# Landscape management for grassland multifunctionality

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., Börshig C., Jung K., Prati D., Schäfer M., Schäfer D., Schöning I., Schrumpf M., Tschapka M., Westphal C. & Manning P.

## Summary

Land-use intensification has contrasting effects on different ecosystem services, often leading to land-use conflicts. Multiple studies, especially within the 'land-sharing versus land-sparing' debate, have demonstrated how landscape-scale strategies can minimise the trade-off between agricultural production and biodiversity conservation. However, little is known about which land-use strategies maximise the landscape-level supply of multiple ecosystem services (landscape multifunctionality), a common goal of stakeholder communities. Here, we combine data collected from 150 grassland sites with a simulation approach to identify landscape compositions, with differing proportions of low-, medium-, and high-intensity grasslands, that minimise trade-offs between the four main grassland ecosystem services demanded by stakeholders: biodiversity conservation, aesthetic value, productivity and carbon storage.

We show that optimisation becomes increasingly difficult as more services are considered, due to varying responses of individual services to land-use intensity and the confounding effects of other environmental drivers. Thus, our results show that simple land-use strategies cannot deliver high levels of all services, making hard choices inevitable when there are trade-offs between multiple services. However, if moderate service levels are deemed acceptable, then strategies similar to the 'land-sparing' approach can deliver landscape multifunctionality. Given the sensitivity of our results on these factors we provide an online tool that identifies strategies based on user-defined demand for each service (https://neyret.shinyapps.io/landscape_composition_for_multifunctionality/). Such a tool can aid informed decision making and allow for the roles of stakeholder demands and biophysical trade-offs to be understood by scientists and practitioners alike.

# Introduction

Habitat conversion and land-use intensification are driving biodiversity loss and changes to ecosystem service supply across the world (IPBES 2019). While high land-use intensity promotes a small number of ecosystem services related to food production, it is often detrimental to biodiversity conservation (Anderson et al., 2009; Bennett et al., 2009; Lavorel et al., 2011; Raudsepp-Hearne et al., 2010) and other regulating or cultural ecosystem services that depend on biodiversity for their delivery (Allan et al., 2015; Cardinale et al., 2012; Clec'h et al., 2019; Foley, 2005; Triviño et al., 2017). Such contrasting responses of different ecosystem services to ecosystem drivers often make it impossible to achieve high levels of all desired services (i.e. ecosystem service multifunctionality, *sensu* Manning et al. (2018)) at a local scale (van der Plas et al., 2019). This has led to land-use conflicts, which are becoming increasingly common across the globe (Eastburn et al., 2017; Goldstein et al., 2012).

To date, much of the work on minimising trade-offs between ecosystem services within landscapes has compared a 'land sparing' strategy, in which semi-natural high-biodiversity areas and intensive farmland are spatially segregated, and a 'land sharing' strategy in which biodiversity conservation and commodity production are co-delivered in a landscape of intermediate intensity (Green, 2005). Within this field, most studies have found that land sparing is the best way to achieve high levels of both biodiversity conservation and commodity production (Feniuk et al., 2019; Phalan et al., 2011; Simons & Weisser, 2017). However, multiple studies have also stressed the limitations of the land sharing versus land sparing concept. The framework focuses on just two extreme strategies, and on only two services - commodity production and biodiversity conservation (Bennett, 2017; Fischer et al., 2014), while in reality, most landscapes are expected to provide multiple services, even within a single ecosystem type. This is the case for semi-natural grasslands (*sensu* Bullock et al. 2011), which supply a wide range of highly demanded ecosystem services including water provision, climate regulation (carbon storage) and recreation services, in addition to food production and biodiversity conservation (Bengtsson et al., 2019). Accounting for these additional ecosystem services could significantly affect which land-use strategies deliver multifunctionality (Knocke, 2020), but the optimal strategy for achieving high levels of multiple services within grassland landscapes remains unknown.

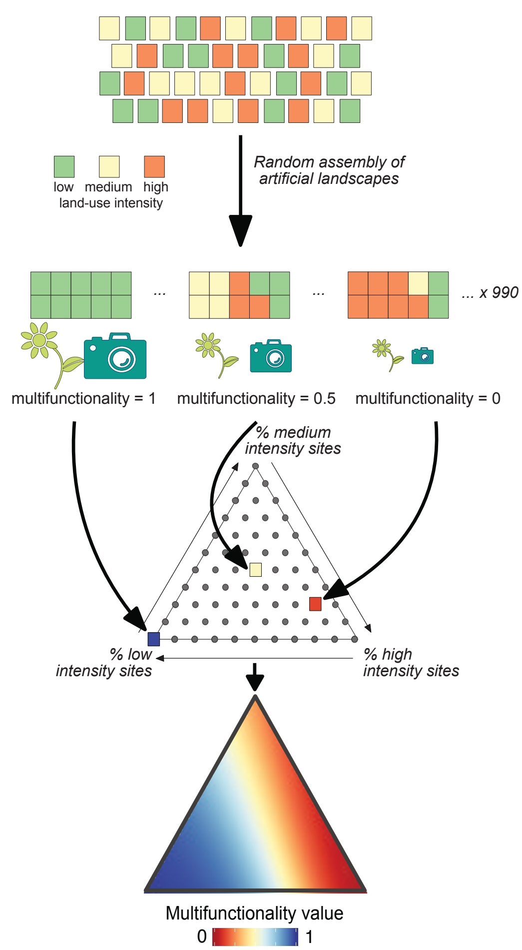
Here, we present a novel approach to identifying the optimal landscape composition for multiple ecosystem services, that involves varying the proportion of land under different intensities in data simulations. We also investigate how the levels of services demanded by land governors affect the optimal strategy. Because trade-offs between services mean that it is unlikely that all services can be maintained at high levels (Bennett et al., 2009; Raudsepp-Hearne et al., 2010; van der Plas et al 2019), managers are often faced with hard choices. To simulate the compromises that can be made we therefore generated two contrasting metrics of multifunctionality. In the first, governors choose to provide a small number of services at high levels, e.g. to meet the needs of a single or few groups to the exclusion of others (hereafter 'threshold scenario'). In the second, governors opt for a compromise situation in which all services are provided at moderate levels but without any guarantee of them being high (hereafter 'compromise scenario'). We base our metrics of multifunctionality on four services which are directly linked to final benefits (*sensu* Fisher & Turner (2008); Mace et al., (2012)): fodder production, biodiversity conservation, climate change mitigation, and aesthetic value. Among the services provided by grasslands in our study region, those four were ranked as most important by the main stakeholder groups, as identified in a social survey (Figure S 1).

The analysis was achieved by combining ecosystem service data collected at 150 grassland sites found in the three regions of the large-scale and long-term Biodiversity Exploratories project, in Germany, with a simulation approach in which artificial 'landscapes' were assembled from site-level data. We then identified the landscape composition with highest multifunctionality in each regional context (Figure 1). For each region, we divided sites into three levels of land-use intensity (Blüthgen et al. 2012). The intensity gradient was mostly driven by fertilisation and cutting frequency in the South-West and Central regions, and by grazing intensity and fertilisation in the North (Figure 2b). We then created 990 different artificial landscapes in each region, that differed in their proportions of high, medium and low intensity grassland. Indicator values for the supply of the four services were then calculated at the landscape level (see Table 1 and Methods) before calculating multifunctionality. We hypothesized that heterogenous landscapes composed of both high- and low-intensity (broadly similar to a land-sparing strategy) sites would have the highest multifunctionality when considering fodder production and biodiversity conservation (van der Plas et al., 2019).

**Table 1 Estimation of the considered ecosystem services from site-scale ecosystem service indicators.** All landscape-scale services were weighted equally within each final benefit category. Services were corrected for the effects of environmental covariates (e.g. soil texture, climate) prior to the calculation of landscape indicators.

| Ecosystem service | | Site-scale ecosystem service indicator | Landscape-scale ecosystem service indicator |
|---|---|---|---|
| | Biodiversity conservation | Number of plant species (alpha diversity)<br><br>Cover of red list species | Number of plant species (gamma diversity)<br><br>Cover of red list plant species |
| | Fodder production | Estimated biomass production (as per Simons & Weisser 2017) x plant nitrogen concentration x 6.25 (Lee, 2018) | Sum of protein production of all sites in the landscape |
| | Aesthetic value | Butterfly abundance<br>Flower cover<br>Bird richness | Average butterfly abundance in the landscape<br>Average flower cover in the landscape<br>Number of bird families (gamma diversity) in the landscape (Hedblom *et al.*, 2014) |
| $CO_2$ | Climate change mitigation (carbon storage) | C stock at 0-10cm depth | Sum of soil C stocks in the landscape |

**Figure 1 Steps of the analysis**




1. Each site from each region is classified into one of the three groups (low, medium or high intensity) based on 33% percentiles of the Land Use Intensity index.

2. For all possible landscape compositions of 10 plots (i.e. from 10 plots at low intensity to 10 plots at medium or high intensity with all possible intermediate compositions, represented by grey dots) 15 artificial landscapes are randomly assembled from the regional pool.

3. The values for multifunctionality (or single services) are calculated at the landscape level (here for two services - see Table 1).

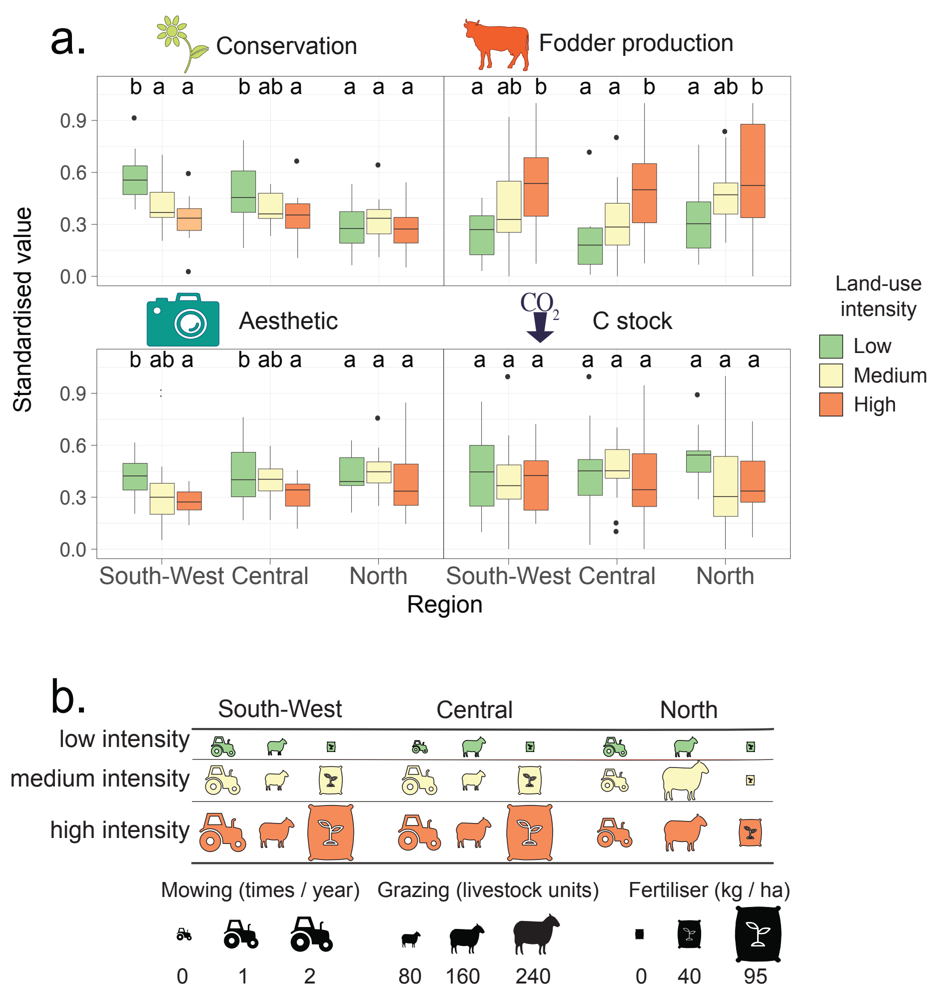
4. All the resulting values are placed in a composition space, then the multifunctionality is smoothed and predicted for all possible combinations based on binomial generalised linear models.

# Results

*Relationships between land-use intensity and ecosystem services*

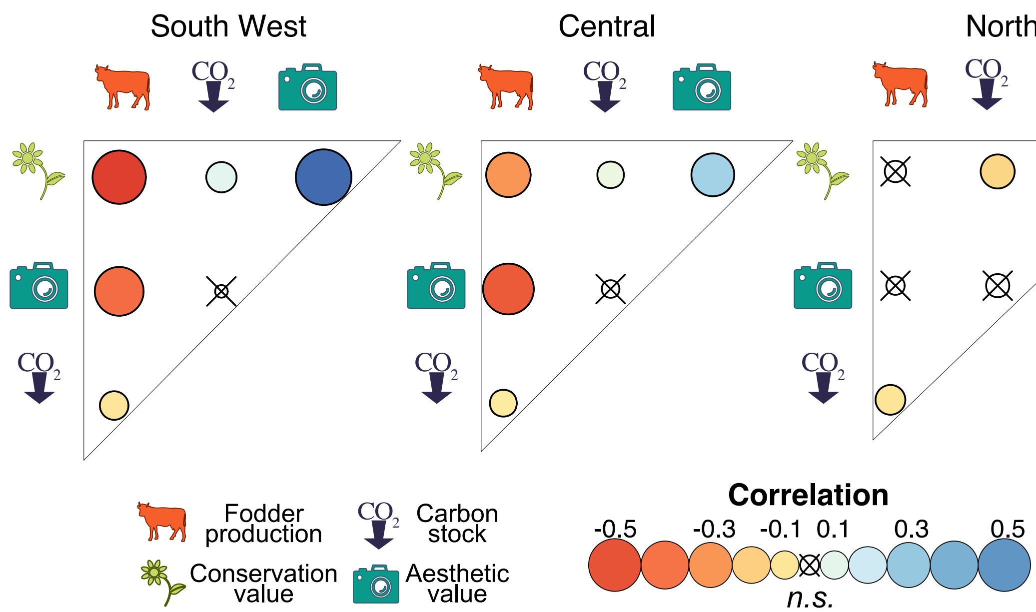
At the single-site scale, the optimal land-use intensity for individual services can be easily identified. Across all regions, fodder production consistently increases with land-use intensity while conservation and aesthetic values respond negatively to land-use intensity (Figure 2). Carbon stocks do not vary with land-use intensity. The trade-offs and synergies observed at the landscape scale (Figure 3) are consistent with these site-scale results (Figure 2). Conservation value is synergic with aesthetic value (Pearson's $r = 0.35$ for all regions, $P < 0.001$) but both

**Figure 2 Relationship between ecosystem service supply and land-use intensity across the study regions.** a. Variation of ecosystem services supplies with land-use intensity. Values shown are calculated at the plot level as the average of their component indicators (see Table 1 and supplementary figures). Values were scaled between 0 and 1. Different letters indicate differences significant at 5% (ANOVA and pairwise comparisons). b. Characterisation of land-use intensity based on mowing, grazing and fertilisation levels in the different regions. The size of the symbols is proportional to the corresponding intensity.



display a trade-off with fodder production (respectively r = - 0.28 and r = -0.32, P < 0.001). Carbon stocks do not show any consistent relationship with the other services.

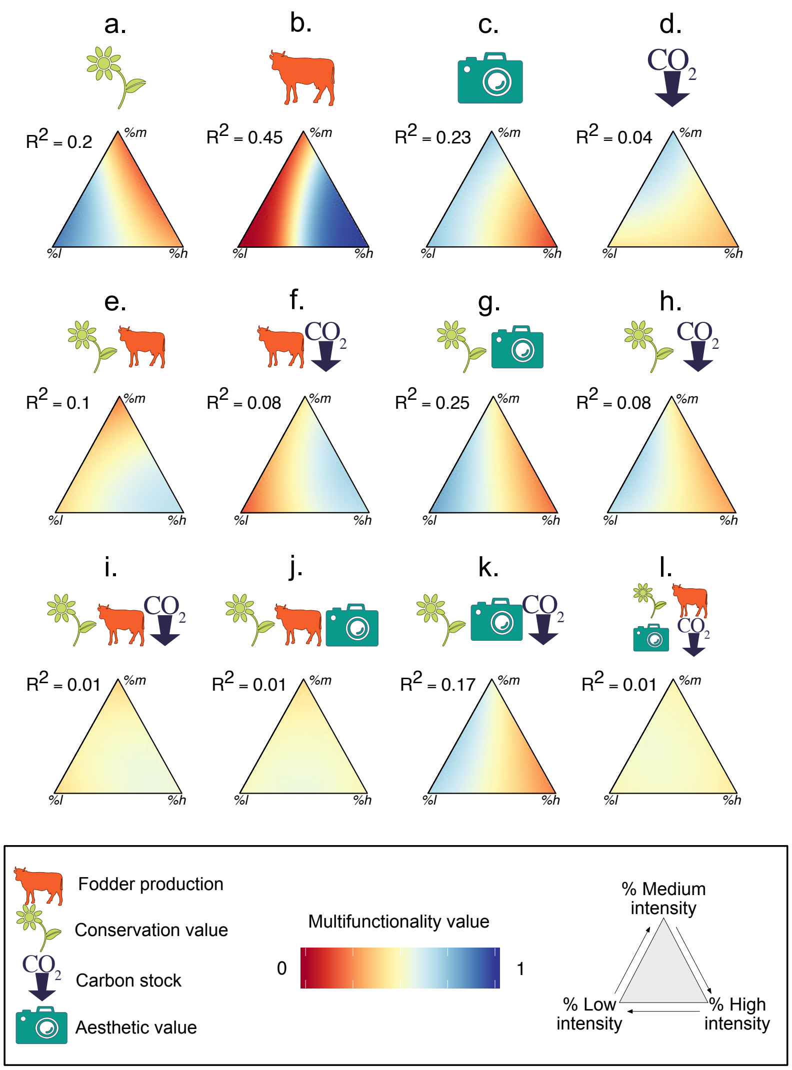
**Figure 3 Trade-offs between landscape-scale ecosystem services.** The colour and size of the circles denote the strength of the correlation between pairs of variables, within each region. Crosses indicate no significant correlations at 5% (Holm correction for multiple testing).




*Optimal land-use allocation at the landscape scale*

At the landscape scale, effective landscape strategies can be identified where only a few services are desired, but optimisation becomes increasingly difficult as more services are considered. This makes hard choices inevitable when there are trade-offs between multiple ecosystem services. The optimal land-use allocation pattern also depends strongly on whether achieving moderate levels of all services, or high levels of a few, is the priority. Given this sensitivity we developed an online tool to allow users to investigate the best management strategy for a given set of ecosystem service demands (https://neyret.shinyapps.io/landscape_composition_for_multifunctionality/). In the text below we highlight a few of the possible combinations of this parameter space, and demonstrate the sensitivity of multifunctionality to multiple factors. We illustrate our results using data collected from three of the main stakeholder groups of the three study regions; farmers, conservationists and the tourism sector. This social survey showed that all groups demanded at least some of each service (Figure S 1), but that conservationists prioritized biodiversity conservation; farmers, food production; and the tourism sector both landscape beauty and biodiversity conservation.

**Figure 4 Dependency of multifunctionality on the services demanded and landscape composition.** Landscape composition is presented in proportions of low, medium and high-intensity sites, for selected combinations of ecosystem services in the Central region of the Exploratories. For single ecosystem services (top row), the value presented corresponds to the probability of the given service being above the median. For combinations of multiple services (middle and bottom rows), multifunctionality is the proportion of services above the median. Blue indicates higher multifunctionality values, orange lower. The full set of service combinations in all regions can be found in Figure S 3. $R^2$ values were calculated from generalised linear models (see Methods).

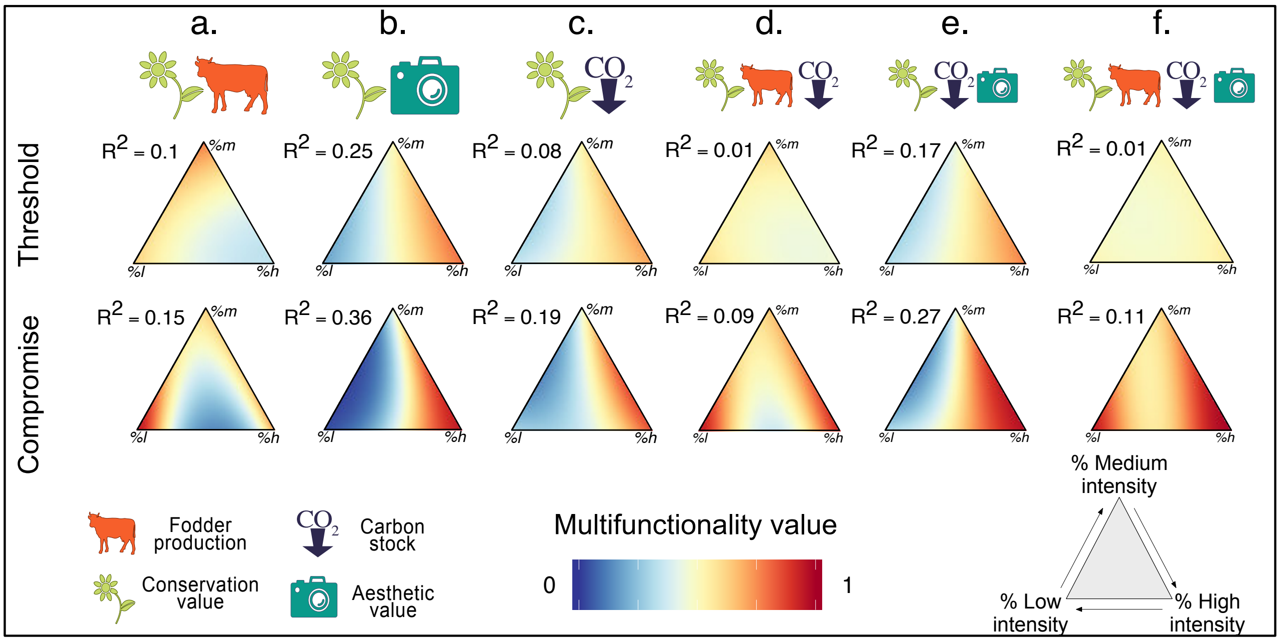



123 In the first set of examples, we present the 'threshold' scenario in which land governors choose to manage the landscape to provide high levels of some services, potentially to the exclusion of some others. This corresponds to a multifunctionality metric calculated as the number of services over the median, and is broadly equivalent to a metric widely used in multifunctionality studies (e.g. Soliveres et al., 2016; van der Plas et al. 2016). Here we find that for individual services, the optimal landscape composition is predictable and consistent with the site-level results, i.e. that the highest service values are found in homogeneous landscapes composed of sites with

land-use intensities favouring that particular service (Figure 4. a-d). Thus, the optimal landscapes for individual stakeholder groups are composed exclusively of one intensity, either high (e.g. fodder production for farmers, Figure 4b), or low (e.g. conservation for conservationists, Figure 4a). The optimal landscape composition when two ecosystem services are considered depends on whether these services have consistent or contrasting responses to land-use intensity. When the two services are synergic, they behave as a single service and optimal landscape composition is found at the common optimum of the two services. For example, a clear optimum can be found for conservation and aesthetic value (Figure 4g), both of which are prioritized by the tourism sector. In contrast, if the two services respond contrastingly to land-use intensity, then whether an optimum could be found depends on the form and strength of their relationship with land-use intensity. For example, a common objective of landscape management is to meet the demands of both the agricultural and conservation sectors, by combining food production with biodiversity conservation (Phalan et al. 2011). As there is a strong trade-off between these services (Figure 3), only a partial optimum with high levels of both services can be found (Figure 4e), with the landscape composition delivering this depending on regional differences in the response of services to land-use intensity (see Figure S 3 for details), and the relative responsiveness of the services considered to land use intensity. For three or four services the identification of an optimal land-use strategy becomes even more challenging. In these cases, multifunctionality varies very little across the full range of landscape composition (maximum $R^2$ for Figure 4i-l: 17%, and often < 10%), with relatively uniform multifunctionality values of about 50%, regardless of the landscape composition.

Next, we explored the ‘compromise’ scenario in which land governors choose to balance the demands of different stakeholder groups, ensuring moderate, but not necessarily high, level of all services. We represented this by creating a multifunctionality metric that is equal to 1 if all services are above the 25$^{th}$ percentile, and 0 otherwise. The use of such a metric strongly affects the outcome of the land management strategy, in comparison to the first ‘threshold’ multifunctionality scenario (for selected service combinations see Figure 5). While the two scenarios give similar results when services were synergic (e.g. Figure 5b), it is easier to identify successful land-use strategies when the considered services display a trade-off for the ‘compromise’ scenario, and especially when there are only two services (Figure 5a). In this case ‘compromise’ multifunctionality is highest in landscapes composed of both high- and low-intensity sites, and with few medium-intensity sites, i.e. - broadly similar to a land-sparing strategy. When multiple services are considered (Figure 5d-f), variation in multifunctionality across the different strategies is also higher in the ‘compromise’ than the ‘threshold’ scenario, ranging from 0 to 0.6.

**Figure 5 Dependency of multifunctionality on stakeholder demand patterns, as represented by 'threshold' and 'compromise' metrics**. Values also depend on landscape composition (in proportions of low, medium and high-intensity sites). In the threshold scenario, multifunctionality is calculated as the number of services above the median (top row, repeated from Fig. 4). In the compromise scenario, multifunctionality equals 1 if all services are above the 25th quantile, and 0 otherwise (bottom row). $R^2$ values were calculated from generalised linear models (see Methods). Only data from the Central region and certain service combinations are presented, other service combinations and regions can be found in Figure S 3 and Figure S 8.

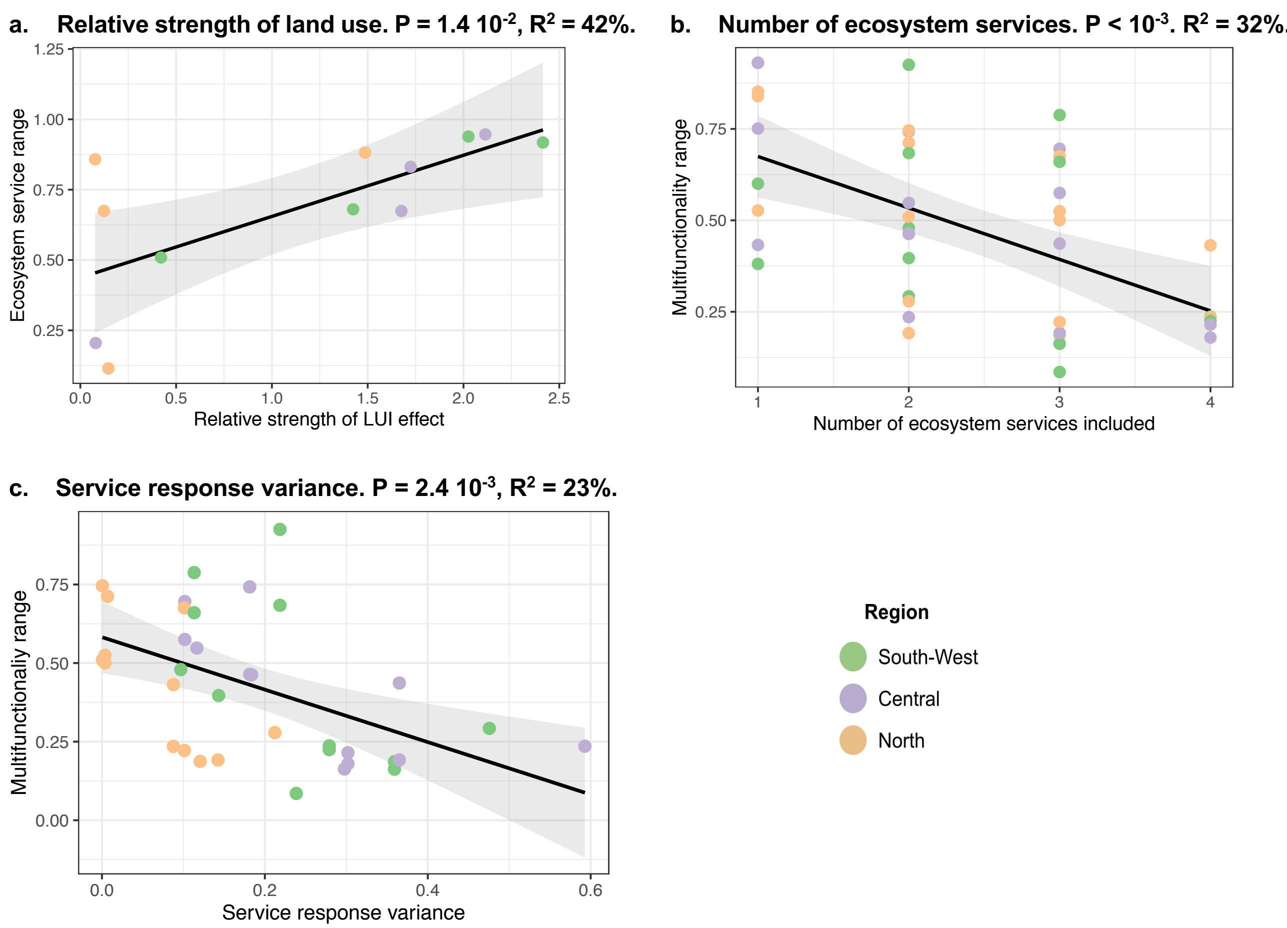



165 *Identifying the drivers of multifunctionality*

To explore why optimal landscape strategies cannot always be identified when multiple services are demanded, we generated and tested several hypotheses. The first was that some services are primarily driven by environmental drivers (e.g. climate and underlying geology) and so respond weakly to landscape land-use composition compared to those predominantly driven by land-use intensity. Second, we predicted that the response of multifunctionality to landscape composition should be weaker when services respond contrastingly to land use intensity (high variance in the response to land use intensity). In such cases, increasing the number of services will weaken the response of multifunctionality to landscape composition by aggregating increasing amounts of variation. We tested whether these two factors determined the responsiveness of multifunctionality to landscape composition, defined as the range of predicted multifunctionality values in the models (appearing as the strength of colour gradient on Figure 4 and Figure 5, see Methods for details). The first hypothesis was supported; in the threshold scenario, multifunctionality range increased if land-use intensity had a relatively large effect on the services included compared to other environmental drivers (Figure 6a, P = 1.4 $10^{-2}$, $R^2$ =

42%). The second hypothesis was also supported; multifunctionality range decreased with increases in the numbers of services included in the analysis (Figure 6b, $P < 0.001$. $R^2 = 32\%$) and the variance in their response to land-use intensity (Figure 6c, $P = 2.4\ 10^{-3}$, $R^2 = 23\%$). In the compromise scenario, multifunctionality was not affected by these factors (Table S 5), due to relatively high multifunctionality ranges for all service combinations, as detailed above.

**Figure 6 Factors explaining the sensitivity of multifunctionality to landscape composition.** Figures show the responsiveness of multifunctionality (range between 5% and 95% quantiles of the predicted values) to landscape composition depending on (a) the strength of each individual ecosystem service's response to landscape composition relative to the effects of land use and environmental covariates (b) the number of ecosystem services included in its calculation (all possible combinations, of 1 to 4 services) and (c) the service-response variance among the included ecosystem services (all possible combinations). Each dot represents individual services (a), or one combination of services (b, c), per region. The lines show the prediction of a linear model, with multifunctionality range as the response and the considered factor as the explanatory variable.



*Additional analyses*

To assess whether the inability to find a clear optimum was due to our simplification of land-use intensity into three categories, we also investigated the response of multifunctionality to the mean and coefficient of variation of land-use intensity at the landscape level. The results of this analysis were largely consistent with the results of the 'three levels' analysis in that unless the services were synergic, no optimum could be found when several services were demanded (Figure S 18). The model fits were also equivalent to those from the aforementioned analyses but are more difficult to translate into simple management recommendations.

In addition to the main cases presented here we identified several other sensitivities including additional metrics for multifunctionality calculation at the landscape level, the number of sites included in each landscape, and the use of raw data instead of that corrected for environmental variation. We encourage readers to explore these sensitivities in the app, although the corresponding figures and specificities are also presented extensively in the supplementary information (Figure S 4 to Figure S 18).

# Discussion

While the land-sharing or -sparing debate has aided our understanding of the trade-offs between commodity production and conservation (Phalan, 2018) we show that neither of these simple strategies can provide high multifunctionality in grassland landscapes, if high levels of multiple ecosystem services are desired. We predict that this difficulty in achieving high multifunctionality is general to many ecosystems and landscapes, as the presence of other drivers and trade-offs or imperfect correlations between services are commonplace (Bennett et al., 2009; Bradford & D'Amato, 2012). Various studies have advocated for the consideration of more complex strategies for balancing commodity production with conservation (Bennett, 2017; Butsic & Kuemmerle, 2015; Fischer et al., 2014; Kamp et al., 2015; Phalan et al., 2011, Simons and Weisser 2017). By employing a rigorous approach based on direct, in-field measurements of ecosystem service indicators, we further show that considering not only trade-offs and synergies between ecosystem services, but also information describing the ecosystem service demand of stakeholders, helps identify land management options that have greater precision and relevance to land users. The approach presented also allows the potential causes of land-use conflicts to be identified, as it can assess whether low multifunctionality is caused by trade-offs in the supply of ecosystem services, or unrealistic and incompatible demands on the ecosystem by stakeholders.

In our study system, ecosystem services showed contrasting responses to land-use intensity, such as the commonly observed trade-off between production and biodiversity or cultural services (Allan et al., 2015; Bradford & D'Amato, 2012; Cordingley et al., 2016; Lavorel et al., 2011; Raudsepp-Hearne et al., 2010). Understanding contrasting responses of ecosystem services to land management is fundamental to identifying landscape-level strategies. Here, we show that strong management-driven trade-offs preclude multifunctionality when high levels of services are required. As a result, even complex landscape strategies can fail to deliver high levels of multiple ecosystem services (Allan et al., 2015) and landscape management is likely to require "hard choices" (Cordingley et al., 2016; Slade et al., 2017) regarding which services to prioritise, and which are secondary. At the same time, we show that it is possible to provide limited levels of multiple services by combining sites at low and high intensities, a strategy broadly similar to land-sparing. In this respect, our results show that the optimal strategy depends heavily on the priorities of landscape managers. While different stakeholders are likely to favour different sets of services, landscape-level governors are faced with a difficult choice: create a landscape with a few services at high value, which will create clear winners and losers among stakeholder groups, or one that minimises the trade-offs among services so that all are present at moderate levels, meaning that all stakeholder groups must accept sub-optimal levels of ecosystem services.

While advancing on previous studies by incorporating multiple services, we acknowledge that our approach to identifying optimal landscape strategies is simple and ignores much of the complexity found in natural systems. Firstly, ecosystem services respond to multiple drivers, and these can be either anthropogenic (e.g. land-use change, overexploitation, Carpenter et al., (2009)) or environmental (e.g. soil (Adhikari & Hartemink, 2016), climate, or elevation (Lavorel et al., 2011)). Failing to account for these drivers can obscure the relationship between land-use composition and multifunctionality. Environmental drivers will differ in their effect on different services, and so can modify their trade-offs (Clec'h et al., 2019). Therefore, the development of strategies to achieve landscape multifunctionality also needs to be informed by regional knowledge (Anderson et al., 2009; Butsic & Kuemmerle, 2015; Clec'h et al., 2019). For instance, in our analysis the North region responded very differently to the other two regions. This was due to regional specificities, such as its uniformly low plant diversity and the association of low-intensity sites with organic soils, which shifted the optimal landscape compositions to different regions of the triangular space compared to the other regions (Figure S 3).

In addition to local drivers, the delivery of many ecosystem services depends on the movement of matter or organisms among landscape units (Mitchell et al., 2014). For instance, pollination, water quality, or pest and disease control are affected by landscape complexity, fragmentation and surrounding land uses (Duarte et al., 2018). Accordingly, we advocate the incorporation of spatial interactions between landscape units (Lindborg et al., 2017) into future models, elements which may modify and expand upon the conclusions presented here.

Our system consists of only one land-use type and does not include unmanaged land, making it only broadly comparable to the land-sparing and -sharing strategies. However, we argue that the methodology presented here could be extended to many different land-use and management regimes, provided that appropriate data on services and drivers is available. Steps must also be taken to ensure that insights from such studies are in a format that can be communicated effectively to land managers. For instance, we argue that the proposed methodology - proportions of land in a number of land-use categories - is more easily transferable than indices of land-use intensity heterogeneity. Strategies for knowledge transfer also need to be developed. We suggest that apps like the one presented here provide a useful demonstration tool for communicating land-use options to land managers and policy makers, as they could be used to explore options, understand the causes of conflicts and trigger discussions, thus helping to support decision-making among different groups of stakeholders. However, full application of findings such as those presented here also requires the existence of structures that aim to identify landscape strategies and operationalise them at a community level, such as the 'landscape approach' (DeFries & Rosenzweig, 2010; Sayer et al., 2013). This aims to balance competing land-use demands to promote environmental conservation and human well-being based on a participatory approach (e.g. the African Forest Landscape Restoration Initiative). Government and corporate policies can also implement such strategies, e.g. via agri-environment schemes that may guide the allocation of different land-use types or land-use intensities to different parts of the landscape (Whittingham, 2011). We suggest that demonstrating of management options via apps such as that presented here, can foster understanding and aid decision making in both of these settings.

Overall, this study shows that landscape strategies are highly sensitive to the identity of the services desired and the type of multifunctionality demanded by stakeholders, making participatory approaches to the development of land management strategies essential. When high levels of all services are required, we show that optimising landscape composition is usually possible for two services. However, when there are strong trade-offs among services or significant effects of other environmental drivers, successful options become increasingly limited unless stakeholders are willing to accept moderate service levels, which can be delivered by strategies akin to land sparing. Across the world, landscapes are increasingly required to provide a wide range of services. This study stresses the need for both theoretical studies and applied social and ecological research into which services are required, at what scale, and how they are affected by environmental drivers. Such knowledge is essential if we are to identify land-use strategies that minimise conflict between stakeholders, and promote the sustainable use of all ecosystem services.

# Material and methods

## Study design

We used data from 150 grassland plots (hereafter sites) studied within the large-scale and long-term Biodiversity Exploratories project in Germany (https://www.biodiversity-exploratories.de/). The sites were located in three regions including the UNESCO Biosphere Area Schwäbische Alb (South-West region), in and around the National Park Hainich (Central region; both are hilly regions with calcareous bedrock), and the UNESCO Biosphere Reserve Schorfheide-Chorin (North of Germany: flat, with a mixture of sandy and organic soils, see Fischer et al. (2010) for details). Sites measured 50 x 50m and were selected to be representative of the whole field they were in, spanning the full range of land-use intensity within the region, while minimising variation in potentially confounding environmental factors.

## Land-use intensity

Data on site management was collected annually from site owners using a questionnaire. We quantified grazing intensity as the number of livestock units × the number of days of grazing (cattle younger than 1 year corresponded to 0.3 livestock units (LU), cattle 2 years to 0.6 LU, cattle older than 2 years to 1 LU, sheep and goat younger than 1 year to 0.05 LU, sheep and goat older than 1 year to 0.1 LU, horse younger than 3 years to 0.7 LU, and horse older than 3 years to 1.1 LU; Fischer et al. 2010), fertilisation intensity as the amount of nitrogen addition excluding on-site animal droppings during grazing events (kg N $ha^{-1}y^{-1}$), and mowing frequency as the annual number of mowing events. For each site these three land-use intensity (LUI) components were standardised, square-root transformed, summed, and then averaged between 2007 and 2012 to obtain an overall LUI value (Blüthgen et al., 2012). We then classified all sites as low-, medium- or high-intensity based on whether their LUI index belonged to the lower, middle or top third (0-33%, 33-66%, 66-100% quantiles) of all LUI indices within the considered region. Confidence intervals for grazing and fertilization intensities for each LUI class in the three regions are presented in Table 2.

**Table 2 Description of the variations of land-use intensity components.** Confidence intervals for fertilisation and grazing intensities in each region, for each land-use intensity (LUI) class. 95% confidence intervals were calculated based on fertilisation and grazing values of individual plots on the period 2007-2012

| | LUI class | South-West | Central | North |
|---|---|---|---|---|
| LUI index | Low | 1 (0.9-1.2) | 1 (0.9-1.1) | 1.2 (1.1-1.3) |
| | Medium | 1.7 (1.6-1.7) | 1.7 (1.6-1.8) | 1.5 (1.4-1.5) |
| | High | 2.2 (2-2.4) | 2.2 (2.1-2.4) | 2.3 (2.1-2.5) |
| Grazing intensity (Livestock units. days.ha$^{-1}$) | Low | 82.2 (49.2-115.3) | 86.5 (64.4-108.6) | 103.4 (52.2-154.6) |
| | Medium | 97.6 (34.5-160.7) | 102.5 (48.3-156.6) | 239.5 (140.9-338.1) |
| | High | 156.7 (24.6-288.8) | 160.7 (39.6-281.8) | 215.9 (80.4-351.4) |
| Mowing intensity (Cut.yr$^{-1}$) | Low | 0.5 (0.1-0.8) | 0.3 (0-0.6) | 0.8 (0.6-1.1) |
| | Medium | 1.4 (0.9-1.8) | 1.2 (0.9-1.5) | 0.8 (0.4-1.2) |
| | High | 1.8 (1.3-2.3) | 1.6 (1.2-2) | 1.2 (0.8-1.7) |
| Fertilisation (kg.N.ha$^{-1}$) | Low | 1.1 (-0.6-2.8) | 1.4 (-1.6-4.4) | 0.4 (-0.4-1.2) |
| | Medium | 38 (23.4-52.7) | 34.9 (18-51.8) | 0.6 (-0.7-2) |
| | High | 95 (67.4-122.6) | 91.1 (65.4-116.9) | 42.8 (25.8-59.7) |

## Ecosystem services demand

A preliminary social survey was conducted among representatives of the main stakeholder groups within each region to identify the most demanded ecosystem services. The participants, one representative per stakeholder group per region, were asked to rank their demand for all possible ecosystem services provided in their region at the landscape level between 1 and 5 (1 being not important and 5 very important). The rankings were then re-normalised by the total number of points attributed by each individual. Of the services identified, we then selected the four most demanded services that are provided by grasslands: biodiversity conservation, fodder production, aesthetic value and climate change mitigation (Figure S 1).

## Ecosystem services

We estimated these services from several indicators (Table 1), measured in each site of the regions. Before estimating the landscape-level services, we imputed missing values for individual indicators using predictive mean matching on the dataset comprising all services (98 out of 1200 values, R mice package). The missing values were mostly found for flower cover, and some for butterfly abundance, but they were equally distributed among regions and land-use intensities. In all following analyses, we used environment-corrected indicators. These were quantified as the residuals from linear models, conducted separately within each region. The four ecosystem service indicators were the response variable and predictors were: pH, soil depth, sand and clay content, topographic wetness index, mean annual temperature and annual rainfall (see Allan et al. (2015) and Hijmans et al. (2005) for details on these measurements) and a topographic wetness index (see supplementary methods). To account for a site's surroundings, we also used the proportion of grassland in a 1km radius as a predictor, as surrounding grassland habitat may act as a source of colonization for local biodiversity (e.g. Henckel et al., 2015; Le Provost et al., 2017; Tscharntke et al., 2012). It was obtained from land-use covers obtained in 2008 data that were mapped QGIS v 3.6 and classified into five broad categories: croplands, grasslands, forests, water bodies, roads and urban areas.

The 'biodiversity conservation' service at the site-level was based on total plant species richness as plant alpha-diversity and the sum of the ground cover of regional red list plant species. Plant species richness has been shown to be a good proxy for diversity at multiple trophic levels at these sites (correlation of 0.67 and 0.68 between the whole ecosystem multidiversity index (Allan et al 2014) and the richness of asterids and rosids respectively, for instance (Manning et al., 2015)). We chose not to include other taxa to prevent co-linearity with the other service measures (see below). Red list plant species included species classified in the following threat categories:

1 (threatened with extinction); 2 (critically endangered); 3 (endangered), by Breunig & Demuth (1999), Korsch & Westhus (2001) and Ristow et al. (2006) (Table S 1). The values of these two indicators were re-calculated at the landscape level (i.e. gamma diversity and the sum of red list species cover in all sites) and then scaled and averaged to calculate the landscape-level service.

The fodder production service was calculated as total fodder protein production, a common agronomical indicator (Lee, 2018) that we calculated based on grassland aboveground biomass production and shoot protein content. Between mid-May and mid-June each year, aboveground biomass was harvested by clipping the vegetation 2 - 3 cm above ground in four randomly placed quadrats of 0.5 m × 0.5 m in each subplot. The plant biomass was dried at 80°C for 48 hours, weighed and summed over the four quadrats. Biomass was then averaged between 2008 and 2012. In order to convert this one-time biomass measurements into estimates of annual field productivity, we used the information on the number of cuts and the number of livestock units in a site to estimate the total biomass production used by farming activities, i.e. converted into fodder or consumed directly by livestock. Details of this estimation process can be found in Simons & Weisser (2017). We then multiplied this productivity by plant shoot protein levels, a common indicator of forage quality (Lee, 2018). Total nitrogen concentrations in ground samples of aboveground biomass were determined using an elemental auto-analyser (NA1500, CarloErba, Milan, Italy), and multiplied by 6.25 to obtain protein content (Lee, 2018). The landscape-scale protein production was then calculated as the sum of the production of all individual sites in the landscape.

Climate change mitigation was quantified as soil organic carbon stocks in the top 10 cm, as deeper stocks are unlikely to be affected strongly by management actions. We sampled composite samples for each plot, prepared by mixing 14 mineral surface soil samples per plot. Soil samples were taken along two 18 m transects in each plot using a split tube auger, 40 cm long and 5 cm wide (Eijkelkamp, Giesbeek, The Netherlands). Composite samples were weighed, homogenized, air-dried and sieved (<2 mm). We then measured total carbon (TC) contents by dry combustion in a CN analyser "Vario Max" (Elementar Analysensysteme GmbH, Hanau, Germany) on ground subsamples. We determined inorganic carbon (IC) contents after combustion of organic carbon in a muffle furnace (450°C for 16 h). We then calculated the soil organic carbon (SOC) content as the difference between TC and IC, and the SOC concentration based on the weight of the dry fine-earth (105°C) and its volume. SOC concentration was then multiplied by soil bulk density to obtain plot-level carbon stock values. The landscape-scale soil carbon stock was calculated as the sum of the soil carbon stock of all individual sites.

The aesthetic value measure integrated flower cover, the number of bird families and abundance of butterflies. The choice of these indicators was led by studies showing people's preference for bird richness over abundance (Cox & Gaston, 2015), including song diversity (Hedblom et al., 2014); and for flower-rich landscapes (Graves et al., 2017). Flowering units were counted between May and September 2009 for all flowering plant species (excluding grasses and sedges) on transects along the four edges of each plot, in a total area of 600m$^2$. For abundant species, the number of flowering units was extrapolated to the whole plot from a smaller area of 112 m$^2$. The total flower cover was calculated at the plot scale as the sum of the individual flower cover of all plant species (see Binkenstein et al. (2013) for details). Butterfly and day-active moths (hereafter termed as Lepidoptera) abundance was measured in 2008 and averaged among sites within each landscape (Börschig et al., 2013). We conducted surveys of Lepidoptera from early May to mid-August. We sampled Lepidoptera during 3 surveys, each along one fixed 300m transect of 30min in each site. Each transect was divided in 50m sections of 5min intervals and we recorded all Lepidoptera within a 5 m corridor. Birds were surveyed by standardized audio-visual point-counts and all birds exhibiting territorial displays (singing and calling) were recorded. We used fixed-radius point counts and recorded all males of each bird species during a five-minute interval per plot. Each plot was visited five times between 15 March and 15 June each year. The data was then aggregated by family. Landscape-scale bird richness was calculated as the total number of bird families found in the landscape (i.e. in at least one site and one year) between 2009 and 2012. These three indicators were then scaled and averaged to estimate landscape-scale aesthetic value. Richness and abundance were usually highly correlated for the three groups (correlation of 0.75 for butterflies, 0.72 for birds, and 0.52 for plants), and the number of families for birds was highly correlated (0.96) to species richness; thus other selections of indicators would have led to similar results (Figure S 2).

## Plot-level analyses

We first analysed the relationship between all plot-level service indicators and land-use intensity class. Within each region, we scaled the services between 0 and 1 and fitted ANOVAs with the land-use class as an explanatory variable; followed by a pairwise mean comparison.

## Landscape simulations

We conducted the simulations separately within each region, as each displayed different relationships between land use and ecosystem services (Figure 2).

We simulated artificial landscapes within each region. Each artificial landscape was composed of ten sites to avoid the high similarity among landscapes composed of more sites. Across the triangular space, and for landscapes made up of 10 plots, there are 66 possible landscape compositions that differ in their proportions of low, medium and high intensity sites; ranging from 100% low intensity to 100% medium or high intensity with all possible intermediates. For each of these compositions, we generated 15 unique artificial landscapes by randomly drawing sites from the regional pool, resulting in 15 x 66 = 990 landscapes. In each simulated landscape, we then calculated landscape-scale ecosystem service indicators, as described above.

Finally, we calculated landscape-scale ecosystem service multifunctionality as described below. We fitted binomial linear models with multifunctionality as a response and with a second-degree polynom of the proportions of low and high land-use intensity as explanatory variables.

## Landscape-level ecosystem multifunctionality

Different multifunctionality scenarios were investigated, corresponding to all the possible combinations of the four main ecosystem services (i.e. single benefits, all the pairs and triplets, or including all four benefits). In each combination, we calculated two measures of multifunctionality. To represent a scenario where high levels of some services are required, multifunctionality was assessed by scoring each final benefit as 1 if it passed a given threshold, equal to the median of the values of the service obtained on all landscapes within the considered region. Multifunctionality was then calculated as the number of services reaching this threshold, divided by the number of services included in the analysis, so that it ranged between 0 and 1.

We also considered an alternative scenario, in which land governors compromise between the needs of multiple stakeholders by maintaining at least intermediate levels of all ecosystem services. Here, we scored the multifunctionality as 1 if all the services were above a 25% threshold (i.e. above the 25% quantile of the service distribution in all landscapes within the region), and 0 otherwise.

## Dependence of multifunctionality range to the number of services included and to environmental covariates

The response of multifunctionality to landscape composition became increasingly complex as more services were added (see Results). Therefore, we performed additional analyses to investigate which factors affected the responsiveness of multifunctionality to landscape composition, which corresponds strength of the colour gradient in the triangle plots presented

(Figure 4 and Figure 5). Multifunctionality responsiveness was calculated as the range (2.5% to 97.5% quantiles) of the fitted values of the models described above (binomial GLMs with the proportion of high and low intensity sites as explanatory variables) over all the possible landscape compositions. Thus, while overall the range of multifunctionality was always 1 (existence of plots with none or all of the services above the threshold), the range of fitted values depended on the fit of the model, i.e. whether the value of multifunctionality depended primarily on landscape composition.

To investigate the relationship between multifunctionality responsiveness and the number of ecosystem services included in its calculation, we regressed it upon the number of ecosystem services included in the landscape-scale assessment (ranging from 1 for individual final benefits to 4 for the multifunctionality measure with all benefits).

Multifunctionality was also hypothesised to depend on contrasting responses to land-use intensity of the different services included in the assessment. In a second analysis we estimated the slope coefficients of the linear regressions between each service and land-use intensity, and calculated the ‘service response variance’ of the considered services as the variance of their slope coefficients (see van der Plas et al., 2019 for details). We then fitted a linear model of multifunctionality range against the service response variance. Finally, we examined the linear relationship between multifunctionality and the relative strength of LUI effect compared to other environmental covariates. For each single ecosystem service and each region, we quantified the relative strength of the effect of land-use intensity ($RS_{LUI}$) as:

$$RS_{LUI} = \left[\frac{corr(ES, LUI)}{max_j(\,corr(ES, EC_j)}\right]$$

Where *corr* is the correlation, *ES* the ecosystem service, *LUI* the value of land-use intensity, and $EC_j$ the environmental covariates (see below). These three models were fitted for all regions together.

## Sensitivity analyses

We complemented the main analyses by extensive sensitivity analyses, which are detailed in the supplementary material of this article as well as in the online app.

We ran the same analyses using indicators that were not corrected for environmental variables. Other sensitivity analyses included changing the number and identity of plots selected to build the landscapes: using only sites with the lowest 20%, highest 20% and medium 20% land-use intensity (i.e. removing sites that are intermediate between two intensity classes) and by including 7, or 13, sites per landscape. We calculated multifunctionality using other threshold values. For the threshold multifunctionality metric, we also run the analysis by setting the threshold at the 40$^{th}$ or 60$^{th}$ percentile, and at 60% of the maximum. For the compromise metric, we investigated thresholds of 20% and 30% in addition to the 25% threshold.

Finally, multifunctionality at the landscape level was also considered as the maximum level observed in the landscape, rather than the sum of all the plots present in the landscape.

# Author contributions

M.N., P.M. and M.F. conceived the study, M.N. and P.M designed and performed the analyses, M.N. and P.M. wrote the manuscript with significant contributions from E.A., N.H., V.H.K., T.K., J.K., G.L.P, S.P., N.S, N.K.S. and F.v.d.P.

Data was contributed by M.F., G.L.P, E.A., J.B., C. B., N. H., K. J., V.H.K., T. K., S. P., D. P., M. S., D. S., N.S., I. S., M. S., N. K. S., M.T. and C. W. Authorship order was determined as follows: (1) core authors; (2) other major contributors (alphabetical); (3) other authors contributing data (alphabetical).

# Acknowledgements

We thank the managers of the three Exploratories Konstans Wells, Swen Renner, Kirsten Reichel-Jung, Sonja Gockel, Kerstin Wiesner, Katrin Lorenzen, Andreas Hemp, Martin Gorke and all former managers for their work in maintaining the plot and project infrastructure; Simone Pfeiffer, Maren Gleisberg and Christiane Fischer for giving support through the central office, Jens Nieschulze and Michael Owonibi for managing the central data base, and Markus Fischer, 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.

The work has been funded by the DFG Priority Program 1374 "Infrastructure-Biodiversity-Exploratories”. Field work permits were issued by the responsible state environmental offices of Baden-Württemberg, Thüringen, and Brandenburg.

# References

Adhikari, K., & Hartemink, A. E. (2016). Linking soils to ecosystem services—A global review. *Geoderma*, *262*, 101–111.

Allan, E., Manning, P., Alt, F., Binkenstein, J., Blaser, S., Blüthgen, N., Böhm, S., Grassein, F., Hölzel, N., Klaus, V. H., Kleinebecker, T., Morris, E. K., Oelmann, Y., Prati, D., Renner, S. C., Rillig, M. C., Schaefer, M., Schloter, M., Schmitt, B., … Fischer, M. (2015). Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. *Ecology Letters*, *18*(8), 834–843. https://doi.org/10.1111/ele.12469

Anderson, B. J., Armsworth, P. R., Eigenbrod, F., Thomas, C. D., Gillings, S., Heinemeyer, A., Roy, D. B., & Gaston, K. J. (2009). Spatial covariance between biodiversity and other ecosystem service priorities. *Journal of Applied Ecology*, *46*(4), 888–896.

Bengtsson, J., Bullock, J. M., Egoh, B., Everson, C., Everson, T., O’Connor, T., O’Farrell, P. J., Smith, H. G., & Lindborg, R. (2019). Grasslands—More important for ecosystem services than you might think.

*Ecosphere*, *10*(2), e02582.

Bennett, E. M. (2017). Changing the agriculture and environment conversation. *Nature Ecology and Evolution*, *1*(1), 1–2.

Bennett, E. M., Peterson, G. D., & Gordon, L. J. (2009). Understanding relationships among multiple ecosystem services. *Ecology Letters*, *12*(12), 1394–1404.

Binkenstein, J., Renoult, J. P., & Schaefer, H. M. (2013). Increasing land-use intensity decreases floral colour diversity of plant communities in temperate grasslands. *Oecologia*, *173*(2), 461–471. https://doi.org/10.1007/s00442-013-2627-6

Blüthgen, N., Dormann, C. F., Prati, D., Klaus, V. H., Kleinebecker, T., Hölzel, N., Alt, F., Boch, S., Gockel, S., & Hemp, A. (2012). A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. *Basic and Applied Ecology*, *13*(3), 207–220.

Börschig, C., Klein, A.-M., von Wehrden, H., & Krauss, J. (2013). Traits of butterfly communities change from specialist to generalist characteristics with increasing land-use intensity. *Basic and Applied Ecology*, 547–554. https://doi.org/10.1016/j.baae.2013.09.002

Bradford, J. B., & D'Amato, A. W. (2012). Recognizing trade-offs in multi-objective land management. *Frontiers in Ecology and the Environment*, *10*(4), 210–216.

Breunig, T., & Demuth, S. (1999). *Rote Liste der Farn—Und Samenpflanzen Baden—Württembergs* (Landesanstalt für Umweltschutz Baden-Württemberg. Naturschutz-Praxis, No. 2; Artenschutz, p. 247).

Butsic, V., & Kuemmerle, T. (2015). Using optimization methods to align food production and biodiversity conservation beyond land sharing and land sparing. *Ecological Applications*, *25*(3), 589–595.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., & Wardle, D. A. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59.

Carpenter, S. R., Mooney, H. A., Agard, J., Capistrano, D., DeFries, R. S., Díaz, S., Dietz, T., Duraiappah, A. K., Oteng-Yeboah, A., & Pereira, H. M. (2009). Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. *Proceedings of the National Academy of Sciences*, *106*(5), 1305–1312.

Clec'h, L., Finger, R., Buchmann, N., Gosal, A., 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*.

Cordingley, J. E., Newton, A. C., Rose, R. J., Clarke, R. T., & Bullock, J. M. (2016). Can landscape-scale approaches to conservation management resolve biodiversity–ecosystem service trade-offs? *Journal of Applied Ecology*, 96–105.

Cox, D. T. C., & Gaston, K. J. (2015). Likeability of Garden Birds: Importance of Species Knowledge & Richness in Connecting People to Nature. *PLOS ONE*, *10*(11), e0141505. https://doi.org/10.1371/journal.pone.0141505

DeFries, R., & Rosenzweig, C. (2010). Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences*, *107*(46), 19627–19632. https://doi.org/10.1073/pnas.1011163107

Duarte, G. T., Santos, P. M., Cornelissen, T. G., Ribeiro, M. C., & Paglia, A. P. (2018). The effects of landscape patterns on ecosystem services: Meta-analyses of landscape services. *Landscape Ecology*, *33*(8), 1247–1257. https://doi.org/10.1007/s10980-018-0673-5

Eastburn, D. J., O'Geen, A. T., Tate, K. W., & Roche, L. M. (2017). Multiple ecosystem services in a working landscape. *PLOS ONE*, 10. https://doi.org/10.1371/journal.pone.0166595

Feniuk, C., Balmford, A., & Green, R. E. (2019). Land sparing to make space for species dependent on natural habitats and high nature value farmland. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1909), 20191483. https://doi.org/10.1098/rspb.2019.1483

Fischer, J., Abson, D. J., Butsic, V., Chappel, M. J., Ekroos, J., Hanspach, J., Kuemmerle, T., Smith, H. G., & von Wehrden, H. (2014). Land sparing versus land sharing: Moving forward. *Conservation Letters*, *7*(3), 149–157.

Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F., Hemp, A., Hessenmöller, D., Korte, G., Nieschulze, J., Pfeiffer, S., Prati, D., Renner, S., Schöning, 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*(6), 473–485. https://doi.org/10.1016/j.baae.2010.07.009

Fisher, B., & Turner, R. K. (2008). Ecosystem services: Classification for valuation. *Biological Conservation*, *141*(5), 1167–1169.

Foley, J. A. (2005). Global Consequences of Land Use. *Science*, *309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Goldstein, J. H., Caldarone, G., Duarte, T. K., Ennaanay, D., Hannahs, N., Mendoza, G., Polasky, S., Wolny, S., & Daily, G. C. (2012). Integrating ecosystem-service tradeoffs into land-use decisions. *Proceedings of the National Academy of Sciences*, *109*(19), 7565–7570.

Graves, R. A., Pearson, S. M., & Turner, M. G. (2017). Species richness alone does not predict cultural ecosystem service value. *Proceedings of the National Academy of Sciences*, *114*(14), 3774–3779. https://doi.org/10.1073/pnas.1701370114

Green, R. E. (2005). Farming and the Fate of Wild Nature. *Science*, *307*(5709), 550–555. https://doi.org/10.1126/science.1106049

Hedblom, M., Heyman, E., Antonsson, H., & Gunnarsson, B. (2014). Bird song diversity influences young people's appreciation of urban landscapes. *Urban Forestry & Urban Greening*, *13*(3), 469–474. https://doi.org/10.1016/j.ufug.2014.04.002

Henckel, L., Borger, L., Meiss, H., Gaba, S., & Bretagnolle, V. (2015). Organic fields sustain weed metacommunity dynamics in farmland landscapes. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1808), 20150002–20150002.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

IPBES (2019): Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondízio E.S., H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. Roy Chowdhury, Y. J. Shin, I. J. Visseren-Hamakers, K. J. Willis, and C. N. Zayas (eds.). IPBES secretariat, Bonn, Germany. 56 pages. https://doi.org/10.5281/zenodo.3553579

Kamp, J., Urazaliev, R., Balmford, A., Donald, P. F., Green, R. E., Lamb, A. J., & Phalan, B. (2015). Agricultural development and the conservation of avian biodiversity on the Eurasian steppes: A comparison of land-sparing and land-sharing approaches. *Journal of Applied Ecology*, *52*(6), 1578–1587.

Knocke, T. (2020). Accounting for multiple ecosystem services in a simulation of land-use decisions: Does it reduce tropical deforestation? *TO DO*.

Korsch, H., & Westhus, W. (2001). Rote Liste der Farn- und Blütenpflanzen (Pteridophyta et Spermatophyta) Thüringens. *Naturschutzreport*, *18*, 273–296.

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*(1), 135–147. https://doi.org/10.1111/j.1365-2745.2010.01753.x

Le Provost, G., Gross, N., Börger, L., Deraison, H., Roncoroni, M., & Badenhausser, I. (2017). Trait-matching and mass effect determine the functional response of herbivore communities to land-use intensification. *Functional Ecology*, *31*(8), 1600–1611. https://doi.org/10.1111/1365-2435.12849

Lee, M. A. (2018). A global comparison of the nutritive values of forage plants grown in contrasting environments. *Journal of Plant Research*, *131*(4), 641–654. https://doi.org/10.1007/s10265-018-1024-y

Lindborg, R., Gordon, L. J., Malinga, R., Bengtsson, J., Peterson, G., Bommarco, R., Deutsch, L., Gren, Å., Rundlöf, M., & Smith, H. G. (2017). How spatial scale shapes the generation and management of multiple ecosystem services. *Ecosphere*, *8*(4), e01741. https://doi.org/10.1002/ecs2.1741

Mace, G. M., Norris, K., & Fitter, A. H. (2012). Biodiversity and ecosystem services: A multilayered relationship. *Trends in Ecology & Evolution*, *27*(1), 19–26.

Manning, Pete, Gossner, M. M., Bossdorf, O., Allan, E., Zhang, Y.-Y., Prati, D., Blüthgen, N., Boch, S., Böhm, S., Börschig, C., Hölzel, N., Jung, K., Klaus, V. H., Klein, A. M., Kleinebecker, T., Krauss, J., Lange, M., Müller, J., Pašalić, E., … Fischer, M. (2015). Grassland management intensification weakens the associations among the diversities of multiple plant and animal taxa. *Ecology*, *96*(6), 1492–1501. https://doi.org/10.1890/14-1307.1

Manning, Peter, van der Plas, F., Soliveres, S., Allan, E., Maestre, F. T., Mace, G., Whittingham, M. J., & Fischer, M. (2018). Redefining ecosystem multifunctionality. *Nature Ecology & Evolution*, *2*(3), 427–436. https://doi.org/10.1038/s41559-017-0461-7

Mitchell, M. G. E., Bennett, E. M., & Gonzalez, A. (2014). Forest fragments modulate the provision of multiple ecosystem services. *Journal of Applied Ecology*, *51*(4), 909–918. https://doi.org/10.1111/1365-2664.12241

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., … Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, *520*(7545), 45–50. https://doi.org/10.1038/nature14324

Phalan, B. (2018). What have we learned from the land sparing-sharing model? *Sustainability*, *10*(6), 1760.

Phalan, B., Onial, M., Balmford, A., & Green, R. E. (2011). Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. *Science*, *333*(6047), 1289–1291.

Raudsepp-Hearne, C., Peterson, G. D., & Bennett, E. M. (2010). Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proceedings of the National Academy of Sciences*, *107*(11), 5242–5247.

Ristow, M., Herrmann, A., Illig, H., Klemm, G., Kummer, V., Kläge, H.-C., Machatzi, B., Rätzel, S., Schwarz, R., & Zimmermann, F. (2006). Rote Liste der etablierten Gefäßpflanzen Brandenburgs. *Naturschutz Und Landschaftspflege in Branden- Burg 15*, *15*(4), 12.

Sayer, J., Sunderland, T., Ghazoul, J., Pfund, J.-L., Sheil, D., Meijaard, E., Venter, M., Boedhihartono, A. K., Day, M., Garcia, C., van Oosten, C., & Buck, L. E. (2013). Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses. *Proceedings of the National Academy of Sciences*, *110*(21), 8349–8356. https://doi.org/10.1073/pnas.1210595110

Simons, N. K., & Weisser, W. W. (2017). Agricultural intensification without biodiversity loss is possible in grassland landscapes. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-017-0227-2

Slade, E. M., Kirwan, L., Bell, T., Philipson, C. D., Lewis, O. T., & Roslin, T. (2017). The importance of species identity and interactions for multifunctionality depends on how ecosystem functions are valued. *Ecology*, *98*(10), 2626–2639.

Soliveres, S., van der Plas, F., Manning, P., & Prati, D. (2016). Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature*, *536*. https://doi.org/doi:10.1038/nature19092

Triviño, M., Pohjanmies, T., Mazziotta, A., Juutinen, A., Podkopaev, D., Le Tortorec, E., & Mönkkönen, M. (2017). Optimizing management to enhance multifunctionality in a boreal forest landscape. *Journal of Applied Ecology*, *54*(1), 61–70. https://doi.org/10.1111/1365-2664.12790

Tscharntke, T., Klein, A. M., Kruess, A., Steffan-Dewenter, I., & Thies, C. (2005). Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters*, *8*(8), 857–874. https://doi.org/10.1111/j.1461-0248.2005.00782.x

Tscharntke, T., Tylianakis, J. M., Rand, T. A., Didham, R. K., Fahrig, L., Batáry, P., Bengtsson, J., 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*(3), 661–685. https://doi.org/10.1111/j.1469-185X.2011.00216.x

van der Plas, F., Allan, E., Fischer, M., Alt, F., Arndt, H., Binkenstein, J., Blaser, S., Blüthgen, N., Böhm, S., Hölzel, N., Klaus, V. H., Kleinebecker, T., Morris, K., Oelmann, Y., Prati, D., Renner, S. C., Rillig, M. C., Schaefer, H. M., Schloter, M., … Manning, P. (2019). Towards the development of general rules describing landscape heterogeneity-multifunctionality relationships. *Journal of Applied Ecology*, *56*(1), 168–179. https://doi.org/10.1111/1365-2664.13260

Whittingham, M. J. (2011). The future of agri-environment schemes: Biodiversity gains and ecosystem service delivery?: Editorial. *Journal of Applied Ecology*, *48*(3), 509–513. https://doi.org/10.1111/j.1365-2664.2011.01987.x

# Landscape management for grassland multifunctionality

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., Börshig C., Jung K., Prati D., Schäfer M., Schäfer D., Schöning I., Schrumpf M., Tschapka M., Westphal C. & Manning P.

## Supplementary information

# Appendix 1. Additional information

**Table S** 1 List of threatened species per region. Lists compiled from Breunig and Demuth (1999); Korsch and Westhus (2001) and Ristow et al. (2006).

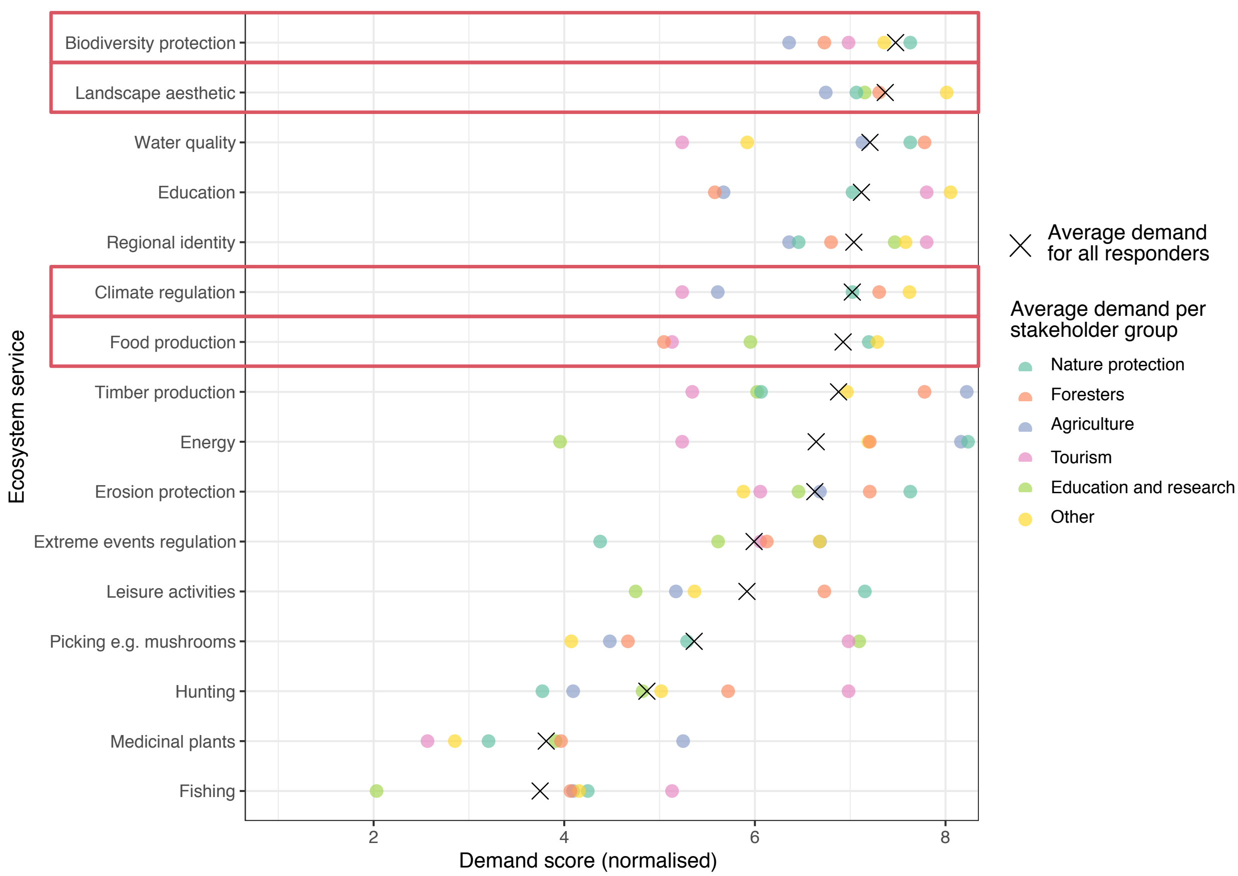
| South-West region (Baden-Württemberg) | Central region (Thuringen) | North region (Brandenburg) | |
|---|---|---|---|
| *Anchusa officinalis* | *Anchusa officinalis* | *Alchemilla vulgaris aggr.* | *Melilotus altissimus* |
| *Antennaria dioica* | *Antennaria dioica* | *Antennaria dioica* | *Myosotis discolor* |
| *Anthemis tinctoria* | *Betonica officinalis* | *Arabis hirsuta aggr.* | *Odontites vernus aggr.* |
| *Botrychium lunaria* | *Botrychium lunaria* | *Asperula cynanchica* | *Onobrychis viciifolia* |
| *Bunium bulbocastanum* | *Bunium bulbocastanum* | *Betonica officinalis* | *Ophioglossum vulgatum* |
| *Eryngium campestre* | *Calamagrostis canescens* | *Bistorta officinalis* | *Orchis militaris* |
| *Gagea pratensis* | *Campanula glomerata* | *Botrychium lunaria* | *Origanum vulgare* |
| *Gentiana verna* | *Euphorbia verrucosa* | *Briza media* | *Orobanche caryophyllacea* |
| *Koeleria macrantha* | *Euphrasia rostkoviana aggr.* | *Carex flacca* | *Phyteuma spicatum* |
| *Lathyrus nissolia* | *Galium verum* | *Carex montana* | *Platanthera bifolia* |
| *Muscari neglectum aggr.* | *Gentiana verna* | *Carum carvi* | *Platanthera chlorantha* |
| *Myosotis discolor* | *Gentianella germanica* | *Centaurium erythraea* | *Polygala amarella* |
| *Myosurus minimus* | *Helianthemum nummularium* | *Chaerophyllum aureum* | *Polygala comosa aggr.* |
| *Ophioglossum vulgatum* | *Koeleria macrantha* | *Cirsium acaule* | *Potentilla heptaphylla* |
| *Orobanche caryophyllacea* | *Lathyrus nissolia* | *Colchicum autumnale* | *Prunus avium* |
| *Phyteuma orbiculare* | *Myosotis discolor* | *Cruciata laevipes* | *Pulsatilla vulgaris* |
| *Pseudolysimachion spicatum* | *Myosurus minimus* | *Cynosurus cristatus* | *Ranunculus auricomus* |
| *Pulsatilla vulgaris* | *Odontites vernus aggr.* | *Eryngium campestre* | *Rhinanthus minor* |
| *Selinum carvifolia* | *Ophioglossum vulgatum* | *Euphrasia rostkoviana aggr.* | *Salvia pratensis* |
| *Seseli annuum* | *Orchis militaris* | *Fragaria viridis* | *Sanguisorba minor* |
| *Stachys arvensis* | *Orobanche caryophyllacea* | *Galium pumilum* | *Selinum carvifolia* |
| *Teucrium montanum* | *Phyteuma orbiculare* | *Geranium dissectum* | *Seseli annuum* |
| *Trifolium montanum* | *Platanthera bifolia* | *Geranium pratense* | *Sherardia arvensis* |
| | *Pseudolysimachion spicatum* | *Geranium sylvaticum* | *Silaum silaus* |
| | *Pulsatilla vulgaris* | *Helictotrichon pratense* | *Stachys arvensis* |
| | *Sedum telephium* | *Koeleria macrantha* | *Trifolium alpestre* |
| | *Seseli annuum* | *Lathyrus nissolia* | *Veronica teucrium* |
| | *Stachys arvensis* | *Listera ovata* | |
| | *Vicia lathyroides* | | |
| | *Viola collina* | | |

**Table S 2 Variation of the final benefits and service indicators, with the land-use intensity class in each region**. Values were first corrected for the environment (see Methods). Different letters indicate differences significant at 5%.

| **Final benefits** or *service indicator* | Land-use intensity | South-West | Central | North |
|---|---|---|---|---|
| Conservation value | Low | 5 ± 4.5 $^{b}$ | 4 ± 7.9 $^{b}$ | -0.1 ± 1.8 $^{a}$ |
| | Medium | -1.2 ± 4.4 $^{a}$ | -0.9 ± 3.7 $^{a}$ | 0.6 ± 2.5 $^{a}$ |
| | High | -3.9 ± 5.1 $^{a}$ | -3.1 ± 4.5 $^{a}$ | -0.4 ± 1.7 $^{a}$ |
| Plant species richness | Low | 8.5 ± 7.8 $^{b}$ | 7.9 ± 15.6 $^{b}$ | 0.1 ± 3.2 $^{a}$ |
| | Medium | -2 ± 6.8 $^{a}$ | -1.8 ± 7.3 $^{a}$ | 0.3 ± 2.9 $^{a}$ |
| | High | -6.6 ± 9.5 $^{a}$ | -6.2 ± 9 $^{a}$ | -0.4 ± 3.3 $^{a}$ |
| Cover by redlist species | Low | 1.4 ± 3 $^{b}$ | 0.1 ± 0.7 $^{a}$ | -0.4 ± 0.8 $^{a}$ |
| | Medium | -0.3 ± 2.5 $^{ab}$ | 0 ± 0.5 $^{a}$ | 0.8 ± 4.8 $^{a}$ |
| | High | -1.1 ± 1.5 $^{a}$ | -0.1 ± 0.9 $^{a}$ | -0.4 ± 0.8 $^{a}$ |
| Fodder production | Low | -313.5 ± 301.5 $^{a}$ | -244.8 ± 279.9 $^{a}$ | -181.8 ± 305.6 $^{a}$ |
| | Medium | 16 ± 488.3 $^{b}$ | -47.5 ± 329.4 $^{a}$ | 28.2 ± 257 $^{ab}$ |
| | High | 298.4 ± 554.8 $^{b}$ | 289.5 ± 431.6 $^{b}$ | 155.3 ± 452.1 $^{b}$ |
| Biomass production | Low | -24.7 ± 19.7 $^{a}$ | -17.6 ± 20.2 $^{a}$ | -11.9 ± 21.7 $^{a}$ |
| | Medium | 0.9 ± 24.2 $^{b}$ | -3.9 ± 22 $^{a}$ | 3.3 ± 16.5 $^{ab}$ |
| | High | 23.9 ± 37.7 $^{c}$ | 21.3 ± 29.3 $^{b}$ | 8.7 ± 22 $^{b}$ |
| Plant protein content | Low | 0 ± 0.2 $^{a}$ | 0 ± 0.2 $^{a}$ | 0 ± 0.5 $^{a}$ |
| | Medium | 0 ± 0.3 $^{a}$ | 0 ± 0.2 $^{a}$ | -0.1 ± 0.4 $^{a}$ |
| | High | 0 ± 0.3 $^{a}$ | 0 ± 0.3 $^{a}$ | 0 ± 0.6 $^{a}$ |
| Aesthetic value | Low | 11.4 ± 28.8 $^{b}$ | 6.2 ± 18.9 $^{a}$ | 3.1 ± 9.6 $^{a}$ |
| | Medium | -0.9 ± 19.3 $^{ab}$ | -0.8 ± 8.5 $^{a}$ | -0.1 ± 5.4 $^{a}$ |
| | High | -10.5 ± 10.7 $^{a}$ | -5.5 ± 8.6 $^{a}$ | -3 ± 4.9 $^{a}$ |
| Flower cover | Low | -0.2 ± 3.5 $^{a}$ | 0.6 ± 3.8 $^{a}$ | -0.3 ± 1.1 $^{a}$ |
| | Medium | 0.2 ± 5 $^{a}$ | 1.1 ± 4.9 $^{a}$ | 0.3 ± 2.5 $^{a}$ |
| | High | 0.1 ± 2.9 $^{a}$ | -1.6 ± 3.9 $^{a}$ | -0.1 ± 2.2 $^{a}$ |
| Butterfly abundance | Low | 33.5 ± 84.8 $^{b}$ | 17.5 ± 53.9 $^{a}$ | 9.3 ± 28.3 $^{a}$ |
| | Medium | -2.5 ± 52.9 $^{a}$ | -3.4 ± 26.7 $^{a}$ | -0.8 ± 15.3 $^{a}$ |
| | High | -31.1 ± 31.6 $^{a}$ | -14.3 ± 23.2 $^{a}$ | -8.5 ± 14.7 $^{a}$ |
| Bird family richness | Low | 0.8 ± 1.9 $^{a}$ | 0.6 ± 3.2 $^{a}$ | 0.4 ± 1.8 $^{a}$ |
| | Medium | -0.2 ± 2.6 $^{a}$ | 0.1 ± 3 $^{a}$ | 0.2 ± 1.8 $^{a}$ |
| | High | -0.6 ± 1.6 $^{a}$ | -0.7 ± 2 $^{a}$ | -0.6 ± 2.2 $^{a}$ |
| C stock | Low | 1.9 ± 10.1 $^{a}$ | 0.3 ± 11.5 $^{a}$ | 18.4 ± 26.9 $^{b}$ |
| | Medium | -0.9 ± 11.3 $^{a}$ | 1.4 ± 8.7 $^{a}$ | -10.1 ± 60.6 $^{a}$ |
| | High | -1.1 ± 9.2 $^{a}$ | -1.6 ± 12.8 $^{a}$ | -9 ± 36.7 $^{a}$ |

702
**Figure S 1. Stakeholder groups' ecosystem preferences.** Over a series of group interviews, 29 responders from multiple stakeholder groups in the three regions of interest were presented with a list of 16 services. For each service, they were asked to quantify their demand for the corresponding service, from 1 (the service is not important) to 5 (very important). The obtained scores were normalised by the total number of points given by each responder, then averaged by stakeholder group**.** a. Mean demand score for all considered services, per stakeholder group (coloured dots) or all groups considered (crosses). We then retained for the analysis only the four main services that can be delivered by grasslands (marked with boxes). b. Mean demand scores for the services included in the analysis of the current paper, per stakeholder group; the demand scores were normalised by the number of points given by responders to only those four services.

a.




b.

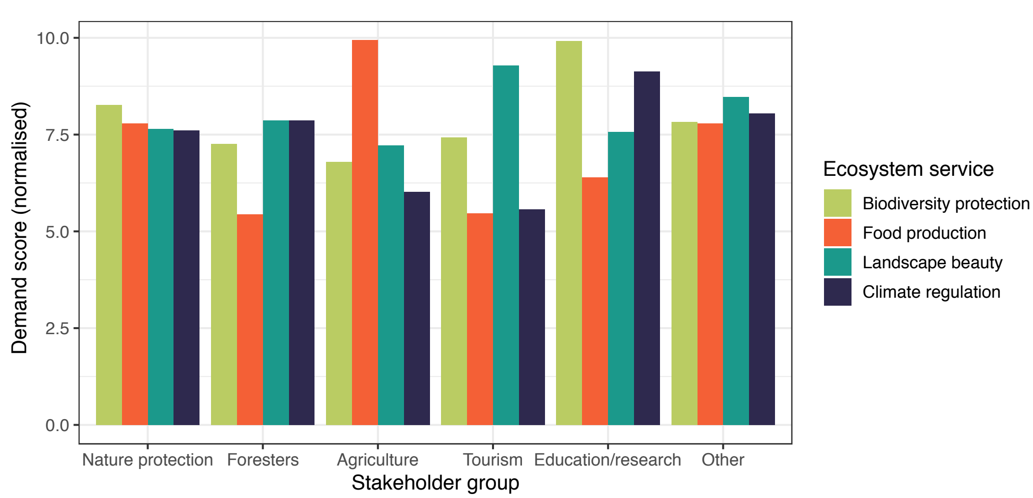
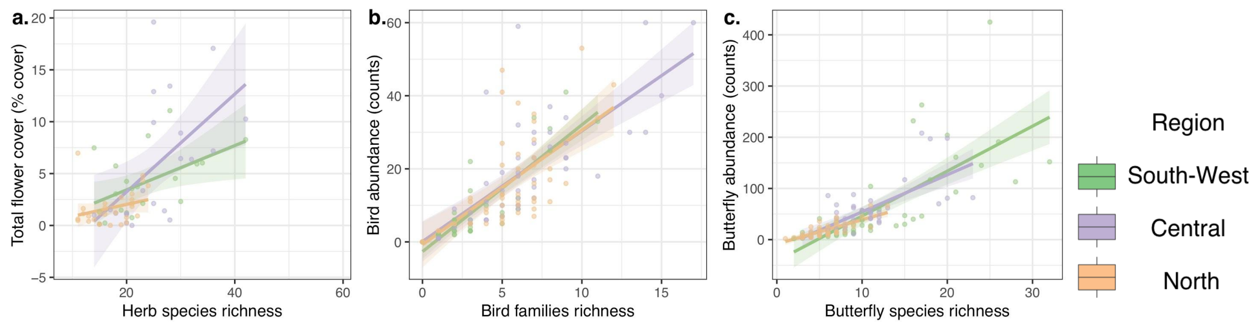


**Figure S 2 Correlations among potential service indicators**




703

704

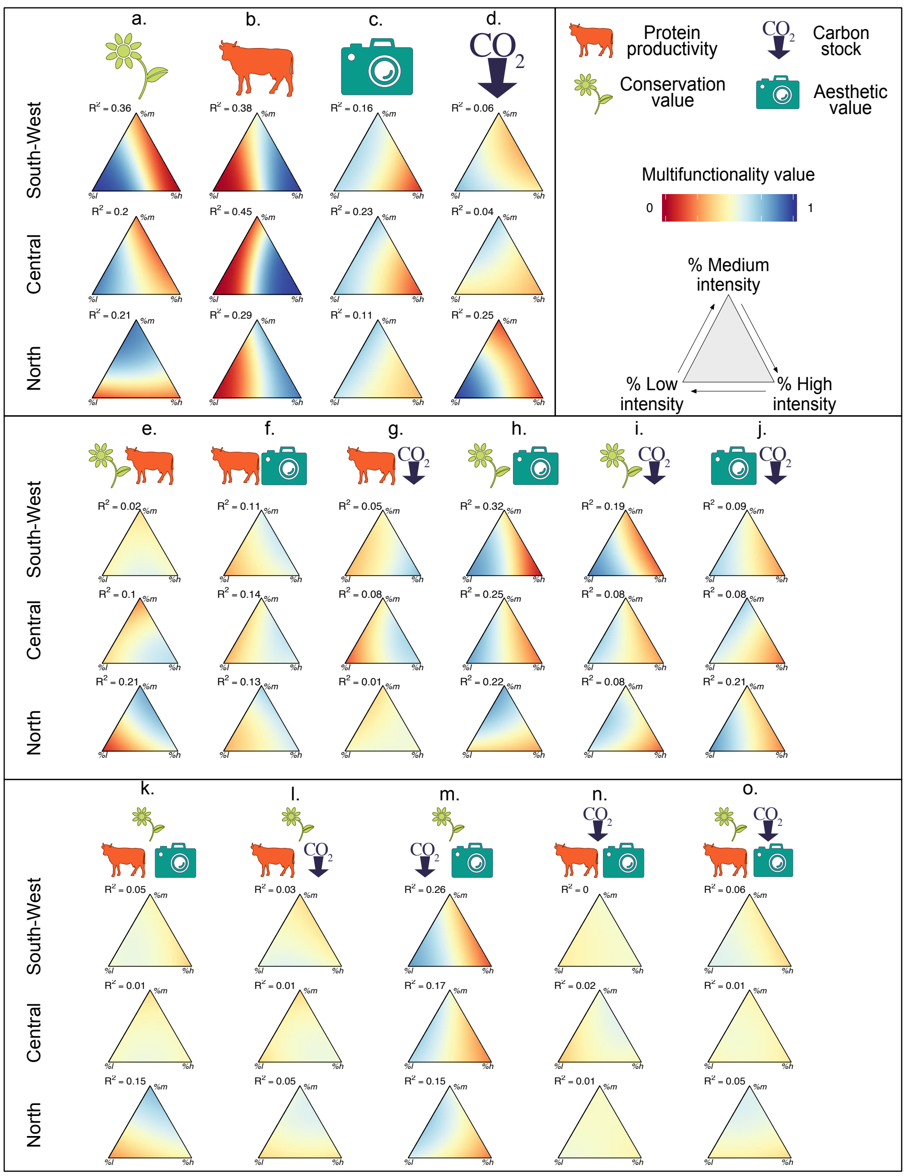
705

706

**Figure S 3** Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites). **This figure shows all the service combinations for the 'threshold' approach, partly shown in Figure 4.**

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the median. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the median. Blue indicates higher multifunctionality values, orange lower.

The specific shape of the response of single services to landscape composition in the region led to slightly different responses for the corresponding combination of services. For example, when considering conservation and production (panel. e), no optimum that maximizes both services could be found when the response of the services to land-use intensity were perfectly opposed and responded with similar intensity to landscape composition ($R^2$ = 36% and 38% respectively), as was the case in the South-West. However, a partial optimum could be found when the responses of services to intensification differed, such as in the Central region, where fodder production responded much more strongly to landscape composition ($R^2$ = 45%) than conservation ($R^2$ = 20%), resulting in slightly higher multifunctionality in landscapes dominated by high-intensity sites.



## Supplementary methods: Topographical Wetness Index calculation

We calculated the Topographic Wetness Index (TWI) of each plot, defined as ln($a$/tan$B$) where $a$ is the specific catchment area (cumulative upslope area which drains through a Digital Elevation Model (DEM, http://www.bkg.bund.de) cell, divided by per unit contour length) and tan$B$ is the slope gradient in radians calculated over a local region surrounding the cell of interest (Gessler et al. 1995; Sørensen et al. 2006). TWI therefore combines both upslope contributing area (determining the amount of water received from upslope areas) and slope (determining the loss of water from the site to downslope areas). TWI was calculated from raster DEM data with a cell size of 25 m for all plots, using ArcGIS tools (flow direction and flow accumulation tools of the hydrology toolset and raster calculator). The TWI measure used was the average value for a 4 × 4 window centred on the plot, i.e. 16 DEM cells corresponding to an area of 100 m ×100 m.

## Supplementary references

Gessler PE, Moore ID, McKenzie NJ, Ryan PJ Soil-landscape modelling and spatial prediction of soil attributes. Int J Geogr Inf Syst 9:421–432

Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol Earth Syst Sci 10:101–112. https://doi.org/10.5194/hess-10-101-2006

# Appendix 2. Details on sensitivity analyses

We explored how multiple choices in the simulations and calculation of multifunctionality could affect the results of our analyses. In the first part of this appendix, we first list all the different sensitivities that we identified and detail how their implementation modified the methodology detailed in the main text. In the second part, we consider in turn each of the main results presented in the paper and describe the potential variations highlighted by the analyses. We describe only analyses departing from the main analysis by one parameter – further combinations can be explored in the online tool.

## 1. Methods

The sensitivities we identified were the following:

a. Analysis on non-environmentally corrected data.

The main analysis was conducted on service indicator values that were first corrected at the plot level by environmental covariates (see Methods). In these sensitivity analyses we also conduct the analyses using raw data.

b. Classification into classes of low, medium and high land-use intensity.

In the main analyses, the plots are classified in the three intensity categories based on 33% quantiles of the intensity values (e.g. lowest, medium and highest thirds) within the regions. In these sensitivity analyses we remove the “intermediate”, potentially confounding plots by using plots within the lowest, medium and highest 20% quantiles (e.g. lowest, medium and highest fifths) of the intensities within each region. This reduced the number of plots available during the landscape simulations, and to avoid running the analysis on identical landscapes we adapted the analyses by lowering the number of landscape replicates for each combination (*).

c. Number of plots per landscape.

In the main analysis, each landscape was composed of 10 sites. We also run similar analyses with 7 and 13 sites per landscapes. This also affected the number of possible landscape combinations, and we consequently adapted the number of landscape replicates for each combination (*).

d. Calculation of landscape-scale service indicator values.

In the main analysis, landscape-scale service indicator values consisted either of gamma diversity (plant and bird diversity) or of the sum of the services provided by all plots in the landscape (other services). We also considered a situation in which high level of each service is expected in only part of the landscape, and calculated the landscape indicator value as the maximum of the indicator in all plots of the landscape.

e. Calculation of landscape-scale multifunctionality

In the main analysis, landscape scale multifunctionality was calculated as the number of services above the median or as 1 if all services were over 25%, 0 otherwise. We also calculated multifunctionality as the number of services above the $40^{th}$ or $60^{th}$ quantiles of the distribution of the values in all landscapes; or as 75% of the maximum (measured as $95^{th}$ quantile to avoid outliers).

We also measured multifunctionality as the average of the (scaled) values for all considered services.

(*) In the main analysis, there were 66 possible landscape compositions (from 0 to 10 sites of each intensity), and 15 random landscape replicates per composition, hence 990 different landscapes. Changing the number of sites per landscapes changed the number of possible combinations (36 possible combinations for 7 sites, 105 for 13 sites) and to keep the total number of simulated landscapes approximately similar, we used 1000/(number of combinations) landscape replicates per combination (i.e. 10 for 13 plots, 28 for 7 plots).

When using only the 20% lowest, medium and highest intensity sites decreased the size of the regional pool from which to build landscapes. In that case we used only 7 sites per landscape and the extreme compositions (e.g. 100% of one intensity) were represented by slightly less landscape replicates.

## 2. Results

### a. Plot-level correlations among indicators and variation of service indicator values with land-use intensity and.

When correcting for the environment, there were positive correlations between flower cover, butterfly abundance, plant species richness and cover of red list plants. These indicators were usually positively correlated with bird richness and negatively correlated with biomass production. Most of these relationships were similar when considering raw, non environmentally-corrected data (Table S 3).

As shown in Table S 4, the variation of plot-level service indicators with land-use intensity was not strongly affected by using raw data instead of environmentally corrected residuals (as is shown in Figure 2 and Table S 2), except that the response of some services (e.g. plant richness) to intensity was more linear (no apparent “threshold”) when it was corrected for the environment.

### b. Landscape-level correlations among final benefits.

Correlations among landscape-scale services was similar when considering services calculated as the maximum value of each service (instead of the sum) in the landscape (**4**), when changing the number of plots per landscape (Figure S 5, Figure S 6) or using data not corrected for the environment (Figure S 7).

c. Multifunctionality response to landscape composition

The following figures present the multifunctionality response to landscape composition as affected by the parameters of the model.

Calculating the landscape-scale services as the maximum instead of the sum (Figure S 9) did not change the direction of the response. It led to weaker responses of single services (top row) and marginally stronger variability when including all services, with slightly higher multifunctionality in landscapes composed of mostly low-intensity plots in the Central and South-West region.

Decreasing (Figure S 10) or increasing (Figure S 11) the number of plots per landscape respectively weakened and strengthened the response of single services to landscape composition, possibly because including more sites made for a higher chance to select sites with very high or low values, especially in the extreme compositions (e.g. 100% low intensity, 100% high intensity). When considering multiple services, including more sites did not change the response of multifunctionality.

For multifunctionality counted as the proportion of services above a given threshold, changing the threshold from the median to the 40$^{th}$ or 60$^{th}$ quantile of the distribution slightly switched the multifunctionality to higher or, respectively, lower values (Figure S 12 and Figure S 13) but the general form of the response was not affected.

The same was observed in the "compromise" scenario, when multifunctionality was calculated as 1 when all services were above a threshold, and 0 otherwise. Changing the threshold from the 25$^{th}$ quantile of the distribution to the 15$^{th}$ or 35$^{th}$ quantiles (Figure S 14 and Figure S 15) made it easier (respectively, more difficult) to provide all services at the required level, resulting in overall higher (resp. lower) multifunctionality values but without affecting the form of the response.

Calculating multifunctionality as the average of all services gave similar results as the main thresholding approach (Figure S 16).

Conversely, changing the threshold from a value based on the distribution of the service (e.g. the median, as was used in the text) to a proportion of the maximum (e.g. 50% in Figure S 17) completely changed the response of individual services, due to different shapes of the distribution among services.

d. Effect of other drivers on the responsivity of multifunctionality to landscape composition

Table S 5 presents the result of the model with the responsivity (i.e. range) of multifunctionality over all landscape compositions as a response, and the ratio of the effect of land-use intensity and other environmental variables; the number of services included; or the service response variance as explanatory variables.

The relative effect of land-use intensity compared to other environmental variables was calculated as the ratio between the slope coefficient between individual services and land-use intensity over the maximum slope coefficient between the service and all other environmental covariates. It was thus calculated only for raw (non environmentally-corrected) values of ecosystem services. It was significantly positive regardless of the model parameters, except when landscape-scale services were calculated based on the maximum of the sites.

The multifunctionality range decreased with the number of services considered in 19 out of the 23 scenarios considered, supporting our main conclusions. However, it increased with the number of services when multifunctionality was calculated as 1 if all the services were above the $15^{th}$ percentile, and did not change when multifunctionality was calculated as 1 if all the services were above the $25^{th}$ percentile. This is because for low thresholds such as these, multifunctionality is expected to be high everywhere if too few services are considered.

Finally, the responsiveness of multifunctionality significantly decreased with the service response variance in 14 of the 23 considered scenarios. There was no positive relationship.

**Table S 3 Plot-level correlation among ecosystem services.** a. corrected for the environment, and b. raw service values. * P < 0.05, ** P < 0.01, *** P < 0.001.

| **a. Data corrected for the environment** | Flower cover | Butterfly abundance | Biomass production | Nitrogen content | Organic C stock | Plant species richness | Bird family richness |
|---|---|---|---|---|---|---|---|
| Butterfly abundance | 0.43 *** | | | | | | |
| Biomass production | -0.04 | -0.38 *** | | | | | |
| Nitrogen content | -0.21 ** | -0.14 | 0.13 | | | | |
| Organic C stock | -0.01 | -0.06 | -0.04 | 0.17 * | | | |
| Plant richness | 0.32 *** | 0.49 *** | -0.54 *** | -0.1 | -0.04 | | |
| Bird richness | 0.11 | 0.24 ** | -0.24 ** | -0.12 | 0 | 0.4 *** | |
| Cover by redlist species | 0.24 ** | 0.37 *** | -0.15 | -0.2 * | -0.2 * | 0.24 ** | 0.11 |

| **b. Raw service indicator values** | Flower cover | Butterfly abundance | Biomass production | Nitrogen content | Organic C stock | Plant species richness | Bird family richness |
|---|---|---|---|---|---|---|---|
| Butterfly abundance | 0.32 *** | | | | | | |
| Biomass production | -0.07 | -0.48 *** | | | | | |
| Nitrogen content | -0.1 | -0.22 ** | 0.19 * | | | | |
| Organic C stock | -0.21 * | -0.17 * | -0.02 | 0.08 | | | |
| Plant richness | 0.35 *** | 0.52 *** | -0.53 *** | -0.17 * | -0.33 *** | | |
| Bird richness | 0.12 | 0.26 ** | -0.29 *** | -0.11 | -0.03 | 0.31 *** | |
| Cover by redlist species | 0.1 | 0.36 *** | -0.2 * | -0.19 * | -0.08 | 0.2 * | 0.1 |

**Table S 4 Variation of the final benefits and service indicators with the land-use intensity class in each region (mean ± sd)**. Indicators were not corrected for the environment (see Methods). Different letters indicate differences significant at 5%.

| **Final benefits** or *service indicator* | Land-use intensity | South-West | Central | North |
|---|---|---|---|---|
| Conservation value | Low | 29.1 ± 5.8 $^{b}$ | 30.2 ± 9.2 $^{b}$ | 13.6 ± 2.3 $^{a}$ |
| | Medium | 18.6 ± 4.6 $^{a}$ | 22.4 ± 3.3 $^{a}$ | 16.4 ± 3 $^{a}$ |
| | High | 18.7 ± 4.2 $^{a}$ | 19.1 ± 3.5 $^{a}$ | 14.7 ± 3.3 $^{a}$ |
| Plant species richness | Low | 55.1 ± 9.2 $^{b}$ | 59.9 ± 18 $^{b}$ | 27.1 ± 4.7 $^{a}$ |
| | Medium | 36.5 ± 7.2 $^{a}$ | 44.7 ± 6.5 $^{a}$ | 31 ± 4.4 $^{a}$ |
| | High | 37.4 ± 8.5 $^{a}$ | 38 ± 6.9 $^{a}$ | 29.3 ± 6.5 $^{a}$ |
| Cover by redlist species | Low | 3.2 ± 5.1 $^{b}$ | 0.4 ± 1 $^{a}$ | 0.1 ± 0.3 $^{a}$ |
| | Medium | 0.6 ± 2.3 $^{a}$ | 0.1 ± 0.4 $^{a}$ | 1.8 ± 5 $^{a}$ |
| | High | 0 ± 0.1 $^{a}$ | 0.2 ± 1 $^{a}$ | 0 ± 0.1 $^{a}$ |
| Fodder production | Low | 325.4 ± 263.5 $^{a}$ | 318.2 ± 293.8 $^{a}$ | 661.9 ± 299.9 $^{a}$ |
| | Medium | 1045 ± 595.8 $^{b}$ | 612.8 ± 342.7 $^{a}$ | 882.5 ± 331.4 $^{ab}$ |
| | High | 1182.2 ± 511.8 $^{b}$ | 1020.4 ± 427.3 $^{b}$ | 1056.5 ± 500.5 $^{b}$ |
| Biomass production | Low | 23.7 ± 17.1 $^{a}$ | 25.8 ± 21.4 $^{a}$ | 48 ± 21.3 $^{a}$ |
| | Medium | 73 ± 31.1 $^{b}$ | 46.4 ± 23.5 $^{a}$ | 64.8 ± 20.8 $^{ab}$ |
| | High | 88.7 ± 34.5 $^{b}$ | 78.4 ± 28.6 $^{b}$ | 72.7 ± 25.2 $^{b}$ |
| Plant protein content | Low | 2.1 ± 0.2 $^{a}$ | 2 ± 0.3 $^{a}$ | 2.2 ± 0.6 $^{a}$ |
| | Medium | 2.2 ± 0.4 $^{a}$ | 2.1 ± 0.3 $^{a}$ | 2.2 ± 0.5 $^{a}$ |
| | High | 2.1 ± 0.3 $^{a}$ | 2.1 ± 0.3 $^{a}$ | 2.3 ± 0.6 $^{a}$ |
| Aesthetic value | Low | 45.2 ± 36.7 $^{b}$ | 31.2 ± 22.3 $^{b}$ | 7.6 ± 3.2 $^{a}$ |
| | Medium | 14.9 ± 17.1 $^{a}$ | 17.3 ± 6.6 $^{a}$ | 8.8 ± 5 $^{a}$ |
| | High | 8.9 ± 5.7 $^{a}$ | 11.8 ± 9.4 $^{a}$ | 8.8 ± 6.7 $^{a}$ |
| Flower cover | Low | 4.1 ± 2.8 $^{a}$ | 4.9 ± 5.4 $^{ab}$ | 2 ± 2.5 $^{a}$ |
| | Medium | 2.5 ± 3.3 $^{a}$ | 7 ± 5.6 $^{b}$ | 1.9 ± 1.9 $^{a}$ |
| | High | 3.9 ± 3.8 $^{a}$ | 2.5 ± 3.3 $^{a}$ | 1.4 ± 1.5 $^{a}$ |
| Butterfly abundance | Low | 126.9 ± 108.2 $^{b}$ | 82.2 ± 62 $^{b}$ | 15.9 ± 9.6 $^{a}$ |
| | Medium | 39.2 ± 47.7 $^{a}$ | 38.6 ± 19.8 $^{a}$ | 20.1 ± 14.4 $^{a}$ |
| | High | 20.3 ± 14.5 $^{a}$ | 28.2 ± 27 $^{a}$ | 20.8 ± 19.2 $^{a}$ |
| Bird family richness | Low | 4.6 ± 2.6 $^{b}$ | 6.5 ± 3.7 $^{a}$ | 5.1 ± 1.7 $^{a}$ |
| | Medium | 2.9 ± 2.5 $^{ab}$ | 6.2 ± 3.1 $^{a}$ | 4.6 ± 2.1 $^{a}$ |
| | High | 2.5 ± 1.8 $^{a}$ | 4.6 ± 2.2 $^{a}$ | 4.1 ± 2.4 $^{a}$ |
| C stock | Low | 63 ± 17.5 $^{a}$ | 46.2 ± 13.4 $^{a}$ | 152.7 ± 107 $^{b}$ |
| | Medium | 65.2 ± 13.6 $^{a}$ | 46.1 ± 9.1 $^{a}$ | 69.9 ± 87.4 $^{a}$ |
| | High | 67.2 ± 10.1 $^{a}$ | 44.6 ± 13.8 $^{a}$ | 92.7 ± 94.9 $^{a}$ |

**Figure S 4 Trade-offs between landscape-scale ecosystem service measures**. This figure differs from Figure 3 as the landscape-scale services were calculated based on the maximum (instead of sum) of the services provided by all sites in the landscape.

The colour and size of the circles denote the strength of the correlation between pairs of variables, within each region. Crosses indicate no significant correlations at 5% (Holm correction for multiple testing).

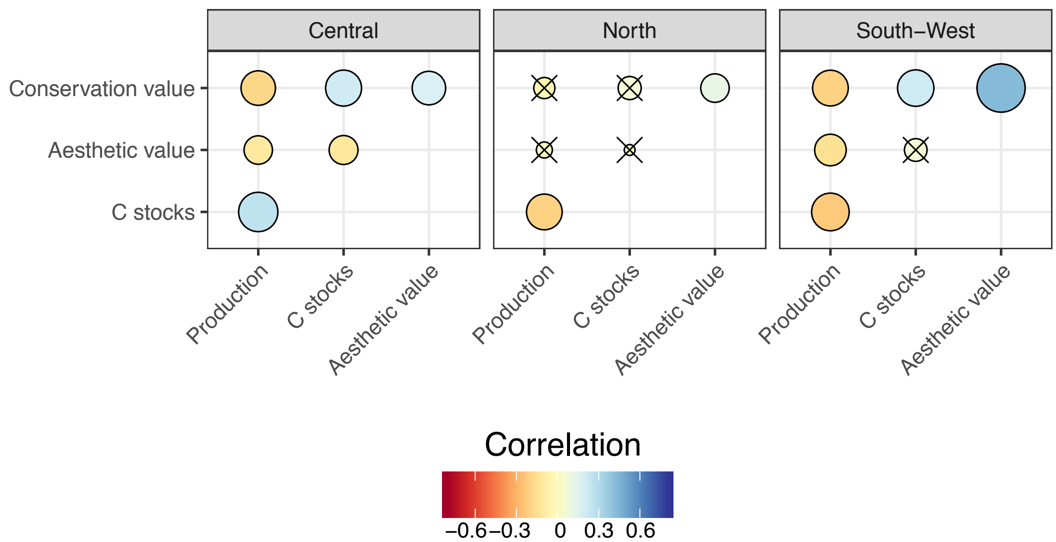



**Figure S 5 Trade-offs between landscape-scale ecosystem service measures.** This figure differs from Figure 3 as the landscapes included 7 (instead of 10) sites.

The colour and size of the circles denote the strength of the correlation between pairs of variables, within each region. Crosses indicate no significant correlations at 5% (Holm correction for multiple testing).

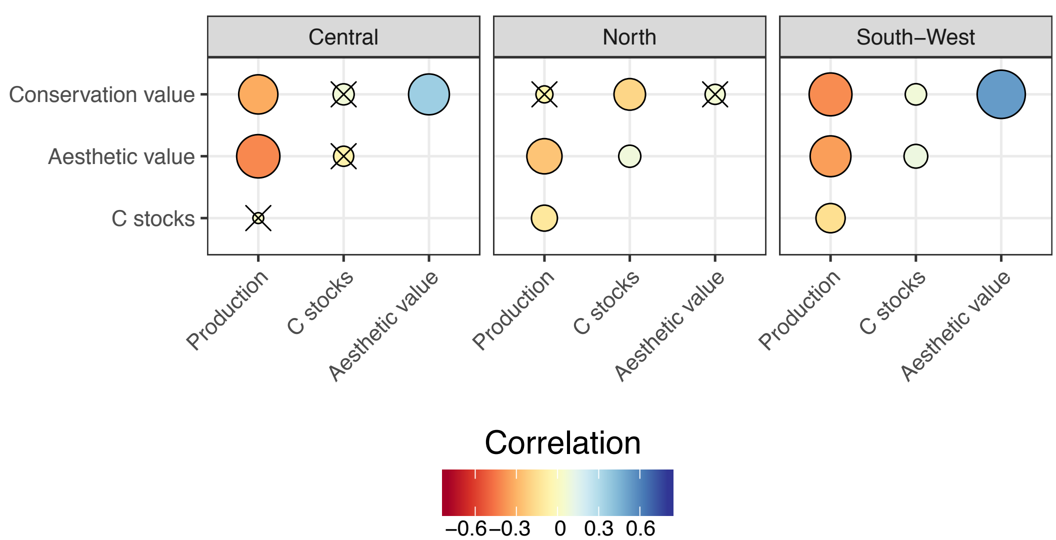


**Figure S 6 Trade-offs between landscape-scale ecosystem service measures.** This figure differs from Figure 3 as the landscapes included 13 (instead of 10) sites.

The colour and size of the circles denote the strength of the correlation between pairs of variables, within each region. Crosses indicate no significant correlations at 5% (Holm correction for multiple testing).

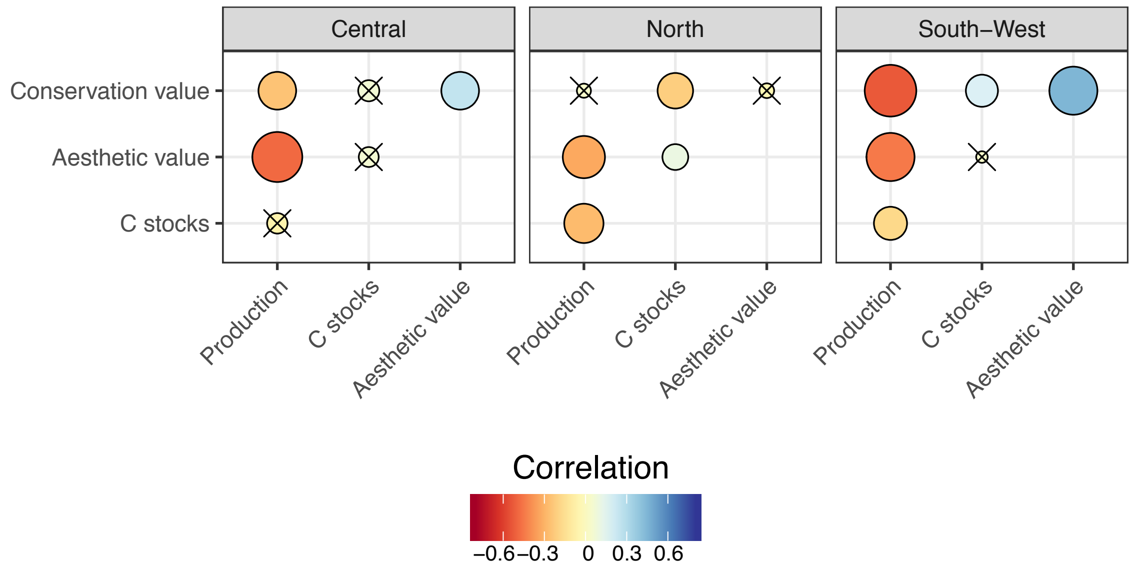



**Figure S 7 Trade-offs between landscape-scale ecosystem service measures.** This figure differs from Figure 3 as the ecosystem service indicators were not corrected for the environment before analysis.

The colour and size of the circles denote the strength of the correlation between pairs of variables, within each region. Crosses indicate no significant correlations at 5% (Holm correction for multiple testing).

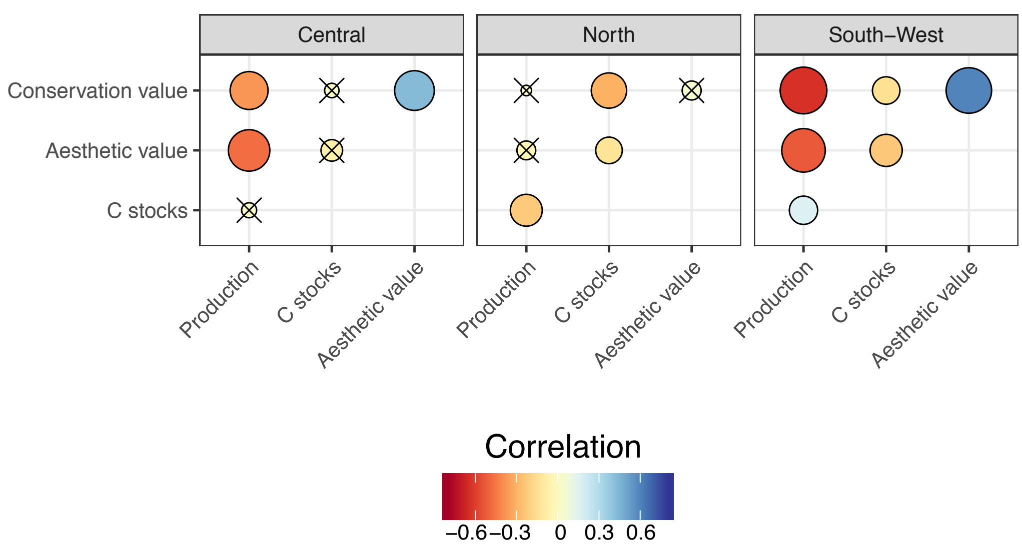


**Figure S 8 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure shows all the service combinations for the 'compromise' approach, partly shown in Figure 5.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the median for combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the median. Blue indicates higher multifunctionality values, orange lower.

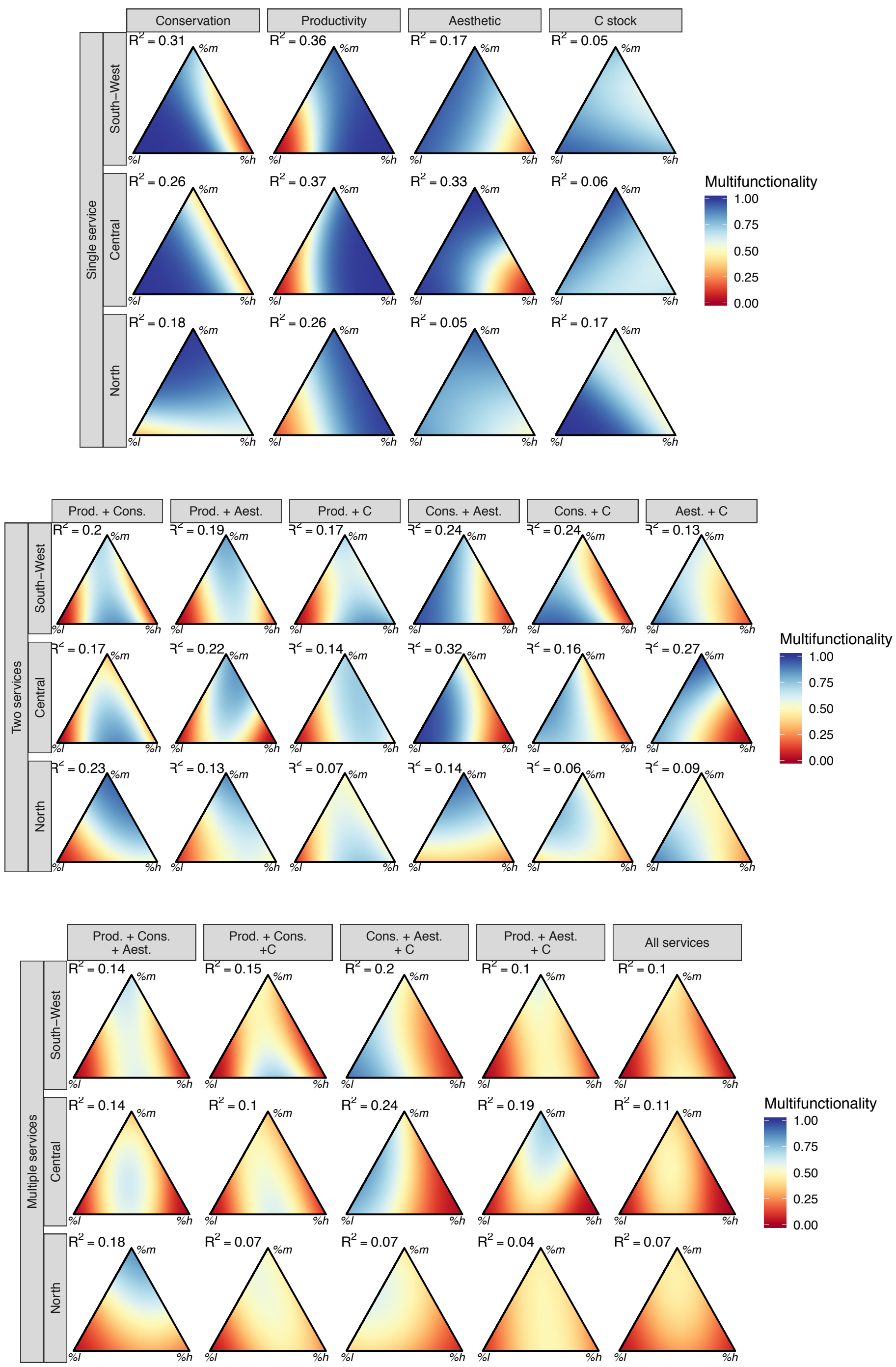


**Figure S 9 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that landscape-scale ecosystem service values were calculated as the maximum, not the sum, of site-level ecosystem services.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the median. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the median. Blue indicates higher multifunctionality values, orange lower.

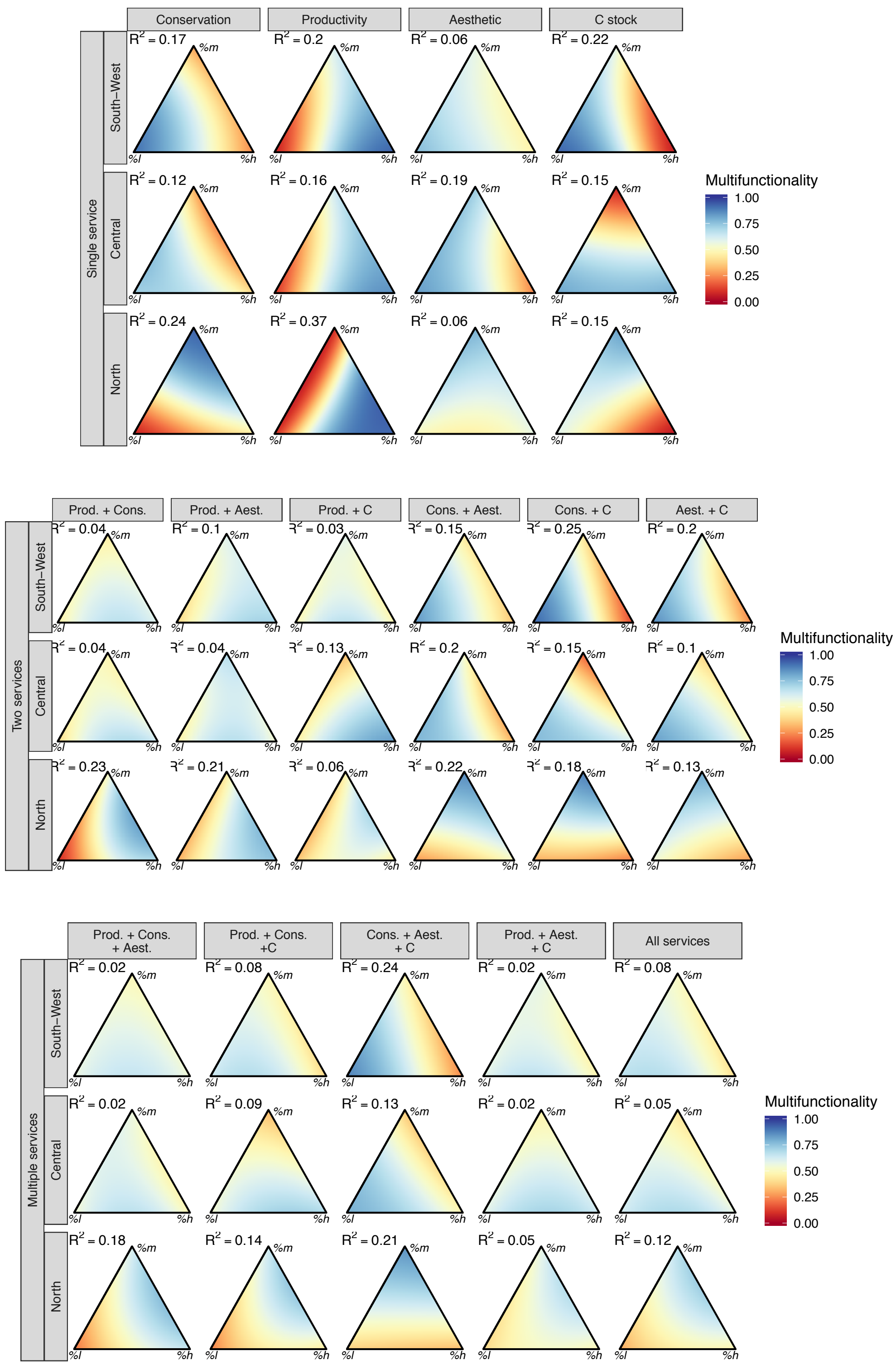


**Figure S 10 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that landscapes were composed of 7 sites, instead of 10.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the median. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the median. Blue indicates higher multifunctionality values, orange lower.

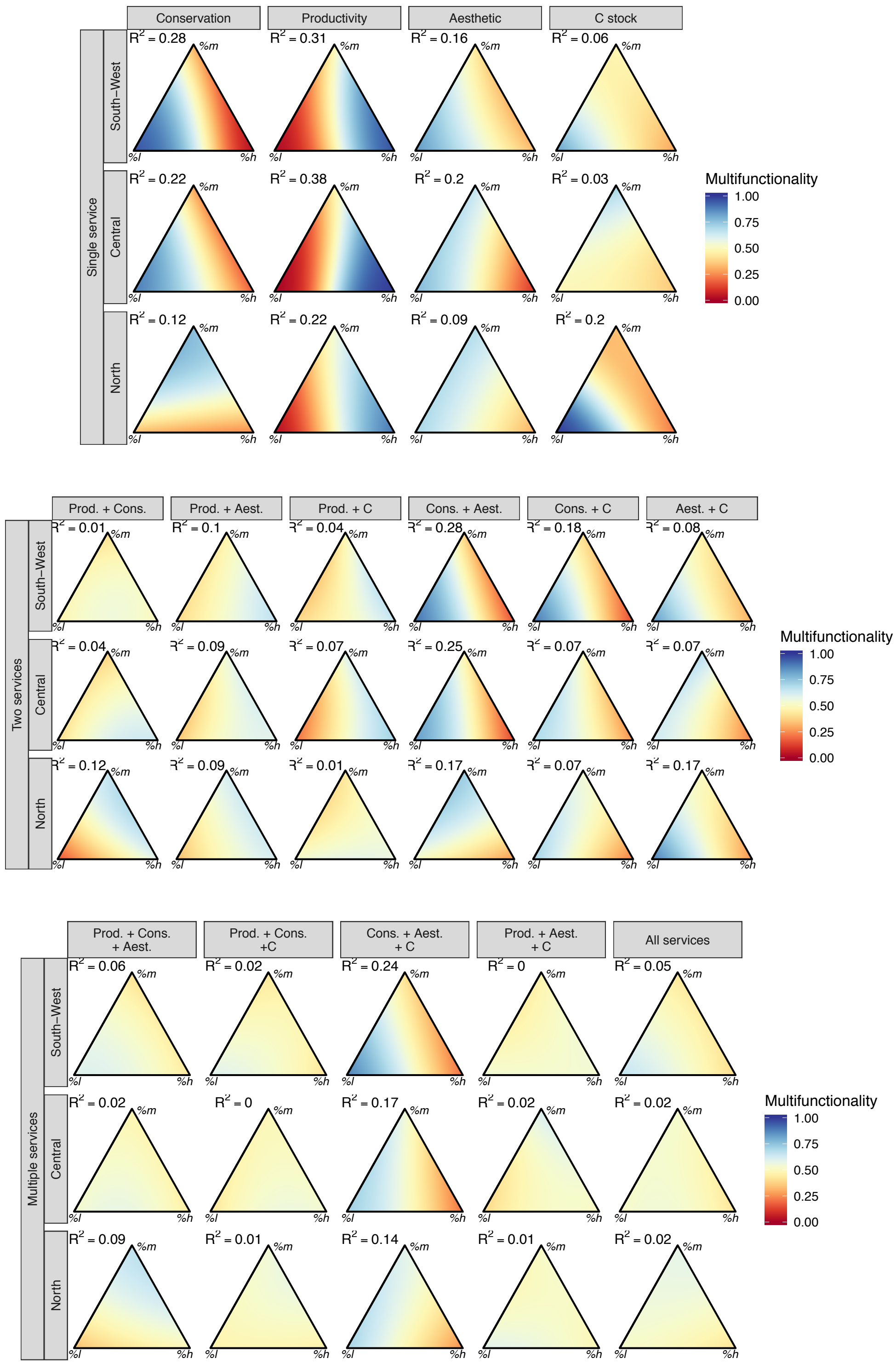


**Figure S 11 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that landscapes were composed of 13 sites, instead of 10.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the median. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the median. Blue indicates higher multifunctionality values, orange lower.

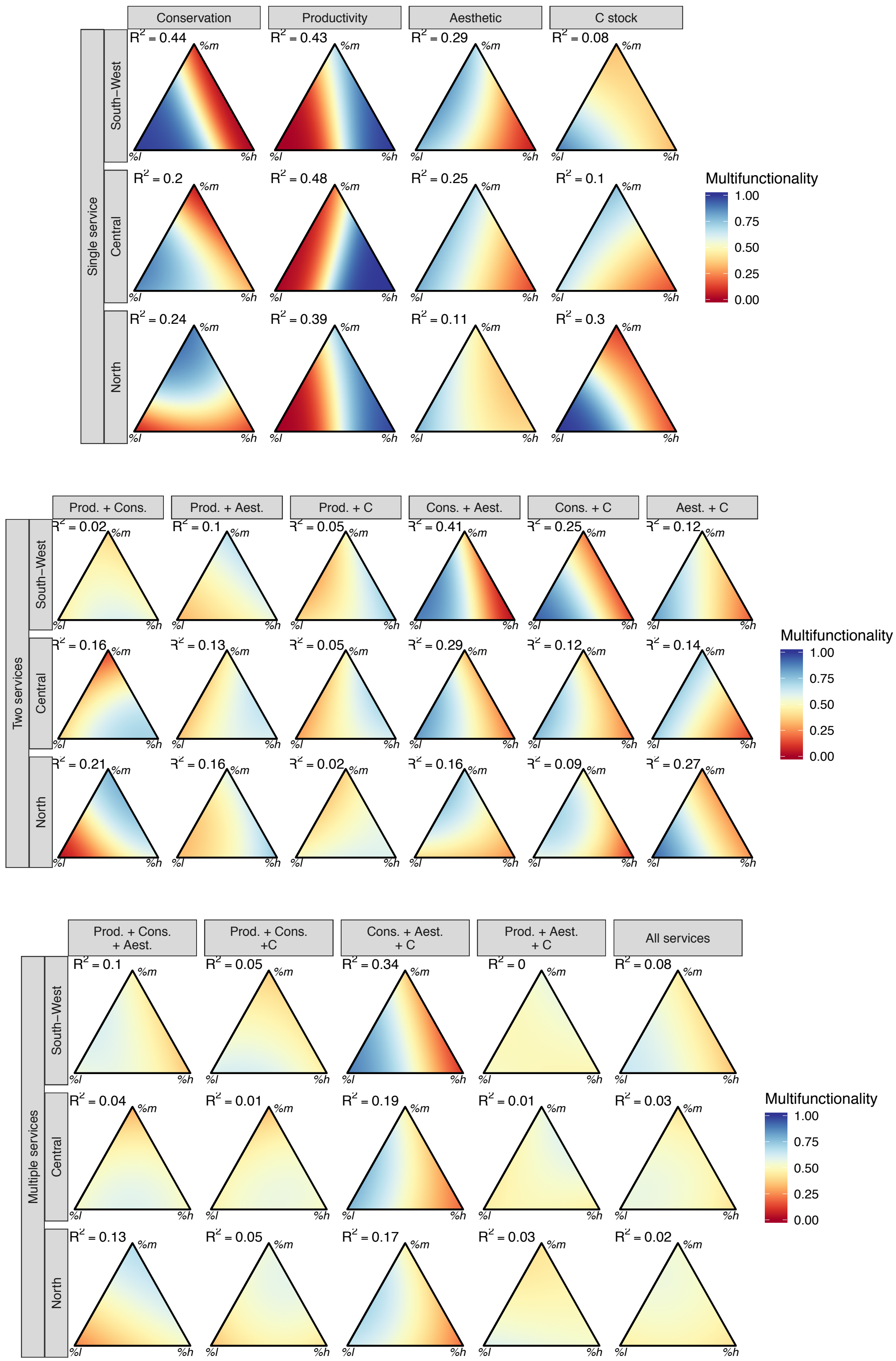
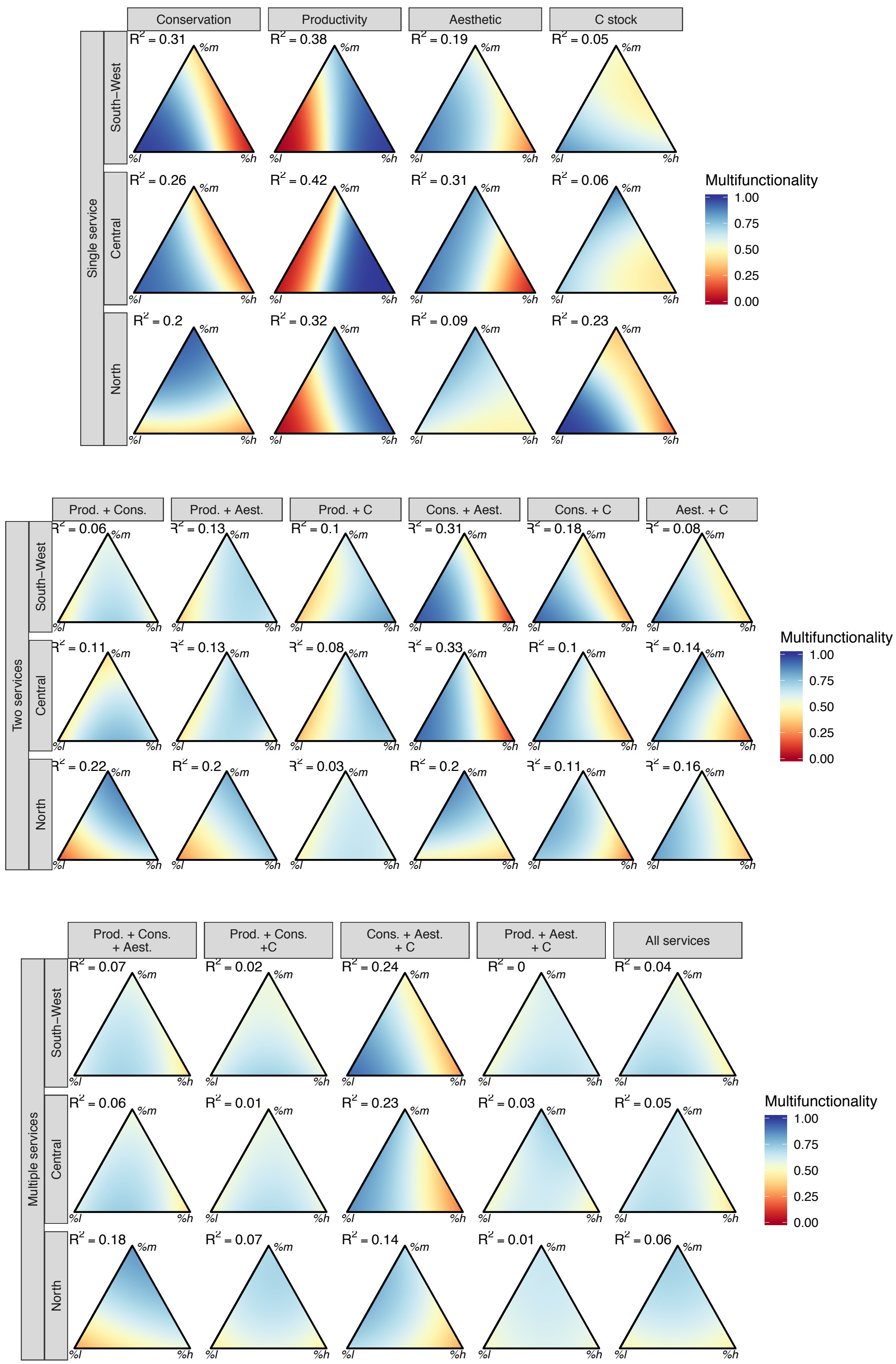


**Figure S 12 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that the threshold was set to the 40th percentile instead of the median.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the threshold. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the threshold. Blue indicates higher multifunctionality values, orange lower.




890

**Figure S 13 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that the threshold was set to the 60th percentile instead of the median.

For single ecosystem services (top row), the value presented corresponds to the probability of the given service to be above the threshold. For combinations of multiple services (middle and bottom rows), multifunctionality is the expected proportion of services above the threshold. Blue indicates higher multifunctionality values, orange lower.

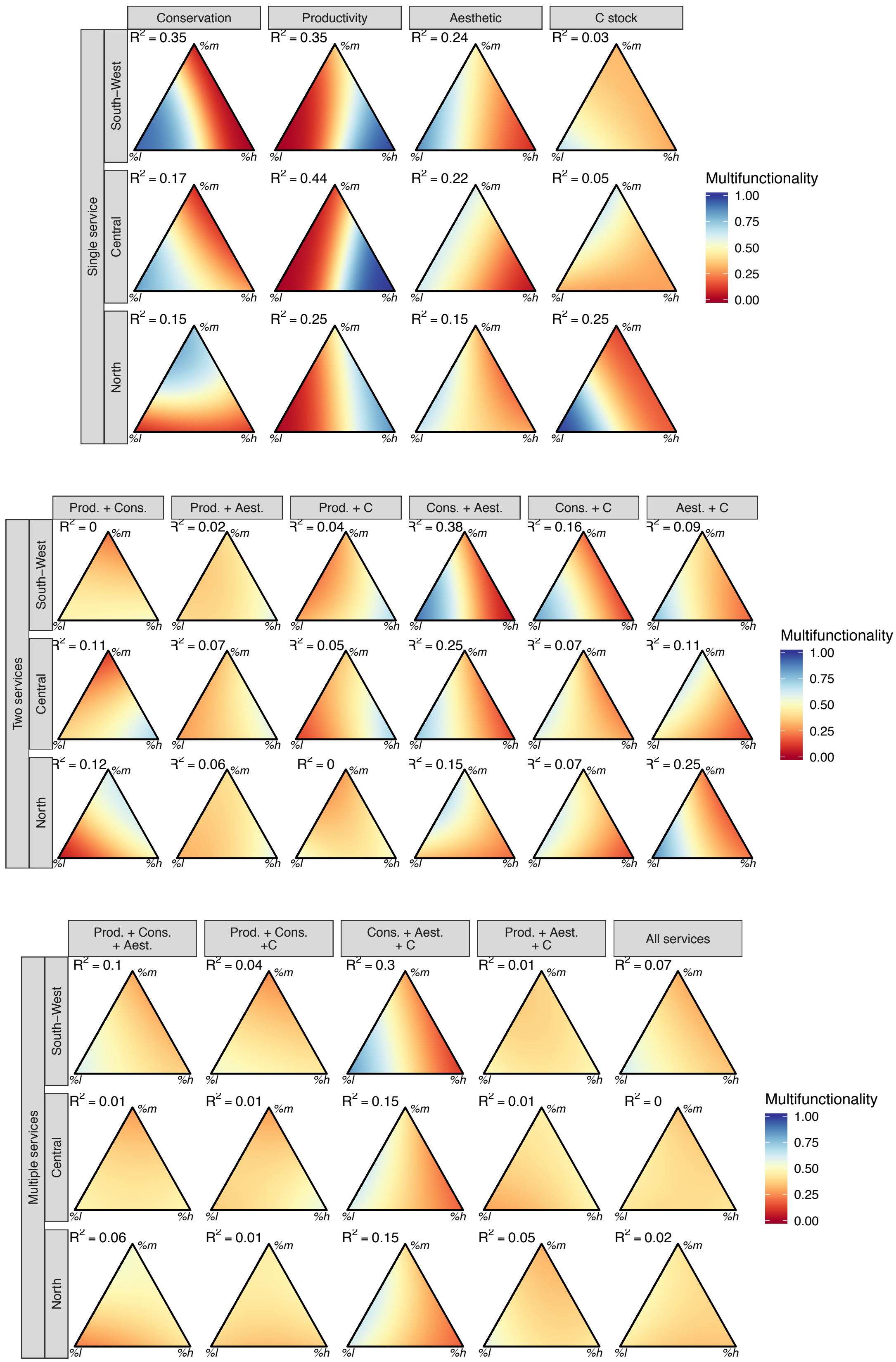
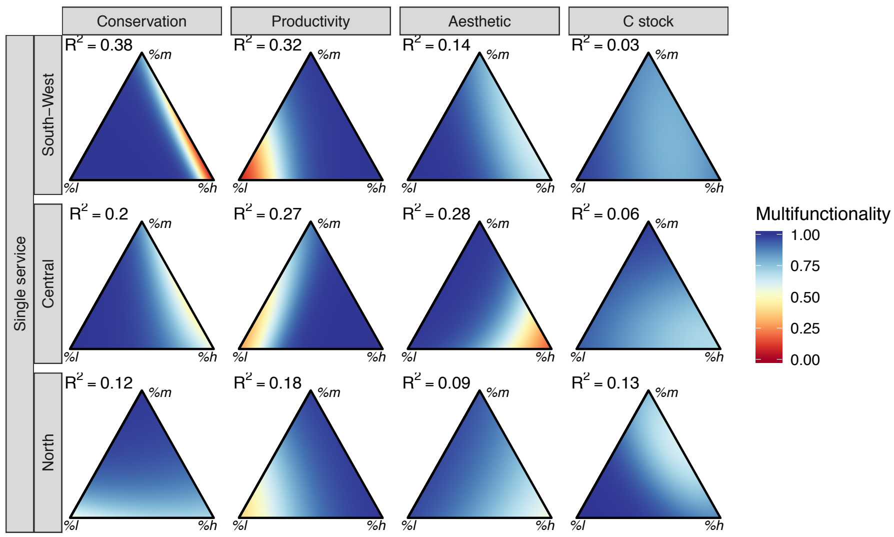
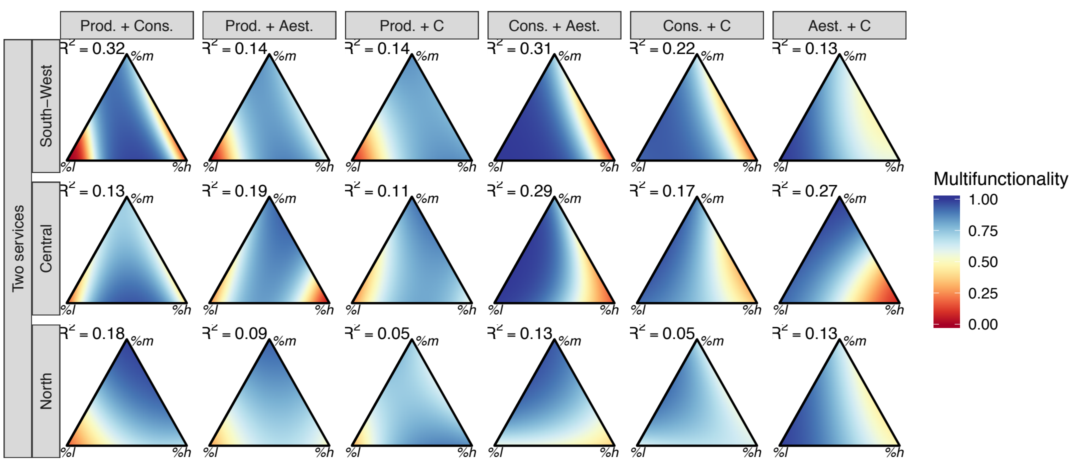


**Figure S 14 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 and Figure 5 that the multifunctionality was calculated as 1 if all the services were above a 15$^{th}$ percentile threshold, and 0 otherwise (instead of a 25$^{th}$ percentile threshold).

The value presented corresponds to the probability that all given services are above the threshold. Blue indicates higher multifunctionality values, orange lower.



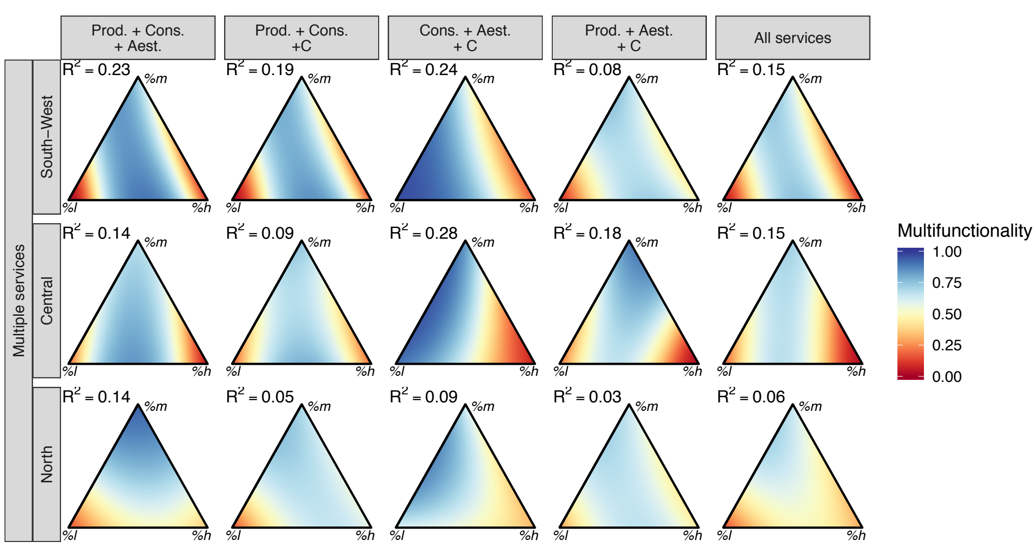



**Figure S 15 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 and Figure 5 that the multifunctionality was calculated as 1 if all the services were above a 35th percentile threshold, and 0 otherwise (instead of a 25th percentile threshold).

The value presented corresponds to the probability that all given services are above the threshold. Blue indicates higher multifunctionality values, orange lower.

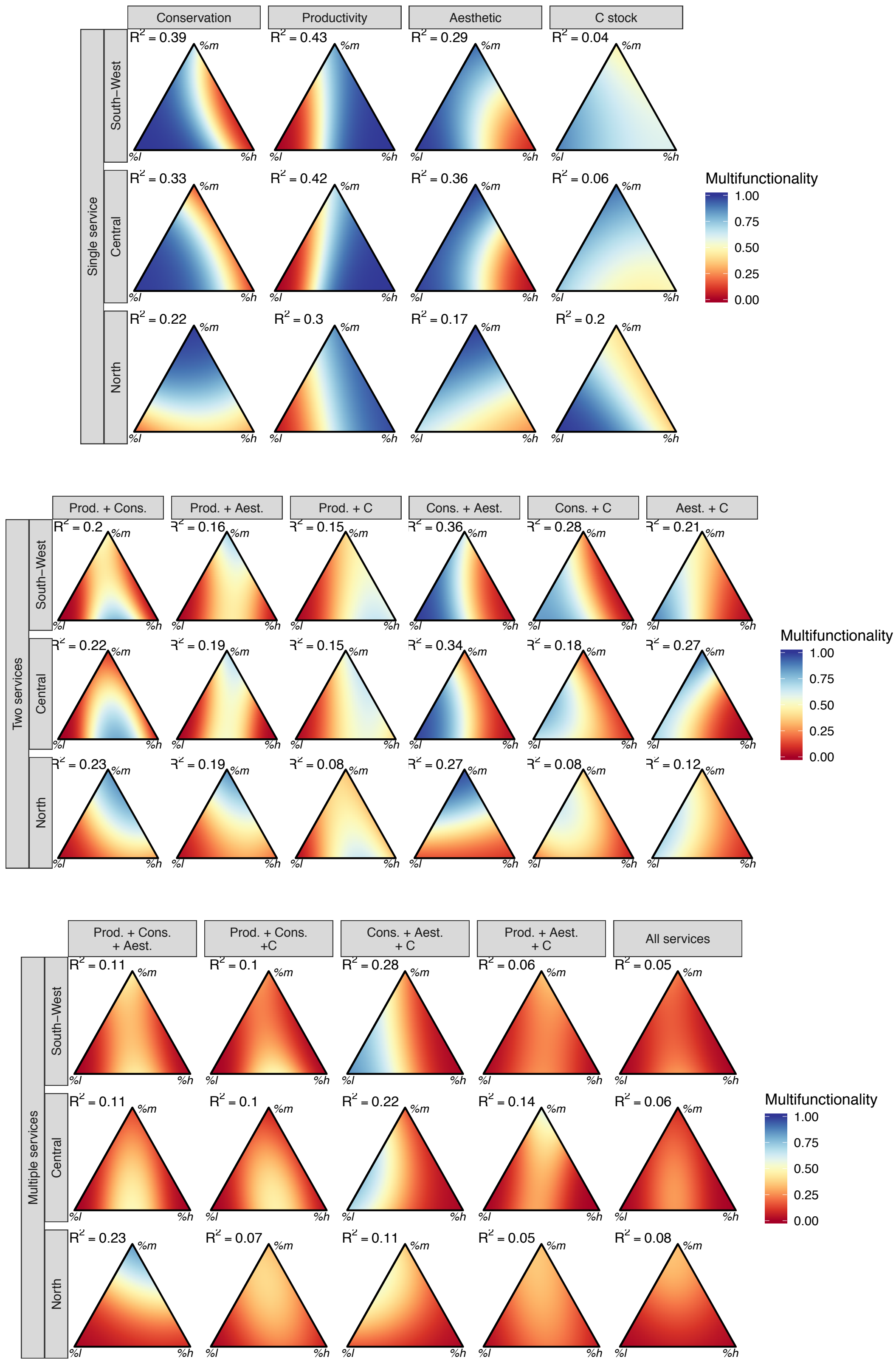


**Figure S 16 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites). T**his figure differs from Figure 4 in that landscapes multifunctionality was calculated as the average of the (scaled) values of all considered services, instead of the number of services above a threshold.

Blue indicates higher multifunctionality values, orange lower.

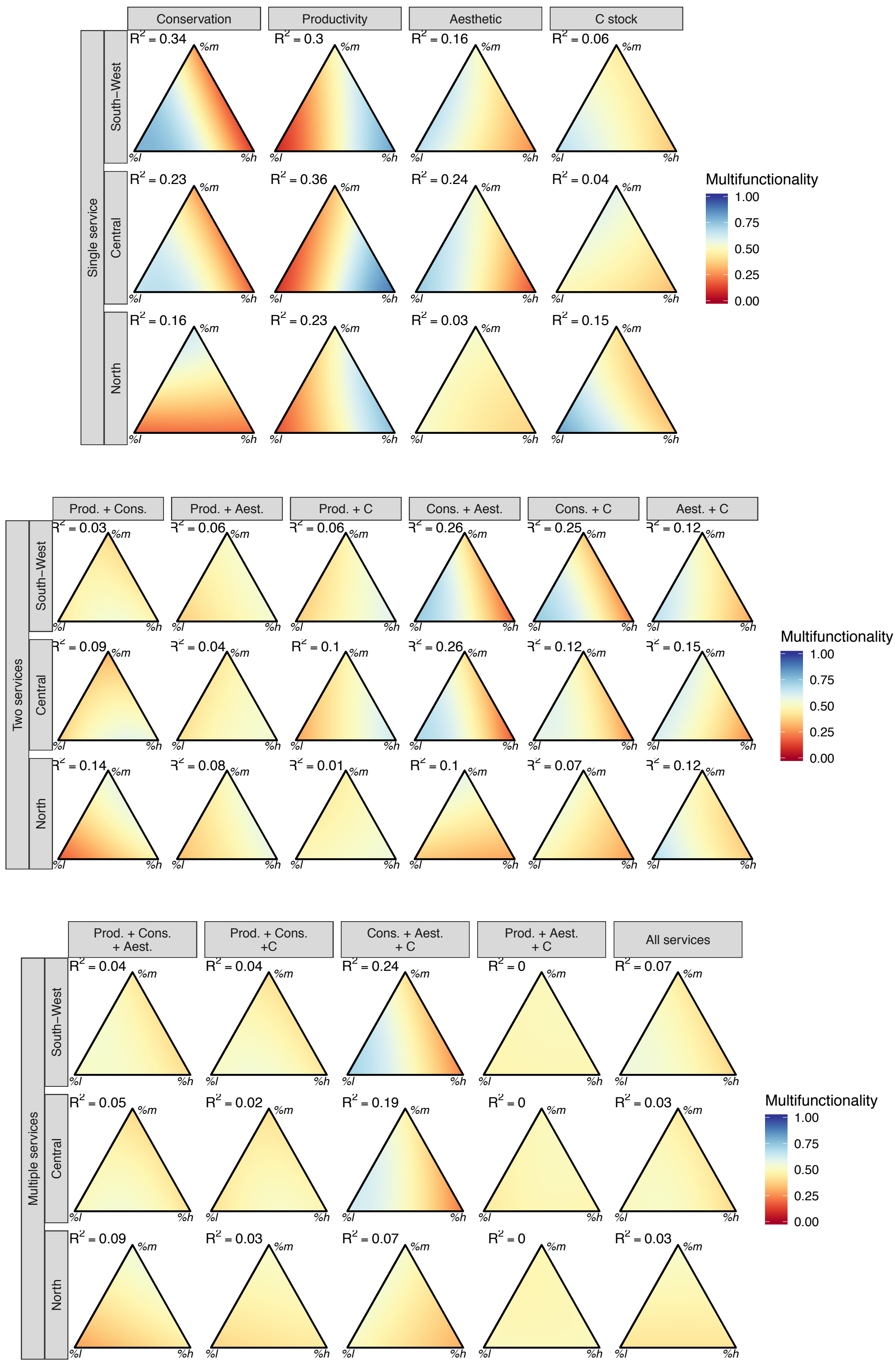


**Figure S 17 Estimated multifunctionality values depending on landscape composition (in proportions of low, medium and high-intensity sites).** This figure differs from Figure 4 in that multifunctionality was calculated as the number of services above a threshold equal to 70% of the maximum (97.5% quantile) observed.

Blue indicates higher multifunctionality values, orange lower.

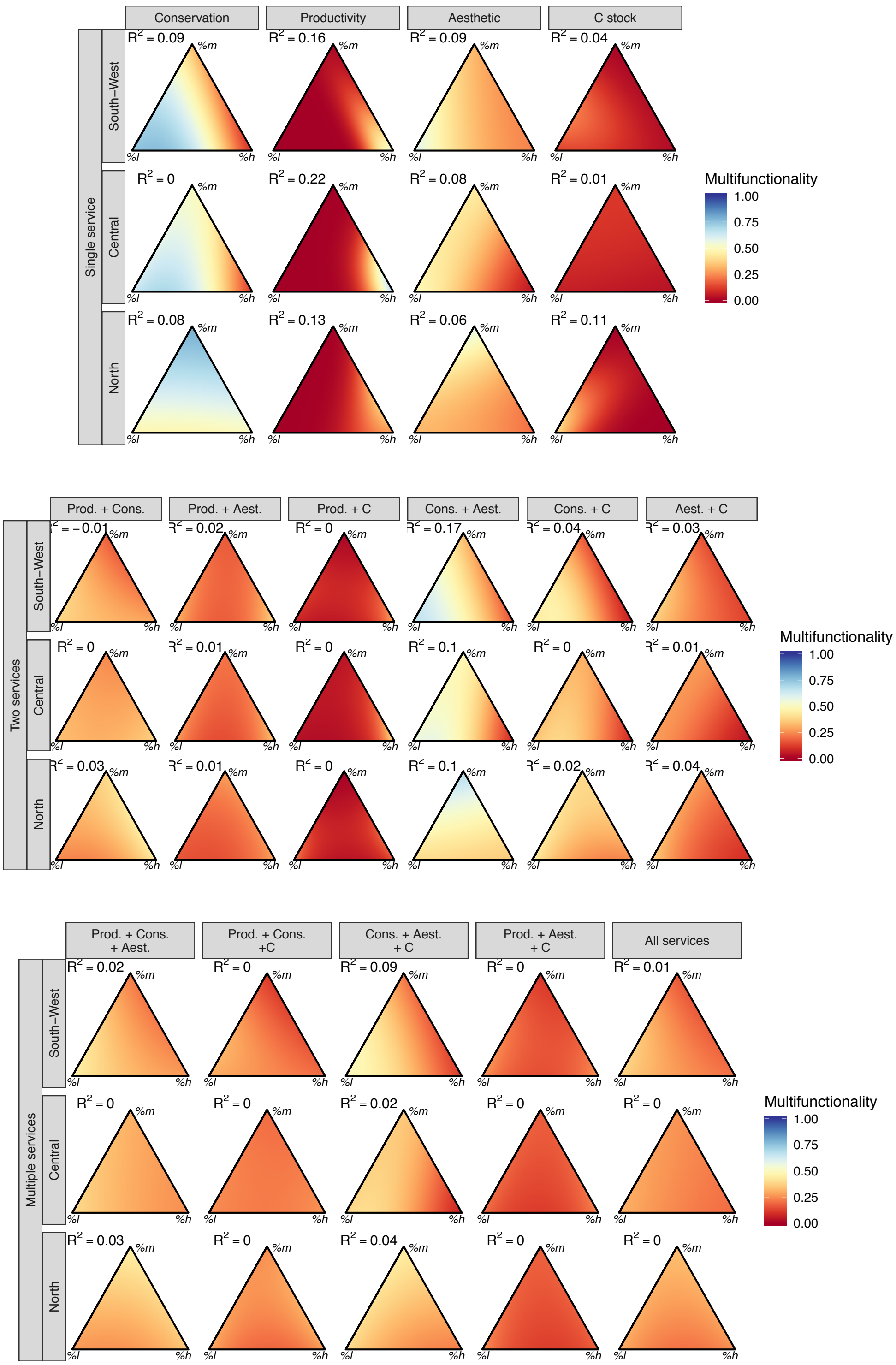


**Figure S 18 Estimated multifunctionality values depending on the mean (x-axis) and coefficient of variation (y-axis) of the land-use intensity in the landscape**. The area outside the coloured represent combinations of intensity mean and variation that were not observed within the region.

Blue indicates higher multifunctionality values, orange lower.

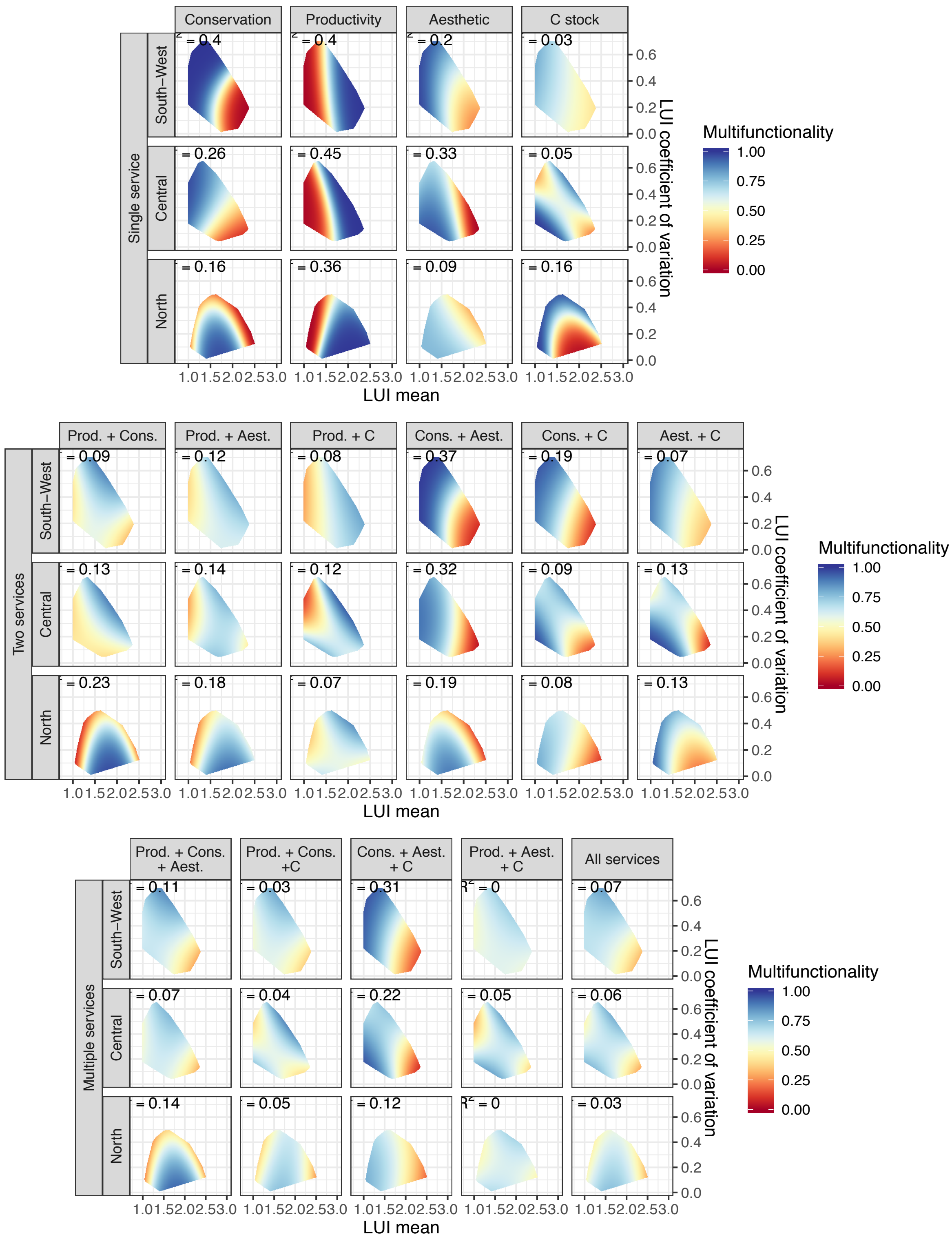


**Table S 5** Variability of multifunctionality responsiveness with the ratio of the effect of intensity and other environmental variables (“Ratio”), the service response variance (“SRV”), and the number of services included in the analysis, depending on omdel parameters. Model results are presented as slope of the relationship ± standard deviation (* P> 0.05, ** P < 0.01, *** P< 0.001). Sensitivity analyses include whether site were classified into low-, medium- and high intensity plots using all intensity values (quantile:30%) or only the lowest, middle and highest fifths (quantile: 20%); the number of sites per landscape (7, 10 or 13); the calculation of services at the landscape scale (either sum or max of the values observed at the site-level). Are also considered the different methodology to mesure multifunctionality (number of services above a given treshold, either based on quantiles of the distribution of percentage of the maiximum observed vales; average of service values; or as 1 if all services are above a treshold, 0 otherwise). (‡) The ratio of the effect of land-use intensity and other environmental covariate is only relevant when the indicators have NOT been previously corrected for the environment.

| Env. correction | Intensity classes | N°of sites | Landscape -scale ES | Multifunctionality | Thres hold | Ratio | SRV | N° of services |
|---|---|---|---|---|---|---|---|---|
| no | 30% quantile | 7 | mean | Threshold based on quantiles | 0,5 | 0,232 ± 0,062** | -0,339 ± 0,171n.s. | -0,139 ± 0,034*** |
| no | 30% quantile | 10 | mean | Threshold based on quantiles | 0,4 | 0,252 ± 0,073** | -0,222 ± 0,163n.s. | -0,119 ± 0,034** |
| no | 30% quantile | 10 | mean | Threshold based on quantiles | 0,5 | 0,217 ± 0,07* | -0,405 ± 0,174* | -0,144 ± 0,035*** |
| no | 30% quantile | 10 | mean | Threshold based on quantiles | 0,6 | 0,197 ± 0,068* | -0,277 ± 0,17n.s. | -0,132 ± 0,033*** |
| no | 30% quantile | 10 | mean | "Compromise" | 0.15 | 0,261 ± 0,04*** | 0,565 ± 0,122*** | 0,068 ± 0,032* |
| no | 30% quantile | 10 | mean | "Compromise" | 0.25 | 0,221 ± 0,049** | 0,195 ± 0,128n.s. | -0,017 ± 0,029n.s. |
| no | 30% quantile | 10 | mean | "Compromise" | 0.35 | 0,188 ± 0,047** | -0,002 ± 0,157n.s. | -0,11 ± 0,028*** |
| no | 30% quantile | 10 | mean | Average | / | 0,192 ± 0,051** | -0,316 ± 0,128* | -0,112 ± 0,028*** |
| no | 30% quantile | 10 | mean | Threshold based on % of maximum | 70% | 0,259 ± 0,091* | -0,087 ± 0,128n.s. | -0,119 ± 0,032*** |
| no | 30% quantile | 10 | max | Threshold based on quantiles | 0,5 | -0,003 ± 0,075n.s. | -0,31 ± 0,143* | -0,124 ± 0,028*** |
| no | 30% quantile | 13 | mean | Threshold based on quantiles | 0,5 | 0,187 ± 0,064* | -0,482 ± 0,174** | -0,144 ± 0,035*** |
| yes | 30% quantile | 7 | mean | Threshold based on quantiles | 0,5 | (‡) | -0,686 ± 0,224** | -0,145 ± 0,03*** |
| yes | 30% quantile | 10 | mean | Threshold based on quantiles | 0,4 | (‡) | -0,623 ± 0,218** | -0,117 ± 0,031*** |
| yes | 30% quantile | 10 | mean | Threshold based on quantiles | 0,5 | (‡) | -0,834 ± 0,239** | -0,141 ± 0,033*** |
| yes | 30% quantile | 10 | mean | Threshold based on quantiles | 0,6 | (‡) | -0,557 ± 0,236* | -0,134 ± 0,03*** |
| yes | 30% quantile | 10 | mean | "Compromise" | 0.15 | (‡) | 0,544 ± 0,153** | 0,067 ± 0,026* |
| yes | 30% quantile | 10 | mean | "Compromise" | 0.25 | (‡) | 0,15 ± 0,169n.s. | -0,025 ± 0,026n.s. |
| yes | 30% quantile | 10 | mean | "Compromise" | 0.35 | (‡) | -0,274 ± 0,231n.s. | -0,096 ± 0,029** |
| yes | 30% quantile | 10 | mean | Average | / | (‡) | -0,439 ± 0,17* | -0,101 ± 0,024*** |
| yes | 30% quantile | 10 | mean | Threshold based on % of maximum | 70% | (‡) | -0,204 ± 0,154n.s. | -0,066 ± 0,02** |
| yes | 30% quantile | 10 | max | Threshold based on quantiles | 0,5 | (‡) | -0,741 ± 0,157*** | -0,111 ± 0,026*** |
| yes | 30% quantile | 13 | mean | Threshold based on quantiles | 0,5 | | -0,688 ± 0,257* | -0,154 ± 0,032*** |
| no | 20% quantile | 7 | mean | Threshold based on quantiles | 0,5 | 0,231 ± 0,033*** | -0,421 ± 0,196* | -0,145 ± 0,037*** |