Check for updates

OPEN

Improving biodiversity protection through artificial intelligence

Daniele Silvestro [1,2,3,8] ✉, Stefano Goria[4], Thomas Sterner [5] and Alexandre Antonelli [3,6,7,8] ✉

Over a million species face extinction, highlighting the urgent need for conservation policies that maximize the protection of biodiversity to sustain its manifold contributions to people's lives. Here we present a novel framework for spatial conservation prioritization based on reinforcement learning that consistently outperforms available state-of-the-art software using simulated and empirical data. Our methodology, conservation area prioritization through artificial intelligence (CAPTAIN), quantifies the trade-off between the costs and benefits of area and biodiversity protection, allowing the exploration of multiple biodiversity metrics. Under a limited budget, our model protects significantly more species from extinction than areas selected randomly or naively (such as based on species richness). CAPTAIN achieves substantially better solutions with empirical data than alternative software, meeting conservation targets more reliably and generating more interpretable prioritization maps. Regular biodiversity monitoring, even with a degree of inaccuracy characteristic of citizen science surveys, further improves biodiversity outcomes. Artificial intelligence holds great promise for improving the conservation and sustainable use of biological and ecosystem values in a rapidly changing and resource-limited world.

Biodiversity is the variety of all life on Earth, from genes to populations, species, functions and ecosystems. Alongside its own intrinsic value and ecological roles, biodiversity provides us with clean water, pollination services, building materials, clothing, food and medicine, among many other physical and cultural contributions that species make to ecosystem services and people's lives[1,2]. The contradiction is that our endeavours to maximize short-term benefits have become unsustainable, depleting biodiversity and threatening the life-sustaining foundations of humanity in the long term[3] (Supplementary Box 1). This can help explain why, despite the risks, we are living in an age of mass extinction[4,5]. The imperative to feed and house rapidly growing human populations, with an estimated 2.4 billion more people by 2050, together with increasing disruptions from climate change, will put tremendous pressure on the world's last remaining native ecosystems and the species they contain. Because not a single one of the 20 Aichi Biodiversity Targets agreed by 196 nations for the period 2011–2020 has been fully met[6], there is now an urgent need to design more realistic and effective policies for a sustainable future[7] that help deliver the conservation targets under the post-2020 Global Biodiversity Framework, the focus of the 15th Conference of the Parties in 2022.

There have been several theoretical and practical frameworks underlying biological conservation[8] since the 1960s. The field was initially focused on the conservation of nature for itself, without human interference, but gradually incorporated the bidirectional links to people, recognizing our ubiquitous influence on nature and the multifaceted contributions we derive from it, including the sustainable use of species[1,8,9]. Throughout this progress, a critical step has been the identification of priority areas for targeted protection, restoration planning, impact avoidance and loss minimization, triggering the development of the fields of spatial conservation prioritization and systematic conservation planning[10–16]. While humans and wild species are increasingly sharing the same space[17], the preservation of largely intact nature remains critical for safeguarding many species and ecosystems, such as tropical rainforests.

Several tools and algorithms have been designed to facilitate systematic conservation planning[18]. They often allow the exploration and optimization of trade-offs between variables, something not readily available in Geographic Information Systems[19], which can lead to substantial economic, social and environmental gains[20]. While the initial focus has been on maximizing the protection of species while minimizing costs, additional parameters can sometimes be modelled, such as species rarity and threat, total protected area and evolutionary diversity[18,21,22]. The most widely used method so far, Marxan[23], seeks to identify a set of protected areas that collectively allow particular conservation targets to be met under minimal costs, using a simulated annealing optimization algorithm. Despite its usefulness and popularity, Marxan and similar methods[18] were designed to optimize a one-time policy, do not directly incorporate changes through time, and assume a single initial gathering of biodiversity and cost data (although temporal aspects can be explored by manually updating and re-running the models under various targets[24]). In addition, the optimized conservation planning does not explicitly incorporate climate change, variation in anthropogenic pressure (although varying threat probabilities are dealt with in recent software extensions of Marxan[25,26]), or species-specific sensitivities to such changes.

In this study we have tackled the challenge of optimizing biodiversity protection in a complex and rapidly evolving world by harnessing the power of artificial intelligence (AI). We have developed an entirely novel tool for systematic conservation planning (Fig. 1) that optimizes a conservation policy based on static or dynamic biodiversity monitoring towards user-defined targets (such as minimizing species loss) and within the constraints of a limited financial budget, and used it to explore, through simulations and empirical analyses, multiple previously identified trade-offs in

¹Department of Biology, University of Fribourg, Fribourg, Switzerland. ²Swiss Institute of Bioinformatics, Fribourg, Switzerland. ³Gothenburg Global Biodiversity Centre, Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. ⁴Thymia, London, UK. ⁵Department of Economics, University of Gothenburg, Gothenburg, Sweden. ⁶Department of Plant Sciences, University of Oxford, Oxford, UK. ⁷Royal Botanic Gardens, Kew, UK. ⁸These authors contributed equally: Daniele Silvestro, Alexandre Antonelli. ✉e-mail: daniele.silvestro@unifr.ch; a.antonelli@kew.org
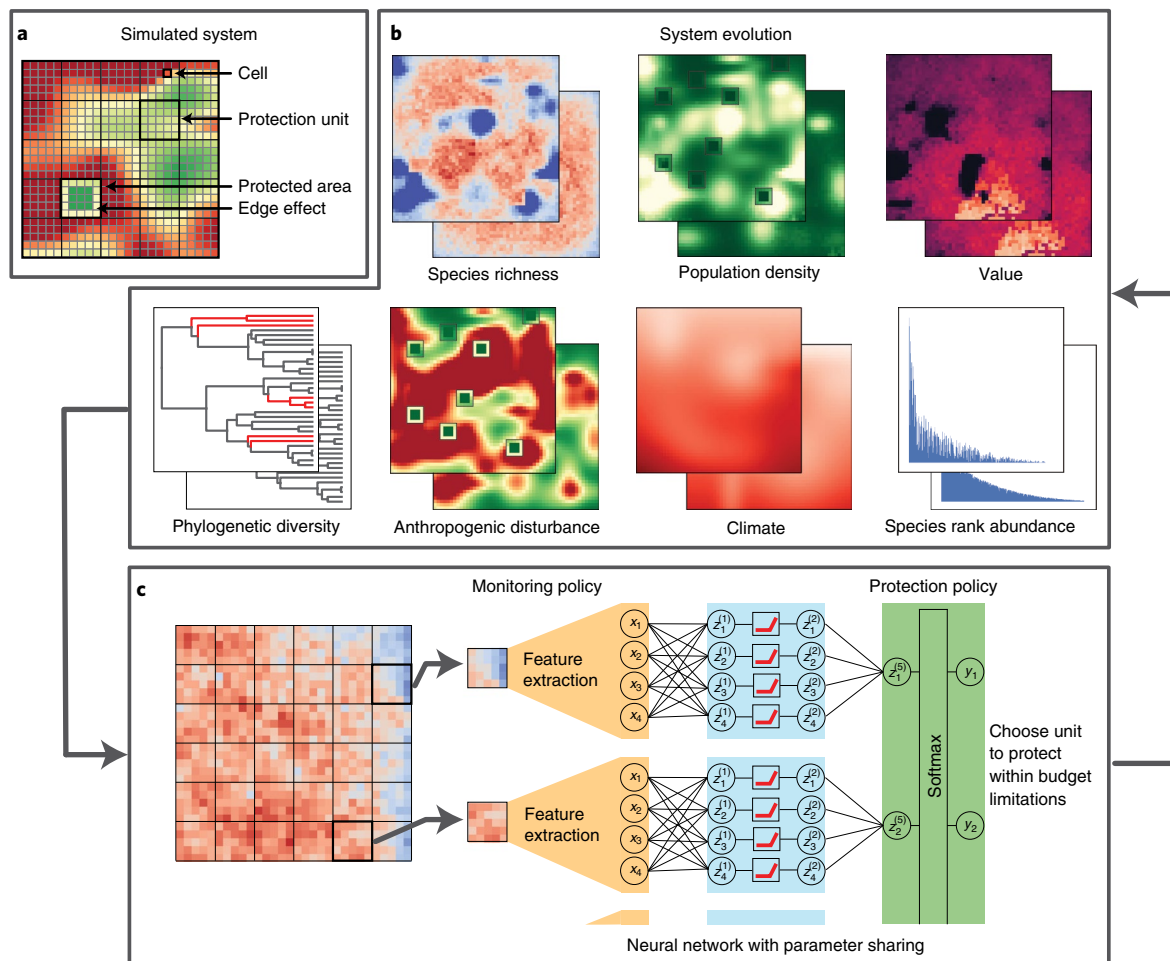
Fig. 1 | The CAPTAIN reinforcement learning framework. a, A simulated system, which could be equivalent to a country an island or a large coral reef, consists of a number of cells, each with a number of individuals of various species. Once a protection unit is identified and protected, its human-driven disturbance (for example, forest logging or sea trawling) will immediately reduce to an arbitrarily low level, except for the well-known edge effect[47] characterized by intermediate levels of disturbance. All simulation settings are provided with initial default values but are fully customizable (Supplementary Tables 1 and 2). Simulated systems evolve through time and are used to optimize a conservation policy using RL. After training the model, the optimized policy can be used to evaluate the model performance based on simulated or empirical data. Using empirical data, the simulated system is replaced with available biodiversity and disturbance data. **b,c**, Analysis flowchart integrating system evolution (**b**) with simulations and AI modules (**c**) to maximize selected outcomes (for example, species richness). The system evolves between two points in time, with several time-dependent variables considered (seven plotted here): species richness, population density, economic value, phylogenetic diversity, anthropogenic disturbance, climate and species rank abundance (see www.captain-project.net for animations depicting these and additional variables). Biodiversity features (species presence per protection unit at a minimum, plus their abundance under full monitoring schemes as defined here; see Methods and Supplementary Box 2 for advances in data gathering approaches) are extracted from the system at regular steps, and are then fed into a neural network that learns from the system's evolution to identify conservation policies that maximize a reward, such as protection of the maximum species diversity within a fixed budget. The vectors of parameters *x*, *z* and *y* represent the nodes of the input, hidden layers and output of the neural network, respectively.

real-world conservation and to evaluate the impact of data gathering on specific outcomes[27]. We have also explored the impact of species-specific sensitivity to geographically varying local disturbances (for example, as a consequence of new roads, mining, trawling or other forms of unsustainable economic activity with negative impacts on natural ecosystems) and climate change (overall temperature increases as well as short-term variations to reflect extreme weather events). We name our framework CAPTAIN, denoting Conservation Area Prioritization Through Artificial INtelligence.

Within AI, we implemented a reinforcement learning (RL) framework based on a spatially explicit simulation of biodiversity and its evolution through time in response to anthropogenic pressure and climate change. The RL algorithm is designed to find an optimal balance between data generation (learning from the

current state of a system, also termed 'exploration') and action (called 'exploitation', the effect of which is quantified by the outcome, also termed 'reward'). Our platform enables us to assess the influence of model assumptions on the reward, mimicking the use of counterfactual analyses[22]. CAPTAIN can optimize a static policy, where all the budget is spent at once, or (more in line with its primary objective) a conservation policy that develops over time, thus being particularly suitable for designing policies and testing their short- and long-term effects. Actions are decided based on the state of the system through a neural network, whose parameters are optimized within the RL framework to maximize the reward. Once a model is trained through RL, it can be used to identify conservation priorities in space and time using simulated or empirical data.
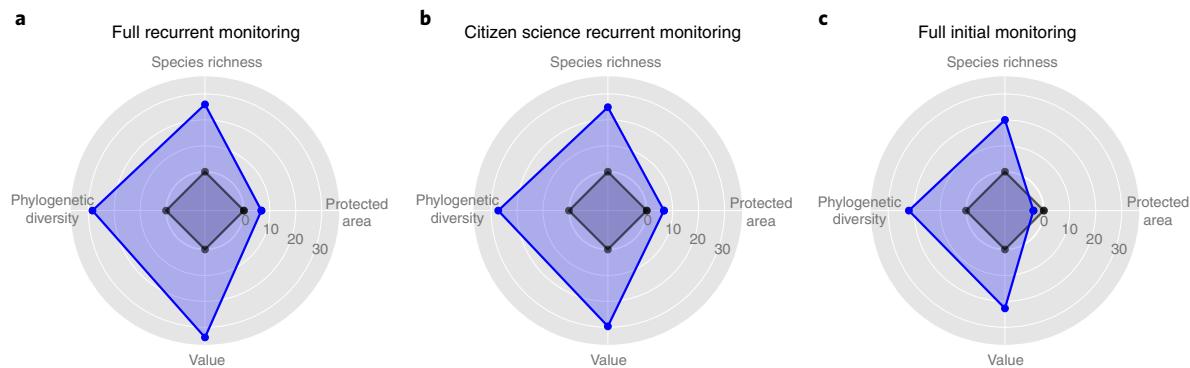
Fig. 2 | Impact of monitoring strategies on biodiversity protection. a–c, Outcome of policies designed to minimize species loss based on different monitoring strategies: full recurrent monitoring (of species presence and abundance at each time step; **a**), citizen science recurrent monitoring (limited to species presence/absence with some error at each time step; **b**) and full initial monitoring (species presence and abundance only at the initial time; **c**). The results show the percentage change in species loss, total protected area, accumulated species value and phylogenetic diversity between a random protection policy (black polygons) and models optimized by CAPTAIN (blue polygons). All results are averaged across 250 simulations, with more details shown in Supplementary Table 3. Each simulation was based on the same budget and resolution of the protection units (5×5 cells) but differed in their initial natural system (species distributions, abundances, tolerances and phylogenetic relationships) and in the dynamics of climate change and disturbance patterns.

Although AI solutions have been previously proposed and to some extent are already used in conservation science[28,29], to our knowledge RL has only been advocated[30] and not yet implemented in practical conservation tools. In particular, CAPTAIN aims to tackle multidimensional problems of loss minimization considered by techniques such as stochastic dynamic programming but proven thus far intractable for large systems[13]. It thus fills an important space in conservation in a dynamic world[31], characterized by heterogeneous and often unpredictable habitat loss[14], which require iterative and regular conservation interventions.

We have used CAPTAIN to address the following questions: (1) What role does the data-gathering strategy have in effective conservation? (2) What trade-offs arise depending on the optimized variable, such as species richness, economic value or total area protected? (3) What can the simulation framework reveal in terms of winners and losers, that is, which traits characterize the species and areas protected over time? (4) How does our framework perform compared with the state-of-the-art model for conservation planning, Marxan[23]? Finally, we demonstrate here the usefulness of our framework and direct applicability of models trained through RL on an empirical dataset of endemic trees of Madagascar.

Results

Impact of data gathering strategy. Using CAPTAIN we found that full recurrent monitoring (where the system is monitored at each time step, including species presence and abundance) results in the smallest species loss: it succeeds in protecting on average 26% more species than a random protection policy (Fig. 2a and Supplementary Table 3). A very similar outcome (24.9% improvement) is generated by the citizen science recurrent monitoring strategy (where only presence/absence of species are recorded in each cell), with a degree of error characteristic of citizen science efforts (Fig. 2b and Methods). These two monitoring strategies outperform a full initial monitoring with no error, which only saves from extinction an average of 20% more species than a random policy (Fig. 2c and Supplementary Table 3).

To thoroughly explore the parameter space of the simulations, each system was initialized with different species composition and distributions and different anthropogenic pressure and climate change patterns (Supplementary Figs. 1–4). Because of this stochasticity, the reliability of the protection policies in relation to species

loss varies across simulations. The policies based on full recurrent monitoring and citizen science recurrent monitoring are the most reliable, outperforming the baseline random policy in 97.2% of the simulations. Both these policies are more reliable than the full initial monitoring, which in addition to protecting fewer species on average (Fig. 2) also results in a slightly lower reliability of the outcome, outperforming the random policy in 91.2% of the simulations.

Optimization trade-offs. The policy objective, which determines the optimality criterion in our RL framework, strongly influences the outcome of the simulations. A policy minimizing species loss based on their commercial value (such as timber price) tends to sacrifice more species to prioritize the protection of fewer, highly valuable ones. This policy, while efficiently reducing the loss of cumulative value, decreases species losses by only 10.9% compared with the random baseline (Supplementary Table 3). Thus, a policy targeting exclusively the preservation of species with high economic value may have a strongly negative impact on the total protected species richness, phylogenetic diversity and even amount of protected area compared with a policy minimizing species loss (Fig. 3a).

A policy that maximizes protected area results in a 27.6% increase in the number of protected cells by selecting those cheapest to buy; however, it leads to substantial losses in species numbers, value and phylogenetic diversity, which are considerably worse than the random baseline, with 13.6% more species losses on average (Supplementary Table 3). The decreased performance in terms of preventing extinctions is even more pronounced when compared with a policy minimizing species loss (Fig. 3b).

As expected, the reliability for optimizations based on economic value and total protected area is high for the respective policy objectives, but they result in highly inconsistent outcomes in terms of preventing species extinctions, with biodiversity losses not significantly different from those of the random baseline policy (Supplementary Table 3).

Winners and losers. Focusing on the policy developed under full recurrent monitoring and optimized on reducing species loss, we explored the properties of species that survived in comparison with those that went extinct, despite optimal area protection. Species that went extinct are characterized by relatively small initial ranges,
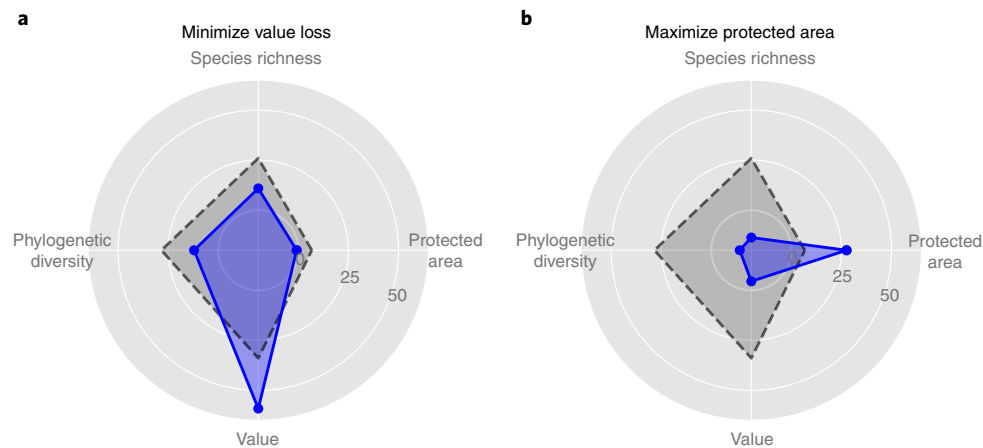
Fig. 3 | Trade-offs in conservation outcomes in relation to policy objectives. a,b, The impact of different policy objectives based on full recurrent monitoring: the policies were designed to minimize value loss (**a**) and maximize the amount of protected area (**b**). The plots show the outcome averaged across 250 simulations (blue polygons). The radial axis shows the percentage change compared with the baseline random policy, and the dashed grey polygons show the outcome of a policy with full recurrent monitoring optimized to minimize biodiversity loss (Fig. 2a).
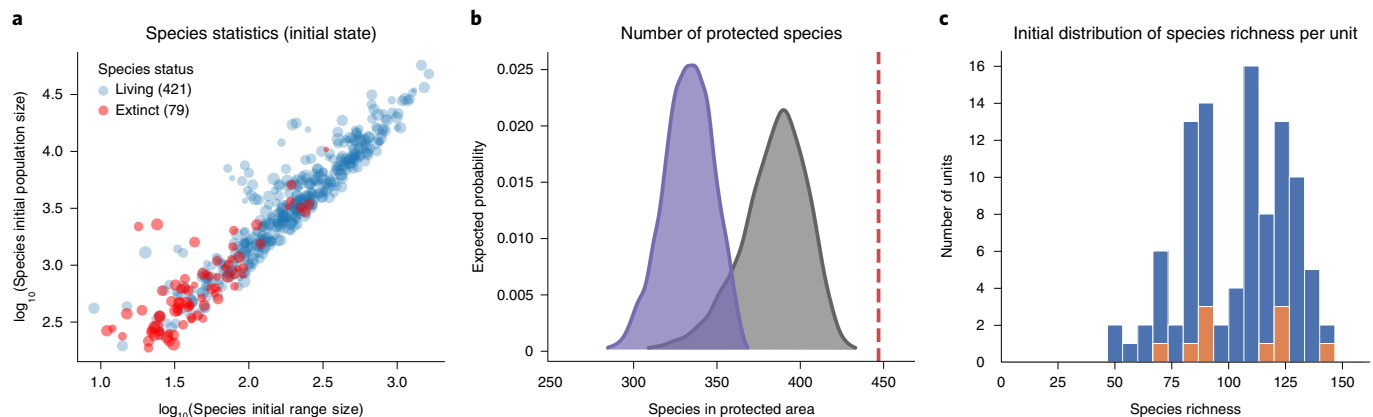
Fig. 4 | Summary statistics for one simulation optimized to reduce species loss and informed through full recurrent monitoring. a, Living (or surviving) and locally extinct species after a simulation of 30 time steps with increasing disturbance and climate change. The x and y axes show the initial range and population sizes of the species ($\log_{10}$ transformed), respectively. The size of the circles is proportional to the resilience of each species to anthropogenic disturbance, with smaller circles representing more sensitive species. **b,** Cumulative number of species encompassed in the ten protected units (5×5 cells) selected on the basis of a policy optimized to minimize species loss. The grey density plot shows the expected distribution from 10,000 random draws, and the purple shaded area shows the expected distribution when protected units are selected 'naively' (here, randomly chosen from among the top 20 most diverse units). The dashed red line indicates the number of species included in the units selected by the optimized CAPTAIN policy, which is higher than in all the random draws. The optimized policy learned to maximize the total number of species included in protected units, thus accounting for their complementarity. Note that fewer species survived (421) in this simulation compared with how many were included in protected areas (447). This discrepancy is due to the effect of climate change, in which area protection does not play a role[47] (see also www.captain-project.net). **c,** Species richness across the 100 protection units included in the area (blue), ten of which were selected to be protected (orange). The plot shows that the protection policy does not exclusively target units with the highest diversity.

small populations and intermediate or low resilience to disturbance (Fig. 4a). In contrast, species that survived have either low resilience but widespread ranges and high population sizes, or high resilience with small ranges and population sizes.

We further assessed what characterizes the grid cells that are selected for protection by the optimized policy. The cumulative number of species included in these cells is significantly higher than the cumulative species richness across a random set of cells of equal area (Fig. 4b). Thus, the model learns to protect a diversity of species assemblages to minimize species loss. Interestingly, the cells selected for protection did not include only areas with the highest species richness (Fig. 4c).

Benchmarking through simulations. We evaluated our simulation framework by comparing its performance in optimizing policies with that of the current state-of-the-art tool for conservation prioritization, Marxan[23]. The methods differ conceptually in that while CAPTAIN is explicitly designed to minimize loss (for example, local species extinction) within the constraints of a limited budget, Marxan's default algorithms minimize the cost of reaching a conservation target (for example, protecting at least 10% of all species ranges). Additionally, Marxan is typically used to optimize the placement of protected units in a single step, while CAPTAIN places the protection units across different time steps.
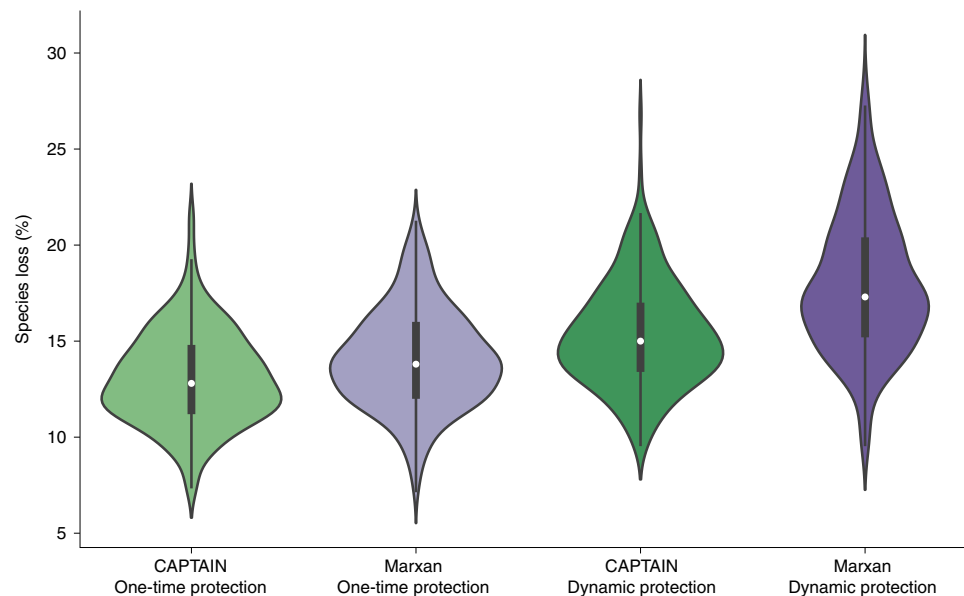
Fig. 5 | Benchmarking of CAPTAIN through simulations. The violin plots show the distribution of species loss outcomes across 250 simulations run under different monitoring policies based on the CAPTAIN framework developed here and on Marxan[23]. Species loss is expressed as a percentage of the total initial number of species in the simulated systems. See text for details of the analyses. The white circle in each box represents the median, the top and bottom of each box represent the 75th and 25th percentiles, respectively, and the whiskers are extended to 1.5× the interquartile range.

To compare the two models, we set up all Marxan analyses with an explicit budget constraint, following other tailored implementations[15,32] (see Methods). In a first comparison, we tested a protection policy in which all protection units (within a predefined budget) are established in one step. To this end, we trained an additional model in CAPTAIN based on a full initial monitoring and on a policy in which all budget for protection is spent in one step. The analysis of 250 simulations showed that CAPTAIN outperforms Marxan in 64% of the cases with an average improvement in terms of prevented species loss of 9.2% (Fig. 5).

In a second comparison, we used CAPTAIN with full recurrent monitoring and allowed the establishment of a single protection unit per time step for both programs (see Methods). Under this condition, our model outperforms Marxan in 77.2% of the simulations with an average reduction of species loss of 18.5% (Fig. 5).

Empirical applications. To demonstrate the applicability of our framework and its scalability to large, real-world tasks, we analysed a Madagascar biodiversity dataset recently used in a systematic conservation planning experiment[33] under Marxan[23]. The dataset included 22,394 protection units (5×5 km) and presence/absence data for 1,517 endemic tree species. The cost of area protection was set proportional to anthropogenic disturbance across cells, as in the original publication[33] (see Methods for more details and Supplementary Fig. 5).

We analysed the data assuming full initial monitoring in a static setting in which all protection units were placed in one step. We limited the budget to an amount that allows the protection of at most 10% of the units (or fewer if expensive units are chosen) and set the target of preserving at least 10% of the species' potential range within protected units. We repeated the Marxan analyses with a boundary length multiplier (BLM; which penalizes the placement of many isolated protection units in favour of larger contiguous areas; BLM = 0.1, as in ref. [33]) and without it (BLM = 0 for comparability with CAPTAIN, which does not include this feature).

The solutions found in CAPTAIN consistently outperform those obtained with Marxan. Within the budget constraints, CAPTAIN

solutions meet the target of protecting 10% of the range for all species in 68% of the replicates, whereas only up to 2% of the Marxan results reach that target (Supplementary Table 4). Additionally, with CAPTAIN, a median of 22% of each species range is found within protected units, well above the set target of 10% and the 14% median protected range achieved with Marxan (Fig. 6c,d and Supplementary Fig. 6c,d). Importantly, CAPTAIN is able to identify priority areas for conservation at higher and therefore more interpretable spatial resolution (Fig. 6b and Supplementary Fig. 6b).

Discussion

We have presented here a new framework to optimize dynamic conservation policies using RL and evaluate their biodiversity outcome through simulations.

Data gathering and monitoring. Our finding that even simple data (presence/absence of species) are sufficient to inform effective policies (Fig. 2 and Supplementary Table 3) is noteworthy because the information required is already available for many regions and taxonomic groups, and could be further complemented by modern technologies such as remote sensing and environmental DNA, and for accessible locations also citizen science[34] in cost-efficient ways (Supplementary Box 2).

The reason why single biodiversity assessments and area protection are often suboptimal is that they ignore the temporal dynamics caused by disturbances, population and range changes of species, all of which are likely to change through time in real-world situations. Although some systems may remain largely static over decades (for example, tree species in old-growth forests), others may change drastically (for example, alpine meadows or shallow-sea communities, where species shift their ranges rapidly in response to climatic and anthropogenic pressures); all such parameters can be tuned in our simulated system and accounted for in training the models through RL. Because current methodologies for systematic conservation planning are static, relying on a similar initial data gathering as modelled here, their recommendations for area protection may be less reliable.
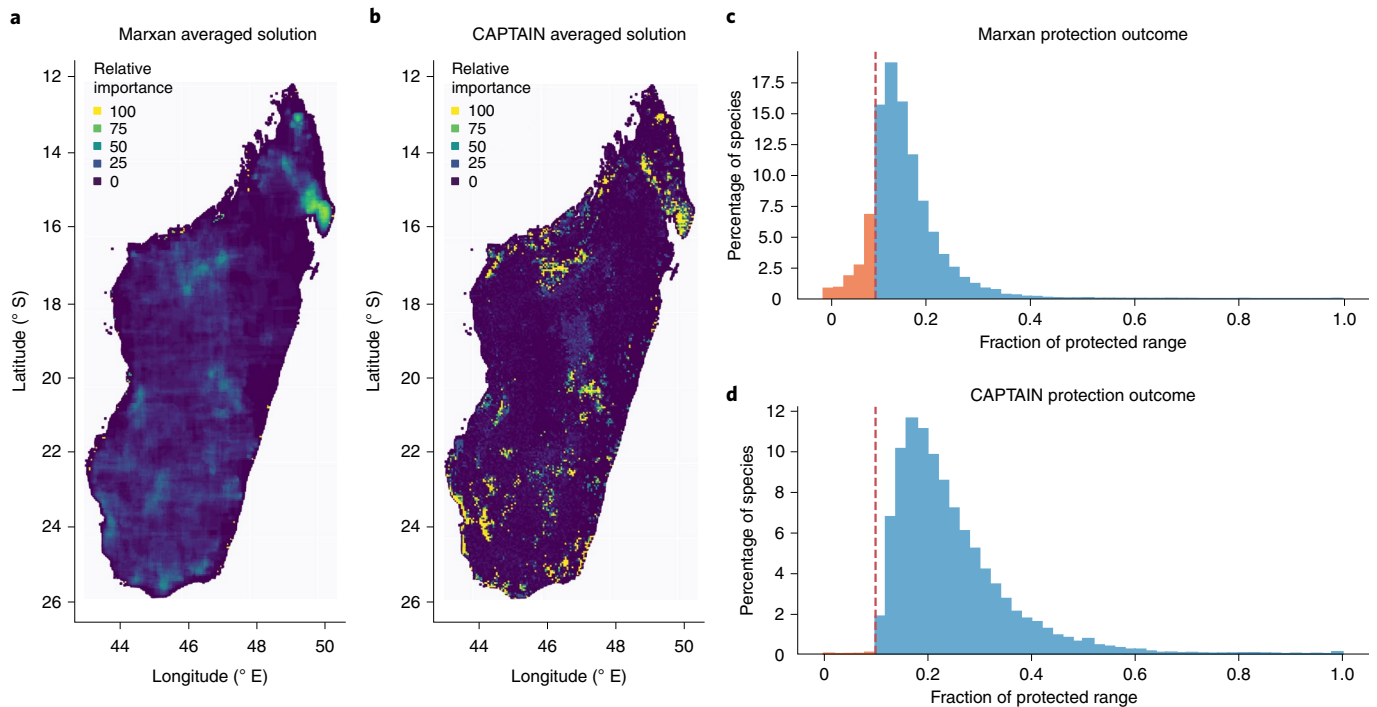
Fig. 6 | Empirical validation of CAPTAIN. a,b, Maps showing the ranking of priority areas for protection across Madagascar based on the distribution of endemic trees and on a limited budget allowing for up to 10% of protected area overall: protection units identified through Marxan optimizations (with BLM = 0.1, as in Carrasco et al.[33]; **a**) and equivalent CAPTAIN results based on a full initial monitoring policy (**b**). The colour scale represents the relative importance of each protected unit, measured as the frequency at which it was included in the optimized solution. **c,d,** Histograms of the fraction of species ranges included in the protected units calculated for each species with Marxan (**c**) and CAPTAIN (**d**). The red dashed lines indicate the 10% threshold set as the target of the policy. Orange bars show species whose protection did not meet the target.

Optimization trade-offs. Our results indicate clear trade-offs, meaning that optimizing one value can be at the cost of another (Fig. 3 and Supplementary Table 3). In particular, our finding that maximizing total protected area can lead to substantial species loss is of urgent relevance, given that total protected area has been at the core of previous international targets for biodiversity (such as the Aichi Biodiversity Targets, https://www.cbd.int/sp/targets) and remains a key focus under the new post-2020 Global Biodiversity Framework under the Convention on Biological Diversity. Focusing on quantity (area protected) rather than quality (actual biodiversity protected) could inadvertently support political pressure for 'residual' reservation[35,36], that is, the selection of new protected areas on land and at sea that are unsuitable for extractive activities, which may reduce costs and risk of conflicts, but are likely suboptimal for biodiversity conservation. Our trade-off analyses imply that economic value and total protected area should not be used as surrogates for biodiversity protection.

Learning from the models. Examination of our results reveals that, perhaps contrary to intuition, protected areas should not be primarily chosen based on high species richness (a 'naive' conservation target; Fig. 4). Instead, the simulations indicate that protected cells should span a range of areas with intermediate to high species richness, reflecting known differences between ecosystems or across environmental gradients. Such selection is more likely to increase protection complementarity for multiple species, a key factor incorporated in our software and some others[10,23,37].

Applications and prospects. Our successful benchmarking against random, naive and Marxan-optimized solutions indicates that CAPTAIN has potential as a useful tool for informing

on-the-ground decisions by landowners and policymakers. Models trained through simulations in CAPTAIN can be readily applied to available empirical datasets.

In our experiments, CAPTAIN solutions outperform Marxan, even when based on the same input data, as in the example of Malagasy trees. Our simulations show that further improvement is expected when additional data describing the state of the system are used, and when the protection policy is developed over time rather than in a single step. These findings indicate that our AI parametric approach can (1) more efficiently use the available information on species distribution and (2) more easily integrate multidimensional and time-varying biodiversity data. As the number of standardized high-resolution biological datasets is increasing (for example, see ref. [38]), as a result of new and cost-effective monitoring technologies (Supplementary Box 2), our approach offers a future-proof tool for research, conservation and the sustainable use of natural resources. Our model can be easily expanded and adapted to almost any empirical dataset and to incorporate additional variables, such as functional diversity and more sophisticated measures of economic value. Similarly, the flexibility of our AI approach allows for the design of custom policy objectives, such as optimizing carbon sequestration and storage.

In contrast to many short-lived decisions by governments, the selection of which areas in a country's territory should be protected will have long-term repercussions. Protecting the right areas, or developing sustainable models of using biodiversity without putting species at risk, will help safeguard natural assets and their contributions for the future. Choosing suboptimal areas for protection, by contrast, could not only waste public funding, but also lead to the loss of species, phylogenetic diversity, socioeconomic value and ecological functions. AI techniques should not replace human

judgement, and ultimately investment decisions will be based on more than just the parameters implemented in our models, including careful consideration of people's manifold interactions with nature[1,8]. It is also crucial to recognize the importance of ensuring the right conditions required for effective conservation of protected areas in the long term[39,40]. However, it is now time to acknowledge that the sheer complexity of sociobiological systems, multiplied by the increasing disturbances in a changing world, cannot be fully grasped by the human mind. As we progress in what many are calling the most decisive decade for nature[9,41], we must take advantage of powerful tools that help us steward the planet's remaining ecosystems in sustainable ways—for the benefit of people and all life on Earth.

Methods

A biodiversity simulation framework. We have developed a simulation framework modelling biodiversity loss to optimize and validate conservation policies (in this context, decisions about data gathering and area protection across a landscape) using an RL algorithm. We implemented a spatially explicit individual-based simulation to assess future biodiversity changes based on natural processes of mortality, replacement and dispersal. Our framework also incorporates anthropogenic processes such as habitat modifications, selective removal of a species, rapid climate change and existing conservation efforts. The simulation can include thousands of species and millions of individuals and track population sizes and species distributions and how they are affected by anthropogenic activity and climate change (for a detailed description of the model and its parameters see Supplementary Methods and Supplementary Table 1).

In our model, anthropogenic disturbance has the effect of altering the natural mortality rates on a species-specific level, which depends on the sensitivity of the species. It also affects the total number of individuals (the carrying capacity) of any species that can inhabit a spatial unit. Because sensitivity to disturbance differs among species, the relative abundance of species in each cell changes after adding disturbance and upon reaching the new equilibrium. The effect of climate change is modelled as locally affecting the mortality of individuals based on species-specific climatic tolerances. As a result, more tolerant or warmer-adapted species will tend to replace sensitive species in a warming environment, thus inducing range shifts, contraction or expansion across species depending on their climatic tolerance and dispersal ability.

We use time-forward simulations of biodiversity in time and space, with increasing anthropogenic disturbance through time, to optimize conservation policies and assess their performance. Along with a representation of the natural and anthropogenic evolution of the system, our framework includes an agent (that is, the policy maker) taking two types of actions: (1) monitoring, which provides information about the current state of biodiversity of the system, and (2) protecting, which uses that information to select areas for protection from anthropogenic disturbance. The monitoring policy defines the level of detail and temporal resolution of biodiversity surveys. At a minimal level, these include species lists for each cell, whereas more detailed surveys provide counts of population size for each species. The protection policy is informed by the results of monitoring and selects protected areas in which further anthropogenic disturbance is maintained at an arbitrarily low value (Fig. 1). Because the total number of areas that can be protected is limited by a finite budget, we use an RL algorithm[42] to optimize how to perform the protecting actions based on the information provided by monitoring, such that it minimizes species loss or other criteria depending on the policy.

We provide a full description of the simulation system in the Supplementary Methods. In the sections below we present the optimization algorithm, describe the experiments carried out to validate our framework and demonstrate its use with an empirical dataset.

Conservation planning within a reinforcement learning framework. In our model we use RL to optimize a conservation policy under a predefined policy objective (for example, to minimize the loss of biodiversity or maximize the extent of protected area). The CAPTAIN framework includes a space of actions, namely monitoring and protecting, that are optimized to maximize a reward R. The reward defines the optimality criterion of the simulation and can be quantified as the cumulative value of species that do not go extinct throughout the timeframe evaluated in the simulation. If the value is set equal across all species, the RL algorithm will minimize overall species extinctions. However, different definitions of value can be used to minimize loss based on evolutionary distinctiveness of species (for example, minimizing phylogenetic diversity loss), or their ecosystem or economic value. Alternatively, the reward can be set equal to the amount of protected area, in which case the RL algorithm maximizes the number of cells protected from disturbance, regardless of which species occur there. The amount of area that can be protected through the protecting action is determined by a budget B_t and by the cost of protection C_t^c, which can vary across cells c and through time t.

The granularity of monitoring and protecting actions is based on spatial units that may include one or more cells and which we define as the protection units. In our system, protection units are adjacent, non-overlapping areas of equal size (Fig. 1) that can be protected at a cost that cumulates the costs of all cells included in the unit.

The monitoring action collects information within each protection unit about the state of the system S_t, which includes species abundances and geographic distribution:

$$S_t = \{H_t, D_t, F_t, T_t, C_t, P_t, B_t\} \tag{1}$$

where H_t is the matrix with the number of individuals across species and cells, D_t and F_t are matrices describing anthropogenic disturbance on the system, T_t is a matrix quantifying climate, C_t is the cost matrix, P_t is the current protection matrix and B_t is the available budget (for more details see Supplementary Methods and Supplementary Table 1). We define as feature extraction the result of a function $X(S_t)$, which returns for each protection unit a set of features summarizing the state of the system in the unit. The number and selection of features (Supplementary Methods and Supplementary Table 2) depends on the monitoring policy π_x, which is decided a priori in the simulation. A predefined monitoring policy also determines the temporal frequency of this action throughout the simulation, for example, only at the first time step or repeated at each time step. The features extracted for each unit represent the input upon which a protecting action can take place, if the budget allows for it, following a protection policy π_Y. These features (listed in Supplementary Table 2) include the number of species that are not already protected in other units, the number of rare species and the cost of the unit relative to the remaining budget. Different subsets of these features are used depending on the monitoring policy and on the optimality criterion of the protection policy π_Y.

We do not assume species-specific sensitivities to disturbance (parameters d_s, f_s in Supplementary Table 1 and Supplementary Methods) to be known features, because a precise estimation of these parameters in an empirical case would require targeted experiments, which we consider unfeasible across a large number of species. Instead, species-specific sensitivities can be learned from the system through the observation of changes in the relative abundances of species (x_3 in Supplementary Table 2). The features tested across different policies are specified in the subsection Experiments below and in the Supplementary Methods.

The protecting action selects a protection unit and resets the disturbance in the included cells to an arbitrarily low level. A protected unit is also immune from future anthropogenic disturbance increases, but protection does not prevent climate change in the unit. The model can include a buffer area along the perimeter of a protected unit, in which the level of protection is lower than in the centre, to mimic the generally negative edge effects in protected areas (for example, higher vulnerability to extreme weather). Although protecting a disturbed area theoretically allows it to return to its initial biodiversity levels, population growth and species composition of the protected area will still be controlled by the death–replacement–dispersal processes described above, as well as by the state of neighbouring areas. Thus, protecting an area that has already undergone biodiversity loss may not result in the restoration of its original biodiversity levels.

The protecting action has a cost determined by the cumulative cost of all cells in the selected protection unit. The cost of protection can be set equal across all cells and constant through time. Alternatively, it can be defined as a function of the current level of anthropogenic disturbance in the cell. The cost of each protecting action is taken from a predetermined finite budget and a unit can be protected only if the remaining budget allows it.

Policy definition and optimization algorithm. We frame the optimization problem as a stochastic control problem where the state of the system S_t evolves through time as described in the section above (see also Supplementary Methods), but it is also influenced by a set of discrete actions determined by the protection policy π_Y. The protection policy is a probabilistic policy: for a given set of policy parameters and an input state, the policy outputs an array of probabilities associated with all possible protecting actions. While optimizing the model, we extract actions according to the probabilities produced by the policy to make sure that we explore the space of actions. When we run experiments with a fixed policy instead, we choose the action with highest probability. The input state is transformed by the feature extraction function $X(S_t)$ defined by the monitoring policy, and the features are mapped to a probability through a neural network with the architecture described below.

In our simulations, we fix monitoring policy π_x, thus predefining the frequency of monitoring (for example, at each time step or only at the first time step) and the amount of information produced by $X(S_t)$, and we optimize π_Y, which determines how to best use the available budget to maximize the reward. Each action A has a cost, defined by the function $\text{Cost}(A, S_t)$, which here we set to zero for the monitoring action (X) across all monitoring policies. The cost of the protecting action (Y) is instead set to the cumulative cost of all cells in the selected protection unit. In the simulations presented here, unless otherwise specified, the protection policy can only add one protected unit at each time step, if the budget allows, that is if $\text{Cost}(Y, S_t) < B_t$.

The protection policy is parametrized as a feed-forward neural network with a hidden layer using a rectified linear unit (ReLU) activation function (Eq. (3)) and

an output layer using a softmax function (Eq. (5)). The input of the neural network is a matrix x of J features extracted through the most recent monitoring across U protection units. The output, of size U, is a vector of probabilities, which provides the basis to select a unit for protection. Given a number of nodes L, the hidden layer $h^{(1)}$ is a matrix $U \times L$:

$$h_{ul}^{(1)} = g\left(\sum_{j=1}^{J} x_{uj} W_{jl}^{(1)}\right) \quad (2)$$

where $u \in \{1, …, U\}$ identifies the protection unit, $l \in \{1, …, L\}$ indicates the hidden nodes and $j \in \{1, …, J\}$ the features and where

$$g(x) = \max(0, x) \quad (3)$$

is the ReLU activation function. We indicate with $W^{(1)}$ the matrix of $J \times L$ coefficients (shared among all protection units) that we are optimizing. Additional hidden layers can be added to the model between the input and the output layer. The output layer takes $h^{(1)}$ as input and gives an output vector of U variables:

$$h_u^{(2)} = \sigma\left(\sum_{l=1}^{L} h_{ul}^{(1)} W_l^{(2)}\right) \quad (4)$$

where σ is a softmax function:

$$\sigma(x_i) = \frac{\exp(x_i)}{\sum_u \exp(x_u)} \quad (5)$$

We interpret the output vector of U variables as the probability of protecting the unit u.

This architecture implements parameter sharing across all protection units when connecting the input nodes to the hidden layer; this reduces the dimensionality of the problem at the cost of losing some spatial information, which we encode in the feature extraction function. The natural next step would be to use a convolutional layer to discover relevant shape and space features instead of using a feature extraction function. To define a baseline for comparisons in the experiments described below, we also define a random protection policy $\hat{\pi}$, which sets a uniform probability to protect units that have not yet been protected. This policy does not include any trainable parameter and relies on feature x_6 (an indicator variable for protected units; Supplementary Table 2) to randomly select the proposed unit for protection.

The optimization algorithm implemented in CAPTAIN optimizes the parameters of a neural network such that they maximize the expected reward resulting from the protecting actions. With this aim, we implemented a combination of standard algorithms using a genetic strategies algorithm[43] and incorporating aspects of classical policy gradient methods such as an advantage function[44]. Specifically, our algorithm is an implementation of the Parallelized Evolution Strategies[43], in which two phases are repeated across several iterations (hereafter, epochs) until convergence. In the first phase, the policy parameters are randomly perturbed and then evaluated by running one full episode of the environment, that is, a full simulation with the system evolving for a predefined number of steps. In the second phase, the results from different runs are combined and the parameters updated following a stochastic gradient estimate[43]. We performed several runs in parallel on different workers (for example, processing units) and aggregated the results before updating the parameters. To improve the convergence we followed the standard approach used in policy optimization algorithms[44], where the parameter update is linked to an advantage function A as opposed to the return alone (Eq. (6)). Our advantage function measures the improvement of the running reward (weighted average of rewards across different epochs) with respect to the last reward. Thus, our algorithm optimizes a policy without the need to compute gradients and allowing for easy parallelization. Each epoch in our algorithm works as:

> **for** every worker p **do**
> $\quad \epsilon_p \leftarrow \mathcal{N}(0, \sigma)$, with diagonal covariance and dimension $W + M$
> $\quad$ **for** $t = 1,...,T$ **do**
> $\quad\quad R_t \leftarrow R_{t-1} + r_t(\theta + \epsilon_p)$
> $\quad$ **end for**
> **end for**
> $R \leftarrow$ average of R_T across workers
> $R_e \leftarrow \alpha R + (1 - \alpha) R_{e-1}$
> **for** every coefficient θ in $W + M$ **do**
> $\quad \theta \leftarrow \theta + \lambda A(R_e, R_T, \epsilon)$
> **end for**

where $\mathcal{N}$ is a normal distribution and $W + M$ is the number of parameters in the model (following the notation in Supplementary Table 1). We indicate with r_t the reward at time t, with R the cumulative reward over T time steps. R_e is the running average reward calculated as an exponential moving average where $\alpha = 0.25$ represents the degree of weighting decrease and R_{e-1} is the running average

reward at the previous epoch. $\lambda = 0.1$ is a learning rate and A is an advantage function defined as the average of final reward increments with respect to the running average reward R_e on every worker p weighted by the corresponding noise ϵ_p:

$$A(R_e, R_T, \epsilon) = \frac{1}{P} \sum_p (R_e - R_T^p)\epsilon_p. \quad (6)$$

Experiments. We used our CAPTAIN framework to explore the properties of our model and the effect of different policies through simulations. Specifically, we ran three sets of experiments. The first set aimed at assessing the effectiveness of different policies optimized to minimize species loss based on different monitoring strategies. We ran a second set of simulations to determine how policies optimized to minimize value loss or maximize the amount of protected area may impact species loss. Finally, we compared the performance of the CAPTAIN models against the state-of-the-art method for conservation planning (Marxan[25]). A detailed description of the settings we used in our experiments is provided in the Supplementary Methods. Additionally, all scripts used to run CAPTAIN and Marxan analyses are provided as Supplementary Information.

Analysis of Madagascar endemic tree diversity. We analysed a recently published[33] dataset of 1,517 tree species endemic to Madagascar, for which presence/absence data had been approximated through species distribution models across 22,394 units of 5×5 km spanning the entire country (Supplementary Fig. 5a). Their analyses included a spatial quantification of threats affecting the local conservation of species and assumed the cost of each protection unit as proportional to its level of threat (Supplementary Fig. 5b), similarly to how our CAPTAIN framework models protection costs as proportional to anthropogenic disturbance.

We re-analysed these data within a limited budget, allowing for a maximum of 10% of the units with the lowest cost to be protected (that is, 2,239 units). This figure can actually be lower if the optimized solution includes units with higher cost. We did not include temporal dynamics in our analysis, instead choosing to simply monitor the system once to generate the features used by CAPTAIN and Marxan to place the protected units. Because the dataset did not include abundance data, the features only included species presence/absence information in each unit and the cost of the unit.

Because the presence of a species in the input data represents a theoretical expectation based on species distribution modelling, it does not consider the fact that strong anthropogenic pressure on a unit (for example, clearing a forest) might result in the local disappearance of some of the species. We therefore considered the potential effect of disturbance in the monitoring step. Specifically, in the absence of more detailed data about the actual presence or absence of species, we initialized the sensitivity of each species to anthropogenic disturbance as a random draw from a uniform distribution $d_s \sim \mathcal{U}(0, 1)$ and we modelled the presence of a species s in a unit c as a random draw from a binomial distribution with a parameter set equal to $p_s^c = 1 - d_s \times D^c$, where $D^c \in [0, 1]$ is the disturbance (or 'threat' sensu Carrasco et al.[33]) in the unit. Under this approach, most of the species expected to live in a unit are considered to be present if the unit is undisturbed. Conversely, many (especially sensitive) species are assumed to be absent from units with high anthropogenic disturbance. This resampled diversity was used for feature extraction in the monitoring steps (Fig. 1c). While this approach is an approximation of how species might respond to anthropogenic pressure, the use of additional empirical data on species-specific sensitivity to disturbance can provide a more realistic input in the CAPTAIN analysis.

We repeated this random resampling 50 times and analysed the resulting biodiversity data in CAPTAIN using the one-time protection model, trained through simulations in the experiments described in the previous section and in the Supplementary Methods. We note that it is possible, and perhaps desirable, in principle to train a new model specifically for this empirical dataset or at least fine-tune a model pretrained through simulations (a technique known as transfer learning), for instance, using historical time series and future projections of land use and climate change. Yet, our experiment shows that even a model trained solely using simulated datasets can be successfully applied to empirical data. Following Carrasco et al.[33], we set as the target of our policy the protection of at least 10% of each species range. To achieve this in CAPTAIN, we modified the monitoring action such that a species is counted as protected only when at least 10% of its range falls within already protected units. We ran the CAPTAIN analysis for a single step, in which all protection units are established.

We analysed the same resampled datasets using Marxan with the initial budget used in the CAPTAIN analyses and under two configurations. First, we used a BLM (BLM = 0.1) to penalize the establishment of non-adjacent protected units following the settings used in Carrasco et al.[33]. After some testing, as suggested in Marxan's manual[45], we set penalties on exceeding the budget, such that the cost of the optimized results indeed does not exceed the total budget (THRESHPEN1 = 500, THRESHPEN2 = 10). For each resampled dataset we ran 100 optimizations (with Marxan settings NUMITNS = 1,000,000, STARTTEMP = −1 and NUMTEMP = 10,000 (ref. [45]) and used the best of them as the final result. Second, because the BLM adds a constraint that does not have a

direct equivalent in the CAPTAIN model, we also repeated the analyses without it (BLM = 0) for comparison.

To assess the performance of CAPTAIN and compare it with that of Marxan, we computed the fraction of replicates in which the target was met for all species, the average number of species for which the target was missed and the number of protected units (Supplementary Table 4). We also calculated the fraction of each species range included in protected units to compare it with the target of 10% (Fig. 6c,d and Supplementary Fig. 6c,d). Finally, we calculated the frequency at which each unit was selected for protection across the 50 resampled datasets as a measure of its relative importance (priority) in the conservation plan.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All data necessary to run the analyses presented here are available in a permanent repository on Zenodo[46].

Code availability
The CAPTAIN software is implemented in Python v.3 and is available at captain-project.net. All scripts and data necessary to run the analyses presented here are available in a permanent repository on Zenodo[46].

References

1. Díaz, S. et al. Assessing nature's contributions to people. *Science* **359**, 270–272 (2018).
2. Chaplin-Kramer, R. et al. Global modeling of nature's contributions to people. *Science* **366**, 255–258 (2019).
3. Dasgupta, P. *The Economics of Biodiversity: The Dasgupta Review* (HM Treasury, 2021).
4. Barnosky, A. D. et al. Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).
5. Andermann, T., Faurby, S., Turvey, S. T., Antonelli, A. & Silvestro, D. The past and future human impact on mammalian diversity. *Sci. Adv.* **6**, eabb2313 (2020).
6. *Global Biodiversity Outlook 5 – Summary for Policy Makers* (Secretariat of the Convention on Biological Diversity, 2020).
7. Sterner, T. et al. Policy design for the Anthropocene. *Nat. Sustain.* **2**, 14–21 (2019).
8. Mace, G. M. Whose conservation? *Science* **345**, 1558–1560 (2014).
9. Leclère, D. et al. Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* **585**, 551–556 (2020).
10. Moilanen, A. Generalized complementarity and mapping of the concepts of systematic conservation planning. *Conserv. Biol.* **22**, 1655–1658 (2008).
11. Margules, C. & Sarkar, S. *Systematic Conservation Planning* (Cambridge Univ. Press, 2007).
12. Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243–253 (2000).
13. Wilson, K. A., McBride, M. F., Bode, M. & Possingham, H. P. Prioritizing global conservation efforts. *Nature* **440**, 337–340 (2006).
14. Visconti, P., Pressey, R. L., Bode, M. & Segan, D. B. Habitat vulnerability in conservation planning—when it matters and how much. *Conserv. Lett.* **3**, 404–414 (2010).
15. Visconti, P., Pressey, R. L., Segan, D. B. & Wintle, B. A. Conservation planning with dynamic threats: the role of spatial design and priority setting for species' persistence. *Biol. Conserv.* **143**, 756–767 (2010).
16. Wilson, K. A. et al. Prioritizing conservation investments for mammal species globally. *Phil. Trans. R. Soc. B* **366**, 2670–2680 (2011).
17. Obura, D. O. et al. Integrate biodiversity targets from local to global levels. *Science* **373**, 746–748 (2021).
18. Moilanen, A., Wilson, K. & Possingham, H. (eds) *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools* (Oxford Univ. Press, 2009).
19. Honeck, E., Sanguet, A., Schlaepfer, M. A., Wyler, N. & Lehmann, A. Methods for identifying green infrastructure. *SN Appl. Sci.* **2**, 1916 (2020).
20. Bateman, I. J. et al. Bringing ecosystem services into economic decision-making: land use in the United Kingdom. *Science* **341**, 45–50 (2013).
21. Carvalho, S. et al. Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nat. Ecol. Evol.* **1**, 0151 (2017).
22. Sacre, E., Weeks, R., Bode, M. & Pressey, R. L. The relative conservation impact of strategies that prioritize biodiversity representation, threats, and protection costs. *Conserv. Sci. Pract.* **2**, e221 (2020).
23. Ball, I. R., Possingham, H. P. & Watts, M. in *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools* (eds Moilanen, A. et al.) 185–195 (Oxford Univ. Press, 2009).
24. Pressey, R. L., Mills, M., Weeks, R. & Day, J. C. The plan of the day: managing the dynamic transition from regional conservation designs to local conservation actions. *Biol. Conserv.* **166**, 155–169 (2013).
25. Watts, M., Klein, C. J., Tulloch, V. J. D., Carvalho, S. B. & Possingham, H. P. Software for prioritizing conservation actions based on probabilistic information. *Conserv. Biol.* **35**, 1299–1308 (2021).
26. Tulloch, V. J. et al. Incorporating uncertainty associated with habitat data in marine reserve design. *Biol. Conserv.* **162**, 41–51 (2013).
27. Grantham, H. S., Wilson, K. A., Moilanen, A., Rebelo, T. & Possingham, H. P. Delaying conservation actions for improved knowledge: how long should we wait? *Ecol. Lett.* **12**, 293–301 (2009).
28. Zizka, A., Silvestro, D., Vitt, P. & Knight, T. M. Automated conservation assessment of the orchid family with deep learning. *Conserv. Biol.* https://doi.org/10.1111/cobi.13616 (2020).
29. Gomes, C. et al. Computational sustainability: computing for a better world and a sustainable future. *Commun. ACM* **62**, 56–65 (2019).
30. Spring, D. A., Cacho, O., MacNally, R. & Sabbadin, R. Pre-emptive conservation versus "fire-fighting": a decision theoretic approach. *Biol. Conserv.* **136**, 531–540 (2007).
31. Meir, E., Andelman, S. & Possingham, H. P. Does conservation planning matter in a dynamic and uncertain world? *Ecol. Lett.* **7**, 615–622 (2004).
32. Adams, V. M. & Setterfield, S. A. Optimal dynamic control of invasions: applying a systematic conservation approach. *Ecol. Appl.* **25**, 1131–1141 (2015).
33. Carrasco, J., Price, V., Tulloch, V. & Mills, M. Selecting priority areas for the conservation of endemic trees species and their ecosystems in Madagascar considering both conservation value and vulnerability to human pressure. *Biodivers. Conserv.* **29**, 1841–1854 (2020).
34. Kosmala, M., Wiggins, A., Swanson, A. & Simmons, B. Assessing data quality in citizen science. *Front. Ecol. Environ.* **14**, 551–560 (2016).
35. Devillers, R. et al. Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquat. Conserv.* **25**, 480–504 (2015).
36. Vieira, R. R. S., Pressey, R. L. & Loyola, R. The residual nature of protected areas in Brazil. *Biol. Conserv.* **233**, 152–161 (2019).
37. Dilkina, B. et al. Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. *Conserv. Biol.* **31**, 192–202 (2017).
38. Bruelheide, H. et al. sPlot – a new tool for global vegetation analyses. *J. Veg. Sci.* **30**, 161–186 (2019).
39. Geldmann, J., Manica, A., Burgess, N. D., Coad, L. & Balmford, A. A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proc. Natl Acad. Sci. USA* **116**, 23209–23215 (2019).
40. Geldmann, J. et al. A global analysis of management capacity and ecological outcomes in terrestrial protected areas. *Conserv. Lett.* **11**, e12434 (2018).
41. Díaz, S. et al. Set ambitious goals for biodiversity and sustainability. *Science* **370**, 411–413 (2020).
42. Barto, A. G. Reinforcement learning and dynamic programming. *IFAC Proc. Volumes* **28**, 407–412 (1995).
43. Salimans, T., Ho, J., Chen, X., Sidor, S. & Sutskever, I. Evolution strategies as a scalable alternative to reinforcement learning. Preprint at https://arXiv.org/abs/1703.03864 (2017).
44. Sutton, R. S., McAllester, D. A., Singh, S. P. & Mansour, Y. in *Advances in Neural Information Processing Systems* (eds Solla, S. et al.) 1057–1063 (MIT Press, 1999).
45. Serra-Sogas et al. *Marxan User Manual: For Marxan Version 2.43 and Above* (The Nature Conservancy and PacMARA, 2021).
46. Silvestro, D., Goria, S., Sterner, T. & Antonelli, A. Supplementary data for: Improving biodiversity protection through artificial intelligence. *Zenodo* https://doi.org/10.5281/zenodo.5643665 (2021).
47. Murcia, C. Edge effects in fragmented forests: implications for conservation. *Trends Ecol. Evol.* **10**, 58–62 (1995).

Author contributions
T.S. initiated the project with A.A. D.S. and S.G. developed the AI model and performed all analyses. A.A., D.S., T.S. and S.G. wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information

nature portfolio

Corresponding author(s): Daniele Silvestro
Alexandre Antonelli

Last updated by author(s): Dec 22, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect data.
Data analysis	The data were analysed using Marxan (v. 4.0.6) and using custom software and scripts provided in the Supplementary Data (doi: 10.5281/zenodo.5643665) and at www.captain-project.net

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data used in this study are available in a permanent open-access repository with doi: 10.5281/zenodo.5643665

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study presents a new method to optimize conservation policies using reinforcement learning
Research sample	Simulated data and previously published species distribution data (Madagascar endemic trees).
Sampling strategy	N/A
Data collection	Simulated data and previously published species distribution data
Timing and spatial scale	Simulated and present data
Data exclusions	N/A
Reproducibility	Scripts and data are provided to make all analyses reproducible.
Randomization	N/A
Blinding	N/A

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging