

# **Spatiotemporal variability and extrapolation from ecological experiments**

**Robert Frühstückl**

robert.fruehstueckl@uni-bielefeld.de

## **Abstract**

In current philosophy of science, extrapolation is seen as an inference from a study to a distinct target system of interest. The reliability of such an inference is generally thought to depend on the extent to which study and target are similar in relevant respects, which is especially problematic when they are heterogeneous. This paper argues that this understanding is underdeveloped when applied to extrapolation in ecology. Extrapolation in ecology is not always well characterized as an inference from a model to a distinct target but often includes inferences from small-scale experimental systems to large-scale processes in nature, i.e., inferences across spatiotemporal scales. For this reason, I introduce a distinction between compositional and spatiotemporal variability. Whereas the former describes differences in entities and causal factors between model and target, the latter refers to the variability of a system over space and time. The central claim of this paper is that our understanding of heterogeneity needs to be expanded to explicitly include spatiotemporal variability and its effects on extrapolation across spatiotemporal scales.

Keywords: extrapolation, ecology, experimentation, heterogeneity, scale

## 1. Introduction

The extent to which we can generalize knowledge acquired in particular situations to new contexts, questions, and environments is a pressing problem in ecology (Spake et al. 2022). Transferring findings from experiments not only supports our predictive ability and forms the basis for management intervention and public policy but is also a part of determining the scope of causal relationships. Recently, this problem has received growing attention in terms of the problem of extrapolation (Baetu 2016; Bareinboim and Pearl 2013; Currie 2020; Guala 2003; Howick, Glasziou, and Aronson 2013; Marcellesi 2015; Reiss 2019). There is already a rich

literature on extrapolation in econometrics and in the biomedical sciences that has produced several different approaches to justifying extrapolations and methodological principles of how to deal with remaining uncertainties (Weber 2005; Cartwright and Hardie 2012; Guala 2005, 2010; Khosrowi 2019, 2022, 2023). These are important and interesting cases, and they have considerably advanced our philosophical understanding of how knowledge transfer across systems works and under what conditions it is problematic. In particular, the recent literature has focused on two problems. Firstly, the problem of heterogeneous populations, according to which extrapolation becomes all the more problematic, the more heterogeneous the experimental and target systems are in themselves. Second, the extrapolator's circle, according to which assessing whether a particular extrapolation is justified, may require substantive research on the target system itself, which undermines the epistemic significance of extrapolation (Steel 2008; Khosrowi 2019).

A pivotal problem case of extrapolation that was also instrumental in shaping large parts of the discussion within the philosophy of biology was the use of animal models (Burian 1993; LaFollette and Shanks 1995; Schaffner 2001; Weber 2005). In addition to the question of whether they can serve as an adequate mechanistic analog for their (primarily human) targets, issues of scale pertaining to body size and differences in metabolic rates, among other things, continue to be pervasive problems in this context (Martignoni, Groothuis, and de Kanter 2006; Reagan-Shaw, Nihal, and Ahmad 2008). Biology is ripe with scaling relations that describe how various properties of organisms change with their size (Spence 2009). This aspect of scaling, or extrapolation across scales, is also significant in transferring causal relationships across ecological systems. However, in this case, it relates not only to the size of the organisms but also to the size of their habitats or, more generally, the extent of the area in which they live.

Extrapolation in ecology often includes inferences from small-scale experimental mesocosms to large-scale natural ecosystems because, in many cases, the patterns and processes of interest cannot be directly manipulated at the large, extensive scales at which they need to be understood (Englund and Cooper 2003; Gardner et al. 2001; Miller et al. 2004). This is especially so in cases where large-scale relationships, such as the effects of global warming, biodiversity loss, or environmental pollution, operating over extensive spatial and temporal extents, need to be estimated from smaller-scale experimental systems because of the often unavoidable mismatch between the scales at which ecological processes can be studied and the scales at which they take place in nature (cf. Miller et al. 2004, 310). However, an ecological system studied at small scales may appear considerably different in composition and behavior than the same system studied at larger scales. Patterns and processes characteristic of a system measured at small spatial and temporal scales may not necessarily hold or be relevant drivers in the large-scale system (Wiens 1989; Schneider 2001). While extrapolation can sometimes be

done with the tacit assumption that differences in scale can be ignored without introducing high epistemic risk, this is rarely the case in ecology.

For this reason, I will argue in this paper that we must consider different forms of variability between the experimental and target systems. Whereas our current understanding of extrapolation primarily emphasizes similarity in terms of causal mechanisms and interactive factors, an important aspect to include in assessing extrapolation in ecology is how relevant factors are distributed across space and time. A crucial piece of the puzzle of understanding the challenging nature of extrapolation in ecology thus lies in the spatiotemporal variability of ecological systems. By clarifying the problem of extrapolation in ecology in this way, it is possible to broaden the philosophical debate on this topic and advance our general understanding of the multifaceted problems that arise when transferring knowledge across different systems of interest. My discussion also has significant consequences for the epistemology of ecological experiments and points towards different strategies to deal with the problem of extrapolation in ecological practice.

The argument will proceed as follows: In section 2, I will review and synthesize recent philosophical analyses of extrapolation and its problems. Section 3 argues that heterogeneity between study- and target system can be understood differently and that this difference matters when extrapolation involves scaling relations. To describe this problem, I will further introduce the notion of spatiotemporal variability and contrast it with what I will call compositional variability. This is followed by a discussion of the relation between spatiotemporal variability and scale dependence (section 4). Section 5 discusses the implications of spatiotemporal variability for experimentation and extrapolation in ecology. Section 6 concludes.

## 2. Extrapolation in current philosophy of science

Extrapolation (Steel, 2008), transferability (Spake et al. 2022), or transportability (Bareinboim and Pearl 2013), as it is sometimes also called in the literature, is an inference from a study- to a target system or population (in the statistical sense), typically but not necessarily involving causal claims. Extrapolating inferences may be necessary if direct investigation of the target system is prohibited for ethical or legal reasons or is not feasible due to practical limitations in data collection and performing direct manipulations of the target. Furthermore, the problem of extrapolation is intricately related to the practice of experimentation. Since experimentation is also about testing hypotheses, how well experimental results generalize is central to whether they can fulfill this function. Finally, in cases where knowledge is supposed to guide intervention in and management of real-world systems, the question is not only whether a causal process can be ascertained under

specific experimental conditions but also to what extent this is significant for the behavior of their natural targets. The general structure of an extrapolating inference is thus the transfer of a known causal relationship from a study system or population to a single or class of target systems or populations about which we have no direct knowledge concerning the transferred relation. Extrapolating inferences are thus related to generalization in that they extend a relationship beyond the measured or ascertained domain, and the problem is how to warrant the validity of such inferences.

Most philosophical accounts of extrapolation pick up on this understanding by essentially following a model-target conceptualization of extrapolation as transferring a causal relationship from some source domain to a target domain of interest and by focusing on the question of how such an inference can be epistemically justified (Guala 2005; Steel 2008; Khosrowi 2019; Marcellesi 2015; Reiss 2019). What source and target are or can be is understood differently among accounts. Steel (2008), Bareinboim and Pearl (2013), and Khosrowi (2019) speak of extrapolation as inferences from one population to a different population. Guala (2010) views it as an "attempt to generalize from an epistemically privileged system (experimental mechanism, animal model, etc.) to a less privileged target of interest" (ibid., 1071). Reiss (2019) finally subsumes all these usages under the terms "model system" and "target system(s)" (ibid., 3105).

The question of how such an inference can be justified ultimately revolves around how similar study- and target systems are. The heuristic principle underlying that thought is that similarity between study and target raises the likelihood that similar processes are occurring in the two systems, which in turn, *prima facie*, supports transferability. However, as the philosophical discussion has made clear, similarity between model and target systems is generally neither the only nor the primary ground on which extrapolation is justified. Furthermore, similarity simpliciter may even lead to erroneous conclusions in cases where outward similar factors have different causal effects in different systems (Cartwright 2011; Cartwright and Hardie 2012).

For this reason, philosophical accounts of extrapolation differ in their respective understanding of similarity and the degree to which it must be satisfied. Depending on the specific goal of the extrapolative inference, what must be shown on most accounts is that model and target are similar in at least some relevant respects (Guala 2010), be they similarities of (certain stages of the) causal mechanisms (LaFollette and Shanks 1995; Steel 2008) or of the distribution of covariates and interactive factors (Bareinboim and Pearl 2013; Cartwright and Hardie 2012).

Assessing the similarity of a study and target system can be more or less challenging and epistemically demanding. This pivotal role of similarity is also evident in recent philosophical accounts that have highlighted two problems that come with extrapolation: the problem of heterogeneous populations and the problem of the

extrapolator's circle. According to the first, extrapolating causal effects from one population to another becomes more uncertain when model and target populations are highly heterogeneous. This is because heterogeneity increases the likelihood of causally relevant dissimilarities between model and target such that the transfer of a causal effect from one to the other can become erroneous (Steel 2008). The second problem is the extrapolator's circle (LaFollette and Shanks 1995; Steel 2008; Khosrowi 2019). The point here is that determining the required similarity between model and target may require extensive investigations of the target itself. We need to ascertain that similar causal processes operate in model and target to ensure that the extrapolative inference is reliable. However, if ascertaining similarity already involves extensive research and knowledge of the causal structure of the target itself, this threatens to make extrapolation from the model superfluous.

While a complete account of extrapolation must be able to explain the feasibility of extrapolation given the heterogeneity of study and target and incomplete knowledge of the latter, my aims in this paper are more modest. I will not attempt to give a comprehensive account of extrapolation but instead focus on the problem of heterogeneity of study and target systems. I will argue that by studying extrapolation in ecology, we can add a dimension of heterogeneity that has not received equal attention in the philosophical literature. For this reason, I will save a discussion of the extrapolator's circle in ecological extrapolation for a future paper and instead dive into the problem of heterogeneous study- and target systems for the rest of this paper.

In the next section, I will examine the problem of heterogeneous populations more closely and argue that the current conception of heterogeneity misses crucial elements relevant to understanding extrapolation from ecological experiments. This will show that the problem of heterogeneous populations is vital to extrapolation in ecology, albeit in a different way than currently assumed.

### 3. Compositional and spatiotemporal variability

The problem of heterogeneous populations describes difficulties when extrapolating a causal effect from a model to a distinct target. According to Steel (2008), the fundamental problem is that we should expect causally relevant differences between model and target in heterogeneous populations since causal effects "are prone to vary according to changes in the distribution of factors in the population that affect the outcome" (Steel, 2008: 18). Khosrowi (2019, 46) builds on this understanding and provides a more fine-grained classification of ways in which model and target can be heterogeneous: i) differences in the probability distribution of interactive factors or moderating variables between the two systems; ii) differences in the coefficients that measure the

influence of an independent variable on a dependent variable; and iii) qualitative differences in the mechanisms or causal processes responsible for the observed effect in the model and the target (ibid.).

Both these views have in common that heterogeneity refers to differences in the internal composition of model and target system. Model and target are regarded as sets of entities with specific, measurable properties engaging in certain activities and processes. The causally relevant context in which these entities act and interact is represented by a set of covariates or interacting factors that modify the effect of one on the other. I call this understanding of heterogeneity compositional variability from now on.

Compositional variability explains well the difficulties we encounter in extrapolations where a causal effect is extrapolated from a model to a distinct target but where considerations of the spatial and temporal scales over which model and target are observed do not play a decisive role. However, when extrapolating across scales in ecology, the relevant heterogeneity to consider is not only about the internal composition of study and target system in terms of entities and causally relevant factors but also about their variability over spatial and temporal extent, what I will call spatiotemporal variability in the following. The reason for this has to do with how ecologically relevant factors are distributed across space and time on the one hand and with how well different scales of observation are suited to capture the resulting spatial structure. Let me start with the first point about the distribution of relevant factors across space and time, returning to the second point in section 4.

A simplifying assumption for studying ecosystems is that a system measured at local scales is similar to the same system measured at regional scales and, likewise, that patterns and processes observed at local scales are similar at regional or even larger scales.<sup>1</sup> According to this assumption, the observed systems are relatively constant over spatial and temporal extent. Study and target systems would not differ in composition when observed on different measurement scales, and for this reason, we were justified in assuming that causal relationships would persist with the same drivers responsible for them. Under this assumption, compositional variability would be sufficient to identify the criteria for justifying extrapolative inferences from ecological experiments. However, in ecological systems, there is little support for this assumption.

---

<sup>1</sup> While spatiotemporal variability principally encompasses both spatial and temporal scales, I will limit myself to the spatial dimension in the following explanation of spatiotemporal variability. I will do this not only to simplify the discussion but also because it presents additional and interesting problems relative to variation in the wake of temporal dynamics.

One of the many aspects of the complexity of ecological systems is their nonstationarity, meaning that statistical parameters like the mean and variance of a variable of interest vary across space and time (Rollinson et al. 2021; Brian and Catford 2023). This naturally has consequences for drawing conclusions about ecological processes that can only be observed over limited spatial and temporal scales. "Nonstationarity in this context is a case where conclusions drawn from a single location or point in time are typically insufficient for explaining large-scale patterns because they only provide glimpses into broad ecological processes that occur over a wider range of conditions" (Rollinson et al. 2021, 66). This means that as we move across space and time in ecological systems, we will likely encounter "a wide range of conditions" which we can understand as variability not only in the composition but also the spatial structure of ecological systems.

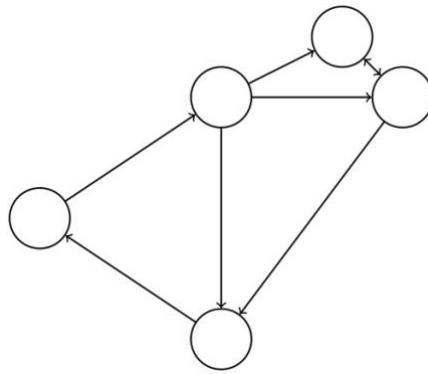
Ecologists have long been aware that the spatial structure of ecological systems matters. As early as 1966, Robert MacArthur and Eric Pianka (1966) incorporated the assumption of a patchy environment into their optimal foraging theory. According to this model, the environment, from the perspective of a particular organism, is structured as a heterogeneous spatial mosaic of favorable and unfavorable patches distributed across space. This distribution, in turn, necessitates different foraging strategies to optimize resource intake while simultaneously minimizing the time spent scrambling through the heterogeneous, patchy environment.

A similar conception of the spatial structuring of ecological systems underlies the use of meta-population models in ecology and conservation biology (Hanski and Simberloff 1997). In the meta-population approach, a population is modeled as a spatially structured set of sub-populations that inhabit favorable habitat patches (favorable from the perspective of the organism) while avoiding the unfavorable patches or areas in between (see Fig. 1). Patches can be of different sizes and different habitat-quality. Migration among patches connects the local sub-populations and partially accounts for the meta-population dynamics.

Letting  $P$  stand for the proportion of occupied patches, and  $c$  and  $e$  stand for colonization and extinction rates, respectively, the basic form of a meta-population model would be:

$$\frac{dP}{dt} = c(1 - P) - eP$$

(Odenbaugh 2019). These models implicitly incorporate spatial structure into the dynamics of the population without explicitly representing space in terms of location or distance. They express that the environment is heterogeneous, that this heterogeneity has a spatial structure, and that this structure, in turn, plays a role in how biotic entities are distributed, moving, and interacting across space.



*Figure 1: Spatial structure of a meta-population. Nodes represent patches of local sub-populations, while arrows indicate migration patterns.*

While consideration of the spatial structure of the environment and its role in ecological processes has thus arguably always been a part of ecological theorizing, further developments in spatial- and landscape ecology have added considerably to our understanding of how spatial variation needs to be accounted for in both the statistical analysis and the conceptual understanding of processes in ecological systems (Tilman and Kareiva 1998; Fletcher and Fortin 2018).

To make that more precise, Wagner and Fortin (2005) distinguish two types of spatial heterogeneity: The first is environmental heterogeneity, which is variation in the distribution of environmental factors that is itself spatially structured. The second is the "patchiness of species and other ecological response variables" (2005, 1975). While the first describes the variability of environmental conditions in space, the second describes patterns that may be due to species' responses to these spatially structured environmental conditions, given their specific resource demands and/or the result of biotic interactions and other ecological processes.

The most obvious examples of these types of heterogeneity is species distributions in geographic space. Organisms of a population can be spatially distributed according to different patterns within a landscape. Theoretically, their distribution might be random, but in reality, it will often be non-random, i.e., either regular or clustered in a particular way. However, the resulting pattern will depend on how favorable environmental conditions are distributed across space and how the suitable habitat patches are connected. This spatial configuration can then, in turn, influence ecological processes such as predator-prey interactions with further potential downstream effects for community structure and dynamics (Borregaard, Hendrichsen, and Nachman 2008). As Borregaard et al. (2008, 3304) point out, "patterns of spatial distribution play an important role in shaping a wide range of ecological dynamics, such as intra- and interspecific competition, mating systems, predation, population genetics, and the spread of contagious diseases."



However, to the extent that spatial patterns play an important role in ecological processes, it becomes a relevant question to what extent these processes can be extrapolated to different scales, where spatial patterns will likely be very different than in the study system even though the system at larger scales might still be relatively similar in terms of composition. In this case, more than compositional variability is needed to approach this problem because it is not only variation in the composition of study and target system but also variation in the spatial distribution of environmental factors and ecological entities that matter.

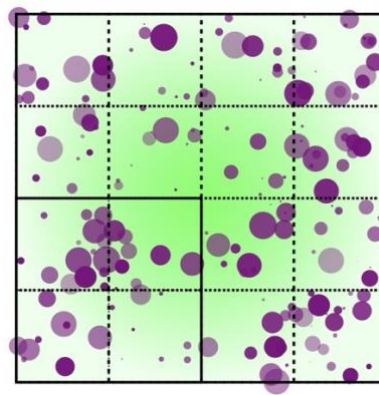
Based on these considerations, I suggest distinguishing compositional variability in the sense defined above from spatiotemporal variability. Spatiotemporal variability is the variation in the distribution of entities or factors of interest over spatial and/or temporal extent reflecting the spatial structure and/or the temporal dynamics within an observed area.

Spatiotemporal variability is heterogeneity that has an irreducible spatial or temporal component. The difference between compositional and spatiotemporal variability is thus related to the difference between static and dynamic perspectives on extrapolation. Under the assumptions of temporal stability and irrelevance of spatial structure, compositional variability is a suitable heuristic for approaching extrapolation. However, in cases where these simplifying assumptions are not met, consideration of spatiotemporal variability becomes an important aspect of justifying extrapolation. To put it in another way, while the variation between two ecosystems is always variation in space and time since ecosystems exist in space and time, spatiotemporal variability is not necessarily always relevant. Depending on the objective of a study, a static perspective, i.e., one that focuses on similarities and differences between ecosystems without an explicit spatial component, may be appropriate for the given purpose. However, in some cases where processes are strongly spatially dependent or differences in scale between study and target systems are significant, the static perspective is not appropriate, and spatial or temporal variation must be considered when assessing transferability to reduce uncertainty and risk of error.

Nevertheless, spatiotemporal variability is generally important for the philosophical justification of extrapolation because it shows that study and target system can be similar or dissimilar in more than one sense, namely, in terms of their composition but also in the way entities of interest and relevant environmental conditions are distributed and structured over spatial and temporal extent which in turn can have consequences for causal inference and subsequently for extrapolation.

#### 4. The relation between variability and scales of measurement

Since ecological systems can vary across spatial extent, as we have just seen, a system measured at local scales can show considerably different levels of variability if compared to the same system studied at more extensive landscape scales. Not only are ecological systems variable across space and time but how well we can capture this variability depends on the scales of observation. Ecologists typically understand scale as encompassing both the grain and extent of observation (Wiens 1989; Schneider 2001; Englund and Cooper 2003). Englund and Cooper (2003) define scale more formally as “the grain, which usually refers to the area or time span covered by individual samples, and the extent, which refers to the total range in time or space over which samples are distributed” (Englund and Cooper 2003, 163).



*Figure 2: The relationship between the scales of observation and spatiotemporal variability of an ecological system. The outer circumference defines the extent of observation of the system, i.e., the total area covered. The small squares with dashed lines represent experimental plots that are either left untreated for controls or on which treatments are carried out. The smaller square with solid lines at the bottom left represents a small-scale measurement of the same system that differs in extent but not grain size. The shades of green represent a gradient of environmental factors (increasing towards the center), while the differently shaded and sized dots represent different (biotic) entities and their distribution in space. Together, the environmental gradient and the distribution of entities make up the system's spatial structure.*

The extent of an experiment is thus the total area it covers. At the same time, the grain describes the size of the individual samples, usually called “plots” in terrestrial systems, observed within the total extent (Fig. 2). Plots are the individual units of experimentation to which treatments are administered (save the control plots). Therefore, the relation between grain and extent can be considered the resolution at which the experiment tracks patterns and processes occurring within the study area (cf. Wiens 1989, 387). Suppose we enlarge the total extent. In that case, we will increase the number of different features of the total area we are investigating and

likely increase spatial variability because larger areas will cover a more extensive range of different environmental conditions.

Suppose we change the grain of the experimental plots. In that case, we will be more or less sensitive to these differences within the studied area (see Fig.2). In general, large grain size will lead to small differences and details within the plots being lost to our view because they get averaged out. In contrast, small grain size allows us more differentiation and covers more detail but also risks missing large-scale patterns of the system (Wiens 1989). Therefore, the spatial resolution determines the observed variability of the system under investigation and, in this way, also our chances to detect ecological patterns (cf. Englund and Cooper 2003, 174-175).

In ecology, this is known as scale dependence. Scale dependence means that ecological patterns and processes observed at small spatial scales do not necessarily translate to larger scales of observation and are thus dependent on the spatial extent (and temporal duration) of the observation or experiment (Kemp, Petersen, and Gardner 2001; Pace 2001; Schneider 2001; Englund and Cooper 2003). The realization of scale-dependent processes arguably brought with it a shift in general theoretical assumptions underlying experimentation in ecology:

“In the 1990s ecologists began to address larger scales again, largely in response to questions about conservation that concern large scales. Initially, the unspoken assumption was that this would conform to the expectation that experiments done at that small scale would provide an understanding of large-scale phenomena. But there was a growing realization that the processes important at larger scales are often fundamentally different from those observed at lower scales.” (Potochnik and McGill 2012, 134; see Mikkelsen 2007 for a similar argument).

Scale dependence can mean several things. It can imply that the spatial distributions of entities in an ecosystem might have different structures depending on the scales of observation. For example, while a population may appear clustered at local scales (bottom left section of Fig. 2), it could appear more evenly distributed at the landscape level. Consequently, processes observed in an experimental system at a particular scale do not (necessarily) translate across spatial or temporal scales to the more extensive natural target systems. Lastly, and most importantly, the relative importance of different drivers of observed causal processes might also change with scale. For example, McGill (2010) lists four processes that have been proposed to determine the spatial distribution of species: (i) random chance dispersal, (ii) habitat preferences, (iii) species interactions, and (iv) climate. However, the relative importance of each process varies with the scale of observation. For instance, while climate is an important driver at small scales, competition is more important at intermediate scales. At the largest biogeographic scales, however, climatic factors become the primary driver again. As McGill points out,

this means that the question of which of the four hypothesized drivers is the most important one is ill-posed. Instead, the correct question would be, which factor drives species distribution at which scale? This also implies that the relevance of specific drivers of species distributions at local scales cannot generally be extrapolated to elucidate the relationship at larger, regional scales.

This points to another important distinction I want to make, namely the one between scaling in the narrower sense and what I have called extrapolation across spatiotemporal scales at the beginning of this paper. So far, I have not formally differentiated between scaling and extrapolation across scales as modes of inference; however, as I will argue here, they are not necessarily identical and should be distinguished.

As already pointed out, scaling relations are ubiquitous in biology and ecology. A pertinent example within the latter is the species-area relationship (SAR). Ecologists have long been aware that levels of species richness increase with the spatial extent of the observed area and that this relation can be approximated using a power function of the form  $S = cAz$ , where  $S$  is the number of species,  $A$  the size of the area and  $c$  and  $z$  are parameters, estimated from empirical data (Harte et al. 1999). We can use this function to predict species richness levels according to the area's size. We can also use it to upscale the relationship between species richness and area size from small to larger scales, thereby calculating the expected saturating shape of the curve. However, as we have just seen, estimating the expected shape of the species-area curve is different from predicting which mechanisms are responsible for certain levels of species richness at small and large spatial scales because different processes drive species distribution at different spatial scales. For this reason, finding a scaling function that approximates the shape of the relationship through different scales is different from trying to extrapolate the mechanism responsible for that relationship from small to larger scales, and to capture this difference, I suggest distinguishing between scaling and extrapolation across scales.

So far, I have pointed out that ecological systems not only differ in their composition but also in the way that these differences are distributed across space and time and that because of this spatiotemporal variability, how similar or dissimilar study and target systems appear to us also depends on the scale of observation. I have further pointed out that patterns and processes in ecological systems are often scale-dependent, meaning that they are observable on some scales but not on others and that the relevance of specific factors as drivers for ecological processes differs across scales.

As should have become clear by now, spatiotemporal variability has two distinct aspects: The first is about the spatial and temporal distribution of entities and environmental factors and the resulting spatial structure or temporal dynamics that can influence ecological processes at specific scales (but not necessarily across smaller or larger scales). The second is about how we can measure and represent this variability in experimental

and observational studies. While the first is ontological in that ecological systems are composed of different entities that inhabit different environments and in that both are distributed according to specific patterns across space and time, the second aspect is epistemological in that it is about how much knowledge of an ecological system and the processes therein we can get by measurements performed at different spatial scales that only provide point estimates of a particular section of a more extensive and essentially open ecological system. I will further discuss this epistemic aspect in the following section.

## 5. The epistemology of spatial variation and scale in ecological experiments

In section 2 of this paper, I emphasized that justifying extrapolation requires identifying relevant similarities between study- and target system. I have also argued that in ecosystems, these similarities often encompass irreducible spatial and temporal dimensions. In this section, I integrate this argument with an analysis of the epistemological underpinnings of experimentation in ecology and how different experimental approaches are suited to address spatiotemporal variability.

Ecologists use various experimental techniques to study ecological phenomena, ranging from laboratory- and mesocosm- to field experiments that differ in how they approach the trade-off between external and internal validity (Naeem 2001). Laboratory or bottle experiments test ecological interactions under controlled conditions using artificial and simplified environments. Field experiments are on the opposite side of the spectrum in studying ecological interactions under realistic conditions in the natural environment. In contrast, mesocosm experiments are located somewhere in between these two endpoints. They simulate field conditions to some extent but also create semi-artificial environments by using tanks, enclosures, and other containers to stabilize environmental conditions and keep out external influences (*ibid.*). Laboratory and bottle experiments in ecology were often accused of being too unrealistic and, therefore, of not being able to provide results that can be extrapolated to populations in natural environments, which at times has led to heated methodological debates (Carpenter 1996; Diamond 1983, 1986; Drenner and Mazumder 1999). Against this background, field experiments were supposed to solve this problem by providing a method of testing hypotheses in the natural or at least semi-natural environment (Englund and Cooper 2003; Hairston 1989; Tilman 1989).

In the following, I will argue that this juxtaposition of non-extrapolable laboratory and extrapolable field experiments is misleading. First, building on arguments by Currie (2020) and Odenbaugh (2006), we can construe the significance of laboratory experiments in ecology differently by dropping the requirement that extrapolation is their primary goal. Second, the assumption that results from field or mesocosm experiments can

be extrapolated to natural systems has been controversial among ecologists themselves, not least because of frequent unexpected variation of ecological relationships between seemingly similar locations (Kemp, Petersen, and Gardner 2001; Price and Billick 2010; Pulliam and Waser 2010; Resetarits and Fauth 1998; Schindler 1998; Tilman 1989; Underwood 1986). Given compositional and especially spatiotemporal variability, the use of field- and mesocosm experiments is therefore far from solving the problem of extrapolation but instead requires justification itself.

Let me begin with the first point about laboratory experiments. Currie (2020) has recently argued that the failure of laboratory experiments to illuminate causal relationships in natural systems and, thus, their low degree of external validity should not lead us to discount them wholesale. Instead, he suggests that these experiments have an additional epistemic value in enhancing our general understanding of ecological processes, specifically of what he calls the “space of ecological possibility” (ibid., 913). Second, laboratory experiments are often used not to test hypotheses derived from field observations of natural systems in the first place but to test hypotheses derived from mathematical models. Accordingly, in these cases, the experimental design is not supposed to represent real ecosystems but to realize the conditions the model specifies (Naeem, 2001). Odenbaugh (2006) takes a complementary view of the role of laboratory experiments in that they can, in turn, provide “evidential constraints” to model building (Odenbaugh 2006, 720).

It is also essential to avoid taking the concept of extrapolation for generality wholesale. In a sense, laboratory experiments are also highly general, albeit only for artificial systems similar to those constructed in the experimental setup. What they achieve is generality in the sense that simple mathematical models are highly general (Elliott-Graves 2022). Thus, what distinguishes laboratory from field experiments is that laboratory experiments prioritize generality without the implication of being representative of particular real-world systems. Consequently, for our discussion of extrapolation, we might be justified to exclude laboratory experiments because extrapolation is not their primary goal in the first place, quite unlike mesocosm and field experiments, where extrapolation can very well be seen as a primary goal (Pace 2001).

Coming then to the second argument of this section, let us see why field experiments alone do not bridge the gap between the laboratory and the real world. The logic underlying the use of field experiments is that greater similarity between study and target system will generally reduce the uncertainty involved in extrapolation.

However, the problem and the crucial point of my argument is that just like variability, so too can similarity be understood in two different senses: On the one hand, similarity can be understood as matching or corresponding properties of study- and target system in terms of entities and environmental factors. On the other

hand, as I have argued throughout this paper, similarity can also be understood in terms of spatial and temporal structure, i.e., the distribution of environmental factors and biotic entities across space and time.

However, maximizing similarity in one sense does not necessarily imply maximizing it in the other. Suppose we aim to extrapolate a causal effect found in one particular system to a different system at roughly similar spatial and temporal scales. In that case, similarity in the first sense is undoubtedly essential and may already be sufficient for describing the general conditions under which such an inference may be justified. However, the situation becomes more complicated if we aim to extrapolate from small- to large-scale systems. In this case, the relevant similarity is not exhausted by matching the properties of the study- and target system without paying attention to whether the scale at which the study system is observed allows for the inference to the spatial and temporal dynamics of the larger scale target system.

Ecologists have suggested different strategies to cope with this uncertainty. The first and most obvious one is to avoid extrapolation if it is unnecessary. For example, Englund and Cooper (2003) suggest this strategy as a remedy for scale dependence, which is doable by “matching the physical size of experimental units with the size of the system of interest or by designing small-scale experimental systems so that the processes of interest are given a realistic representation” (Englund and Cooper, 2003: 162). If extrapolation cannot be avoided, it may be possible to directly up-scale an observed relationship from the small-scale study system to the larger scales of the natural system by calculating the proportionate increase. Of course, this only works under the assumption that “the relationship of a variable to changes in scale is linear or additive” (Miller et al., 2004: 314). Scaling relationships between ecological variables from small to large is more complicated if the scaling function is non-linear or non-continuous, i.e., when it exhibits breaks, jumps, or other thresholds (Miller et al. 2004; Wiens 1989, 2001). For example, Wiens (2001) shows that even the already mentioned species-area relationship exhibits thresholds at some points, in which case up-scaling could easily lead to aggregation error and mistaken extrapolations.

If avoiding extrapolation and direct upscaling is impossible, ecologists can develop new research strategies and experimental designs sensitive to spatiotemporal variability and its effects. Instead of relying on single studies of randomly chosen locations, experimenters could pursue designs that incorporate variability by studying interactions at multiple scales, locations, and along significant environmental gradients (Hewitt et al. 2007; Snelgrove et al. 2014). In all these cases, however, understanding the problems and limits of extrapolation requires consideration of spatiotemporal variability.

## 6. Conclusion

In this paper, I have argued that extrapolation from ecological experiments is an interesting case for philosophical analyses of knowledge transfer because it presents a variety of relatively specific problems that have hitherto not been considered as much. In particular, I have argued that the problem of heterogeneous populations is incomplete if only considered along the lines of compositional variability. Likewise, philosophical analyses of extrapolation fall short when only referring to differences in composition without considering how compositional variability is distributed and structured over space and time, i.e., without considering spatiotemporal variability. In this sense, spatiotemporal variability does not replace compositional variability but complements it. Spatiotemporal variability is another way in which study- and target systems in an extrapolative inference can be heterogeneous.

Finally, in referring to arguments advanced by Currie and Odenbaugh, I have pointed out that different experiments in ecology can serve different epistemic functions and that extrapolation is not necessarily one of them. However, I have also claimed that even in experimental practices where extrapolation is a relevant epistemic function, this is not always straightforward and presents epistemic challenges to ecologists. While this is not sufficient to discard extrapolation at all, general limits to its application need to be understood.

#### Acknowledgements

I want to thank two anonymous reviewers for helpful comments on an earlier draft of this paper. I am also very grateful to Alkistis Elliott-Graves, Marie Kaiser, Phyllis Illari, Anton Killin, Hanna Metzen, Alexander Linsbichler, and the Philosophy of Life Sciences group at Bielefeld University for their generous feedback.

#### Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project 254954344/GRK2073.

#### Competing Interests

The author has no competing interests to declare that are relevant to the content of this article.

#### References



- Baetu T 2016. “The ‘Big Picture’: The Problem of Extrapolation in Basic Research.” *Br J Philos Sci* 67 (4): 941–964. <https://doi.org/10.1093/bjps/axv018>.
- Bareinboim E, Pearl J 2013. “A General Algorithm for Deciding Transportability of Experimental Results.” *J. Causal Inference* 1 (1): 107–134. <https://doi.org/10.1515/jci-2012-0004>.
- Borregaard, M. K., Hendrichsen D, Nachman G 2008. “Spatial Distribution.” In *Encyclopedia of Ecology*, 1st ed., edited by Sven Erik Jørgensen and Brian D. Fath, 3304–3310. Amsterdam: Elsevier. isbn: 978-0-444-52033-3.
- Brian, J, Catford J 2023. “Ecological Scale and Context Dependence.” In *Effective Ecology. Seeking Success in a Hard Science*, edited by Roger D. Cousens, 63–79.
- Burian R 1993. “How the Choice of Experimental Organism Matters: Epistemological Reflections on an Aspect of Biological Practice.” *J Hist Biol* 26 (2): 351–367. <https://doi.org/10.1007/BF01061974>.
- Carpenter S 1996. “Microcosm Experiments Have Limited Relevance for Community and Ecosystem Ecology.” *Ecology* 77 (3): 677–680. <https://doi.org/10.2307/2265490>.
- Cartwright N 2011. “Predicting ‘It Will Work for Us’: (Way) beyond Statistics.” In *Causality in the Sciences*, edited by Phyllis McKay Illari, Federica Russo, and Jon Williamson, 750–768. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199574131.003.0035>.
- Cartwright N, Hardie J 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.
- Currie A 2020. “Bottled Understanding: The Role of Lab Work in Ecology.” *Br J Philos Sci* 71 (3): 905–932. <https://doi.org/10.1093/bjps/axy047>.
- Diamond J 1983. “Ecology: Laboratory, Field and Natural Experiments.” *Nature* 304 (5927): 586–587. <https://doi.org/10.1038/304586a0>.
- Diamond J 1986. “Overview: Laboratory Experiments, Field Experiments, and Natural Experiments.” In *Community Ecology*, edited by Jared M. Diamond and Ted J. Case, 3–22. New York: Harper & Row. isbn: 978-0-06-041202-9.
- Drenner R, Mazumder A 1999. “Microcosm Experiments Have Limited Relevance for Community and Ecosystem Ecology: Comment.” *Ecology* 80 (3): 1081–1085. Accessed November 30, 2023. <https://doi.org/10.2307/177042>.
- Elliott-Graves A 2022. “What Are General Models About?” *Euro J Phil Sci* 12, no. 4 (December): 74. <https://doi.org/10.1007/s13194-022-00502-9>.

- Englund G, Cooper S. 2003. "Scale Effects and Extrapolation in Ecological Experiments." *Adv. Ecol. Res.* 33:161–213. Accessed May 25, 2023. [https://doi.org/10.1016/S0065-2504\(03\)33011-9](https://doi.org/10.1016/S0065-2504(03)33011-9).
- Fletcher R, Fortin M 2018. *Spatial Ecology and Conservation Modeling*. Cham: Springer International Publishing, January. isbn: 978-3-030-01988-4.
- Gardner R, Kemp M, Kennedy V, Petersen J, eds. 2001. *Scaling Relations in Experimental Ecology*. New York Chichester, West Sussex: Columbia Univ. Press. Accessed October 11, 2022. <https://doi.org/10.7312/gard11498>.
- Guala F 2003. "Experimental Localism and External Validity." *Philos Sci* 70 (5): 1195–1205. Accessed October 11, 2022. <https://doi.org/10.1086/377400>.
- Guala F 2005. *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511614651>.
- Guala F 2010. "Extrapolation, Analogy, and Comparative Process Tracing." *Philos of Sci* 77 (5): 1070–1082. <https://doi.org/10.1086/656541>.
- Hairston N 1989. *Ecological Experiments: Purpose, Design and Execution*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511608513>.
- Hanski I, Simberloff D 1997. "The Metapopulation Approach, Its History, Conceptual Domain, and Application to Conservation." In *Metapopulation Biology: Ecology, Genetics, and Evolution*, edited by Ilkka Hanski and Michael E. Gilpin, 5–26. San Diego, Calif.: Academic Press. isbn: 978-0-12-323445-2.
- Harte J, McCarthy S, Taylor K, Kinzig A, Fischer M 1999. "Estimating Species-Area Relationships from Plot to Landscape Scale Using Species Spatial-Turnover Data." *Oikos* 86 (1): 45–54. 2023. <https://doi.org/10.2307/3546568>.
- Hewitt, Judi ., Simon F. Thrush, Paul K. Dayton, and Erik Bonsdorff. 2007. "The Effect of Spatial and Temporal Heterogeneity on the Design and Analysis of Empirical Studies of Scale-Dependent Systems." *Am Nat* 169 (3): 398–408. 2023. <https://doi.org/10.1086/510925>.
- Howick J, Glasziou P, Aronson J 2013. "Problems with Using Mechanisms to Solve the Problem of Extrapolation." *Theor Med Bioeth* 34 (4): 275–291. <https://doi.org/10.1007/s11017-013-9266-0>.
- Kemp M, Petersen J, Gardner R 2001. "Scale-Dependence and the Problem of Extrapolation." In *Scaling Relations in Experimental Ecology*: edited by Robert H. Gardner, W. Michael. Kemp, Victor S. Kennedy, and John E. Petersen, 3–57. New York Chichester, West Sussex: Columbia University Press. isbn: 978-0-231-11499-8. <https://doi.org/10.7312/gard11498>.

- Khosrowi D 2019. “Extrapolation of Causal Effects – Hopes, Assumptions, and the Extrapolator’s Circle.” *J Econ Methodol* 26 (1): 45–58. <https://doi.org/10.1080/1350178X.2018.1561078>.
- Khosrowi D 2022. “What’s (Successful) Extrapolation?” *Journal of Economic Methodology* 29 (2): 140–152. <https://doi.org/10.1080/1350178X.2021.1952290>. 2023.
- Khosrowi D “Extrapolating from Experiments, Confidently.” *Euro Jnl Phil Sci* 13 (2): 18. Accessed October 6, 2023. <https://doi.org/10.1007/s13194-023-00520-1>.
- LaFollette H, Shanks N 1995. “Two Models of Models in Biomedical Research.” *Philos Q* 45 (179): 141–160. <https://doi.org/10.2307/2220412>.
- MacArthur R, Pianka E 1966. “On Optimal Use of a Patchy Environment.” *Am Nat* 100 (916): 603–609.
- Marcellesi A 2015. “External Validity: Is There Still a Problem?” *Philos Sci* 82 (5): 1308–1317. <https://doi.org/10.1086/684084>.
- Martignoni M, Groothuis G, de Kanter R 2006. “Species Differences between Mouse, Rat, Dog, Monkey and Human CYP-mediated Drug Metabolism, Inhibition and Induction.” *Expert Opin Drug Metab Toxicol* 2 (6): 875–894. <https://doi.org/10.1517/17425255.2.6.875>.
- McGill, B 2010. “Matters of Scale.” *Science* 328 (5978): 575–576. <https://doi.org/10.1126/science.1188528>.
- Mikkelsen, G 2007. “Ecology.” In *The Cambridge Companion to the Philosophy of Biology*, edited by David L. Hull and Michael Ruse, 372–387. Cambridge: Cambridge University Press. isbn: 978-0-521-61671-3.
- Miller J, Turner M, Smithwick E, Dent L, Stanley E (2004) *Spatial Extrapolation: The Science of Predicting Ecological Patterns and Processes*. *BioScience* 54:310–320. [https://doi.org/10.1641/0006-3568\(2004\)054\[0310:SETSOP\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0310:SETSOP]2.0.CO;2)
- Naeem, S 2001. “Experimental Validity and Ecological Scale as Criteria for Evaluating Research Programs.” In *Scaling Relations in Experimental Ecology*, edited by Robert H. Gardner, W. Michael. Kemp, Victor S. Kennedy, and John E. Petersen, 223–250. Columbia University Press. <https://doi.org/10.7312/gard11498-011>.
- Odenbaugh, J 2006. “Message in the Bottle: The Constraints of Experimentation on Model Building.” *Philos Sci* 73 (5): 720–729. <https://doi.org/10.1086/518773>.
- Odenbaugh J 2019. *Ecological Models*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108685283>.
- Pace, M 2001. “Getting It Right and Wrong. Extrapolations Across Experimental Scales.” In *Scaling Relations in Experimental Ecology*: edited by

- Potochnik A, McGill B 2012. “The Limitations of Hierarchical Organization” *Philos Sci* 79 (1): 120–140.  
<https://doi.org/10.1086/663237>.
- Reagan-Shaw S, Minakshi N, Ahmad N 2008. “Dose Translation from Animal to Human Studies Revisited.”  
*FASEB J* 22 (3): 659–661. <https://doi.org/10.1096/fj.07-9574LSF>.
- Reiss, J 2019. “Against External Validity.” *Synthese* 196 (8): 3103–3121. <https://doi.org/10.1007/s11229-018-1796-6>.
- Rollinson, C O Finley A, Alexander R, Banerjee S, Dixon Hamil K, Koenig L, Locke D, et al. 2021. “Working across Space and Time: Nonstationarity in Ecological Research and Application.” *Frontiers in Ecol & Environ* 19 (1): 66–72. <https://doi.org/10.1002/fee.2298>.
- Schaffner K 2001. “Extrapolation from Animal Models. Social Life, Sex, and Super Models.” In *Theory and Method in the Neurosciences*, edited by Peter K. Machamer, Rick Grush, and Peter McLaughlin, 200–230. Pittsburgh, Pa: University of Pittsburgh Press.
- Schneider D 2001. “The Rise of the Concept of Scale in Ecology.” *BioSci* 51 (7): 545–553.  
[https://doi.org/10.1641/0006-3568\(2001\)051\[0545:TROTCO\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0545:TROTCO]2.0.CO;2).
- Snelgrove P, Thrush S, Wall D, Norkko A 2014. “Real World Biodiversity–Ecosystem Functioning: A Seafloor Perspective.” *Trends Ecol Evol* 29 (7): 398–405. <https://doi.org/10.1016/j.tree.2014.05.002>.
- Spake R, O’Dea R, Nakagawa S, Doncaster P, Ryo M, Callaghan C, Bullock J. 2022. “Improving Quantitative Synthesis to Achieve Generality in Ecology.” *Nat Ecol Evol* 6:1818–1828.  
<https://doi.org/10.1038/s41559-022-01891-z>.
- Spence A 2009. “Scaling in Biology.” *Curr Biol* 19 (2): R57–61. <https://doi.org/10.1016/j.cub.2008.10.042>.
- Steel D 2008. *Across the Boundaries: Extrapolation in Biology and Social Science*. Oxford ; New York: Oxford University Press.
- Tilman D. 1989. “Ecological Experimentation: Strengths and Conceptual Problems.” In *Long-Term Studies in Ecology: Approaches and Alternatives*, edited by Gene E. Likens, 136–157. New York, NY: Springer.  
[https://doi.org/10.1007/978-1-4615-7358-6\\_6](https://doi.org/10.1007/978-1-4615-7358-6_6).
- Tilman D, Peter K 1998. *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Edited by Peter Kareiva and David Tilman. Vol. 30. *Monographs in Population Biology*. Princeton, NJ: Princeton University Press. <https://doi.org/10.23943/9780691188362>.
- Wagner H., Marie-Josée F. 2005. “Spatial Analysis of Landscapes: Concepts and Statistics.” *Ecology* 86 (8): 1975–1987. <https://doi.org/10.1890/04-0914>.

Weber M. 2005. *Philosophy of Experimental Biology*. Cambridge: Cambridge University Press.

<https://doi.org/10.1017/CBO9780511498596>.

Wiens J 1989. "Spatial Scaling in Ecology." *Funct Ecol* 3 (4): 385–397. <https://doi.org/10.2307/2389612>.

Wiens J 2001. "Understanding the Problem of Scale in Experimental Ecology." In *Scaling Relations in Experimental Ecology*: edited by Robert H. Gardner, W. Michael. Kemp, Victor S. Kennedy, and John E. Petersen, 61–88. New York Chichester, West Sussex: Columbia University Press.

<https://doi.org/10.7312/gard11498>.