



Taming the tyranny of scales: models and scale in the geosciences

Alisa Bokulich¹ 

Received: 25 January 2021 / Accepted: 9 September 2021 / Published online: 1 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

While the predominant focus of the philosophical literature on scientific modeling has been on single-scale models, most systems in nature exhibit complex multiscale behavior, requiring new modeling methods. This challenge of modeling phenomena across a vast range of spatial and temporal scales has been called the *tyranny of scales* problem. Drawing on research in the geosciences, I synthesize and analyze a number of strategies for taming this tyranny in the context of conceptual, physical, and mathematical modeling. This includes several strategies that can be deployed in physical (table-top) modeling, even when strict dynamical scaling fails. In all cases, I argue that having an adequate conceptual model—given both the nature of the system and the particular purpose of the model—is essential. I draw a distinction between depiction and representation, and use this research in the geosciences to advance a number of debates in the philosophy of modeling.

Keywords Multiscale · Modeling · Simulation · Conceptual model · Physical model · Mathematical model · Scale · Dynamical scaling · Geosciences · Geomorphology · Geohydrology · Representation · Thresholds · Adequacy for purpose · Hierarchy · Similitude · Reductionism · Emergence · Universality · Tyranny of scales

1 Introduction: the tyranny of scales

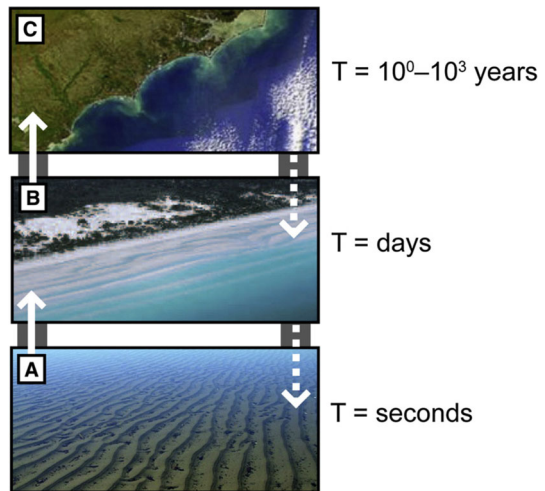
The tyranny of scales problem is the recognition that many phenomena of interest span a wide range of spatial and temporal scales, where the dominant features and physical processes operating at any one scale are different from those operating at both smaller (shorter) and larger (longer) scales. This physical fact then poses the following

This article belongs to the topical collection “Multiscale Modeling and Active Materials”, edited by Patrick McGivern and Nick Brancazio.

✉ Alisa Bokulich
abokulich@bu.edu

¹ Department of Philosophy, Boston University, Boston, MA 02215, USA

Fig. 1 Evolution of a sandy coastline involves a vast range of spatial and temporal scales. Ripples (A) that form from sand-grain motions on the order of seconds influence the formation of sandbars and channels (B) on the order of days, which over time adds up to a net alongshore movement of sand that shapes coastline evolution (C) on the scale of years to millennia. The larger scales then set the context in which the smaller-scale dynamics must operate (dashed arrows). (From Murray et al., 2009, Fig. 4, with permission from Elsevier.)



methodological problem: How does one go about modeling such a phenomenon of interest, especially when that phenomenon can be causally influenced by—and in turn influence—the entities and processes at the scales both above and below it?

Consider, for example, the evolution of a sandy coastline. As those who live by the coast are often painfully aware, coastlines are not static, but rather are continually changing—eroding in some areas and accreting in others. There is thus great scientific and practical interest in better understanding how coastlines dynamically evolve. But how should one go about modeling the evolution of a sandy coastline? If one examines the problem at the smallest (fastest) scale, sand grains respond to waves on the timescale of seconds; this leads to small bedforms like ripples over hours. These ripples then influence the development of sandbars and channels in the surf zone, which form on the timescale of days. Channels and sandbars influence the movements of sediment over the timescale of weeks, and the net transport of sediment along the shoreline shapes coastlines over years to millennia.¹ To further complicate the situation, the shape of the coastline sets the context and causally influences the dynamics of sandbar and channel formation; and those channels and sandbars then set the context and causally influence the formation of ripples (see Fig. 1). In short, attempts to model and understand coastline evolution are plagued by the tyranny of scales.

The tyranny of scales problem, though widespread across the sciences, has only recently begun to attract scientific and philosophical attention. The scientific document often credited with naming the problem was a 2006 National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science, chaired by J. Tinsley Oden. In this report, the panel notes the following:

¹ I am grateful to Brad Murray (personal communication) for this example.

[Researchers] have run into a formidable roadblock: *the tyranny of scales*. . . . [C]onventional methods . . . cannot cope with physical phenomena operating across large range of scale [e.g., 10 orders of magnitude]. . . . Confounding matters further, the principle physics governing events often changes with scale, so that the models themselves must change in structure. (Oden, 2006, pp. 29–30; emphasis added)

As the panel goes on to note, these problems, which bedevil scientific research spanning from atomic physics to astrophysics, require that new modeling approaches be developed.

Although the tyranny of scales problem is often taken to be intuitively clear from the description provided by Oden et al. above, it is helpful to explicitly draw out three components that typically characterize the tyranny of scales problem: First, it concerns phenomena that span a wide range of spatial and/or temporal scales. Second, these phenomena involve fundamentally different entities and processes operating at different scales. In other words, the effective dynamics (or “principle physics” to use Oden’s phrase) dominating at one scale will be different from the effective dynamics that dominate at a different scale. Finally, the third component of the tyranny of scales problem is that there can be complex dependencies or feedbacks between the entities and processes operating on these different scales. Although any given phenomenon will manifest these three aspects to varying degrees, collectively they define what is meant by the *tyranny of scales* problem. Moreover, this phrase can be used to describe the physical features of a system themselves, or the challenges that these features give rise to in the context of modeling.

In the philosophical literature, there has been a growing interest in the implications of tyranny of scales problem for modeling specifically. In his Oxford Handbook entry on “The Tyranny of Scales” Robert Batterman argues that

much philosophical confusion about reduction, emergence, atomism, and antirealism follows from the absolute choice between bottom-up and top-down modeling that the tyranny of scales apparently forces upon us. (Batterman, 2013, p. 257)

Batterman argues that the purely reductionistic, bottom-up approach is naive and unsupported by the actual modeling approaches of scientists. He emphasizes instead the importance of mesoscale structures, which “provide the bridges that allow us to model across the scales” (p. 285). More recently, Mark Wilson has also highlighted the tyranny of scales problem and the emergence of multiscale, or what he calls multiscale, modeling. Using the example of modeling the material properties of a steel bar, he notes the problem is not just managing the large number of parameters required to describe the system, but also dealing with what he calls the problem of the “syntactic inconsistency” between models used to describe the system at different scales (Wilson, 2017, pp. 225–226). The tyranny of scales problem has also been recognized to pose an obstacle to reductive explanations in biology (Green & Batterman, 2017). As these papers note, there remains much philosophical work to do in understanding the implications of the tyranny of scales for scientific modeling methodologies.

In this paper I expand philosophical discussions about multiscale modeling to include new examples from the geosciences, and show how this work in the geosciences points to new strategies for taming the tyranny of scales. Before turning to the geoscience examples, however, a couple of preliminary remarks are in order to situate my philosophical approach to scientific modeling. First, although the models-as-tools (e.g., Currie, 2017; Knuuttila, 2005) and models-as-representations (e.g., Giere, 1999) views are often juxtaposed, there is not as much opposition between these views as one might be led to believe. Indeed, on the view I will adopt here, models are both: they are *representational tools* (Parker, 2010). Second, the use of non-veridical elements in modeling (e.g., fictional properties, states, processes and entities) does not render such models non-representational. In keeping with my account of fictional models (2009, 2011) and the eikonic conception (2016, 2018), a scientific model doesn't need to be a veridical representation of its target in order to be a good scientific representation;² it can be a non-veridical representation that nonetheless captures the relevant dependencies and licenses the relevant correct inferences.³

In order to drive this point home, I introduce a terminological distinction between representation and depiction: I use *representation* in its broad sense to mean simply one thing standing in for—or being about—another.⁴ By *depiction* I mean a veridical, or nearly veridical, representation that tries to mirror as closely as possible its target.⁵ This distinction will be important when it comes to the discussion of the different—often conflicting—representations that can be simultaneously deployed in a multiscale model. Recognizing that not all successful representations are depictions reorients the methodological strategies used in multiscale modeling. The aim of scientific modeling need not be to produce as detailed or as fundamental of a veridical depiction as is, say, computationally feasible; rather the aim is to construct a representation that is adequate to the purpose that the model will be used for (e.g., Wimsatt [1987] 2007; Murray, 2013; Parker, 2020).

The central sections of the paper will be organized around three different classes of models: conceptual, physical, and mathematical. Section 2 will focus on conceptual models, which have received little attention from philosophers of science. I will highlight three strategies for taming the tyranny of scales at the level of conceptual modeling, namely, attention to thresholds, hierarchy theory, and adequacy-for-purpose. In Sect. 3, I turn to physical modeling, where the tyranny of scales problem

² This element of my view has often been misunderstood: Both Massimi (2019) and Frigg & Hartmann (2020) have mistakenly assumed that the examples of fictional models I discuss must be a case of “targetless” modeling (2019, p. 870) or a “non-representational” account of model explanation (2020). As a more careful reading makes clear, this is not the case: one can have a fictional representation of real entities and processes.

³ A striking example of this from the history of science is James Clerk Maxwell's (veridical) inference that light is electromagnetic radiation from his fictional (non-veridical) vortex-idle wheel model (see Bokulich 2015 for a discussion).

⁴ There is, of course, a huge literature on representation in scientific modeling, a review of which would take us too far beyond the scope of this paper. For an overview of some prominent philosophical approaches to representation, see Stephen Downes (2009), who offers a cogent argument that there is no one unified account of representation for scientific modeling.

⁵ In the context of art, ‘depiction’ is used to mean *pictorial* representation, though there are debates regarding precisely what that means (see Shech 2016 for a philosophical discussion). I am instead using the term to signal a veridical, rather than specifically pictorial, representation.

typically manifests itself as scale effects. I will examine the conditions required for dynamical scaling to hold, and discuss four different strategies for dealing with the problems that arise when dynamical scaling fails. The philosophy of science literature on multiscale modeling has focused predominantly on mathematical multiscale modeling, which is the topic of Sect. 4. However, what has not been adequately appreciated is that there are many different types of multiscale behavior in nature—involving different spatial and temporal dependencies between the scales—which require fundamentally different kinds of multiscale models.

2 Taming the scales in conceptual modeling

Although there are many different ways one could taxonomize the wide variety of types of scientific models, in the context of the geosciences it is helpful to distinguish the following three broad categories: conceptual models, physical models, and mathematical models. The vast majority of the philosophical literature on scientific modeling has been about mathematical models, with physical models coming in a distant second, and conceptual models receiving comparatively little philosophical attention. In the geosciences, one might argue that conceptual models are the most important category: in addition to being an important class of models on their own, they typically underlie, and are a prerequisite for, both mathematical and physical modeling.

2.1 Conceptual models

A conceptual model, as the name implies, is an abstract (and typically idealized and incomplete) conceptualization of the target system that involves specifying what the key entities, processes, and interactions are operating in that system. It can be conveyed in narrative form or as a diagram. Conceptual models in the geosciences, just like other models, must contend with the tyranny of scales problem. However, there are certain strategies that modelers can make use of at the level of the conceptual model to help tame this tyranny of scales. There are three such strategies I want to highlight in this section: these involve attention to thresholds, hierarchy theory, and adequacy-for-purpose. To illustrate these central philosophical points about conceptual models, and how the tyranny of scales affects them, I will draw on examples from fluvial geomorphology, which is the study of the processes and morphology of rivers in their associated landscapes.

One of the most well-known conceptual models in fluvial geomorphology is Emory Lane's (1955) "stable channel balance" model of a river channel. This conceptual model relates in a qualitative way four variables: amount of water discharge, the sediment supply, the grain size of the sediment, and the river slope, thereby describing how a river channel will either degrade (i.e., erode or incise into the river valley) or aggrade (i.e., fill in the river valley through the deposition of sediment). Although Lane originally conveyed this conceptual model purely descriptively, in 1960 another

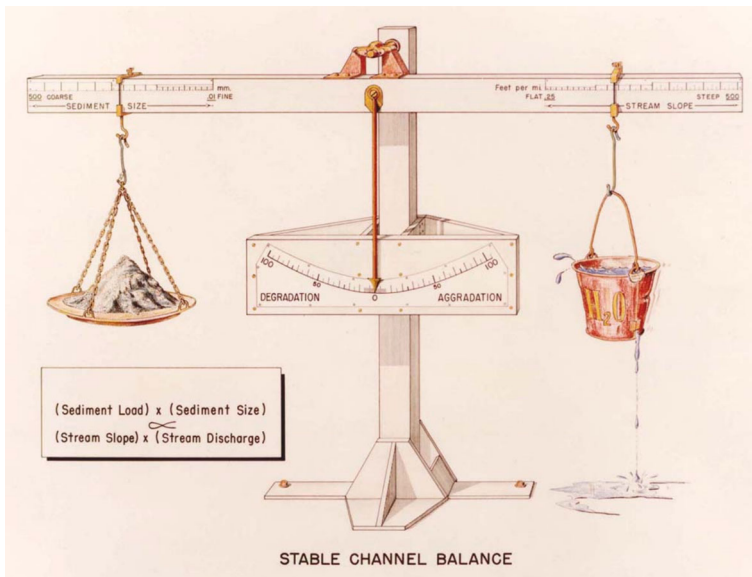


Fig. 2 Diagram of the Lane-Borland stable channel balance conceptual model, from Borland (1960), illustration by James Vitaliano (US Bureau of Reclamation)

civil engineer, Whitney Borland, portrayed Lane’s model with a diagram of a scale balance (Fig. 2).

The ruler on the beam has increasing sediment grain size going to the left and increasing stream slope to the right; the amount of sediment supply is in the left balance pan, and the amount of water discharge is in the bucket on the right. As the balance shifts, the indicator needle will point to the amount of stream channel degradation or aggradation. This simple conceptual model conveys how a river channel will respond to changing environmental conditions: If climate change leads to drought, then the river channel will aggrade; if vegetation slows hillslope erosion, reducing the sediment supply to system, the river will degrade; if the sediment source changes to a finer sediment, that will also cause the channel to degrade. The conceptual model here identifies what the relevant ontological elements, processes, and causal dependences are that are needed in order to understand (or explain, or qualitatively predict, etc.) what will happen to the system of interest (e.g., whether there will be channel degradation, aggradation, or stability).

Although conceptual models on their own are conveyed only narratively or via a diagram, more often in the geosciences today, conceptual models will be turned into either physical models or mathematical models. Despite these different realizations, however, the underlying conceptual model remains an important part of that physical or mathematical model. Thus when a physical or mathematical model fails, it is important to ask whether it was due to the underlying conceptual model, or just its particular implementation in a system of equations, or in a computational algorithm, or in a particular table-top hardware set up.

2.2 Thresholds in conceptual modeling

With this better understanding of conceptual models in hand, we can now turn to how the tyranny of scales problem is manifested in conceptual models, and the first strategy to manage it: thresholds. In the geosciences, the importance of thresholds for taming the tyranny scales was first recognized by the geomorphologist Stanley Schumm. In a landmark paper written with Robert Lichty in the mid-1960s, Schumm calls attention to the tyranny of the scales problem in geomorphology, its implications for understanding cause and effect in geomorphic systems, and the need for better conceptual models to reflect this. Although they don't use the phrase "tyranny of scales" (which is a twenty-first-century label), it is clearly this problem that motivates their discussion of how to investigate and model river channel morphology. They write,

The distinction between cause and effect in the development of landforms is a function of time and space . . . because the factors that determine the character of landforms can be either dependent or independent variables as the limits of time and space change. During moderately long periods of time, for example, river channel morphology is dependent on the geologic and climatic environment, but during a shorter span of time, channel morphology is an independent variable influencing the hydraulics of the channel. (Schumm & Lichty, 1965, p. 110)

They are here calling attention to the fact that river channels involve a wide range of spatial and temporal scales, and importantly that the relevant processes to be modeled—and how they are to be modeled—can change as different spatial and temporal scales are considered.

To aid in the development of more adequate models, they describe how the relationships between the variables of a river system change as different temporal scales are considered. In particular, they consider three different time scales: geologic, modern, and present; and three statuses that river variables can take: indeterminate (i.e., unmeasurable), dependent, and independent (Table 1).

The first time scale they call "geologic," which is on the order of ten thousand to a million years, and begins during the Pleistocene Epoch when glacial discharges established the width and depth of many river valleys, which on this time scale are considered the dependent variables. The second time scale they call "modern," which encompasses the last 1000 years. During this time span, river valley width and depth shift from being dependent variables to independent variables, since they are "inherited" from the paleohydrology of geologic time. The mean water and sediment discharges go from being indeterminate (unmeasurable) to being measurable and the relevant independent variables. It is these independent variables that determine channel morphology, where channel morphology is now a dependent variable. Finally, for the shortest time scale (the "present"), which is on the order of a year, channel morphology changes from being a dependent variable to being an independent variable, and the observed (as opposed to mean) water and sediment discharges now go from being indeterminate (unmeasurable) to being the dependent variables.

In a subsequent paper, Schumm (1979) brings together these insights about the importance of multiscale considerations in the conceptual modeling of fluvial systems with a discussion of thresholds. A threshold can be understood as a marked

Table 1 The status of river variables (rows) during time spans of decreasing duration (columns) redrawn after Schumm and Lichty (1965), Table 2

| River variables | Status of variables during designated time spans | | |
|---|--|---------------|--------------|
| | Geologic | Modern | Present |
| 1. Time | Independent | Not relevant | Not relevant |
| 2. Geology (lithology and structure) | Independent | Independent | Independent |
| 3. Climate | Independent | Independent | Independent |
| 4. Vegetation (type and density) | Dependent | Independent | Independent |
| 5. Relief | Dependent | Independent | Independent |
| 6. Paleohydrology (long-term discharge of water and sediment) | Dependent | Independent | Independent |
| 7. Valley dimension (width, depth, and slope) | Dependent | Independent | Independent |
| 8. Mean discharge of water and sediment | Indeterminate | Independent | Independent |
| 9. Channel morphology (width, depth, slope, shape, and pattern) | Indeterminate | Dependent | Independent |
| 10. Observed discharge of water and sediment | Indeterminate | Indeterminate | Dependent |
| 11. Observed flow characteristics (depth, velocity, turbulence, etc.) | Indeterminate | Indeterminate | Dependent |

transformation of state or behavior as a result of an incremental, ongoing process. A threshold crossing can be precipitated by an external process (extrinsic threshold) or by an internal process of the system (intrinsic threshold); it can be abrupt or transitional; and it can be transitive (reverse crossing restores original state) or intransitive (e.g., Church, 2017). Schumm argues that an adequate conceptual modeling of rivers requires recognizing that river morphology doesn't change continuously, but rather has two critical thresholds: “[T]here is not a continuous change in stream patterns with increasing slope from straight through meandering to braided, but rather the changes occur essentially at two threshold values of slope, or stream power” (Schumm, 1979, p. 491). River morphology can be classified into essentially three types: straight, meandering, or braided. The importance of these thresholds for an adequate understanding of rivers can be seen in Fig. 3.

As Grant and colleagues note,

Adding the concept of thresholds introduces . . . new conceptual models. . . [T]hese newer conceptual models that explicitly recognize a hierarchy of control acting over different timescales with the potential for abrupt changes in system behavior offer hope that future states of the river can be predicted. (Grant et al., 2013, p. 13)

Traditionally geomorphology, like paleontology, was thought to be restricted to being a purely historical and idiographic discipline, that is, confined to “the description of unique, unrepeated events” (Gould, 1980, p. 113). However, with Schumm's and other's work developing new conceptual models that incorporate thresholds within a

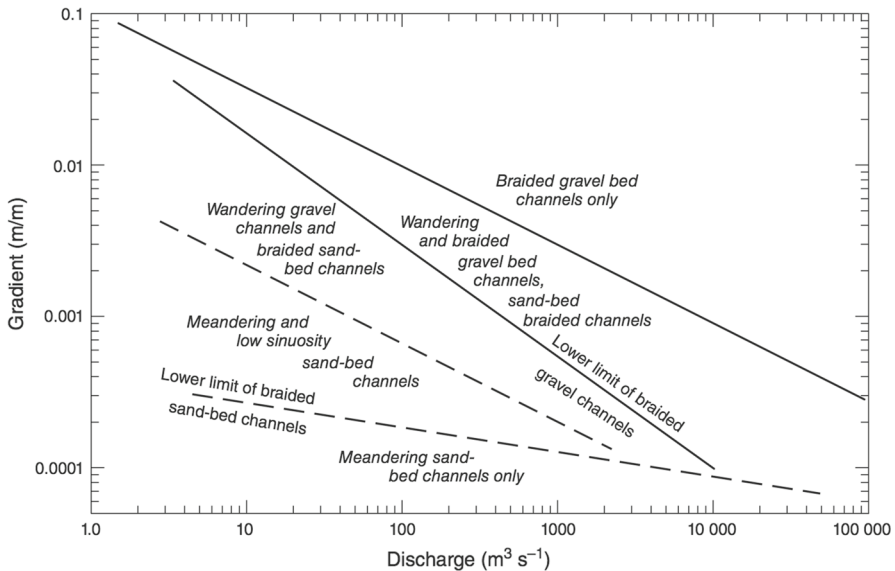


Fig. 3 Threshold diagram for river morphology. (From Church, 2002, Fig. 4, reproduced with permission from John Wiley & Sons.)

multiscale framework, the promise of geomorphology as a nomothetic and predictive science began to be seen as within reach.

2.3 Hierarchy theory in conceptual modeling

A second strategy for taming the tyranny of scales at the level of conceptual modeling draws on what is known as hierarchy theory. Hierarchy theory was first developed in the context of ecology as a method for studying complex, multiscale ecological systems (Allen & Starr, 1982). The basic idea is to develop a conceptual model that locates the phenomenon of interest within a hierarchy of levels. Using the reference scale of the phenomenon of interest, the hierarchy should include the relevant higher levels at a larger spatial or longer temporal scales above it, and the relevant lower levels at a smaller or shorter scale below it. Hierarchy theory focuses attention more specifically on the ways in which the higher (larger or slower) levels act so as to constrain the dynamics of the phenomenon of interest at the target level.

Before further elaborating this strategy for taming the tyranny of scales in conceptual modeling, it is worth pausing to briefly address some common philosophical worries about the notion of a hierarchy of levels. As Potochnik and McGill (2012) argue, many difficulties plague the “classical” notion of a universal set of discrete ontological levels. In particular they rightfully target the classical series of levels built on simple mereology or “composition.” As they note, this does not, however, mean that no hierarchy of levels can be constructed, and they point to scale and causation as the appropriate basis from which to construct a hierarchy of what they call “quasi” levels. They draw on the work of Rueger and McGivern (2010), who point to a different notion of a hierarchy of levels:

When physicists talk about levels, they often do not have in mind a mereological ordering of entities. Instead, what they describe is best understood as a stratification of reality into processes or behaviors at different scales. (Rueger & McGivern, 2010, p. 382)

As Potochnik and McGill clarify, the focus is on the scales at which particular causal processes are operating, rather than composition. I will largely follow these philosophers of science in identifying this as the relevant notion of level to be used in hierarchy theory.⁶

The strategy, on this approach, is to use observations of the different timescales on which processes unfold to identify the functionally relevant scales and to organize the hierarchy. As Brad Werner explains, “variables with disparate intrinsic time scales, when nonlinearly coupled, can develop an asymmetrical relationship: Fast variables become... [tied]⁷ to slow variables and lose their status as independent dynamical variables” (Werner, 1999, p. 103). The variables and processes operating on these slower time scales (the higher level) thus set the context and constraints within which the phenomena of interest operate, while the variables and processes at faster time scales (at lower levels) are abstracted, providing the effective dynamics underlying the phenomenon of interest. These hierarchies do not need to be nested (Allen & Starr, 1982); a non-nested hierarchy is one in which the higher, constraining level is not an aggregation of (composed of) the lower levels. Recall the example of the sandy coastline in the first section: the sand bars are not composed of the ripples, but they set the dynamical context (or constraints) in which the ripples must form and evolve.

Hierarchy theory helps tame the tyranny of scales by focusing attention on the functionally relevant scales. Which scales are functionally relevant is not universal, but depends on the question you are trying to answer. Asking the right question at the wrong scales is what the landscape ecologist Kevin McGarigal describes as a “type III error” (2018). Saying that it is purpose-relative, however, does not mean that it is not objective. The facts that determine which scales are functionally relevant for a given purpose (hence which levels have to be modeled), which processes and variables are the “effective” ones operating at those scales, and the causal relations between those relevant scales and variables is something that has to be discovered empirically. Since we will discuss hierarchical modeling further in the section on mathematical models, let us turn now to the final strategy to be discussed for taming the tyranny of scales in conceptual modeling, namely the importance noted above of paying attention to the *purpose* of the model.

2.4 Adequacy for purpose in conceptual modeling

Before building or selecting a model, one of the most important questions to ask is what purpose you want that model to serve. As Ron Giere cogently argues “[t]here is no best scientific model of anything; there are only models more or less good for different purposes” (Giere, 2001, p. 1060). The central idea is that it is better to have a plurality of

⁶ There are just a few points on which my account will diverge slightly from theirs.

⁷ The problematic traditional term ‘slaved’ has here been replaced with ‘tied’.

different models of a system of interest, each tailored to answering a particular cluster of questions or performing a particular subset of tasks, rather than trying to construct a single model that is supposed to do it all.⁸ What has not, perhaps, been adequately recognized is that this also has implications for the nature of the representations used in the models. The presumption that one can construct a single model that will answer all questions about the system of interest and perform all tasks is bound up with the idea that a model needs to be as accurate and detailed a representation as possible—i.e., what in Sect. 1 I called a *depiction*. Rejecting this presumption of a single model then opens up the possibility of using multiple models, each with different (and perhaps incompatible) representations of a system, that is, representations that are not perfectly accurate but are nonetheless tailored to be adequate for a specific purpose.

This notion of designing and evaluating models as fit for particular purposes has been worked out in greatest philosophical detail by Parker (2010, 2020). Whether or not a model is adequate for purpose depends on more than just the target phenomenon being modeled. She elaborates as follows:

[F]or a model to be adequate-for-purpose, it must stand in a suitable relationship . . . with a target T, user U, methodology W, circumstances B, and goal P jointly. (Parker, 2020, p. 464)

She collectively refers to this as a *problem space*, which is constituted by the problem (P) and constrained by the target, user, methodology, circumstances. The model is then designed (or selected) so as to be a “solution” in this problem space (p. 465). Depending on the problem space, different modes of representation that are more or less veridical⁹ may be called for. Consider the most familiar liquid on the planet: water. If you are working in biology on molecular simulations, then the classical atomistic (“tinker-toy”) model of water may be adequate. If you are trying to understand the high surface tension of water, then a quantum–mechanical model of water might be more appropriate. On the other hand, if you are studying how water flows through an aquifer (within the context of geophysical fluid dynamics), then you must use a continuum representation of water; both the classical atomistic and quantum representations of water will be inadequate.¹⁰

The key point is that the tyranny of scales problem is exacerbated by attempts to construct what we might call the “one-best-depiction” model of a target system, which strives to answer any question about the target that one might pose by making the model as realistic a depiction as possible. This one-best-depiction model approach is often attempted in a “reductionistic” manner by trying to model even largescale (or

⁸ Some have argued that there is a necessary tradeoff in modeling purposes, such as Richard Levins classic (1966) paper arguing that there is a necessary tradeoff between the modeling goals of generality, realism, and precision. For a discussion of this issue in the context of geomorphology see Bokulich (2013) and references therein.

⁹ The issue of veridicality is arguably nontrivial here. At one end of the spectrum of interpretation, veridicality means a representation that uses the ontology of our best current fundamental theory of that domain. Though at the other end of the spectrum, James Nguyen (2020) argues that a *prima facie* fictional representation can be interpreted as veridical if there is an appropriate translation key.

¹⁰ I elaborate this example of different representations of water, discussing how some of them do a better job of allowing researchers to answer certain kinds of questions than others, and arguing that the most “veridical” representation is not always the best in Bokulich (2016).

long timescale) phenomena by beginning with a detailed small spatial scale and short timescale description of the ontology and dynamics.¹¹ So, for example, even if one is trying to understand how regional coastlines evolve over centuries, the one-best-depiction approach would try to do so by beginning with how individual grains of sand are moved in the fluid.

In this context of coastal geomorphology, Jon French and colleagues argue against this one-best-depiction approach, and in favor of what they call *appropriate complexity* modeling. They write,

It should be fundamental to any modelling endeavour to match the level of mechanistic understanding with the scale of the problem. In the context of understanding and managing coastal change, that scale of interest is often as much determined by applications and stakeholder needs than by any intrinsic organisational property that the geomorphological systems involved are dependent upon. (French et al., 2016, p. 14)

Here French and colleagues are calling attention to the importance of what Parker describes as the problem space of the modeling endeavor, and how that problem space determines the way in which the particular coastal geomorphology system is to be modeled. Their notion of appropriate complexity modeling captures this emphasis on adequacy for purpose.

Interestingly French and colleagues also allude to the other two strategies we have described here for taming the tyranny of scales in conceptual modeling: thresholds and insights from hierarchy theory. They write that models should be “formulated with a view to resolving critical state changes” (p. 3), that is, the dynamical or behavioral thresholds in the geomorphic system. They note that while “reductionist” approaches to modeling are often able to resolve incremental changes in the system, they are ill-suited to capturing the qualitative changes of state characteristic of thresholds that emerge at mesoscales. They also emphasize that modeling coastal geomorphology on the coastal tract cascade approach is developed with explicit reference to hierarchy theory (p. 8). The concept of a “coastal tract” was introduced by Peter Cowell and colleagues as a “new overarching morphological entity” (2003, p. 813), with the corresponding notion of a coastal-tract cascade as a way to formalize concepts for separating coastal processes and behaviors into a scale hierarchy. They explain,

The purpose of the cascade is to provide a systematic basis for distinguishing, on any level of the hierarchy, those processes that must be included as internal variables in modeling coastal change, from those that constitute boundary conditions [at larger scales], and yet others that may be regarded as unimportant ‘noise’ [at smaller scales]. (Cowell et al., 2003, p. 825)

The concept of a coastal tract thus provides a framework for aggregating processes into mesoscale models. As we will see in Sect. 4, this “aggregating” is not a flat-footed averaging, but rather a synthesist approach that focuses on identifying the emergent variables or effective degrees of freedom.

¹¹ The label “reductionistic” modeling is widespread in the geomorphology literature and is often contrasted with what is called “reduced complexity” modeling. For a philosophical discussion of these types, see Bokulich (2013).

To summarize, we have examined three strategies for taming the tyranny of scales at the level of conceptual modeling: First, the importance of identifying the relevant dynamic thresholds in the system where there are qualitatively discontinuous changes in the system's behavior. Second, the deployment of hierarchy theory, which identifies the functionally relevant scales given the application of interest, as well as the constraints coming from the higher levels of the hierarchy. Third, the exhortation to develop models with a focus on their adequacy for particular purposes (or more generally their suitability to a particular problem space), rather than develop models as all-purpose, realistic depictions of their target systems. I have spent considerable time on conceptual models because, in addition to being an important (yet neglected) class of models in their own right, they also underlie the other two broad classes of models we will discuss (*viz.*, physical models and mathematical models). With this deeper understanding of conceptual modeling in hand, let us now turn to the second-least philosophically discussed category of models: physical models.

3 Taming the scales in physical modeling

Physical models have long played a central role in the geosciences. Physical models are known by many names, including table-top models, concrete models (e.g., Weisberg, 2013), material models (e.g., Frigg & Nguyen, 2018), and hardware models; they are typically either a scaled-down version of the system of interest, hence are often referred to as scale models (e.g., Bokulich & Oreskes, 2017), or they are analogue physical systems, hence also sometimes referred to as analog models (e.g., Sterrett, 2017b). Physical modeling in the geosciences faces several dimensions of the tyranny of scales problem. First, many geoscience systems of interest, such as river valleys, coastlines, and mountain ranges, develop over large spatial and long temporal scales; and moreover, the forces that shape these geological systems also can involve pressures and temperatures beyond our reach. These conditions make the investigation of such systems difficult and their manipulation nearly impossible.¹² To overcome these limitations, scale physical models are developed using dynamical scaling and the formal theory of similitude (discussed below). However, true scaling is often difficult, if not impossible, to achieve in all the relevant respects; hence, researchers often must make do with what are called “distorted” scale models. Such distorted scale models exhibit “scale effects,” which are artefacts of the modeling process due to problems of scale, and these can lead the behavior of the model system to deviate from the behavior of the target. As we will see, recent scaling studies in the geosciences have revealed opportunities for addressing some of these tyranny of scales problems in physical modeling. Each of these scale issues in physical modeling will be discussed in turn, beginning with a brief introduction to the formal theory of similitude and dynamical scaling.

¹² I say ‘nearly’ because human activity through dam-building, extensive mining, deforestation, and climate change are manipulating geomorphic systems, though these projects are not typically undertaken with the aim of advancing scientific knowledge. Most of the experimentation that takes place in the geosciences is what we might call process-based experiments which focus on understanding individual geological processes, such as wind-tunnel studies of sand abrasion on different types of rock.

In general, simply shrinking down a system into a smaller scale physical model will not be informative about the target system. In order to have a model that is scientifically useful, one typically needs to design a physical model that has the relevant similitude. There are three key notions of similitude. The first is *geometric similitude*. Intuitively, two systems are geometrically similar if they have the same shape. More precisely, geometric similitude obtains when the length ratios of model and target along every dimension have the same linear scale ratio (λ), which implies that all angles preserved. Susan Sterrett (2021) gives the example of a common school exercise to indirectly measure the height of a tree by measuring the length of its shadow, along with measuring your own height and shadow. Since the angle of the sun is the same, the ratios of heights to shadows should be the same; thus by knowing three quantities in the equated ratios one can calculate the fourth quantity. The two triangles formed by object height, shadow length, and light ray have geometric similitude.

The second notion of similitude in formal scaling is *kinematic similitude*. Here the physical model and target system must have the same displacement ratios and velocity ratios at their respective homologous points. The third, and often most important, notion of similarity is *dynamic similitude*, which will be the focus of our discussion below. When dynamic similitude can be established between the physical model and target system, one can be confident that the investigations carried out on the physical model will be a reliable guide to inferences drawn about the target system.

The subtleties of dynamic similitude and scaling were first recognized by Galileo and were the subject of the first two days of his *Dialogues Concerning Two New Sciences*.¹³ Galileo writes,

[Y]ou can plainly see the impossibility of increasing the size of structures to vast dimensions either in art or in nature; . . . it would be impossible to build up the bony structures of men, horses, or other animals so as to hold together and perform their normal functions . . . for this increase in height can be accomplished only by employing a material that is harder and stronger than usual. (Galilei [1638] 1991, p. 130)

Galileo's insights on scaling seem to have not been widely known or appreciated, and it was not until the late 1930s that the geologist M. King Hubbert would introduce the quantitative theory of dynamic scaling into the geosciences, citing Galileo as his inspiration.¹⁴

Hubbert's groundbreaking work on scaling was motivated by the need for a new approach to physical modeling in the geosciences. His two key papers laying this foundation are "Theory of Scale Models as Applied to the Study of Geologic Structures" (1937) and "Strength of the Earth" (1945). Hubbert's latter paper begins with

¹³ Galileo's theory of scaling, discussed in relation to the strength of materials, is the first of the "two new sciences." For an interesting account of his discovery of scaling principles that locates the inspiration for his insights in his lectures on the spatial dimensions of hell in Dante's *Inferno*, see Peterson (2002). For a more complete history of physical similitude modeling, see Sterrett 2017a.

¹⁴ Hubbert cites Galileo in his landmark (1937) paper on scaling and in his AIP oral history interview (Hubbert 1989, <https://www.aip.org/history-programs/niels-bohr-library/oral-histories/5031-5>). In the context of biology, Haldane ([1926] 1985) had resurrected Galileo's insights on scaling a decade earlier in his paper "On Being the Right Size".

the following puzzle: “Among the most perplexing problems in geologic science has been the persistent one of how an earth whose exterior is composed of hard rocks can have undergone repeated deformations *as if* composed of very weak and plastic materials” (Hubbert, 1945, p. 1630). He notes, citing John Hutton, that it is not that the forces in the past were any more “spectacular” than the geologic forces experienced today. He then shows that a proper accounting of scale differences, through the formal theory of dynamical scaling, can resolve this paradox: “By means of the principles of physical similarity it is possible to translate geological phenomena whose length and time scales are outside the domain of our direct sensory perception into physically similar systems within that domain” (p. 1630). Reducing spatial scales down to 1 to 10 million (reduction ratio of 10^{-7}) and temporal scales down to 1 min to 10 million years (reduction ratio of 10^{-13}), he calculates that the corresponding reduction ratio for the viscosity of the Earth would be 10^{-20} , making the hard rock of our experience comparable to that of “very soft mud or pancake batter” (Hubbert, 1945, p. 1651). The theory of dynamical scaling not only resolves this puzzle, but shows how one can construct a physical model such that it bears the relevant physical similarity (dynamic similitude) to the system of interest, thus providing a critical tool for taming the vast scales of geologic phenomena and bringing them within the reach of scientific investigation.

Dynamic similitude is the idea that the dynamics that govern one system are equivalent to the dynamics that govern another system, which implies that the two systems will behave the same way. The standard way to secure this dynamic similitude is by making sure that for all relevant forces, the ratio of forces in one system is equal to the corresponding ratio of forces in the other system. Examples of force ratios that are relevant when fluid dynamics are involved include the ratio of the inertial force to the gravitational force (known as the Froude number “F”), the ratio of the inertial force to the viscous force (known as the Reynolds number “R”), the ratio of the inertial force to the force of surface tension (known as the Weber number “W”), and the ratio of the inertial force to the elastic compression force (known as the Cauchy number “C”). When any of these ratios is satisfied, the model and target are said to have the corresponding similarity, such as “Froude similarity.” To achieve full dynamic similitude, *all* of these ratios in the model would need to take on the same values as the ratios in the real-world target system.

Although the theory of dynamical scaling may be clear, the challenge is that, in practice, it is highly nontrivial to get all the force ratios in the physical model to have the same values as the force ratios in the target system. Indeed, if one is using the same fluid in the physical model as in nature (e.g., water) then complete dynamical similarity is nearly impossible, because the viscosity of water cannot be appropriately reduced.¹⁵ In such cases, only one of the force ratios can be satisfied and the modeler has to choose which one is the most important for a given study.

As Valentin Heller (2011) explains, for phenomena where gravity and inertial forces are dominant, Froude similarity is the most important to satisfy in the physical model;

¹⁵ Few liquids have lower viscosity than water, though isopropyl alcohol and even air have been used as substitutes in models to help achieve similarity. However, such substitutes will often lose similarity in another respect (e.g., air models can recover viscosity effects, but will fail to reproduce gravity effects (Heller 2011, p. 301)). The gravitational force is another example of a quantity that is difficult to scale, though some physical models have done so by placing the apparatus within a centrifuge.

for phenomena where viscous and inertial forces are dominant, Reynolds similarity is the most relevant. As we saw in the context of conceptual modeling, a key strategy when perfect similarity between model and target is impossible is to try to make the model adequate to the particular purpose for which it will be used—a point to which we will return below.

When force ratios in the model are not the same as in the target, then scale effects occur. Scale effects are an important class of artefacts in the modeling process, arising from differences of scale that cause the behavior of the model to deviate from the behavior of the target.¹⁶ Scale effects can occur when forces that are not relevant in the target system, such as surface tension or cohesive forces, become dominant at the scale of the physical model. Following the terminology introduced above, physical models that are *Froude similar*, will have scale effects arising from the inequivalent Reynolds, Weber, and Cauchy numbers, while physical models that are *Reynolds similar* will have scale effects arising from inequivalent Froude, Weber, and Cauchy numbers, and so forth.

It is important to recognize that scale effects are one manifestation of the tyranny of scales problem. Recall that we identified three aspects of the tyranny of scales problem: first, the systems involve a wide range of spatial and temporal scales; second, the processes or dynamics dominating at one scale will be different from those dominating at a different scale; and, third, there can be complex dependencies between entities and properties operating at these different spatial and temporal scales. Scale effects, as we've seen here, arise when the effective forces that dominate the system at the scale of the physical model are not the same as the effective forces that dominate at the scale of the target system (e.g., when cohesive forces dominate at the scale of the physical model, but are not significant at the larger scale of the target system).

Although scale effects are a manifestation of the tyranny of scales problem that can invalidate the inferences drawn from physical scale models, there are a number of strategies for managing such effects. However, successfully managing scale effects requires first understanding the magnitude and direction of their influence on the variables of interest. If one is dealing with a very simple system that is theoretically well understood, then one might be able to calculate the influence of the scale effects a priori. Typically, however, the complexity of systems in the geosciences makes this impossible, and the detection and influence of scale effects must be investigated empirically. An ingenious strategy for determining the influence of scale effects experimentally is to construct what is called a *scale series* of physical models.

A scale series is a sequence of several physical models for the same target system, with each one constructed at a different scale in order to investigate the influence of scale effects as the model is shrunk down. An example from the geosciences is the work Heller and colleagues who studied how landslides and avalanches can generate

¹⁶ Scale effects are just one subset of artefacts that can arise in the physical modeling process. There are also more general “model effects” which can, for example, arise from modeling a 3D system as 2D, or artefacts arising from the boundary conditions of the model, etc. Heller discusses a third category of artefacts, which he calls “measurement effects” that arise from different measurement and data sampling techniques being used in the model and the target system (2011, p. 293).

large impulse waves, such as the 1958 landslide into Lituya Bay in Alaska that generated a megatsunami with a wave height of 524 m (or 1,720 feet).¹⁷ Heller et al. (2008) constructed a series of seven different scale physical models (using a pneumatic landslide generator and wave channel tank) all of which were Froude similar (meaning the ratio of inertial force to gravitational force was essentially the same in both the Lituya Bay landslide and in the scale models). Because the physical models used water, it was not possible for the other force ratios (Reynolds, Weber, and Cauchy numbers) to simultaneously be satisfied.¹⁸ The failure to satisfy these other force ratios means that scale effects will influence the behavior of the model, making it deviate from the target system. The key insight behind using a scale series is to investigate how those scale effects manifest themselves in a sequence of scale models so their influence on various variables of interest can be determined.

Once one has a quantitative understanding of the relevant scale effects and their influence on various quantities, then there are various strategies one can deploy to manage them. Heller (2011) groups these strategies into three categories, which he labels *avoidance*, *correction* and *compensation*. The first strategy of avoidance takes advantage of two key strategies highlighted in the context of conceptual modeling: thresholds and adequacy for purpose. Although scale effects cannot be entirely avoided, one can design the scale model such the scale effects are within a regime where their influence on the variable of interest (i.e., for a particular purpose) is negligible. For example, Heller's scale series of Froude similar models was used to define threshold values of Reynolds and Weber numbers for which scale effects impacting the maximum wave amplitude can be neglected. He writes, "considering all seven scale series, scale effects are negligibly small (< 2%) relative to the maximum wave amplitude a_m if $R_I = g^{1/2} h^{3/2} / \nu \geq 300,000$ and $W_I = \rho g h^2 \sigma \geq 3,000$ " (Heller, 2011, p. 299).¹⁹ In other words, if the Reynolds and Weber numbers are above a certain threshold, then their associated scale effects for wave amplitude can be neglected. These threshold studies give rise to various "rules of thumb" for how to construct a scale model such that the impact of scale effects on the variables of interest is minimized. As Heller explains, however, whether various threshold values will be adequate depends on which feature of the target one is trying to draw inferences about:

If one parameter, such as discharge . . . is not considerably affected by scale effects, it does not necessarily mean that other parameters, such as the jet air

¹⁷ This case is also discussed in Pincock 2020, as will be mentioned below.

¹⁸ Chris Pincock has argued such models are *essentially idealized*, by which he means that "there is no known way to develop a concrete model... without having that model generate some known falsehoods about its intended target" (2020, p. 12). Some comments on Pincock's notion of 'essential' are in order here: first, the inability to satisfy all the force ratios is predicated on the assumption that the same fluid (water) is used, and as noted above different fluids can be used to improve the Reynolds and Weber numbers. However, more generally I would argue that *all* models are "essentially idealized", since no model is identical to the target and hence there will always be some way to generate a falsehood about the target using the model. Models in science typically have an implicit or explicit set of guidelines for what kinds of inferences are—or are not—licensed about the target system on the basis of the model, though determining *which* category a particular inference falls into can sometimes be a substantive scientific question requiring further theoretical or empirical investigation.

¹⁹ The subscript I on R and W refers to the impulse Reynolds and impulse Weber numbers, g is gravitational acceleration, h is the still water depth, ν is kinematic viscosity, ρ is density and σ is the surface tension of water.

concentrations, are also not affected. Each involved parameter requires its own judgement regarding scale effects. (Heller, 2011, p. 296)

That is, the avoidance strategy for managing scale effects in physical models requires paying attention not only to key thresholds, but also to the particular purpose for which that scale model will be used. It cannot be assumed that satisfying the thresholds that render a physical model adequate for one purpose, will also make that model adequate for other purposes. There will not typically be “one best” scale physical model, from which all inferences of interest about a target system can be drawn.

Chris Pincock has analyzed Heller’s avoidance strategy for scale modeling in terms of James Woodward’s notion of conditional causal irrelevance (Woodward, 2021). Pincock’s adaptation specifies that “a set of variables. Y_k is irrelevant to variable E conditional on additional variables X_i each exceeding [or falling below] some specified threshold” when the X_i and Y_k sets of variables are both unconditionally relevant to E , but when the X_i exceed (or fall below) the threshold, then changes to Y_k are irrelevant to E (Pincock, 2020, p. 18). So for the effect variable a_m (the maximum wave amplitude), when R and W exceed the above thresholds, the “mismatch between model and target with respect to causal consequences of ν [kinematic viscosity] and σ [surface tension] can be discounted as their actual values are conditionally causally irrelevant [to a_m]” (Pincock, 2020, p. 18). Strictly speaking, to call them conditionally causally irrelevant is misleading because the scale effects arising from the inequivalent R and W numbers (related to the viscous force and surface tension force respectively) do in fact causally influence the value of the variable of interest—in this example the maximum wave amplitude (a_m)—they just don’t change its value by more than 2%, as noted above. Given Heller’s aim, this inaccuracy in the value of the effect variable (arising from these scale effects) is not enough to thwart the purpose of the study, though there could, of course, be other scientific projects for which this 2% difference (arising from these residual scale effects within this threshold regime) is relevant.

Heller identifies two other strategies for taming the tyranny of scales in physical modeling that Pincock does not discuss; these are *correction* and *compensation*. If one is able to quantitatively determine the influence of the scale effects on a given variable, then one might be able to correct for them mathematically after the data is collected. I have elaborated this sort of approach in my philosophical discussions of model-based data correction (Bokulich, 2018, 2020; Bokulich & Parker, 2021). The idea here is to try to control scale effects, not physically during the construction of the model, but rather vicariously during data reduction; this is an extension of Norton and Suppe’s (2001) notion of *vicarious control* from their general context of experimentation to the present case of scale effects in physical modeling. Correction is, I argue, a central tool for taming tyranny of scales in physical modeling.

An example of the correction approach for managing scale effects in physical modeling is found in the work of Cuomo et al. (2010). They are concerned with studying wave impact pressures on sea walls and the scale effects that arise for their Froude-scaled physical models. They note that such models overestimate the impact pressure in the target system and set out to develop a scale correction factor, drawing on theoretical and experimental work from the 1930s by the geomorphologist Ralph Bagnold. They determine a quantity called the Bagnold number that can be calculated

for both the scale physical model and the target system in the world. Using these two Bagnold numbers as the axes, they construct a correction factor graph that determines the amount by which the impact pressure of the wave in the model needs to be adjusted in order to infer the correct value in the real-world scenario. Thus despite the inability to construct a physical model that preserves the relevant similitude, they nonetheless are able to use the imperfectly scaled physical model *plus correction factor* to draw correct inferences. This approach can only be used, however, if there is a sufficient theoretical understanding of the scale effects and how they influence the quantities of interest—knowledge that can be nontrivial to obtain.

The third category of strategies for taming the tyranny of scale effects is known as *compensation*. On this approach one intentionally gives up on one aspect of physical similarity in order to achieve a greater degree of physical similarity in another respect. In fluvial geomorphology, for example, although one could construct a physical scale model of a river where geometric similarity is preserved, geomorphologists will often intentionally distort the geometry of the river. This is because a geometrically down-scaled river model will exhibit increased friction effects, making its flow behavior deviate from that of a natural river. By giving up perfect geometric similitude, and making the river width and height scale factor larger than the length scale factor, the friction scale effects can be compensated for and the similarity in flow behavior improved. Compensation is particularly useful in those situations where the strategy of avoiding scale effects is difficult to implement and the more complete theoretical understanding necessary for the strategy of correction is lacking.

So far we have been discussing successful physical modeling in terms of achieving dynamic similitude or a full physical similarity. As we have seen, this is not always possible, but nonetheless modeling can still be successful with partial or “incomplete” similarity, or even a distorted similarity (sometimes called affinity) that can be corrected or compensated for either physically or vicariously. When it comes to physical modeling in the geosciences, the requirements for formal scaling can rarely be achieved. This might lead one to be pessimistic about the utility of physical modeling in the geosciences. A review by Chris Paola and colleagues argues this pessimism is unfounded. In their paper “The ‘Unreasonable Effectiveness’ of Stratigraphic and Geomorphic Experiments” they write,

The principal finding of our review is that stratigraphic and geomorphic experiments work surprisingly well. By this we mean that morphodynamic self-organization in experiments creates spatial patterns and kinematics that resemble those observed in the field. . . [despite the fact that they] fall well short of satisfying the requirements of dynamic scaling. . . . Similarity in landforms and processes in the absence of strict dynamic scaling is what we mean by the ‘unreasonable effectiveness’. (Paola et al., 2009, pp. 33–34)

The reference here, of course, is to Eugene Wigner’s classic 1960 paper “The Unreasonable Effectiveness of Mathematics in the Natural Sciences,” and the analogous point is that improperly scaled physical models in the geosciences, which should not be able to generate the relevant patterns found in nature, nonetheless seem to be able to do so. However, the key question, as they note, is whether this *pattern* similarity is indicative of a broader underlying *process* similarity that could support other scientific

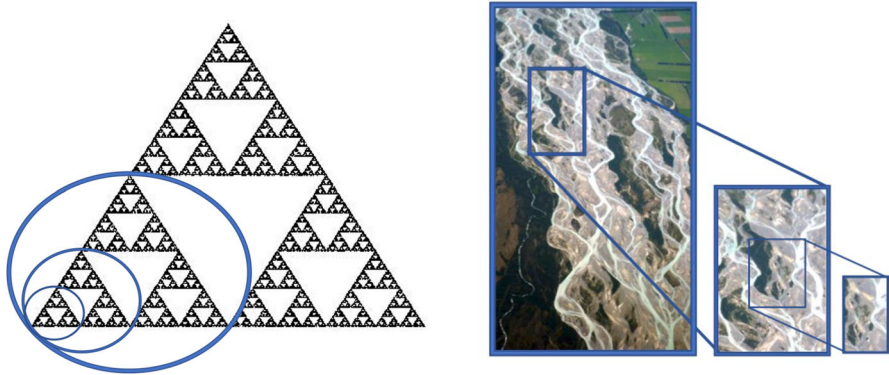


Fig. 4 Left: Sierpinski triangle illustrating self-similarity. Right: Braided river exhibiting idea of statistical self-affinity

inferences. Determining the answer to this question suggests yet another approach to taming the tyranny of scales in physical modeling: scale independence.

The approach taken by Paola and colleagues is to search for aspects of natural phenomena that are scale independent over a certain scale range. They write, “our goal is to refocus the discussion of geomorphic experimentation away from formal scaling and toward the causes, manifestations, and limits of scale independence” (Paola et al., 2009, p. 34). They go on to identify three different ways that a limited kind of scale independence can arise in a physical system: these can be labeled *self-similarity*, *convergence*, and *decoupling*. Self-similarity is the idea that a part of a system can bear a certain similarity to the whole. A Sierpinski triangle is an object that has an exact self-similarity—the object looks the same whether you zoom in or out (see Fig. 4). More often in nature, self-similarity is not exact, but rather is statistical. Statistical self-similarity holds when the “statistical properties of the phenomenon at one scale relate to the statistical properties at another scale via a transformation which involves only the ratio of the two scales” (Sapozhnikov & Fofoula-Georgiou, 1996, p. 1429). One can also relax the notion of self-similarity to self-affinity, meaning the system scales differently along different dimensions. In a series of papers Sapozhnikov and Fofoula-Georgiou (1996, 1997) quantitatively show that braided rivers exhibit statistical self-affinity. They further show that this statistical self-affinity was not just morphological (a small part of a braided river spatially resembles (in a statistical sense) a larger part of the river) but also dynamical (a small part of a braided river evolves in a statistically identical way to a larger part of the river). Paola et al. note that these studies of self-similarity in fluvial geomorphology have important implications for modeling.²⁰ They contrast self-similarity, which they call *internal similarity*, with *external similarity*, by which they mean that “a small version of a large system is similar to the large system” (Paola et al., 2009, p. 34). They argue that internal similarity is an indicator of external similarity, meaning that systems in nature that exhibit internal similarity (such as braided rivers) are able to be successfully investigated with smaller-scale physical

²⁰ See also Murray (2007b) for a discussion of implications of self-similarity for modeling.

models. Although they take internal similarity to imply external similarity, they note that the reverse is not the case: systems that exhibit external similarity need not exhibit internal similarity. This is because there are other ways to achieve external similarity, or what we more generally call (limited) scale independence.²¹

A second way to achieve scale independence is through what has been called *convergent physics*. In their study of alluvial rivers,²² Eric Lajeunesse and colleagues note that because turbulent flows are ubiquitous in such rivers, it has been assumed that the rhythmic bedform morphologies (ripples, subaqueous dunes, etc.) of such rivers are caused by turbulence. They show, however, that such morphologies can also be formed by laminar flows. This is surprising because laminar flows—that is, orderly streamline flows where flow properties remain constant at each point in the fluid—are dramatically different from turbulent flows, which undergo mixing and chaotic changes in flow pressure and velocity, including the formation of vortices. The transition from laminar to turbulent flow is an important threshold, marked by a disparity in Reynolds numbers. Thus “laminar-flow analogues of turbulent-flow morphologies cannot... be expected to satisfy dynamic similarity in terms of all the relevant dimensionless parameters” (Lajeunesse et al., 2010, p. 1). Nonetheless, despite this failure of dynamic scaling, “microscale experiments with laminar flow... create analogues of morphodynamic phenomena that also occur in turbulent flow at field scales” (p. 2). They are careful to note that this does *not* imply that turbulence is irrelevant to the formation of these features—both the timescales of development and the spatial scales of expression differ “because laminar and turbulent flows obey different friction relations, one is Reynolds dependent and the other is not” (p. 21). The reason that the same bedform morphologies form in the laminar flow of the physical model and the turbulent flow of the real river is that there is what they term a *convergence of the physics*. As Paola et al. explain in more detail,

the relation between shear stress and topography, and that between shear stress and bedload flux—is similar enough that the morphodynamics is surprisingly consistent across this major, scale-dependent transition. (Paola et al., 2009, p. 36)

Thus convergent physics is another way to achieve the “external similarity” necessary for successful physical modeling, without satisfying the formal requirements of dynamical scaling, and without a system exhibiting self-similarity.

Yet a third way to achieve the external similarity needed for physical modeling is through a decoupling of the scales. Decoupling of scales occurs when the dynamics at the scale of interest is insensitive to the details of the behavior at smaller scales. As Paola et al. note, “[t]he fluid and sediment processes that are the focus of classical dynamical scaling occur at the fine-scale end of most landscape experiments, so insensitivity to fine-scale dynamics translates into scale independence” (p. 36). In other words, there is a decoupling of the large scale dynamics of interest from the underlying details of smaller-scale processes, so getting those small-scale processes

²¹ For a more mathematical discussion of the relationship between scaling and self-similarity see Barenblatt (2003).

²² Alluvial rivers are just rivers whose channels are formed by moveable sediment.

“right” is not as critical.²³ This is because some features of landscape evolution are only sensitive to general properties of the flow that can be multiply realized in a wide variety of ways at the micro-level.

The “unreasonable effectiveness” of small-scale physical models in successfully representing large-scale stratigraphic and geomorphic systems is thus explained by the much wider variety of ways in which external similarity can be achieved, beyond the confines of strict dynamic scaling. They thus lay out an alternative research program for taming the tyranny of scales in physical modeling that involves refocusing “the discussion geomorphic experimentation away from formal scaling and toward the causes, manifestations, and limits of scale independence” (Paola et al., 2009, p. 34). The three paths to scale independence outlined here, namely *self-similarity*, *convergence*, and *decoupling* are an important step in that direction.

4 Taming the scales in mathematical modeling

So far we have discussed methods for taming the tyranny of scales in multiscale systems in the context of both conceptual modeling and physical modeling. Here we turn to multiscale approaches in mathematical modeling and simulation, which has been the dominant focus of the philosophical literature. By and large, philosophical discussions have tended to lump all multiscale models together, without recognizing that there is a plurality of multiscale methods. My aim here is both to draw apart some of the different approaches to mathematical modeling in the geosciences (Sect. 4.1) and to show how different kinds of multiscale systems in nature (Sect. 4.2) require different kinds of multiscale models (Sect. 4.3).

4.1 Reductionist, universality, and synthesist approaches

There is a variety of approaches to mathematical modeling in the geosciences, which have been usefully grouped into three general categories: the ‘reductionist’²⁴ approach, the universality approach, and the synthesist approach (e.g., Murray, 2003; Werner, 1999).²⁵ The first, so-called *reductionist modeling* approach tries to remain firmly grounded in classical mechanics, invoking laws such as conservation of mass, conservation of momentum, classical gravity, and the like. In a field like aeolian geomorphology (which studies how wind-driven sand dunes and dune fields evolve) these conservation laws may be applied to the motion of individual grains of sand and how they bounce, impact, and roll (saltate) along other grains of sand as they are moved by the wind. This approach moreover seeks to represent in the model as many of the physical processes known to be operating in the target system and in as much detail as is computationally feasible. The reductionist approach to modeling fits nicely with the

²³ I will return to this idea in Sect. 4.2, given its central role in hierarchical modeling.

²⁴ As will become clear, geomorphologists are using the term ‘reductionist’ in way different from philosophers of science.

²⁵ These three approaches are conceptually useful to distinguish, even though they are something of an idealization, with a continuum of cases graduating between them.

view (discussed in Sect. 2) that models should be *depictions* of their target systems. Traditional climate models (e.g., GCMs—global circulation models) can also be seen as an example of this approach to modeling. One concern with reductionist modeling is that, as the complexity of the model approaches the complexity of the target system, the model becomes almost as opaque as the real-world system you are trying to understand. Moreover, as Brad Murray notes, “[w]hen modeling large-scale behaviors of multi-scale systems, explicitly representing processes at much smaller time and space scales unavoidably runs the risk that imperfections in parameterizations at those scales cascade up through the scales, precluding reliable results” (Murray, 2007a, p. 189). In other words, in some cases a reductionist approach may result in a *worse* model than a model that represented the effective variables and processes at a higher level of description.

At the other end of the spectrum is the *universality* approach, which tries to find the simplest model that belongs to the same universality class as the target system of interest. Such models are devoid of the physical details that distinguish one type of system in this class from another. Batterman and Rice (2014) have described these sorts of models as minimal models. The notion of a universality class was developed in context of condensed matter physics to describe why very different fluids all behave the same way near their critical points, and has since been extended to other fields, such as biology and the geosciences. Universality modeling has the advantage of simplicity, but is often at a level of abstraction too far removed for predicting the behavior of the sorts of real-world systems one finds in the geosciences, and hence of limited value.

A third approach to modeling in the geosciences has been termed *synthesist modeling*, which is often deployed within a hierarchical modeling framework. The synthesist approach focuses attention on the emergent variables, or what might better be described as the effective degrees of freedom of the system at the scale of interest. Paola introduces the synthesist approach as follows:

The crux of the new approach to modelling complex, multi-scale systems is that behaviour at a given level in the hierarchy of scales may be dominated by only a few crucial aspects of the dynamics at the next level below. . . . Crudely speaking, it does not make sense to model 100% of the lower-level dynamics if only 1% of it actually contributes to the dynamics at the (higher) level of interest. (Paola, 2001, p. 12)

Of course, not even the reductionist can model 100% of the lower level dynamics, so a key question is *how* one should go about reducing the degrees of freedom. The reductionist and synthesist differ not just in the amount detail they include in their respective models, but also in how they arrive at what detail is included. The reductionist typically tries to reduce the degrees of freedom through a straightforward lumping or averaging procedure. This is sometimes called traditional upscaling, where one posits a microscale model and then invokes a formal averaging procedure, such as volume averaging or homogenization to arrive at a macroscale model (Scheibe & Smith, 2015, Sect. 2.1).

The synthesist, by contrast, tries to identify *emergent structures*.²⁶ Paola illustrates this difference as follows:

[W]hat I termed ‘synthesism’ really does represent a major departure from traditional reductionism. . . . [S]ynthesism as I understand it is not the same as lumping or averaging—for instance, the Reynolds and St. Venant equations are certainly not synthesist in either letter or spirit, although they represent two successive levels of averaging. For instance, a synthesist approach to turbulence might not involve formal averaging at all, but instead centre on coherent structures as the crucial emergent feature of the dynamics. (Paola, 2001, p. 43)

In the context of turbulence, such coherent structures might be things like quasi-stable ring vortices or hairpin vortices. In the context of aeolian geomorphology, emergent coherent structures could be things such as pattern “defects” in a field of ripples or even sand dunes themselves, which can maintain a structural coherence even as they move for miles, exchanging sand.²⁷ The synthesist idea is to describe the dynamics of the system in terms of the evolution and interaction of these larger-scale coherent structures, or more generally in terms of emergent variables and dynamics, rather than a simple averaging of the lower-level dynamics. This synthesist approach of identifying emergent structures at different levels of the dynamics can then be combined with a hierarchical or multiscale approach to modeling.

4.2 Behavior across the scales and hierarchical modeling

The recognition of the many different manifestations of the tyranny of scales in the geosciences has led to the development of many different multiscale approaches to mathematical modeling and simulation. Before engaging in mathematical modeling, however, it is important to ask what *kind* of hierarchical or multiscale behavior your system of interest is exhibiting. Physical systems can exhibit multiscale behavior in spatial dimensions, temporal dimensions, or both. And the sort of multiscale behavior within each of these dimension can vary too.

Timothy Scheibe and colleagues have developed an analysis platform for distinguishing various types of multiscale behavior within hydrogeological systems. Although developed in the context of the geosciences, their conceptual taxonomy is much more widely applicable. They note that a system may have an evolving sequence of characteristic length scales. Consider an effective parameter, γ , which in a hydrogeological context could represent a field quantity such as porosity (the average amount of pore space in a volume of rock or soil). They illustrate two different possible multiscale behaviors for γ in Fig. 5 below.

There are different thresholds (in increasing r) for different key length scales: sub-pore, Darcy scale, mesoscale, and field scale. Consider curve (1) in Fig. 5: at the sub-pore scale, measures of γ will be constant (because one is sampling within a

²⁶ ‘Emergent’ is here being used in the sense of weak emergence (see, e.g., Chalmers 2006 for this distinction). One could also call them effective dynamical structures.

²⁷ On defect dynamics in ripples see Werner & Kocurek (1999), and for a philosophical discussion Bokulich (2018); for dune dynamics see, for example, Kocurek et al. (2010) and Worman et al. (2013).

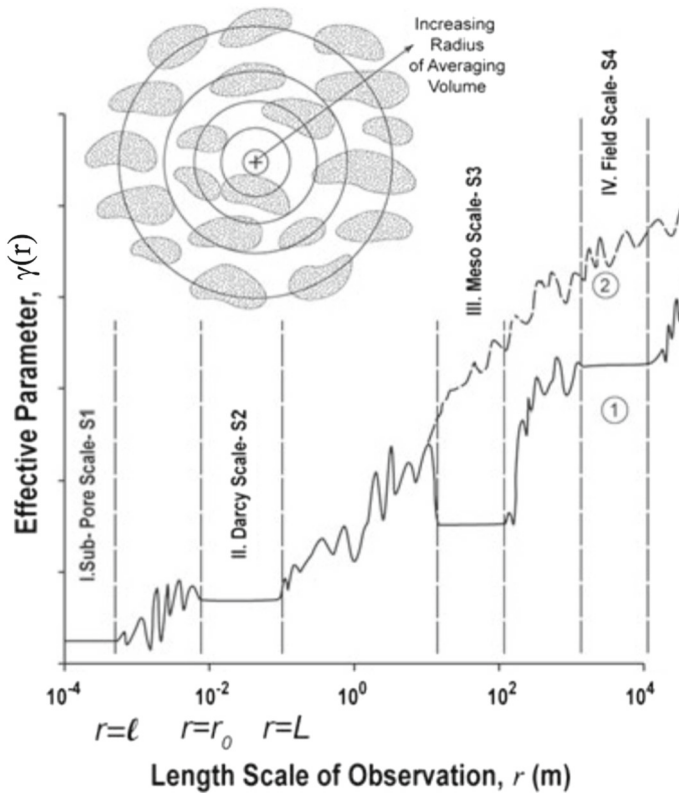


Fig. 5 Two different kinds of hierarchical or multiscale behavior for some quantity γ (e.g., porosity). Curve (1) is a classical discretely hierarchical system. Curve (2) shows a system with mixed behavior, exhibiting discretely hierarchical behavior at small scales and continuously evolving or fractal behavior at larger spatial scales. (From Scheibe et al. 2015, Fig. 2, with permission from John Wiley & Sons.)

pore, for example); as one increases the length scale (increasing r) variations in the measurement of γ will be recorded until finally a new threshold is reached, the Darcy scale, which is the smallest averaging volume where measurements of γ will again reach a constant value (this is called a representative elementary volume or REV). As one further increases the length scale, however, the measured value of γ will once again begin to fluctuate as one reaches the scale at which geological heterogeneity is sampled, and so on. The system described by curve (1) is discretely hierarchical, with a clear separation between the characteristic length scales.

Contrast this with a different system described by curve (2): this system has one regime where it is discretely hierarchical and another regime (at larger length scales) where it is continuously evolving. This kind of behavior or system, they explain,

has been recognized comparatively more recently, corresponding roughly with the discovery of fractal or power law structures in nature. . . . For such systems

there may be arbitrarily-long space and time correlation structures in the subsurface materials; thus, the behavior at any point in the system, may, in principle, be a function of the time-space behavior at all other points in the domain. (Scheibe et al., 2015, pp. 40–41)

For the sort of systems described by curve (2), conventional upscaling will not work and more sophisticated multiscale modeling methods will have to be used, such as multiscale hybrid models.

4.3 Many kinds of multiscale models

As we saw in the previous sections, many complex systems are plagued by the tyranny of scales problem, insofar as they have many critical processes simultaneously operating across a vast range of spatial and temporal scales. This multiscale behavior can thwart traditional modeling methods, requiring the development of more sophisticated multiscale models. A *multiscale* model can most broadly be defined as any model that attempts to represent a particular system at two or more scales within a single modeling or simulation framework. Multiscale models are, of course, to be contrasted with single-scale models, where one attempts to model a complex system by considering a single scale only, such as the most “fundamental” scale (as in reductionistic modeling) or even a single high-level scale. While single-scale models have been the norm throughout most of the history of scientific modeling, since the turn of the twenty-first century there has been an explosion of new multiscale modeling approaches. Multiscale models can involve representing multiple spatial scales, multiple temporal scales, or both—and they can invoke different, even strictly incompatible, theories to describe each of these different scales.

Because of the wide variety of multiscale behavior in nature, mathematical modelers need to develop a variety of different kinds of multiscale simulation models. One can distinguish different types of multiscale models depending on the nature and degree of coupling between the processes at different scales. Scheibe and colleagues have developed a multiscale analysis platform (MAP) that classifies the different multiscale simulation model approaches into a set of motifs. Before discussing more complex multiscale models, it is helpful to begin by thinking about cases where there is a separation of spatial or temporal scales.

As we saw in Sect. 2.3, within geomorphology and ecology, hierarchical models are defined as a special kind of multiscale model, where the component models are decoupled: “A model can be constructed across a broad range of temporal scales as a hierarchy of dynamically uncoupled models, ordered by characteristic time” (Werner, 1999, 103). Here we will refer to these as *discrete* hierarchical models. These sorts of models work when the macroscale phenomenon is insensitive to the details of the microscale processes (where ‘macro’ and ‘micro’ are being used to describe any two levels in a scale hierarchy). Philosophers sometimes describe these as situations where the microlevel is ‘screened-off’ from the macrolevel, the macrolevel is ‘multiply realized’ by the microlevel, or the macrolevel exhibits a kind of ‘universality’.²⁸

²⁸ There are of course subtle distinctions to be made among these terms, but they are not relevant here.

One can ask, however, what is it that actually allows one to replace the enormous number degrees of freedom at the microlevel with a smaller number of effective parameters at the macrolevel? The answer is that one must assume—or empirically determine—that certain physical facts hold, which Brian Wood calls *scaling postulates* (Wood, 2009; Wood & Valdés-Parada, 2013).²⁹ Scaling postulates allow you to filter out redundant—or low value—information. For example, in the familiar example of the ideal-gas model, the relevant scaling postulates are that (i) inter-particle potential energy can be neglected relative to kinetic energy, (ii) all valid microstates are equiprobable, (iii) statistics of the system are spatially homogeneous, and (iv) some kind of ergodic principle holds. As Wood explains, in “adopting these scaling [postulates] one can... reduce the roughly 54×10^{23} degrees of freedom down to essentially 1 degree of freedom (the pressure)” (Wood, 2009, p. 724). Turning to the hydrogeology example of mass transport in a heterogeneous porous media, there are six scaling postulates to consider: (1) influence of boundaries on the representative elementary volume (REV), (2) the statistical homogeneity of parameter fields, (3) separation of macro and micro length scales, (4) quasi-ergodicity, (5) smallness of the variance of the parameter fields, and (6) separation of the macroscopic and averaging volume length scales (Wood, 2009, p. 724). Importantly, scaling postulates are not just a mathematical trick of averaging, but rather require appropriate conceptual modeling of the system—both in terms of the empirical facts of the system and the purposes of the modeler.

When such scaling postulates do not hold, then one needs a multiscale model that allows for coupling between models at different scales. In some cases there may be a loose coupling between scales, in which case one might use a simple parameterization and upscaling. One example is the use of cloud-resolving models within the subgrid elements of a global circulation model (GCM) in climate science. Strictly speaking, however, “the local processes (e.g., cloud formation) depend on larger-scale driving forces... which often invalidates the closure approximations intrinsic to the parameterizations” (2015, p. 8). Instead of conventional parameterizations, which do not respond to changes in global forcing, one might use, for example, a 2D subgrid model of cloud physics to simulate (statistically sample) a portion of the 3D GCM grid space; this results in a ‘super-parameterization’ that allows for a two-way coupling between the different models (Randall et al., 2003).³⁰

More generally, when there is a tight coupling between a macro- and microscale, then one needs a proper multiscale model simulation. Recall that a multiscale model is defined as one that combines two or more models defined at different length scales—each typically having different ways of representing the relevant physical, chemical or biological processes—into a single simulation framework over the same spatio-temporal domain. Examples of using two different representations of the same target are discussed, for example, in my eikonic approach (Bokulich, 2018). What is novel in the context of multiscale modeling is that this plurality of representations

²⁹ In his earlier work, Wood calls these scaling ‘laws’, but his later term scaling ‘postulates’ is more apt and will be used here.

³⁰ As Randall et al. note, although this approach is 10^2 to 10^3 times more computationally expensive, than standard GCMs with conventional parameterizations, it is still less than a global cloud system resolving model (CSRMs) would be at 10^6 times more computationally expensive.

is being used simultaneously, such as when a single multiscale simulation combines both molecular dynamic and continuum mechanics representations. This underscores again the point made in Sect. 1 of seeing models not as depictions, but as more general representations.

As noted above, Scheibe et al. (2015) provide a nice taxonomy for capturing the different kinds of multiscale models depending on the nature and degree of spatial and temporal coupling, which we follow here. If the microscale conditions rapidly equilibrate to changes in the macroscale conditions, then there is an effective *temporal scale separation* that can be exploited. These multiscale models use short bursts of a microscale simulation to inform the macroscale simulation over the large time step. This approach can be further divided into two scenarios: those situations where there are formal equations defined at the macroscale (leading to a “top-down” multiscale approach) and those contexts where no such macroscale equations exist (requiring a “bottom-up” approach).

If on the other hand there is no temporal scale separation, such as when the microscale behavior equilibrates slowly relative to the timescales over which the macroscale conditions change, then one might look for ways to restrict the spatial domain of the microsimulation to just those limited locales where the conditions for decoupling break down. These lead to *concurrent multiscale models* that perform simultaneous microscale and macroscale simulations over different spatial subdomains, which are then linked with a “handshake” at the boundaries or areas of overlap. Winsberg (2010) provides an insightful discussion of these sorts of multiscale simulations and handshaking algorithms in the context of crack propagation in nanoscale solids. An alternative to the concurrent multiscale approach is the “gap-toothed” approach, with microscale simulations taking place in the teeth with boundary conditions derived from the macroscale solutions at the boundaries of each tooth. As they explain, “the gap-toothed method enables ‘large space, small time’ simulations through ‘small space, small time simulations’ (Scheibe et al., 2015, p. 11). This approach is useful when no macroscale equation exists.

My aim has not been to give a comprehensive taxonomy and discussion of all the different kinds of multiscale models, but rather to highlight that a wide diversity of multiscale modeling approaches in fact exists and continues to be developed. In all cases, it is important to have an adequate conceptual model of the system at all the functionally relevant scales for the phenomenon and purpose of interest. This conceptual modeling can involve representing systems in ways we know they are not (such as continuum mechanics representations of discrete systems) and can involve the use of different—and indeed strictly incompatible—theories and representations in the different levels of the hierarchy. In arriving at a model representation at any scale, there are different approaches one can take: a bottom-up reductionist modeling approach with a straight averaging or lumping, a top-down universality modeling, or a synthesist approach that relies on identifying relevant emergent structures when possible. Moreover, as we saw in detail, there are then substantive questions about how these multiple models at different scales are to be linked in a simulation or other theoretical framework, depending on the degree of coupling (or lack thereof) between the different spatial and temporal scales.

When modeling multiscale systems, it is important to remember that uncertainties can arise at every scale; these uncertainties should be quantified and communicated to the extent possible. Returning for a moment to the example of multiscale transport models in hydrogeology, Scheibe et al. note that at the pore scale (recall Fig. 5) uncertainty can arise from unknown pore geometry and rate constants; and at the Darcy scale “uncertainty can be due to variable properties of porous media (e.g., permeability, porosity, dispersion coefficients) and insufficient data” (2015, p. 13). Thus multiscale models can present more challenges to uncertainty quantification than single-scale modeling, though potentially involving less overall uncertainty. The topic of uncertainty quantification in modeling is an important one, worthy of more philosophical attention, but would take us beyond the scope of this paper.

5 Conclusion

The vast majority of work in the philosophy of scientific modeling has been focused on single-scale models, with only a handful of papers on multiscale modeling and the tyranny of scales problem. This problem is pervasive in the sciences, however, and poses formidable challenges, which scientists have only recently begun to learn how to address. Here I have focused on methods for taming the tyranny of scales in the geosciences, which must manage spatial scales ranging from grains of sand to entire planets, and temporal scales from seconds to billions of years. These scale challenges arise not only in mathematical modeling, but also in the context of conceptual and physical modeling.

As I argued in Sect. 1, one of the first steps in creating an adequate multiscale model is recognizing that models need not be realistic depictions, but rather should be thought of as representational tools. Multiscale models often need to invoke different theories with different ontologies at different levels in the scale hierarchy; hence they typically involve different, and even strictly incompatible, model representations of their target. For example, multiscale models can combine both molecular dynamics and continuum mechanics representations. Freeing oneself from the veridical-depiction view of models can thus allow one to develop more successful multiscale models.

In Sect. 2, I called attention to the critical role that conceptual models play in scientific practice, whether they are ends-in-themselves or a prelude to a physical or mathematical model. I argued that even at the level of conceptual modeling, there are several strategies one can adopt to help tame the tyranny of scales. The first is paying attention to the various dynamical thresholds in the target system; second, utilizing hierarchy theory to identify the functionally relevant scales; and third, tailoring the multiscale model to be adequate for particular purposes. The intended purpose will determine which levels, features, and processes are essential to include in the model, and which can be neglected.

When it comes to the context of physical modeling, the strategies become more complex. In this context the tyranny of scales problem typically manifests itself as scale effects, which arise because the force ratios in the model are not the same as the ratios in the target system. The strategy first designed to address this problem is the formal theory similitude, or dynamical scaling, which gives the instructions for designing

a physical model whose force ratios match those in nature. Strict dynamical scaling is often difficult—if not impossible—to achieve, and hence other strategies must be deployed. Here we examined three strategies for addressing incomplete dynamic similitude: avoidance, correction, and compensation. Yet a fourth strategy involves identifying those aspects of natural phenomena that are scale independent within a certain range, and I discussed three ways this scale independence can be achieved: self-similarity, convergent physics, and decoupling of scales.

Finally, we turned to mathematical modeling in Sect. 4. Here we defended the synthesist approach to climbing the scales, which involves reducing the degrees of freedom by identifying the relevant emergent structures and variables that arise from the collective behaviors of the smaller-scale degrees of freedom, rather than simply lumping or averaging. These emergent structures are useful in identifying the functionally relevant scales to include in the multiscale model, in accordance with hierarchy theory. A central lesson here is that there are many different kinds of multiscale behavior in nature, which can involve hierarchies in both spatial and temporal dimensions, with a variety of dependence relations between the levels. Thus an adequate conceptual model, given both the nature of the system and a particular purpose, is a prerequisite for building a successful mathematical multiscale model.

Geoscience research, ranging from hydrogeology to geomorphology, has been at the forefront of multiscale-modeling efforts to develop the conceptual, physical, and mathematical tools needed to tame the tyranny of scales. Although the examples discussed here are drawn from the geosciences, many of these methodological strategies are widely applicable to other sciences. In particular, I have tried to show how this work gives us deeper insight into what it means to be a multiscale system, involving different behaviors at different scales, and how it reveals the variety of possible dependencies (or lack thereof) between those various scales. These insights also have implications for traditional questions in metaphysics related to reduction and emergence, which, though outside the scope of this paper, have their foundation laid here. Finally, this paper shows that the project of taming the tyranny of scales is not quite as hopeless as one might have first supposed.

Acknowledgements The research for this paper began while I was a visiting researcher in the Earth and Climate Sciences Program at Duke University. I would like to express my deep gratitude to Brad Murray and the other researchers there for providing such a stimulating and welcoming environment in which to do the philosophy of geosciences and for helpful continuing discussions. I am also grateful to the participants at the Active Matter 2 Workshop at Georgetown University where this material was first presented.

References

- Allen, T., & Starr, T. (1982). *Hierarchy: Perspectives for ecological complexity*. Chicago University Press.
- Barrenblatt, G. (2003). *Scaling. Cambridge Texts in Applied Mathematics*. Cambridge University Press.
- Batterman, R. (2013). The tyranny of scales. In R. Batterman (Ed.), *The Oxford handbook of philosophy of physics* (pp. 255–286). Oxford University Press.
- Batterman, R., & Rice, C. (2014). Minimal model explanations. *Philosophy of Science*, 81(3), 349–376.
- Bokulich, A. (2013). Explanatory models versus predictive models: Reduced complexity modeling in geomorphology. In V. Karakostas & D. Dieks (Eds.), *EPSA11 Perspectives and Foundational 115 Problems in Philosophy of Science, The European Philosophy of Science Association Proceedings*. (Vol. 2). Springer. https://doi.org/10.1007/978-3-319-01306-0_10

- Bokulich, A. (2015). Maxwell, Helmholtz, and the unreasonable effectiveness of the method of physical analogy. *Studies in History and Philosophy of Science*, 50, 28–37.
- Bokulich, A. (2016). Fiction as a vehicle for truth: Moving beyond the ontic conception. *The Monist*, 99(3), 260–279.
- Bokulich, A. (2018). Representing and Explaining: The Eikonic Conception of Explanation. *Philosophy of Science*, 85(5), 793–805.
- Bokulich, A., & Oreskes, N. (2017). Models in the geosciences. In L. Magnani & T. Bertolotti (Eds.), *Handbook of model-based science* (pp. 891–911). Springer.
- Bokulich, A. (2020). Towards a taxonomy of the model-ladenness of data. *Philosophy of Science*, 87(5): 793–806. <https://doi.org/10.1086/710516>.
- Bokulich, A., & Parker, W. (2021). Data models, representation, and adequacy-for-purpose. *European Journal for the Philosophy of Science*, 11(31), 1–26. <https://doi.org/10.1007/s13194-020-00345-2>
- Borland, W. (1960). *Stream channel stability*. U.S. Bureau of Reclamation.
- Chalmers, D. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The re-emergence of emergence* (pp. 244–256). Oxford University Press.
- Church, M. (2002). Geomorphic thresholds in riverine landscapes. *Freshwater Biology*, 47, 541–557.
- Church, M. (2017). Geomorphic thresholds. In D. Richardson, N. Castree, M. Goodchild, A. Kobayashi, L. Weidong, & R. Marston (Eds.), *The International Encyclopedia of Geography*. Wiley.
- Cowell, P., Stive, M., Niedoroda, A., de Vriend, H., Swift, D., Kaminsky, G., & Capobianco, M. (2003). The coastal-tract (part 1): A conceptual approach to aggregated modeling of low-order coastal change. *Journal of Coastal Research*, 19(4), 812–827.
- Cuomo, G., Allsop, W., & Takahashi, S. (2010). Scaling wave impact pressures on vertical walls. *Coastal Engineering*, 57, 604–609.
- Currie, A. (2017). From models-as-fictions to models-as-tools. *Ergo*, 4(27), 759–781. <https://doi.org/10.3998/ergo.12405314.0004.027>
- Downes, S. (2009). Models, pictures, and unified accounts of representation: Lessons from aesthetics for philosophy of science. *Perspectives on Science*, 17(4), 417–428.
- French, J., Payo, A., Murray, B., Orford, J., Eliot, M., & Cowell, P. (2016). Appropriate complexity for the prediction of coastal and estuarine geomorphic behaviour at decadal to centennial scales. *Geomorphology*, 256, 3–16.
- Frigg, R., & Hartmann, S. (2020). Models in science. In Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition). <https://plato.stanford.edu/archives/spr2020/entries/models-science>.
- Frigg, R., & Nguyen, J. (2018). The turn of the valve: Representing with material models. *European Journal for the Philosophy of Science*, 8, 205–224.
- Galilei, G. ([1638] 1991). *Dialogues concerning two new sciences [Discourses and Mathematical Demonstrations Relating to Two New Sciences]*. Buffalo, NY: Prometheus Books.
- Giere, R. (1999). Using models to represent reality. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery*. New York: Kluwer Academic.
- Giere, R. (2001). The nature and function of models. *Behavioral and Brain Sciences*, 24(6), 1060.
- Gould, S. (1980). The promise of paleobiology as a nomothetic, evolutionary discipline. *Paleobiology*, 6(1), 96–118.
- Grant, G. E., O'Connor, J. E., & Wolman, M. G. (2013). A river runs through it: Conceptual models in fluvial geomorphology. In J. Shroder & E. Wohl (Eds.), *Treatise on geomorphology* (Vol. 9, pp. 6–21). Academic Press.
- Green, S., & Batterman, R. (2017). Biology meets physics: Reductionism and multi-scale modeling of morphogenesis. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 61, 20–34.
- Haldane, J. B. S. ([1926] 1985). “On being the right size” *Harper’s Magazine* (March). Reprinted in J. Maynard Smith (ed.) *On Being the Right Size and Other Essays*. Oxford: Oxford University Press.
- Heller, V. (2011). Scale effects in physical hydraulic engineering models. *Journal of Hydraulic Research*, 49(3), 293–306.
- Heller, V., Hager, W., & Minor, H.-E. (2008). Scale effects in subaerial landslide generated impulse waves. *Experiments in Fluids*, 44, 691–703. <https://doi.org/10.1007/s00348-007-0427-7>
- Hubbert, M. K. (1989). Interview of M. King Hubbert by Ronald Doel on 1989 January 20 (Session V), Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA. www.aip.org/history-programs/niels-bohr-library/oral-histories/5031-5.

- Hubbert, M. K. (1937). Theory of scale models as applied to the study of geologic structures. *Geological Society of America Bulletin*, 48(10), 1459–1520. <https://doi.org/10.1130/GSAB-48-1459>
- Hubbert, M. K. (1945). Strength of the earth. *Bulletin of the American Association of Petroleum Geologists*, 29(11), 1630–1653.
- Knuutila, T. (2005). Models, representation, and mediation. *Philosophy of Science*, 72, 1260–1271.
- Kocurek, G., Ewing, R., & Mohrig, D. (2010). How do bedform patterns arise? New views on the role of bedform interactions within a set of boundary conditions. *Earth Surface Processes and Landforms*, 35, 51–63.
- Lajeunesse, E., Malverti, L., Lancien, P., Armstrong, L., Métivier, F., Coleman, S., Smith, C., Davies, T., Cantelli, A., & Parker, G. (2010). Fluvial and submarine morphodynamics of laminar and near laminar flows: A synthesis. *Sedimentology*, 57, 1–26.
- Lane, E. (1955). The importance of fluvial morphology in hydraulic engineering. *Proceedings of the American Society of Civil Engineers*, 81(Paper No. 745), 1–17.
- Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, 54(4), 421–431.
- Massimi, M. (2019). Two kinds of exploratory models. *Philosophy of Science*, 86(5), 869–881.
- McGarigal, K. (2018). “Concepts of Scale” Lecture Notes for Landscape Ecology Course at University of Massachusetts, Amherst. Version downloaded June 28th, 2018. http://www.umass.edu/landeco/teaching/landscape_ecology/schedule/chapter2_scale.pdf.
- Murray, A. B. (2013). Which models are good (enough), and when? In: *Treatise on geomorphology* (Editor-in-Chief: JF Shroder). Oxford: Elsevier-Academic Press. ISBN 0123747392.
- Murray, A. B. (2003). Contrasting the goals, strategies, and predictions associated with simplified numerical models and detailed simulations. *Prediction in Geomorphology, Geophysical Monograph*, 135, 1–15.
- Murray, A. B. (2007a). Reduced model complexity for explanation and prediction. *Geomorphology*, 90, 178–191.
- Murray, A. B. (2007b). Two paradigms in landscape dynamics: Self-similar processes and emergence. In A. A. Tsonis & J. B. Elsner (Eds.), *Nonlinear dynamics in geophysics* (pp. 17–35). Springer.
- Murray, A. B., Lazarus, E., Ashton, A., Baas, A., Coco, G., Coulthard, T., Fonstad, M., Haff, P., McNamara, D., Paola, C., Pelletier, J., & Reinhardt, L. (2009). Geomorphology, complexity, and the emerging science of the earth’s surface. *Geomorphology*, 103, 496–505.
- Nguyen, J. (2020). It’s not a game: Accurate representation with toy models. *British Journal for the Philosophy of Science*, 71, 1013–1041. <https://doi.org/10.1093/bjps/axz010>
- Norton, S., & Suppe, F. (2001). Why atmospheric modeling is good science. In P. Edwards & C. Miller (Eds.), *Changing the atmosphere: Expert knowledge and environmental governance* (pp. 67–106). MIT Press.
- Oden, J. T. (Chair) (2006). *Simulation based engineering science—An NSF Blue Ribbon Report*. www.nsf.gov/pubs/reports/sbes_final_report.pdf.
- Paola, C. (2001). Modelling stream braiding over a range of scales, and reply. *New Zealand Hydrological Society* (pp. 11–38, 42–46). Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/164368>.
- Paola, C., Straub, K., Mohrig, D., & Reinhardt, L. (2009). The unreasonable effectiveness of stratigraphic and geomorphic experiments. *Earth Science Reviews*, 97, 1–43.
- Parker, W. (2010). Scientific models and adequacy-for-purpose. *Modern Schoolman*, 87(3–4), 285–293.
- Parker, W. (2020). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science*, 87(3), 57–77.
- Peterson, M. (2002). Galileo’s discovery of scaling laws. *American Journal of Physics*, 70(6), 575–580.
- Pincock, C. (2020). Concrete scale models, essential idealization, and causal explanation. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axz019>.
- Potochnik, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science*, 79, 120–140.
- Randall, D. M., Khairoutdinov, A. A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. *American Meteorological Society*, 84(11), 1547–1564. <https://doi.org/10.1175/BAMS-84-11-1547>
- Rueger, A., & McGivern, P. (2010). Hierarchies and levels of reality. *Synthese*, 176, 379–397.
- Scheibe, T., & J. Smith. (2015). *Multiscale computation: Needs and opportunities for BER science*. Report on the Workshop held August 26, 2014. Pacific Northwest National Laboratory.
- Sapozhnikov, V., & Fofoula-Georgiou, E. (1996). Self-affinity in braided rivers. *Water Resources Research*, 32(5), 1429–1439.

- Sapozhnikov, V., & Foufoula-Georgiou, E. (1997). Experimental evidence of dynamic scaling and indications of self-organized criticality in braided rivers. *Water Resources Research*, 33(8), 1983–1991.
- Scheibe, T., Murphy, E., Chen, X., Rice, A., Carroll, K., Palmer, B., Tartakovsky, A., Battiato, I., & Wood, B. (2015). An analysis platform for multiscale hydrogeologic modeling with emphasis on hybrid multiscale methods. *Groundwater*, 53(1), 38–56.
- Schumm, S. (1979). Geomorphic thresholds: The concept and its applications. *Transactions of the Institute of British Geographers*, 4(4), 485–515.
- Schumm, S., & Lichty, R. (1965). Time, space, and causality in geomorphology. *American Journal of Science*, 263, 110–119.
- Shech, E. (2016). Fiction, depiction, and the complementarity thesis in art and science. *The Monist*, 99, 311–332.
- Sterrett, S. (2017a). Physically similar systems—A history of the concept. In L. Magnani & T. Bertolotti (Eds.), *Handbook of model-based science* (pp. 377–410). Springer.
- Sterrett, S. (2017b). Experimentation on analog models. In L. Magnani & T. Bertolotti (Eds.), *Handbook of model-based science* (pp. 857–878). Springer.
- Sterrett, S. (2021). Scale Modeling. In D. Michelfelder & N. Doorn (Eds.), *The Routledge Handbook of the philosophy of engineering* (pp. 394–407). Routledge.
- Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. Oxford University Press.
- Werner, B. (1999). Complexity in natural landform patterns. *Science*, 284, 102–104.
- Werner, B., & Kocurek, G. (1999). Bedform spacing from defect dynamics. *Geology*, 27, 727–730.
- Wilson, M. (2017). *Physics Avoidance: Essays in Conceptual Strategy*. Oxford University Press.
- Wimsatt, W. ([1987] 2007). False models as means to truer theories. In *Re-engineering philosophy for limited beings: Piecewise approximations to reality* (pp. 94–132). Cambridge, MA: Harvard University Press.
- Winsberg, E. (2010). *Science in the age of computer simulation*. Chicago University Press.
- Wood, B. (2009). The role of scaling laws in upscaling. *Advances in Water Resources*, 32, 723–736.
- Wood, B., & Valdés-Parada, F. (2013). Volume averaging: Local and nonlocal closures using a green's function approach. *Advances in Water Resources*, 51, 139–167.
- Woodward J. (2021). Explanatory autonomy: The role of proportionality stability and conditional irrelevance. *Synthese*, 198(1): 237–265. <https://doi.org/10.1007/s11229-018-01998-6>.
- Worman, S., Murray, A. B., Littlewood, R., Andreotti, B., & Claudin, P. (2013). Modeling emergent large-scale structures of barchans dune fields. *Geology*, 41, 1059–1062.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.