RESEARCH ARTICLE

Predicting the optimal amount of time to spend learning before designating protected habitat for threatened species

Abbey E. Camaclang[1,2] | Iadine Chadès[3] | Tara G. Martin[2] | Hugh P. Possingham[1]

[1]Centre for Biodiversity and Conservation Science, University of Queensland, St Lucia, Qld, Australia

[2]Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, Canada

[3]CSIRO Land and Water, Ecosciences Precinct, Dutton Park, Qld, Australia

Correspondence
Abbey E. Camaclang
Email: abbey.camaclang@uqconnect.edu.au

Present address
Abbey E. Camaclang, Department of Forest and Conservation Sciences, 3041-2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

Abstract

1. Deciding when to protect threatened species habitat when complete knowledge about the habitat extent is uncertain is a common problem in conservation. More accurate habitat mapping improves conservation outcomes once that habitat is protected. However, delaying protection to improve accuracy can lead to species decline or, at worst, local extinction when threats to that habitat continue unabated before protection is implemented. Hence, there is a trade-off between gaining knowledge and taking conservation action.

2. We quantified this trade-off and determined the optimal time to spend learning about a species' habitat before protecting that habitat. We used a range of hypothetical learning curves to model improvements in the accuracy of predicted habitat over time, and receiver operating characteristic (ROC) curves to model the corresponding increase in the proportion of habitat protected. We used rates of habitat loss to model the impact of delaying habitat protection and derived analytical solutions to the problem for different types of learning curves.

3. We illustrate our approach using two threatened species, the koala *Phascolarctos cinereus* in Australia and northern abalone *Haliotis kamtschatkana* in Canada. Our approach confirms that when impacts of threatening processes are incurred rapidly, the need for timely protection is high, and the optimal time to spend learning is short for all learning curves. When the rate of habitat loss is low, we benefit from better habitat identification, and the optimal time to protect is sensitive to assumptions about how we learn and the proportion of non-habitat we are willing to protect unnecessarily.

4. Navigating the trade-off between information gain and timely action is a common problem in conservation. By optimizing the trade-off between the benefits of improving mapping accuracy and the costs of delaying protection, we provide guidelines on the effective allocation of resources between habitat identification and habitat protection. Importantly, by explicitly modelling this trade-off with a range of learning curves and estimates of the rates of habitat loss or other threatening processes, we can predict the optimal time to spend learning even when relatively little is known about a species and its habitat.

1 | INTRODUCTION

Protection of habitat from the impact of threatening processes is the single most important action for conserving species (Leclère et al., 2020). However, designating habitat for protection is difficult and controversial due to competing interests for land and natural resources. Ensuring that only areas critical to the recovery and persistence of species are protected can help minimize conflict and reduce the costs of establishing and managing protected areas (Bohorquez et al., 2019). Greater accuracy of habitat identification can improve the effectiveness of habitat protection and reduce the likelihood of unnecessarily protecting the wrong places. However, delays in conservation action can have severe consequences, including the irreversible loss of species (Martin et al., 2012; Naujokaitis-Lewis et al., 2018).

Knowing when we have sufficient information on species or threats to switch focus to implementation of management actions is a common challenge in conservation. Additional knowledge to inform decision-making should be balanced with the need for timely responses to ensure improved outcomes for biodiversity (Lindenmayer et al., 2013; Martin et al., 2012). For management actions that can be easily changed, an adaptive management framework may be useful in balancing this trade-off (Williams, 2011). However, some actions are irreversible or cannot be easily changed with improved knowledge. Actions that require changes in policies or regulations, such as the establishment of protected areas, legal designation of critical habitats for listed species or adoption of financial incentive schemes for private land conservation, are difficult, costly to modify and often controversial, and therefore potentially irreversible in practice (Pindyck, 2002). Furthermore, deciding not to protect an area may be irreversible once the area has been converted to other land uses (Perrings & Brock, 2009). Therefore, for those actions, striking the optimal balance between the need for better information and the need to act quickly is important for maximizing conservation outcomes and minimizing the potential political and economic costs of habitat protection.

Decision science can help optimize the trade-off between the need for accuracy and for timely action regarding the protection of species habitat, particularly for controversial decisions that are likely to be contested. Decision science approaches have been increasingly used by government agencies and conservation organizations to inform decisions about conservation and resource management (e.g. Chadès et al., 2019; Pascal et al., 2020; Runge et al., 2020). The application of decision science to determine when to prioritize information gain over management action has been explored within the context of detection and control of invasive species (Baxter & Possingham, 2011), land acquisition (McDonald-Madden et al., 2008), design of protected area networks (García-Barón et al., 2021; Grantham et al., 2009), threatened species management (Chadès et al., 2008) and designation of critical habitats for listed threatened species (Martin et al., 2017).

However, the modelling approaches used in these studies require substantive data inputs or high levels of technical expertise that may not be readily available (Field et al., 2007; Marescot et al., 2013). For example, Martin et al. (2017) used habitat suitability and population models to estimate the optimal time to identify critical habitat for northern abalone in Canada. Data required to develop such models are often not available for many species, hence the need to delay protection to collect more data. Simpler tools or decision rules are needed to assist managers in deciding how long to spend on learning before taking management action.

Here, we used a decision science approach to determine the optimal time to spend on research before taking action when the action is irreversible, either technically or in practice. We used a range of hypothetical learning curves and estimated rates of threatening processes to predict the optimal amount of time to spend learning before designating protected habitat, and applied our approach to two species considered to be at risk of extinction: the koala *Phascolarctos cinereus* in Australia and the northern abalone *Haliotis kamtschatkana* in Canada. Our method estimates the optimal amount of time to allocate to primary data collection for habitat identification before protecting habitat when there is insufficient information to develop habitat or population models, for example, at the start of a new research and conservation programme.

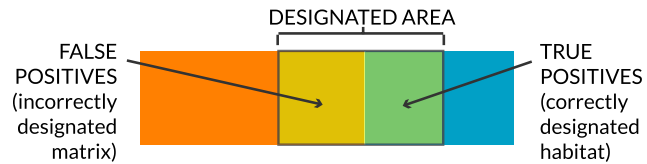## 2 | MATERIALS AND METHODS

2.1 | Problem definition

We consider the scenario wherein populations of a species in a given region are declining, and the relevant government agency has a legal responsibility to identify and designate the species' habitat for protection. The aim is to prevent declines by protecting as much of the species' habitat, and the populations within it, as possible. We assume that the resources needed for designation will be made available to meet this legal requirement, and that once designated, the habitat and the populations within are protected against threats and no further action is necessary.

We modelled the area of species' habitat that is identified and protected at the time of designation as a response to the dynamics of two variables: (a) the amount of habitat available (x_H) and (b) the accuracy (*a*) of the habitat model used for identification (Figure 1). Until designation occurs, the species and its habitat remain unprotected and vulnerable to threats, leading to a decrease in population or amount of habitat available for protection over time. Designating early helps prevent further declines and maximize the area of species' habitat protected. However, imperfect knowledge about species and their habitat requirements leads to errors in habitat identification, resulting in the unintended exclusion
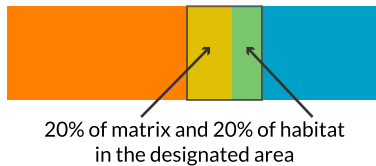
A given region may consist of areas that are habitat (H) for a given species and non-habitat, or matrix (M), areas.

MATRIX (NON-HABITAT)	HABITAT

Areas designated for the protection of the species and its habitat will include both habitat and matrix.

DESIGNATED AREA

FALSE POSITIVES (incorrectly designated matrix)

TRUE POSITIVES (correctly designated habitat)

Without much knowledge, designation will have *low accuracy* and tends to be *random*.

20% of matrix and 20% of habitat in the designated area

When designation is random (i.e. when accuracy = 1), the proportion of true positives (δ) will be roughly equal to the proportion of false positives (β).

Accuracy = 10
Accuracy = 3
Accuracy = 1
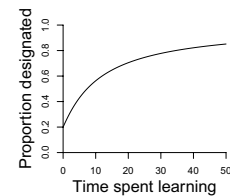
True positive rate, δ
False positive rate, β
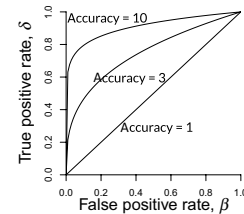
We can *delay designation* and *spend more time learning* about the species' habitat.

With learning, designation will have GREATER ACCURACY.

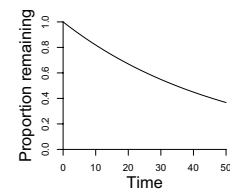76% of habitat could be designated

A higher proportion of total habitat is included in the designated area.

Proportion designated
Time spent learning

However, delays could lead to more HABITAT LOSS.

25% of habitat has been lost

Less habitat is available to protect when designation occurs.

Proportion remaining
Time

There is often a *trade-off* between *improving accuracy* and *minimizing habitat loss* before designation.

The time spent learning affects the AREA OF HABITAT DESIGNATED.

76% of remaining habitat, or 57% of the initial habitat area, is designated

We can maximize the area of habitat designated by determining the OPTIMAL TIME TO SPEND LEARNING.

Proportion designated
Maximum area designated
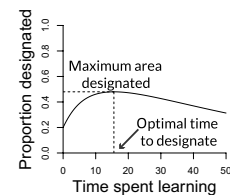Optimal time to designate
Time spent learning

FIGURE 1 Effect of improved accuracy and habitat loss over time on the area of species' habitat protected at designation, expressed as a proportion of the initial habitat area available. To achieve the most benefits for conservation, we need to determine the optimal amount of time to spend learning that maximizes the area of habitat protected

of portions of the species' habitat (i.e. false-negative errors) and/or the unnecessary inclusion of non-habitat areas (i.e. false-positive errors; Fielding & Bell, 1997). It may therefore be desirable to delay designation and collect more data to improve the accuracy of habitat models. As accuracy improves with learning, the ability to correctly identify species' habitat increases, thus reducing false-negative and false-positive errors. Improved accuracy results in a greater proportion of the species' habitat that is correctly protected than if designation occurred immediately, and reduces the associated social or economic costs of unnecessary designation, which may be important considerations in the context of limited budgets and competing demands for space and natural resources.

2.2 | State model and objective

Maximizing the area of a species' habitat that is protected requires determining the value of the control variable—the amount of time to spend learning—that optimizes the trade-off between improving accuracy and minimizing habitat loss. To determine the optimal time to spend learning (t^*) before designating habitat, we consider a set of sites that are either habitat (H) or matrix (M) (Figure 1). Here, we define habitat as having the conditions and resources that allow a species to occupy a site or location (Hall et al., 1997)—that is, all areas that are suitable and are needed for the population to persist,

including areas required for connectivity between patches or to buffer against edge effects, are considered as part of the 'habitat' for that species, while the matrix consists of the surrounding non-habitat areas that are unsuitable for the species (Fahrig, 2001).

When designation occurs, areas presumed to be habitat, based on information available at the time, are protected (P). The conservation objective of maximizing the area of habitat protected at designation can be expressed as.

$$\max_{t \in [0,T]} x_{H \cap P}(t), \qquad (1)$$

where $x_{H \cap P}(t)$ represents the total area (x) of habitat that is correctly identified and protected ($H \cap P$) if designation occurs at time t. The value of t that satisfies Equation 1 is the optimal time to spend learning before designating protected habitat (t^*). To find this optimal time, we need to determine how $x_{H \cap P}$ changes over time due to learning and habitat loss.

2.3 | Model dynamics

2.3.1 | Improved accuracy

We assumed that knowledge of the habitat of a species improves over time as more data are collected—that is, that the accuracy (a) of habitat models improves over time with more research. In the absence of information about the rate at which we gain this knowledge, we used a range of hypothetical learning curves (Table 1; Figure 2a) to model the change in a as a function of the time spent learning. We used functions similar in shape and value ranges as those in Martin et al. (2017), which were demonstrated to encompass empirical estimates of a for several bird and mammal species (Martin et al., 2017). We used a linear function to represent a proportional increase in model accuracy with the amount of time spent learning. We used hyperbolic curves to describe the rate of learning about species, such as habitat specialists, for which reasonably high model accuracy may be achieved quickly with relatively little investment (Franklin et al., 2009; Stockwell & Peterson, 2002). Finally, we used s-shaped or sigmoid curves to represent slow initial rates of learning followed by more rapid increase, which may occur once some threshold number of species observations has been reached. This may be more likely for widely distributed or more cryptic species for which evaluating habitat suitability may be more difficult or require more data (Franklin et al., 2009).

We used receiver operating characteristic (ROC) curves (Figure 2b) of the form $\delta = \beta^{1/a}$ to model how map accuracy (a) influences the proportion of the total habitat area correctly identified and protected

(true-positive rate, δ), where β represents the proportion of non-habitat matrix incorrectly identified as habitat (false-positive rate). The total area of a species' habitat that is correctly protected at time t is determined by:

$$x_{H \cap P}(t) = \delta(t) \cdot x_H(t), \qquad (2)$$

where x_H is the total area of habitat available. We set the false-positive rate β to a constant value that represents the decision-maker's tolerance for the amount of non-habitat unnecessarily protected through designation. Setting higher β values will result in higher δ values for the same value of a and therefore greater conservation benefit, but potentially at higher financial costs or other social or political consequences. In contrast, with low β values, greater map accuracy and therefore more time spent learning is required to achieve the same δ (Figure 2b). In this way, β acts as a constraint that represents the costs—including non-monetary costs—of incorrect designation that a decision-maker is willing to tolerate in the effort to protect the species and its habitat.

Receiver operating characteristic curves have been used to describe the accuracy of species habitat models, with the area under the ROC curve (AUC) as a relative measure of the predictive performance of the model (Pearce & Ferrier, 2000). AUC values range from 0.5 for a model that fails to discriminate between habitat and non-habitat areas (i.e. little or no better than a random classification), to 1.0 for perfect discrimination (Fielding & Bell, 1997; Pearce & Ferrier, 2000). The AUC for the ROC curves used here can be calculated as AUC(t)=a(t)/(a(t)+1) (Baxter & Possingham, 2011), which yields an AUC value of 0.5 when a(t) = 1. Therefore, to simulate a random habitat designation when little or no information is initially available, we defined the learning curve functions (Table 1) such that a = 1 at time t = 0, before any learning occurs.

2.3.2 | Ongoing habitat loss

When protection is delayed to improve accuracy, the species and its habitat are vulnerable to ongoing threats. We used an exponential decay function to model habitat loss over time as proportional to the current area, according to:

$$\frac{dx_H}{dt} = -\lambda x_H, \qquad (3)$$

where λ is the proportional rate of habitat loss (Figure 3). The total area of habitat available at the time of designation is therefore predicted by:

$$x_H(t) = x_H(0) e^{-\lambda t}, \qquad (4)$$

TABLE 1 Hypothetical learning curves that model the change in information accuracy (a) over time (t)

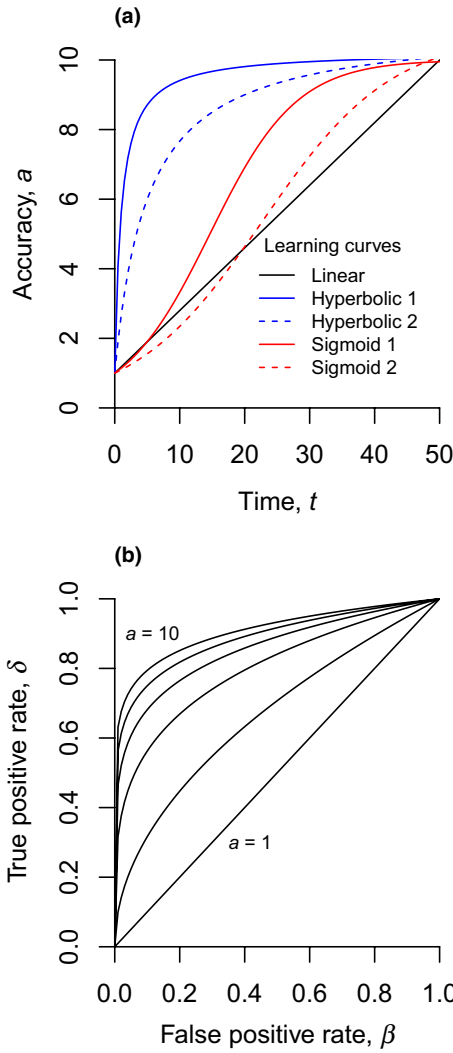| General equation	Curve type	Shape parameters	$\dfrac{da}{dt}$
$a(t) = mt + 1$	Linear	$m = 9/50$	m
$a(t) = \dfrac{bt}{t+m} + 1$	Hyperbolic	1) $m = 1, b = 9.25$ 2) $m = 5, b = 10$	$\dfrac{bm}{(t+m)^2}$
$a(t) = \dfrac{b}{1 + (b-1)e^{-mt}}$	Sigmoid	1) $m = 0.15, b = 10$ 2) $m = 0.10, b = 10.75$	$\dfrac{be^{-mt}(b-1)m}{\left(1 + e^{-mt}(b-1)\right)^2}$

(a)

(b)

FIGURE 2 Hypothetical curves modelling (a) the increase in the accuracy, *a*, of habitat identification over time as learning occurs, and (b) the rate of true positives, δ, relative to the rate of false positives, β, for different values of *a*, based on ROC curves of the form $\delta = \beta^{1/a}$

where $x_H(0)$ is the area of species' habitat initially available (henceforth, 'initial habitat area') at the time when the need to designate habitat is first recognized (i.e. $t = 0$).

2.4 | Trade-off between improving accuracy and minimizing habitat loss

When we account for both improved accuracy and habitat loss over time, the area of habitat designated at time *t* becomes:

$$x_{H\cap P}(t) = \beta^{\frac{1}{a(t)}} \cdot x_H(0) \, e^{-\lambda t}. \tag{5}$$

Finding t^* requires knowledge of the initial habitat area, $x_H(0)$, which we do not know. However, as $x_H(0)$ is a constant, maximizing the proportion of the initial habitat area that is protected, or, $x_{H\cap P}(t)/x_H(0)$, also
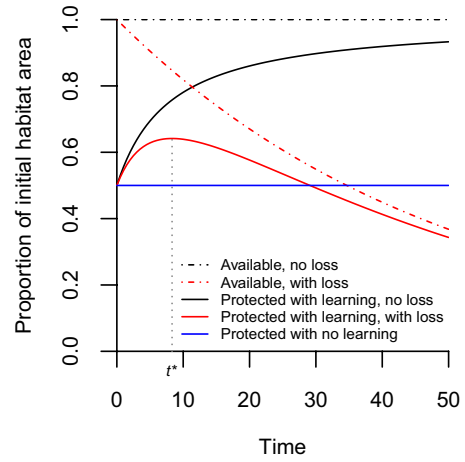

FIGURE 3 Proportion of initial habitat area correctly identified and protected over time when no habitat loss is occurring (black solid line) and when habitat area is decreasing by $\lambda = 0.02$ (red solid line), assuming a linear learning curve and false-positive rate $\beta = 0.5$. The optimal time, t^*, is the amount of time to spend learning that maximizes the proportion of initial habitat area that is correctly identified and protected

maximizes the value $x_{H\cap P}(t)$. We can express the objective function instead as:

$$\max_{t \in [0,T]} s(t) = \beta^{\frac{1}{a(t)}} \cdot e^{-\lambda t}, \tag{6}$$

where $s(t) = x_{H\cap P}(t)/x_H(0)$. The optimal time to spend learning (t^*) can then be determined by:

$$t^* = \arg\max_{t \in [0,T]} \beta^{\frac{1}{a(t)}} \cdot e^{-\lambda t}, \tag{7}$$

for different functions of $a(t)$, where values of λ and β reflect context-specific estimates of the rate of habitat loss and the rate of false-positive errors acceptable to the decision-maker, respectively.

2.5 | Analytical solution

The optimal time (t^*) can be determined numerically by calculating $s(t)$ for each $t \epsilon [0,T]$ to find the value of *t* that maximizes $s(t)$ (Figure 3). For some learning curve functions, such as those used here (Table 1), t^* may also be found analytically (Supporting Information). For each type of learning curve, $a(t)$, and its derivative, da/dt, the optimal time to designate habitat for protection is:

$$\text{Linear} \quad t^* = \begin{cases} \dfrac{\sqrt{-\frac{m\ln\beta}{\lambda}} - 1}{m}, & 0 < \lambda < -m\ln\beta \\ 0, & \lambda \geq -m\ln\beta \\ T, & \lambda \leq 0 \end{cases} ,$$

$$\text{Hyperbolic} \quad t^* = \begin{cases} \dfrac{\sqrt{\frac{-bm\ln\beta}{\lambda}} - m}{b+1}, & 0 < \lambda < -\dfrac{b}{m}\ln\beta \\ 0, & \lambda \geq -\dfrac{b}{m}\ln\beta \\ T, & \lambda \leq 0 \end{cases},$$

$$\text{Sigmoid} \quad t^* = \begin{cases} -\dfrac{\ln\left(\frac{-\lambda b}{(b-1)m\ln\beta}\right)}{m}, & 0 < \lambda < -\dfrac{(b-1)m}{b}\ln\beta \\ 0, & \lambda \geq -\dfrac{(b-1)m}{b}\ln\beta \\ T, & \lambda \leq 0 \end{cases},$$

where b and m are parameters that determine the shape of the learning curve (Table 1).

2.6 | Accounting for uncertainty in how we learn about habitat

So far, we have determined the optimal amount of time to spend learning for a range of plausible learning curves. However, in practice, determining which learning curve best represents the increase in accuracy of our habitat model over time can be difficult without prior knowledge about the species and its habitat use. When we use the wrong learning curve to determine the optimal time to spend learning—for example, if we used a sigmoid curve to estimate t^* when the true learning curve is hyperbolic—designation of habitat for protection may occur too early or too late, resulting in a lower proportion of initial habitat area correctly designated due to lower map accuracy or habitat loss. To account for this uncertainty, we propose the use of the minimax regret decision rule (Kouvelis & Yu, 1997) to identify which learning curve minimizes the worst possible consequence, or regret, of using the wrong learning curve. Here, we define regret (Δs_{ij}) as the proportion of species' habitat we might fail to protect when we assume learning occurs following curve j when the true learning curve ($a(t)$) is i, relative to what we could protect when we know the true learning curve. That is,

$$\Delta s_{ij} = s\left(t_i^* \mid a(t) = i\right) - s\left(t_j^* \mid a(t) = i\right), \tag{8}$$

where t_i^* and t_j^* are the optimal time to spend learning before designating habitat for each learning curve i or j, respectively. The function $s(t_i^* \mid a(t) = i)$ gives the proportion of initial habitat area that would be protected when we designate at the optimal time for the correct learning curve i, while $s(t_j^* \mid a(t) = i)$ is the proportion protected when we designate at the optimal time for learning curve j instead.

For each learning curve under consideration, we used the analytical solutions to determine t^* for a given λ and β. For each learning curve j, we then identified the maximum regret of using that curve across the different scenarios of true learning curve i, and determined which learning curve j will result in the least maximum regret:

$$\min_j \max_i \left(\Delta s_{ij}\right). \tag{9}$$

2.7 | Designating protected habitat areas for koalas and northern abalone

We demonstrate the use of the analytical solutions to determine the optimal time to spend learning before protecting habitat for two threatened species, the koala in South East Queensland, Australia, and the northern abalone in the west coast of Canada. Given the lack of information about the rates of learning for either species or the rates of false-positive errors considered acceptable, we used a range of hypothetical learning curves (Table 1) and β values of 0.2 and 0.5.

The koala is a charismatic marsupial endemic to eastern Australia, and is currently listed as vulnerable in Australia. Declines in koala populations have been attributed primarily to habitat loss and fragmentation, either directly through the loss of food resources or indirectly through higher incidences of vehicle strikes and dog attacks at higher land-use intensities (Natural Resource Management Ministerial Council, 2009). There is also some evidence to suggest that disease mortality in koalas may be partially attributed to increased physiological stress and reduced body condition resulting from habitat loss and fragmentation (McAlpine et al., 2017). We looked at identifying and designating protected habitat areas for koala populations in the Koala Coast, an area in southeast Queensland with a high concentration of koalas (Rhodes et al., 2011). The Koala Coast is currently under threat from habitat loss and urbanization (Rhodes et al., 2011); between 1997 and 2003, an estimated 4.7% of koala habitat within this area was lost (Preece, 2007). Surveys conducted between 1996 and 2008 also indicate a decline in the koala populations in this region. We derived a value of $\lambda = 0.008$ from the estimated rate of habitat loss, and used this value to determine the optimal amount of time that could be spent learning before identifying habitat areas for protection.

The northern abalone is a marine mollusc found along the west coast of North America, and is listed as endangered in Canada. The northern abalone has suffered considerable population declines due to commercial fishing during the last century, resulting in fishery closures in the 1990s (Sloan & Breen, 1988). However, populations continue to decline due to illegal harvest, or poaching, with instantaneous poaching mortality estimated to range from 0.1 to 0.7 (Campbell, 2000), corresponding to about 10%–50% mortality per year. Here, we assumed that (a) poaching mortality is directly proportional to the area poached and (b) designating abalone habitat as protected will result in successful prevention of further poaching at that site due to an increase in monitoring and other anti-poaching activities, thus resulting in a corresponding decrease in poaching mortality. Therefore, we used poaching mortality estimates of 0.1–0.5 as λ to simulate habitat loss.

3 | RESULTS

3.1 | General trends

We used three variables to predict the optimal time to spend learning (t^*) before designating protected habitat: (a) the rate of habitat loss (λ); (b) the rate of false-positive errors (β) considered as acceptable; and (c)

the change in accuracy over time ($a(t)$) due to learning (i.e. the shape of the learning curves). When no habitat loss or other threatening processes are occurring (i.e. $\lambda \leq 0$), there is no cost to waiting and the optimal strategy is to delay designation until the maximum possible map accuracy is achieved or until a specified maximum time period T. As the rate of habitat loss increases, t^* decreases (Figures 4 and 5). Accepting higher threshold β values in habitat mapping also results in lower t^* values (Figures 4 and 5), as the level of accuracy required is lower and can be achieved sooner. Finally, t^* is shorter when initial learning rates are fast, as in hyperbolic curves, than when initial learning rates are slow, as in the sigmoid or linear learning curves. Moreover, earlier habitat designation means less opportunity for habitat loss to continue, resulting in a higher $s(t^*)$ when learning rates are fast.

The optimal time to spend learning was more sensitive to the shape of the learning curve and the false-positive rate at low rates of habitat loss compared to high rates of habitat loss. At low rates of habitat loss, setting a lower threshold for false-positive error rate, for example, $\beta = 0.2$, resulted in a wider range of optimal learning periods for the different learning curves (Figures 4 and 5). However, at high rates of habitat loss (e.g. $\lambda \geq 0.5$), t^* tends to converge to 0 years regardless of either the learning curve or β value (Figure 5). Specifically, the analytical solutions indicate that the optimal decision is to protect habitats immediately when the rate of habitat loss exceeds a value determined by the initial rate of learning (i.e. da/dt at $t = 0$) and the amount of time needed to reduce the risk of protecting

non-habitat unnecessarily (i.e. the false-positive error rate) to an acceptable rate (i.e. $\ln \beta$).

The wide range in t^* for different learning curves (Figure 4) means that using the incorrect learning curve leads to a suboptimal outcome: a lower proportion of the initial habitat area that is correctly identified and protected, relative to the maximum that could be protected when we use the true learning curve to determine t^* (Table 2). At low λ values, the proportion of habitat missed is lower when we incorrectly assume a slow learning curve than in the converse scenario (Table 2). However, at higher λ, the opposite is true—the proportion of habitat missed is lower when we assume a fast learning curve when the true rate is slow (Table 2).

Based on the minimax regret decision rule, the best strategy at lower rates of habitat loss ($\lambda = 0.01$) is to use a Sigmoid 1 curve, which results in a maximum of 0.102 proportion of habitat missed (Table 2). With higher rates of habitat loss ($\lambda = 0.1$), however, the best strategy is to assume a Hyperbolic 2 curve, resulting in a maximum proportion missed of 0.032 (Table 2). When we are uncertain about both learning curves and the rate of habitat loss, assuming a linear curve will minimize the proportion of habitat missed across all possible scenarios.

3.2 | Example 1: Koala

Based on the range of learning curves specified in Table 1 and the estimated rate of habitat loss in the Koala Coast in 2003, the optimal
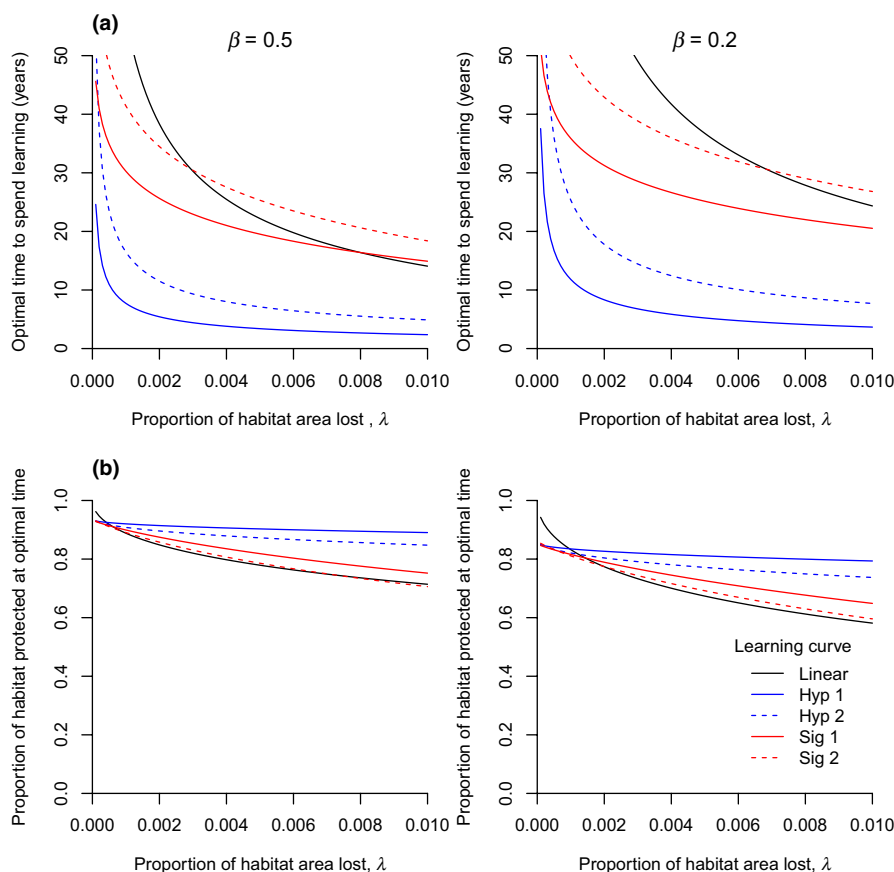

FIGURE 4 (a) Optimal amount of time to spend learning before protecting habitat for the koala at different rates of habitat loss for five different learning curves and for false-positive rates $\beta = 0.5$ and $\beta = 0.2$, and (b) the corresponding proportion of the initial habitat area protected when the optimal amount of time is spent learning before designation

FIGURE 5 (a) Optimal amount of time to spend learning before protecting habitat for the northern abalone at different rates of poaching for five different learning curves and for false-positive rates $\beta = 0.5$ and $\beta = 0.2$, and (b) the corresponding proportion of the initial habitat area protected when the optimal amount of time is spent learning before designation
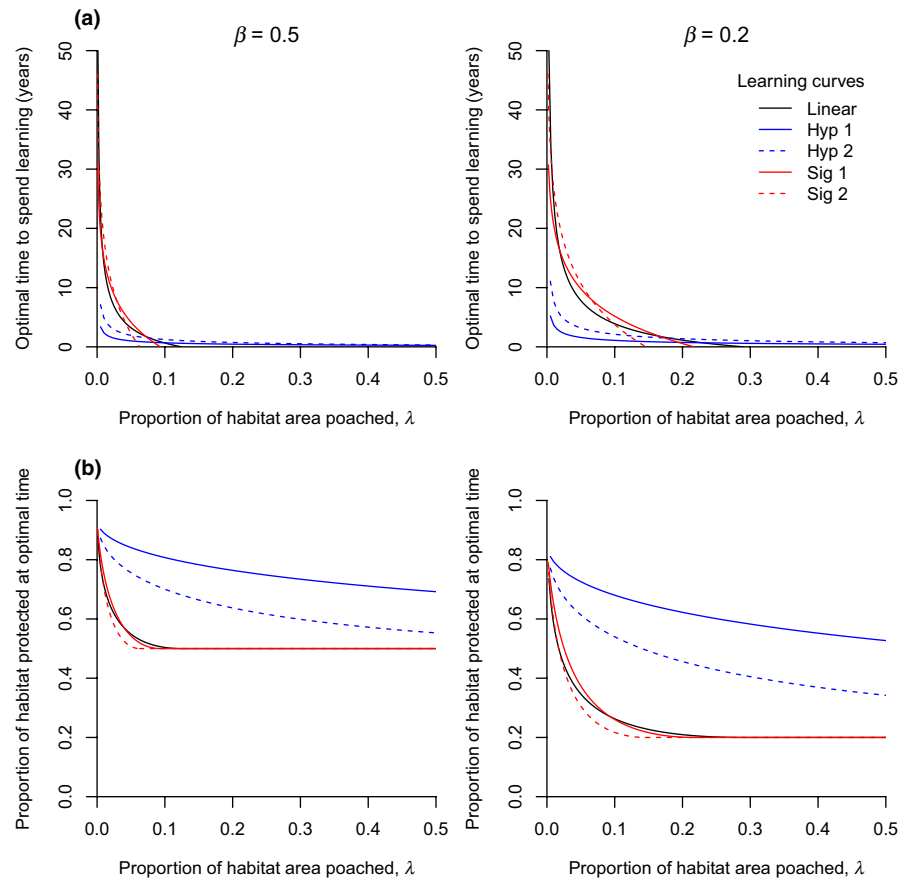

TABLE 2 The proportion of habitat missed (Δs_{ij}), given the true learning curve (*i*, rows) and $\beta = 0.2$, when the wrong learning curve (*j*, columns) is assumed, relative to the maximum proportion that could have been protected when the true learning curve is used (values along the diagonal, in parentheses). For a given λ, values in bold face indicate the worst outcome (maximum regret) of assuming curve *j*, while bold italicized values indicate the best of the worst outcomes across all possible learning curves (the minimum maximum regret)

		Assumed, *j*				
λ	True curve, *i*	Hyperbolic 1	Hyperbolic 2	Sigmoid 1	Sigmoid 2	Linear
0.01	Hyperbolic 1	(0.793)	0.016	*0.102*	**0.143**	**0.127**
	Hyperbolic 2	0.029	(0.737)	0.056	0.092	0.078
	Sigmoid 1	0.293	0.149	(0.649)	0.014	0.006
	Sigmoid 2	**0.294**	**0.190**	0.014	(0.596)	0.002
	Linear	0.216	0.110	0.003	0.001	(0.581)
0.1	Hyperbolic 1	(0.680)	*0.032*	**0.184**	**0.116**	**0.122**
	Hyperbolic 2	0.036	(0.540)	0.082	0.034	0.038
	Sigmoid 1	**0.037**	0.020	(0.261)	0.004	0.003
	Sigmoid 2	0.008	0.003	0.002	(0.217)	0.000
	Linear	0.029	0.011	0.004	0.000	(0.263)

time to spend learning before protecting koala habitat may vary from 0 to 35 years (Figure 4). Assuming $\beta = 0.2$, and without prior knowledge about the rate at which we learn about koala habitat, the minimax regret strategy for $\lambda = 0.008$ would be to use a Sigmoid 1 learning curve, which predicts an optimal time of 29 years and result in up to around 71% of initial habitat area correctly designated, with a maximum regret (Δs_{ij}) of 0.09 if the true learning curve is different.

Given that the koala is a specialist folivore (Phillips et al., 2000), however, the shape of the learning curve is more likely to be hyperbolic, which predicts much shorter optimal learning times—around 6 years or less when $\beta = 0.5$, or <10 years when $\beta = 0.2$ (Figure 4). Constraining the range of possible learning curves to the two hyperbolic curves, the minimax regret strategy at $\beta = 0.2$ would be to assume a Hyperbolic 2 learning curve. This reduces the maximum

regret to around 0.015 if the true learning curve is Hyperbolic 1, and results in the designation of up to 78% of initial habitat area.

3.3 | Example 2: Northern abalone

When poaching mortalities were used directly as a measure of habitat loss, the optimal time to spend learning about northern abalone at low poaching levels, $\lambda = 0.1$, varied from 0 to around 6 years, depending on the chosen false-positive rate β value and learning curve (Figure 5). At this poaching rate, the minimax regret strategy when $\beta = 0.2$ would be to assume a Hyperbolic 2 curve and designate after 2 years, resulting in the correct designation of up to 65% of the initial habitat area, and a maximum of around 0.03 of habitat missed if Hyperbolic 2 is not the true learning curve. The range in the optimal learning period decreased rapidly to between 0 and 2 years at $\lambda = 0.3$ and to between 0 and 1 year at high poaching levels, $\lambda = 0.5$. At these higher levels of poaching, the minimax regret strategy would be to assume a Hyperbolic 1 learning curve, and designate as quickly as possible—around 1 year or less.

4 | DISCUSSION

Once enacted, making substantive changes to the designation of protected habitat can be challenging due to the amount of effort and financial cost involved in designation, and potential opposition from various stakeholders. This provides a strong incentive to get it right the first time. However, we can rarely afford to wait for perfect information before protecting habitats. Determining the optimal timing for habitat protection is important in maximizing the effectiveness of protection in the long term. Here, we demonstrated the use of estimated rates of habitat loss and hypothetical learning curves to determine the amount of time we can afford to spend learning, before learning no longer results in better protection because too much habitat has been lost, or the threats within the habitat have left the population perilously low. Our results corroborate those of Martin et al. (2017), with the added advantages that our model is particularly useful when little information is currently known about species and their habitat use or preferences, and that it can be solved analytically. Furthermore, by explicit consideration of the trade-off between the benefit of better accuracy and the cost of waiting, our approach contributes to improving the transparency and defensibility of decisions regarding habitat protection.

4.1 | Rates of learning

Given the variation in the optimal time to spend learning for different learning curves, particularly at lower rates of habitat loss or population decline, quantifying the rate in which we learn about habitat, for instance, by evaluating the improvements in the accuracy of species distribution maps as a result of increased survey effort (Grantham

et al., 2009; Stockwell & Peterson, 2002) can improve our models of initial learning over time. When uncertainties over the correct form of the learning curve cannot be resolved, we could apply a minimax regret decision rule to identify which learning curve would minimize the maximum amount of habitat that we might 'miss', or fail to designate, by making the wrong assumption about the learning curve. At lower rates of habitat loss, the minimax regret strategy is to assume a slower initial rate of learning, such as in a linear or sigmoidal curve. However, at high rates of habitat loss, the minimax regret strategy would be to use a fast learning curve, such as the hyperbolic curves considered here and to direct effort towards learning as much and as quickly as possible until designation occurs.

The functional forms used to model the change in accuracy over time all assume that no prior knowledge about habitat exist, that is, that $a = 1$ when $t = 0$ (Figure 4). These functions can be modified, or different ones can be used, to reflect scenarios where some information is already available to accurately identify habitat with greater than random probability, or when $a > 1$ at $t = 0$. Such scenarios will likely predict shorter optimal times for learning. At high rates of habitat loss, however, the shape of the learning curve no longer has a strong influence on the optimal time to spend learning; therefore, having prior information about species and their habitat is unlikely to change the predictions significantly.

The functional forms also assume that learning can occur—that is, that the will and the capacity to conduct additional research exists. When the ability to conduct research is constrained, for instance, due to mandatory timelines or limited research budgets, the amount of time to spend on research can be optimized subject to the constraint—by setting the maximum length of time (T) that can be spent on research based on the specified timelines or the amount of money available for research. However, when additional research is not feasible, no improvement in accuracy over time can be expected, and there is no benefit of delaying habitat protection.

4.2 | False-positive errors

Our approach requires the specification of the rate of false-positive errors (β)—that is, the proportion of non-habitat we are willing to protect unnecessarily in exchange for an earlier designation. When a high β value can be tolerated, habitat protection can occur sooner. In this scenario, the level of accuracy required is lower and may be achieved faster. In contrast, if a low β value is required, a higher level of accuracy is needed to correctly identify the same proportion of species habitat (δ), thus requiring a longer learning period.

Despite a number of studies on evaluating the accuracy of predictive maps, and particularly of maps that predict species occurrences (Allouche et al., 2006), there are no prescriptive guidelines for selecting acceptable error rates. Instead, decisions depend on the intended purpose of mapping and the relative costs of the two types of errors (Fielding & Bell, 1997). For instance, when the consequences of losing habitat and species are deemed unacceptable, a decision-maker may be willing to allow a higher proportion

of non-habitat to be included in the designation (i.e. a higher β) in exchange for protecting habitats sooner. In contrast, one who is averse to incurring potential economic and political costs of inaccurate designation may choose to set a lower β. Our approach makes the selection of a false-positive error threshold explicit, which helps encourage greater consideration of the costs of different types of errors and associated preferences of decision-makers for different outcomes.

4.3 | Rates of habitat loss

We used estimated rates of habitat loss or other threatening processes to simulate the impact of these threats over time on the long-term persistence of a given species or population. The main assumptions are as follows: that all habitat areas are of equal quality; that unprotected habitat areas are equally likely to be lost, regardless of the type or quality; and that a decrease in habitat area will result in a proportional decrease in the probability of persistence of the species. However, the relationship between the amount of habitat available and the long-term persistence of species may not be as straightforward—the effect of habitat loss on persistence, for instance, may be greater as the amount of habitat remaining to support the species becomes smaller (Fahrig, 2001; Swift & Hannon, 2010). Similarly, the loss of a given habitat area may have a greater negative impact on species persistence if that habitat is of higher quality (Heinrichs et al., 2010; Thomas et al., 2001). Furthermore, our model does not evaluate whether the resulting area of habitat protected at the time of designation will be of sufficient size and quality to allow for the long-term persistence of the species. When possible, quantifying the impact of habitat loss or other threats on population persistence can help provide a more accurate estimate of the optimal time to spend on learning, as well as the total amount and quality of habitat that should be protected, that will maximize the probability of species persistence.

5 | CONCLUSIONS

Our work confirms that timely decisions about conservation interventions can save species from extinction (Martin et al., 2012). Recently, there has been growing recognition of the implicit trade-off that occurs when conservation interventions are delayed due to calls for more data. We demonstrate that while delaying habitat protection may sometimes be beneficial, it is often better to protect habitats immediately rather than wait for more information when rates of habitat loss are high. The optimization approach presented here predicts the optimal time to spend learning before protecting habitat by using habitat loss as the cost of learning and a threshold false-positive rate as a constraint on the amount of non-habitat included in designation. In particular, the use of simple functional forms to model the rates of learning and habitat loss makes it possible to do so when little information is available

to develop more complex models. Moreover, this approach can easily be generalized to address the impact of other threatening processes, such as harvesting or the invasion of a species with negative impacts on a species of conservation concern, provided that the same metric is used for both the costs and the benefits of learning. Our method explicitly acknowledges the trade-off between the benefit of delaying protection to improve accuracy and the costs of such delays, and provides evidence to support timely decisions to protect habitat, despite potential uncertainty about what constitutes that habitat.

CONFLICT OF INTEREST
The authors declare that there are no conflicts of interest.

AUTHORS' CONTRIBUTIONS
A.E.C. and H.P.P. conceived the ideas and developed the model and the analytical solutions; A.E.C. also developed the case studies, conducted the analysis and led the writing of the manuscript; I.C. contributed to the development of the model and analytical solutions, and provided critical feedback on the application to the case studies; and T.G.M. contributed to the selection and development of the case studies. All authors contributed to the structure and content of the manuscript, and gave final approval for publication.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13770.

DATA AVAILABILITY STATEMENT
An R package developed to implement the analytical solutions, and the R scripts used to generate the results and figures for this analysis are available from an online repository https://github.com/aecamaclang/opttiming (Camaclang, 2021).

ORCID
Abbey E. Camaclang https://orcid.org/0000-0001-7334-3483
Iadine Chadès https://orcid.org/0000-0002-7442-2850
Tara G. Martin https://orcid.org/0000-0001-7165-9812
Hugh P. Possingham https://orcid.org/0000-0001-7755-996X

REFERENCES

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Baxter, P. W. J., & Possingham, H. P. (2011). Optimizing search strategies for invasive pests: Learn before you leap. *Journal of Applied Ecology*, 48(1), 86–95. https://doi.org/10.1111/j.1365-2664.2010.01893.x

Bohorquez, J. J., Dvarskas, A., & Pikitch, E. K. (2019). Filling the data gap – A pressing need for advancing MPA sustainable finance. *Frontiers in Marine Science*, 6, 45. https://doi.org/10.3389/fmars.2019.00045

Camaclang, A. E. (2021). opttiming: Predicting the optimal time to spend learning. *Zenodo*, https://doi.org/10.5281/zenodo.5701680

Campbell, A. (2000). Review of northern abalone, *Haliotis kamtschatkana*, stock status in British Columbia. *Canadian Special Publication of Fisheries and Aquatic Sciences*, 130, 41–50.

Chadès, I., McDonald-Madden, E., McCarthy, M. A., Wintle, B., Linkie, M., & Possingham, H. P. (2008). When to stop managing or surveying cryptic threatened species. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 13936–13940. https://doi.org/10.1073/pnas.0805265105

Chadès, I., Ponce Reyes, R., Nicol, S., Pascal, L., Fletcher, C., Cresswell, I., & Carwardine, J. (2019). *An integrated spatial prioritisation plan for the Saving our Species program*. CSIRO. https://doi.org/10.25919/5d03eeb896667

Fahrig, L. (2001). How much habitat is enough? *Biological Conservation*, 100(1), 65–74. https://doi.org/10.1016/S0006-3207(00)00208-1

Field, S. A., O'Connor, P. J., Tyre, A. J., & Possingham, H. P. (2007). Making monitoring meaningful. *Austral Ecology*, 32(5), 485–491. https://doi.org/10.1111/j.1442-9993.2007.01715.x

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1), 38–49. https://doi.org/10.1017/S0376892997000088

Franklin, J., Wejnert, K. E., Hathaway, S. A., Rochester, C. J., & Fisher, R. N. (2009). Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. *Diversity and Distributions*, 15(1), 167–177. https://doi.org/10.1111/j.1472-4642.2008.00536.x

García-Barón, I., Giakoumi, S., Santos, M. B., Granado, I., & Louzao, M. (2021). The value of time-series data for conservation planning. *Journal of Applied Ecology*, 58(3), 608–619. https://doi.org/10.1111/1365-2664.13790

Grantham, H. S., Wilson, K. A., Moilanen, A., Rebelo, T., & Possingham, H. P. (2009). Delaying conservation actions for improved knowledge: How long should we wait? *Ecology Letters*, 12(4), 293–301. https://doi.org/10.1111/j.1461-0248.2009.01287.x

Hall, L. S., Krausman, P. R., & Morrison, M. L. (1997). The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, 25(1), 173–182.

Heinrichs, J. A., Bender, D. J., Gummer, D. L., & Schumaker, N. H. (2010). Assessing critical habitat: Evaluating the relative contribution of habitats to population persistence. *Biological Conservation*, 143(9), 2229–2237. https://doi.org/10.1016/j.biocon.2010.06.009

Kouvelis, P., & Yu, G. (1997). *Robust discrete optimization and its applications*. Springer. https://doi.org/10.1007/978-1-4757-2620-6

Leclère, D., Obersteiner, M., Barrett, M., Butchart, S. H. M., Chaudhary, A., De Palma, A., DeClerck, F. A. J., Di Marco, M., Doelman, J. C., Dürauer, M., Freeman, R., Harfoot, M., Hasegawa, T., Hellweg, S., Hilbers, J. P., Hill, S. L. L., Humpenöder, F., Jennings, N., Krisztin, T., ... Young, L. (2020). Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature*, 585(7826), 551–556. https://doi.org/10.1038/s41586-020-2705-y

Lindenmayer, D. B., Piggott, M. P., & Wintle, B. A. (2013). Counting the books while the library burns: Why conservation monitoring programs need a plan for action. *Frontiers in Ecology and the Environment*, 11(10), 549–555. https://doi.org/10.1890/120220

Marescot, L., Chapron, G., Chadès, I., Fackler, P. L., Duchamp, C., Marboutin, E., & Gimenez, O. (2013). Complex decisions made simple: A primer on stochastic dynamic programming. *Methods in Ecology and Evolution*, 4(9), 872–884. https://doi.org/10.1111/2041-210X.12082

Martin, T. G., Camaclang, A. E., Possingham, H. P., Maguire, L. A., & Chadès, I. (2017). Timing of protection of critical habitat matters. *Conservation Letters*, 10(3), 308–316. https://doi.org/10.1111/conl.12266

Martin, T. G., Nally, S., Burbidge, A. A., Arnall, S., Garnett, S. T., Hayward, M. W., Lumsden, L. F., Menkhorst, P., McDonald-Madden, E., & Possingham, H. P. (2012). Acting fast helps avoid extinction. *Conservation Letters*, 5(4), 274–280. https://doi.org/10.1111/j.1755-263X.2012.00239.x

McAlpine, C., Brearley, G., Rhodes, J., Bradley, A., Baxter, G., Seabrook, L., Lunney, D., Liu, Y., Cottin, M., Smith, A. G., & Timms, P. (2017). Time-delayed influence of urban landscape change on the susceptibility of koalas to chlamydiosis. *Landscape Ecology*, 32(3), 663–679. https://doi.org/10.1007/s10980-016-0479-2

McDonald-Madden, E., Bode, M., Game, E. T., Grantham, H., & Possingham, H. P. (2008). The need for speed: Informed land acquisitions for conservation in a dynamic property market. *Ecology Letters*, 11(11), 1169–1177. https://doi.org/10.1111/j.1461-0248.2008.01226.x

Natural Resource Management Ministerial Council. (2009). *National koala conservation and management strategy 2009–2014*. Department of the Environment, Water, Heritage and the Arts. https://www.environment.gov.au/biodiversity/threatened/publications/national-koala-conservation-mgt-strategy-2009-2014

Naujokaitis-Lewis, I., Pomara, L. Y., & Zuckerberg, B. (2018). Delaying conservation actions matters for species vulnerable to climate change. *Journal of Applied Ecology*, 55(6), 2843–2853. https://doi.org/10.1111/1365-2664.13241

Pascal, L., Memarzadeh, M., Boettiger, C., Lloyd, H., & Chadès, I. (2020). A Shiny R app to solve the problem of when to stop managing or surveying species under imperfect detection. *Methods in Ecology and Evolution*, 11(12), 1707–1715. https://doi.org/10.1111/2041-210X.13501

Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133(3), 225–245. https://doi.org/10.1016/S0304-3800(00)00322-7

Perrings, C., & Brock, W. (2009). Irreversibility in economics. *Annual Review of Resource Economics*, 1(1), 219–238. https://doi.org/10.1146/annurev.resource.050708.144103

Phillips, S., Callaghan, J., & Thompson, V. (2000). The tree species preferences of koalas (*Phascolarctos cinereus*) inhabiting forest and woodland communities on Quaternary deposits in the Port Stephens area, New South Wales. *Wildlife Research*, 27(1), 1–10. https://doi.org/10.1071/WR98054

Pindyck, R. S. (2002). Optimal timing problems in environmental economics. *Journal of Economic Dynamics and Control*, 26(9), 1677–1697. https://doi.org/10.1016/S0165-1889(01)00090-2

Preece, H. J. (2007). *Monitoring and modelling threats to koala populations in rapidly Urbanising landscapes: Koala Coast, South East Queensland, Australia* (PhD Thesis). The University of Queensland. https://doi.org/10.14264/uql.2015.742

Rhodes, J. R., Ng, C. F., de Villiers, D. L., Preece, H. J., McAlpine, C. A., & Possingham, H. P. (2011). Using integrated population modelling to quantify the implications of multiple threatening processes for a rapidly declining population. *Biological Conservation*, 144(3), 1081–1088. https://doi.org/10.1016/j.biocon.2010.12.027

Runge, M. C., Converse, S. J., Lyons, J. E., & Smith, D. R. (2020). *Structured decision making: Case studies in natural resource management*. Johns Hopkins University Press. https://doi.org/10.1353/book.74951

Sloan, N. A., & Breen, P. A. (1988). Northern abalone, *Haliotis kamtschatkana*, in British Columbia: Fisheries and synopsis of life history information. *Canadian Special Publication of Fisheries and Aquatic Sciences*, *103*, 1–46.

Stockwell, D. R. B., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, *148*(1), 1–13. https://doi.org/10.1016/S0304-3800(01)00388-X

Swift, T. L., & Hannon, S. J. (2010). Critical thresholds associated with habitat loss: A review of the concepts, evidence, and applications. *Biological Reviews*, *85*(1), 35–53. https://doi.org/10.1111/j.1469-185X.2009.00093.x

Thomas, J. A., Bourn, N. A. D., Clarke, R. T., Stewart, K. E., Simcox, D. J., Pearman, G. S., Curtis, R., & Goodger, B. (2001). The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1478), 1791–1796. https://doi.org/10.1098/rspb.2001.1693

Williams, B. K. (2011). Adaptive management of natural resources—Framework and issues. *Journal of Environmental Management*, *92*(5), 1346–1353. https://doi.org/10.1016/j.jenvman.2010.10.041

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Camaclang, A. E., Chadès, I., Martin, T. G., & Possingham, H. P. (2022). Predicting the optimal amount of time to spend learning before designating protected habitat for threatened species. *Methods in Ecology and Evolution*, 13, 722–733. https://doi.org/10.1111/2041-210X.13770