Improving biodiversity protection through artificial intelligence

Daniele Silvestro#1,2,* Stefano Goria3, Thomas Sterner4, Alexandre Antonelli#2,5,6,*

1Department of Biology, University of Fribourg and Swiss Institute of Bioinformatics, 1700 Fribourg, Switzerland
2Gothenburg Global Biodiversity Centre, Department of Biological and Environmental Sciences, University of Gothenburg, 40530 Gothenburg, Sweden
3Thymia Limited, London, United Kingdom
4Department of Economics, University of Gothenburg, 40530 Gothenburg, Sweden
5Department of Plant Sciences, University of Oxford, Oxford, OX1 3RB, United Kingdom
6Royal Botanic Gardens, Kew, TW9 3AE, United Kingdom

# These authors contributed equally to this work.

Abstract

Over a million species face extinction, urging the need for conservation policies that maximize the protection of biodiversity to sustain its manifold contributions to people. 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, CAPTAIN (Conservation Area Prioritization Through Artificial INtelligence), 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 substantially 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, substantially 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 resourcelimited world.
Keywords
Citizen science; Climate change; Human impact; Marxan; Reinforcement learning; Remote sensing; Sustainable development; Systematic conservation planning

Introduction

Biodiversity is the variety of all life on Earth, from genes through to populations, species, functions and ecosystems. Alongside its own intrinsic value and ecological roles, biodiversity provides us with clean water, pollination, 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-run\(^3\) (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 massively 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. Since not a single 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 15\(^{th}\) Conference of the Parties in 2022 (https://www.cbd.int/).

There have been several theoretical and practical frameworks underlying biological conservation since the 1960s\(^8\). 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 multi-faceted 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 field of spatial conservation prioritization or 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.
its usefulness and popularity, Marxan and similar methods\(^{18}\) are 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.

Here we tackle the challenge of optimizing biodiversity protection in a complex and rapidly evolving world by harnessing the power of artificial intelligence (AI). We develop 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. We use it to explore – through simulations and empirical analyses – multiple previously identified trade-offs in real-world conservation, and to evaluate the impact of data gathering on specific outcomes\(^{27}\). We also explore the impact of species-specific sensitivity to geographically varying local disturbances (e.g., 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 (Conservation Area Prioritization Through Artificial INtelligence).

Within AI, we implement 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 estimated 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.

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\(^{30}\) but not yet implemented in practical conservation tools. In particular, CAPTAIN aims to tackle multi-dimensional 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 of conservation in a dynamic world\(^{31}\), characterized by heterogeneous and often unpredictable habitat loss\(^{14}\) which require iterative and regular conservation interventions.
We use CAPTAIN to address the following questions: i) What role does data gathering strategy play for effective conservation? ii) What trade-offs exist depending on the optimized variable, such as species richness, economic value or total area protected? iii) What can the simulation framework reveal in terms of winners and losers – i.e., which traits characterize the species and areas protected over time? and iv) How does our framework perform compared with the state-of-the-art model for conservation planning Marxan? Finally, we demonstrate the usefulness of our framework and direct applicability of models trained through RL to an empirical dataset of endemic trees of Madagascar.

Results

Impact of data gathering strategy

We find 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; 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; see 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; Supplementary Table 3).

To thoroughly explore the parameter space of 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 those 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, significantly influences the outcome of the simulations. A policy minimizing the loss species based on their commercial value (such as timber price) tends to sacrifice more species in order to prioritize the protection of fewer, highly valuable ones. This policy, while efficiently reducing the loss of cumulative value, decreases species losses decrease 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 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, by resulting in 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 on economic value and total protected area is high for the respective policy objectives, but 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 range, 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 characterize 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 against the current state-of-the-art tool for conservation prioritization, Marxan\(^\text{23}\). The methods differ conceptually in that, while CAPTAIN is explicitly designed to minimize loss (e.g., local species extinction) within the constraints of a limited budget, Marxan’s default algorithms minimize the cost of reaching a conservation target (e.g., 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.

To compare the two models, we set up all Marxan analyses with an explicit budget constraint, following other tailored implementations\(^\text{15,33}\) (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 shows 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 setting, 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 analyzed a Madagascar biodiversity dataset recently used in a systematic conservation planning experiment\(^3^4\) under Marxan\(^2^3\). The dataset included 22,394 protection units (5 x 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\(^3^4\) (see Methods for more details; Supplementary Fig. 5).

We analyzed the data assuming a Full Initial Monitoring in a static setting where 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 (which penalizes the placement of many isolated protection units in favour of larger contiguous areas; BLM = 0.1 as in\(^3^4\) and without it (BLM = 0 for comparability to CAPTAIN that does not include that feature).

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

**Discussion**

We presented a new framework to optimize dynamic conservation policies using reinforcement learning and evaluate their biodiversity outcome through simulations.

**Data gathering and monitoring**

Our finding that even simple data (presence/absence of species) is sufficient to inform effective policies (Fig. 2, Supplementary Table 3) is noteworthy because the information it requires 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\(^3^5\) 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 – all of which are likely to change through time in real-world situations. While
some systems may remain largely static over decades (e.g., tree species in old-growth forests), others may change drastically (e.g., 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. Since current methodologies for systematic conservation planning are static – relying on a similar initial data gathering as modelled here – this means their recommendations for area protection may be less reliable.

Optimization trade-offs

Our results indicate clear trade-offs, meaning that optimizing one value can be at the cost of another (Fig. 3, 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 Aichi; 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’ reservation36,37 – the selection of new protected areas on land and at sea that are unsuitable for extractive activities, which may reduce costs and risks for 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, these simulations indicate that protected cells should span a range of areas with intermediate to high species richness, reflecting known differences among ecosystems or across environmental gradients. Such selection is more likely to increase protection complementarity for multiple species, a key factor incorporated by our software and some others10,23,38.

Applications and prospects

Our successful benchmarking against random, naive, and Marxan-optimized solutions indicate 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 is 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 i) more efficiently use the available information of species distribution and ii) more easily integrate multidimensional and time-varying biodiversity data. As the number of standardized high-resolution biological datasets is increasing (e.g.,39), thanks to new and cost-effective monitoring technologies
(Supplementary Box 2), our approach offers a future-proof tool for research, conservation and 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, socio-economic value and ecological functions. AI techniques should not replace human judgment, 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\textsuperscript{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\textsuperscript{40,41}. However, it is now time to acknowledge that the sheer complexity of socio-biological 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\textsuperscript{9,42}, 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 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 a reinforcement learning 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 (i.e., 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 sizes 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 a reinforcement learning algorithm to optimize how to perform the protect 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

We use reinforcement learning to optimize a conservation policy under a pre-defined policy objective (e.g., 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 time frame evaluated in the simulation. If the value is set equal across all species, the reinforcement learning algorithm will minimize overall species extinctions. However, different definitions of value can be used to minimize loss based on evolutionary distinctiveness of species (e.g., 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 $C_t$, which can vary across cells and through time.

The granularity of monitoring and protection 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), which 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, which includes species abundances and geographic distribution,

\[ S_t = \{H_t, D_t, F_t, T_t, C_t, P_t, B_t\} \]  

(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 Table 2) depends on the monitoring policy (\( \pi_X \)), which is decided a priori in the simulation. A predefined monitoring policy also determines the temporal frequency of this action throughout the simulation, e.g. only at the first time step or repeated at each time step. The features extracted for each unit represent the input upon which a protect action can take place, if the budget allows for it, following a protection policy \( \pi_Y \). These features (listed in Supplementary Table 2) include the number of species which 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 (\( \pi_Y \)).

We do not assume species-specific sensitivities to disturbance (\( d_s, f_s \)) to be known features, since 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 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 protect 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, while 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 compared to the centre, to mimic the generally negative edge effects in protected areas (e.g., higher vulnerability to extreme weather). While 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 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 protect 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 thorough time. Alternatively, it can be defined as a function of the current level of anthropogenic disturbance in the cell. The cost of each protection 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 $\pi_Y$. 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 to all possible protect 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 architecture described below.

In our simulations we fix $\pi_X$, thus pre-defining the frequency of monitoring (e.g. at each time step or only at the first time step) and the amount of information produced by $X(S_t)$ and we optimize $\pi_Y$, which determines how to best use the available budget to maximize the reward. Each action has cost, defined by the function $\text{Cost}(A, S_t)$, which here we set to a constant for monitoring and equal across all monitoring policies. The cost of the protect action 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, i.e. 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 ReLU activation function (Eq. 3) and an output layer using a softMax function (Eq. 5). The input of the neural network is a matrix 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. The hidden layer $h^{(1)}$ is a matrix $J \times L_1$

$$h^{(1)}_{nl1} = g\left(\sum_j x_{nj} W^{(1)}_{fj1}\right)$$

(2)

where

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

(3)

is the ReLU activation function and the matrix $W^{(1)}$ is the matrix of coefficients we are optimizing. Additional hidden layers can be added to the model between the input and the output layer. The output layer takes the $h^{(1)}$ as input and gives an $N$ vector as output:

$$h^{(2)}_n = \sigma\left(\sum_{l1} h^{(1)}_{nl1} W^{(2)}_{l1}\right)$$

(4)

with $\sigma$ a softMax function.
\[ \sigma(x) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \]  

We interpret the output vector of \( N \) variables as the probability of protecting the cell \( n \).

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 protect policy, \( \pi \), 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) 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 protect actions. To this aim, we implemented combination of standard algorithms, using a genetic strategies algorithm\(^44\) and incorporating aspects of classical Policy Gradient methods such as an Advantage function\(^45\). Specifically, our algorithm is an implementation of the Parallelized Evolution Strategies\(^44\), 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, i.e. 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\(^44\). We perform several runs in parallel on different workers (e.g. CPUs) and aggregate the results before updating the parameters. To improve the convergence we follow the standard approach used in policy optimization algorithms\(^45\), 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:

\begin{verbatim}
for every worker \( p \) do
    \( \epsilon_p \leftarrow N(0, \sigma) \) with diagonal covariance and dimension \( W + M \)
    for \( t = 1, \ldots, T \) do
        \( R_t \leftarrow R_{t-1} + r(\theta + \epsilon_p) \)
    end for
end for

\( R \leftarrow \) average of \( R_T \) across workers
\end{verbatim}
\[ \begin{align*}
R_e & \leftarrow \alpha R + (1 - \alpha)R_{e-1} \\
& \text{for every coefficient } \theta \text{ in } W + M \text{ do} \\
\theta & \leftarrow \theta + \lambda A(R_e, R_T, \epsilon) \\
& \text{end for}
\end{align*} \]

where \( R \) is the cumulative reward over \( T \) time steps, \( \lambda = 0.1 \) 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 \( \epsilon_p \):

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

(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 CAPTAIN models against that of a state-of-the-art method for conservation planning (Marxan\textsuperscript{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 analyzed a recently published\textsuperscript{34} 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-analyzed these data within a limited budget allowing for a maximum of 10% of the units with the lowest cost to be protected (i.e. 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. Since the dataset did not include abundance data, the features only included species presence/absence information in each unit and the cost of the unit. Since the presence of a species in the input data represents a theoretical expectation based on species distribution modeling, it does not consider the fact that strong anthropogenic pressure on a unit (e.g., 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 $d_s \sim \mathcal{U}(0, 1)$ 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 $p_{sc} = 1 - d_s \times D^c$, where $D^c \in [0, 1]$ is the disturbance (or “threat” sensu Carrasco et al.\cite{34}) 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 analyzed 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 an model pre-trained 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.\cite{34}, we set as the target of our policy to protect 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, which in which all protection units are established.

We analyzed the same resampled datasets using Marxan with the initial budget used in the CAPTAIN analyses and under two configurations. First we used a boundary length multiplier ($BLM = 0.1$), to penalize the establishment of non adjacent protected units following the settings used in Carrasco et al.\cite{38}. After some testing, as suggested in Marxan’s manual\cite{46}, 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 settings: $NUMTNS = 1,000,000$; $STARTTEMP = -1$; $NUMTEMP = 10000$) and used the best among them as the final result. Second, since the boundary length multiplier 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; 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.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgements**

We thank Kerstin Johannesson for the initial discussions that sparked this project; Ben Groom, Atte Moilanen, Bob Pressey, Kiran Dhanjal-Adams and Ian Bateman for invaluable feedback on an earlier version of this manuscript; Jesús Carrasco for providing the Madagascar tree distribution data; Justin Moat for input on remote sensing technologies; Weston Testo for suggesting the acronym CAPTAIN; Rhian Smith for editorial support; Gytis Dudas for coding support; many colleagues at our respective research groups for discussions. AA is funded by the Swedish Foundation for Strategic Research (FFL15-0196), the Swedish Research Council (VR: 2019-05191) and the Royal Botanic Gardens, Kew. TS and AA acknowledge funding from the Strategic Research Area Biodiversity and Ecosystem Services in a Changing Climate, BECC, funded by the Swedish government. DS received funding from the Swiss National Science Foundation (PCEFP3_187012) and from the Swedish Research Council (VR: 2019-04739).

**Data availability**

All scripts and data necessary to run the analyses presented here are available in a permanent repository on zenodo.org.

**Code availability**

The CAPTAIN software is implemented in Python v.3 and available here: captain-project.net.

**References**

1. Díaz S, et al. Assessing nature’s contributions to people. Science. 2018; 359: 270–272. DOI: 10.1126/science.aap8826 [PubMed: 29348221]
2. Chaplin-Kramer R, et al. Global modeling of nature’s contributions to people. Science. 2019; 366: 255–258. DOI: 10.1126/science.aaw3372 [PubMed: 31601772]
3. Dasgupta, P. The Economics of Biodiversity: The Dasgupta Review. HM Treasury; London: 2021.
4. Barnosky AD, et al. Has the Earth’s sixth mass extinction already arrived? Nature. 2011; 471: 51–57. DOI: 10.1038/nature09678 [PubMed: 21368823]
5. Andermann T, Fauryby S, Turvey ST, Antonelli A, Silvestro D. The past and future human impact on mammalian diversity. Science Advances. 2020; 6 eabb2313 doi: 10.1126/sciadv.abb2313 [PubMed: 32917612]
6. Diversity, S. o. t. C. o. B. Global Biodiversity Outlook 5 – Summary for Policy Makers. 2020.
7. Sterner T, et al. Policy design for the Anthropocene. Nature Sustainability. 2019; 2: 14–21. DOI: 10.1038/s41893-018-0194-x
8. Mace GM. Whose conservation? Science. 2014; 345: 1558–1560. DOI: 10.1126/science.1254704 [PubMed: 25258063]
9. Leclère D, et al. Bending the curve of terrestrial biodiversity needs an integrated strategy. Nature. 2020; 585: 551–556. DOI: 10.1038/s41586-020-2705-y [PubMed: 32908312]
10. Moilanen A. Generalized Complementarity and Mapping of the Concepts of Systematic Conservation Planning. Conservation biology: the journal of the Society for Conservation Biology. 2008; 22: 1655–1658. DOI: 10.1111/j.1523-1739.2008.01043.x [PubMed: 18778265]
11. Margules, C, Sarkar, S. Systematic conservation planning. Cambridge University Press; 2007.
12. Margules CR, Pressey RL. Systematic conservation planning. Nature. 2000; 405: 243–253. DOI: 10.1038/35012251 [PubMed: 10821285]

13. Wilson KA, McBride MF, Bode M, Possingham HP. Prioritizing global conservation efforts. Nature. 2006; 440: 337–340. [PubMed: 16541073]

14. Visconti P, Pressey RL, Bode M, Segan DB. Habitat vulnerability in conservation planning—when it matters and how much. Conservation Letters. 2010; 3: 404–414.

15. Visconti P, Pressey RL, Segan DB, Wintle BA. Conservation planning with dynamic threats: the role of spatial design and priority setting for species’ persistence. Biol Conserv. 2010; 143: 756–767.

16. Wilson KA, et al. Prioritizing conservation investments for mammal species globally. Philos Trans R Soc Lond B Biol Sci. 2011; 366: 2670–2680. [PubMed: 21844046]

17. Obura DO, et al. Integrate biodiversity targets from local to global levels. Science. 2021; 373: 746–748. DOI: 10.1126/science.abh2234 [PubMed: 34385386]

18. Moilanen, A, Wilson, K, Possingham, H. Spatial conservation prioritization: quantitative methods and computational tools. Oxford University Press; 2009.

19. Honeck E, Sanguet A, Schlaepfer MA, Wyler N, Lehmann A. Methods for identifying green infrastructure. SN Applied Sciences. 2020; 2: 1916. doi: 10.1007/s42452-020-03575-4

20. Bateman IJ, et al. Bringing Ecosystem Services into Economic Decision-Making: Land Use in the United Kingdom. Science. 2013; 341: 45–50. DOI: 10.1126/science.1234379 [PubMed: 23828934]

21. Carvalho S, et al. Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. Nature Ecology & Evolution. 2017; 1 0151 doi: 10.1038/s41559-017-0151

22. Sacre E, Weeks R, Bode M, Pressey RL. The relative conservation impact of strategies that prioritize biodiversity representation, threats, and protection costs. Conservation Science and Practice. 2020; 2 e221 doi: 10.1111/csp2.221

23. Ball IR, Possingham HP, Watts M. Marxan and relatives: software for spatial conservation prioritisation. Spatial conservation prioritisation: Quantitative methods and computational tools. 2009. 185–195.

24. Pressey RL, Mills M, Weeks R, Day JC. The plan of the day: Managing the dynamic transition from regional conservation designs to local conservation actions. Biol Conserv. 2013; 166: 155–169. DOI: 10.1016/j.biocon.2013.06.025

25. Watts M, Klein CJ, Tulloch VJD, Carvalho SB, Possingham HP. Software for prioritizing conservation actions based on probabilistic information. Conserv Biol. 2021; 35: 1299–1308. DOI: 10.1111/cobi.13681 [PubMed: 33305882]

26. Tulloch VJ, et al. Incorporating uncertainty associated with habitat data in marine reserve design. Biol Conserv. 2013; 162: 41–51. DOI: 10.1016/j.biocon.2013.03.003

27. Grantham HS, Wilson KA, Moilanen A, Rebelo T, Possingham HP. Delaying conservation actions for improved knowledge: how long should we wait? Ecol Lett. 2009; 12: 293–301. DOI: 10.1111/j.1461-0248.2009.01287.x [PubMed: 19243409]

28. Zizka A, Silvestro D, Vitt P, Knight TM. Automated conservation assessment of the orchid family with deep learning. Conserv Biol. 2020; doi: 10.1111/cobi.13616

29. Gomes C, et al. Computational sustainability: computing for a better world and a sustainable future. Commun ACM. 2019; 62: 56–65. DOI: 10.1145/3339399

30. Spring DA, Cacho O, Mac Nally R, Sabadin R, Pre-emptive conservation versus “fire-fighting”: A decision theoretic approach. Biol Conserv. 2007; 136: 531–540. DOI: 10.1016/j.biocon.2006.12.024

31. Meir E, Andelman S, Possingham HP. Does conservation planning matter in a dynamic and uncertain world? Ecol Lett. 2004; 7: 615–622.

32. Murcia C. Edge effects in fragmented forests: implications for conservation. Trends Ecol Evol. 1995; 10: 58–62. [PubMed: 21236953]

33. Adams VM, Setterfield SA. Optimal dynamic control of invasions: applying a systematic conservation approach. Ecol Appl. 2015; 25: 1131–1141. [PubMed: 26465047]
34. 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. 2020; 29: 1841–1854.

35. Kosmala M, Wiggins A, Swanson A, Simmons B. Assessing data quality in citizen science. Front Ecol Environ. 2016; 14: 551–560. DOI: 10.1002/fee.1436

36. Devillers R, et al. Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? Aquat Conserv: Mar Freshwat Ecosyst. 2015; 25: 480–504. DOI: 10.1002/aqc.2445

37. Vieira RRS, Pressey RL, Loyola R. The residual nature of protected areas in Brazil. Biol Conserv. 2019; 233: 152–161. DOI: 10.1016/j.biocon.2019.02.010

38. Dilkina B, et al. Tradeoffs and efficiencies in optimal budget-constrained multispecies corridor networks. Conserv Biol. 2017; 31: 192–202. [PubMed: 27677418]

39. Bruehlheide H, et al. sPlot – A new tool for global vegetation analyses. Journal of Vegetation Science. 2019; 30: 161–186. DOI: 10.1111/jvs.12710

40. Geldmann J, Manica A, Burgess ND, Coad L, Balmford A. A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. Proc Natl Acad Sci USA. 2019; 116: 23209–23215. DOI: 10.1073/pnas.1908221116 [PubMed: 31659036]

41. Geldmann J, et al. A global analysis of management capacity and ecological outcomes in terrestrial protected areas. Conservation Letters. 2018; 11 e12434 doi: 10.1111/conl.12434

42. Díaz S, et al. Set ambitious goals for biodiversity and sustainability. Science. 2020; 370: 411–413. DOI: 10.1126/science.abe1530 [PubMed: 33093100]

43. Barto AG. Reinforcement Learning and Dynamic Programming. IFAC Proceedings Volumes. 1995; 28: 407–412. DOI: 10.1016/S1474-6670(17)45266-9

44. Salimans T, Ho J, Chen X, Sidor S, Sutskever I. Evolution strategies as a scalable alternative to reinforcement learning. arXiv preprint. 2017. arXiv:1703.03864

45. Sutton, RS, McAllester, DA, Singh, SP, Mansour, Y. NIPs. Citeseer; 1057–1063.

46. Serra-Sogas. Marxan User Manual. For Marxan version 2.43 & above. 2021.

47. Silvestro D, Goria S, Sterner T, Antonelli A. Supplementary data for: Improving biodiversity protection through artificial intelligence. 2021; doi: 10.5281/zenodo.5643665
Fig. 1 | The CAPTAIN reinforcement learning framework

a, A simulated system – which could be the equivalent to a country state, 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 (e.g., forest loggings or sea trawling) will immediately reduce to an arbitrarily low level, except for the well-known edge effect\(^3\) characterized by intermediate levels of disturbance. All simulation settings are provided with initial default values but are fully customizable (see Supplementary Tables 1–2). Simulated systems evolve through time (b), and are used to optimize a conservation policy using RL (c). After training the model, the optimized policy can be used to evaluate the model performance based on simulated or empirical data. In empirical, the simulated system is replaced with available biodiversity and disturbance data. b-c, Analysis flowchart integrating simulations and AI modules to maximize selected outcomes (e.g., species richness). b, System evolution between two points in time, in relation to six variables: species richness, population density, economic value, phylogenetic diversity, anthropogenic disturbance, climate and species-rank abundance (see www.captain-
project.net for a time-lapse video 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, which 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.
Fig. 2 |. Impact of monitoring strategies on biodiversity protection

a-c, Outcome of policies designed to minimize species loss based on different monitoring strategies: a, Full Recurrent Monitoring (of species presence and abundance at each time step); b, Citizen science Recurrent Monitoring (limited to species presence/absence with some error at each time step); c, Full Initial Monitoring (species presence and abundance only at initial time). The blue polygons show the reduced loss of species, total protected area, accumulated species values and phylogenetic diversity between a random protection policy (grey shaded area) and models optimized by CAPTAIN. 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 (5x5 cells) but differed in their initial natural system (species distributions, abundances, tolerances, phylogenetic relationships) and in the dynamics of climate change and disturbance patterns.
Fig. 3 | Trade-offs in conservation outcomes in relation to policy objectives

The plots show the outcome (averaged across 250 simulations) of different policy objectives based on Full Recurrent Monitoring. The policies were designed to  

a, minimize value loss and  
b, maximize the amount of protected area. The radial axis shows the percentage change compared with the baseline random policy, while 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 (blue circles) and locally extinct (red circles) species after a simulation of 30 time-steps with increasing disturbance and climate change. The X and Y axes show the initial range size and population size of the species (log10 transformed) and the size of the circles is proportional to the resilience of each species to anthropogenic disturbance and climate change, with smaller circles representing more sensitive species.

b, Cumulative number of species encompassed in the ten protected units (5x5 cells) selected based on a policy optimized to minimize species loss. The grey density plot shows the expected distribution from 10,000 random draws, the purple shaded area shows the expected distribution when protected units are selected ‘naively’ (here, randomly chosen among the top 20 most diverse ones), while 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 to 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 (Supplementary Animation 1).

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.
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\textsuperscript{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.
Fig. 6 | Empirical validation of CAPTAIN.
The maps show a ranking of priority areas for protection across Madagascar based on the distribution of endemic trees based on a limited budget allowing for up to 10% of protected area overall. The color scale is proportional to the relative importance of each protected unit measured as the frequency at which it was included in the optimized solution. a, protection units identified through Marxan optimizations (with BLM = 0.1, as in Carrasco et al.34); b, equivalent CAPTAIN results based on a Full Initial Monitoring policy; c–d, histograms of the fraction of species ranges included in the protected units calculated for each species. 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.