Meeting European conservation and restoration targets under future land-use demands

Melissa Chapman[1,2,*], Martin Jung[2], David Leclère[3], Carl Boettiger[4], Andrey L. D.Augustynczik[3], Mykola Gusti[3], Leopold Ringwald[3], and Piero Visconti[2]

[1]National Center for Ecological Analysis and Synthesis, Santa Barbara, CA, USA
[2]Biodiversity, Ecology and Conservation Research Group, International Institute for Applied Systems Analysis (IIASA), Vienna, Austria
[3]Integrated Biosphere Futures Research Group, International Institute for Applied Systems Analysis (IIASA), Vienna, Austria
[4]Department of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley, CA, USA
[*]mchapman@nceas.ucsb.edu

ABSTRACT

The European Union is committed to achieving ambitious area-based conservation and restoration targets in the upcoming decade. Yet, there is concern that these targets may conflict with societal needs, particularly food and timber production. Ensuring that competing demands for land are balanced, while the underlying objectives of these targets to mitigate climate change and reduce biodiversity loss are achieved, will require strategic planning across land uses, management measures, and jurisdictional boundaries. Here, we use an integrated spatial planning approach to identify restoration and conservation measures that maximize benefits to species and contributions to climate mitigation, while ensuring projected 2030 crop, pasture, and forestry production demands are met across the EU. We find that pursuing 2030 land restoration targets, when planned in concert with conservation goals and production requirements, could improve the conservation status of 23-42% of species while also increasing terrestrial carbon stocks. Through a series of policy implementation scenarios, we highlight the potential impacts of multilateral cooperation and future land demands on the expected biodiversity and climate mitigation outcomes, equity, and feasibility of 2030 restoration targets in the EU. Our analysis demonstrates how the ambitious targets and objectives of EU Biodiversity policies such as the Nature Restoration law can be met without major impacts to the bioeconomy.

Keywords: biodiversity, climate mitigation, spatial planning, burden sharing, Nature Restoration Law

In an effort to halt biodiversity loss and safeguard nature's contributions to people (NCP)[1], the Convention on Biological Diversity (CBD) recently adopted the Kunming-Montreal post-2020 global biodiversity framework (GBF)[2]. This framework is influencing hundreds of regional, national, and sub-national conservation policies worldwide as governments pledge to improve the protection of their biodiversity assets[3].

In accordance with the global post-2020 biodiversity goals, the European Union (EU) has committed to ambitious land conservation targets. Complemented by the Nature Restoration Law, these decadal policies seek to expand the EU's protected area network to cover at least 30% of land and waters as well as enact a commitment to restore up to 20% of EU area by 2030. Yet, most of the EU territory is under some form of human use, and given anticipated geopolitical and climate impacts on the bioeconomy[4–6], it is crucial that biodiversity objectives be achieved in ways that do not compromise economic needs. Additionally, because biodiversity and carbon are not uniformly distributed across the EU, planning the implementation of these policies to maximize their potential benefits will likely require intergovernmental cooperation and coordination to ensure effective responsibility sharing between Member States (MS) of the EU. These challenges necessitate strategic, integrated planning across land use and management measures while also considering future production needs.

Previous conservation planning studies have sought to address the question of where to allocate new protected areas or restoration efforts independently; while often overlooking the importance of natural, semi-natural and managed ecosystems outside legal protected area designations for achieving biodiversity targets and climate mitigation objectives[7,8]. Both the Global Biodiversity Framework (GBF) and the EU Biodiversity Strategy call for establishing integrated biodiversity-inclusive spatial plans that complement protected areas in addressing land-use change to bring the loss of areas of biodiversity importance close to zero, and support the recovery of species and ecosystems (e.g., Target 1 of the GBF and Green Infrastructure Strategy of the EU, and the cross-cutting principle of 'Do no significant harm" [to the environment] that applies to all policies following the

EU Green Deal). Achieving this goal requires shifting from planning protected areas towards integrated spatial planning across natural and managed landscapes to conserve biodiversity, mitigate climate change, and ensure societal needs are met through sustainable land management, e.g. provision of food, fiber, timber and energy.

Strategically planning the implementation of EU biodiversity policies is particularly important because expanding conservation or restoration efforts at the proposed scale has potentially widespread consequences on the bioeconomy[9,10]. Given the current EU context, such as the war in Ukraine and its consequent increased food and energy prices, hasty reactions and resistance from European policymakers towards increased biodiversity conservation efforts have dominated conversations, despite these concerns being widely criticized from the scientific community[9,10]. However, quantitative evidence on the feasibility of ambitious Restoration Targets such as those in the NRL and the GBF Target 2 is still lacking in the EU, where 38% of the land area is farmed, and another 26% is production forest. Therefore the question remains open: are these international biodiversity policy targets achievable and, if so, at what costs[11]?

Here we assess how the Nature Restoration Law (NRL) and other cross-cutting policies of the EU Biodiversity strategy could be realized across the EU, without sacrificing needs for current or future food and timber production. Moreover, we assess the capacity of different implementations of these policies to meet their underlying biodiversity and carbon sequestration objectives. The overarching target of the NRL is to restore 20% of EU lands, of which, according to European Commission's estimate, up to 15% would come from more biodiversity-friendly management of agro-ecosystems and forests, and the remainder from improving the ecological condition of natural ecosystems and urban areas (Article 9 and 10 of the NRL)[12]. Without EU-wide spatially-explicit information on ecosystem condition, we focus our analyses on restoration priorities in managed ecosystems, thus constraining our restoration priorities not to exceed approximately 15% of EU land area.

Leveraging the best available data on species distributions and land-based carbon sequestration, we identify priority areas for the allocation of conservation, restoration, and production across land-cover types to optimally achieve 2030 biodiversity objectives. We integrate existing and future demand on land for timber and food production across sub-national administrative regions (NUTS2) and production intensity levels. While most previous restoration prioritization studies[8,13] ignored the benefits provided by habitat conservation and the policy commitments associated with them - thus potentially over-allocating restoration efforts to habitats that are well represented in protected area systems and under-allocating them in those that do not- we simultaneously identify priority areas for conservation and restoration that would maximize the achievement of species conservation and carbon sequestration targets. This methodological advancement allows us to explore the trade-offs between food and timber security, climate mitigation and biodiversity conservation, while assessing the feasibility of alternative realization of European restoration plans. While we do not focus on legal land protection priorities and targets, we optimize the allocation of restoration targets in the context of natural habitat conservation.

While the NRL outlines country-level outcome targets (i.e., species benefits and carbon sequestration) for each Member State, it does not necessarily require equal-area restoration efforts due to uneven land area with the potential for restoration and heterogeneous carbon and biodiversity benefits per unit area restored. This suggests that while at the EU level, it is expected that approximately 15% of land will be farmland or managed forest under some form of more biodiversity-friendly practice by 2030, the fraction of land under restoration measure could vary significantly among countries. This poses two questions: 1) what are the equity implications (unevenness of burden share) of targeting restoration plans at the EU scale to maximize conservation status improvements of species? 2) what are the efficiency losses in terms of carbon and biodiversity recovery if the restoration effort is equally apportioned to each country?

To answer these questions, we simulate three scenarios of area-based burden sharing under two future production land-use scenarios. Future land use scenarios are based on "Business As Usual" (BAU) land-use policies and the "Fit for 55" EU policy package which includes land-based mitigation measures to increase the land-use sector carbon sink as a means to reduce overall net emissions by 55% compared to 1990 (see methods for additional details). Further, we assess two levels of restoration ambition, each aiming to restore 15% of EU land area: first, a 'high nature' ambition allowing up to 6% of expansion of natural ecosystems through active or passive restoration, and the remainder of the target met through deintensification of production landscapes, and a second a 'baseline' ambition exclusively allowing deintensification of production landscapes and $<1\%$ wetland restoration. Through 120 different scenarios, we assess (i) the feasibility of restoration targets against competing demand for land for agriculture and forestry under two futures of production requirements (BAU and "Fit for 55"), (ii) the potential implications of multilateral burden-sharing on the coherence and efficiency of possible realizations of the EU's conservation and restoration targets, and (iii) the synergies and tradeoffs across species conservation and climate change mitigation policy objectives.

Results

Feasibility of restoration targets aided by integrated spatial planning

Planning the distribution of restoration, conservation, and production in concert (Figure 1B, Figure S2), allows us to assess the feasibility of European restoration measures and assess areas of highest priority given other land requirements. Our results show

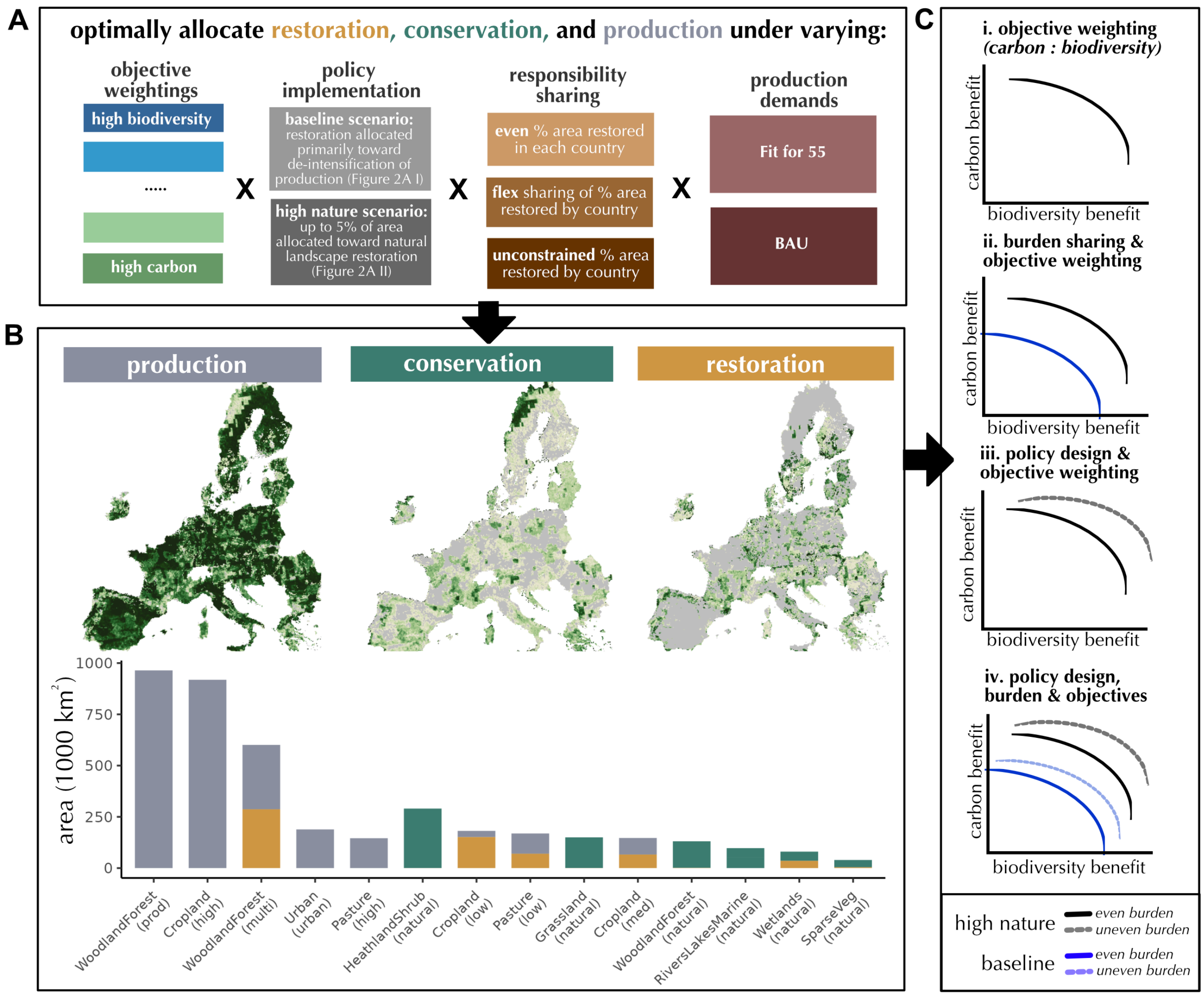

Figure 1. A European-wide implementation of the Nature Restoration Law that achieves restoration targets without sacrificing production. (A) We allocate conservation, restoration, and production measures to maximize the extent to which targets for species habitat and carbon sequestration are met through optimal zoning across land-use types. This is done for each of 120 scenarios, with different weightings of objectives, policy implementation, and responsibility sharing (see methods, Table S1 and Figure S1 for additional details). (B) Represents a single solution scenario where the carbon targets are weighted as equal to the sum of all species targets, restoration policy targets are met at the member state scale under fit for 55 production demands and baseline implementation scenario (see methods for additional details). The three maps in (B) are further broken down by land cover class and/or management intensity spatially, combined accounting for 100% of the landscape. The full spatial solution can be seen in SI Figure S2. Note that "conservation" here only includes unmanaged natural landscapes (SI figure 1). Low intensity production (e.g., multi functional forests) may, in many cases, overlap with notions of conservation and/or priorities for legal protection not addressed here. (C) Conceptual representation of comparing the outcomes of each of the 120 scenarios along objective axes (Table S1). Each scenario from A is solved (resulting in a distinct spatial solution, represented in B and Figure S2) and assessed for contributions to improving species conservation status and increasing land-based climate mitigation. We compare these solutions and their synergies and tradeoffs, providing insight into the interacting effects of objective weighting, responsibility sharing, and policy design on outcomes (i-iv, Figures 2-3), feasibility (Table 1), and multilateral equity (Figure 3).

burden sharing	restoration scenario	production constraints	% total area restored	% deintensification	% nature	% rewetting
Even	Baseline	Fit for 55	13.7	12.7	0.1	0.9
Even	Baseline	BAU	14.0	13.0	0.1	0.9
Even	HN	Fit for 55	15.1	12.1	2.1	0.9
Even	HN	BAU	15.1	12.1	2.2	0.9
Flexible	Baseline	Fit for 55	14.2	13.2	0.1	0.9
Flexible	Baseline	BAU	15.0	14.0	0.1	0.9
Flexible	HN	Fit for 55	15.1	12.0	2.2	0.9
Flexible	HN	BAU	15.1	12.0	2.2	0.9
Unconstrained	Baseline	Fit for 55	14.3	13.3	0.1	0.9
Unconstrained	Baseline	BAU	15.1	14.1	0.1	0.9
Unconstrained	HN	Fit for 55	15.1	12.0	2.2	0.9
Unconstrained	HN	BAU	15.1	12.0	2.2	0.9

Table 1. Restoration feasibility in each scenario while ensuring sufficient future cropland, pasture, and forestry production (SI Table 1). Imposing even burden sharing in restoration efforts between countries reduces the capacity of area-based targets to be met, particularly for deintensification of agriculture. Allowing flexible sharing of restoration burden (MS allowed up to 25% priority area) increases the feasible extent of restoration by up to 1.4% of EU land area in baseline implementation scenarios.

that the majority of area based restoration targets as envisaged by the European Nature Restoration Law can be achieved while improving biodiversity and carbon outcomes and maintaining a functioning bioeconomy regardless of different implementation and constraint scenarios (Table 1). Overall, our analyses suggest that ambitious restoration targets (up to 15.1% of EU land area under some form of restoration actions) can be met without jeopardizing future food and timber production areas. The extent of managed ecosystems and natural ecosystems that are under some form of restoration increases up to 1.4% in absence of climate policies and with a slightly uneven burden share among countries (Table 1).

It is noteworthy that jointly planning for all major land management options, compared to allocating restoration priorities alone (hereby "naive" restoration prioritization), thus ignoring the demand on land for production as well priority areas for habitat conservation and the associated carbon and biodiversity benefits, implies significant differences in restoration priority, both spatially and across management measures (i.e., "what" is prioritized to be restored) (Figure S3). For example, naive prioritization of restoration measures and joint prioritization (Figure S1) only have a 19% spatial overlap in their restoration priority areas (Figure S3). Because the joint prioritization framework integrates information about the entire decision space (including maintained natural landscapes and production landscapes) within which restoration decisions are made, it is difficult to directly compare outcomes on species and carbon sequestration objectives. However, integrated planning presents a more holistic and informed perspective for prioritizing both restoration and conservation.

Implementing restoration targets between nature and production

High nature restoration scenarios, where up to 6% of EU land area could be placed under restoration measures aimed at re-creating natural ecosystems, enhance biodiversity outcomes significantly, improving the conservation status of up to 42% of species with an insufficient area of habitat within some part of their range (hereby area shortfall), compared to 26% in equivalent baseline restoration scenarios where except for re-wetting, restoration focuses exclusively on reducing management intensity in production landscapes (Figure 2A). The climate mitigation differences between these two scenarios are more nuanced: carbon sequestration in the "high nature" scenario increases only slightly (1-3%) compared to equivalent baseline scenarios (Figure S3).

Under the high nature restoration scenario meeting European land-use demands by 2030 constrained the maximum potential restoration of natural landscapes (excluding wetlands) to 2.2% of the EU land area plus an additional 0.9% of rewetting (Table 1). Natural landscape restoration priority and feasibility is not evenly distributed across the EU, with Hungary having the largest

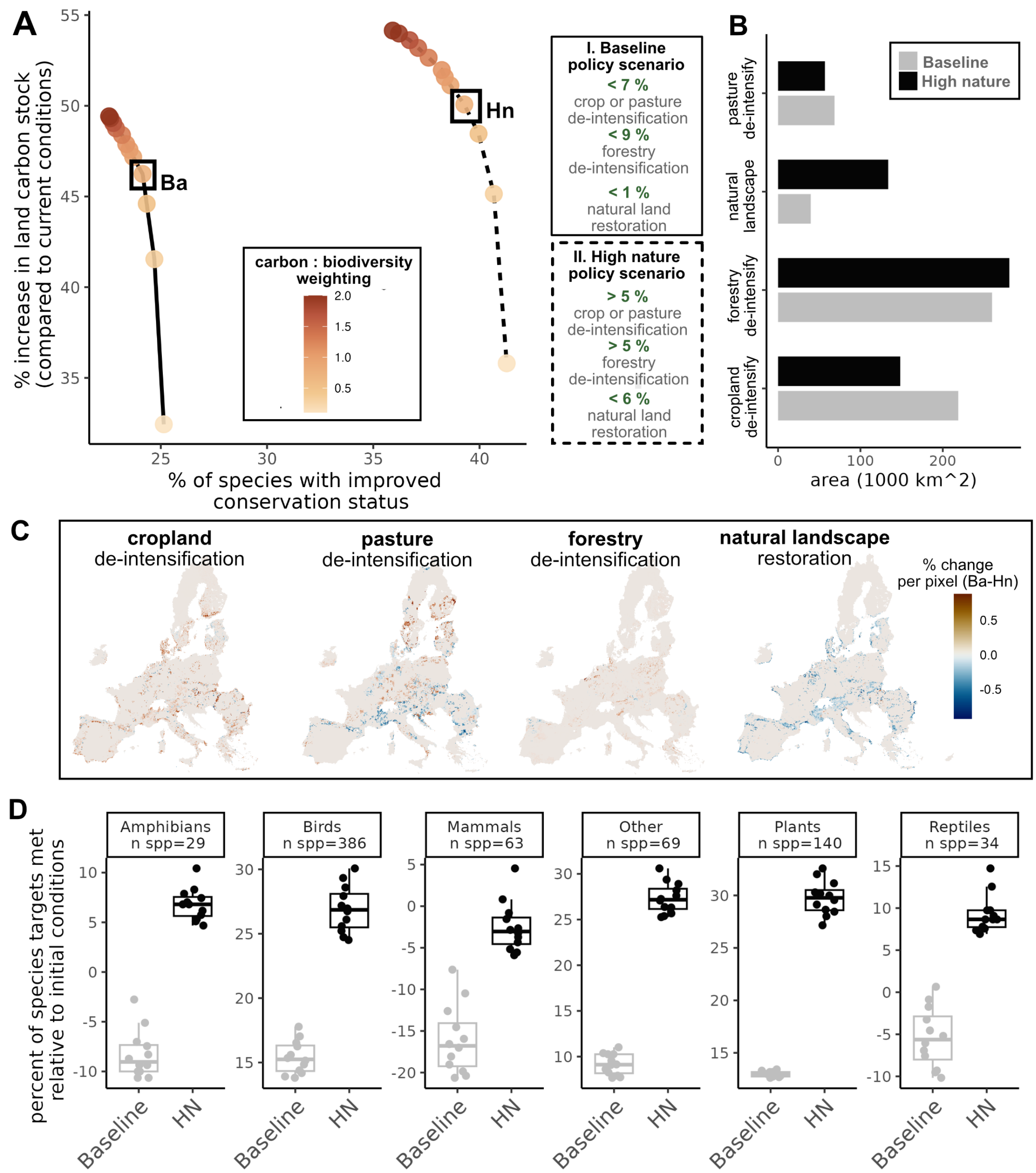

Figure 2. Impact of the allocation of restoration targets between deintensification and natural landscape restoration. We assess two scenarios of implementation of restoration targets, assuming flexible area implementation of targets in each member state (see methods and 3 for additional details on burden sharing scenarios) and "Fit for 55" production constraints. (A) Shows that a "High Nature" (Scenario II) implementation of restoration targets can improve the conservation status of 15% more species compared to "Baseline" (Scenario I) implementation, suggesting that small amounts of natural landscape restoration can have large impacts on species outcomes. (B) Shows the difference in land area of different restoration typologies (see methods and Figure S1 for full list of restoration zones) (Ba and Hn scenario) (C) Displays the spatial differences in restoration priorities across the 4 typologies of restoration. (D) Shows how biodiversity benefits break down across species groups (includes all scenarios of carbon to biodiversity weighting under fit for 55 scenarios of future production). While all species are impacted significantly by changes in restoration scenario, some species such as amphibians, reptiles, and mammals, which might otherwise be negatively affected by production demands over the coming decade, are able to have positive outcomes in the high nature solutions.

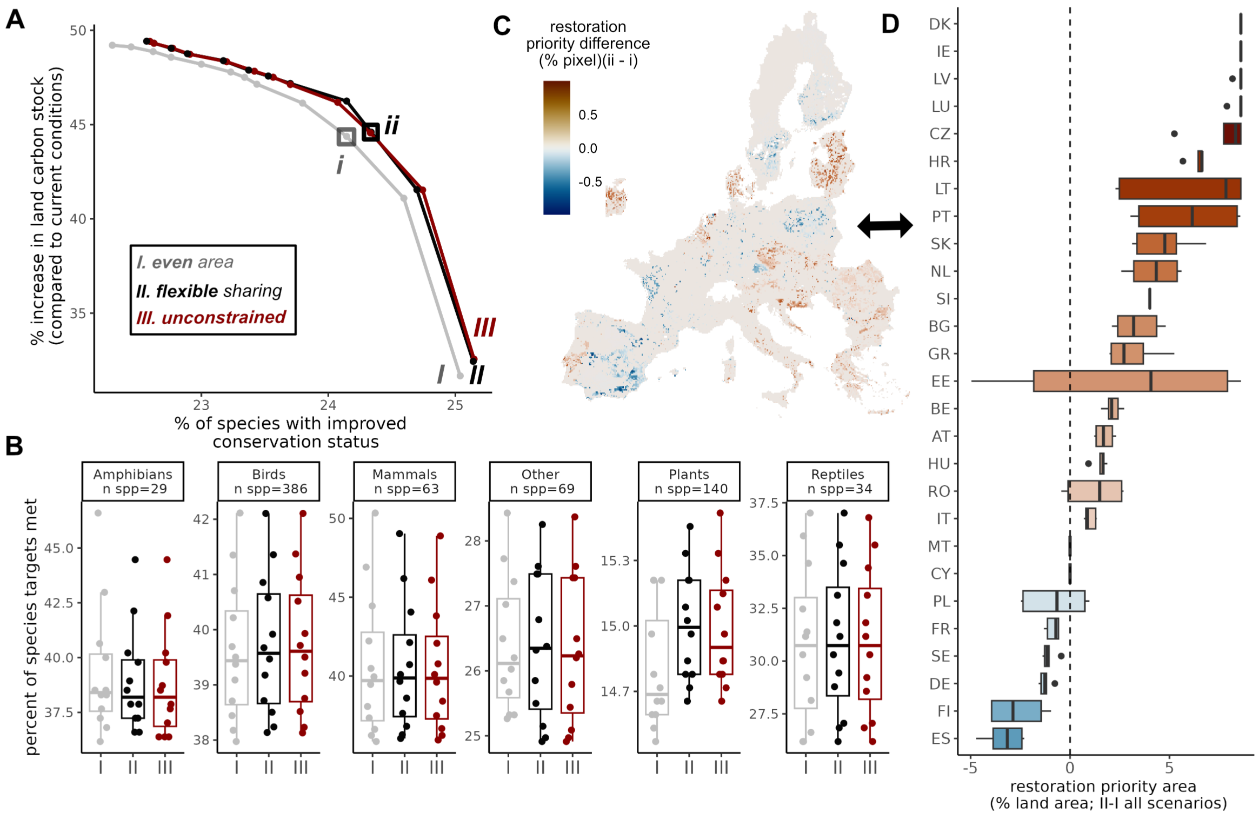

Figure 3. Impact of international cooperation on restoration outcomes. Using the baseline restoration, fit for 55 production scenarios, we assess the implications of three different scenarios of burden sharing: (I) "Even" burden sharing, requiring even distribution of area based restoration targets between countries (each country restores up to the target % area of their jurisdiction for each restoration category) (II) "Flexible" burden sharing, where countries can have up to 10% additional restoration priority within their jurisdiction but the overall EU target remains constant. And (III) and "Unconstrained" version where restoration can occur anywhere in the EU and is not constrained by jurisdictional boundaries. (A) Unconstrained prioritization (III) improves carbon and species (B) outcomes relative to even burden-sharing between countries. However, these differences are minor and insignificant relative to the differences in outcomes resulting from carbon weightings and restoration scenario implementation (Figure 2) (C) We can observe how burden equity shifts spatial priorities (solutions i-ii) and (D) how those spatial shifts aggregate to change the multilateral equity of solutions regardless of carbon-biodiversity objective weighting (lines II-I, all carbon to biodiversity weightings from A).

priority for increased restoration efforts under high nature scenarios with 11.1% of the country under some form of restoration, while other MS, such as Sweden, having only small margin for natural land restoration (0.6% of their extent) (Figure 2C). It is noteworthy that for some species groups (amphibians, mammals, and reptiles) meeting projected production demands over the next decade without investment in some amount of natural landscape restoration might result in continued habitat losses for species of conservation concern (Figure 2D). For example, we expect a 5-11% increase in unmet species targets for reptiles included in this analysis under future scenarios of production demands in absence of restoration of natural habitat as simulated in the Baseline restoration scenarios (Figure 2D). Ambitious and strategic placement of restoration areas is therefore not only supporting biodiversity recovery, but also likely essential to avoid losses.

The impact of international cooperation and burden sharing

Plant and animal species and carbon stocks are not uniformly distributed, implying an unequal geographic responsibility for their conservation and recovery across EU MS. The EU Nature Restoration Law proposal does not prescribe a specific maximum burden share to be shouldered by each country. Therefore, a plausible simulation of the NRL is that of an 'Unconstrained scenario' in which conservation and restoration areas can be placed without limits to the amount of restoration shouldered by any country. This gives, by design, more freedom to optimize their placement where the carbon and biodiversity gains are highest and results in higher performance (lower area of habitat and carbon stocks shortfalls) than a scenario where each country shoulders the same responsibility ('Even scenario') where each country can have up to 15% of their land restored (Figure 3A). Effective European cooperation ('Unconstrained scenario') between MS could increase the conservation status of 4-9% more species and store 1-4% more carbon (Figure 3A) compared to equal area implementation of targets.

Our results suggest that some amount of flexibility in restoration efforts across countries ("flexible" scenarios) may realize most of the biodiversity and carbon benefits of unconstrained cooperation (Figure 3A) and be necessary to keep the high ambitions of NRL while maintaining production needs (Table 1). This provides extra capacity to prioritize restoration areas that least collide with timber and food production needs, compared to an implementation scenario where each country must restore approximately 15% of their land (Table 1).

We find no clear "winners" (species groups that benefit most) or "losers" (species groups that benefit least) of different planning and burden sharing scenarios (Figure 3B). However, shifting the distribution of responsibility to reach restoration targets does affect the equity of spatial priorities of restoration (figure 3D). Large countries with extensive area of land under management such as Germany, Spain, Finland and Sweden require a lower contribution towards achieving EU restoration target under a flexible approach (uneven but bounded to maximum 25% of the country), than under an even burden-sharing approach (figure 3D).

Despite the benefits of international cooperation, unconstrained restoration targets result in some countries bearing a more substantial burden for meeting EU restoration targets (up to 26% of land prioritized for restoration in a single country, as opposed to a maximum of 16% area per country in the even burden sharing scenarios) (figure 3D). Our results emphasize the potential impact of different restoration burdens but further exploring the social, financial, and equity implications of these different scenarios will be critical for assessing their political feasibility and fairness.

Discussion

Assessing a suite of plausible assumptions about the EU Nature Restoration Law implementations, our analysis provides insight into the optimal allocation of restoration and conservation (of natural ecosystems), while maintaining sufficient production area across land-cover types in the context of current and proposed EU policies (Figure 1, Figure SI 2). We specifically assess the interactions between the NRL targets and future production demands under both a Business as usual future and the Fit for 55 package the EU established to contribute to the Climate Targets of the Paris Agreement (Figure 2, Figure SI 5). We show that 15.1% of European land can be feasibly placed under restoration, and 15.2% allocated to conservation of unmanaged natural habitats without constraining current or future food and timber production under any considered land-use and climate policy scenario (Figure 1B).

Regardless of the relative weight placed on reaching biodiversity or climate policy objectives, all scenarios contribute to an increase in total carbon stocks and improvement in conservation status >22% of the species assessed in the annexes of the Habitats Directive and Birds Directive. This strongly supports the win-win expectations in the Nature Restoration Law. All our scenarios put some weight on meeting biodiversity outcomes and constrain restoration measures to actions which are expected to improve the ecological conditions of managed agroecosystems and forests. However, other land-use management aimed at increasing carbon stocks or reducing reliance on fossil fuels (e.g. monoculture tree planting, or bio-energy crops) would result in significant trade-offs[14] that are not addressed here. Nevertheless, our results show that there is a wide range of biodiversity and carbon gains to be expected from implementation of the NRL, depending on the specific emphasis given to either environmental goal. Because a large win-win across objectives should not always be assumed, we recommend managers and policy-makers concerned with designing local restoration efforts use the best available data and models appropriate for their context to explore the restoration options available and their expected returns for biodiversity, carbon, and bio-economy (Figure 2).

Aligning decadal conservation and restoration targets and actions with the multiple ambitions of the EU Green Deal will require an integrated perspective and coordinated spatial planning effort. Unlike previous approaches to spatial planning of conservation and restoration measures - which have to-date been implemented separately[7,8,15] - our analysis jointly optimizes the allocation of multiple area-based measures (i.e., conservation and restoration) to mitigate biodiversity loss and draw down carbon together with production land use contributing to food, feed and energy use (including for climate mitigation purposes). Importantly, we show that strategic conservation planning allied with natural landscape restoration enables meeting future demands for food and fibers without compromising the habitat of species of concern (Figure 2B).

We found only a small efficacy loss in both biodiversity and carbon benefits of the NRL when redistributing efforts equitably between EU MS (Figure 3). Here our results demonstrate that European policies aiming at conservation and restoration can be successfully implemented in various possible configurations (Table 1, Figure S4), highlighting both the option space towards future inclusive planning, and the need for adequately assessing trade-offs with expansions of the European conservation network. Sharing the burden of reducing greenhouse gas emissions, conserving biodiversity, and promoting sustainable practices requires an assessment of countries' respective capabilities and contributions. Because our analysis optimizes biodiversity and land-based climate mitigation portfolios across the entire EU, even in scenarios where restoration targets are constrained at the MS scales (Figure 3A; scenario I), the modeling assumes high-level coordination across MS. Our analysis, therefore, provides insights on how MS implementations of the Nature Restoration Law can be strategized to maximize benefits in the context of continental biodiversity portfolios and climate mitigation opportunities, regardless of the precise amount of burden sharing deemed acceptable between countries. There are social, political, and financial implications of multilateral burden sharing. However, EU legal frameworks in other sectors have navigated burden sharing through principles such as the "polluter pay principle", "principle of subsidiarity", or "capacity to pay principle", enabling translation of EU targets to country-level targets for meeting, for example, the Emission sharing regulation. A more thorough exploration of the unequal distribution of current degradation and potential benefits might help inform metrics of fairness in proposed solutions and scenarios.

Investments in habitat restoration have been shown to generate several direct and indirect potential benefits for biodiversity[8], people[16] and economies[17,18]. However, restoration, as postulated by the Nature Restoration Law, could also have several trade-offs with the bioeconomy as natural landscapes are restored or de-intensified. Previous studies have shown that any expansions of nature conservation have displacement effects on existing and future land uses[10,19] or can create off-site impacts in surrounding landscapes[20]. Our planning model attempts to minimize these tradeoffs through integrated planning. In this work, we directly account for future production constraints in European working landscapes (Figure 2B), which can help to identify areas with higher potential for implementation. However, it should be noted that our assessment of restoration critically ignores any opportunity costs of establishing restoration, climate impacts on biodiversity (e.g. ecosystems resilience, species range shifts) as well as impacts and feedback on land-use (e.g. changes in crop yields, changes in fire regimes, forest and agricultural pests). As a result, the actual biodiversity and climate mitigation benefit of restoration might be lower than estimated. Future studies should aim to explore the consequences of coupled climate and land-use impacts on biodiversity and LULUCF to guide conservation and restoration efforts[21].

While our analysis provides restoration priorities that attempt to align with area-based targets outlined by the EU Nature Restoration Law, we do not claim or propose any areas for legal or strict protection as outlined by the EU 2030 Biodiversity Strategy (30% of land area in protected areas and 10% under strict protection). Such decisions need to be taken following in-depth and bottom-up investigations, also considering factors that were not available for inclusion in this work and at this scale (for example, land tenure or management costs). Rather, our solutions reflect the feasibility of allocating conservation and restoration efforts to a given area with legal designations of protection only being a subset of the areas identified as meaningfully contributing to biodiversity and climate mitigation targets, including, for example, areas outside of traditional protected areas ("Other Effective Conservation Measures" (OECMs)[22]). Further work integrating compensation schemes in the light of the recently published guidelines by the European Commission on the "Development of Public and Private Payment Schemes for Forest Ecosystem Service" might help realize the implementation of restoration in high-priority areas. Even still, some of the conservation, restoration, and production (e.g., multi-functional forests) priority areas identified in our solutions would likely benefit from protected area designation pending further investigation under EU Member state laws and national criteria of designation (high biodiversity value, economic pressure, high risk of low).

In conclusion, this study explores the potential of restoration targets in the EU to contribute to land management approaches to reducing biodiversity loss and mitigating climate change. Through a suite of scenarios, we highlight the implications of multilateral burden sharing and policy implementation on expected biodiversity and climate mitigation outcomes of decadal restoration and conservation targets across the EU while also assessing the feasibility of these policies. Our results provide insight into the strategic implementation of the EU restoration law and 2030 biodiversity policy commitments while creating a baseline for EU MS to evaluate the potential of their conservation and restoration pledges. Our methods are broadly relevant to contexts worldwide and showcase how integrated spatial planning can meaningfully optimize land allocation in a way that benefits both nature and people.

Methods

In this work, we aimed to assess whether the goals of the EU Biodiversity Strategy and Nature Restoration Law could still be reached given future food production constraints. We leverage data on the current and potential distributions of (1) species, (2) land-based carbon sequestration, and (3) land cover, to identify priority areas for the allocation of conservation, restoration, and production across land-cover types to optimally achieve 2030 biodiversity targets in the EU (Figure S1). Note that conservation here is distinct from legal protection, instead indicating natural land not allocated to production, urban area, or restoration. Importantly, this conceptualization of conservation allows us to capture the habitat contributions of all natural land in a given solution to optimize the biodiversity benefits of restoration allocation.

Data

Biodiversity data

We estimated the current distribution of 721 species included in the EU habitats directive using an integrated species distribution modeling approach (iSDM,[23]) where best-available data sources (occurrence, preference, expert information) are harmonized into one joint prediction using different types of linear and non-linear models (full methodology for current SDMs in SI). We also leverage species distribution models to depict the potential distribution of the same set of species (sensu[24]), understanding potential in this context as the contemporary climatic, soil and natural vegetation conditions that would allow a species to persist in an area (full methodology for potential SDMs available in SI). Critically, and opposed to mapping current suitable habitat, the potential modeling approach considers only contemporary differences in climate and soil, and not any land-cover or land-use (aligning with the concept of the potential natural vegetation of Europe[25]) although we acknowledge that both climate and soil can to some extent also be shaped by land-use practices. The predictions from the species distribution models used here thus aim to depict where a species might exist under current conditions, while also allowing modest inter- and extrapolation from its current distribution to assess the potential added habitat benefits of ecosystem restoration.

Land cover and management intensity data

Dynamics of the LULUCF sector are modelled with the GLOBIOM/G4M framework[26,27], where the partial equilibrium model GLOBIOM represents land-use change in cropland, managed grassland, bioenergy production as well as other natural vegetation and the biophysical and economic model G4M covers the forestry sector. In combination, these models allow future projections in 10-year time steps from 2000 onwards of scenario-specific changes in land-use and correspondingly demand, production, trade and prices for agricultural and forest products as well as greenhouse gas emissions from the LULUCF sector at national level in the EU. Initial conditions for the prioritization were aligned with GLOBIOM/G4M circa 2020. To generate the initial distribution of both protected and unprotected land-cover within each planning unit, we harmonize data from GLOBIOM (at NUTS2 level circa 2020), G4M (at 30 arcmin level circa 2020,) and the Corine Landcover Accounting Layer ("CLC" circa 2018) aggregated to 5-arcminute resolution. This entailed first approximating implied thematic transitions from CLC land-cover data to calibrated and simulated GLOBIOM/G4M land-use data at their NUTS2 aggregates by an iterative proportional fitting algorithm[28]. The thematic resolution of the reported land use states was then modified to match the Ecosystem Types level 2 classification code (MAES), in which grassland is split into managed and unmanaged grassland and transitional woodland–shrub reported separately from heathland and shrub. We next disaggregated protected lands from unprotected lands by rasterizing and masking protected areas in the Natura 2000 network with CLC data remapped to MAES classification at 100-meter resolution. Finally, we aggregate the protected areas per MAES class to shares of 5-arcminute area and intersect it with the harmonized 5-arcmin land-cover at (MAES classification). The resulting harmonized data of each land cover inside and outside current protected areas is reprojected to 10km resolution.

We further disaggregated the circa 2020 conditions from above by management intensity classes for cropland (high, medium, low intensity), forests (production, mutifunctional, natural), and pasture (high and low intensity) in each grid cell. Information on the initial distribution of production, multifunctional, natural forests is simulated by the G4M model. Pasture intensity was simulated using zoning information on low and high-intensity pasture areas from[29]. Cropland information on tillage practices and fertilizer applications were obtained from GLOBIOM. From the above initial state map (MAES level) production areas were disaggregated by intensity using the percentage of a given intensity type expected in that 10km pixel from one of the above three datasets (forestry, cropland, and pasture were disaggregated separately). In total, we distinguish between 16 land-cover types in the initial state map, where cropland and forest are separated into three different levels of management intensity (low, medium and high, Figure S1 A) and pasture into two levels of management intensity (low and high). Because our prioritization is focused on terrestrial restoration and its impact on species and carbon, we aggregate rivers, lakes, and marine inlets into one conservation zone in the results.

To identify potentially restorable natural land cover types (particularly for the "High Nature" restoration scenarios and to constrain wetland restoration to feasible locations), we followed the concept of potential natural vegetation (sensu[24]). We define potential in this context as the contemporary climatic, topography, soil, and natural vegetation conditions that allow for a specific type of natural habitat (e.g., Forest, Wetland) to occur in an area. The full methodology and input data used for estimating potential land cover are available in the SI.

Carbon data and emissions

For current carbon stocks, we used data on above-ground, below-ground, and soil organic carbon at risk from land-use change from[7]. These data were created by selecting and integrating the best available carbon maps for different vegetation types aligned with the land cover classification used. All data are in units of tC/ha. For the analysis, we combined the current carbon layers by calculating the combined sum of above- and below-ground and soil organic carbon for Europe. Similarly, for allocating restoration priorities with regard for carbon contributions, we needed to identify areas with high carbon sequestration potential. Here followed an approach that combined the different techniques from[8] and[30] for potential carbon estimation, allowing for an estimation of carbon potential per planning unit and natural land-cover types (see SI for full details on methodology).

To capture the benefits of transitioning between land cover and land use intensity through the means of restoration, we relied on emission factors as used by IPCC and G4M/GLOBIOM (see SI for additional details). For cropland and pasture we used IPCC emission factors from the 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (Table S2-S3). For forestry, we used average estimates of timber increments and stocking rates simulated by the G4M model for European forestry for each of the different production intensities (Table S4;[31]).

Problem formulation

We formulate the systematic land allocation decision as a linear programming problem, where the objective is to maximize the extent to which targets for species habitat and carbon sequestration are met through optimal zoning ("decisions", Figure S1) of conservation, restoration, and production measures across land-cover types in the EU in each of 41,046 10 km2 grid cells ("plannning units"):

$$min[\sum_{s}^{S} w_s((t_s - \sum_{p}^{P}\sum_{z}^{Z} r_{s,p} k_{z,s} x_{p,z})/t_s)] \tag{1}$$

In the minimal shortfall objective (eq.), $p \in P$ indicates a given 100 km^2 planning unit (10km grid cell), $z \in Z$ indicates a management zone (Figure S1A), and $x_{p,z}$ is the area of the planning unit p allocated to zone z in the solution. $r_{s,p}$ is the amount of feature $s \in S$ in the planning unit p and has the same unit as the respective target, t_s. Features include species area of suitable habitat (split by biome) and carbon. $k_{z,s}$ is a zone-specific parameter that defines the proportional contribution of zone z to achieving the target for feature s and is determined by the habitat preferences and land-use specific threats for the species (see SI and table S4 for additional details), or the estimated carbon sequestration of that zone (Table S2-4). w_s is a feature-specific weighting that gives higher or lower priority to achieving a given target when not all targets can be met (see scenarios for additional details).

We use the prioritzr package[32] and the Gurobi solver (v9.5)[?] to build and solve each of our spatial optimization problem scenarios.

Management zones

We consider 25 management zones (decision possibilities) that can be allocated proportionally within each of the 41,046 planning units across the EU in each of the different scenarios (see below). Restoration zones consist of both natural land cover restored from production landscapes and low-intensity production landscapes restored from high-intensity production landscapes (Figure S1, Table S1). Production zones consist of varying intensity crop, pasture, and forestry areas that can be maintained or created from other land-cover types. Finally, conservation zones capture the maintenance of natural land-cover types (i.e., land not allocated to production or restoration). We bound the allocation of each of these zones in a given planning unit based on the current and potential land cover, improving the realism of our proposed solution space:

(1) For the set of production zones, the lower bound is 0 (no production is required in any given planning unit), and the upper bound is the land area of the planning unit minus the distribution of the natural land cover in currently protected areas. For the set of conservation zones in each planning unit, we consider the lower bound to be the current area of a given land-cover type within the planning unit that is presently within protected areas (Natura 2000 site).

(2) The upper bound of conservation zones is the total extent of a given natural land-cover type in the planning unit. This helps ensures that production and conservation are not unecessarily reshuffling - natural landscapes can not be "created" and thus are conserved to the extent possible given projected production demands in a given jurisdiction.

(3) For the set of restoration zones, the lower bound is 0 (no restoration is required in any given planning unit). The upper bound of restoration potential within each 10 km^2 planning unit was refined using the land use potential and the existing land cover distributions within each PU (see above). Restoring current land cover to potential land cover is defined as possible only up to the probability of a given planning unit to be of that land-cover type (using outputs from potential distribution models, see SI for additional information). The restoration potential was allowed only for managed land cover types that are feasible for restoration to a given potential land cover. Meaning the upper bound of the potential area of restoration for a given land cover j in a given planning unit is given by eq. 2:

$$Pr(j)\sum_{i}^{I} a_i T_{i,j} \tag{2}$$

Where $Pr(j)$ is the probability of a given potential land cover in the PU, a_i is the percent of the PU area that is a given current land cover i and $T_{i,j}$is the logical transition between the current land cover and potential land cover, 1 if a managed landscape to a lower intensity managed landscape or natural ecosystem.

While we minimize proposing unnecessary or unrealistic "reshuffling" of land use (e.g., high-intensity cropland is restored to lower-intensity cropland in some planning units while being established in other planning units) by bounding the possible land use transitions and their extents, some reshuffling is allowed because in some jurisdictions future bio-economy demands will require increasing high-intensity production, while in other jurisdictions the opposite is true and restoration is possible and beneficial.

Habitat and carbon targets

Species targets are defined at the country-bioregional scale and relative to the minimum habitat area (km^2) necessary to qualify the species for the Least Concern conservation status following IUCN criteria[7,33]. We equate this to the proportional current area of suitable habitat (AOH) within the range of the species and bioregion, or the proportion of the species range within a bioregion multiplied by 2200 km^2 in the case that the EU AOH is less than 2200 km^2 (eq.3). To calculate the current AOH, we assess the proportion of each species range that is within corresponding MAES land cover preferences ("preferred" or "suitable" habitat) for the given species.

Moreover, we set a maximum area target for any given species to 10$\hat{6}$ km^2 to avoid infeasibly large AOH targets following[7].

$$t_s = \min(\max(2200, AOH_s), 10^6) \tag{3}$$

We define a zone-specific parameter $k_{z,s}$ that identifies the proportional contribution of zone z to achieving the target for feature s and is determined by the habitat preferences for the species. $k_{z,s}$ is 1 if zone z (e.g. mountain coniferous forest) fully contributes to achieving the conservation target for species s (e.g. the three-toed woodpecker). $k_{z,s}$ is 0 if the zone z does not contribute to the species' habitat. $k_{z,s}$ can be between 0 and 1 for species that persist in agricultural habitat but are sensitive to threats associated with the intensity and type of agriculture in zone z (see Table S5-S7 for more details). To differentiate the habitat provided by low and high-intensity production landscapes, we leverage frequently reported pressures on habitats and species associated with agriculture from the EEA State of Nature Dataset[34] (Table S5-S7). We map EEA land use threats onto our seven production zones. Species are then mapped to threats that they are sensitive to. If a species is sensitive to the highest quantile of threats but is considered suitable to land use, $k_{s,z}$ in eq. is set to 0.1. If a species is sensitive to mid qualtile of threats but considered suitable to land use, $k_{s,z}$ in eq.eq. is set to 0.5. If a species is sensitive to the lowest quantile of threats but considered suitable to land use, $k_{s,z}$ is set to 1.

We define a species with an improved conservation status in the solutions (Figure 2-3) as one in which a part of the range has gone from an unmet to a met habitat target.

We set the carbon sequestration target to the sum of the maximum potential value of carbon in each planning unit and similarly defined zone-specific parameters to understand the unique contributions of different land-cover types to that target.

We jointly optimize for both biodiversity and carbon sequestration targets, and since we include only one target for carbon compared to multiple targets for biodiversity outcomes, the weighting of target shortfalls is set accordingly (see information on feature weighting scenarios below and Table S1).

Meeting 2030 production demands

To minimize conflicts between conservation and restoration targets and projected demands for timber and agricultural products, we included production area requirements for each subnational (EU NUTS2) jurisdiction of the seven production zones (high-intensity cropland, reduced intensity cropland, low-intensity cropland, high-intensity pasture, low-intensity pasture, production forestry, and multi-functional forestry). Future (year 2030) land-use targets were derived from a version of the GLOBIOM/G4M economic modeling framework of the agriculture, forestry and bioenergy sectors[35]. The model was specifically adapted for EU contexts and policy problems[26,27,36] and embedded in a broader EU-focused climate mitigation modeling framework[37] that includes bioenergy biomass demand projections from the PRIMES model[38] for various feedstocks.

For the 2030 area targets for production land uses, two alternative scenarios were run with the GLOBIOM/G4M modeling framework: the "Business as usual" (BAU) and FitFor55 scenarios. The FitFor55 scenario is broadly consistent with the EU land use pathways used by the EU commission in the development of its climate targets through its long-term vision “A Clean Planet for All” and the Fit for 55 Package[39]. The BAU scenario projects expected land use developments in the absence of new policies and is broadly consistent with the official EU Reference scenario[39]. The FitFor55 scenario adds to the BAU scenario interventions (additional energy biomass demand, carbon tax applied to LULUCF CO2 emissions of 10 EUR/tCO2/yr), allowing to reach the 2030 LULUCF emission target of -310 MtCO2/yr).

The 2030 area targets are then used as constraints in the optimization, imposing the solution to contain at least as much area as projected for 2030 by GLOBIOM/G4M for each NUTS2 and production land use class. For example, to ensure the allocation of sufficient mid-intensity cropland to meet projected 2030 production needs, we require that the sum of the area allocated to the relevant cropland zones across planning units within each NUTS2 jurisdiction is greater than or equal to the specified 2030 target of that jurisdiction. We formally implement these constraints using eq. :

$$\sum_{p}^{P}\sum_{z}^{Z}(D_{i,k,p,z}\times X_{p,z})\leq t_{i,k} \qquad \forall i\in I \quad \text{and} \quad \forall k\in K \tag{4}$$

where $i \in I$ denotes a set of planning units p (e.g., all planning units within a NUTS2 jurisdiction), and $k \in K$ a set of zones relevant to a given target $t_{i,k}$. $D_{i,k,p,z}$ denotes the constraint data associated with planning units $i \in I$ for zones $k \in K$. Due to planning units being classified as a single NUTS2 region, some some NUTS2 jurisdictions required small reductions in production areas to ensure problem feasibility.

Scenarios

We explore 120 different scenarios of optimal land allocation (SI Table S1). First, we vary the weighting of minimizing carbon shortfalls relative to biodiversity shortfalls (eq. parameter w_s; 10 different weighting scenarios) to understand the synergies and trade-offs between these two broad objectives. To assess the impact of burden sharing between countries across different carbon/biodiversity objective weighting scenarios, for each of these weightings we set constraints for meeting restoration targets at (1) the EU scale (unconstrained) and (2) the Member State scale (even % area contribution per country) and (3) with up to 10% increase in the amount of priority burden allocated to a given country. We run each of the above scenarios for both a (1) “baseline” restoration law implementation where >5% of EU land area is allocated to cropland or pasture deintensification, >7% to forestry deintensification, and <1% to natural landscape restoration (including wetland restoration) and (2) a “high nature” restoration law implementation where >5% of EU land area is allocated to cropland or pasture deintensification, >6% to forestry deintensification, and <6% to natural landscape restoration (including wetland restoration). Each of these is run under production constraints aligned with a BAU and Fit for 55 scenario. Finally, we compare a single scenario (mid carbon-biodiversity weight, no burden sharing, baseline restoration implementation) that naively allocates restoration to the same scenario that prioritizes restoration allocation in the context of conservation and broader landscape constraints (Figure S3).

Supplementary information

Supplemental methods and figures are provided in a separate document.

Data and code availability

All code and processed data to solve the optimization problem and create figures is available at: https://github.com/milliechapman/EU-restoration-prioritization

Acknowledgments

MC would like to acknowledge financial support from the International Institute for Applied Systems Analysis, Laxenburg (Austria) and the National Member Organizations that support the institute in taking part in the Young Scientists Summer Program. PV, DL, MG, AL and MJ acknowledge funding from the European Union Biodiversity and Climate Strategies Assessment (EU BIOCLIMA).

References

1. Díaz, S. *et al.* Assessing nature's contributions to people. *Science* **359**, 270–272, DOI: 10.1126/science.aap8826 (2018).
2. Convention on Biological Diversity. Kunming-Montreal Global Biodiversity Framework (2022).
3. Xu, H. *et al.* Ensuring effective implementation of the post-2020 global biodiversity targets. *Nat. Ecol. & Evol.* **5**, 411–418, DOI: 10.1038/s41559-020-01375-y (2021).
4. Alexander, P. *et al.* High energy and fertilizer prices are more damaging than food export curtailment from Ukraine and Russia for food prices, health and the environment. *Nat. Food* **4**, 84–95, DOI: 10.1038/s43016-022-00659-9 (2022).
5. Arneth, A. *et al.* Making protected areas effective for biodiversity, climate and food. *Glob. Chang. Biol.* **29**, 3883–3894, DOI: 10.1111/gcb.16664 (2023).
6. Strange, N., Geldmann, J., Burgess, N. D. & Bull, J. W. Policy responses to the Ukraine crisis threaten European biodiversity. *Nat. Ecol. & Evol.* **6**, 1048–1049, DOI: 10.1038/s41559-022-01786-z (2022).
7. Jung, M. *et al.* Areas of global importance for conserving terrestrial biodiversity, carbon and water. *Nat. Ecol. & Evol.* **5**, 1499–1509, DOI: 10.1038/s41559-021-01528-7 (2021).
8. Strassburg, B. B. N. *et al.* Global priority areas for ecosystem restoration. *Nature* **586**, 724–729, DOI: 10.1038/s41586-020-2784-9 (2020).
9. Pe'er, G. *et al.* Scientists support the EU's Green Deal and reject the unjustified argumentation against the Sustainable Use Regulation and the Nature Restoration Law. DOI: 10.5281/ZENODO.8128624 (2023). Publisher: Zenodo Version Number: Full Version 9.7.2023 (preprint).
10. Staccione, A. *et al.* Exploring the effects of protected area networks on the European land system. *J. Environ. Manag.* **337**, 117741, DOI: 10.1016/j.jenvman.2023.117741 (2023).
11. Searchinger, T., James, O., Dumas, P., Kastner, T. & Wirsenius, S. EU climate plan sacrifices carbon storage and biodiversity for bioenergy. *Nature* **612**, 27–30, DOI: 10.1038/d41586-022-04133-1 (2022).
12. Commission, E. Commission staff working document impact assessment accompanying the proposal for a Regulation of the European Parliament and of the Council on nature restoration. (2022).
13. Yoshioka, A., Akasaka, M. & Kadoya, T. Spatial Prioritization for Biodiversity Restoration: A Simple Framework Referencing Past Species Distributions: Simple Biodiversity Restoration Prioritization. *Restor. Ecol.* **22**, 185–195, DOI: 10.1111/rec.12075 (2014).
14. Doelman, J. C. *et al.* Quantifying synergies and trade-offs in the global water-land-food-climate nexus using a multi-model scenario approach. *Environ. Res. Lett.* **17**, 045004, DOI: 10.1088/1748-9326/ac5766 (2022).
15. Adame, M. F., Hermoso, V., Perhans, K., Lovelock, C. E. & Herrera-Silveira, J. A. Selecting cost-effective areas for restoration of ecosystem services: Cost-Effective Restoration. *Conserv. Biol.* **29**, 493–502, DOI: 10.1111/cobi.12391 (2015).
16. Chaplin-Kramer, R. *et al.* Global modeling of nature's contributions to people. *Science* **366**, 255–258, DOI: 10.1126/science.aaw3372 (2019).
17. De Groot, R. S. *et al.* Benefits of Investing in Ecosystem Restoration: Investing in Ecosystem Restoration. *Conserv. Biol.* **27**, 1286–1293, DOI: 10.1111/cobi.12158 (2013).
18. Martin, D. M. & Lyons, J. E. Monitoring the social benefits of ecological restoration. *Restor. Ecol.* **26**, 1045–1050, DOI: 10.1111/rec.12888 (2018).
19. Malkamäki, A. *et al.* A systematic review of the socio-economic impacts of large-scale tree plantations, worldwide. *Glob. Environ. Chang.* **53**, 90–103, DOI: 10.1016/j.gloenvcha.2018.09.001 (2018).
20. Buckley, M. C. & Crone, E. E. Negative Off-Site Impacts of Ecological Restoration: Understanding and Addressing the Conflict. *Conserv. Biol.* **22**, 1118–1124, DOI: 10.1111/j.1523-1739.2008.01027.x (2008).
21. Pörtner, H.-O. *et al.* Overcoming the coupled climate and biodiversity crises and their societal impacts. *Science* **380**, eabl4881, DOI: 10.1126/science.abl4881 (2023).
22. Maxwell, S. L. *et al.* Area-based conservation in the twenty-first century. *Nature* **586**, 217–227, DOI: 10.1038/s41586-020-2773-z (2020).
23. Jung, M. An integrated species distribution modelling framework for heterogeneous biodiversity data. *Ecol. Informatics* **76**, 102127, DOI: 10.1016/j.ecoinf.2023.102127 (2023).

24. Hengl, T. *et al.* Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential. *PeerJ* **6**, DOI: 10.7717/peerj.5457 (2018).

25. Bohn, U. Map of the Natural Vegetation of Europe. Tech. Rep. (2004).

26. Frank, S. *et al.* Dynamics of the land use, land use change, and forestry sink in the European Union: the impacts of energy and climate targets for 2030. *Clim. Chang.* **138**, 253–266, DOI: 10.1007/s10584-016-1729-7 (2016).

27. Frank, S., Gusti, M., Valin, H., Forsell, N. & Havlík, P. Documentation for estimating LULUCF emissions / removals and mitigation potentials with GLOBIOM/G4M. Tech. Rep. (2020).

28. Fienberg, S. E. An Iterative Procedure for Estimation in Contingency Tables. *The Annals Math. Stat.* **41**, 907–917, DOI: 10.1214/aoms/1177696968 (1970).

29. Dou, Y. *et al.* A new European land systems representation accounting for landscape characteristics. *Landsc. Ecol.* **36**, 2215–2234, DOI: 10.1007/s10980-021-01227-5 (2021).

30. Walker, W. S. *et al.* The global potential for increased storage of carbon on land. *Proc. Natl. Acad. Sci.* **119**, e2111312119, DOI: 10.1073/pnas.2111312119 (2022).

31. Kindermann, G. E. *et al.* Potential stocks and increments of woody biomass in the European Union under different management and climate scenarios. *Carbon Balanc. Manag.* **8**, 2, DOI: 10.1186/1750-0680-8-2 (2013).

32. Hanson, J. O. *et al.* prioritizr: Systematic Conservation Prioritization in R (2023).

33. Fastré, C., Van Zeist, W.-J., Watson, J. & Visconti, P. Integrated spatial planning for biodiversity conservation and food production. *One Earth* **4**, 1635–1644, DOI: 10.1016/j.oneear.2021.10.014 (2021).

34. Roscher, S., Conde, S. & Maitre, J. Final database on linkages between species/habitat-types and broad ecosystems. Tech. Rep. (2015).

35. Havlík, P. *et al.* Global land-use implications of first and second generation biofuel targets. *Energy Policy* **39**, 5690–5702, DOI: 10.1016/j.enpol.2010.03.030 (2011).

36. Forsell, N. *et al.* Impact of modelling choices on setting the reference levels for the EU forest carbon sinks: how do different assumptions affect the country-specific forest reference levels? *Carbon Balanc. Manag.* **14**, 10, DOI: 10.1186/s13021-019-0125-9 (2019).

37. Weitzel, M. *et al.* Model-based assessments for long-term climate strategies. *Nat. Clim. Chang.* **9**, 345–347, DOI: 10.1038/s41558-019-0453-5 (2019).

38. Capros, P. *et al.* Description of models and scenarios used to assess European decarbonisation pathways. *Energy Strateg. Rev.* **2**, 220–230, DOI: 10.1016/j.esr.2013.12.008 (2014).

39. European Commission. Directorate General for Energy., European Commission. Directorate General for Climate Action. & European Commission. Directorate General for Mobility and Transport. *EU reference scenario 2020: energy, transport and GHG emissions : trends to 2050.* (Publications Office, LU, 2021).

Supplemental Information

Melissa Chapman[1,2], Martin Jung[2], David Leclère[3], Carl Boettiger[4], Andrey L. D.Augustynczik[3], Mykola Gusti[3], Leopold Ringwald[3], and Piero Visconti[2]

[1]National Center for Ecological Analysis and Synthesis, Santa Barbara, CA, USA
[2]Biodiversity, Ecology and Conservation Research Group, International Institute for Applied Systems Analysis (IIASA), Vienna,Austria
[3]Integrated Biosphere Futures Research Group, International Institute for Applied Systems Analysis (IIASA), Vienna,Austria
[4]Department of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley, CA, USA

Meeting European conservation and restoration targets under future land-use demands

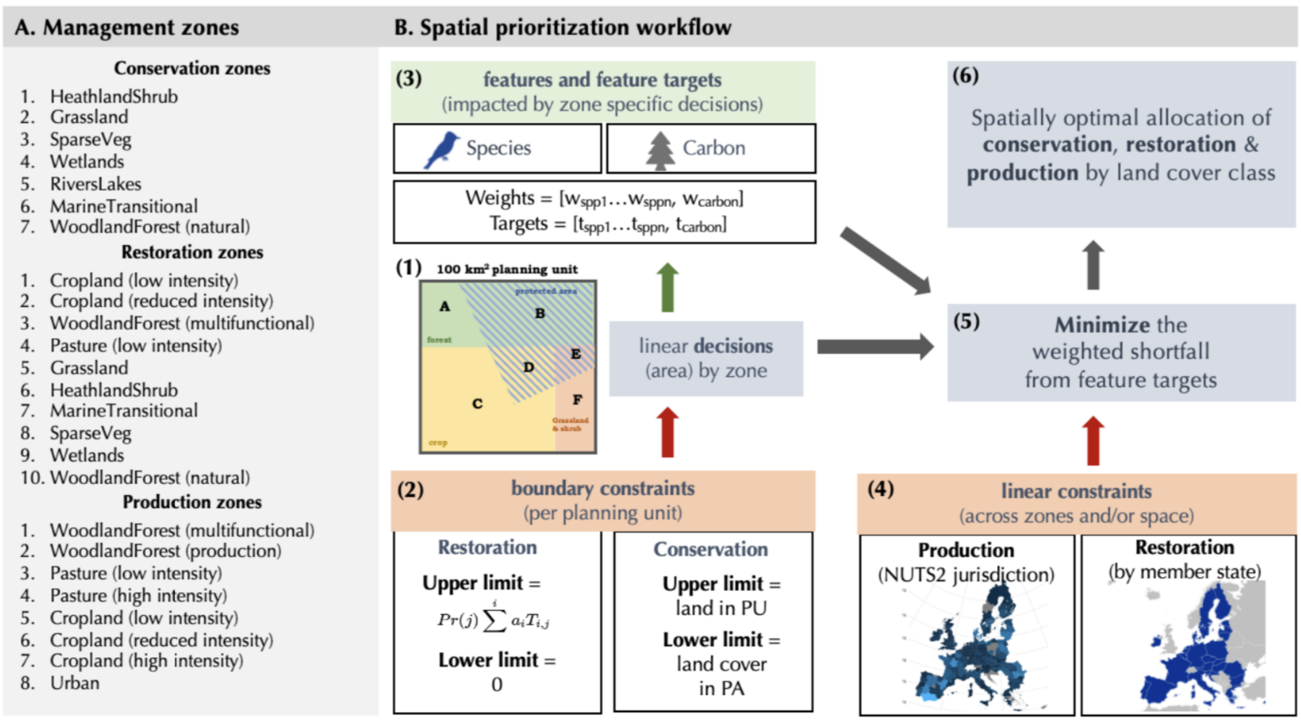

Figure 1. Schematic diagram of the prioritization analyses. (A) List of management zones, whose spatial allocation is optimized under a suite of different scenarios (Table S1) in the analyses. (B) For each species, we set an extinction-risk informed target to be met by conserving or restoring habitat types for the species within their potential range. For carbon, the aim was to maximize the amount of carbon stored in conserved or restored areas. Depending on the scenario variants, species were weighted in importance differently relative to carbon to explore the implications of putting more emphasis on different objectives. The area allocation of a planning unit to a given management zone was bounded depending on the planning unit and zone (B2). In addition, the optimization included constraints on the area under restoration and how much area needed to be under production (grazing, farming, timber harvesting, B4) in 2030. The result of this constrained optimization is a series of maps identifying priorities for conservation, restoration, and production of food and timber products (B6).

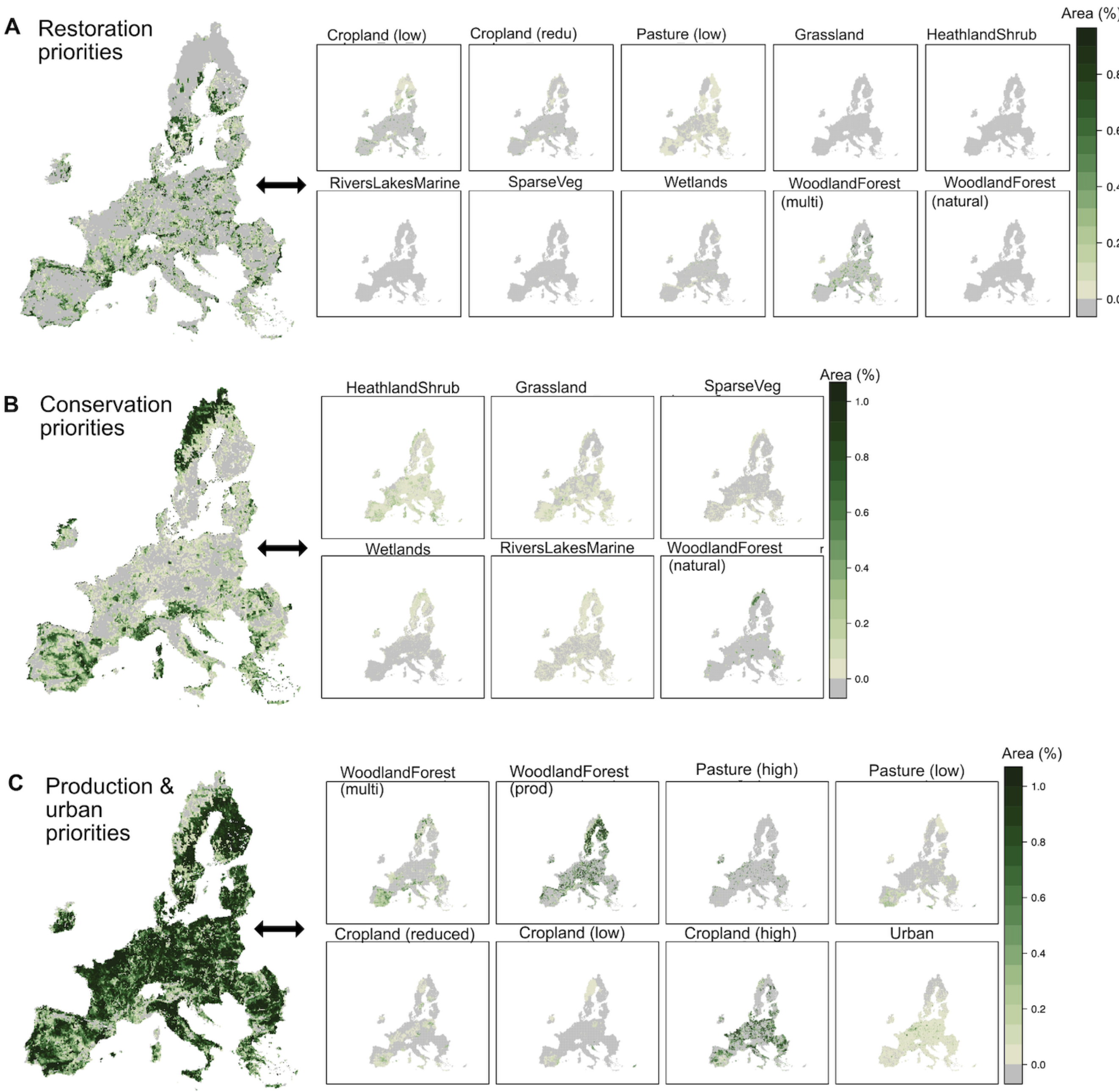

Figure 2. **Optimal allocation of conservation, restoration, and production across land-cover types for one scenario** (carbon target weight set to equal the sum of all species targets and restoration budgets defined at the member state scale). (A) The optimal allocation of conservation (maintained natural land) across the EU and the breakdown of conservation by land cover type. (B) The optimal allocation of restoration across the EU and the breakdown of restoration by land cover type. (C) Production area targets, while constant throughout scenarios at the sub-national (NUTS2) level, vary in their spatial distribution as the result of the conservation and restoration priorities of that given scenario. Urban areas remain constant throughout all scenario solutions and are set to match 2018 urban area distributions in each planning unit.

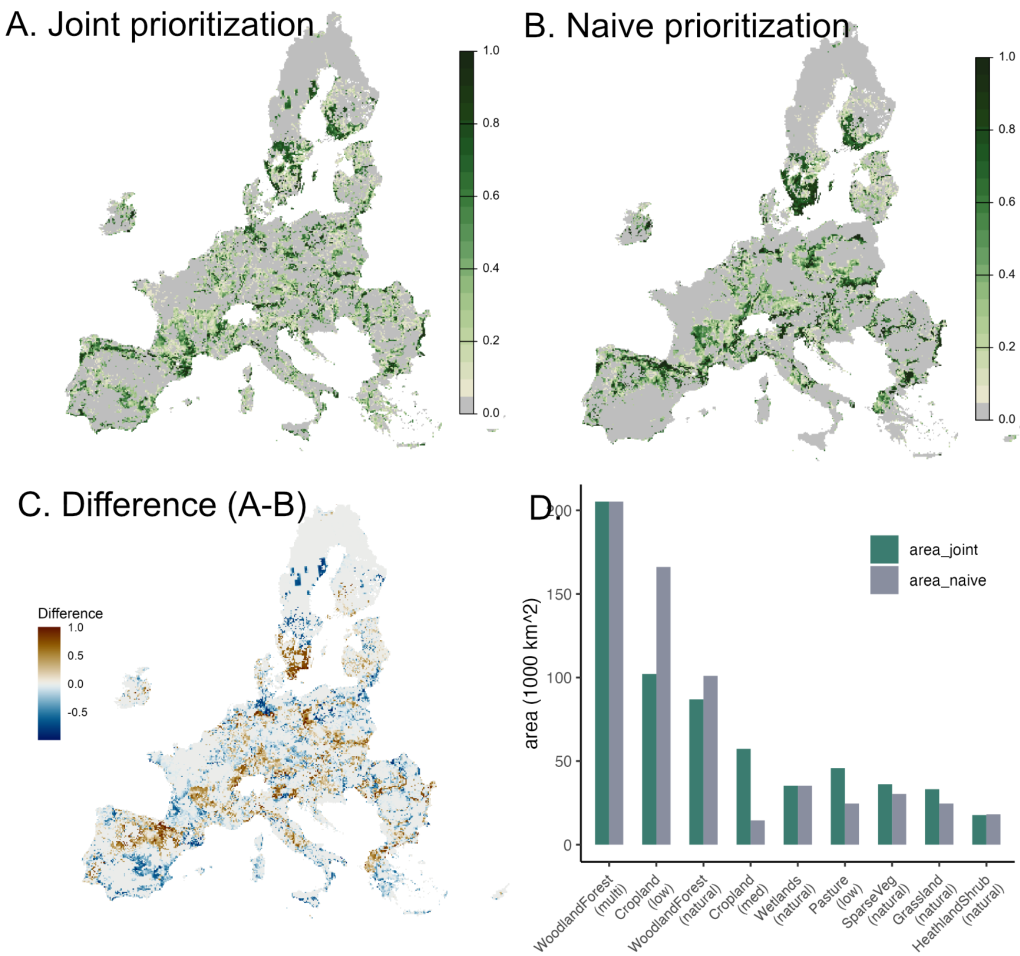

Figure 3. Implications of jointly optimizing the allocation of restoration and conservation. Using the same planning unit constraints and restoration transition matrix from figure 2, we compare solutions of restoration prioritization when (A) jointly optimized to meet restoration targets in the context of the optimal allocation of conservation and production lands (same as figure 2) and (B) only considering restoration priorities. (C) Significant spatial differences and (D) land type differences in restoration priority emerge between the two solutions.

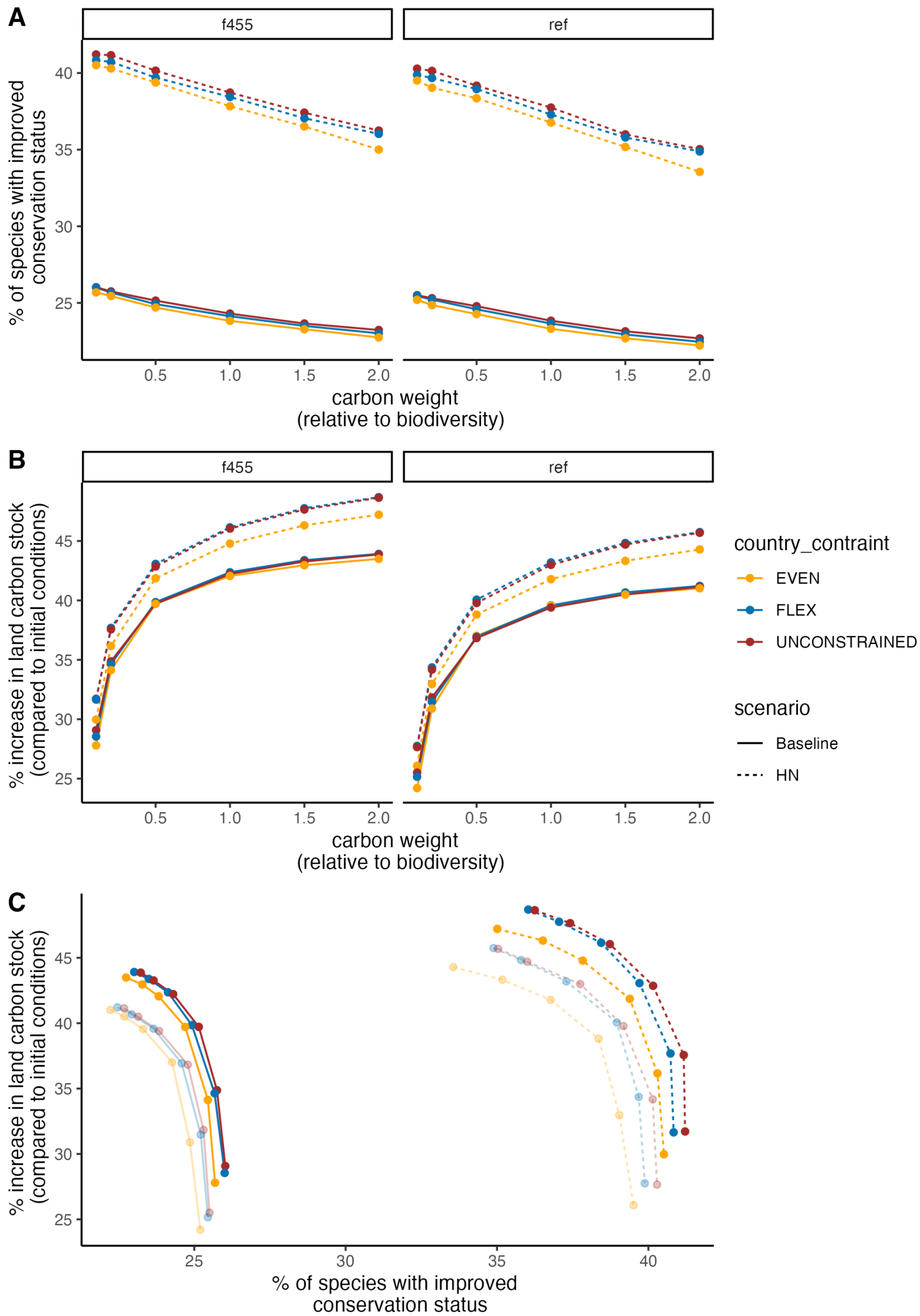

Figure 4. Each point represents a different weighting carbon relative to biodiversity objectives (A) Contributions of solutions to improving the conservation status of species have a linear response to the weighting of species targets relative to carbon targets across all scenarios. (B) By contrast, carbon stock increases are non-linearly with carbon weight. (C) The carbon and biodiversity impacts of all scenarios show that objective weighting and restoration scenarios have a larger impact than burden sharing on the expected 2030 outcomes.

Table 1. We consider 120 different solution scenarios (Figure 1A) below and throughout the manuscript.

Restoration scenario	**Burden Sharing**	**Carbon Weight**	**Production constraints**
Baseline	Even	10 different weighting (0.1-2)	F455
Baseline	Flex (restoration up to 25% per MS)	10 different weighting (0.1-2)	F455
Baseline	Unconstrained	10 different weighting (0.1-2)	F455
Baseline	Even	10 different weighting (0.1-2)	REF
Baseline	Flex (restoration up to 25% per MS)	10 different weighting (0.1-2)	REF
Baseline	Unconstrained	10 different weighting (0.1-2)	REF
High Nature	Even	10 different weighting (0.1-2)	F455
High Nature	Flex (restoration up to 25% per MS)	10 different weighting (0.1-2)	F455
High Nature	Unconstrained	10 different weighting (0.1-2)	F455
High Nature	Even	10 different weighting (0.1-2)	REF
High Nature	Flex (restoration up to 25% per MS)	10 different weighting (0.1-2)	REF
High Nature	Unconstrained	10 different weighting (0.1-2)	REF

Table 2. Cropland carbon adjustments by intensity level. These values are used as multipliers to refine the spatial carbon estimates which are not intensity specific. The mean IPCC default value for different production intensities are mapped to zone classification and normalized to 1.

Level	Temperature regime	Zone classification	IPCC default
full	all	high	1
reduced	cool temperate	mid	0.98
reduced	cool temperate	mid	1.04
reduced	warm temperate	mid	0.99
reduced	warm temperate	mid	1.05
no-till	cool temperate	low	1.03
no-till	cool temperate	low	1.09
no-till	warm temperate	low	1.04
no-till	warm temperate	low	1.1

Table 3. Pasture carbon adjustments by intensity level. These values are used as multipliers to refine the spatial carbon estimates which are not intensity specific. The mean IPCC default value per zone classification is normalized to 1.

	Level	Temperature regime	Zone classification	IPCC Default
Management (FMG)	Nominally managed (non – degraded)	all	High intensity	1
Management (FMG)	High intensity grazing	all	High intensity	0.9
Management (FMG)	Severly degraded	all	High intensity	0.7
Management (FMG)	Improved	temperate	Low intensity	1.14

Table 4. Forestry carbon adjustments by intensity level. These values are calculated using outputs from the G4M model. We average values across all countries for simplicity and use as a multiplier to spatially explicit carbon maps.

Intensity	**Value**	**Emission factor**
production	194.70	0.40
primary	484.72	1
multi	339.70	0.70

Table 5. Mapping EEA threats onto crop production intensity zones. Threat codes align with species included in the prioritization, allowing for the differentiation of biodiversity contributions of different intensities of production.

Code	Pressure/threat	Crop (high)	Crop (redu)	Crop (low)
A01	Conversion into agricultural land	3	2	1
A02	Conversion from one type of agricultural land use to another	2	1	0
A03	Conversion from mixed farming/ agroforestry to specialised production	2	1	0
A04	Changes in terrain and surface of agricultural areas	2	1	0
A05	Removal of small landscape features for agricultural land parcel consolidation	1	1	0
A07	Abandonment of management/use of other agricultural/agroforestry systems	1	1	0
A08	Mowing or cutting of grasslands	1	1	0
A15	Tillage practices (e.g. ploughing) in agriculture	2	1	0
A16	Other soil management practices in agriculture	2	1	0
A17	Harvesting of crops and cutting of croplands	2	1	0
A18	Irrigation of agricultural land	2	1	1
A19	Application of natural fertilisers on agricultural land	2	1	1
A20	Application of synthetic (mineral) fertilisers on agricultural land	2	1	1
A21	Use of plant protection chemicals in agriculture	2	1	1
A22	Use of physical plant protection in agriculture	2	1	1
A23	Use of other pest control methods in agriculture (excluding tillage)	1	1	2
A24	Waste management practices in agriculture	1	1	1
A25	Agricultural activities generating point source pollution to surface/ground waters	3	2	1
A26	Agricultural activities generating diffuse pollution to surface or ground waters	3	2	1
A29	Agricultural activities generating soil pollution	3	2	1
A30	Active abstractions from groundwater/ surface water for agriculture	3	2	1
A31	Drainage for use as agricultural land	3	2	1
A34	Introduction and spread of new crops (including GMOs)	2	1	0
A35	Agricultural crops for renewable energy production	3	2	1

Table 6. Mapping EEA threats onto pasture production intensity zones. Threat codes align with species included in the prioritization, allowing for the differentiation of biodiversity contributions of different intensities of production.

Code	Pressure/threat	PastureLow	PastureHigh
A01	Conversion into agricultural land	1	2
A02	Conversion from one type of agricultural land use to another	0	1
A03	Conversion from mixed farming/ agroforestry systems to specialised production	0	1
A04	Changes in terrain and surface of agricultural areas	1	2
A05	Removal of small landscape features for agricultural land parcel consolidation	0	1
A06	Abandonment of grassland management	0	1
A07	Abandonment of management/use of other agricultural and agroforestry systems	0	1
A08	Mowing or cutting of grasslands	1	2
A09	Intensive grazing or overgrazing by livestock	1	2
A10	Extensive grazing or undergrazing by livestock	0	0
A13	Reseeding of grasslands and other semi-natural habitats	1	1
A14	Livestock farming (without grazing)	1	2

Table 7. Mapping EEA threats onto forestry production intensity zones. Threat codes align with species included in the prioritization, allowing for the differentiation of biodiversity contributions of different intensities of production.

Code	Pressure/threat	Forestry (multi)	Forestry (production)
B01	Conversion to forest from other land uses, or afforestation (excluding drainage)	0	0
B02	Conversion to other types of forests including monocultures	1	1
B03	Replanting with or introducing non-native or non-typical species	1	1
B04	Abandonment of traditional forest management	1	2
B05	Logging without replanting or natural regrowth	1	1
B06	Logging (excluding clear cutting) of individual trees	1	0
B07	Removal of dead and dying trees, including debris	1	2
B08	Removal of old trees (excluding dead or dying trees)	1	1
B09	Clear-cutting, removal of all trees	0	1
B10	Illegal logging	1	1
B11	Cork extraction and forest exploitation excluding logging	0	0
B12	Thinning of tree layer	1	1
B13	Burning for forestry	0	0
B14	Suppression of fire for forestry	0	0
B15	Forest management reducing old growth forests	0	1
B16	Wood transport	1	1
B17	Tillage practices in forestry and other soil management practices in forestry	0	1
B18	Application of natural fertilisers	1	1
B19	Application of synthetic fertilisers in forestry, including liming of forest soils	1	2
B20	Use of plant protection chemicals in forestry	1	2
B21	Use of physical plant protection in forestry, excluding tree layer thinning	1	1
B22	Use of other pest control methods in forestry	1	1
B23	Forestry activities generating pollution to surface or ground waters	1	2
B24	Forestry activities generating air pollution	1	1
B25	Forestry activities generating marine pollution	1	2
B26	Forestry activities generating soil pollution	1	2
B27	Modification of hydrological conditions	1	1
B28	Forests for renewable energy production	1	1
B29	Other forestry activities, excluding those relating to agro-forestry	1	2

Supplemental Methods

1 Data

1.1 Biodiversity data

1.1.1 Biodiversity data collation

Openly available biodiversity data sources in Europe are heterogeneous in type, format and purpose; and to be able to use them in an integrated SDM type of approach, a considerable amount of data harmonization and format control is necessary.

Throughout we followed the taxonomic "backbone" of GBIF and codified functions to harmonize and match taxonomic names from different data sources to the GBIF taxonomy backbone of 2021, (GBIF Secretariat (2021)). We primarily focused throughout on terrestrial species listed in the EU Article 12 (Birds directive) and Article 17 (Habitats directive), or which are assessed by European Redlist of species. A complete list of all included species can be found at `https://github.com/milliechapman/EU-restoration-prioritization/blob/main/figures/updated/spp_list.csv`.

Firstly, we obtained presence-only records from GBIF for all animal and plant species in the database ((GBIF Secretariat (2021)). We excluded fossil specimens, and those with invalid spatial coordinates, and applied standard data pre-preprocessing steps for unstructured citizen science data using the 'CoordinateCleaner' R package[1]. We removed duplicate points (those within a 2-km distance within the same year) and highly uncertain records.

Additionally, we collated taxonomic group specific data for bird and plant species. For birds we made use of eBird data[2], which we processed similarly as GBIF records above. Potential absence data were inferred from sites where full communities were assessed and for which the focal species had never been recorded. To further limit the number of total absence sites, we first took eBird sites where the species had been recorded, and spatially buffered these by 200 km, excluding any sites within this buffer zone. From the remaining potential absence sites, we randomly selected an equal number of absences as presence sites in which the species was recorded, up to a maximum of 500 per species. For plant species, we also obtained presence-absence data from comprehensive vegetation plot inventories collated and made available through the SPlotOpen database[3]. We filtered these data to Europe, and the representative subset species, as well as excluding any observations prior the year 2000 as above, and to records with a positional uncertainty of less than 2 km. We inferred absence data similarly to eBird but using a lower maximum of 100 absences at maximum, because of the smaller size of this dataset.

We further obtained polygonal global, European, and Mediterranean species ranges from the IUCN Red List version 2021-2 (IUCN 2021) and from BirdLife International (BirdLife International and Handbook of the Birds of the World 2020). These data were filtered to only include areas where species were recorded as extant, possibly extant, or possibly extinct, and included all seasonal occurrences. Where existing for a given species, we further compiled habitat preference (land-cover and elevation) and threat information using data from the IUCN Red List (IUCN 2021) and (EEA preferences). Those estimates were linked to species-specific priors (see below).

Finally, we also obtained the 2020 spatial distribution data for birds listed in the Article 12 Birds Directive and for animals and plants listed in the Article 17 Habitats Directive, excluding sensitive species. Although these data do include population estimates recorded at the (sometimes sub-) Member State level, for this work we used these data only as occurrence data recorded as presence-only atlas data on a 10-km grid across Europe. For the course of this work, we treated the species information reported at various sites (polygons) in the Natura2000 network as presence-absence information, recognizing that surveys in some sites might be incomplete or outdated and not all species are necessarily recorded during. A R-package was created to format these data for the modelling (https://github.com/iiasa/rN2000). We supplemented these data with species checklists for Important Bird Areas (IBAs,[4]) across Europe, adding presence-absence records per species for the polygon area of the IBAs where the species was recorded or not.

1.1.2 Predictor Variables

We prepared a series of environmental predictors related to topography, soil conditions, climate and land cover.

In both planning and species distribution modeling, we make use of land cover and land-use data that is consistent with the European Ecosystem accounting framework such as the Mapping and Assessment of Ecosystems and their Services (MAES) system. For the current distribution of land cover, we used data from the (`https://land.copernicus.eu/pan-european/corine-land-cover/clc2018`). For mapping the potential natural distribution of a species (see below) we additionally considered data on the potential distribution of land cover[5] matched to the same legend. The thematic legend of the Corine land-cover data was then recategorized into different MAES categories through a crosswalk, however, differed from the MAES categories as we split the class Pasture (2.3.1) from other Grasslands as considered by the MAES Grassland class.

In addition to land cover and land use we also considered data on topography of European landscapes making use of the EU DEM ver1.1 (`https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1`). We considered the mean elevation (in m), the topographic position index (TPI,[6]) and the aspect of the topography, which we transformed into eastness and northness estimates through a sinusoidal and cosinus transformation respectively. This transformation is necessary to avoid circularities in units (0 degree being closer to 359 degrees) caused by the degrees-based characterization of an aspect layer.

We used long-term average climatic conditions in Europe. Specifically, we leveraged downscaled bioclimatic ERA5 indicators over the last 40 years (1979 to 2018) from the European Copernicus program. These climatic indicators represent Essential Climate Variables (ECV) such as the surface energy, drought or moisture all of which are known to be important factors in delineating the range and environmental niche. Specifically, we made use of the BIOCLIM data BIO01 to BIO19, average aridity and cloud cover, the annual sum of frost days, potential evaporation and volumetric soil water as well as different characterizations of the number for growing degree days and the start, end and length of the growing season[7]. For those parts of European member states which are missing in the European downscaled Copernicus Climate products (Such as the Spanish Canary islands) we used the average values of the global rather than the downscaled climate product instead.

For both current and potential species projections of species distributions, we considered only variables related to land cover as well as temporally static variables that are unlikely to change in a future world such as for instance altitude. On the other hand, for predictions of the potential natural distribution of a species we made use of all predictors excluding those related to land cover and land use. We added species-group-specific sets of predictors to the modeling to be included. For example, for bird species, we additionally included a layer depicting the Euclidean distance to the ocean from each land grid cell, given the importance of marine waters to many onshore nesting grassland and wetland species. For plant species we included spatial-explicit predictors related to groundwater and soil conditions, specifically data on the Ph value and Calcium Carbonate content of groundwater resources as well as estimates of the depth to groundwater in meters from[8]. In particular, groundwater Ph and Calcium carbonate do not cover small islands (Madeira, Canary Islands) as well as Cyprus, which is why we filled any remaining missing values of the predictors with a spatial prediction from a random forest model, using spatial proximity and climatic variables as covariates[5]. We also included data from a thematic layer of a European soil lithology classification system owing to the importance of different soil types to growth and niche space of plant species.

Most observational species occurrence points in Europe are known to be biased towards areas with higher accessibility and wealth, with critically Eastern European member states having a comparably lower density of records compared to western and northern European member states. Besides the integration of multiple datasets, we attempted to control for such sampling biases through a model-based control following[9]. To do so, we first took the presence-only and presence-absence localities of all biodiversity sources considered in this work (see above) and rasterizing them for the target background. This resulted in a counted number of points for each grid cell, which we then aggregated overall as the total sum of all occurrences. Furthermore, we prepared data on the accessibility of land ([10]) and the human population density per grid cell using data from the GHSL product for Europe ([11]). The biased background grid was then created by first calculating an adjusted log transformation of each individual layer through the following equation ($log(1+x-min(x)*w)$), where w stands for a numerical weight reflective of the direction of bias (-1 for accessibility, 1 for all others), and afterwards the individual layers normalized to a range of 0 to 1 and averaged.

Finally, all layers were aggregated from their original resolution to a 10-km (Lamberts Equal Area projection) grain size determined by the background modelling domain layer. We did so by either calculating the proportion of grid cells in each coarser grain if these binary data, by calculating the bilinear resampled average of all values within the coarser grain, or – in the case of multinomial categorical data – calculating the mode of all finer grained classes. For the modelling all continuous variables were standardized to the mean and divided by their standard deviation, thus ensuring compatibility in terms of units. All covariates were matched against the modelling extent and made consistent with a NUTS2 representation of European countries. Any missing data at the pixel level not filled or extrapolated at this stage was filled with missing data across all covariates. All calculations and variable preparations were done using GDAL and R packages such as "Raster" or "sf".

1.1.3 Current distribution of a species

We estimated the current distribution of species using an integrated species distribution modelling (iSDM) approach where different best-available data sources (occurrence, preference, expert information) are integrated into one joint prediction using different types of linear and non-linear modelling approaches[12].

We collated for each species the available suitable data (see above), separating between different types, namely presence-only, presence-absence and polygon presence-only data, the latter of which is used in the form of spatial-explicit offsets (e.g. expert range). For species occurring in Natura 2000 or Important Bird area (IBA) sites, we assumed that the species occur in all Natura 2000 sites in which a presence was indicated and that biophysical conditions are relatively homogeneous within Natura 2000 sites (which, given the small size of many sites and the SDM modelling grain size of 10km is a reasonable assumption).

We sampled at random across all Natura 2000 and IBA sites presence point estimates up to a number of two-times as many as there are other occurrence observations (see above), to broadly characterize the environmental conditions prevalent in those sites. We furthermore created an equal number of absence point data which we sampled at random across all sites excluding the ones where the species has been recorded as present (thus resulting in species-specific contrasts).

Because of computational reasons and to further reduce sampling biases, we applied thinning on point data across all datasets for species with more than 200 records. The process of spatial thinning removes occurrence points at random from areas that are oversampled, for example because of sampling or spatial biases in the database[13]. Notably thinning only removes points from grid cells where there are multiple and never removes all points from any grid cell completely. For presence-only records from GBIF – usually the largest data source by size – we first applied a bias thinning, where we preferentially removed observations from 10km grid cells considered as biased (based on occurrence information across all species). In addition, and across all point occurrence datasets, we also removed at random observations until a minimum number of 10 points at maximum has been reached. This approach ensures that presence and absence information (where existing) are relatively homogeneously distributed in density across the European land area, thus representing average conditions representing the suitable species habitats across the modelling period.

Whenever presence-only atlas or expert-delineated information on the occurrence of species existed, such as for example from the global, Mediterranean or European IUCN assessments, the Amphibian and Reptile atlas[14], the Atlas Hymenoptera for bumblebee species (http://www.atlashymenoptera.net/default.aspx) or for polygon information from the EU Habitats directive or Bird directive data, we included this information as spatial offsets. For IUCN we only used those parts of the range where the species is permanently resident or which are part of its breeding distribution. Spatial offsets can be included in species distribution model as spatial priors, thus increase the probability of any given grid cell to be suitable for a species[15]. We first binarized the different range estimates and then calculated the Euclidean distance from the boundary of the range to all other grid cells in the modelling background, after which we applied a negative exponential kernel with a average dispersal distance estimates to account for the decreasing suitable value of a grid cell[16]. The resulting distance layer was then rescaled to a range from 0 (furthest away) to 1 (within the range). Notably we used different distance transforms depending on whether bird or non-bird species were estimated, using either infinite and an average 20-km distance transformations for non-bird species respectively. All offsets created in this way were log-transformed before adding them to any model using presence-only information and in the case multiple offsets were supplied, these were combined first via simple multiplication.

To avoid overprediction into novel areas, the predictions were spatially constrained by a broad environmental zoning layer[?] and the expanded offsets highlighted above. This was done by removing broad zones in which there are no contemporary occurrence points to avoid, for instance, extrapolations from a Mediterranean into boreal climatic zones, while also allowing modest extrapolations within similar environmental conditions. It should be noted that this zoning was only included for current projections and not for any potential distribution. We also tested for collinearity between included variables, removing those that were highly collinear (Pearson's r > 0.7) unless they were known to be of particular importance to a species (e.g. have a set prior, see below).

We estimated the potential distribution of the species through an ensemble modelling approach (stacked SDM,[17, 18]) using state-of-the-art machine learning and Bayesian algorithms that complement each other's strengths. Model structure and response were determined based on data type, with Poisson Process models being fitted for presence-only datasets and logistic regressions for presence-absence data. We fitted tree-based regressions using the XGBoost modelling approach[19]. XGBoost makes use of gradient descent boosting, supports variable regularization and also non-linear tree-based regressions. XGBoost Models were fitted using a total of 10000 boosting iterations, a learning rate of 0.001 and Gamma parameter of 4 (larger is more conservative) for regularization. We also used another gradient descent boosting algorithm (GDB) available from the 'mboost' R-package[20]. GDB models makes use of non-linear baselearners (splines) for additive inference similar to the popular Generalized Additive Models (GAMs), however in contrast to GAMs it also supports variable regularization directly through boosting and additional baselearners (see below on priors). Here models were fitted using a total of 2500 boosting iterations and a learning rate of 0.001 per iteration. Bayesian regularized regressions were fitted using the 'BoomSpikeSlab' R-package[21] and 10 000 MCMC iterations and four MCMC chains. Lastly we used regularized linear regression models fitted with the "glmnet" package[22]. In a Poisson-Process modelling framing these type of regressions are statistically equivalent to the popular maxent package[23].

Models were only fitted for those species for which at least – after thinning - 20 data points were available, assuming that species with fewer records have not been sampled comprehensively enough to make inferences about their current distribution. Further, we made use of simple rules to avoid fitting overly complex models for a limited number of observations. Only linear models (boosted and non-boosted) were fitted for species with fewer than 100 observational points and for species with point observations fewer than 1.5 times the minimum data size of 20, we did only fit linear Bayesian Poisson Process models and not use any non-linear or boosted approaches to avoid overfitting to limited datasets[24]. Linear regressions, compared to non-linear ones, usually fare better when the goal is extrapolation and are less prone to model overfitting. In case only presence-only information from GBIF was available for a species, we furthermore included spatial effects as covariate using polynomial-transformed coordinates[16].

Integrating prior information on species habitat and elevational preferences and distances to known occurrences can improve range estimates. We obtained information on the susceptibility of species to certain habitat and elevational preferences from the EEA habitat preference database and IUCN[25]. Preferences to certain land-cover types in the IUCN habitat preferences or respectably Corine land-cover categories were remapped to MAES categories. Priors can help to stabilize and avoid mapping unrealistic response functions to certain covariates[26]. Elevational preferences were included as specific threshold transforms on the raw elevation data and were used instead of the raw elevation data instead. This approach thus creates two separate discrete elevational bounds in which a species might or might not exist. We specified monotonically constrained baselearners for both XGBoost and GDB[27], which are prior constraints placed on the linear and non-linear effects to follow certain directions. Here we used priors assuming either increasing, in case the habitat was highly preferred or suitable, or alternatively positive constraints when the species was just known to occur in the habitat. Previous studies have shown that the use of such monotonicity constraints in SDMs can results in more ecological plausible response functions[27]. For Bayesian Poisson-Process models we used Zellner-style spike and slab priors with two parameters, a coefficient for a Gaussian prior on the mean coefficient of the covariate and a inclusion probability which states the probability by which a certain variable is to be used and thus avoided to be regularized out from the model[28]. For habitats preferred by a species we used mean coefficients of 3 and an inclusion probability of 1, for suitable habitats we used a coefficient of 1 and 0.5 respectively and for occasionally occupied habitats we used 0.1 in both cases. Similarly, for BART models we specified priors as transition probabilities for the variable so that the regression tree is forced - with a certain probability, here 0.75 – to generate a split for a given variable.

On the full point occurrence dataset, we then applied a spatial block cross validation scheme using the blockCV R-package[18,29]. Specifically, we created three spatial folds of training and testing data to evaluate each of the three algorithms on. However for species with very few occurrence records overall (< 50) and where the creation of spatial folds failed (owing to points being too close in distance), we instead implemented default randomized folds where 25% of data was removed at random. All predictors were scaled prior to model fitting by subtracting the mean from each value and dividing by their standard deviation to ensure comparable unit scales. We included among the final predictors also the bias variable (see above), which was controlled during the prediction[9], thus helping to reduce some of the spatial biases in available occurrence datasets. Ensemble of different datasets per species were integrated and thus separate models were estimated for each spatial block and for each data type[30].

Each separate model prediction was binarized using a 0.05 percentile threshold and then validated using the withheld data to obtain an estimate of the True-Skill-Statistic (TSS). We used the TSS values to create a weighted mean ensemble of all predictions[12]. Individual predictions from different models were first normalized owing to the different units (relative rate of occurrence compared to relative occurrence probability). We then thresholded all ensembles using a 5% minimum percentile threshold on all observed data points (across datasets), thus creating a conservative estimate of where suitable habitat for a species might or might not persist in Europe[24]. We used percentile-based thresholds[31] opposed to approaches maximizing any performance metric since they can be applied across different predictions and dataset types (presence-only and presence-absence). Furthermore it assumes that the least suitable habitat at which the species is known to occur is the minimum suitability value for the species, while allowing for some flexibility so that outliers do not bias the threshold. For each species the validation statistics, the predicted suitability and the thresholded map is then retained. All SDMs were fitted using the integrated species distribution modelling framework ibis.iSDM coded for R[12].

1.1.4 Potential distribution of a species

The goal of this modelling is to obtain a depiction of the potential distribution of the species (sensu[5]). We understand potential in this context as the contemporary climatic, soil and natural vegetation conditions that would allow a species to persist in an area. Critically, and opposed to the mapping of current suitable habitat, this approach considers only contemporary differences in climate and soil, and not any land-cover or land-use, aligning with the concept of potential natural vegetation of Europe[32]. The predictions from the species distribution models used here thus aim to depict where a species might exist under contemporary conditions, while also allowing modest inter- and extrapolation from its current distribution.

While for the current estimation of species distribution (see above) we considered each biodiversity data type separately in a model, for the potential distribution of the species we merged datasets with presence-only data and presence-absence data, adding pseudo-absence points to the former[33,34]. This is a widely applied approach for SDM mapping, which however is scale dependent and can result in an overestimation of the niche[?][26], yet in this context is acceptable given that our aim is to map the widest possible potential distribution of species (although we modestly constrain the prediction, see below). Although there can be benefits in modelling these datasets jointly for more constrained predictions[?], our aim is to identify and characterize the maximum potential extent of the environmental niche of a species given contemporary conditions. We first combined all cleaned and filtered point occurrence data into one joint dataset, removing duplicates per grid cell in the process. For each presence-only dataset we created pseudo-absence points randomly distributed within the modelling domain, but spatially weighted them so that pseudo-absences preferentially fall into areas with high bias defined by human population density.

We used a similar ensemble modelling approach as for the current estimates of species ranges (see above), however used binomial distributed responses throughout, adding pseudo-absence points for presence-only datasets. For validating and thresholding the ensemble models we evaluated the predictions in terms of their accuracy through the F1 score, which is calculated as the ratio of the model precision (true positives) and the recall (also known as sensitivity). We specifically chose the F1 score for evaluation since it maximizes correct predictions and thus can help to ensure that most training occurrence points are retained. The final ensemble prediction was then created as a weighted mean of the nine different F1 scores (3 spatial blocked subsets per algorithm). All modelling was done in the integrated species distribution modelling framework ibis.iSDM coded for R[12].

1.2 Land cover data

1.2.1 Potential distribution of natural land cover

For the identification of potentially restorable land we followed the concept of potentially natural vegetation (sensu[5]). To map potential land cover, we first assembled a Europe-wide database on the distribution of habitats in Europe where we followed the thematic legend of the MAES habitat classification system at level 2, while ignoring any strictly anthropogenic habitats (e.g. Urban, Cropland, Pasture) as well as Rivers and Lakes. Different data sources on the distribution of habitats differ in terms of their geographic spread and biases. In order to not rely on any single data source of European ecosystems we integrated habitat data from three different sources collated for Europe:

We took habitat information from the European Habitat Directive which gives the occurrence of all EUNIS habitats listed in the Article 17 of the habitats directive at a 10km resolution. We used a crosswalk developed by the European Environment Agency and Biodiversity Topic Centre to translate the EUNIS types to Corine CLC and subsequently to MAES level 2 categories[35]. We further made use of point occurrence datasets for key habitats in the new EEA suitability predictions[36]. Here the habitat categories were reclassified into the respective MAES types, following the CLC to MAES crosswalk. Finally, we prepared point occurrence data from the openly available land-cover and land-use database LUCAS[37]. The LUCAS database contains stratified and repeated survey records of local land-cover and land-use types for Europe[37] with the latest one being available for the year 2018. We took the LUCAS survey records and reclassified the land-cover type (”$lc1_{l}able$”) to the natural MAES ecosystem categories, discarding all anthropogenic created habitats (Cropland, Urban, ...) in the process.

As predictors for the potential habitat modelling, we considered data on the potential distribution of land cover ([38],[32]) as well as long-term average climatic conditions in Europe where we used downscaled bioclimatic ERA5 indicators over the last 40 years (1979 to 2018) from the European Copernicus program. These climatic indicators represent Essential Climate Variables (ECV) such as the surface energy, drought or moisture all of which are known to be important factors in delineating the range and environmental niche. Specifically, we made use of the BIOCLIM data BIO01 to BIO19, average aridity and cloud cover, the annual sum of frost days, potential evaporation and volumetric soil water as well as different characterizations of the number of growing degree days and the start, end and length of the growing season. We also included a predictor that quantified the Euclidean distance to the ocean from each terrestrial grid cell, given the importance of some wetland habitats to brackish water and coastal conditions. We prepared data on groundwater and soil conditions, specifically data on the Ph value and Calcium Carbonate content of groundwater resources as well as estimates of the depth to groundwater in meters[8,39]. We also included data from a thematic layer of a European soil lithology classification system[40] owing to the importance of difference soil types. The individual lithology classes were included as factorial combinations in the modeling.

We estimated the potential distribution of the habitat by relating presence-only habitats with pseudo-absence points[33,34]. This is a widely applied approach usually for species distribution modelling, which although it can result in “overconfident” extrapolations, is in this context desirable since our aim is to map the potential natural distribution of a habitat. Although it is possible to create predictions of potential natural habitat as multi-nominal problem, e.g. where each class has a different and exclusive probability to potentially occur[5], we decided instead to estimate the distribution of the habitat separately, since in many areas of Europe there is a potential for more than one habitat to potentially occur under natural conditions especially when succession trajectories are unknown. Thus, for each habitat dataset we created an equal balanced number of pseudo-absence points randomly distributed within the modelling domain. We furthermore rasterized the 10-km Article 17 data and applied an Euclidean distance transform to them, e.g. there is a monotonically decreasing probability of a habitat type to occur outside the Article 17 reporting data. The resulting layer was then included as an additional covariate.

We used non-linear and tree-based Bayesian Regression Trees (BART) for projecting the probability of any potential habitat. The BART algorithm has the benefit of being able to quantify complex non-linear interactions between variables as well as being able to consider prior information in a Bayesian framework[41]. For the regression trees we used a logistic model formulation of the response by assuming the habitat presence and pseudo-absences to be Bernoulli distributed, e.g.

$y_{habitat}$ $Pr(y=1|x|)$.

We fitted the BART models with 500 tree and 50 burn-in iterations across four MCMC chains through the use of the ‘dbarts’ R-package[41]. From the resulting posterior of the fitted model and for each grid cell, we summarized the median and lower (5%) and upper (95%) percentile of the posterior, thus allowing us to spatially represent the uncertainty of each individual habitat type prediction. The resulting predictions thus contain an estimate of the probability of a potential occurrence of a given natural habitat for each 10km grid cell.

1.3 Carbon data

1.3.1 Current distribution of carbon

For current carbon stocks we used data on above-ground, below-ground and soil organic carbon at risk from land-use change from[31]. These data were created by selecting and integrating best available carbon maps for different vegetation classes. For more detail on the integration and handling of individual data layers see[31]. We used the carbon data at the original resolution of 100m and intersected them with the current distribution of land cover according to the Corine landcover data for 2018, which we reclassified to the MAES legend. The intersected individual carbon estimates were then aggregated (arithmetic mean) to a grain size of 10km used for the prioritization. All data are in units of tC/ha and for the analysis we combined the current carbon layers by calculating the combined sum of above- & below-ground and soil organic carbon for Europe and included it as an additional feature in the prioritization.

1.3.2 Potential carbon

To spatially allocate specific restoration priorities, we needed to identify areas with high carbon sequestration potential. Here we used an approach that combined the different techniques from[42] and[43]. First, we created a regular sampling grid at 1km for each current MAES ecosystem type reclassified from Corine 2018 (see above) and extracted the fraction of the respective land cover type. We then extracted estimates of current reference carbon data (in tC/ha) from the above, below and soil organic carbon data layers from[31] as well as from the European-specific JRC Biomass map[?]For each of the different types of carbon products (below, above and soil carbon) we then calculated a consensus estimate (arithmetic mean) per gridded 1km point. We further extracted estimates on whether a given point locality was situated in a peat land as considered by the European peatland map[?] or land covered by primary forest[44]

From the resulting extracted estimates, we then selected for each natural land cover type (e.g. Grassland, Heathland, Marine inlets, Sparsely vegetated land, Wetland and Forest and Woodland) a total of 10000 reference points for modelling training. We ensured that (a) the respective land cover type currently covers at least 50% of a given 1km grid cell, (b) average carbon density estimates are in the largest 95% percentile of values, (c) the reference points were preferentially sampled in remaining European peat and primary forest sites for the Wetland, Marine Inlets and Forest & Woodland classes, (d) points were geographically representative by covering each European biogeographical regions (adjusted for area) and (e) that extracted mean carbon density estimates were corrected for the fraction of non-natural land contained within them. Instead of using a single reference value for the carbon contained in non-natural systems[42], we calculated the average estimate of all non-natural land cover types in MAES (e.g. cropland and urban).

As above for potential species distributions and land cover, we then subjected these reference estimates to a spatial extrapolation approach. Here we followed an approach set out by[43] and estimated potential carbon density as $CDensPo = f(S,T,C)$, where potential carbon density is predicted as a function of soil, S, topographic, T, and climatic, C, factors. We used the same predictors as for potential land cover and species occurrences. Finally, we corrected each estimate by the amount of potentially occurring natural land cover.

References

1. Zizka, A. *et al.* CoordinateCleaner : Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751, DOI: 10.1111/2041-210X.13152 (2019).

2. of Ornithology, C. L. eBird Basic Dataset. Version: EBD_relmar-2021. (2021).

3. Sabatini, F. M. & et al. sPlotOpen – An environmentally balanced, open-access, global dataset of vegetation plots. DOI: 10.1111/geb.13346 (2021).

4. Donald, P. F. *et al.* Important Bird and Biodiversity Areas (IBAs): the development and characteristics of a global inventory of key sites for biodiversity. *Bird Conserv. Int.* **29**, 177–198, DOI: 10.1017/S0959270918000102 (2019).

5. Hengl, T. *et al.* Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential. *PeerJ* **6**, DOI: 10.7717/peerj.5457 (2018).
6. De Reu, J. *et al.* Application of the topographic position index to heterogeneous landscapes. *Geomorphology* **186**, 39–49, DOI: 10.1016/j.geomorph.2012.12.015 (2013).
7. Copernicus Climate Change Service. Downscaled bioclimatic indicators for selected regions from 1979 to 2018 derived from reanalysis, DOI: 10.24381/CDS.FE90A594 (2021).
8. Hájek, M. *et al.* A European map of groundwater pH and calcium. *Earth Syst. Sci. Data* **13**, 1089–1105, DOI: 10.5194/essd-13-1089-2021 (2021).
9. Warton, D. I., Renner, I. W. & Ramp, D. Model-Based Control of Observer Bias for the Analysis of Presence-Only Data in Ecology. *PLoS ONE* **8**, e79168, DOI: 10.1371/journal.pone.0079168 (2013).
10. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336, DOI: 10.1038/nature25181 (2018).
11. Pesaresi, M. *et al.* A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **6**, 2102–2131, DOI: 10.1109/JSTARS.2013.2271445 (2013).
12. Jung, M. An integrated species distribution modelling framework for heterogeneous biodiversity data. *Ecol. Informatics* **76**, 102127, DOI: 10.1016/j.ecoinf.2023.102127 (2023).
13. Steen, V. A., Tingley, M. W., Paton, P. W. C. & Elphick, C. S. Spatial thinning and class balancing: Key choices lead to variation in the performance of species distribution models with citizen science data. *Methods Ecol. Evol.* **12**, 216–226, DOI: 10.1111/2041-210X.13525 (2021).
14. Sillero, N. *et al.* Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-Reptilia* **35**, 1–31, DOI: 10.1163/15685381-00002935 (2014).
15. Merow, C., Allen, J. M., Aiello-Lammens, M., Silander, J. A. & Fortin, M. Improving niche and range estimates with Maxent and point process models by integrating spatially explicit information. *Glob. Ecol. Biogeogr.* **25**, 1022–1036, DOI: 10.1111/geb.12453 (2016).
16. Domisch, S., Wilson, A. M. & Jetz, W. Model-based integration of observed and expert-based information for assessing the geographic and environmental distribution of freshwater species. *Ecography* **39**, 1078–1088, DOI: 10.1111/ecog.01925 (2016).
17. Biber, M. F., Voskamp, A., Niamir, A., Hickler, T. & Hof, C. A comparison of macroecological and stacked species distribution models to predict future global terrestrial vertebrate richness. *J. Biogeogr.* **47**, 114–129, DOI: 10.1111/jbi.13696 (2020).
18. Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J. & Elith, J. Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecol. Monogr.* **92**, DOI: 10.1002/ecm.1486 (2022).
19. Chen, Y. *et al.* Comparison of feature selection methods for mapping soil organic matter in subtropical restored forests. *Ecol. Indic.* **135**, 108545, DOI: 10.1016/j.ecolind.2022.108545 (2022).
20. Hofner, B., Mayr, A., Robinzonov, N. & Schmid, M. Model-based boosting in R: a hands-on tutorial using the R package mboost. *Comput. Stat.* **29**, 3–35, DOI: 10.1007/s00180-012-0382-5 (2014).
21. Scott, S. L. BoomSpikeSlab: MCMC for Spike and Slab Regression (2022).
22. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, DOI: 10.18637/jss.v033.i01 (2010).
23. Phillips, S. J. & Dudík, M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161–175, DOI: 10.1111/j.0906-7590.2008.5203.x (2008).
24. Merow, C., Smith, M. J. & Silander, J. A. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069, DOI: 10.1111/j.1600-0587.2013.07872.x (2013).
25. Agency, E. E. Linkages of species and habitat types to MAES ecosystems (2017).
26. Hannemann, H., Willis, K. J. & Macias-Fauria, M. The devil is in the detail: unstable response functions in species distribution models challenge bulk ensemble modelling: Unstable response functions in SDMs. *Glob. Ecol. Biogeogr.* **25**, 26–35, DOI: 10.1111/geb.12381 (2016).
27. Hofner, B., Hothorn, T., Kneib, T. & Schmid, M. A Framework for Unbiased Model Selection Based on Boosting. *J. Comput. Graph. Stat.* **20**, 956–971, DOI: 10.1198/jcgs.2011.09220 (2011).

28. Cui, W. & George, E. I. Empirical Bayes vs. fully Bayes variable selection. *J. Stat. Plan. Inference* **138**, 888–900, DOI: 10.1016/j.jspi.2007.02.011 (2008).

29. Meyer, H., Reudenbach, C., Wöllauer, S. & Nauss, T. Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. *Ecol. Model.* **411**, 108815, DOI: 10.1016/j.ecolmodel.2019.108815 (2019).

30. Fletcher, R. J. *et al.* A practical guide for combining data to model species distributions. *Ecology* e02710, DOI: 10.1002/ecy.2710 (2019).

31. Jung, M. *et al.* Areas of global importance for conserving terrestrial biodiversity, carbon and water. *Nat. Ecol. & Evol.* **5**, 1499–1509, DOI: 10.1038/s41559-021-01528-7 (2021).

32. Bohn, U. Map of the Natural Vegetation of Europe. Tech. Rep. (2004).

33. Guisan, A. & Thuiller, W. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009, DOI: 10.1111/j.1461-0248.2005.00792.x (2005).

34. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many?: *How to use pseudo-absences in niche modelling? Methods Ecol. Evol.* **3**, 327–338, DOI: 10.1111/j.2041-210X.2011.00172.x (2012).

35. on Biological Diversity, E. T. C. Crosswalk between EUNIS habitats classification and Corine land cover.

36. Agency, E. E. EUNIS habitat classification (2019).

37. d'Andrimont, R. *et al.* Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union. *Sci. Data* **7**, 352, DOI: 10.1038/s41597-020-00675-z (2020).

38. Hengl, T., Jung, M. & Visconti, P. Potential distribution of land cover classes (Potential Natural Vegetation) at 250 m spatial resolution, DOI: 10.5281/ZENODO.3631254 (2020).

39. Fan, Y., Li, H. & Miguez-Macho, G. Global Patterns of Groundwater Table Depth. *Science* **339**, 940–943, DOI: 10.1126/science.1229881 (2013).

40. European Commission. Joint Research Centre. Institute for Environment and Sustainability. *Soils of the European Union.* (Publications Office, LU, 2008).

41. Carlson, C. J. embarcadero: Species distribution modelling with Bayesian additive regression trees in r. *Methods Ecol. Evol.* **11**, 850–858, DOI: 10.1111/2041-210X.13389 (2020).

42. Strassburg, B. B. N. *et al.* Global priority areas for ecosystem restoration. *Nature* **586**, 724–729, DOI: 10.1038/s41586-020-2784-9 (2020).

43. Walker, W. S. *et al.* The global potential for increased storage of carbon on land. *Proc. Natl. Acad. Sci.* **119**, e2111312119, DOI: 10.1073/pnas.2111312119 (2022).

44. Sabatini, F. M. *et al.* European primary forest database v2.0. *Sci. Data* **8**, 220, DOI: 10.1038/s41597-021-00988-7 (2021).