Computing Robust Strategies for Managing Invasive Plants

Andreas Lydakis[1], **Jenica M. Allen**[2], **Marek Petrik**[1], **Tim Szewczyk**[2]
[1] Department of Computer Science
[2] Department of Natural Resources and the Environment
University of New Hampshire, Durham, NH

Abstract

The significant threat that invasive species pose to native ecosystems can be reduced by targeted management actions. Land managers are risk-averse and are reluctant to rely on observational data which can be flawed and limited, but this can lead to bad decisions. In this paper, we propose and evaluate methods for data-driven risk-averse decision support systems that can work with flawed data. We use distributionally robust optimization combined with the Wasserstein metric to compute management strategies that not only work well on average but are also robust to data problems. We evaluate our methods using the case of glossy buckthorn, an invasive shrub in the Northeast United States.

1 Introduction

Invasive species constitute a major threat to native organisms and ecosystems [8, 30]. In the United States alone, there are over 50,000 invasive species that are responsible for more than $100 billion in environmental damage annually [25]. Although they cannot be entirely eliminated, the impact of invasive species on ecosystems may be reduced by effective land management strategies. Developing cheap and effective strategies can be challenging because natural systems are complex, difficult to model, and expensive to observe. Even the best available data sets on the distribution of invasive species are small, biased, and sometimes incorrect [16].

Common land-management actions that eliminate invasive plants include cutting the invasive plant, burning, flooding, or planting native vegetation [18, 19, 28]. Committing to such an action is expensive and can have long-term consequences. Land managers, who do not like taking risks, are often uncertain about effects of mitigation actions and wait too long before addressing an invasion [5, 10]. This ultimately inflates the mitigation costs. To make more effective and timely decisions, land managers need decision support systems that recommend safe actions in the face of limited and flawed data [4].

In this paper, we propose and evaluate new *robust* methods for computing invasive management strategies from imperfect observational data. Such methods are a crucial component of developing practical decision support systems. They can compute a management strategy that is likely to work well even if observations of invasive plant presences are sparse and deviate from the true distribution. To address data-induced uncertainties in a principled and tractable way, we use *distributionally robust* optimization [6, 9, 11, 31].

We model land management as a resource allocation problem. The land manager must decide on how to distribute the available mitigation resources, such as pesticide application, over their land. More extensive invasions require more resources, but the precise distribution of the target plant is rarely available. Instead, the resource allocation must be based on sparse data of actual plant observations. A standard approach is to estimate the distribution of the invasive plant, using the machine learning method *MaxEnt* [21, 24, 26] for example, and then compute the optimal resource allocation with respect to this distribution. This approach can fail because it ignores areas which have a low expected concentration of the invasive plant, but may have a high level of uncertainty.

A robust method must trade off between allocating resources to the locations in which the invasive species is most likely, with locations in which presence is uncertain. *Distributionally robust optimization* negotiates this trade-off by modeling uncertainties as a zero-sum game against nature. The adversarial nature chooses a worst *plausible* realization of the uncertainties. To achieve a solution quality that is both robust and works well on average, the distributions available to nature, also known as the *ambiguity set*, must be chosen carefully. We propose and evaluate several possible choices.

As one of the contributions, we also propose and evaluate a new approach to selecting ambiguity sets in distributionally robust optimization. We use the result of MaxEnt to choose the *shape* of the ambiguity set. Our empirical results show that this approach can be superior to standard techniques [9].

The remainder of the paper is organized as follows. In Section 2, we describe and motivate the general formulation of the management problem. Then in Section 3, we summarize methods for adding robustness to this optimization problem, considering both the simple ℓ_1-norm and the popular Wasserstein distance metric. Section 4 describes our experimental

results on glossy buckthorn data and Section 5 outlines future work.

2 Problem Formulation

In this section, we formally describe the resource allocation problem. We assume that the goal of a land manager is to optimally distribute available resources across a given area in order to contain or remove the undesirable plant. These resources can represent manual removal by volunteers, pesticide application, or controlled burning of an area, among others. Since our goal is to study effects and the mitigation of uncertainty, we abstract from non-essential management constraints.

The area of interest is discretized into a discrete number of regions, or cells, which are represented by a set $\mathscr{L}$. Each *region* can span an area from a few hundred square meters up to several square kilometers. The observation data can be used to estimate—using MaxEnt or a similar method—a distribution $p \in [0,1]^{\mathscr{L}}$ of the invasive plant across the discrete regions $\mathscr{L}$. We assume that p is a probability distribution and thus sums to one. That means that p_l for any region $l \in \mathscr{L}$ represents the relative prevalence of the species and would have to be multiplied by the total number of individuals in the area of interest to get the actual prevalence. In terms of management resources, we assume that there is a total available budget of C, such as the number of volunteers available for manual pulling, or the available pesticide application budget.

The aim is to allocate c_l to each region $l \in \mathscr{L}$ to maximize some pre-defined management goals. We assume that the benefit of allocating resources to a region grows linearly with 1) the amount of resources c_l and 2) the prevalence $\hat{p}_l$ of the plant. This essentially models the case where the benefit is proportional to the amount of the plant removed and it is possible to apply resources to a given region without a significant increase or decrease in efficiency. While this assumption is rarely satisfied, the dependence on resource allocation and prevalence can be relaxed to be piecewise linear concave. The objective of computing the best allocation can be modeled as the following mathematical optimization problem:

$$\max_{\boldsymbol{c} \in \mathbb{R}_+^{\mathscr{L}}} \sum_{l \in \mathscr{L}} \hat{p}_l \cdot c_l = \hat{\boldsymbol{p}}^\mathsf{T} \boldsymbol{c} \quad \text{s.t.} \quad \sum_{l \in \mathscr{L}} c_l = \mathbf{1}^\mathsf{T} \boldsymbol{c} \leq C. \quad (1)$$

Here, the bold letters denote vectors, $\mathbf{1}$ represents a vector of all ones of a size appropriate to its context, and $\mathbb{R}_+$ is the set of *non-negative* real numbers. This formulation can be easily augmented by additional constraints, such as constraining the lower and upper bounds on the resource allocation $c_{\min} \leq c \leq c_{\max}$ for some $c_{\max}$ and $c_{\min}$. The objective can also be generalized to $\hat{\boldsymbol{p}}^\mathsf{T} Q \boldsymbol{c}$ for some matrix Q that defines proportional benefits of allocation to different regions. Note that in the absence of additional constraints in (1), we can assume that $C = 1$ without loss of generality.

The species distribution $\hat{\boldsymbol{p}}$ is typically estimated from observational data. As mentioned above, the estimation must often be performed from presence-only data, in which only reported observations of the species are available without information of where the species does not occur. Numerous methods have been proposed for estimating species distributions; see for example [3] for an overview. One of the most popular methods is MaxEnt [20]. MaxEnt uses bio-relevant geographic features, such as temperature, rainfall, and land cover to estimate a species distribution that is both consistent with the reported observations and has maximal possible entropy, or in other words is as close to the uniform distribution as possible.

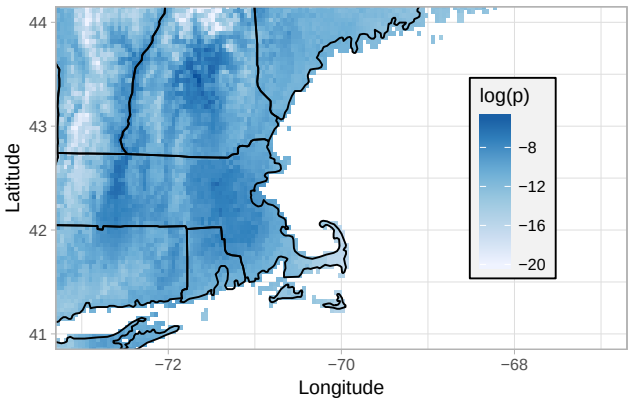

Figure 1: Estimated distribution of glossy buckthorn in the New England region of United States. The color represents the logarithm of the distribution probability p estimated using MaxEnt.

Fig. 1 depicts the estimated distribution of glossy buckthorn, an invasive species in North America. The environmental variables are based on biological variables from WorldClim [14] with a 30s spatial resolution. The distribution is estimated from historical glossy buckthorn presence observations available from EDDMaps [1].

As mentioned above, management decisions must consider the uncertainty from limited presence observations. Observations about plant species distributions rarely come from organized and systematic surveys. Rather, they are comprised primarily of incidental observations from federal and state agencies or third-party observers. As a result, it is not the individual observations that are uncertain, but rather the uncertainty is driven by 1) the small number of total observations, 2) a lack of observations of plant absences, and 3) a sample that is biased to the distribution of observers [12, 27, 29].

Using the optimization formulation in (1) leads to the optimal allocation resources, as long as the estimate $\hat{p}$ is known. Unfortunately, as mentioned before, the distribution $\hat{p}$ is rarely known with any kind of precision. When the estimate is wrong, the allocation may be significantly suboptimal. In the remainder of the paper, we describe and evaluate methods that use distributionally robust optimization to minimize risk of an incorrect allocation when the the distribution $\hat{p}$ is known only approximately.

3 Robust Resource Allocation

In this section, we leverage *distributionally robust optimization* to compute resource allocation that is more robust with

respect to an imprecise estimate of the species distribution p. Distributionally robust optimization is a methodology for formulating and solving stochastic optimization problems in which the probabilities are known only approximately. The main idea is to compute a resource allocation that is good for a range of species distributions that are close to its estimate.

The distributionally robust formulation of (1) is:

$$\max_{\boldsymbol{c}\in\mathbb{R}_+^{\mathscr{L}}} \min_{p\in\mathscr{P}} \boldsymbol{p}^\mathsf{T}\boldsymbol{c} \quad \text{s.t.} \quad \mathbf{1}^\mathsf{T}\boldsymbol{c} \le C. \tag{2}$$

The set $\mathscr{P}$ represents plausible species distributions and is often referred to as an *ambiguity set*. It is called distributionally robust optimization because the inner minimization is over a set of distributions and could be alternatively written as:

$$\min_{\boldsymbol{p}\in\mathscr{P}} \mathbb{E}_{\boldsymbol{p}}[c], \tag{3}$$

where c is the random variable of resource allocations to each cell. Note that a probability distribution over $\mathscr{L}$ can be also interpreted as a vector defined over $\mathscr{L}$.

The control resources need to be allocated in a way that satisfies two conflicting objectives. Resources should be applied in a way that maximizes the coverage over the infestation, but at the same time accounts for the uncertainty of the data.

One of the main challenges in distributionally robust optimization are how to design the ambiguity set $\mathscr{P}$ such that the solution is not overly conservative and the optimization problem is still easy to solve. We consider ambiguity sets that are defined based on a distance from the nominal distribution $\hat{p}$ according to some distance metric D:

$$\mathscr{P} = \left\{\boldsymbol{p}\in\Delta^{\mathscr{L}} \;:\; D(\boldsymbol{p}||\hat{\boldsymbol{p}}) \le \psi\right\}. \tag{4}$$

Here, the symbol $\Delta^{\mathscr{L}}$ denotes the set of all valid probability distributions over $\mathscr{L}$. The value ψ above is known as an *ambiguity budget* and determines the robustness of the resource allocation. A value of ψ that is close to 0 means that the allocation needs to be good only with respect to the nominal probability distribution, resulting in an allocation that will not be robust to deviations from $\hat{p}$. A large value of ψ means that the allocation is expected to be robust with respect to species distributions that can be very different from $\hat{p}$. The metric D measures the distance between two probability distributions; one common example is the KL divergence.

Note that our focus is on robustness and risk aversion with respect to model error. This is different from the traditional approach to risk aversion, in which risk is taken with respect to a probabilistic outcome. In a traditional risk model, one would consider replacing $\mathbb{E}[c]$ in (3) by a risk measure.

For the distributionally robust formulation to be useful, one needs to decide on the shape and size of the set $\mathscr{P}$. The shape of the set is determined by the metric D, and the size by the budget. We consider two distance metrics: the L_1 norm, also known as the Manhattan distance, and the Wasserstein distance, also known earthmover's distance. We focus on these two measures because then (2) can be formulated as a linear program and solved for relatively large problems. To solve (2) as a linear program, it is sufficient to dualize the inner minimization problem. KL-divergence is another common measure, but we do not consider it because it yields difficult optimization problems and can become intractable when $|\mathscr{L}|$ is large.

3.1 L_1 Distance

The L_1 distance metric D_1 between the two distributions is defined as $D_1(\boldsymbol{p}||\boldsymbol{q}) = \sum_{l\in\mathscr{L}}|p_l - q_l|$. This is a common metric used in robust Markov decision processes and reinforcement learning; see for example [2, 7, 15, 22, 23] and references therein. We consider this distance metric because of its simplicity and lack of parameters, and use it as a baseline measure.

3.2 Wasserstein Distance

The Wasserstein distance for discrete probability distributions $\boldsymbol{p}$ and $\boldsymbol{q}$ over regions $\mathscr{L}$ is defined as follows:

$$D_W(\boldsymbol{p}||\boldsymbol{q}) =$$

$$= \min_{T\in\mathbb{R}_+^{\mathscr{L}\times\mathscr{L}}} \left\{ \sum_{l_1,l_2\in\mathscr{L}} \mathrm{dst}(l_1,l_2)\,T(l_1,l_2) \;:\; \begin{matrix} T\mathbf{1} = \boldsymbol{p} \\ \mathbf{1}^\mathsf{T}T = \boldsymbol{q}^\mathsf{T} \end{matrix} \right\} \tag{5}$$

The Wasserstein distance is also known as earthmover's distance because it can be seen as representing the minimal amount of work that it takes to move probabilities in order to transform the distribution $\boldsymbol{p}$ to the distribution $\boldsymbol{q}$. The matrix T is known as a transportation plan and determines the cheapest transport. Specifically, each $T(l_1, l_2)$ represents how much probability mass should be transported from region l_1 to region l_2. The weight $\mathrm{dst}(l_1, l_2)$ denotes the *distance*, or cost, by which the mass should be transported and its choice makes it possible to highly customize it. The Wasserstein metric has gained popularity in the robust optimization community in recent years because it is very flexible, can be estimated from data, and is relatively easy to optimize [9].

When using the Wasserstein distance, the solution quality depends on the choice of the distance metric dst. The distance metric is most often the norm of the difference between features. In our setting, the features $\boldsymbol{\phi_l}$ for each region l are the biological indicators derived from WorldClim, leading to the following distance metric:

$$\mathrm{dst}_{L1}(l_1, l_2) = \|\boldsymbol{\phi_{l_1}} - \boldsymbol{\phi_{l_2}}\|_1$$

However, not all features are equally relevant in quantifying the biologically relevant distance. We propose to use MaxEnt to identify the relevant features as follows. The MaxEnt distribution $\tilde{\boldsymbol{p}}$ satisfies that:

$$\log \tilde{p}_l = \boldsymbol{\phi_l}^\mathsf{T}\beta, \tag{6}$$

for some coefficients β. Using the values β we propose to weight the features by the corresponding coefficients:

$$\mathrm{dst}_{ME}(l_1, l_2) = \|\boldsymbol{\phi_{l_1}} - \boldsymbol{\phi_{l_2}}\|_{1,\beta}\,, \tag{7}$$

where $\|\cdot\|_{1,\beta}$ is a β-weighted L_1 norm. We also consider reciprocally weighted features:

$$\mathrm{dst}_{IME}(l_1, l_2) = \|\boldsymbol{\phi_{l_1}} - \boldsymbol{\phi_{l_2}}\|_{1,1/|\beta|}\,. \tag{8}$$

These formulations attempt to bias the ambiguity set to account for important and unimportant features.

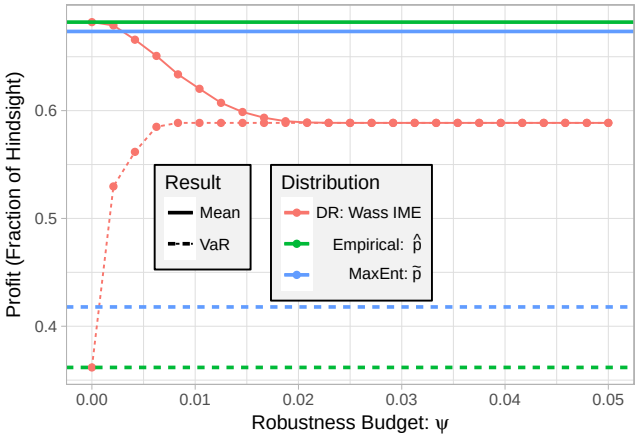

Figure 2: Response to uncertainty budget allocated.

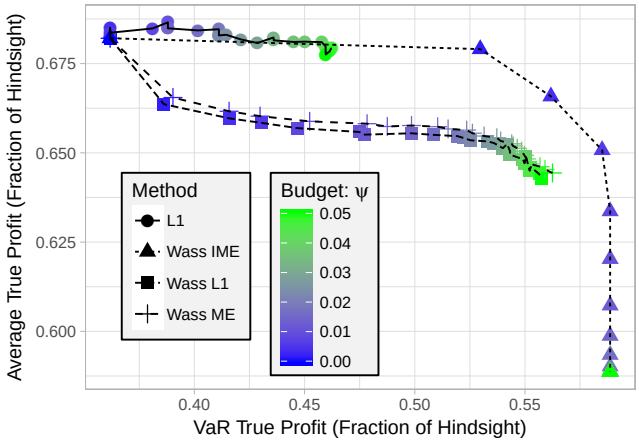

Figure 3: Tradeoff between average-case and worst-case performance.

4 Evaluation: Glossy Buckthorn

To experimentally evaluate our framework, we consider optimizing management allocations to limit the spread of glossy buckthorn (*Frangula alnus*) in New England. Glossy buckthorn is a woody shrub that was introduced from Europe in the 1800s and has invaded North American forests throughout New England [17].

We aim to evaluate the quality of the management strategy in the context of the ground truth. Unfortunately, no large-scale ground truth data is available for invasive plant species of the Northeast. We instead use observational data from EDDMapS [1], biological features from WorldClim [13], and MaxEnt to construct a hypothetical ground-truth distribution $\boldsymbol{p}^\star$. We restrict our observations to 100 randomly chosen locations to generate a set of artificial *presence-only* observations according to $\boldsymbol{p}^\star$. Using these artificial observations, we denote the empirical distribution as $\hat{\boldsymbol{p}}$.

Solving the optimization problem (2) returns an allocation of resources over the individual regions. Let $\hat{\boldsymbol{c}}$ be such an allocation for an estimated distribution $\hat{\boldsymbol{p}}$. To evaluate the quality of the allocation, we compare the efficiency of the resources with respect to the ground truth species distribution $\boldsymbol{p}^\star$, and compute the *profit* $\hat{\boldsymbol{c}}^\mathsf{T}\boldsymbol{p}^\star$.

The goal of the robust solution is to compute an allocation that works not only when data sets are good, but also in cases where the observed data have problems. To evaluate the robustness of the method, we generate n observation sets, estimate distributions $\hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2, \hat{\boldsymbol{p}}_3, \ldots, \hat{\boldsymbol{p}}_n$, and compute the corresponding $\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, \ldots, \boldsymbol{c}_n$. In order to be robust, $\boldsymbol{c}_i^\mathsf{T}\boldsymbol{p}^\star$ must be good not only on average, but also for the majority of the data sets. The average performance over the data sets is computed as $1/n \sum_{i=1}^{n} \boldsymbol{c}_i^\mathsf{T}\boldsymbol{p}^\star$. The worst-case is evaluated using value-at-risk (VaR) at 95% level, representing the wost 5% of datasets.

Fig. 2 shows how the true profit varies with the robustness budget. The profit is represented as a fraction of what is possible when the true distribution $\tilde{p}$ is known. Only the distributionally robust solution using dst_{IME} depends on the budget; the solutions for the empirical and MaxEnt distributions ($\hat{\boldsymbol{p}}$ and $\tilde{\boldsymbol{p}}$) are independent of the budget size. The empirical solution actually works best on average. In this solution, the resources are allocated only to areas in which the plant has been observed. However, when the observations are not aligned well with the true distribution, it can achieve a very low profit. Using Maximum Entropy improves the VaR, but the distributionally robust solution allows to smoothly trade off the average profit with robustness.

Fig. 3 compares how different robust methods trade off between the average-case solution quality and the VaR. Because the budget values have different meanings for different methods, it does not make sense to compare their values against the same budget. The results show that using dst_{IME} leads to best performance. The other Wasserstein metrics surprisingly under-perform even the simple robust optimization with $D(\boldsymbol{p}||\hat{\boldsymbol{p}}) = |\boldsymbol{p} - \hat{\boldsymbol{p}}||$.

5 Conclusion and Future Work

We demonstrate that distributionally robust optimization is a viable method for improving robustness of resource allocation when managing invasive species. The proposed methods significantly increase the robustness of solutions to imperfect data while only modestly degrading the average performance. In future work, we plan to study a theoretical grounding for using the IME metric, and we will explore more realistic species distribution models based on an integral projection model (IPM) framework.

Acknowledgments

This work was supported, in part, by the National Science Foundation under Grant No. IIS-1717368. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] EDDMapS. 2016. Early Detection & Distribution Mapping System. The University of Georgia - Center for Invasive Species and Ecosystem Health, 2016.

[2] P Auer, Thomas Jaksch, and R Ortner. Near-optimal regret bounds for reinforcement learning. *Journal of Machine Learning Research*, 11(1):1563–1600, 2010.

[3] Bethany A. Bradley. Predicting abundance with presence-only models. *Landscape Ecology*, 31(1):19–30, 2016.

[4] L R Carrasco, R Baker, a Macleod, J D Knight, and J D Mumford. Optimal and robust control of invasive alien species spreading in homogeneous landscapes. *Journal of the Royal Society, Interface / the Royal Society*, 7(September):529–540, 2010.

[5] Alisha D. Davidson, Chad L. Hewitt, and Donna R. Kashian. Understanding acceptable level of risk: Incorporating the economic cost of under-managing invasive species. *PLoS ONE*, 10(11):1–12, 2015.

[6] Eric Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data driven problems. *Operations Research*, 58(3):595–612, 2010.

[7] TG Dietterich, MA Taleghan, and Mark Crowley. PAC optimal planning for invasive species management: Improved exploration for reinforcement learning from simulator-defined MDPs. *National Conference on Artificial Intelligence (AAAI)*, 2013.

[8] Joan G Ehrenfeld and Diego Rodriguez. Ecosystem Consequences of Biological Invasions. *Annu. Rev. Ecol. Evol. Syst*, 41:59–80, 2010.

[9] Peyman Mohajerin Esfahani and Daniel Kuhn. *Data-driven distributionally robust optimization using the Wasserstein metric : performance guarantees and tractable reformulations*. Springer Berlin Heidelberg, 2017.

[10] David Finnoff, Jason F. Shogren, Brian Leung, and David Lodge. Take a risk: Preferring prevention over control of biological invaders. *Ecological Economics*, 62(2):216–222, 2007.

[11] Joel Goh and Melvyn Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4):902–917, 2010.

[12] Andrew M Gormley, David M Forsyth, Peter Griffioen, Michael Lindeman, David SL Ramsey, Michael P Scroggie, and Luke Woodford. Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology*, 48(1):25–34, 2011.

[13] Robert J. Hijmanns and Catherine H. Graham. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, 12(12):2272–2281, dec 2006.

[14] Robert J. Hijmans, Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978, 2005.

[15] Garud N. Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2):257–280, may 2005.

[16] Koji Kotani, Makoto Kakinaka, and Hiroyuki Matsuda. Optimal invasive species management under multiple uncertainties. *Mathematical Biosciences*, 233(1):32–46, 2011.

[17] Thomas D. Lee and Jennifer H. Thompson. Effects of logging history on invasion of eastern white pine forests by exotic glossy buckthorn (Frangula alnus P. Mill.). *Forest Ecology and Management*, 265:201–210, 2012.

[18] Cory Merow, Jenica M. Allen, Matthew E. Aiello-Lammens, and John. A. Silander Jr. Improving niche and range estimates with Maxent and point process models by integrating spatially explicit information. *Global Ecology and Biogeography*, 25:1022–1036, 2016.

[19] Cory Merow, Nancy LaFleur, John A. Silander Jr., Adam M. Wilson, and Margaret Rubega. Developing Dynamic Mechanistic Species Distribution Models: Predicting Bird-Mediated Spread of Invasive Plants across Northeastern North America. *The American Naturalist*, 178(1):30–43, jul 2011.

[20] Cory Merow, Matthew J. Smith, and John A. Silander. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10):1058–1069, 2013.

[21] John Mount. The equivalence of logistic regression and maximum entropy models. *URL: http://www. win-vector. com/dfiles/LogisticRegressionMaxEnt. pdf*, 2011.

[22] Marek Petrik, Mohammad Ghavamzadeh, and Yinlam Chow. Safe Policy Improvement by Minimizing Robust Baseline Regret. In *Advances in Neural Information Processing Systems (NIPS)*, 2016.

[23] Marek Petrik and Dharmashankar Subramanian. RAAM : The benefits of robustness in approximating aggregated MDPs in reinforcement learning. In *Neural Information Processing Systems (NIPS)*, 2014.

[24] SJ Phillips, M Dudík, and RE Schapire. Maxent software for modeling species niches and distributions version 3.4. 0, 2017.

[25] David Pimentel, Rodolfo Zuniga, and Doug Morrison. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52(3):273–288, feb 2005.

[26] Ian W Renner and David I Warton. Equivalence of maxent and poisson point process models for species distribution modeling in ecology. *Biometrics*, 69(1):274–281, 2013.

[27] J Andrew Royle, Richard B Chandler, Charles Yackulic, and James D Nichols. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution*, 3(3):545–554, 2012.

[28] Majid Alkaee Taleghan, Thomas G. Dietterich, Mark Crowley, Kim Hall, and H. Jo Albers. PAC Optimal MDP Planning with Application to Invasive Species Management. *Journal of Machine Learning Research*, 16:3877–3903, 2015.

[29] Gill Ward, Trevor Hastie, Simon Barry, Jane Elith, and John R Leathwick. Presence-only data and the em algorithm. *Biometrics*, 65(2):554–563, 2009.

[30] David S. Wilcove, David Rothstein, Jason Dubow, Ali Phillips, and Elizabeth Losos. Quantifying threats to imperiled species in the United States. *BioScience*, 48(8):607–615, 1998.

[31] Insoon Yang. A Convex Optimization Approach to Distributionally Robust Markov Decision Processes With Wasserstein Distance. *IEEE Control Systems Letters*, 1(1):164–169, 2017.