

Scaling methods in ecological modelling

Matthias Fritsch¹ | Heike Lischke² | Katrin M. Meyer¹ 

¹Ecosystem Modelling, Faculty of Forest Sciences and Forest Ecology, University of Göttingen, Göttingen, Germany

²Dynamic Macroecology, Land Change Science, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

Correspondence

Katrin M. Meyer
Email: kmeyer5@uni-goettingen.de

Funding information

Deutsche Forschungsgemeinschaft

Handling Editor: David Murrell

Abstract

1. Modelling is often confronted with scaling problems, because modelling—directly or indirectly—always implies scaling. This is because models simplify. Simplification usually means aggregation, and aggregation is a scaling process. As scaling cannot be avoided in modelling, it should carefully be addressed and resolved, at least to the degree possible.
2. In this paper, we give an overview of scaling approaches in ecological modelling. We propose to classify scaling approaches into pre-model scaling, in-model scaling and post-model scaling depending on the timing of the scaling relative to the main modelling process.
3. We show general approaches, examples and potential application problems for each category. We suggest that scaling problems might be more widespread than previously thought. These scaling problems are matched with a range of solutions, but often these solutions will have to be adapted and tailored to the specific scaling case.
4. Thus, we recommend that ecologists be aware of scaling challenges especially where models do not explicitly aim at scaling. Developing general test systems for scaling methods may help to broaden and enhance the application of scaling methods in ecology.

KEYWORDS

aggregation, ecological model, levels of scale, meta-model, pattern-process relationship, scale transition theory, scaling-down, scaling-up

1 | INTRODUCTION

Scaling is ubiquitous and persistent in ecology. Following the acclaimed concept of pattern and scale (Levin, 1992), patterns and processes at a certain spatial or temporal scale or organizational level emerge from patterns and processes at finer scales or levels and these, in turn, are influenced by the large-scale patterns (Figure 1, Lischke, Löffler, Thornton, & Zimmermann, 2007). Due to this, scaling, that is, changing from one scale to another, is not always straightforward, and sometimes can cause problems due to scale

breaks, nonlinearities, feedbacks and heterogeneity in such pattern-process relationships (Snell et al., 2014). Additionally, scaling is sometimes not explicit, and confusion in terminology adds to scaling-related problems. Here, we (1) address scaling terminology; (2) define three categories of scaling approaches, (a) pre-model scaling, (b) in-model scaling and (c) post-model scaling; and (3) explore examples, problems, and, where available, potential solutions in each category. We also elaborate our main claim that modelling is often confronted with scaling challenges, because modelling—directly or indirectly—always implies scaling.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society

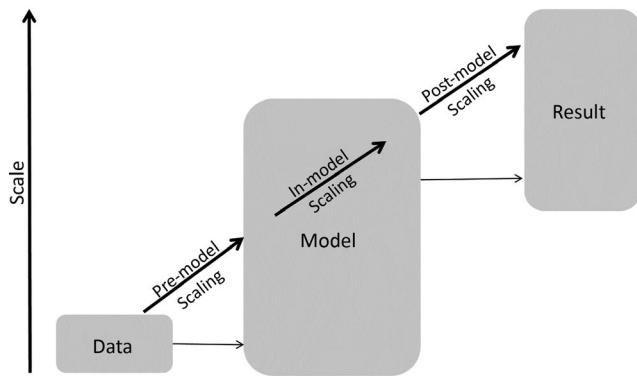


FIGURE 1 Concept of the relationship between models and scaling (bold arrows) in the context of scaling-up. Models are scaling methods in themselves, because they aggregate data to answer questions. We call this in-model scaling. Pre-model scaling methods can precede the actual model where model parameterization is based on statistical models of empirical data. Post-model scaling translates model results from one scale to another and reflects what traditionally is considered as scaling. Input data can also operate on the same scale as models and models can also produce same-scale results (thin arrows)

The term 'scale' has several meanings in different scientific fields. Moreover, scale is not a property of the observed system, but of the observer, so that the observed systems may appear different at different scales, although the system does not change. In the ecological context, scale usually refers to the combination of extent and grain. Extent is the overall area of space or overall duration in time of the study, divided into sub-units with dimensions defined by the grain. Sub-units with small dimensions have a fine grain (also 'fine scale'), and sub-units with large dimensions have a coarse grain (also 'coarse scale'). Ecological scales are usually explicit (spatially and or temporally) and thus can be quantified with SI units [e.g. time in (s) or area in (m^2)]. On the contrary, hierarchical levels or entities such as individual, population and community or local, regional and landscape usually cannot be converted, combined or divided in the same way. Unfortunately, hierarchical or organizational levels and entities are also referred to as 'scales'. This is a common source of terminological confusion yielding imprecise descriptions such as 'at the landscape scale' instead of an exact scale quantification in terms of extent and grain or simply 'at the landscape level'. Furthermore, different scientific fields accept different terminologies, which may prevent agreement on a standard definition (Schneider, 2009) and hinder interdisciplinary collaborations. For example, cartographers use 'large scale' to describe maps with small extent (e.g. with a side length of 1 km), whereas ecologists often use 'large scale' to refer to large extents of areas (e.g. with a side length of 1,000 km). Here, we follow the ecologists' approach and use large scale and small scale to refer to large and small extents if not indicated otherwise.

Scaling or process scaling is defined as translating information from one scale to the other. These scales are called source and target scale and differ in grain, extent or both. The direction of the information transfer defines whether it is scaling-up or scaling-down.

In ecology, scaling is inevitable because ecological studies typically involve a range of scales. Ecosystem patterns and processes cover a wide range of space and time and often have multiple drivers that act at different scales.

Scaling problems arise basically from the loss of fine-scale information involved in scaling-up the grain or scaling-down the extent and the lack of information involved in scaling-up the extent or scaling-down the grain (Figure 2). They are reinforced by the complexity of the involved systems or models. In scaling-up, problems can be caused by fine-scale heterogeneity, unknown factors, temporally changing relationships, interactions, nonlinearity and feedbacks, the latter three leading to scale dependence (partly reviewed by Lischke et al., 2007; Snell et al., 2014). These problems can originate at the fine scale, at the coarse scale, or in between. Fine-scale problems are due to heterogeneity or interactions. These scaling problems occur if the quantity to scale is not proportional to the ratio of area in spatial scaling or the time step ratio in temporal scaling. This means that proportionality to these scale ratios is only given if the process forming the quantity is homogeneous over the scale range and no interactions distort the proportionality. There are a few cases where this proportionality to scale ratios holds true, for example, for allometric relationships in tropical forests (Chave et al., 2005). Allometric relationships include nonlinear relationships that are still monotonic so that scaling problems may be overcome through the derivation of power equations, with further examples ranging from body-size allometry to species-area curves (Miller, Turner, Smithwick, Dent, & Stanley, 2004). However, most real situations are more complex than this (e.g. Muller-Landau et al., 2006) and span several domains of scale, which are characterized by abrupt changes or critical thresholds in the target process. Scaling via power laws (and in fact via most available techniques) is ill-advised across domains of scale. Here, inappropriate scaling equations would result in an aggregation error. At the coarse scale, unknown factors may play a role that were not considered important enough to be included at the fine scale. Scaling-up without considering these factors may thus lead to erroneous coarse-scale results. The same holds for scaling-down, where, for example, microclimate can have a huge effect not visible at the coarse scale. Between the scales, scaling problems arise from feedbacks between processes or patterns at different levels and from scale-dependent processes. Scale dependence occurs where a process or pattern is significantly different at different scales due to local interactions and feedbacks. Such processes at broader scales emerge from the finer-scale processes. For example, following the gap-dynamics theory, the succession in a natural forest stand emerges from the stochastic dynamics of tree groups on many small patches, which show very different trajectories (Botkin, Janak, & Wallis, 1972).

Modelling is often confronted with scaling challenges, because scaling is an integral part of all modelling; otherwise, every process would have to be described on the spatial, temporal and organizational scale of sub-atomic particles. Here, we broadly define models as purposeful simplifications, including conceptual, statistical,

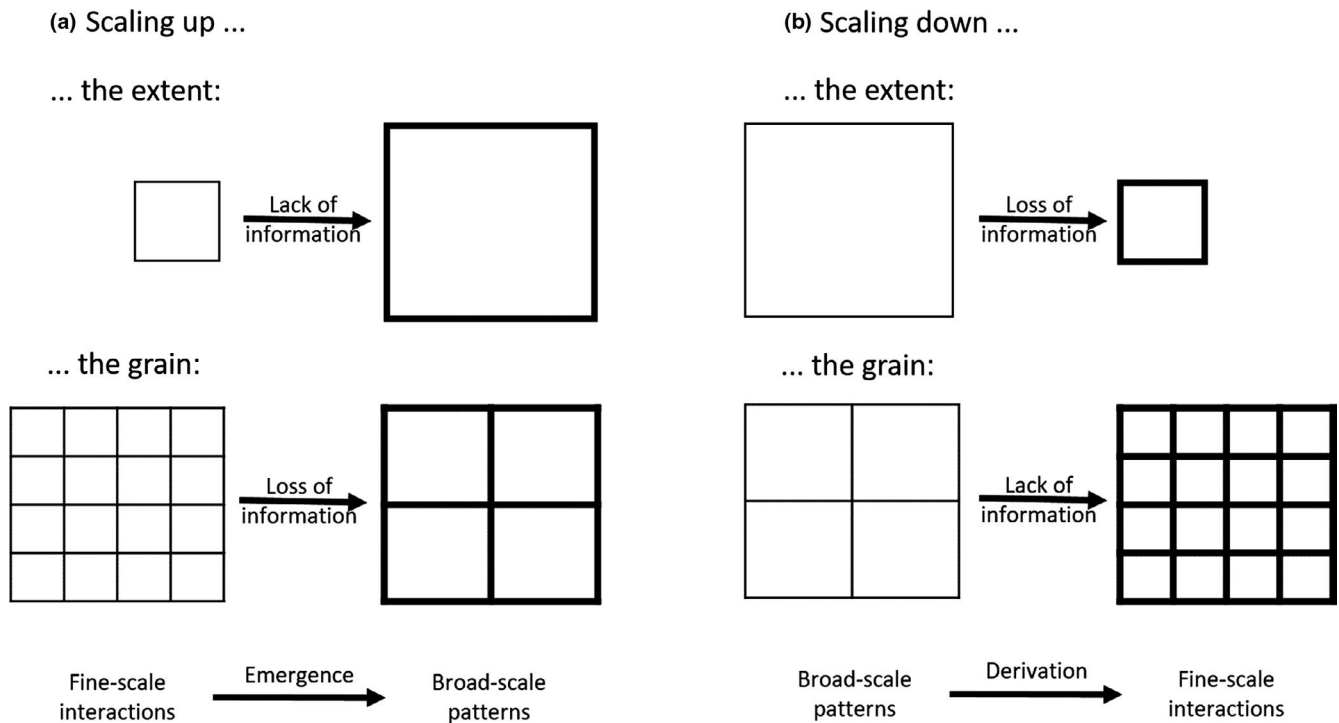


FIGURE 2 (a) During scaling-up of the extent, there may be a lack of information regarding areas not covered at the fine scale. Scaling-up of fine grain to coarse grain requires aggregation and thus loss of information. Scaling-up methods can directly (e.g. via individual-based models) or indirectly exploit the emergence of broad-scale patterns from fine-scale interactions. (b) Scaling-down of the extent involves loss of information about the areas not anymore covered by the model. Scaling-down of the grain may require fine-grain information that is not available. Lacking fine-scale information such as fine-scale interactions can sometimes be derived from broad-scale patterns

equation-based, process-based, rule-based, simulation-based or agent-based models. Essentially, models have inputs (data, parameters), do something (calculation, simulation) and then produce an output (values, statistics, patterns). Even in the simple case of a statistical model that is calculating the mean value of a sample, the scale or level of the output differs from the scale or level of the input because the model provides aggregated information about the sample. We call this scaling in the broad sense, whereas scaling in the strict sense is translating a result obtained at one scale to another scale (e.g. Barraquand & Murrell, 2013; Denny & Benedetti-Cecchi, 2012; Englund & Leonardsson, 2008; Johst, Lima, & Berryman, 2013). In summary, models, by definition, simplify and simplification is a scaling process.

Scaling can, thus, not be avoided in modelling and should carefully be addressed, at least to the degree possible. Scaling may occur at different stages during the modelling process (Figure 1). We have pointed out that models in themselves are scaling procedures. Such 'in-model scaling' can also be done on purpose, for example, for a controlled simplification (Lischke, Löffler, & Fischlin, 1998), or for including finer-scale processes, for example, hourly transpiration in a vegetation model acting on a yearly time step (Speich, Zappa, & Lischke, 2018; Speich, Zappa, Scherstjanoi, & Lischke, 2019). However, scaling can also happen before and after the actual modelling process. For example, when models are parametrized with empirical data, the parameter values usually are the result of statistical models of these data such as slopes and intercepts of a linear

regression model. Since statistical analysis can be an aggregation and thus a scaling process, we suggest to call this 'pre-model scaling'. Pre-model scaling includes any scaling that is done prior to running a model, including the data collection process. It can involve the use of other models, such as statistical models in the example, or climatic relationships in statistical climate downscaling (Karger et al., 2017). 'Post-model scaling' then captures scaling based on model outputs, that is, scaling in the strict sense. Neither pre-model nor post-model scaling are essential, because a model can do without empirical input data if it addresses theoretical questions or without outputs scaled up in space or time if outputs at the spatial and temporal scales of the model are sufficient to answer the model question. We think that the classification into pre-model, in-model and post-model scaling methods facilitates revealing hidden scaling procedures and disentangling some of the confusion and ambiguity in current treatments of scaling methods. Where the application of this classification reveals hidden scaling methods, it becomes apparent that scaling methods and thus scaling problems might be more widespread than previously thought. We assigned some commonly used scaling techniques to those new categories, describing the purpose of the scaling methods and whether they accept heterogeneous input, visualizing their main principle and giving at least one example application from the literature for each of them (Table 1).

In the following, we will specify scaling problems and examples for approaches from pre-model scaling via in-model scaling to post-model scaling. Finally, we will make some practical

TABLE 1 Overview of scaling methods in the categories pre-model scaling, in-model scaling and post-model scaling (see Figure 1), together with their main purpose and their type of input and including a visualization of the principle of the method and one or more example studies addressing the method (see Supporting Information for enlarged versions of the figures)


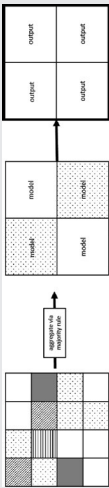
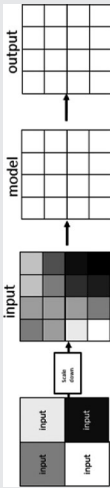
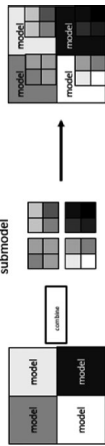
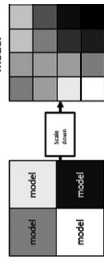
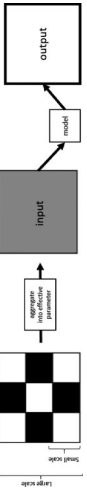

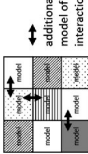
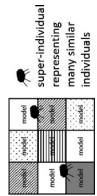
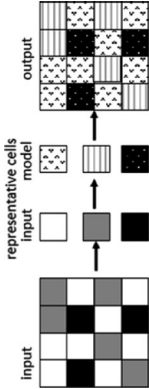
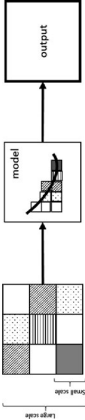
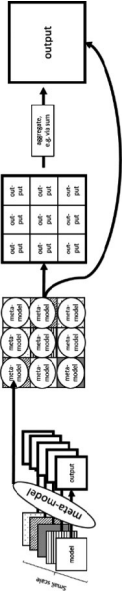
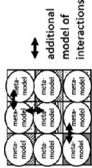

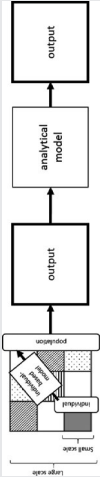

Scaling method	Purpose	Input ^a	Visualization ^b	Examples
No scaling	To create a model with input and output at the same scale ^c	Het		Only as 'null' reference
Pre-model scaling				
Statistical models	To aggregate values to make them usable as model inputs	n.a.	- No spatial visualization possible -	Simple averaging, linear and nonlinear regression modelling, Bayesian analysis, generalized linear mixed effects models
Coarsening gridded maps	To aggregate spatial information	Het		Majority rule, simple averaging, area-weighted mean, kriging
Refining gridded maps	To derive finer-scale spatial patterns	Het		Interpolation, statistical scaling-down, with(out) additional information (Karger et al., 2017)
In-model scaling				
Process refinement	To simulate certain processes on a finer resolution			Smith, Prentice, and Sykes (2008)
State variable refinement	To apply aggregated model to finer-scale state variables	Het		Hickler et al. (2012)
Effective parameters	To derive large-scale parameters directly from small-scale parameters to use as input to small-scale model	Non-het		Wu, Jones, Li, and Loucks (2006)
Brute force—classic	To run copies of the small-scale model in each cell of a large-cell grid	Het		Peters et al. (2004)
Brute force—with interactions	To add a model of interactions between cells	Het		Cipriotti, Wiegand, Pütz, Bartoloni, and Paruelo (2016)
Brute force—with super-individuals	To aggregate individuals of similar age or size into super-individuals that are used as entities in large-scale model	Het		Xavier, Grose, and Whetton (2007)
(Continues)				

TABLE 1 (Continued)

Scaling method	Purpose	Input ^a	Visualization ^b	Examples
Representative cells	To run the simulation only once for each cell representing many cells with same conditions	Het		Lehsten, Mischurow, Lindström, Lehsten, and Lischke (2019) and Nabel (2015)
Scale transition theory	To mathematically or heuristically derive larger-scale processes from small-scale processes, interactions and heterogeneity, for example, to use nonlinear averaging to scale-up nonlinear population growth	Het		Bolker and Pacala (1997) and Melbourne and Chesson (2005)
Meta-model—classic	To fit a meta-model (dynamic, statistical, or neural network) to the results of running a small-scale model at a range of relevant conditions and only applying the simpler meta-model to obtain large-scale results; in contrast to analytical approximations (see post-model scaling), the small-scale model is only of interest as basis for scaling-up	Het		Acevedo, Ablan, Urban, and Pamarti (2001), Acevedo, Urban, and Ablan (1995), Acevedo, Urban, and Shugart (1996), Meier, Lischke, Schmatz, and Zimmermann (2012), Rammer and Seidl (2019) and Urban, Acevedo, and Garman (1999)
Meta-model—with interactions	To add an interaction model to the meta-model	Het		ZelStage model (Urban et al., 1999)
Post-model scaling				
Approximation by analytical model in general	To approximate small-scale model results by an analytical model and apply only analytical model at large extent; similar to meta-models, but applied to models whose original aim did not include scaling-up	Non-het		Moorcroft, Hurtt, and Pacala (2001)
Approximation by analytical model for individual-based models	To approximate large-scale results of individual-based models with analytical model and only apply analytical model in other large-scale scenarios	Het		Fahse, Wissel, and Grimm (1998) and Johst et al. (2013)
Spatial statistics	To combine scale-dependent pair correlation functions and species abundance distributions to predict species numbers across scales via species–area curves	Non-het		Azeale et al. (2015)

^aType of model input: Het—spatially heterogeneous, Non-het—no spatially heterogeneous input possible.
^bVisualizations show only examples. For simplicity, visualizations refer to spatial scaling only. This can be transferred analogously to temporal and hierarchical scaling.
^cThis is only theoretical, because in practice, according to our claim in this paper, all modelling implies scaling.

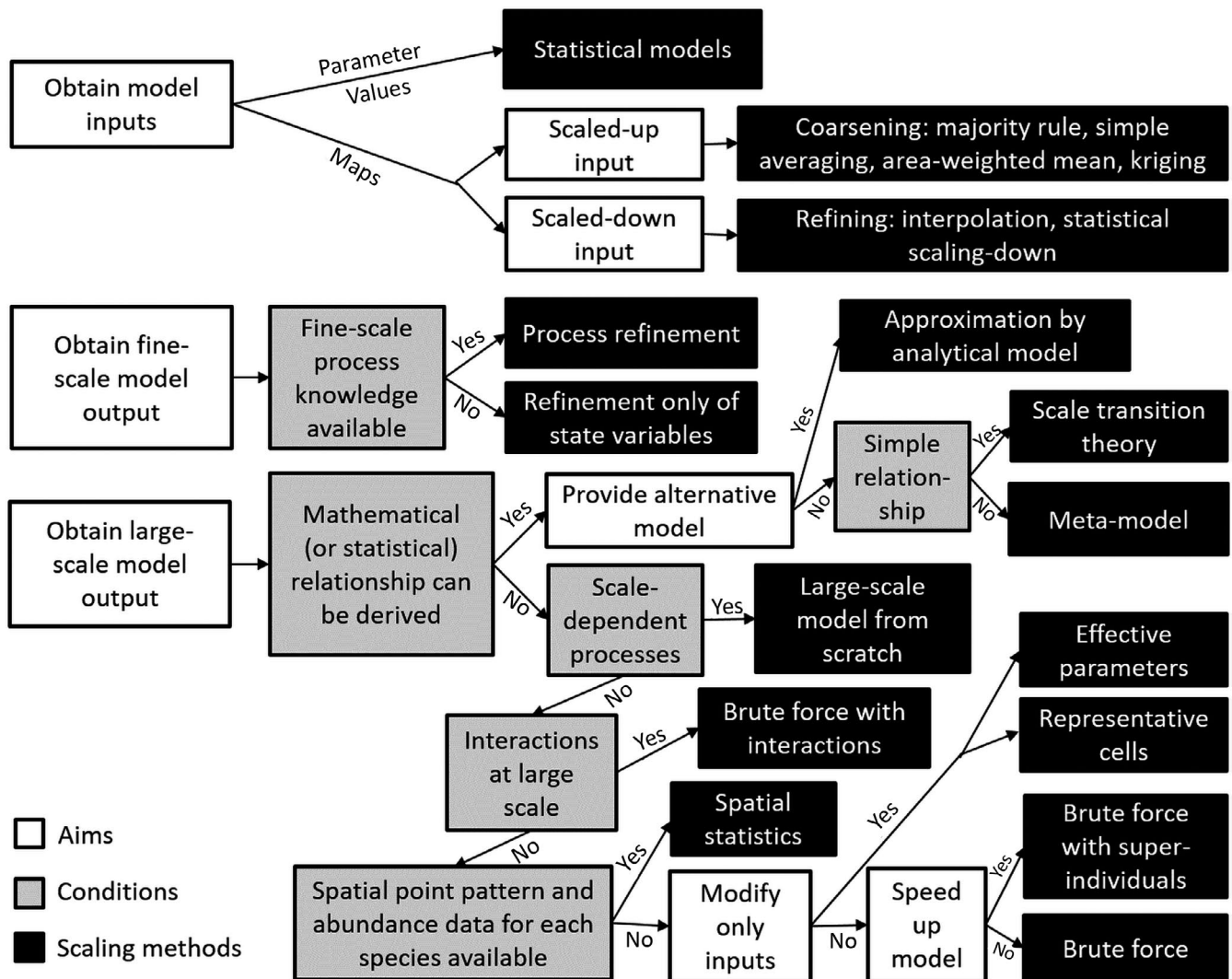


FIGURE 3 Suggestion on an identification key for scaling approaches (black boxes) based on the aims of the modeller (white boxes) and conditions set by the available data and knowledge (grey boxes). For example, if a modeller would like to obtain large-scale model output (i.e. do scaling-up), she first checks whether she can (or wants to) derive a mathematical or statistical relationship; if yes and she is not interested to provide an alternative model formulation for an existing large-scale model, but use the relationship to do the actual scaling-up, she might select scale transition theory if the relationship is simple, i.e. without extra assumptions (see also main text and Table 1)

recommendations to guide the application of scaling approaches (see also Figure 3) and point out theoretical avenues towards minimizing some related problems in scaling.

2 | SCALING APPROACHES ALONG THE MODEL CHAIN

2.1 | Pre-model scaling

Pre-model scaling refers to aggregation or disaggregation of values to make them usable as model inputs. This (dis)aggregation is not necessarily related to a scaling procedure across space or time (note that if parameters were aggregated across spatial or temporal scales, the result would be called an effective parameter and the procedure would fall into the category in-model scaling). Pre-model scaling

often involves statistical modelling if parameter values are determined by parameterization from detailed empirical data instead of by calibration with coarse empirical data (e.g. via pattern-oriented modelling; Grimm et al., 2005). Parameters of a model usually have a standard or reference value that needs to be extracted from empirical data. These data can be obtained from literature sources, own measurements, expert knowledge or their combination. Even if only one source is used, the parameter value is in almost all cases the result of some statistical calculation to aggregate over a sample. This ranges from simple averages and estimates of linear regression models or generalized linear mixed effects models to results of more advanced techniques such as Bayesian analysis. All these approaches do some kind of aggregation, that is, scaling-up. This is more obvious where values from several sources are combined into one model parameter value, again using statistics and thus scaling-up. Fortunately, pre-model scaling problems such as nonlinear

relationships, heterogeneity in the form of non-normal distributions, and interactions can be addressed with a broad array of statistical methods including nonlinear regression models, generalized linear models or mixed models (Crawley, 2012).

Model inputs such as spatial maps or time series can also be subject to pre-model scaling. The spatial resolution of input maps can significantly and differentially affect model behaviour. For instance, scaling-up is always associated with a loss of information or aggregation error, for example with respect to landscape homogenization, when animal dispersal and animal population dynamics are simulated (Bocedi, Pe'er, Heikkinen, Matsinos, & Travis, 2012). This can lead to a bias depending on the aim of the model. Thus, the spatial resolution of input maps needs to be adapted to the requirements of the model by scaling-up and scaling-down. Common coarsening methods for aggregation of information in gridded maps apply the majority rule (use the most common state of fine cells for the aggregated cell), simple averaging (use the average of the fine-scale information), area-weighted mean (weigh the average by the area a state is covering inside the aggregated cell) and kriging (fit a spatial function of the information; Chiles & Delfiner, 1999). Scaling-down in contrast increases the information content by generating finer-scale patterns based on interpolating large-scale patterns, sometimes using additional information and models. One example is climate regionalization, where climatic variables are scaled down (Ekstroem, Grose, & Whetton, 2015). This includes that temperatures are interpolated and refined by fine-scale elevation data, and precipitation patterns by additional data about prevailing wind fields resulting in luv and lee effects of mountain ranges (Karger et al., 2017). Similarly, temporal scaling-down adds values between given values, by more or less explicit assumptions about the underlying process. For example, daily temperature data are often created by interpolation based on linear or sine-wave functions between monthly values (Lischke, Löffler, & Fischlin, 1997).

2.2 | In-model scaling

Modelling is a robust way to bridge scale discrepancies and to include multiple spatial and temporal interactions when analysing complex ecological phenomena (Urban, 2005). This is reflected in the broad range of model-based scaling-up methods. Reviews of ecological model-based scaling methods are available for general ecology (Rastetter, Aber, Peters, Ojima, & Burke, 2003; Urban, 2005; Urban et al., 1999; Wu et al., 2006), species distribution models (Miller et al., 2004), landscape research (Lischke et al., 2007), forest ecosystem management (Seidl et al., 2013) and dynamic vegetation models (Snell et al., 2014); thus, we only give a short overview here (see also Table 1). Scaling-down refers to the refinement of processes within the model, sometimes by coupling a submodel for some processes to the original model. For example, the dynamic global vegetation model LPJ, which was based on representative (average) individuals of plant functional types, was merged with an individual-based model of tree species, significantly improving the simulated vegetation dynamics (Smith et al., 2008).

The purposes of in-model scaling-up are deriving simpler models for (a) computationally more efficient simulations (in terms of memory and time), (b) mathematical tractability and (c) detecting and formalizing emergent dynamics. A few in-model scaling-up methods will be reviewed in the following (see also Table 1).

Averaged or 'effective' parameters (Wu et al., 2006) can sometimes be derived from the small scale for the large scale and directly used as input to the original small-scale model to generate large-scale output (Cipriotti et al., 2016). This only works in the absence of scale-dependent processes or interactions at the large scale.

Brute force approaches work with copies of a small-scale model that are run in each cell of a large-scale grid using the local conditions of the respective cell as inputs (Cipriotti et al., 2016; Peters et al., 2004). To work trustworthily, they also require the absence of additional large-scale processes and large-scale interactions. A brute force approach can be combined with interactions at the large scale to broaden its applicability. However, this comes at the cost of huge computing power requirements and is often not feasible. In individual-based modelling, computing power constraints of brute force approaches can be alleviated using super-individuals or representative individuals. In individual-based models (and some complex physiological models), population- or community-level dynamics are not imposed, but emerge from individual-level interactions (Grimm & Railsback, 2005). Aggregating such individuals into super-individuals that share similar properties such as age or size is very useful to save computing time (Scheffer, Bavaco, DeAngelis, Rose, & van Nes, 1995). For example, Xavier et al. (2007) successfully implemented a multiscale individual-based model of microbial and bioconversion dynamics with bacterial groups clustered into super-individuals called granules. Representative individuals are also typical of forest gap models (Bugmann, 2001). Similarly, grid cells can be clustered into similar representative cells. For example, the forest landscape models LANDIS II (Scheller et al., 2007) and iLand (Seidl, Rammer, Scheller, & Spies, 2012) work with ecoregions that are considered homogeneous over the entire simulation time. In their dynamic global vegetation model, Lehsten et al. (2019) simulate local forest dynamics only along transects, interpolate between them, but simulate dispersal in all fine-scale grid cells. Nabel (2015) dynamically assign cells of similar species composition to representative cells in which local forest dynamics are simulated and then redistributed into space. However, the aggregation bias associated to a super-individual or representative cell approach arising from the clustering of similar but not necessarily identical entities can advocate against this approach and in favour of non-aggregated approaches (Parry & Evans, 2008).

A very elegant but also very difficult way of scaling-up is to derive the broad-scale process functions of the model given the scaling-up of the state variables, ideally mathematically or partly heuristically (see Figure 1 in Lischke et al., 2007). For example, Lischke et al. (1998) aggregated tree populations of a forest gap model into height classes. The main heuristic assumption was that a crucial element of the gap models—namely the between-gap heterogeneity

resulting from the small-scale demographic stochasticity in each gap—can be mimicked by a random spatial distribution of the trees. Based on this assumption, it was possible to mathematically derive the main process functions.

Scale transition theory also provides analytical methods for such a mathematical scaling-up of population dynamics based on nonlinear population growth and spatial or temporal variation in population densities as well as abiotic factors (Chesson, Donahue, Melbourne, & Sears, 2005). These methods are based on a mathematical manipulation, the Laplace transform, and draw on the phenomenon of nonlinear averaging (Chesson, 1978, 1981). It has, for example, been applied to scale-up the population dynamics of caddisflies from riffles to whole streams (Melbourne & Chesson, 2005). Scale transition theory has been elaborated and widely extended over the past decades (Benedetti-Cecchi et al., 2012; Chesson, 1998, 2009, 2012; Chesson et al., 2005; Melbourne & Chesson, 2006).

Meta-models are doubly simplified, because they are models of models (Urban, 2005; reviewed in Pietzsch et al., 2020). To obtain a meta-model, first a small-scale simulation model is run under all sets of conditions that are relevant for large-scale model analysis. Then, the meta-model is fitted to the results of these simulations given the respective conditions. The meta-model is then used to make predictions at the large scale that are either directly used or entered into a simplified large-scale simulation model. In contrast to the method of analytical approximations (see Section 2.3), in meta-modelling, the small-scale model is not of interest in itself, but only as a basis for scaling-up with the meta-model. Typically, a meta-model is a statistical model such as a multiple regression, mixed effects or generalized linear model (Meier et al., 2012). More recent meta-modelling techniques include neural networks. Rammer and Seidl (2019) fitted a neural network to the IBM iLand, run in many combinations of environmental conditions. The meta-model then gave the trajectories of the forest dynamics. In forest gap modelling, Markov models have also been used as meta-models (Acevedo et al., 1995, 1996) and can under certain conditions be generated automatically (Acevedo et al., 2001). Urban et al. (1999) presented a semi-Markov meta-model based on a cellular automaton. These meta-models interpolate over the whole parameter space given by the simulated conditions. In other cases, meta-models consist of libraries or lookup tables that contain results of small-scale simulations only for the simulated sets of conditions without interpolation (Cipriotti et al., 2016; Seidl et al., 2012). These libraries or lookup tables are then used in large-scale simulations to save ad hoc simulation time. If meta-models are used to avoid running many costly simulations of the original model, then they can be 'black boxes', as in the case of the neural networks. If the meta-model structure is based on ecological theory, such as the Lotka-Volterra competition model, meta-modelling could additionally give ecological insights by relating the processes and parameters of the original complex model to that of a mathematically well studied simpler theoretical model.

2.3 | Post-model scaling

Post-model scaling addresses all scaling procedures that are applied to the output of models. Post-model scaling procedures comprise analytical approximations as well as statistical methods such as those described in the section on pre-model scaling but also advanced spatial statistical approaches such as the pair correlation function. Usually, simulation results are presented at different degrees of aggregation, for example, as total biomass versus environmental gradients or as biodiversity in time and space, all of them representing a simple post-model scaling-up. In contrast, if representative-cell approaches have been used, the results are then redistributed in space, that is, scaled-down again.

Analytical approximations are elegant tools to simplify scaling-up approaches by fitting an analytical model to the large-scale output of a small-scale simulation model (Grimm & Railsback, 2005, chapter 11). If a suitable analytical expression can be found, it is a huge advantage, because analytical models are computationally much simpler than corresponding simulation models (Tietjen & Huth, 2006). Hence, it is not a scaling method per se, but a method to improve computing times during scaling. In contrast to meta-modelling (see Section 2.2), the analytical approximation is not necessary to fulfil the aim of the original model, but contributes an additional post-hoc scaling-up component. For instance, Moorcroft et al. (2001) used partial differential equations to approximate the results of a stochastic individual-based simulation model of ecosystem dynamics. Their small-scale simulation model operates at the scale of 15 m × 15 m, representing a forest gap of a few individuals. With their analytical approximation of time-dependent model behaviour, they were able to scale up to one-degree grid cells. Analytical approximations can also help to gain insight from scaling-up between individuals and populations. Fahse et al. (1998) were able to show that the results of their individual-based simulation model on bird flocking corresponded well with the analytical model of logistic growth. This was perceived as very beneficial during model analysis. However, applications of analytical approximations are not always unequivocal. In their analytical approximations of the population-level results of an individual-based competition model, Johst et al. (2013) found that the same type of resource competition was best described by different functional forms depending on the type of exogenous fluctuation, that is, food availability or weather. Fitting analytical models should thus not only follow statistical criteria, but be tailored to the ecological case study.

Advanced spatial statistics can also be used for scaling-up simulation results. Azaele et al. (2015) applied the pair correlation function to scale spatial species abundance distributions from a 1 m radius to a 400 m radius. The pair correlation function describes spatial point patterns (e.g. of study subject locations) across scales by determining how the density of points varies depending on the distance to reference points. However, predicting species abundance distributions from pair correlation functions relies on restrictive assumptions that exclude scale dependence in the functional form of the species abundance distribution, require the random environmental

variability in the region to be moderate and prohibit spatial heterogeneity. This strongly restricts the applicability of the pair correlation function for scaling-up.

3 | CONCLUSIONS AND FUTURE DIRECTIONS

This review takes the perspective that scaling is inherent to modelling and elaborates how the scaling approaches that are available can be classified into pre-model, in-model and post-model scaling methods. This implies that scaling and the associated problems are probably more widespread than previously thought, since they cover so many different areas of modelling. Thus, we recommend that ecologists be aware of scaling problems especially where models do not explicitly aim at scaling.

In terms of practical recommendations, we have derived a rough identification key for scaling approaches based on the approaches reviewed here (Figure 3). Given the aims and conditions of an ecological modelling study, the key provides decision support which scaling approach may be most suitable in which cases. However, the small number of aims and conditions covered by the key also reveals that the available scaling approaches do not often provide generally applicable solutions. Often, scaling approaches will have to be tailored to the specific case or new methods will have to be developed.

Future directions for theoretical and methodical advancement include four steps: First, the potential of existing scaling approaches should be more fully exploited. For example, in meta-modelling, many more statistical techniques may be explored. Second, underexplored scaling techniques such as Bayesian approaches (Clark, 2003; Ellison, 2004; Wikle, Berliner, & Cressie, 1998) and spatial statistics (Azaele et al., 2015; Illian, Penttinen, Stoyan, & Stoyan, 2008) should be further developed. For example, a Bayesian approach has already been successfully applied to scale-up from the individual level to the population level in an animal movement model (Hooten, Buderman, Brost, Hanks, & Ivan, 2016). Third, gaps in the coverage of current approaches (Figure 3), for example, dealing with scale dependence, should be filled by developing entirely new scaling approaches. Finally and importantly, scaling-up results need to be evaluated (Moorcroft et al., 2001; Schindler, 1998). Such an evaluation requires knowledge of the true pattern for comparison, at least for parts of the target scale, and should ideally control for environmental variability (Azaele et al., 2015), which is difficult to achieve. One avenue to address this evaluation challenge, for example, in vegetation modelling is to take advantage of data from mega-plots such as the ForestGEO plots with extents of more than 25 ha or of new methods in remote sensing that provide data at almost any desired grain and extent and compare them with scaling results. Another avenue is to design virtual test systems where the true patterns are known at any scale so that the relative performance of different scaling methods can be evaluated (e.g. case study 2 in Sciaini, Fritsch, Scherer, & Simpkins, 2018). Following these directions may help to improve

current scaling solutions and diversify the portfolio of available scaling methods in the future.

ACKNOWLEDGEMENTS

The authors thank Daniel S. Esser, Craig E. Simpkins and Sebastian Hanß for fruitful discussions and critical comments on the manuscript. This research was part of the Research Training Group 1644-Scaling Problems in Statistics funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). Open access funding enabled and organized by Projekt DEAL.

AUTHORS' CONTRIBUTIONS

M.F. and K.M.M. conceived the ideas supported by previous ideas of H.L., all authors conceptualized the paper; M.F. and H.L. did the literature research; M.F. and K.M.M. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13466>.

DATA AVAILABILITY STATEMENT

There are no data to be archived, since this is a review article.

ORCID

Katrin M. Meyer  <https://orcid.org/0000-0002-9990-4047>

REFERENCES

- Acevedo, M. F., Abla, M., Urban, D. L., & Pamarti, S. (2001). Estimating parameters of forest patch transition models from gap models. *Environmental Modelling & Software*, 16, 649–658. [https://doi.org/10.1016/S1364-8152\(01\)00034-2](https://doi.org/10.1016/S1364-8152(01)00034-2)
- Acevedo, M. F., Urban, D. L., & Abla, M. (1995). Transition and gap models of forest dynamics. *Ecological Applications*, 5, 1040–1055. <https://doi.org/10.2307/2269353>
- Acevedo, M. F., Urban, D. L., & Shugart, H. H. (1996). Models of forest dynamics based on roles of tree species. *Ecological Modelling*, 87, 267–284. [https://doi.org/10.1016/0304-3800\(94\)00208-8](https://doi.org/10.1016/0304-3800(94)00208-8)
- Azaele, S., Maritan, A., Cornell, S. J., Suweis, S., Banavar, J. R., Gabriel, D., & Kunin, W. E. (2015). Towards a unified descriptive theory for spatial ecology: Predicting biodiversity patterns across spatial scales. *Methods in Ecology and Evolution*, 6, 324–332. <https://doi.org/10.1111/2041-210X.12319>
- Barraquand, F., & Murrell, D. J. (2013). Scaling up predator–prey dynamics using spatial moment equations. *Methods in Ecology and Evolution*, 4, 276–289. <https://doi.org/10.1111/2041-210X.12014>
- Benedetti-Cecchi, L., Tamburello, L., Bulleri, F., Maggi, E., Gennusa, V., & Miller, M. (2012). Linking patterns and processes across scales: The application of scale-transition theory to algal dynamics on rocky shores. *Journal of Experimental Biology*, 215, 977–985. <https://doi.org/10.1242/jeb.058826>
- Bocedi, G., Pe'er, G., Heikkinen, R. K., Matsinos, Y., & Travis, J. M. J. (2012). Projecting species' range expansion dynamics: Sources of systematic biases when scaling up patterns and processes: Scaling species' range expansion dynamics. *Methods in Ecology and Evolution*, 3, 1008–1018. <https://doi.org/10.1111/j.2041-210X.2012.00235.x>

- Bolker, B., & Pacala, S. W. (1997). Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52, 179–197. <https://doi.org/10.1006/tpbi.1997.1331>
- Botkin, D. B., Janak, J. F., & Wallis, J. R. (1972). Rationale, limitations, and assumptions of a Northeastern forest growth simulator. *IBM Journal of Research and Development*, 16, 101–116. <https://doi.org/10.1147/rd.162.0101>
- Bugmann, H. (2001). A review of forest gap models. *Climatic Change*, 51, 259–305.
- Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., ... Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145, 87–99. <https://doi.org/10.1007/s00442-005-0100-x>
- Chesson, P. (1978). Predator-prey theory and variability. *Annual Review of Ecology and Systematics*, 9, 323–347. <https://doi.org/10.1146/annurev.ev.es.09.110178.001543>
- Chesson, P. L. (1981). Models for spatially distributed populations: The effect of within-patch variability. *Theoretical Population Biology*, 19, 288–325. [https://doi.org/10.1016/0040-5809\(81\)90023-X](https://doi.org/10.1016/0040-5809(81)90023-X)
- Chesson, P. L. (1998). Making sense of spatial models in ecology. In J. Bascompte & R. Sole (Eds.), *Modelling spatiotemporal dynamics in ecology* (pp. 151–166). Berlin, Germany & New York, NY: Springer.
- Chesson, P. (2009). Scale transition theory with special reference to species coexistence in a variable environment. *Journal of Biological Dynamics*, 3, 149–163. <https://doi.org/10.1080/17513750802585491>
- Chesson, P. (2012). Scale transition theory: Its aims, motivations and predictions. *Ecological Complexity*, 10, 52–68. <https://doi.org/10.1016/j.ecocom.2011.11.002>
- Chesson, P. L., Donahue, M. J., Melbourne, B. A., & Sears, A. L. W. (2005). Scale transition theory for understanding mechanisms in metacommunities. In M. Holyoak, M. A. Leibold, & R. Holt (Eds.), *Metacommunities: Spatial dynamics and ecological communities* (pp. 279–306). Chicago, IL: University of Chicago Press.
- Chiles, J.-P., & Delfiner, P. (1999). *Geostatistics: Modeling spatial uncertainty, Wiley series in probability and statistics. Applied probability and statistics section*. New York, NY: Wiley.
- Cipriotti, P. A., Wiegand, T., Pütz, S., Bartoloni, N. J., & Paruelo, J. M. (2016). Nonparametric upscaling of stochastic simulation models using transition matrices. *Methods in Ecology and Evolution*, 7, 313–322. <https://doi.org/10.1111/2041-210X.12464>
- Clark, J. S. (2003). Uncertainty and variability in demography and population growth: A hierarchical approach. *Ecology*, 84, 1370–1381. [https://doi.org/10.1890/0012-9658\(2003\)084\[1370:UAVIDA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[1370:UAVIDA]2.0.CO;2)
- Crawley, M. J. (2012). *The R book* (2nd ed.). Chichester, UK: Wiley.
- Denny, M., & Benedetti-Cecchi, L. (2012). Scaling up in ecology: Mechanistic approaches. *Annual Review of Ecology, Evolution, and Systematics*, 43, 1–22. <https://doi.org/10.1146/annurev-ecolsys-102710-145103>
- Ekstroem, M., Grose, M. R., & Whetton, P. H. (2015). An appraisal of downscaling methods used in climate change research. *Ecology Letters*, 7(6), 301–319.
- Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters*, 7(6), 509–520. <https://doi.org/10.1111/j.1461-0248.2004.00603.x>
- Englund, G., & Leonardsson, K. (2008). Scaling up the functional response for spatially heterogeneous systems. *Ecology Letters*, 11, 440–449. <https://doi.org/10.1111/j.1461-0248.2008.01159.x>
- Fahse, L., Wissel, C., & Grimm, V. (1998). Reconciling classical and individual-based approaches in theoretical population ecology: A protocol for extracting population parameters from individual-based models. *The American Naturalist*, 152, 838–852. <https://doi.org/10.1086/286212>
- Grimm, V., & Railsback, S. F. (2005). *Individual-based modeling and ecology*. Princeton, NJ: Princeton University Press.
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., ... DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, 310, 987–991. <https://doi.org/10.1126/science.1116681>
- Hickler, T., Vohland, K., Feehan, J., Miller, P. A., Smith, B., Costa, L., ... Sykes, M. T. (2012). Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model: Future changes in European vegetation zones. *Global Ecology and Biogeography*, 21, 50–63. <https://doi.org/10.1111/j.1466-8238.2010.00613.x>
- Hooten, M. B., Buderman, F. E., Brost, B. M., Hanks, E. M., & Ivan, J. S. (2016). Hierarchical animal movement models for population-level inference. *Econometrics*, 27, 322–333. <https://doi.org/10.1002/env.2402>
- Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Chichester, UK & Hoboken, NJ: John Wiley & Sons.
- Johst, K., Lima, M., & Berryman, A. A. (2013). Scaling up: How do exogenous fluctuations in individual-based resource competition models re-emerge in aggregated stochastic population models? *Population Ecology*, 55, 173–182. <https://doi.org/10.1007/s10144-012-0345-2>
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Lehsten, V., Mischurow, M., Lindström, E., Lehsten, D., & Lischke, H. (2019). LPJ-GM 1.0: Simulating migration efficiently in a dynamic vegetation model. *Geoscientific Model Development*, 12, 893–908. <https://doi.org/10.5194/gmd-12-893-2019>
- Levin, S. A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. *Ecology*, 73, 1943–1967. <https://doi.org/10.2307/1941447>
- Lischke, H., Löffler, T. J., & Fischlin, A. (1997). Calculating temperature dependence over long time periods: A comparison and study of methods. *Agricultural and Forest Meteorology*, 86, 169–181. [https://doi.org/10.1016/S0168-1923\(97\)00015-4](https://doi.org/10.1016/S0168-1923(97)00015-4)
- Lischke, H., Löffler, T. J., & Fischlin, A. (1998). Aggregation of individual trees and patches in forest succession models: Capturing variability with height structured, random, spatial distributions. *Theoretical Population Biology*, 54, 213–226. <https://doi.org/10.1006/tpbi.1998.1378>
- Lischke, H., Löffler, T. J., Thornton, P. E., & Zimmermann, N. E. (2007). Model up-scaling in landscape research. In F. Kienast, O. Wildi, & S. Ghosh (Eds.), *A changing world. Challenges for landscape research* (pp. 259–282). Dordrecht, The Netherlands: Springer.
- Meier, E. S., Lischke, H., Schmatz, D. R., & Zimmermann, N. E. (2012). Climate, competition and connectivity affect future migration and ranges of European trees: Future migration and ranges of European trees. *Global Ecology and Biogeography*, 21, 164–178. <https://doi.org/10.1111/j.1466-8238.2011.00669.x>
- Melbourne, B. A., & Chesson, P. (2005). Scaling up population dynamics: Integrating theory and data. *Oecologia*, 145, 178–186. <https://doi.org/10.1007/s00442-005-0058-8>
- Melbourne, B. A., & Chesson, P. (2006). The scale transition: Scaling up population dynamics with field data. *Ecology*, 87, 1478–1488. [https://doi.org/10.1890/0012-9658\(2006\)87\[1478:TSTSU\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1478:TSTSU]2.0.CO;2)
- Miller, J. R., Turner, M. G., Smithwick, E. A. H., Dent, C. L., & Stanley, E. H. (2004). Spatial extrapolation: The science of predicting ecological patterns and processes. *BioScience*, 54, 310. [https://doi.org/10.1641/0006-3568\(2004\)054\[0310:SETSOP\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0310:SETSOP]2.0.CO;2)
- Moorcroft, P., Hurtt, G., & Pacala, S. W. (2001). A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs*, 71, 557–586. [https://doi.org/10.1890/0012-9615\(2001\)071\[0557:AMFSDV\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSDV]2.0.CO;2)

- Muller-Landau, H. C., Condit, R. S., Chave, J., Thomas, S. C., Bohlman, S. A., Bunyavejchewin, S., ... Ashton, P. (2006). Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters*, 9, 575–588. <https://doi.org/10.1111/j.1461-0248.2006.00904.x>
- Nabel, J. E. M. S. (2015). Upscaling with the dynamic two-layer classification concept (D2C): TreeMig-2L, an efficient implementation of the forest-landscape model TreeMig. *Geoscientific Model Development*, 8, 3563–3577. <https://doi.org/10.5194/gmd-8-3563-2015>
- Parry, H. R., & Evans, A. J. (2008). A comparative analysis of parallel processing and super-individual methods for improving the computational performance of a large individual-based model. *Ecological Modelling*, 214, 141–152. <https://doi.org/10.1016/j.ecolmodel.2008.02.002>
- Peters, D. P. C., Pielke, R. A., Bestelmeyer, B. T., Allen, C. D., Munson-McGee, S., & Havstad, K. M. (2004). Cross-scale interactions, nonlinearities, and forecasting catastrophic events. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 15130–15135. <https://doi.org/10.1073/pnas.0403822101>
- Pietzsch, B., Fiedler, S., Mertens, K. G., Richter, M., Scherer, C., Widayastuti, K., ... Berger, U. (2020). Metamodels for evaluating, calibrating and applying agent-based models: A review. *Journal of Artificial Societies and Social Simulation*, 23, 9. <https://doi.org/10.18564/jasss.4274>
- Rammer, W., & Seidl, R. (2019). A scalable model of vegetation transitions using deep neural networks. *Methods in Ecology and Evolution*, 10, 879–890. <https://doi.org/10.1111/2041-210X.13171>
- Rastetter, E. B., Aber, J. D., Peters, D. P. C., Ojima, D. S., & Burke, I. C. (2003). Using mechanistic models to scale ecological processes across space and time. *BioScience*, 53, 68. [https://doi.org/10.1641/0006-3568\(2003\)053\[0068:UMMTSE\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2003)053[0068:UMMTSE]2.0.CO;2)
- Scheffer, M., Baveco, J. M., DeAngelis, D. L., Rose, K. A., & van Nes, E. H. (1995). Super-individuals a simple solution for modelling large populations on an individual basis. *Ecological Modelling*, 80, 161–170. [https://doi.org/10.1016/0304-3800\(94\)00055-M](https://doi.org/10.1016/0304-3800(94)00055-M)
- Scheller, R. M., Domingo, J. B., Sturtevant, B. R., Williams, J. S., Rudy, A., Gustafson, E. J., & Mladenoff, D. J. (2007). Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling*, 201, 409–419. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>
- Schindler, D. W. (1998). Replication versus realism: The need for ecosystem-scale experiments. *Ecosystems*, 1, 323–334. <https://doi.org/10.1007/s100219900026>
- Schneider, D. C. (2009). *Quantitative ecology: Measurement, models, and scaling* (2nd ed.). Amsterdam, The Netherlands and Boston, MA: Elsevier/Academic Press.
- Sciaini, M., Fritsch, M., Scherer, C., & Simpkins, C. E. (2018). NLMR and landscapetools: An integrated environment for simulating and modifying neutral landscape models in R. *Methods in Ecology and Evolution*, 9, 2240–2248. <https://doi.org/10.1111/2041-210X.13076>
- Seidl, R., Eastaugh, C. S., Kramer, K., Maroschek, M., Reyher, C., Socha, J., ... Hasenauer, H. (2013). Scaling issues in forest ecosystem management and how to address them with models. *European Journal of Forest Research*, 132, 653–666. <https://doi.org/10.1007/s10342-013-0725-y>
- Seidl, R., Rammer, W., Scheller, R. M., & Spies, T. A. (2012). An individual-based process model to simulate landscape-scale forest ecosystem dynamics. *Ecological Modelling*, 231, 87–100. <https://doi.org/10.1016/j.ecolmodel.2012.02.015>
- Smith, B., Prentice, I. C., & Sykes, M. T. (2008). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space: Vegetation dynamics in ecosystem models. *Global Ecology and Biogeography*, 10, 621–637. <https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x>
- Snell, R. S., Huth, A., Nabel, J. E. M. S., Boci, G., Travis, J. M. J., Gravel, D., ... Lischke, H. (2014). Using dynamic vegetation models to simulate plant range shifts. *Ecography*, 37, 1184–1197. <https://doi.org/10.1111/ecog.00580>
- Speich, M. J. R., Zappa, M., & Lischke, H. (2018). Sensitivity of forest water balance and physiological drought predictions to soil and vegetation parameters – A model-based study. *Environmental Modelling & Software*, 102, 213–232. <https://doi.org/10.1016/j.envsoft.2018.01.016>
- Speich, M. J. R., Zappa, M., Scherstjanoi, M., & Lischke, H. (2019). FORHYCS v1.0: A spatially distributed model combining hydrology and forest dynamics. *Geoscientific Model Development*, 1–48. <https://doi.org/10.5194/gmd-2019-117>
- Tietjen, B., & Huth, A. (2006). Modelling dynamics of managed tropical rainforests – An aggregated approach. *Ecological Modelling*, 199, 421–432. <https://doi.org/10.1016/j.ecolmodel.2005.11.045>
- Urban, D. L. (2005). Modeling ecological processes across scales. *Ecology*, 86, 1996–2006. <https://doi.org/10.1890/04-0918>
- Urban, D. L., Acevedo, M., & Garman, S. (1999). Scaling fine-scale processes to large-scale patterns using models derived from models: Meta-models. In D. Mladenoff & W. Baker (Eds.), *Spatial modeling of forest landscape change: Approaches and applications* (pp. 70–98). Cambridge, UK: Cambridge University Press.
- Wikle, C. K., Berliner, L. M., & Cressie, N. (1998). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, 5, 117–154.
- Wu, J., Jones, K. B., Li, H., & Loucks, O. L. (2006). *Scaling and uncertainty analysis in ecology*. Dordrecht, The Netherlands: Springer, Netherlands. <https://doi.org/10.1007/1-4020-4663-4>
- Xavier, J. B., de Kreuk, M. K., Picioreanu, C., & van Loosdrecht, M. C. M. (2007). Multi-scale individual-based model of microbial and bioconversion dynamics in aerobic granular sludge. *Environmental Science and Technology*, 41, 6410–6417. <https://doi.org/10.1021/es070264m>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Fritsch M, Lischke H, Meyer KM. Scaling methods in ecological modelling. *Methods Ecol Evol*. 2020;11:1368–1378. <https://doi.org/10.1111/2041-210X.13466>