Received: 19 January 2025 | Accepted: 27 January 2026

DOI: 10.1111/2041-210x.70281

RESEARCH ARTICLE

Methods in Ecology and Evolution

Accounting for spatial interactions in the upscaling of ecosystem services

Andrea Larissa Boesing[1] | Gaëtane Le Provost[2] | Margot Neyret[3] | Anja Linstädter[4] | Javier Muro[5] | Jörg Müller[6] | Kirsten Jung[7] | Markus Fischer[8] | Maximilian Lange[9] | Olena Dubovyk[10] | Paul Magdon[11] | Ralph Bolliger[8] | Sophia Leimer[12] | Steffen Boch[13] | Swen Renner[14] | Till Kleinebecker[15] | Ute Hamer[16] | Valentin H. Klaus[17,18] | Wolfgang Wilcke[12] | Peter Manning[19]

Correspondence
Andrea Larissa Boesing
Email: andrea-larissa.boesing@senckenberg.de

Peter Manning
Email: peter.manning@uib.no

Funding information
Deutsche Forschungsgemeinschaft, 1374 'BiodiversityExploratories', Grant/Award Number: DU1596/1-1 BEF-UP II, LI1842/4-1 SEBAS and MA7144/1-2 SEBAS; Agroscope, Indicate

Handling Editor: Maria Grazia Pennino

Abstract

1. Maps of ecosystem service (ES) supply are frequently used to guide spatial planning, policymaking and ecosystem management. However, these are typically based upon coarse land-cover proxies. This approach lacks a strong mechanistic basis and neglects spatial biodiversity dynamics and interactions among landscape properties that can modify ES provision.
2. We present an analytical framework for ES upscaling that incorporates spatial interactions between landscape properties, which determine ES supply. The resulting models can be viewed as a spatially informed ES production function. Key aspects of our synthetic framework include (i) the systematic assessment of multiple drivers across many levels of abiotic and biotic organization to formulate the statistical ES production function, (ii) the inclusion of spatial interactions with the surrounding environment into the ES production function and (iii) the use of expert input to inform ES production functions.
3. We demonstrate the approach using two example ES from German grasslands: biodiversity conservation and water supply. We show that the inclusion of spatial interactions in the upscaling model improved model predictions by 15%–33%, depending on the ES evaluated. In addition, inclusion of spatial interactions led to reduced error associated with the upscaled estimates.
4. By overcoming several shortcomings of existing upscaling approaches, we generate maps of ES supply that can more reliably inform spatial planning. Further, the underlying models allow for simulation of changes in the drivers of ES supply and estimation of respective outcomes. These advantages have the potential to better link detailed local-scale ecological understanding and land management with large-scale ES supply mapping, and thus better inform decision making.

For affiliations refer to page 13.

KEYWORDS
causal inferences, ecosystem services mapping, landscape processes, scaling up, spatial planning, statistical modelling

1 | INTRODUCTION

Within ecology and land management, it is often necessary to upscale small-scale, often plot-level, ecological measurements to the larger spatial scales at which most ecosystem management decisions are made (Maes et al., 2012). This is particularly true for ecosystem services (ES), which are mapped in both research projects and in national and regional assessments that inform environmental planning and communicate ecological information to stakeholders (IPBES, 2019; Nelson et al., 2009).

ES mapping is conducted using several methods. Perhaps the simplest and most widely applied is the use of land use proxies. Here, ES supply is derived from land cover maps, with each unit of land cover typically having a fixed ES supply rate across its entire area (McGuire et al., 2001; Nelson et al., 2009). This approach is often used in the absence of detailed field data but can be inaccurate as many other factors including abiotic factors, specific biodiversity components or landscape structure also drive ES supply (Eigenbrod et al., 2010). Errors associated with this approach can further propagate if ES valuation applies an estimate of economic value to a land unit based solely on its land cover (Nelson et al., 2009; Plummer, 2009). While this simple approach is common, significant progress in mapping ES has been made by integrating more detailed biological data, such as vegetation indices or functional traits (e.g. Lavorel et al., 2011; Lavorel & Grigulis, 2012), and incorporating multiple drivers operating at different spatial scales (González-Chaves et al., 2022; Martínez-Harms et al., 2016; Spake et al., 2019).

ES mapping may also be achieved via upscaled data. Upscaling uses local-level ES data to predict/infer ES at larger spatial scales. Recent years have seen the emergence and rapid development of machine learning techniques for ES upscaling, in which automated procedures select and generate high performing models from many candidate variables (Scowen et al., 2021; Willcock et al., 2018). Further, the emergence of platforms such as ARIES (Artificial Intelligence for ES) has also facilitated the generation of ES predictions (Villa et al., 2014). Although machine learning techniques were designed to capture complex systems, the resulting models often lack a clear mechanistic/causal link to the service in question, mostly due to lack of a hypothesis-based selection of variables. This also leads to issues regarding the interpretability of the selected model and its predictions (Perry et al., 2022). A shift towards a more mechanistic representation of the relationships between ES and their drivers should provide more reliable predictions (Spake et al., 2019) and allow the causes of trade-offs and synergies between multiple ES to be identified (Neyret et al., 2021), resulting in better informed decision making.

An important aspect missing from most ES upscaling and mapping approaches is the incorporation of spatial processes and interactions between landscape components. We propose that this is achievable by incorporating landscape ecology concepts into modelling (see Metzger et al. (2021), for a conceptual overview). In recent years, the field of landscape ecology has provided significant conceptual and empirical insights into the impact of landscape structure and spatial interactions in driving ecological processes (Tscharntke et al., 2012; Mitchell et al., 2015) and ES (Le Provost et al., 2023; Martin et al., 2019). Unfortunately, this knowledge remains poorly integrated into ES upscaling approaches (but see González-Chaves et al. (2020); Spake et al. (2019); Grafius et al. (2018); Verhagen et al. (2016) and O'Brien et al. (2025), for examples of landscape ecology perspectives applied on ES modelling). Spake et al. (2019) also highlighted the spatial-context dependencies of ecological responses, aiming to identify where management actions are most effective for enhancing natural capital, the ecosystem components that deliver ES. While existing approaches discuss the integration of multiple scales of effects when upscaling ES (e.g. Spake et al., 2017, 2019) very few consider spatial interactions and processes explicitly (González-Chaves et al., 2022; Le Clec'h et al., 2019; Spake et al., 2019), and none that we are aware of provide clear conceptual and methodological guidelines on how such processes must be represented in upscaling models. This needs to be addressed as upscaling has much to gain; when spatial interactions are accounted for in ES upscaling, they are often very important, with ES supply estimates changing by ~20%–700% (min–max) depending on the ES considered (Verhagen et al., 2016). Further, for spatial interactions to be accurately incorporated into upscaling, the underlying mechanisms should ideally be represented by landscape metrics that reflect them. This requires expert understanding of the physical and ecological processes involved in supplying the ES.

Here, we present an ES upscaling framework that is based on statistical ES production functions (i.e. statistical models that contain variables representing the fundamental drivers of ES supply; Bruins et al., 2017; Figure 1) rather than proxies of ES supply. It synthesizes theoretical principles from other key literature in the field (e.g. Díaz et al., 2007; Metzger et al., 2021; Spake et al., 2019) and additionally advances the field of ES upscaling by tackling challenges in model transferability and applicability, through an application-oriented framework. We also provide practical tools for implementation. Key aspects of the framework include (i) the systematic assessment of multiple drivers across many levels of abiotic and biotic organization to formulate the statistical ES production function; (ii) the inclusion of landscape variables representing spatial interactions with the surrounding environment into the ES production function; and (iii) use of expert input to inform ES production functions. While we recognize

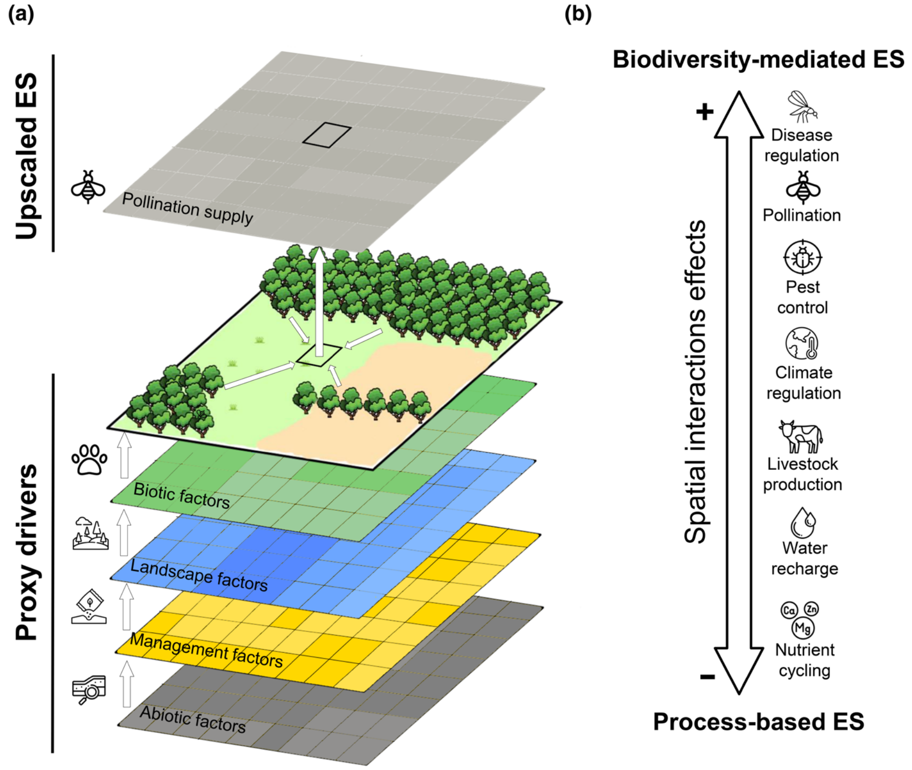

FIGURE 1 Representation of (a) spatially informed ecosystem service (ES) production function and (b) the expected importance of spatial interactions in driving the supply of different ES. The hierarchy of proxy drivers in (a) includes abiotic factors, management factors, landscape factors, and biotic factors. We hypothesize that biodiversity-mediated ES (e.g. pollination, disease regulation) are typically more influenced by spatial interactions than abiotic process-based ES (dominated by abiotic drivers; e.g. nutrient cycling, water recharge).

the importance of considering ES demand when studying ES, its upscaling and mapping requires different approaches (Wolff et al., 2015) than the upscaling of potential ES supply, which we focus on here.

Our framework encompasses seven key steps. We first describe the rationale behind each step and then demonstrate the application of the framework with two ES examples: biodiversity conservation and water supply. We used data from the large-scale and long-term German Biodiversity Exploratories (Fischer et al., 2010; https://www.biodiversity-exploratories.de/en) to train the models and upscale.

2 | SEVEN STEPS TO UPSCALE ES

The seven steps of our framework (Figure 2) are as follows: (1) Define the area for which upscaling is required (*Define the area*). (2) Identify which ES are demanded by stakeholders and appropriate indicators of these (*Identify ecosystem services*). (3) Identify spatially explicit indicators of the potential biotic and abiotic drivers of each ES (*Identify multiple proxy drivers*). (4) Select a spatially explicit statistical model for upscaling (*Select the statistical model*). (5) Conduct expert evaluation of the selected models' validity and adjust accordingly (*Expert evaluation*). (6) Cross-validate the model to assess performance (*Validate the re-fitted model*). (7) Upscale, acknowledging uncertainty in model predictions (*Upscale*) (Figure 2). The following sections describe these steps in detail.

2.1 | Define the area

Upscaling ES involves using existing local-level ES point data to predict/infer ES in larger units or unmeasured locations. In some cases, predictive models are applied outside the original study area to new geographic contexts that might differ from the training data in their environmental properties (Meyer & Pebesma, 2021), potentially leading to erroneous predictions. Thus, before upscaling, the boundaries of the area to be predicted must be defined to ensure that appropriate indicators, proxies and baseline data are used. If ES supply measures need to align with demand measures or other human factors, as is often the case in national and regional assessments, then it is appropriate to define this boundary as that of the socio-ecological system. Whilst ecological boundaries are usually defined by interfaces between different ecosystems, social boundaries are typically set by institutional forces, and represented as cultural and/or administrative borders (Dallimer & Strange, 2015). For further discussion of area delimitation, see Supporting Information S1.

2.2 | Identify ESs

ESs are linked to human demands; therefore, mapping the ES that are actually demanded (defined here as those prioritized by people; Wolff et al., 2015; Linders et al. 2021), along with the underlying

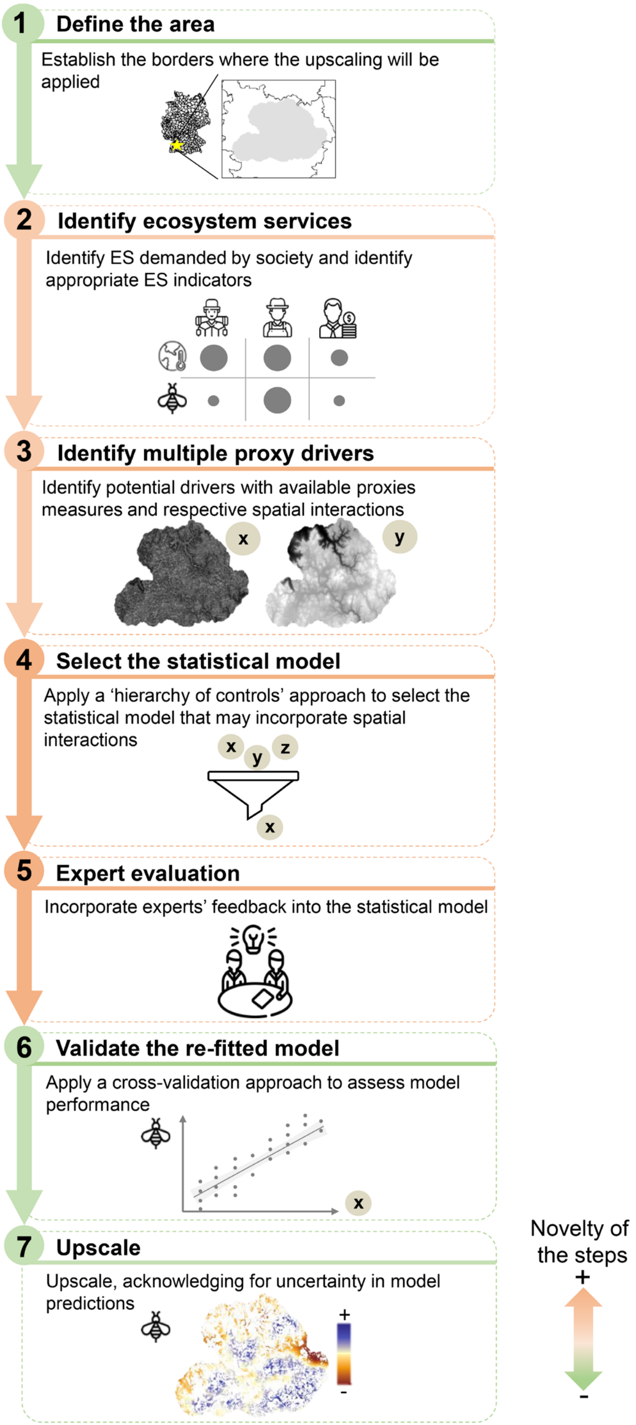

FIGURE 2 Seven-step analytical framework for spatially informed ecosystem service (ES) production function upscaling, indicating the novelty of the steps compared to standard upscaling approaches. The procedures are appropriate when high-quality plot-level data on ES supply and its drivers are available.

regulating services that support them (e.g., nutrient cycling and pollination), helps ensure the relevance of the resulting measures. This requires a multidisciplinary approach involving social scientists to accurately assess stakeholder's ES preferences (Neyret et al., 2023; Peter et al., 2022). However, we acknowledge in some cases specific services that may not be acknowledged by stakeholders, for example, due to a lack of ecological understanding, may be the priority for upscaling.

Once the demanded ESs are identified, it is necessary to identify appropriate *indicators* (Müller & Burkhard, 2012) that capture the characteristics and status of each ES, and which can be measured and upscaled appropriately. For instance, carbon storage and/or sequestration rates can indicate climate regulation (Smith et al., 2013); pollinator abundance can indicate pollination (Woodcock et al., 2019), and species richness in priority taxa can indicate biodiversity conservation value (Veach et al., 2017). Such indicators are usually only proxies of ES supply and already introduce uncertainty in the modelling process. Thus, good indicators, that is, those that closely approximate the ES they represent, both conceptually and quantitatively, minimize uncertainty. ES indicators are ideally collected at the plot level in a well-replicated study design, which encompasses the full environmental variability of the defined region (Redlich et al., 2022).

2.3 | Identify multiple proxy drivers

Before creating the statistical model for upscaling, multiple potential *drivers* of the ES indicator must be identified. The first step in this process is to review existing knowledge of the indicator, and select candidate variables that may drive it, including, where appropriate, abiotic drivers, land use, landscape features (e.g. surrounding land cover, and its configuration) and biotic properties (e.g. species richness and composition). This approach is hypothesis-driven, which helps limit the selection of closely fitting but mechanistically unrelated variables, but requires consideration and knowledge of the ES. Next, we must identify reliable *proxy* measures of these driver variables, for example remotely sensed measures, or those mapped conventionally. These should be available at the landscape scale, but also accurately represent the true observed local drivers. Advances in remote sensing data in recent years have substantially improved the resolution, accuracy, and temporal range of the available data, for example to derive landscape structure drivers (Cavender-Bares et al., 2022).

We advocate using proxies from at least four variable classes, which form a 'hierarchy of controls' (Figure 1a). These are: abiotic factors, management factors, landscape factors and biotic factors. *Abiotic factors* including climate and soil attributes (e.g. soil type, pH, texture and depth) underpin many ES (e.g. nutrient cycling, soil carbon storage) and data are increasingly available at regional and national levels (e.g. Copernicus, European Soil Data Centre). Topographic and geological properties, such as slope, aspect and topographic wetness index, may also be important abiotic drivers and can derived from digital terrain models and geological maps. *Management factors* such as management intensity and type are often important local-level predictors of ES, especially for provisioning services such as crop, livestock and fodder production. Proxy measures of management factors may be unavailable but remotely sensed anthropogenic activity measures are constantly improving: reliable geospatial information describing forest disturbance, crop and grassland type and land use intensity now exists (e.g. Griffiths et al., 2020; Lange et al., 2022). *Landscape factors* characterize the surroundings of the focal plots from which the ES are measured. They include landscape composition (i.e. land-cover types) and configuration (i.e., land-cover arrangement within a landscape), and landscape scale land management measures derived from remote-sensing (Griffiths et al., 2020; Lange et al., 2022). These variables should represent hypothetically important spatial processes such as spillover effects or colonization and extinction rates (see Metzger et al., 2021; Tscharntke et al., 2012) that may affect ES supply by influencing the movement, diversity and abundance of ES providing organisms (Le Provost et al., 2023; O'Brien et al., 2025), and the flow of abiotic materials such as water and pollutants (Cole et al., 2020). Finally, *biotic factors* include variables such as species richness and composition, biomass and functional composition. The availability and quality of remote sensed data for biotic factors has dramatically improved in several domains, including plant species richness and composition (e.g. Civantos-Gómez et al., 2021; Muro et al., 2022), but also functional traits (Martínez et al., 2016). The examples cited here provide a flavour of the many variables that can be employed, and which together form the main drivers of ES supply (Potschin & Haines-Young, 2011). Importantly, these variables must be selected using a knowledge-based approach and tailored to the ES and social-ecological system under evaluation. It is paramount to assess the correlation among variables, both within and across hierarchical levels, to ensure only uncorrelated variables are retained (details in Supporting Information S4 and Tutorials).

The next step is to determine the 'hierarchy of effects' of the chosen proxy drivers, based upon the principles of hierarchy theory (Allen & Starr, 1982; Müller, 1992). As part of this, it must be decided how each level's dynamics (i.e. drivers) are controlled by, and nested within the larger system (Müller, 1992). In general, lower level hierarchy is often influenced by the upper level hierarchy, but in most cases, not vice versa (Figure 1a). We also propose that deciding the driver variables used and the order of the hierarchy should be similar in philosophy to structural equation model formulation (see Grace 2006), in that each potential causal driver should be previously established via theoretical or empirical evidence, and the direction of causality should also be known, to determine the order of inclusion in the modelling process. To determine the hierarchy, we suggest that variables of climate, soil, geology or hydrology (*which are inherent ecosystem properties and relatively time invariant*) are likely to be basal in the hierarchy, while more temporally dynamic variables, such as land cover, landscape structure and biodiversity (*which are typically manageable*) are likely to be more apical in the hierarchy. These apical properties might feed-back upon each other. In such cases, we suggest assessing the relative strength of feedback arrows (i.e. which direction of the loop is dominant) based on scientific and expert knowledge to assign the subsequent order.

To exemplify the establishment of a hierarchy, consider a grassland ecosystem where abiotic factors (e.g. slope and hydrology) are basal variables (more static) that set the long-term conditions for soil depth and fertility. These abiotic factors (basal-levels) may influence management factors (apical-levels) (e.g. mowing or fertilization regimes) and landscape factors (such as land cover of other habitat types or grassland connectivity), but the opposite is rarely true (even though in some cases management practices may influence some soil aspects locally for instance). All levels described above would influence local biodiversity and community structure (Le Provost et al., 2021), but these biological properties will rarely influence the other properties significantly, meaning they are apical within this hierarchy. Thus, in this case, the hierarchical order would be: abiotic factors, management factors, landscape factors, and finally, biodiversity.

While the establishment of these hierarchies may oversimplify the complexity of ecological systems, their purpose is not to capture the full complexity of it, but rather to provide a logical and plausible structure that best identifies the true causal drivers and thus guide the establishment of a final upscaling model that approximates the true drivers of the ES. A caveat of the hierarchical approach is that apical levels may be underrepresented, or with conservative effect size, if they correlate strongly with more basal terms. In such cases, the expert assessment (see below) should remove drivers (often basal) that are unlikely to represent true mechanistic relationships, and assist in the retention of the most proximal drivers. Another limitation is that feedback between the levels, for example between biodiversity and soil properties, cannot be incorporated, although this may be important in some cases (Thakur et al., 2021).

Modelling ES supply as driven by a hierarchy of controls is consistent with the widely used ES cascade model (Potschin & Haines-Young, 2011), in which abiotic features alter the biota and together with external inputs drive ES supply (Bateman & Mace, 2020; Potschin & Haines-Young, 2011). The hierarchy of controls approach has been successfully applied in a range of studies aiming to predict ecosystem functions from biotic factors (e.g. Díaz et al., 2007; Lavorel et al., 2011; Manning et al., 2015), and to class drivers into different scales of influence within a natural capital framework (Spake et al., 2019), but to our knowledge, it has not been used in upscaling as a tool that informs the sequential stepwise construction of statistical ES production functions (see Section 2.4 below).

2.3.1 | Identifying spatial interactions driving the supply of ES

At the landscape level (Figure 1), we can identify at least three main categories of effects that are relevant to landscape drivers of ES provision, and thus worth including in our upscaling framework: (i) proximity effects; (ii) composition effects; and (iii) configuration effects (for more details, see Metzger et al. (2021)).

Proximity effects (i) can strongly influence some regulating ES such as pollination and pest control that depend on mobile agents. These organisms often inhabit native habitat patches surrounding agricultural areas and move into crops for foraging via spillover

(Tscharntke et al., 2012). Therefore, their ES can be delivered at a distance of a few meters to a few kilometres, depending on the movement capacity of the organisms and landscape connectivity. Similarly, the amount and arrangement of different land use types might also limit ES provision via *composition effects* (ii) which determine the ES providers' abundance and diversity, and their colonization and extinction rates. Similarly, *configuration effects* (iii), which affect the movement—for example via habitat corridors, and thus the distance ES providers might reach, their source-sink dynamics and their colonization and extinction rates. As an example, structurally complex agricultural landscapes tend to enhance local biodiversity and the abundance of ES providers as landscape heterogeneity positively influences habitat availability and species movement (Le Provost et al., 2023). This can offset the negative consequences of local high-intensity management and lead to higher ES provision in diverse and well-connected landscapes (Boesing et al., 2024; Le Provost et al., 2023)—which are related to both *composition* and *configuration effects*.

In operational terms, these spatial processes mentioned above might be represented by landscape metrics that can be easily retrieved from land use maps. For instance, *proximity effects* might be represented by using metrics such as edge density, aggregation and isolation indices; for *composition effects*, metrics such as the percentage of different land use types, landscape heterogeneity, core area indices or the ratio between different land uses can be used, and for *configuration effects*, metrics that represent the connectivity or isolation between land use types, and the shape of landscape elements can be used. See Hesselbarth et al. (2019) for a list of measurable metrics.

For such landscape drivers, we must determine the *scale of effect*, that is the scale over which the underlying processes operate (Jackson & Fahrig, 2015). Since the scale of effect of a given driver is often unknown, a multi-scale approach should be performed. Here, variables incorporating a range of scales are used in modelling (e.g. land cover measures in radii of 200, 500 and 1000 m) and the 'best' is selected as an approximation of the scale of influence that a driver operates over (Jackson & Fahrig, 2015). This approach is widely used to infer the scale over which mechanisms operate in landscape ecology research, and a fixed radii extent may be used where the scale of operation is well established from many studies (Mohamed et al., 2024). Selection of the scale of effect can be based on the strength of the relationship between the landscape property and the ES indicator (see Jackson & Fahrig, 2015).

2.4 | Select the statistical model

Next, we employ the hierarchy of controls approach to statistical model selection. In this, proxy drivers are fitted in a fixed sequential order that follows the hierarchy of effects determined in Step 2.3 to identify and select the final drivers of ES supply. In practical terms, groups of terms are added to a regression model in fixed sequential order, determined by their influence across levels and their impact on model likelihood is assessed. Only terms found to significantly improve model likelihood in a likelihood ratio deletion test (LRTs) are retained. After conducting this procedure for all variables of a hierarchical level, an interim model is selected, and the modelling process moves to the next level of the hierarchy, where another set of variables is added and the process is repeated (Figure 3). This continues until all levels are added and the final model is selected. Interactions between variables within and across levels should be tested at each stage in modelling where these have a hypothetical basis (Spake et al., 2019), since many ES are driven by multiple drivers and their interactions. To avoid the retention of many weakly significant factors and to fit a robust model, we suggest fitting only low-order interaction terms.

After selecting the statistical upscaling model, we suggest checking whether spatial autocorrelation, and related issues of pseudo-replication, impacts the upscaling (Dormann et al., 2007). Spatial autocorrelation (i.e. the higher similarity of closer samples) is common in ecology and can appear for several reasons, including biotic processes such as dispersal (Miller, 2012) or abiotic factors, such as climate, soils, topography and disturbances (Ranta et al., 1997). When spatial autocorrelation is related to the driver of interest, inclusion of the variable causing the autocorrelation, for example as fixed factor, may be sufficient. In other cases, ignoring spatial autocorrelation can lead to biased parameter estimates. Detection of autocorrelation can be done in several ways including Moran's *I* correlograms, Geary's correlograms and semi-variograms (Dormann et al., 2007). The inclusion of appropriate landscape factors in the modelling process should minimize spatial autocorrelation, but if it is detected, we suggest using autocovariate regression and spatial eigenvector mapping to capture the spatial configuration of covariates and adding a term for this to the upscaling model.

2.5 | Expert evaluation

After determining the statistical model, we suggest consulting experts in the ES of interest for model evaluation. Here, they should check whether the driver variables included are likely to be true mechanistic drivers, or simply correlates of them. This step is important when dealing with multiple ES, as it is unlikely that the person performing upscaling will have expert knowledge of all relevant ES. Experts should also assess whether other important drivers are missing, if the data used for the calculations is suitable, for example in its spatial and temporal resolution. After consultation, models must be revised and adapted, for example by substituting spurious drivers with closely correlated variables that represent the true drivers. While this step adds complexity and costs time, model outputs are likely to be greatly improved, especially where the model is used to make predictions. If possible, expert evaluation may also be applied when identifying the levels and order of the hierarchy (Step 2.3) and ideally along the whole process, though we note that each step of expert involvement adds effort and complexity to the modelling process.

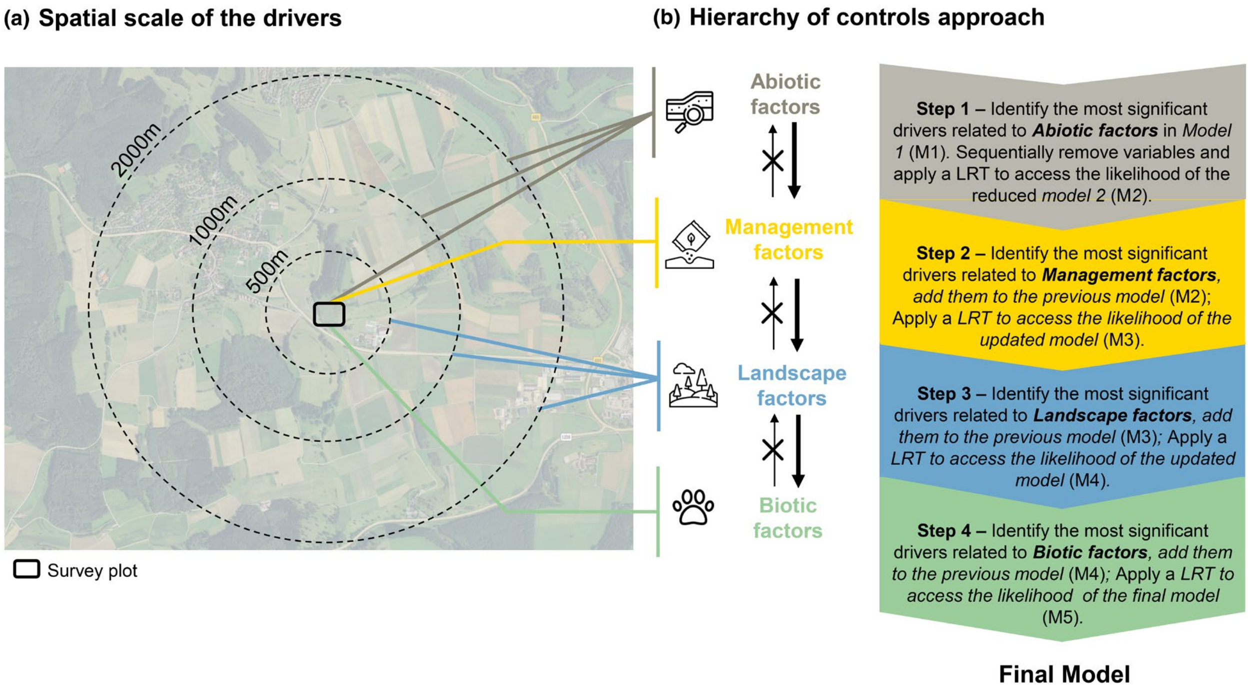

FIGURE 3 Representation of (a) the spatial scale of the drivers, and (b) the steps used in the hierarchy of controls approach to select the final statistical upscaling model.

2.6 | Validate the re-fitted model

After expert assessment and revision, the next step is to assess model performance via cross-validation and estimate the accuracy of model predictions. We suggest this is done by setting aside a random set of test observations at the model calibration stage and using these to quantify prediction error (true validation step). Validation techniques include 'validation set aside', '*k*-fold cross validation' and 'repeated *k*-fold cross-validation' (Browne, 2000), each of which has advantages and drawbacks. One limitation of cross-validation is that it is only possible for large datasets. For smaller datasets, we recommend (repeated) *k*-fold cross-validation to estimate the prediction error rate as this can use the entire dataset for generating estimates. Knowing data limitations is crucial when estimating error, and thus communicating uncertainty. A model should be considered reliable when it both retains the true drivers of ES change and has low error of estimates.

Finally, regarding the statistical setup of the models, both linear and generalized linear models may be used. Trying different model distributions and fitting models to both transformed and untransformed data are also recommended for finding models with good fit, performance, and lower error associated with predictions. Special attention must be taken when upscaling nonlinear models and for the presence of outliers in the training or proxy driver data as these can result in over or under-estimation or result in higher errors. When outliers are present, one simple solution is to limit predictions to areas within the range of the training data. This reduces prediction coverage but prevents extrapolation into areas where error may be high.

2.7 | Upscale

Once the final model is selected and its performance assessed, the layers containing the drivers must be prepared and the landscape interactions incorporated when necessary. For non-spatial drivers the model is applied to each pixel of the driver raster, while for spatially interacting drivers represented by landscape variables, a calculation needs to be made prior to upscaling of the 'landscape surroundings' value for each pixel (e.g. the proportion of a habitat type within a certain radius of the pixel). The model is then applied to this layer of calculated values and the other proxy drivers for each point in the landscape to generate estimates (Figure 1). Creating spatially explicit layers may be time-consuming and computationally demanding, so we advise the integrated use of different tools, for example QGIS, ArcGIS and R. In principle, the resulting maps represent ES supply as function of mechanistic drivers, including those involving spatial processes.

3 | APPLYING THE UPSCALING FRAMEWORK IN REAL LANDSCAPES

Our approach to upscaling ES incorporates novel elements that will likely not only improve the ES predictions but also improve their use as a tool for decision-making. These include the hierarchical approach to identify likely ES drivers, the representation of spatial processes, for example via interactions between landscape units, and expert consultation. Next, we demonstrate how our framework can be applied in practice with examples from the Schwäbische Alb region of southwest Germany.

3.1 | Step 1—Define the area

In this case, the study area was defined by a fundamental research project, the large-scale and long-term Biodiversity Exploratories project (Fischer et al., 2010), which was established to investigate local scale land-use-biodiversity-ecosystem function relationships. The boundaries of the region were therefore delimited to encompass the project's focal study plots, which were selected to span the full range of land use intensity of each region. We performed the upscaling only for grasslands, where plot-level ES indicator data were collected, meaning predictions outside this land use type (e.g. into forests) would be unreliable.

3.2 | Step 2—Identify ESs

We applied the framework in two case studies: biodiversity conservation and water supply. These ES were chosen based on their identified priority in a local social survey (see Peter et al., 2022) and to demonstrate the applicability of our approach to ES that are influenced by surrounding landscape factors. The indicator of biodiversity conservation was above-ground multidiversity, a synthetic measure of the overall diversity of 14 above-ground taxa (Table S2.1) (Allan et al., 2014) that is related to ecosystem multifunctionality (Soliveres et al., 2016). Data were from 50 grassland plots sampled between 2008 and 2021.

The second example is groundwater recharge, an indicator of water supply to people, and a regulating service for crop, fodder and timber production, which were all identified as important ES in the region. Values were derived from a soil water balance model developed to calculate vertical soil water fluxes from the top layer (0–0.15 m) (Kreutziger, 2006; Leimer et al., 2014, 2018) (Supporting Information S1, S2). As input data, we used biweekly rainfall data from 2010 to 2016 in the 50 grassland plots and net water fluxes from soil to deeper layers, aggregated to annual values. The average value across the 7 years was used.

3.3 | Step 3—Identify multiple proxy drivers

We collected data for several proxy drivers that potentially influence these ESs. These were organized in a hierarchy of: (1) abiotic factors including topographic variables: slope, elevation, topographic wetness index and aspect, and soil attributes: soil texture, pH, clay content, bulk density and soil depth, (2) local management factors: land use intensity (LUI; Blüthgen et al., 2012), grazing, fertilization, and mowing. (3) landscape factors (500–2000 m scale; including the landscape composition and configuration of land-cover), including forest, grasslands and crops, and spatial patterns of landscape-management, represented by amount and configuration of low and high LUI grasslands at the best fitting scale of response (1500 m surrounding the plots for biodiversity conservation, and 2000 m for water recharge), and biotic factors (as mean alpha plant richness). Details of the ecological mechanisms these drivers represent, their respective metrics, spatial scale selection and the sources of remote sensing data for the proxy drivers can be found in Supporting Information S2, S3 and S4. Correlations across variables were checked (Supporting Information S4) and variables with $r = < 0.7$ were used, since correlations above this threshold may inflate regression coefficients (Dormann et al., 2013).

3.4 | Step 4—Select the statistical model

We next applied the 'hierarchy of controls' approach to identify the upscaling models. For each level (Figure 3), we tested which variables improved model likelihood in a LRT, before advancing to the next step of the modelling process. For both ESs, we used linear models with Gaussian error distribution. Model fit was evaluated by examining residuals (Hartig, 2020) (Supporting Information S3) and we tested for spatial autocorrelation using Moran's *I* correlograms (Paradis & Schliep, 2019) (Supporting Information S5). The initially selected model for biodiversity conservation explained 37% of the variation and contained terms related to local abiotic factors (soil phosphorus content), landscape factors (landscape heterogeneity, measured as the diversity of land use types at 1500-m scale), and local biotic factors (estimated plant species richness). For water supply, the initial model explained 72% of the variance and contained terms for local abiotic factors (pH and bulk density), and landscape factors (% coniferous cover within 2000 m, and % high LUI cover at 1500-m scale).

3.5 | Step 5—Expert evaluation

We sent the initially selected models to experts from the Biodiversity Exploratories Consortium (Fischer et al., 2010) who were the original data collectors of the ES indicators in the form of a spreadsheet containing the models, and requested feedback on model plausibility, including appropriateness and absence of driver variables, and direction of effects. For *biodiversity conservation*, three changes were made following this consultation: first, we updated landscape LUI measures as improved data had become available (Lange et al., 2022); second, experts noted potential circularity in using remote sensed plant richness estimates as proxy driver of biodiversity conservation since related measures were included in the multi-diversity index; thus, it was removed. Finally, we moved topography measures to the first level of hierarchy, before soil attributes. For water supply, two changes were performed: the model initially contained bulk density and LUI at the landscape level (1500 m) as drivers, but experts identified the direction of effects as spurious so these were removed. Second, we moved topography measures to the first level of hierarchy as well.

3.6 | Step 6—Validate the re-fitted model

The final model after revision for *biodiversity conservation* had a R^2 of 0.47, and contained effects of landscape scale land-use diversity (1500 m) (explaining 15% of variance), slope (14%), soil depth (10%) and soil phosphorus content (8%) (Table S3.1). The landscape factor indicates the importance of spatial processes in driving biodiversity, consistent with detailed ecological studies of the same grasslands (Le Provost et al., 2021). The final mechanistic upscaling model for *water supply* had a R^2 of 0.56 and contained effects of soil type (explaining 26% of the variance) and the cover of coniferous forest within 2000 m (25% of variance) (Table S3.1). These factors influence water permeability and thus infiltration rates; coniferous forests are typically denser and intercept more water than broad-leaved forests (Adane & Gates, 2015; Bellot et al., 1999). See Supporting Information S3 for a detailed description of the mechanisms driving the supply of both ES. While R^2 here was not particularly high for a predictive model we note that these models predict variation within a single ecosystem type and local region. Broader environmental gradients could be expected to lead to more closely fitting models.

We used a *K*-cross fold validation to assess model predictability power, where the entire dataset is used to calculate the root-mean-squared error (RMSE), since our training dataset is relatively small ($N=50$). RMSE indicates the correspondence between observed and predicted values, with lower values representing lower prediction error. For *biodiversity conservation*, RMSE was 0.07 (observed: 0.13–0.58, predicted: 0.0001–0.71). Given the range of training data and predicted values, an RMSE this can be considered acceptable error (i.e. predictions may vary ±0.07 from actual values). For *water supply*, we obtained a RMSE of 53.11 (observed: 477.69–791.51 mm, predicted mm: 339.39–737.39 mm). In both cases, RMSE is considerably smaller than the range of predicted values. Therefore, while exact estimates may sometimes not be possible along the gradient, areas of relatively high and low values are identified with confidence.

3.7 | Step 7—Upscale

Remote sensed inputs were harmonized and stacked, and model coefficients were applied to each pixel of the stack to make predictions (details in the Tutorial; Supporting Information S7). First, we illustrate how landscape variables, which we expect to represent known spatial factors driving local biodiversity such as landscape-level habitat heterogeneity (Le Provost et al., 2021), affect levels of *biodiversity conservation*. The model predicts that grasslands surrounded by a high diversity of land uses, on steeper slopes, and with deep and phosphorus-poor soils, possess the highest above-ground biodiversity (Figure 4).

Second, for water supply, areas embedded in low coniferous cover landscapes, and with soil of type Braunerde-Terra Fusca, have a higher groundwater recharge (Figure 5). In this case, the landscape component (low coniferous cover) was expected to relate to forest density, and thus interception rates and larger scale water availability (see Supporting Information S3).

4 | UPSCALING WITH VERSUS WITHOUT SPATIAL INTERACTIONS

Integrating spatial interactions should provide more reliable estimates of ES supply if these processes are important, which we demonstrate in the following section. Since the landscape drivers were added towards the end of the hierarchy, we did this by comparing the best models with and without spatial interaction terms. For *biodiversity conservation*, removing landscape heterogeneity reduced explained variance from 47% to 32% and increased RMSE from 0.07 to 0.09. In the absence of the spatial term, the model underestimated multidiversity in more heterogeneous landscapes and overestimated multidiversity in more homogenous landscapes (Figure 6a,b).

For *water supply*, removing the landscape term reduced the explained variance by 25% and increased the RMSE from 53.1 to 67.1. In the absence of spatial drivers, the model underestimates groundwater recharge by up to ~100 mm in lower coniferous cover landscapes (Figure 6c,d), and overestimates it in those with higher cover, by as much as ~200 mm (Figure 6c,d), an approximate overestimate of 32%. Spatial model predictions ranged from 319.4 to 739.4 mm, while in the non-spatial model, only two values are predicted, based on soil type (646.41 mm in Braunerde-Terra fusca soil type, and 549.80 mm in Rendzina-Braunerde; Figure 6c). These examples demonstrate that incorporating spatial interactions can strongly improve model performance to a level that would impact ES estimates and consequent decision-making (e.g. the prioritization of conservation areas).

5 | FRAMEWORK TRANSFERABILITY, ACCESSIBILITY AND SCALABILITY

While the detailed examples above illustrate the use of the framework in practice, we also investigated the general applicability of our approach to upscale a wide range of ES by applying it to 11 ES represented by 17 ES indicators in the same study region. We found that the framework produces reliable outputs across different ES types (details in Supporting Information S6). As expected, the identity and scale of influence of drivers varied considerably (Figure S6.2; Table S6.1). The errors associated with the estimates were generally acceptable (average 14%), but for cultural ES, relatively higher (average 54%; Figure S6.1). For the cultural services, the selected biophysical drivers may be weak predictors of the indicators as many cultural ES are strongly determined by human preferences (e.g., one indicator was 'culturally important birds'). In such cases, it may be preferable to first upscale ecological

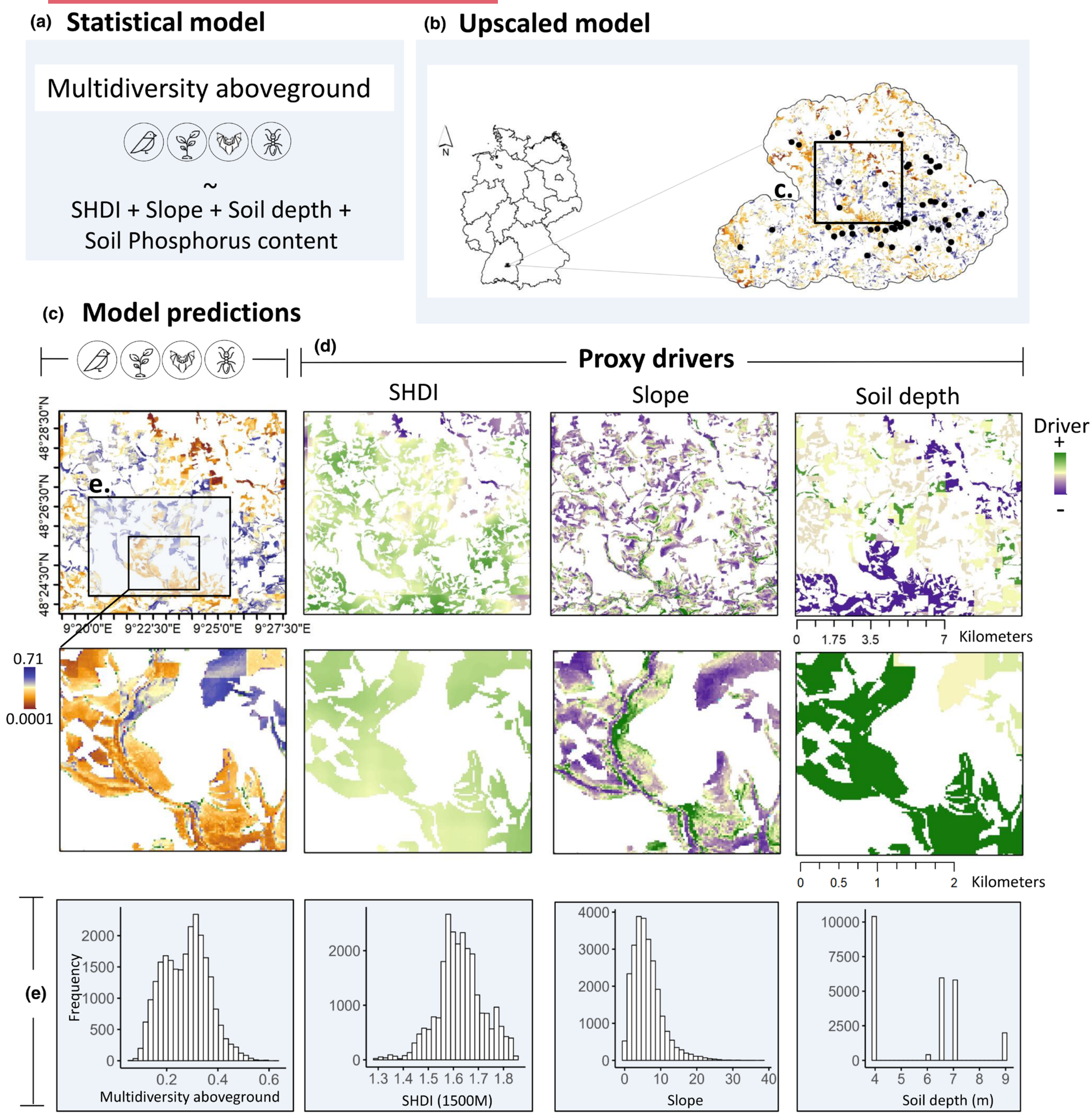

FIGURE 4 Upscaled predictions for biodiversity conservation (indicated by grassland above-ground multiversity). We present the final upscaling model in (a). In (b) we present the region and upscaled biodiversity conservation with a zoom in (c). The dots in (b) represent training data plot locations. (d) Three of the drivers used to model multidiversity. The middle row shows a closer zoom to show how spatial context, in the case of landscape heterogeneity (SHDI; Shannon's diversity index) alters multidiversity. The additive effect of SHDI and local factors (slope, soil depth, and soil phosphorus content) are noticeable, and landscapes with higher diversity of land uses have higher multidiversity. The bottom row (e) shows the data extracted from the closer zoom (larger square), and predicted values of multidiversity and its drivers (SHDI, slope and soil depth). Soil phosphorus content is not shown in the figure but included in the model. Predictions for areas mapped white were not calculated (e.g., arable land and forest).

properties (in this case for instance, individual cultural bird species abundances) and then derive ES measures from these. Based on these results, we expect that our framework is applicable to a wide range of social-ecological systems, ecosystem types and regulating, provisioning and supporting ES.

As with any upscaling approach, it is essential to critically assess the predictability of the underlying model and to communicate uncertainties transparently. In our case, statistical models serve as the basis for upscaling, meaning that models with low performance may lack generalizability and carry high uncertainty—particularly

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

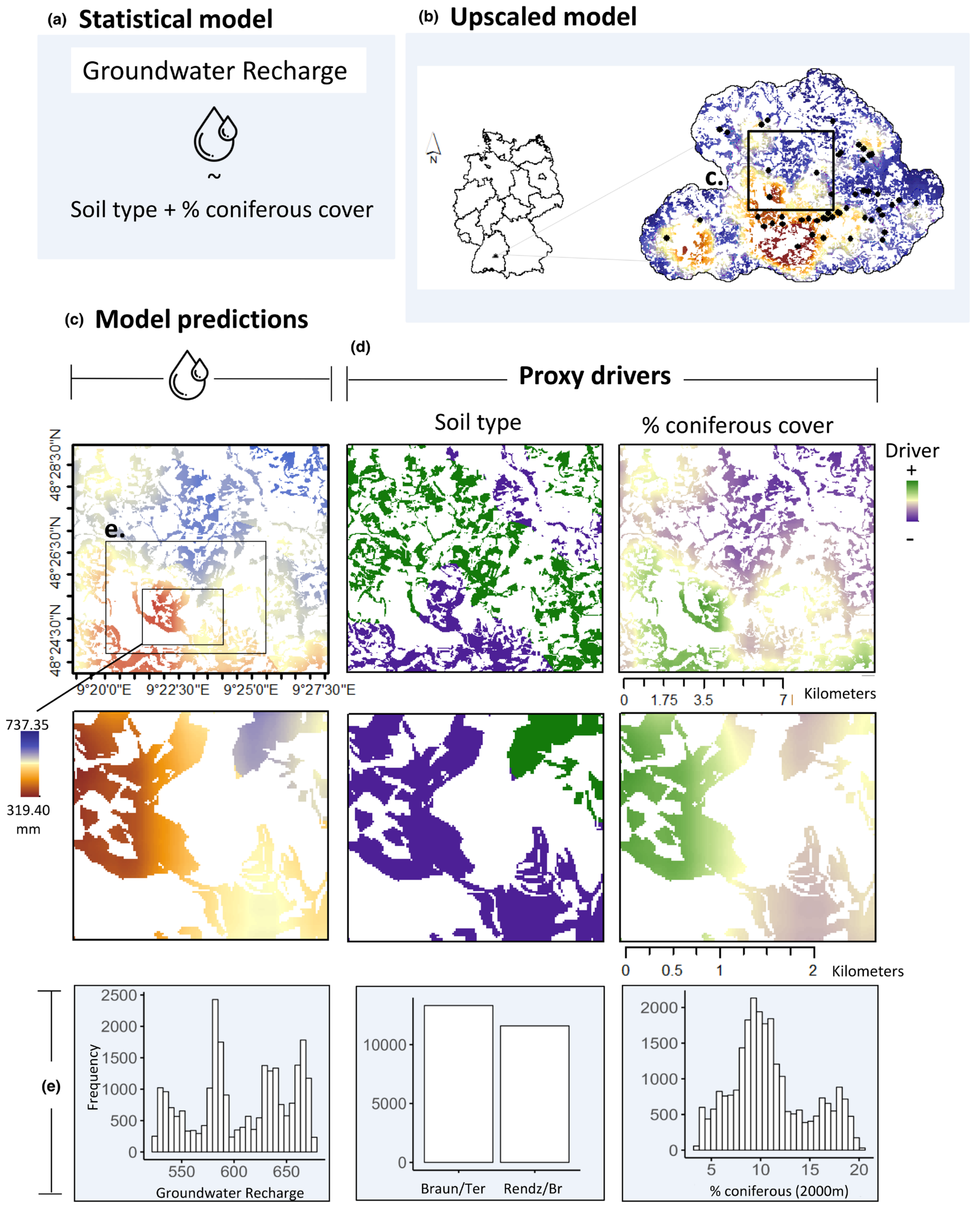

FIGURE 5 Upscaled predictions for water supply using grassland groundwater recharge as indicator. We present the final upscaling model in (a). In (b), we present the region and the upscaled water supply with a zoom in (c). The dots in (b) represent the training data plot locations. (d) Presents the drivers used to model groundwater recharge. The middle row shows a closer zoom to show how surrounding coniferous forest cover alters groundwater recharge. The bottom row (e) shows the data extract from the closer zoom (larger square), with predicted values of groundwater recharge and its drivers (soil type, as Braunerde-Terra Fusca [Brau/Terr] and Rendzine/Braunerde [Rend/Br], and coniferous cover). Predictions for areas mapped white were not calculated (e.g. arable land and forest).

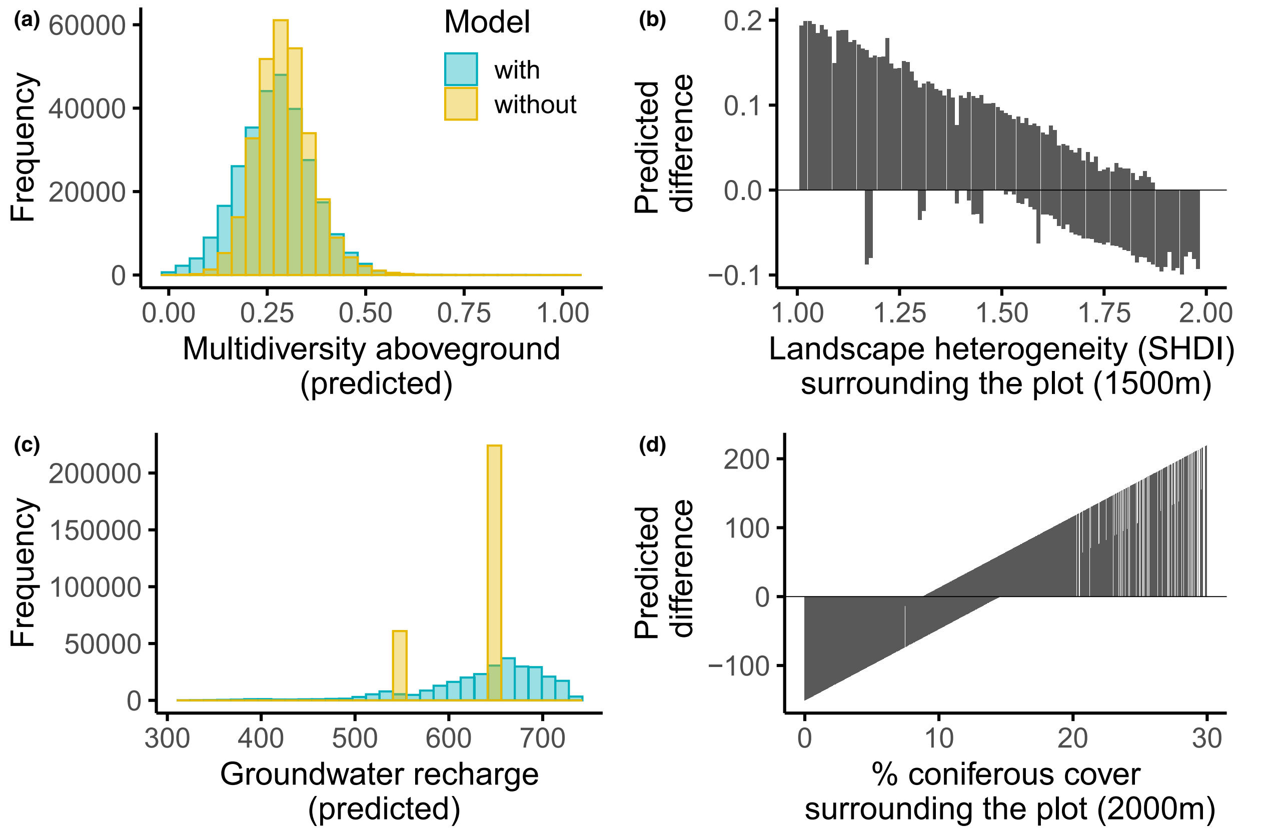

FIGURE 6 Predicted estimates including and excluding spatial interaction terms. (a) Predictions of multidiversity above-ground (biodiversity conservation) with and without spatial interactions, as represented by landscape variables, on predicted estimates of multidiversity above-ground along a landscape heterogeneity change gradient (SHDI); (c) predicted estimates of groundwater recharge (water supply) with and without spatial interactions; (d) difference (without spatial interactions—with spatial interactions) on predicted estimates along the gradient of coniferous cover.

when applied beyond the range of the training data. Therefore, we recommend that predictions be constrained to the same sampling domain, even if this limits the spatial extent. Equally important is the clear communication of prediction limitations, which helps prevent overinterpretation and supports more responsible ecological inference.

In terms of broader transferability, we expect that our framework is applicable to a wide range of social-ecological systems, ecosystem types and regulating, provisioning and supporting ES. We advise caution when addressing cultural services as the drivers and indicators for these often require explicit consideration of human perspectives (details in Supporting Information S6). Application of our framework is likely to be most effective when detailed plot-level information relating to ES supply and its drivers are available to inform upscaling. Such data are increasingly available (e.g. NutNet https://nutnet.org/; Biodiversity Exploratories https://www.biodiversity-exploratories.de/en/; FunDivEurope http://project.fundiveurope.eu/) but we also acknowledge that such data is absent from many regions.

A lack of high-quality data on the drivers might restrict evaluation to a small set of potential drivers, and omit the most important—a common drawback of upscaling approaches. In addition, where the spatial resolution of landscape predictor data is too large (e.g. >100 m), spatial interactions will be hard to assess since heterogeneity is averaged across, and data may fail to capture important features (e.g. small habitat patches). In such cases, the hierarchy of controls approach can still be applied, but terms representing local fine-scale spatial interactions are unlikely to be included.

The spatial interactions and hierarchy of drivers aspects of our methodological framework may also be adapted to cases beyond upscaling where ES maps are sought but high-quality ES indicator data is absent. Here, it may be possible to formulate ES production functions based on scientific literature and other forms of information (e.g. reported yields and market values) and/or expert knowledge instead of selecting variables in statistical modelling. For instance, there is evidence from global meta-analysis that many ESs, including pollination, respond positively to natural/semi-natural habitat cover (Mohamed et al., 2024; Ulyshen et al., 2023). Thus, instead of assigning the same values of pollination for all semi-natural areas, the areas with higher cover of natural/semi-natural habitat within an appropriate radius can be assigned higher values (for similar approaches using other ES; see Chan et al., 2006; Egoh et al., 2008). Another option would be downscaling Earth Observation measures (EO-based ES) for some indicators, but further methods development and testing would be required. The framework and approach is also in principle scalable to larger scales (such as national or continental scales) but modelling might become both time and computationally demanding if fine-scale spatial interactions are to be included.

In a broader comparison to alternative techniques to obtaining mapped ES supply estimates, we believe that our upscaling approach forms a half-way house between biophysical simulation models (e.g. INVEST; Grafius et al., 2018) and machine learning based model formulation (e.g. Pichler & Hartig, 2023), since it has elements of the mechanistic realism of the former but also the tailoring of model structure to fit local data of the latter. If done well, it may possess both these advantages but if performed without thought it may offer little advantage over these approaches, which may be better suited to very well understood ecosystem process driven services (e.g. carbon storage), and poorly understood services, respectively.

Several avenues for future research were identified during the framework development. First, although the inclusion of spatial interactions aims to represent the underlying mechanisms of ES provision, it may be possible that in some cases, these can only be accurately represented using process-based models. At the other extreme in the absence of good mechanistic knowledge, machine learning models may outperform our approach. Therefore, it would be insightful to systematically test the performance of the approach across different social ecological systems and ES with varying data resolution and quality, and to compare its performance to alternative upscaling approaches (e.g. machine learning and process-based models). Second, we propose that using models that incorporate causal effects between drivers across the hierarchy, such as in a structural equation model (Grace, 2006) could potentially lead to improved models that are better suited to forecasting the impact of environmental change.

6 | CONCLUSION

Our upscaling framework advances upon existing ES upscaling approaches by modelling ES in a way can capture the true underlying mechanisms and which incorporates spatial interactions between landscape components. The resulting maps have the potential to more reliably inform policy decisions, and spatial planning, for example in prioritizing areas for grants or subsidies to support ES provision, or spatial targeting of monitoring schemes. The framework may be particularly appropriate when the upscaling of ES supply aims to inform decision management, since the approach allows users to simulate the impact of change in the drivers on ES supply by substituting predictor layers in the GIS (e.g. adding a future climate or land cover layer).

This approach, when incorporating spatial interactions, also allows assessment and quantification of the impact of compositional and configurational changes within landscapes at larger scales than they are usually assessed in landscape ecology research. For instance, in the case of our examples, the diversity of land-use types could be changed within a GIS model to assess the impact on biodiversity conservation, and changes in coniferous forest cover could also be simulated to predict changes in groundwater recharge. If widely applied, the proposed approach may form a bridge between detailed local scale ecological studies in which mechanisms are well-understood, and larger scale ES research, which tends to be performed at scales relevant to policy and planning.

AUTHOR CONTRIBUTIONS

Andrea Larissa Boesing and Peter Manning conceived the idea and methodology. Andrea Larissa Boesing analysed the data and led the manuscript writing; Gaëtane Le Provost, Kirsten Jung, Maximilian Lange, Sophia Leimer, Steffen Boch, Till Kleinebecker, Ute Hamer, Valentin H. Klaus and Wolfgang Wilcke contributed discussing the models; Anja Linstädter, Javier Muro, Jörg Müller, Markus Fischer, Margot Neyret, Maximilian Lange, Olena Dubovyk, Paul Magdon, Ralph Bolliger, Sophia Leimer, Steffen Boch, Swen Renner, Till Kleinebecker and Wolfgang Wilcke contributed with data. All authors commented upon the paper and contributed discussing the models.

AFFILIATIONS

[1]Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany; [2]INRAE, Bordeaux Science Agro, ISVV, SAVE, Villenave d'Ornon, France; [3]Alpine Ecology Laboratory, Université Grenoble Alpes—CNRS—Université Savoie Mont Blanc, Gières, Grenoble, France; [4]Department of Biodiversity Research/Systematic Botany, University of Potsdam, Potsdam, Germany; [5]Institute of Farm Economics, Thünen Institut, Braunschweig, Germany; [6]Heinz Sielmann Foundation, Elstal, Germany; [7]Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany; [8]Institute of Plant Sciences, University of Bern, Bern, Switzerland; [9]Department of Remote Sensing, Helmholtz-Centre for Environmental Research UFZ, Leipzig, Germany; [10]Earth System Department, University of Hamburg, Hamburg, Germany; [11]Hochschule für Angewandte Wissenschaft und Kunst (HAWK), Göttingen, Germany; [12]Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany; [13]WSL Swiss Federal Research Institute, Birmensdorf, Switzerland; [14]Natural History Museum Vienna, Wien, Austria; [15]Institute for Landscape Ecology and Resources Management, Justus Liebig University Giessen, Giessen, Germany; [16]Institute of Landscape Ecology, University of Münster, Münster, Germany; [17]Forage Production and Grassland Systems, Agroscope, Zurich, Switzerland; [18]Institute of Geography, Ruhr University Bochum, Bochum, Germany and [19]Department of Biological Sciences, Centre for Sustainable Area Management (CeSAM), University of Bergen, Bergen, Norway

ACKNOWLEDGEMENTS

We thank the current and former managers of the Schwäbische Alb Exploratory Konstans Wells, Swen Renner, Kirsten Reichel-Jung, Iris Steitz, Sandra Weithmann, Florian Staub, Julia Bass and Max Müller for their work in maintaining the plot and project infrastructure; the current and former BEO coordinators Victoria Grießmeier, Christiane Fischer, Jule Mangels, Simone Pfeiffer and

Maren Gleisberg for giving support through the central office; the current and former data manager coordinators Andreas Ostrowski, Michael Owonibi and Jens Nieschulze for managing the central database; Eric Allan for supporting with both tools and data, and Eduard Linsenmair, Dominik Hessenmöller, Daniel Prati, Ingo Schöning, François Buscot, Ernst-Detlef Schulze, Wolfgang W. Weisser and the late Elisabeth Kalko for their role in setting up the Biodiversity Exploratories project. We thank the administration of the UNESCO Biosphere Reserve Swabian Alb as well as all land owners for the excellent collaboration. The work has been (partly) funded by the DFG Priority Program 1374 'BiodiversityExploratories', specifically the BEF-Up II DFG contributing project (MA 7144/1-2). AL and OD acknowledge funding via the SEBAS DFG contributing project (LI 1842/4-1 and DU 1596/1-1). VHK acknowledges funding by the Agroscope research program *Indicate*. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest. They confirm the work is their own and has not been published or submitted to any other journal. The manuscript contains ideas from their own research project and data collection and analysis meet scientific quality standards.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70281.

DATA AVAILABILITY STATEMENT

This work is based on data from several projects of the Biodiversity Exploratories program (DFG Priority Program 1374). The assembled data and code to reproduce the framework are available in the Biodiversity Exploratories Information System (Boesing, 2025; https://www.bexis.uni-jena.de/ddm/data/Showdata/32138).

ORCID

Andrea Larissa Boesing https://orcid.org/0000-0002-7467-4281
Gaëtane Le Provost https://orcid.org/0000-0002-1643-6023
Margot Neyret https://orcid.org/0000-0001-9435-1634
Kirsten Jung https://orcid.org/0000-0002-9449-2215
Ralph Bolliger https://orcid.org/0000-0001-5383-9713
Sophia Leimer https://orcid.org/0000-0001-6272-204X
Till Kleinebecker https://orcid.org/0000-0003-1121-2861
Peter Manning https://orcid.org/0000-0002-7940-2023

REFERENCES

Adane, Z. A., & Gates, J. B. (2015). Determining the impacts of experimental forest plantation on groundwater recharge in the Nebraska Sand Hills (USA) using chloride and sulfate. *Hydrogeology Journal*, *23*, 81–94.

Allan, E., Bossdorf, O., Dormann, C. F., Prati, D., Gossner, M. M., Tscharntke, T., Blüthgen, N., Bellach, M., Birkhofer, K., Boch, S., Böhm, S., Börschig, C., Chatzinotas, A., Christ, S., Daniel, R., Diekötter, T., Fischer, C., Friedl, T., Glaser, K., … Fischer, M. (2014). Interannual variation in land-use intensity enhances grassland multidiversity. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 308–313.

Allen, T. F. H., & Starr, T. B. (1982). *Hierarchy: Perspectives for ecological complexity* (2nd ed.). University of Chicago Press.

Bateman, I. J., & Mace, G. M. (2020). The natural capital framework for sustainably efficient and equitable decision making. *Nature Sustainability*, *3*, 776–783.

Bellot, J., Sanchez, J. R., Chirino, E., Hernandez, N., Abdelli, F., & Martinez, J. M. (1999). Effect of different vegetation type cover on the soil water balance in semi-arid areas of south eastern Spain. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, *24*, 353–357.

Blüthgen, N., Dormann, C. F., Prati, D., Klaus, V. H., Kleinebecker, T., Hölzel, N., Alt, F., Boch, S., Gockel, S., Hemp, A., Müller, J., Nieschulze, J., Renner, S. C., Schöning, I., Schumacher, U., Socher, S. A., Wells, K., Birkhofer, K., Buscot, F., … Weisser, W. W. (2012). A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. *Basic and Applied Ecology*, *13*, 207–220.

Boesing, A. L. (2025). *Data and code: Accounting for spatial interactions in the upscaling of ecosystem services. Version 8*. Biodiversity Exploratories Information System. https://www.bexis.uni-jena.de/ddm/data/Showdata/32138

Boesing, A. L., Klaus, V. H., Neyret, M., Le Provost, G., Peter, S., Fischer, M., & Manning, P. (2024). Identifying the optimal landscape configuration for landscape multifunctionality. *Ecosystem Services*, *67*, 101630.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*, 108–132.

Bruins, R. J. F., Canfield, T. J., Duke, C., Kapustka, L., Nahlik, A. M., & Schäfer, R. B. (2017). Using ecological production functions to link ecological processes to ecosystem services. *Integrated Environmental Assessment and Management*, *13*, 52–61.

Cavender-Bares, J., Schneider, F. D., Santos, M. J., Armstrong, A., Carnaval, A., Dahlin, K. M., Fatoyinbo, L., Hurtt, G. C., Schimel, D., Townsend, P. A., Ustin, S. L., Wang, Z., & Wilson, A. M. (2022). Integrating remote sensing with ecology and evolution to advance biodiversity conservation. *Nature Ecology & Evolution*, *6*, 506–519.

Chan, K. M. A., Shaw, M. R., Cameron, D. R., Underwood, E. C., & Daily, G. C. (2006). Conservation planning for ecosystem services. *PLoS Biology*, *4*, 2138–2152.

Civantos-Gómez, I., García-Algarra, J., García-Callejas, D., Galeano, J., Godoy, O., & Bartomeus, I. (2021). Fine scale prediction of ecological community composition using a two-step sequential machine learning ensemble. *PLoS Computational Biology*, *17*, 1–20.

Cole, L. J., Stockan, J., & Helliwell, R. (2020). Managing riparian buffer strips to optimise ecosystem services: A review. *Agriculture, Ecosystems and Environment*, *296*, 106891.

Dallimer, M., & Strange, N. (2015). Why socio-political borders and boundaries matter in conservation. *Trends in Ecology & Evolution*, *30*, 132–139.

Díaz, S., Lavorel, S., De Bello, F., Quétier, F., Grigulis, K., & Robson, T. M. (2007). Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 20684–20689.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., García Marquéz, J. R., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr,

F. M., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, *30*, 609–628.

Egoh, B., Reyers, B., Rouget, M., Richardson, D. M., Le Maitre, D. C., & van Jaarsveld, A. S. (2008). Mapping ecosystem services for planning and management. *Agriculture, Ecosystems and Environment*, *127*, 135–140.

Eigenbrod, F., Armsworth, P. R., Anderson, B. J., Heinemeyer, A., Gillings, S., Roy, D. B., Thomas, C. D., & Gaston, K. J. (2010). The impact of proxy-based methods on mapping the distribution of ecosystem services. *Journal of Applied Ecology*, *47*, 377–385.

Fischer, M., Bossdorf, O., Gockel, S., Hansel, F., Hemp, A., Hessenmoller, D., Korte, G., Nieschulze, J., Pfeiffer, S., Prati, D., Renner, S. C., Schoning, I., Schumacher, U., Wells, K., Buscot, F., Kalko, E. K. V., Linsenmair, K. E., Schulze, E.-D., & Weisser, W. W. (2010). Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. *Basic and Applied Ecology*, *11*, 473–485.

González-Chaves, A., Jaffe, R., & Metzger, J. P. (2020). Forest proximity rather than local forest cover affects bee diversity and coffee pollination services. *Landscape Ecology*, *35*, 1841–1855.

González-Chaves, A. G., Carvalheiro, L. G., Garibaldi, L. A., & Metzger, J. P. (2022). Positive forest cover effects on coffee yields are consistent across regions. *Journal of Applied Ecology*, *59*, 330–341.

Grace, J. B. (2006). *Structural equation modeling and natural systems* (p. 365). Cambridge University Press.

Grafius, D. R., Corstanje, R., & Harris, J. A. (2018). Linking ecosystem services, urban form and green space configuration using multivariate landscape metric analysis. *Landscape Ecology*, *33*, 557–573.

Griffiths, P., Nendel, C., Pickert, J., & Hostert, P. (2020). Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. *Remote Sensing of Environment*, *238*, 1–12.

Hartig, F. (2020). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. Retrieved November 16, 2020, from https://cran.r-project.org/package=DHARMa

Hesselbarth, M. H. K., Sciaini, M., With, K. A., Wiegand, K., & Nowosad, J. (2019). Landscapemetrics: An open-source R tool to calculate landscape metrics. *Ecography*, *42*, 1648–1657.

IPBES. (2019). *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. IPBES secretariat.

Jackson, H. B., & Fahrig, L. (2015). Are ecologists conducting research at the optimal scale? *Global Ecology and Biogeography*, *24*, 52–63.

Kreutziger, Y. (2006). *Rückkopplungseffekte verschieden diverser Grünlandökosysteme auf die Komponenten des Bodenwasserhaushalts an einem Auestandort der Saale* [Dissertation]. Jena.

Lange, M., Feilhauer, H., Kühn, I., & Doktor, D. (2022). Mapping land-use intensities of grasslands in Germany with machine learning algorithms using Copernicus Sentinel-2 optical satellite data. *Remote Sensing of Environment*, *277*, 112888.

Lavorel, S., & Grigulis, K. (2012). How fundamental plant functional trait relationships scale-up to trade-offs and synergies in ecosystem services. *Journal of Ecology*, *100*, 128–140.

Lavorel, S., Grigulis, K., Lamarque, P., Colace, M.-P., Garden, D., Girel, J., Pellet, G., & Douzet, R. (2011). Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology*, *99*, 135–147.

Le Clec'h, S., Finger, R., Buchmann, N., Gosal, A. S., Hörtnagl, L., Huguenin-Elie, O., Jeanneret, P., Lüscher, A., Schneider, M. K., & Huber, R. (2019). Assessment of spatial variability of multiple ecosystem services in grasslands of different intensities. *Journal of Environmental Management*, *251*, 109372.

Le Provost, G., Schenk, N. V., Penone, C., Thiele, J., Westphal, C., Allan, E., Ayasse, M., Blüthgen, N., Boeddinghaus, R. S., Boesing, A. L., Bolliger, R., Busch, V., Fischer, M., Gossner, M. M., Hölzel, N., Jung, K., Kandeler, E., Klaus, V. H., Kleinebecker, T., … Manning, P. (2023). The supply of multiple ecosystem services requires biodiversity across spatial scales. *Nature Ecology & Evolution*, *7*, 236–249.

Le Provost, G., Thiele, J., Westphal, C., Penone, C., Allan, E., Neyret, M., Le Provost, G., van der Plas, F., Ayasse, M., Bardgett, R. D., Birkhofer, K., Boch, S., Bonkowski, M., Buscot, F., Feldhaar, H., Gaulton, R., Goldmann, K., Gossner, M. M., Klaus, V. H., … Manning, P. (2021). Contrasting responses of above- and belowground diversity to multiple components of land-use intensity. *Nature Communications*, *12*, 3918.

Leimer, S., Bischoff, S., Boch, S., Busch, V., Escher, P., Fischer, M., Hänsel, F., Hölzel, N., Kerber, K., Klaus, V. H., Kleinebecker, T., Michalzik, B., Nauss, T., Schäfer, D., Schöning, I., Schwarz, M. T., Siemens, J., Thieme, L., Wöllauer, S., & Wilcke, W. (2018). Does plant diversity affect the water balance of established grassland systems? *Ecohydrology*, *11*, e1945.

Leimer, S., Kreutziger, Y., Rosenkranz, S., Beßler, H., Engels, C., Hildebrandt, A., Oelmann, Y., Weisser, W. W., Wirth, C., & Wilcke, W. (2014). Plant diversity effects on the water balance of an experimental grassland. *Ecohydrology*, *7*, 1378–1391.

Linders, T. E. W., Schaffner, U., Alamirew, T., Allan, E., Choge, S. K., Eschen, R., Shiferaw, H., & Manning, P. (2021). Stakeholder priorities determine the impact of an alien tree invasion on ecosystem multifunctionality. *People and Nature*, *3*, 658–672.

Maes, J., Egoh, B., Willemen, L., Liquete, C., Vihervaara, P., Schägner, J. P., Grizzetti, B., Drakou, E. G., Notte, A. L., Zulian, G., Bouraoui, F., Luisa Paracchini, M., Braat, L., & Bidoglio, G. (2012). Mapping ecosystem services for policy support and decision making in the European Union. *Ecosystem Services*, *1*, 31–39.

Manning, P., de Vries, F. T., Tallowin, J. R. B., Smith, R., Mortimer, S. R., Pilgrim, E. S., Harrison, K. A., Wright, D. G., Quirk, H., Benson, J., Shipley, B., Cornelissen, J. H. C., Kattge, J., Bönisch, G., Wirth, C., & Bardgett, R. D. (2015). Simple measures of climate, soil properties and plant traits predict national-scale grassland soil carbon stocks. *Journal of Applied Ecology*, *52*, 1188–1196.

Martin, E. A., Dainese, M., Clough, Y., Báldi, A., Bommarco, R., Gagic, V., Garratt, M. P. D., Holzschuh, A., Kleijn, D., Kovács-Hostyánszki, A., Marini, L., Potts, S. G., Smith, H. G., al Hassan, D., Albrecht, M., Andersson, G. K. S., Asís, J. D., Aviron, S., Balzan, M. V., … Steffan-Dewenter, I. (2019). The interplay of landscape composition and configuration: New pathways to manage functional biodiversity and agroecosystem services across Europe. *Ecology Letters*, *22*, 1083–1094.

Martínez, O. J. A., Fremier, A. K., Günter, S., Bendaña, Z. R., Vierling, L., Galbraith, S. M., Bosque-Pérez, N. A., & Ordoñez, J. C. (2016). Scaling up functional traits for ecosystem services with remote sensing: Concepts and methods. *Ecology and Evolution*, *6*, 4359–4371.

Martínez-Harms, M. J., Quijas, S., Merenlender, A. M., & Balvanera, P. (2016). Enhancing ecosystem services maps combining field and environmental data. *Ecosystem Services*, *22*, 32–40.

McGuire, A. D., Sitch, S., Clein, J. S., Dargaville, R., Esser, G., Foley, J., Heimann, M., Joos, F., Kaplan, J., Kicklighter, D. W., Meier, R. A., Melillo, J. M., Moore, B., Prentice, I. C., Ramankutty, N., Reichenau, T., Schloss, A., Tian, H., Williams, L. J., & Wittenberg, U. (2001). Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO_2, climate and land use effects with four process-based ecosytem models. *Global Biogeochemical Cycles*, *15*, 183–206.

Metzger, J. P., Villarreal-Rosas, J., Suárez-Castro, A. F., López-Cubillos, S., González-Chaves, A., Runting, R. K., Hohlenwerger, C., & Rhodes, J. R. (2021). Considering landscape-level processes in ecosystem service assessments. *Science of the Total Environment*, *796*, 149028.

Meyer, H., & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, *12*, 1620–1633.

Miller, J. A. (2012). Species distribution models: Spatial autocorrelation and non-stationarity. *Progress in Physical Geography*, *36*, 681–692.

Mitchell, M. G. E., Suarez-Castro, A. F., Martinez-Harms, M., Maron, M., McAlpine, C., Gaston, K. J., Mitchell, M. G., Johansen, K., & Rhodes, J. R. (2015). Reframing landscape fragmentation's effects on ecosystem services. *Trends in Ecology & Evolution*, *30*, 190–198.

Mohamed, A., DeClerck, F., Verburg, P. H., Obura, D., Abrams, J. F., Zafra-Calvo, N., Rocha, J., Estrada-Carmona, N., Fremier, A., Jones, S. K., Meier, I. C., & Stewart-Koster, B. (2024). Securing nature's contributions to people requires at least 20%–25% (semi-)natural habitat in human-modified landscapes. *One Earth*, *7*, 59–71.

Müller, F. (1992). Hierarchical approaches to ecosystem theory. *Ecological Modelling*, *63*, 215–242.

Müller, F., & Burkhard, B. (2012). The indicator side of ecosystem services. *Ecosystem Services*, *1*, 26–30.

Muro, J., Linstädter, A., Magdon, P., Wöllauer, S., Männer, F. A., Schwarz, L.-M., Ghazaryan, G., Schultz, J., Malenovský, Z., & Dubovyk, O. (2022). Predicting plant biomass and species richness in temperate grasslands across regions, time, and land management with remote sensing and deep learning. *Remote Sensing of Environment*, *282*, 113262.

Nelson, E., Mendoza, G., Regetz, J., Polasky, S., Tallis, H., Cameron, D. R., Chan, K. M. A., Daily, G. C., Goldstein, J., Kareiva, P. M., Lonsdorf, E., Naidoo, R., Ricketts, T. H., & Shaw, M. R. (2009). Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment*, *7*, 4–11.

Neyret, M., Fischer, M., Allan, E., Hölzel, N., Klaus, V. H., Kleinebecker, T., Krauss, J., Le Provost, G., Peter, S., Schenk, N., Simons, N. K., van der Plas, F., Binkenstein, J., Börschig, C., Jung, K., Prati, D., Schäfer, D., Schäfer, M., Schöning, I., … Manning, P. (2021). Assessing the impact of grassland management on landscape multifunctionality. *Ecosystem Services*, *52*, 101366.

Neyret, M., Peter, S., Le Provost, G., Boch, S., Boesing, A. L., Bullock, J. M., Le Provost, G., Hölzel, N., Klaus, V. H., Kleinebecker, T., Krauss, J., Müller, J., Müller, S., Ammer, C., Buscot, F., Ehbrecht, M., Fischer, M., Goldmann, K., Jung, K., … Manning, P. (2023). Landscape management strategies for multifunctionality and social equity. *Nature Sustainability*, *6*, 391–403.

O'Brien, S. A., Tylianakis, J. M., Anderson, D. P., Boesing, A. L., Lai, H. R., Le Provost, G., Le Provost, G., Manning, P., Neyret, M., Blüthgen, N., Jung, K., Magdon, P., Müller, S., Schenk, N. V., Scherer-Lorenzen, M., & Lavorel, S. (2025). Mobile species' responses to surrounding land use generate trade-offs and synergies among nature's contributions to people. *Proceedings of the National Academy of Sciences*, *122*, e2505401122.

Paradis, E., & Schliep, K. (2019). Ape 5.0: An enviroment for modern phylogenetics and evolutionary analysis in R. *Bioinformatics*, *35*, 526–528.

Perry, G. L. W., Seidl, R., Bellvé, A. M., & Rammer, W. (2022). An outlook for deep learning in ecosystem science. *Ecosystems*, *25*, 1700–1718.

Peter, S., Le Provost, G., Mehring, M., Müller, T., & Manning, P. (2022). Cultural worldviews consistently explain bundles of ecosystem service prioritisation across rural Germany. *People and Nature*, *4*, 218–230.

Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, *14*, 994–1016.

Plummer, M. L. (2009). Assessing benefit transfer for the valuation of ecosystem services. *Frontiers in Ecology and the Environment*, *7*, 38–45.

Potschin, M. B., & Haines-Young, R. H. (2011). Ecosystem services: Exploring a geographical perspective. *Progress in Physical Geography*, *35*, 575–594.

Ranta, E., Lindström, J., Kaitala, V., Kokko, H., Linden, H., & Helle, E. (1997). Solar activity and hare dynamics: A cross-continental comparision. *American Naturalist*, *149*, 765–775.

Redlich, S., Zhang, J., Benjamin, C., Dhillon, M. S., Englmeier, J., Ewald, J., Fricke, U., Ganuza, C., Haensel, M., Hovestadt, T., Kollmann, J., Koellner, T., Kübert-Flock, C., Kunstmann, H., Menzel, A., Moning, C., Peters, W., Riebl, R., Rummler, T., … Steffan-Dewenter, I. (2022). Disentangling effects of climate and land use on biodiversity and ecosystem services—A multi-scale experimental design. *Methods in Ecology and Evolution*, *13*, 514–527.

Scowen, M., Athanasiadis, I. N., Bullock, J. M., Eigenbrod, F., & Willcock, S. (2021). The current and future uses of machine learning in ecosystem service research. *Science of the Total Environment*, *799*, 149263.

Smith, P., Ashmore, M. R., Black, H. I. J., Burgess, P. J., Evans, C. D., Quine, T. A., Thomson, A. M., Hicks, K., & Orr, H. G. (2013). REVIEW: The role of ecosystems and their management in regulating climate, and soil, water and air quality. *Journal of Applied Ecology*, *50*, 812–829.

Soliveres, S., Van Der Plas, F., Manning, P., Prati, D., Gossner, M. M., Renner, S. C., Alt, F., Arndt, H., Baumgartner, V., Binkenstein, J., Birkhofer, K., Blaser, S., Blüthgen, N., Boch, S., Böhm, S., Börschig, C., Buscot, F., Diekötter, T., Heinze, J., … Allan, E. (2016). Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature*, *536*, 456–459.

Spake, R., Bellamy, C., Graham, L. J., Watts, K., Wilson, T., Norton, L. R., Wood, C. M., Schmucki, R., Bullock, J. M., & Eigenbrod, F. (2019). An analytical framework for spatially targeted management of natural capital. *Nature Sustainability*, *2*, 90–97.

Spake, R., Lasseur, R., Crouzat, E., Bullock, J. M., Lavorel, S., Parks, K. E., Schaafsma, M., Bennett, E. M., Maes, J., Mulligan, M., Mouchet, M., Peterson, G. D., Schulp, C. J. E., Thuiller, W., Turner, M. G., Verburg, P. H., & Eigenbrod, F. (2017). Unpacking ecosystem service bundles: Towards predictive mapping of synergies and trade-offs between ecosystem services. *Global Environmental Change*, *47*, 37–50.

Thakur, M. P., van der Putten, W. H., Wilschut, R. A., Veen, G. F. C., Kardol, P., van Ruijven, J., Allan, E., Roscher, C., van Kleunen, M., & Bezemer, T. M. (2021). Plant–soil feedbacks and temporal dynamics of plant diversity-productivity relationships. *Trends in Ecology & Evolution*, *36*, 651–661.

Tscharntke, T., Tylianakis, J. M., Rand, T. A., Didham, R. K., Fahrig, L., Bengtsson, J., Batáry, P., Clough, Y., Crist, T. O., Dormann, C. F., Ewers, R. M., Fründ, J., Holt, R. D., Holzschuh, A., Klein, A. M., Kleijn, D., Kremen, C., Landis, D. A., Laurance, W., … Westphal, C. (2012). Landscape moderation of biodiversity patterns and processes - Eight hypotheses. *Biological Reviews*, *87*, 661–685.

Ulyshen, M., Urban-Mead, K. R., Dorey, J. B., & Rivers, J. W. (2023). Forests are critically important to global pollinator diversity and enhance pollination in adjacent crops. *Biological Reviews*, *98*, 1–24.

Veach, V., Di Minin, E., Pouzols, F. M., & Moilanen, A. (2017). Species richness as criterion for global conservation area placement leads to large losses in coverage of biodiversity. *Diversity and Distributions*, *23*, 715–726.

Verhagen, W., Van Teeffelen, A. J. A., Baggio Compagnucci, A., Poggio, L., Gimona, A., & Verburg, P. H. (2016). Effects of landscape configuration on mapping ecosystem service capacity: A review of evidence and a case study in Scotland. *Landscape Ecology*, *31*, 1457–1479.

Villa, F., Bagstad, K. J., Voigt, B., Johnson, G. W., Portela, R., Honzák, M., & Batker, D. (2014). A methodology for adaptable and robust ecosystem services assessment. *PLoS One*, *9*, e91001.

Willcock, S., Martínez-López, J., Hooftman, D. A. P., Bagstad, K. J., Balbi, S., Marzo, A., Prato, C., Sciandrello, S., Signorello, G., Voigt, B., Villa, F., Bullock, J. M., & Athanasiadis, I. N. (2018). Machine learning for ecosystem services. *Ecosystem Services*, *33*, 165–174.

Wolff, S., Schulp, C. J. E., & Verburg, P. H. (2015). Mapping ecosystem services demand: A review of current research and future perspectives. *Ecological Indicators*, *55*, 159–171.

Woodcock, B. A., Garratt, M. P. D., Powney, G. D., Shaw, R. F., Osborne, J. L., Soroka, J., Lindström, S. A. M., Stanley, D., Ouvrard, P., Edwards, M. E., Jauker, F., McCracken, M. E., Zou, Y., Potts, S. G., Rundlöf, M.,

Methods in Ecology and Evolution

Noriega, J. A., Greenop, A., Smith, H. G., Bommarco, R., … Pywell, R. F. (2019). Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. *Nature Communications*, *10*, 1–10.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting Information S1. Area delimitation.

Supporting Information S2. ES indicators used to measure Biodiversity Conservation and Groundwater recharge.

Supporting Information S3. Mechanistic drivers of Ecosystem Services supply (Biodiversity conservation and Groundwater recharge).

Supporting Information S4. Drivers and correlations.

Supporting Information S5. Spatial autocorrelation.

Supporting Information S6. Applying the upscaling approach for a range of ecosystem services.

Supporting Information S7. R tutorials.

How to cite this article: Boesing, A. L., Le Provost, G., Neyret, M., Linstädter, A., Muro, J., Müller, J., Jung, K., Fischer, M., Lange, M., Dubovyk, O., Magdon, P., Bolliger, R., Leimer, S., Boch, S., Renner, S., Kleinebecker, T., Hamer, U., Klaus, V. H., Wilcke, W., & Manning, P. (2026). Accounting for spatial interactions in the upscaling of ecosystem services. *Methods in Ecology and Evolution*, *00*, 1–17. https://doi.org/10.1111/2041-210x.70281