant ecology, where many dependencies are strictly based on logical and statistical arguments, will act as a guide into where knowledge is missing and suggest future research agendas. Moreover, it is important to capture the causal mechanisms at a sufficiently high level of aggregation to allow statistical treatment of available ecological data in a structural equation model. It is not necessary, or even desirable, to model all the known mechanistic details, such as modelling every single birth and death event using individual based models.

In my opinion, the image of plant ecology as a “soft” science, which plays a disproportionally small role in political decisions, hinges on our failure as a scientific community to make credible predictions where uncertainties are quantified. Currently, the majority of plant ecological predictions, from local management plans of natural habitats to global assessments of biodiversity, are broad verbal statements without any attempts to quantify the uncertainty of the predictions. To communicate credibility of the verbal predictions, a common practice is to call on the consensus of experts. For example, in a report from IPBES (2019) 150 selected experts extracted the content of more than 15,000 scientific publications to make a global assessment of biodiversity, and the obtained conclusions, including the recommended actions, were approved at a plenary meeting. How is such a process even possible and what is the role of the assembled ecological data in the 15,000 scientific publications in reaching a consensus for the recommended actions? No body of even distinguished experts can grasp the complexity of ecosystem dynamics in a changing environment and fairly assess the wealth of relevant spatial and temporal data in round-table and panel discussions. When good data are available, it would be better to apply a quantitative modelling approach, where the uncertainty of the effects of the recommended actions may be assessed in a systematic way. The results of the quantitative analyses can then be interpreted by experts for decision makers.

Acknowledgements

Thanks to Jacob Weiner for valuable discussions and suggestions in preparing this paper

Figures

Fig. 1. : Hierarchical model of competitive plant growth from plants at an early and later growth stage. The true, but unknown, plant size at the early and later growth stage is modelled by latent variables (squares) and denoted X, Y, respectively. The latent variables are logically linked to the observed data (circles), which are denoted x, y, respectively, with the corresponding latent variables as mean values. The sampling and measurement uncertainty are modelled with σ_x and σ_y, respectively. The competitive growth process is modelled by the structural function, F, which depends on X and possibly on some measured environmental variables z_m. The effect of unmeasured variables affecting both earlier and later growth stages is modelled by σ_{xy} as the part of the covariance between the early and later growth stages that is not explained by the independent factors in the structural model.

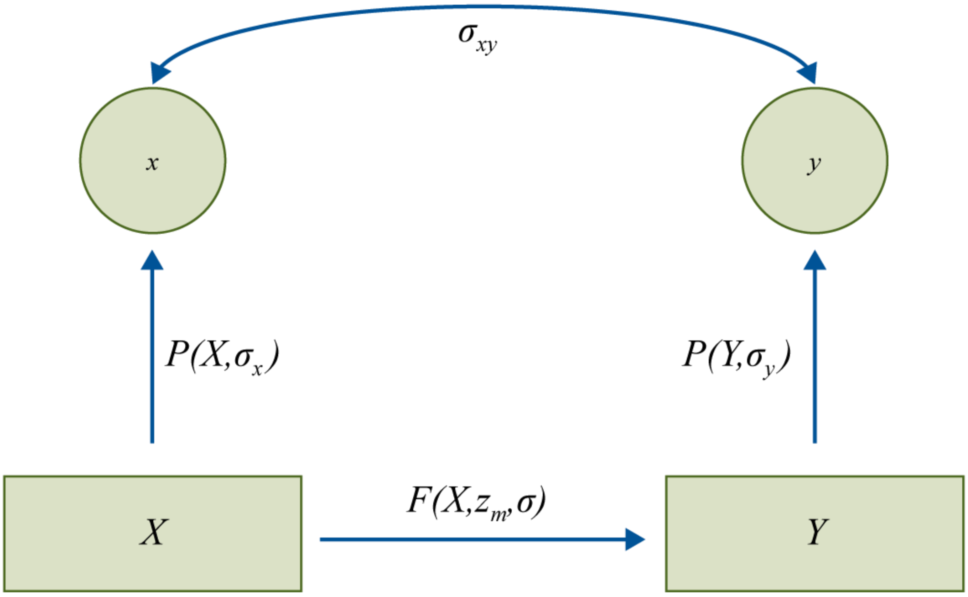
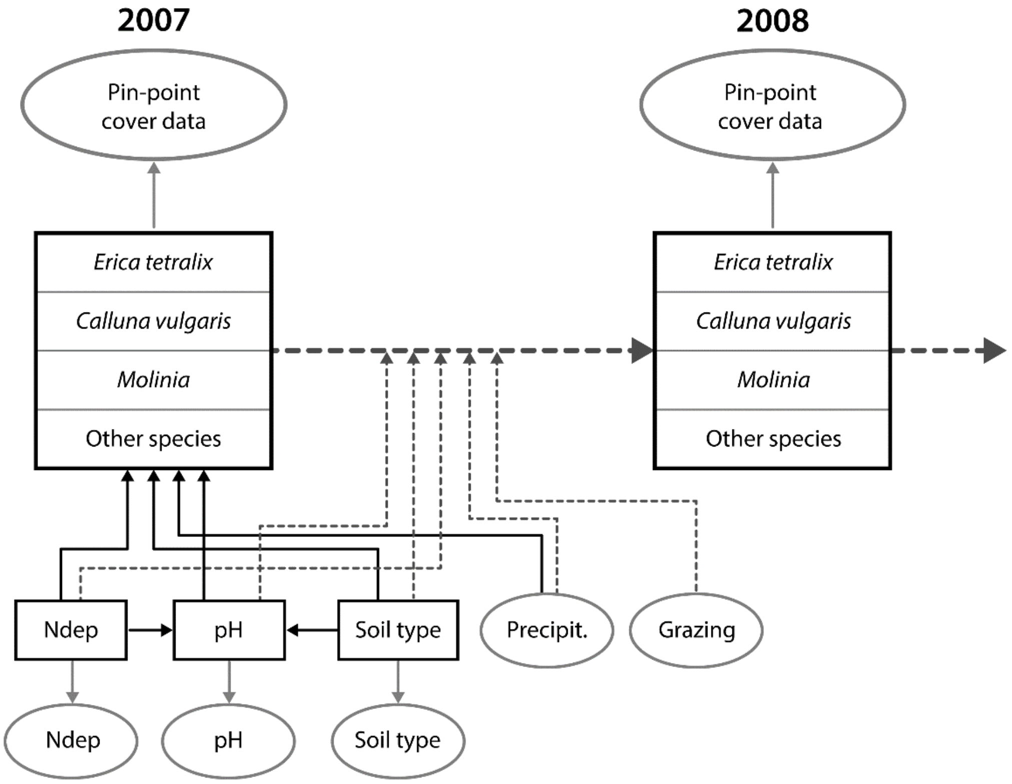


Fig. 2. Outline of a structural equation model where the spatial and temporal processes are separated. The spatial variation in vegetation cover in 2007 is modelled by nitrogen deposition (Ndep), soil pH (pH), soil type and precipitation (Precipit.). The yearly change in vegetation cover from 2007 to 2014 (only a single yearly change is shown in the figure) is modelled by all the former variables as well as grazing. The square boxes are latent variables and the oval boxes are data. The full black arrows denote spatial processes, the dotted black arrows denote temporal processes and the grey arrows denote the modelling of sampling and measurement error (Damgaard 2019b).



References

Aldorfová A, Dostálek T, Münzbergová Z. 2022. Effects of soil conditioning, root and shoot litter addition interact to determine the intensity of plant–soil feedback. Oikos. 2022(6):e09025.

Arif S, MacNeil MA. 2022. Predictive models aren't for causal inference. Ecol Lett. 25(8):1741-1745.

Benedetti-Cecchi L, Bulleri F, Dal Bello M, Maggi E, Ravaglioli C, Rindi L. 2018. Hybrid datasets: Integrating observations with experiments in the era of macroecology and big data. Ecology. 99(12):2654-2666.

Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu C. 2006. Measurement error in nonlinear models: A modern perspective. Boca Raton: CRC.

Chesson P. 2000a. General theory of competitive coexistence in spatially-varying environments. Theoretical Population Biology. 58(3):211-237.

Chesson P. 2000b. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics. 31:343-366.

Chesson PL, Warner RR. 1981. Environmental variability promotes coexistence in lottery competitive systems. Am Nat. 117(6):923-943.

Chisholm RA, Fung T. 2020. Janzen-connell effects are a weak impediment to competitive exclusion. Am Nat. 196(5):649-661.

Clark AT, Ann Turnbull L, Tredennick A, Allan E, Harpole WS, Mayfield MM, Soliveres S, Barry K, Eisenhauer N, de Kroon H et al. 2020. Predicting species abundances in a grassland biodiversity experiment: Trade-offs between model complexity and generality. J Ecol. 108(2):774-787.

Clark JS. 2005. Why environmental scientist are becoming bayesians. Ecol Lett. 8:2-14.

Clark JS. 2007. Models for ecological data. Princeton: Princeton University Press.

Clark JS, LaDeau S, Ibanez I. 2004. Fecundity of trees and the colonization–competition hypothesis. Ecological Monographs. 74(3):415-442.

Connell JH, Tracey JG, Webb LJ. 1984. Compensatory recruitment, growth, and mortality as factors maintaining rain forest tree diversity. Ecological Monographs. 54(2):141-164.

Cordonnier T, Courbaud B, Franc A. 2006. The effect of colonization and competition processes on the relation between disturbance and diversity in plant communities. Journal of Theoretical Biology. 243:1-12.

Cox RT. 1946. Probability, frequency and reasonable expectation. American Journal of Physics. 14(1):1-13.

Cushman SA. 2010. Space and time in ecology: Noise or fundamental driver? In: Cushman SA, Huettmann F, editors. Spatial complexity, informatics, and wildlife conservation. Springer

Damgaard C. 2005. Evolutionary ecology of plant-plant interactions - an empirical modelling approach. Aarhus, Denmark: Aarhus University Press.

Damgaard C. 2015. Modelling pin-point cover data of complementary vegetation classes. Ecol Inform. 30:179-184.

Damgaard C. 2018. The joint distribution of pin-point plant cover data: A reparametrized dirichlet - multinomial distribution. arXiv e-prints. [accessed August 01, 2018]https://ui.adsabs.harvard.edu/\#abs/2018arXiv180804582D.

Damgaard C. 2019a. A critique of the space-for-time substitution practice in community ecology. Trends Ecol Evol. 34(5):416-421.

Damgaard C. 2019b. Spatio-temporal structural equation modeling in a hierarchical bayesian framework: What controls wet heathland vegetation? Ecosystems. 22:152-164.

Damgaard C. 2020a. Adaptive management plans rooted in quantitative ecological predictions of ecosystem processes: Putting monitoring data to practical use. bioRxiv.2020.2010.2011.334789.

Damgaard C. 2020b. Measurement uncertainty in ecological and environmental models. Trends Ecol Evol. 35:871-873.

Damgaard C, Holmstrup M, Schmidt IK, Beier C, Larsen KS. 2018. On the problems of using linear models in ecological manipulation experiments: Lessons learned from a climate experiment. Ecosphere. 9(6):e02322.

Damgaard C, Weiner J. 2008. Modeling the growth of individuals in crowded plant populations. Journal of Plant Ecology. 1:111–116.

Damgaard C, Weiner J. 2017. It's about time: A critique of macroecological inferences concerning plant competition. Trends Ecol Evol. 32:86-87.

Damgaard C, Weiner J. 2021. The need for alternative plant species interaction models. Journal of Plant Ecology.

Detto M, Visser MD, Wright SJ, Pacala SW. 2019. Bias in the detection of negative density dependence in plant communities. Ecol Lett. 22(11):1923-1939.

Eigentler L. 2021. Species coexistence in resource-limited patterned ecosystems is facilitated by the interplay of spatial self-organisation and intraspecific competition. Oikos. 130(4):609-623.

Gelman A, Hill J. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press.

Gelman A, Hwang J, Vehtari A. 2014. Understanding predictive information criteria for bayesian models. Stat Comput. 24(6):997-1016.

Grace JB. 2021. Instrumental variable methods in structural equation models. Methods in Ecology and Evolution. 12(7):1148-1157.

Grace JB, Anderson TM, Olff H, Scheiner SM. 2010. On the specification of structural equation models for ecological systems. Ecological Monographs. 80:67–87.

Grace JB, Schoolmaster Jr DR, Guntenspergen GR, Little AM, Mitchell BR, Miller KM, Schweiger EW. 2012. Guidelines for a graph-theoretic implementation of structural equation modeling. Ecosphere. 3(8):art73.

Greig-Smith P. 1979. Pattern in vegetation. J Ecol. 67(3):755-779.

Heinen R, Hannula SE, De Long JR, Huberty M, Jongen R, Kielak A, Steinauer K, Zhu F, Bezemer TM. 2020. Plant community composition steers grassland vegetation via soil legacy effects. Ecol Lett. 23(6):973-982.

Herben T, During HJ, Law R. 2000. Statio-temporal patterns in grassland communities. In: Dieckmann U, Law R, Metz JAJ, editors. The geometry of ecological interactions: Simplifying spatial complexity. Cambridge: Cambridge University Press. p. 48-64.

Huston M. 1979. A general hypothesis of species diversity. Am Nat. 113(1):81-101.

IPBES. 2019. Global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services. Bonn, Germany: IPBES secretariat.

Jaynes ET. 2003. Probability theory: The logic of science. Cambridge: Cambridge University Press.

Kant I. 1781. Kritik der reinen vernunft.

Kratz TK, Deegan LA, Harmon ME, Lauenroth WK. 2003. Ecological variability in space and time: Insights gained from the us lter program. Bioscience. 53(1):57-67.

Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D et al. 2004. The metacommunity concept: A framework for multi-scale community ecology. Ecol Lett. 7:601-613.

Lenton TM, Kohler TA, Marquet PA, Boyle RA, Crucifix M, Wilkinson DM, Scheffer M. 2021. Survival of the systems. Trends Ecol Evol. 36(4):333-344.

Lindley DV. 2002. Seeing and doing: The concept of causation. International Statistical Review / Revue Internationale de Statistique. 70(2):191-197.

MacArthur RH, Wilson EO. 1967. The theory of island biogeography. Princeton: Princeton University Press.

Martyn TE, Stouffer DB, Godoy O, Bartomeus I, Pastore AI, Mayfield MM. 2020. Identifying “useful” fitness models: Balancing the benefits of added complexity with realistic data requirements in models of individual plant fitness. Am Nat. 197(4):415-433.

Mayfield MM, Stouffer DB. 2017. Higher-order interactions capture unexplained complexity in diverse communities. Nature Ecology & Evolution. 1(3):0062.

Mazzoleni S, Bonanomi G, Incerti G, Chiusano ML, Termolino P, Mingo A, Senatore M, Giannino F, Cartenì F, Rietkerk M et al. 2015a. Inhibitory and toxic effects of extracellular self-DNA in litter: A mechanism for negative plant–soil feedbacks? New Phytologist. 205(3):1195-1210.

Mazzoleni S, Cartenì F, Bonanomi G, Senatore M, Termolino P, Giannino F, Incerti G, Rietkerk M, Lanzotti V, Chiusano ML. 2015b. Inhibitory effects of extracellular self-DNA: A general biological process? New Phytologist. 206(1):127-132.

Merlin A, Bonis A, Damgaard CF, Mesléard F. 2015. Competition is a strong driving factor in wetlands, peaking during drying out periods. PLoS ONE. 10(6):e0130152.

Muff S, Riebler A, Held L, Rue H, Saner P. 2015. Bayesian analysis of measurement error models using integrated nested laplace approximations. J R Stat Soc C-Appl. 64(2):231-252.

Neill C, Daufresne T, Jones CG. 2009. A competitive coexistence principle? Oikos. 118(10):1570-1578.

Pacala S, Levin SA. 1997. Biological generated spatial pattern and the coexsistence of competing species. In: Tilman D, Kareiva P, editors. Spatial ecology the role of space in population dynamics and interspecific interactions. Princeton: Princeton University Press.

Pearl J. 2009. Causal inference in statistics: An overview. Statistics Surveys. 3:96-146.

Pickett STA. 1989. Long-term studies in ecology. In: Likens GE, editor. Long-term studies in ecology. Springer-Verlag.

Rees M, Condit R, Crawley M, Pacala S, Tilman D. 2001. Long-term studies of vegetation dynamics. Science. 293:650-655.

Ricklefs RE. 2004. A comprehensive framework for global patterns in biodiversity. Ecol Lett. 7(1):1-15.

Rinella MJ, Strong DJ, Vermeire LT. 2020. Omitted variable bias in studies of plant interactions. Ecology. 101(6):e03020.

Schippers P, van Teeffelen AJA, Verboom J, Vos CC, Kramer K, WallisDeVries MF. 2014. The impact of large herbivores on woodland–grassland dynamics in fragmented landscapes: The role of spatial configuration and disturbance. Ecol Complex. 17:20-31.

Shipley B. 2016. Cause and correlation in biology. Cambridge Cambridge University Press.

Simberloff D. 2004. Community ecology: Is it time to move on? Am Nat. 163(6):787-799.

Stoll P, Weiner J. 2000. A neighborhood view of interactions among individual plants. In: Dieckmann U, Law R, Metz JAJ, editors. The geometry of ecological interactions. Cambridge: Cambridge University Press. p. 11-27.

Svenning J-C, Sandel B. 2013. Disequilibrium vegetation dynamics under future climate change. Am J Bot. 100:1–21.

Svenning J-C, Skov F. 2004. Limited filling of the potential range in european tree species. Ecol Lett. 7(7):565-573.

Tansley AG. 1935. The use and abuse of vegetational concepts and terms. Ecology. 16:284-307.

van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, Kardol P, Klironomos JN, Kulmatiski A, Schweitzer JA et al. 2013. Plant–soil feedbacks: The past, the present and future challenges. J Ecol. 101(2):265-276.

Vellend M. 2010. Conceptual synthesis in community ecology. The Quarterly Review of Biology. 85(2):183-206.

von Humboldt A, Bonpland A. 1805. Essai sur la géographie des plantes. Chicago and London: University of Chicago Press.

Walker SG. 2013. Bayesian inference with misspecified models. Journal of Statistical Planning and Inference. 143(10):1621-1633.

Watt AS. 1947. Pattern and process in the plant community. J Ecol. 35:1-22.

Weiner J. 1995. On the practice of ecology. J Ecol. 83(1):153-158.

Wikle CK. 2003. Hierarchical bayesian models for predicting the spread of ecological processes. Ecology. 84(6):1382-1394.

Wittgenstein L. 1922. Tractatus logico-philosophicus. London: KEGAN PAUL, TRENCH, TRUBNER & CO., LTD.

Wright S. 1921. Correlation and causation. Journal of Agricultural Research. 20:557-585.

Yanai RD, See CR, Campbell JL. 2018. Current practices in reporting uncertainty in ecosystem ecology. Ecosystems. 21(5):971-981.