Robust Pest Management Using Reinforcement Learning

Talha Siddique
Department of Natural Resources and the Environment
University of New Hampshire
Durham, NH 03824
ts1121@wildcats.unh.edu

Jia Lin Hau
Department of Computer Science
University of New Hampshire
Durham, NH 03824
jh1111@wildcats.unh.edu

Shadi Atallah
Department of Natural Resources and the Environment
University of New Hampshire
Durham, NH 03824
shadi.atallah@unh.edu

Marek Petrik
Department of Computer Science
University of New Hampshire
Durham, NH 03824
mpetrik@cs.unh.edu

Abstract

Developing effective decision support systems for agriculture matters. Human population is likely to peak at close 11 billion and changing climate is already reducing yields in the Great Plains and in other fertile regions across the world. With virtually all arable land already cultivated, the only way to feed the growing human population is to increase yields. Making better decisions, driven by data, can increase the yield and quality of agricultural products and reduce their environmental impact.

In this work, we address the problem of an apple orchardist who must decide how to control the population of codling moth, which is an important apple pest. The orchadist must decide when to apply pesticides to optimally trade off apple yields and quality with the financial and environmental costs of using pesticides. Pesticide spraying decisions are made weekly throughout the growing season, with the yield only observed at the end of the growing season. The inherent stochasticity driven by weather and delayed rewards make this a classical reinforcement learning problem.

Deploying decision support systems in agriculture is challenging. Farmers are averse to risk and do not trust purely data-driven recommendations. Because weather varies from season to season and ecological systems are complex even a decade worth of data may be insufficient to get good decisions with high confidence.

We propose a robust reinforcement learning approach that can compute good solutions even when the models or rewards are not known precisely. We use Bayesian models to capture prior knowledge. Our main contribution is that we evaluate which model and reward uncertainties have the greatest impact on solution quality.

Keywords: Natural Resources, Small Data, Markov Decision Process, Robust Optimization.

1 Introduction

Apple is the most consumed fruit in the USA. In 2015, an average person in the USA consumed about 115.4 pounds of fresh and processed fruits and 24.7 pounds of apple in other forms such as juice, canned, frozen etc (USDA ERS, 2015). One of the most problematic apple pests is the codling moth (Cydia Pomonella). If left unchecked, it can claim up to 95 percent of the seasons apple crops [8]. Common pest management actions are mostly dependent on chemical pesticides. Avoiding the use of a pesticide improves the quality of an apple crop and increases profits but waiting too long after the pest is detected can lead to a crop failure [4, 10, 11].

There is a need for effective decision support tools. However, developing cheap and effective strategies can be challenging because natural systems are complex, difficult to model, and expensive to observe [6]. Recently, reinforcement learning has been used with great success to model such complex domains [7]. The drawback of using such techniques is that they require data sets of appropriate size and accuracy. Even though thousands of samples can be generated using domain simulators but any form of available data sets on the distribution of invasive species like pests tend to suffer from biases and inaccuracy [5]. To make more effective and timely decisions, there is a need for decision support systems that recommend safe actions in the face of limited and flawed data [1].

In this paper, we address the problem of an orchard manager who must decide, given a pest population level, whether to apply pesticides, in order to maximize the value of the orchard over a finite time horizon. We can formulate the management problem using the reinforcement learning methodology. The pesticide application policies must be based on data consisting of imperfect observations of the pest population level. For our solutions to be immune to such data and parameter uncertainties, there is a need for our method to be robust. A robust method will compute solutions that trade-off brittle optimality for increased confidence.

We develop and evaluate a new method for robust reinforcement learning, which can reliably compute pest management policies from imperfect observational data. They can determine pest control policies that are likely to work well even if observations of pest population levels are scattered and does not resemble the true distribution accurately. To address such uncertainties from data and parameters in a tractable way, we combine the flexibility of Bayesian modeling with the computational tractability of robust optimization. We will be using the Bayesian modeling tool Stan to implement our approach.

Unfortunately, planning directly with posterior distributions leads to intractable optimization problems [2, 3, 9]. We are, instead, proposing to use the posterior distributions to construct the ambiguity set and then use tractable robust optimization methods.

The contributions of this paper is to advance the understanding of uncertainty in reinforcement learning. Our research provides new insights into the benefits and drawbacks of considering uncertainty given a specific problem structure.

2 Experiments and Results

Given our model, we compare the performance of different pesticide application schedules using threshold policies. A threshold policy applies the pesticide only when the pest level exceeds the given threshold.

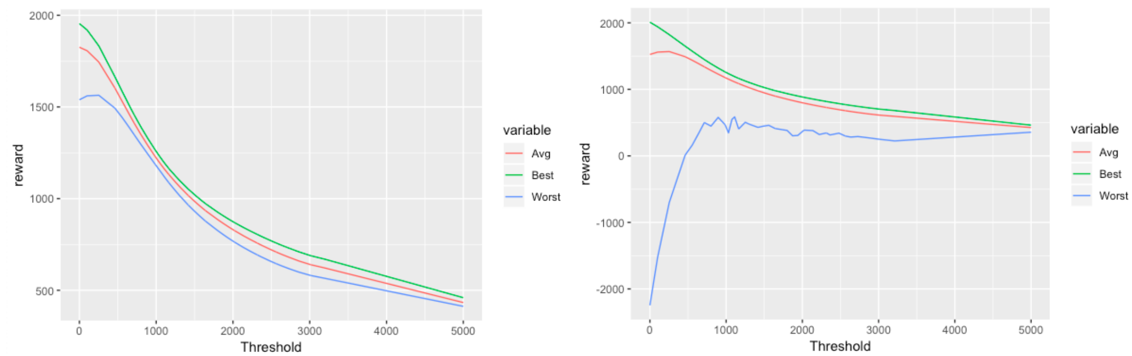
For each threshold policy π, we determine the best case, worst case, and average case scenario returns, when we apply it to the set of different pest growth rates: λ. The pest growth rates λ are sampled from our Bayesian model. In addition, we will also determine the robust return given our π and λ.

Let $\rho(\pi, \lambda)$ be the return of policy π under the growth rate λ. Following are the mathematical definitions of each of our objectives:

1. Best case scenario return: $\max_{\lambda,\epsilon,r} \rho(\pi, \lambda)$
2. Worst case scenario return: $\min_{\lambda,\epsilon,r} \rho(\pi, \lambda)$
3. Average case scenario return: $\mathbb{E}_{\lambda,\epsilon,r} \rho(\pi, \lambda)$
4. Robust return: $\max_{\pi} \min_{\lambda,\epsilon,r} \rho(\pi, \lambda)$

According to Figure 1(a), we plot the graph where we only consider the uncertainty of the growth rate λ. Since the best, worst, and average cases have similar optimal threshold policies, the robustness does not really make a difference compared to the regular approach. All objectives suggest the farmer to always spray the plant because the lower threshold has a higher reward. The intuition of the policy is, if a farmer always sprays pesticide to their apples, the number of moths will be the least and they will likely to get a good apple and a high reward (revenue).

However, this is not realistic, spraying pesticides frequently could damage the soil or your plant which is the cost that our model does not capture. The second issue of the simulation and our model did not capture, is when an apple turns bad, some of the damage of an apple is not reversible, we could kill the moth in the apple but the apple will still remain damage. Therefore, to make the reward function more realistic, we included 2 variables (i) cost of spraying and



(a) Uncertainty of Growth Rate λ (b) Added Uncertainty of Cost and Penalty

Figure 1: Returns as a Function of Application Threshold

(ii) irreversible financial penalty for infected apples. Since we do not know the real underlying cost for spraying and irreversible penalty we created a distribution of possible cost and explore the effect of the cost.

In Figure 1(b), we depict the costs of spraying and an irreversible penalty for a bad apple to make the model more realistic. After we added the cost of spraying and penalty for irreversible apple damage, we can see that the robust policy which maximizes the worst case and the regular approach which maximize the average case are totally different to each other.

3 Conclusion

To conclude, robust optimization is not always necessary if we have uncertainty on a linear function. However, robust optimization is crucial if we have uncertainty in other types of function. From our experiment, we show the importance to identify and define the uncertainties of all important factors in a model, which otherwise could lead to a very different solution. An example is the solutions determined by our model where after defining the cost of spraying and penalty for infected apples, we can see an obvious change in policy.

References

[1] 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.

[2] Michael Castronovo, Damien Ernst, Adrien Couëtoux, and Raphael Fonteneau. Benchmarking for Bayesian reinforcement learning. *PLoS ONE*, 11(6):1–25, 2016.

[3] Arthur Guez, David Silver, and Peter Dayan. Efficient Bayes-Adaptive Reinforcement Learning using Sample-Based Search. In *Advances in Neural Information Processing Systems (NIPS)*, 2012.

[4] Darwin C. Hall and Richard B. Norgaard. On the Timing and Application of Pesticides: Reply. *American Journal of Agricultural Economics*, 56(3):644, 2006.

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

[6] Andreas Lydakis, Jenica M. Allen, Marek Petrik, and Tim M. Szewczyk. Robust strategies for managing invasive plants. In *IJCAI Workshop: Artificial Intelligence for Wildlife Conservation*, 2018.

[7] Matthew H. Meisner, Jay A. Rosenheim, and Ilias Tagkopoulos. A data-driven, machine learning framework for optimal pest management in cotton. *Ecosphere*, 7(March):1–13, 2016.

[8] M.Gregory Peck and Ian A. Merwin. A Grower ' s Guide to Organic Apples. *Agriculture*, 208(NYS IPM Publication No 223):43–56, 2016.

[9] Pascal Poupart, Nikos Vlassis, Jesse Hoey, and Kevin Regan. An Analytic Solution to Discrete Bayesian Reinforcement Learning. In *International Conference of Machine Learning (ICML)*, 2006.

[10] W. D. Shaw. Environmental and Natural Resource Economics: Decisions Under Risk and Uncertainty. *Encyclopedia of Energy, Natural Resource, and Environmental Economics*, 3-3(979):10–16, 2013.

[11] Richard T. Woodward and David Tomberlin. Practical Precautionary Resource Management Using Robust Optimization. *Environmental Management*, 54(4):828–839, 2014.