Stay Ahead of Poachers: Illegal Wildlife Poaching Prediction and Patrol Planning Under Uncertainty with Field Test Evaluations (Short Version)

Lily Xu*[1], Shahrzad Gholami*[2], Sara Mc Carthy[2], Bistra Dilkina[2], Andrew Plumptre[3], Milind Tambe[1], Rohit Singh[4], Mustapha Nsubuga[5], Joshua Mabonga[5], Margaret Driciru[6], Fred Wanyama[6], Aggrey Rwetsiba[6], Tom Okello[6], Eric Enyel[6]
[1]*Harvard University,* [2]*University of Southern California,* [3]*Key Biodiversity Area Secretariat,*
[4]*World Wide Fund for Nature,* [5]*Wildlife Conservation Society,* [6]*Uganda Wildlife Authority*
`lily_xu@g.harvard.edu, sgholami@usc.edu`

***Abstract*—Illegal wildlife poaching threatens ecosystems and drives endangered species toward extinction. However, efforts for wildlife protection are constrained by the limited resources of law enforcement agencies. To help combat poaching, the Protection Assistant for Wildlife Security (PAWS) is a machine learning pipeline that has been developed as a data-driven approach to identify areas at high risk of poaching throughout protected areas and compute optimal patrol routes. In this paper, we take an end-to-end approach to the data-to-deployment pipeline for anti-poaching. In doing so, we address challenges including extreme class imbalance (up to 1:200), bias, and uncertainty in wildlife poaching data to enhance PAWS, and we apply our methodology to three national parks with diverse characteristics. (i) We use Gaussian processes to quantify predictive uncertainty, which we exploit to improve robustness of our prescribed patrols and increase detection of snares by an average of 30%. We evaluate our approach on real-world historical poaching data from Murchison Falls and Queen Elizabeth National Parks in Uganda and, for the first time, Srepok Wildlife Sanctuary in Cambodia. (ii) We present the results of large-scale field tests conducted in Murchison Falls and Srepok Wildlife Sanctuary which confirm that the predictive power of PAWS extends promisingly to multiple parks. This paper is part of an effort to expand PAWS to 800 parks around the world through integration with SMART conservation software.**

***Index Terms*—wildlife protection, data mining, predictive modeling, patrol route planning, poaching, uncertainty, gaussian processes**

I. Introduction

Illegal wildlife poaching is an international problem that threatens biodiversity, ecological balance, and ecotourism [1]. Timely detection and deterrence of illegal poaching activities in protected areas are critical to combating poaching. Artificial intelligence frameworks can advance wildlife protection efforts by learning from past poaching activity to prescribe actionable recommendations to park managers.

Assessing poaching risk through a protected area and prescribing patrol plans to rangers requires in-depth knowledge of the poachers' behavior. Learning the poachers' behavior is a challenging machine learning problem since (a) the wildlife crime datasets are typically extremely imbalanced, with up to 99.6% negative labels; (b) negative labels indicating absence of illegal activity are not reliable due to the inherent difficulty of detecting well-hidden poaching signs; (c) historical poaching observations are not collected thoroughly and uniformly, resulting in bias; and (d) poaching patterns and landscapes vary between protected areas, preventing a universal model.

Fig. 1: Rangers in Srepok Wildlife Sanctuary with snares they removed during our field tests. Photo: WWF Cambodia.

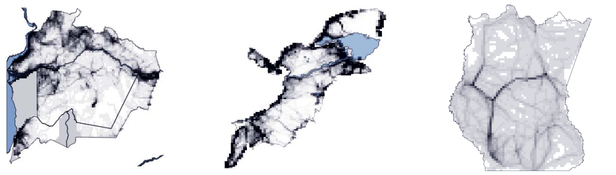
In this paper, we present an integrated approach to the data-to-deployment pipeline†. The entire pipeline has been designed with deployment in mind, where the ultimate goal is to maximize the number of snares removed—corresponding to animal lives saved. Hence, rangers want to patrol areas with highest expected returns; this domain insight translates to risk-averse use of our predictive models, taking into account uncertainty. To do so, (i) We use Gaussian processes to quantify uncertainty in predictions of poaching risk and exploit the uncertainty in our patrol route optimization to increase the robustness to prediction errors. We evaluate our approach on historical poaching data using three real-world datasets from Uganda and Cambodia, which have differ in both ecological and data quality characteristics. (ii) We present results from large-scale field tests conducted in two parks.

The historical patrol data used in this paper are managed with SMART, a database system used in over 800 protected areas across 55 countries [2]. Although SMART records significant amounts of historical data, its current capabilities

*Equal contribution.
The full version of this paper is available at https://arxiv.org/abs/1903.06669

†Our code is available online at https://github.com/lily-x/paws-public

2375-026X/20/$31.00 ©2020 IEEE
DOI 10.1109/ICDE48307.2020.00198



(a) MFNP, Uganda (b) QENP, Uganda (c) SWS, Cambodia

Fig. 2: The protected areas used in this study. Visualized are the historical patrol effort for each protected area.

TABLE I: About the datasets

	MFNP	QENP	SWS	SWS dry
Number of features	22	19	21	21
Number of 1×1 km cells	4,613	2,522	3,750	3,750
Number of points (6 years)	18,254	19,864	43,269	30,569
Number of positive labels	2,602	937	155	76
Percent positive labels	14.3%	4.7%	0.36%	0.25%
Avg. patrol effort (km/cell)	1.75	2.08	3.96	3.03

are limited to managing data. To leverage that data to inform patrol strategy, PAWS is being integrated into the SMART software and will become available to park managers around the world. The enhancements and field tests in this paper outline significant steps to expand PAWS on a global scale.

II. Related Work

In the anti-poaching predictive modeling literature, [3] introduced PAWS as a stochastic behavioral model; the project has since grown into a multi-year, multi-university effort.

Conceptual thread of PAWS: The origins of PAWS were developed in conceptual papers using game theory and machine learning. CAPTURE [4] used a two-layered Bayesian network with latent variables to model imperfect detection of poaching activity. In contrast, INTERCEPT [5] offered an ensemble of decision trees that did not assume imperfect detection of poaching activities but achieved better runtime and performance than CAPTURE. The patrol planning component of PAWS is developed through a game theoretical framework [6], which extends the models of Stackelberg Security Games [7], [8] to the setting of environmental sustainability and wildlife protection, known as Green Security Games (GSGs), where adversaries may not behave as perfectly rational utility maximizers [9]. The iWare-E ensemble model, which we build upon in this paper, was introduced in [10].

Practical thread of PAWS: In the first field test of predictive modeling for poaching, [5] worked with park rangers in Queen Elizabeth for a one-month trial. However, the results of that trial are inconclusive as there was no control group. [11] extended field tests in Queen Elizabeth, but these tests were also limited: they had only two experiment groups (high and low risk), and only 9% of the areas used in the field test were considered high-risk of poaching.

III. Wildlife Crime Domain and Data

We study Murchison Falls National Park (MFNP) and Queen Elizabeth National Park (QENP) in Uganda, and Srepok Wildlife Sanctuary (SWS) in Cambodia.

We discretize the protected areas into 1×1 km grid cells. Each cell is associated with static geospatial features unique to each park such as rivers, roads, elevation, slope, forest cover, and animal density. We partition time into three-month time intervals. To combat poaching, rangers conduct patrols through protected areas and use GPS trackers to record their observations [12]. We rebuild historical patrol effort from these observations by using sequential waypoints to calculate patrol trajectories. For each time interval, we aggregate the patrol effort at each cell in terms of distance patrolled (in km).

The datasets $\mathcal{D} = (\mathbf{X}, \mathbf{y})$ are built from these historical patrol observations. The input matrix is $\mathbf{X} \in \mathbb{R}^{T \times N \times k}$ with T time steps, N locations, and k features. Each feature vector $x_{t,n}$ contains time-invariant geospatial features associated with each location (described above) and one time-variant covariate: $c_{t-1,n}$, the patrol coverage in cell n during the previous time step $t-1$, which models the potential deterrence effect of past patrols. The labels $\mathbf{y}$ are a binary indicator of whether illegal poaching activity was observed in a cell at a given time step. We assign a positive label $y_{t,n} = 1$ if rangers observed poaching-related activity and negative label $y_{t,n} = 0$ otherwise.

IV. Predictive Modeling with Uncertainty

The first stage of PAWS is a predictive model to identify the relative risk of poaching throughout a protected area. We build on the **i**mperfect observation-a**ware** **E**nsemble model (iWare-E) proposed by [10]. The iWare-E ensemble generates subsets of the data by filtering patrol effort across different thresholds to produce an ensemble of learners. Each weak learner $\mathcal{C}_{\theta_i^-}$ is trained on a subset of the dataset $\mathcal{D}_{\theta_i^-}$ with patrol effort c_v below a threshold θ_i. Due to label imbalance, we discard only negative samples and keep all positive samples, even those below the specified threshold; this is one of the key insights of the iWare-E approach.

We enhance iWare-E to explicitly reason about uncertainty of the predictions. We augment the ensemble with Gaussian process (GP) classifiers [13] as the weak learners. GPs are given by the function $f(\mathbf{x}_i) \sim \mathcal{GP}\left(\mu(\mathbf{X}), \Sigma(\mathbf{X})\right)$, with mean $\mu(\mathbf{X})$ and covariance matrix $\Sigma(\mathbf{X})$. GPs compute a variance value associated with each prediction based on confidence from the training data. Later on in patrol planning, we make use of these variance values as the metric for uncertainty to plan more informed patrol routes.

V. Evaluation of the Predictive Model

We study the predictive performance across different datasets by varying the weak learner used in iWare-E. We use three base classifiers: bagging ensembles of SVMs (SVB-iW), decision trees (DTB-iW), and Gaussian process classifiers (GPB-iW). We compare these models to baseline models, using those same bagging weak learners but without iWare-E (SVB, DTB, and GPB).

Table II presents the performance of the different predictive models evaluated on MFNP, QENP, and SWS. We generate

TABLE II: Comparing performance (AUC) of each model across all datasets

		without iWare-E			with iWare-E		
		SVB	DTB	GPB	SVB	DTB	GPB
MFNP	2014	0.518	0.587	0.626	0.695	**0.711**	0.685
	2015	0.504	0.620	0.670	0.683	0.706	**0.726**
	2016	0.513	0.589	0.603	0.672	0.680	**0.681**
	Avg	0.512	0.599	0.633	0.683	**0.699**	0.697
QENP	2014	0.505	0.654	0.693	0.619	**0.735**	0.717
	2015	0.501	0.589	0.600	0.632	0.696	**0.713**
	2016	0.502	0.635	0.611	0.644	0.728	**0.733**
	Avg	0.503	0.626	0.635	0.632	0.720	**0.721**
SWS	2016	0.815	0.774	0.656	**0.870**	0.865	0.825
	2017	0.672	0.670	0.673	0.744	0.742	**0.847**
	2018	0.518	0.499	0.527	0.510	0.541	**0.680**
	Avg	0.668	0.648	0.619	0.708	0.716	**0.784**
SWS dry	2016	0.500	0.490	0.648	0.501	0.610	**0.771**
	2017	0.500	0.486	0.615	0.581	0.681	**0.827**
	2018	0.500	0.526	0.615	0.500	0.576	**0.674**
	Avg	0.500	0.501	0.626	0.527	0.622	**0.757**

predictive poaching models with four years of data for each park, training on the first three years and testing on the fourth. This setup simulates the ability of each model to predict future incidences of poaching. The year listed is the test set; e.g., MFNP (2016) indicates that 2013–2015 were used for training and 2016 for testing. SWS experiences significant seasonality, so we also analyze SWS using data from just the dry season (November–April), as used for the field tests in Section VII-A.

iWare-E consistently improves AUC across all models, with an average increase of 0.100. Specifically, GPB-iW yields the best performance in over half the cases. Note that GPs achieve significant performance gains in SWS dry season, which has the strongest class imbalance with only 0.25% positive labels (see Table I).

The predicted probabilities and the corresponding uncertainties in MFNP as computed by the GPB-iW model are shown in Fig. 3. The riskmaps and uncertainty maps are displayed along with the historical patrol data over three years: total distance (km) that rangers patrolled in each cell (Fig 3a), and the number of incidences of illegal activity detected during those patrols (Fig 3b). The uncertainty is often higher in areas with less historical effort.

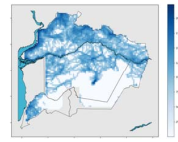
VI. Patrol Planning

Having assessed the relative risk of poaching activity with our predictive model, we incorporate uncertainty to plan optimal patrol routes. Conducting risk-adverse patrols enables us to increase detection of snares by an average of 30%.

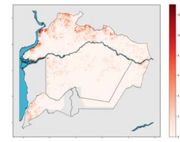
A. Prescriptive Modeling with Uncertain Crime Predictions

The planning model in [10] allows us to create paths, but does not account for uncertainty in the predictions. To account for this, we augment the model by using the uncertainty from the GPs to plan patrol routes. The GPB-iW model gives for each cell v a variance function $\nu_v : c_v \to \mathcal{V}_v$, where $\mathcal{V}_v$ is the variance for each prediction $g_v(c_v)$ at a given patrol effort c_v. The utility of patrolling a cell v then becomes:

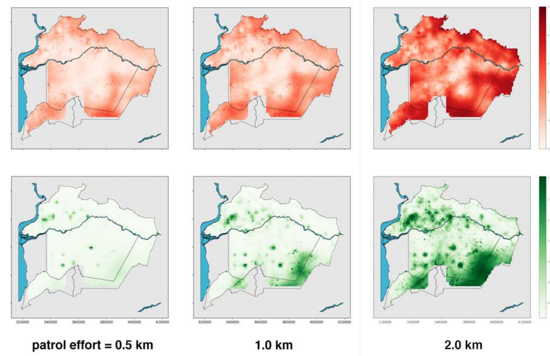
$$U_v(c_v) = g_v(c_v) - \beta g_v(c_v)\nu_v(c_v) \ . \qquad (1)$$



(a) Historical patrol effort from 2014–2016.



(b) Historical illegal activity detected from 2014–2016.



(c) Predicted probability of detecting of poaching activity (above, red) and corresponding uncertainty of the predictions (below, green).

Fig. 3: Predictions and uncertainties for different levels of patrol effort in MFNP in Uganda in the first quarter of 2017.

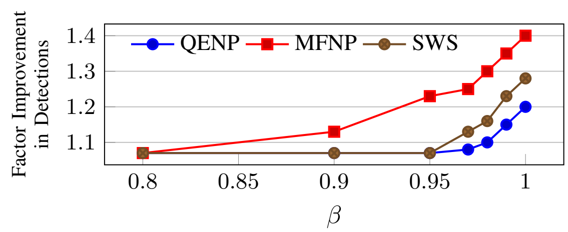


Fig. 4: Improvement in solution quality measured by the ratio $U_\beta(C_\beta)/U_\beta(C_{\beta=0})$ as a function of the tuning parameter β.

The uncertainty ν_v is squashed to the range $[0, 1]$ through a logistic function. $\beta > 0$ corresponds to a pessimistic, risk-averse strategy of patrolling less in cells with greater uncertainty. β thus becomes a parameter that enables us to tune the robustness of our plans; larger values allow us to optimize for plans that are more risk averse. We compute the optimal patrol by substituting

$$\sum_{v \in N} U_v^{\mathrm{PWL}}(c_v) = \sum_{v \in N} \left(g_v^{\mathrm{PWL}}(c_v) - \beta g_v^{\mathrm{PWL}}(c_v) \nu_v^{\mathrm{PWL}}(c_v) \right) \quad (2)$$

into our objective function, where, as in [10] we construct piecewise-linear approximations to g_v and ν_v so that the optimization problem is expressible as a MILP.

B. Evaluation of the Prescriptive Model

We compare the patrols computed with and without uncertainty scores by evaluating them on the ground truth given by the objective with uncertainty. We refer to the plan computed using uncertainty as $C_\beta = \mathrm{argmax}_c \sum_v g_v(c_v) - \beta g_v(c_v)\nu_v(c)$,

TABLE III: Field test results

Risk group	# Obs.	# Cells	Effort	# Obs. / # Cells
	SWS trial 1: Dec 2018–Jan 2019			
High	14	42	103.8	0.34
Medium	5	40	111.36	0.13
Low	0	36	84.12	0
	SWS trial 2: Feb–Mar 2019			
High	7	22	71.52	0.31
Medium	2	25	69.58	0.08
Low	0	34	89.07	0
	MFNP trial 1: Nov–Dec 2017			
High	6	18	71.6	0.33
Medium	5	21	31.9	0.24
Low	2	10	12.6	0.20
	MFNP trial 2: Jan–Mar 2018			
High	17	36	197.4	0.47
Medium	7	34	83.4	0.21
Low	1	13	45.1	0.08

where $C_{\beta=0}$ is a plan which does not account for uncertainty and $C_{\beta=1}$ is a fully robust plan. We then evaluate each of the plans using a utility function $U_\beta(C)$ and compute the ratio of the solution quality of the plan at a given β to the baseline of $\beta = 0$, $U_\beta(C_\beta)/U_\beta(C_{\beta=0})$. The results are shown in Fig. 4.

VII. Field Tests

Before scaling PAWS globally to over 800 parks worldwide through integration with SMART, it is critical to test these algorithms on the ground to ensure strong performance across diverse environments.

A. Field Tests in Srepok Wildlife Sanctuary

In partnership with WWF Cambodia, we conducted field tests of our anti-poaching predictive algorithm in SWS, the first deployment in Southeast Asia.

We built an iWare-E model with GPs as the weak learner trained on dry season data from 2015 through 2018. From the predictions of poaching risk, we discarded all 3×3 blocks with historical patrol effort above the 50th percentile, to ensure we were assessing the ability of our model to make predictions in regions with limited data rather than relying on past patterns, as [5] did. We then identified high-, medium-, and low-risk areas by considering blocks with risk predictions within the 80–100, 40–60, and 0–20 percentiles. In total, we selected five 3×3 km blocks from each of the three risk categories. To prevent bias, we did not reveal the risk category of each region to the rangers.

In December 2018, 72 park rangers in teams of eight began conducting patrols throughout the park, focusing on our suggested areas. As shown in Table III, our predictive model effectively evaluates the poaching threat across the park, with great success at discriminating between high-risk and low-risk areas. *Park rangers found absolutely no poaching activity in low-risk areas, despite exerting a comparable amount of effort in those regions.* This result suggests that revealing our risk predictions would grant them valuable insight into their patrol strategy so as to more effectively allocate their limited resources, as they can confidently spend less time patrolling these low-risk regions.

B. Field Tests in Murchison Falls National Park

Field tests in MFNP were executed with a similar process as SWS. During the five-month deployment with the Uganda Wildlife Authority, rangers detected poaching activity in 38 cells. As shown in Table III, the algorithm again effectively learns to discriminate poaching threat.

VIII. Conclusion

To assist rangers in combating illegal wildlife poaching, we take an end-to-end approach to the data-to-deployment pipeline by identifying an unaddressed need: to conduct risk-averse patrols and maximize the number of snares removed, a challenge exacerbated by uncertainty and bias in the historical data. During our field tests, rangers detected 38 attacked cells in MFNP and confiscated over 1,000 snares in SWS in a single month. We are enthusiastic about these real-world successes, which indicate that a data-to-deployment pipeline can be effectively implemented for wildlife protection and ought to be extended to other domains.

Acknowledgment: This work was supported by the Army Research Office (MURI W911NF1810208).

References

[1] R. Cooney, D. Roe, H. Dublin, J. Phelps, D. Wilkie, A. Keane, H. Travers, D. Skinner, D. W. Challender, J. R. Allan *et al.*, "From poachers to protectors: engaging local communities in solutions to illegal wildlife trade," *Conservation Letters*, vol. 10, no. 3, pp. 367–374, 2017.

[2] SMART, "Spatial monitoring and reporting tool," http://smartconservationtools.org/, 2013.

[3] R. Yang, B. Ford, M. Tambe, and A. Lemieux, "Adaptive resource allocation for wildlife protection against illegal poachers," in *AAMAS*, 2014, pp. 453–460.

[4] T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow *et al.*, "Capture: A new predictive anti-poaching tool for wildlife protection," in *AAMAS*, 2016, pp. 767–775.

[5] D. Kar, B. Ford, S. Gholami, F. Fang, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga *et al.*, "Cloudy with a chance of poaching: Adversary behavior modeling and forecasting with real-world poaching data," in *AAMAS*, 2017, pp. 159–167.

[6] F. Fang, T. Nguyen, B. Ford, N. Sintov, and M. Tambe, "Introduction to green security games," in *IJCAI*, 2015.

[7] M. Tambe, *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, 2011.

[8] D. Korzhyk, V. Conitzer, and R. Parr, "Solving stackelberg games with uncertain observability," in *AAMAS*, 2011, pp. 1013–1020.

[9] M. P. Johnson, F. Fang, and M. Tambe, "Patrol strategies to maximize pristine forest area," in *AAAI*, 2012.

[10] S. Gholami, S. Mc Carthy, B. Dilkina, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga, J. Mabonga, T. Okello, and E. Enyel, "Adversary models account for imperfect crime data: Forecasting and planning against real-world poachers," *AAMAS '18*, 2018.

[11] S. Gholami, B. Ford, F. Fang, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga, and J. Mabonga, "Taking it for a test drive: a hybrid spatio-temporal model for wildlife poaching prediction evaluated through a controlled field test," in *ECML*, 2017.

[12] R. Critchlow, A. J. Plumptre, B. Alidria, M. Nsubuga, M. Driciru, A. Rwetsiba, F. Wanyama, and C. M. Beale, "Improving law-enforcement effectiveness and efficiency in protected areas using ranger-collected monitoring data," *Conservation Letters*, 2016.

[13] C. E. Rasmussen, "Gaussian processes in machine learning," in *Advanced Lectures on Machine Learning*. Springer, 2004, pp. 63–71.