

Reducing Poaching in Elephant Populations on Random Forest Algorithm

Ye Jin^{*, #}, Yuxuan Liu[#] and Ji'an Wang[#]

Department of International, Beijing University of Posts and Telecommunications, Beijing, China

*Corresponding author: 1343799958@qq.com

#These authors contributed equally to this work

Abstract. Although most countries have established bans on the **ivory trade**, the illegal ivory trade continues to be a serious threat to the lives of populations due to, among other things, the existence of legal stockpiles and the desire for artifacts. To address this challenge, we have initiated a five-year data-focused program, the Ecological Tracking and Intelligent Prevention System (ETIPS). By utilizing **random forest algorithms** to predict areas where poachers are likely to be active, the project aims to open up new avenues for elephant conservation. This paper describes in detail the conceptualization and implementation plan of the project, and provides an in-depth analysis of the effectiveness of the project after implementation based on existing data. By providing theoretical support for the model's tracking effects and conservation outcomes, we expect to be able to provide more effective strategies and solutions for protecting these precious lives.

Keywords: ivory trade; elephant conservation; random forest algorithms.

1. Introduction

Wildlife is an important part of biodiversity. **Their existence adds to the variety of life on Earth and contributes to maintaining ecosystem** stability and sustainable development. But the illegal wildlife trade has exacerbated the biodiversity crisis, including genetic, ecological and behavioral diversity, at an alarming rate. This is particularly true of the illegal ivory trade. Although the 《Convention on International Trade in Endangered Species of Wild Fauna and Flora》 (CITES) voted to "ban" international trade in ivory in 1973, a 2014 article in 《Conservation Biology》 showed that "The global illegal ivory trade has doubled since 2007 [1]." This indicates that international trade in illegal ivory is still rampant. Of the more than 120,000 species monitored on the IUCN Red List of Threatened Species, the African savannah elephant is now classified as endangered and the African forest elephant has been classified as critically endangered in 2021 [2].

Therefore, in order to further combat the illegal ivory trade and protect elephants, we have proposed a five-year data-driven project, the Ecological Tracking and Intelligent Prevention System (ETIPS), based on existing literatures and analyses. Targeting international organisations such as wildlife conservation organisations as well as technology companies, this project aims to establish a global wildlife detection network using the latest technological tools. In the following, we will take African elephant herds as an example and develop a dynamic analysis model based on the Random Forest algorithm, which predicts the poaching of African elephant herds so that local wildlife conservation organisations can take timely action to restrict poachers, contributing to the fight against illegal wildlife trade.

2. Introduction to the ETIPS project

ETIPS is a data-driven, 5-year project that requires careful planning, substantial data and a strong capacity for innovation. To that purpose, we provide a brief description of its core ideas, potential effects of its role, and the ideal plan for conducting it.

2.1. General Description of ETIPS

In designing the Eco-Tracking and Smart Prevention System (ETIPS) project, we developed an innovative and practical approach: a detailed classification of global wild elephant habitats and migration routes. Through scientific data analysis and risk assessment, we classify areas into high-risk and low-risk zones. For areas where migration routes frequently pass through and other locations identified as high-risk, we plan to implement 'high-cost' measures, including the installation of ground-based bionic surveillance cameras, drone patrols, and the addition of manned outposts. For areas assessed as low risk, we will implement minimal protection measures to balance the overall cost-effectiveness of the project.

As described in Section 3, the core of the project is a dynamic analytical model we have developed that identifies and predicts, in real time, the high-risk areas where illegal poaching is most likely to occur, i.e., the priority prevention areas. The model requires a large amount of data to support it, such as the migration paths of herds of elephants and the density of villages, etc. Based on artificial intelligence, big data analytics and satellite tracking technology, we collect and analyze a large amount of environmental, climatic, and human activity data to accurately predict hotspots for the illegal trade and quickly locate poachers. Once the model identifies key prevention areas, we will implement comprehensive and effective protection measures for these areas, thus greatly improving the efficiency of the fight against illegal international ivory trade. This not only protects elephants, but also brings huge social benefits to the clients involved in the project.

On the basis of program implementation, we will conduct a detailed cost-benefit analysis to ensure the cost-effectiveness of the project. Through the cost assessment, we are able to provide potential business partners with accurate payback data, including the long-term economic and social benefits of the program. These data will provide an important basis for our clients' decision-making, helping them to realize that joining the ETIPS program can not only effectively help solve the problem of illegal ivory trade, but also enhance the social influence of international organizations, thus raising people's awareness of elephant conservation. It can also enhance the brand image of the organization and improve its innovation ability.

2.2. A Data-Driven 5-year Project

As in **Fig. 1**, the ETIPS project broadly follows the framework:

Planning - development - optimization - expansion - evaluation

Year 1 (Planning and Initial R&D): conducting market research and demand analysis. Select key technologies such as big data analytics and satellite tracking technology to form a cross-disciplinary team to commence development of the data collection platform and algorithms.

Year 2 (System Development and Testing): Technology companies complete development of smart prevention systems, including data processing and user interfaces. Internal testing as well as small-scale field pilot testing and system optimization are conducted with the help of the elephant herd data collected by international organizations in the African region.

Year 3 (System Deployment and Optimization): The team optimizes the system based on pilot feedback. At the same time, the system is fully deployed to key prevention zones and continuously monitored and dynamically optimized.

Year 4 (Expansion and Cooperation): Expand the scope of monitoring and establish cooperation with governments, international organizations and environmental NGOs to raise public awareness of protection.

Year 5 (Evaluation and Sustainable Development): Conduct a comprehensive evaluation, address strengths and weaknesses, formulate a future development plan, explore new application scenarios and business models, and ensure the sustainable development of the project.

Jan—Mar market survey	Jan—June	Jan—Mar adjustment and optimization	Jan—Dec	Jan—June
Apr—June technical research	The main development of the system	Apr—June overall action plan	Expand the system	Conduct a comprehensive evaluation
July—Sep Build project team	July—Sep ST (system testing)	July—Dec Dynamic optimization	Propaganda and public education	July—Dec
Oct—Dec research and development work	Oct—Dec Field pilot test	feature updates	Add new monitoring points	Make plans for future development
Timeline → 2025 → 2026 → 2027 → 2028 → 2029				

Fig. 1 Work division

3. Dynamic Analytical Model

This model is based on the Random Forest algorithm, and its advantages can be summarized as highly parallelized data, strong generalization ability of the constructed model, small variance, and insensitivity to the lack of data features. In the following, we will introduce this model in detail and simulate the effect of the project after implementation based on the collected data.

3.1. Data Collection and Pre-processing

3.1.1. Data Collection

It is estimated that the elephant population in Tanzania had declined to about 109,051 in 2009; the latest population survey conducted in 2013 shows that the shrinking of the elephant population has reached devastating proportions. The elephant population in Cerros has plummeted by 66% in just four years, from 38,975 in 2009 to 13,084 in 2013. This is the lowest level since 1976, when more than 100,000 elephants inhabited Cerros [3][4]. Since then, this downward trend has continued at an alarming rate.

In addition, not limited to data, we also drew information from distribution and migration maps.

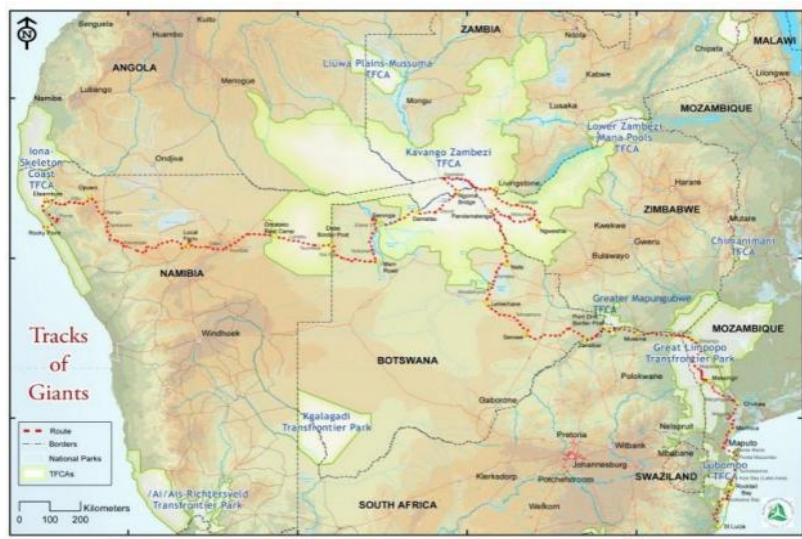


Fig. 2 Migration of elephant herds

Migration path data: Based on our collection of elephant migration maps (Fig. 2), it is clear that every year during the migration season there are large numbers of African savannah elephants spread across the area shown by the red line, and they are perfect targets for poachers.

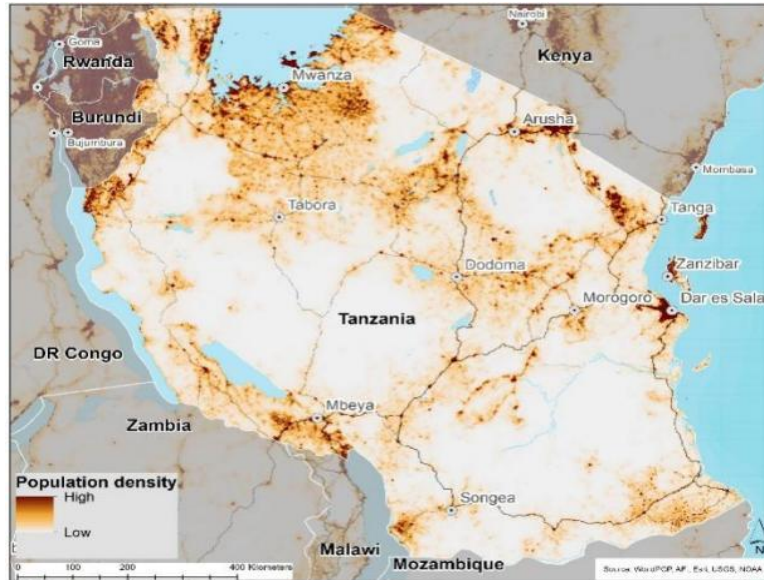


Fig. 3 Village distribution

Population and village distribution: Since we consider that the probability of poaching behavior is also related to the density of villages and population around the range of the herd, we obtained data on population and village distribution in the area shown in the Fig. 3.

3.1.2. Pre-processing

From the various types of maps presented in this section, as well as the data collected, we can extract the following metrics information, which is needed for model predictions:

a. **Mainstream migration paths of elephant herds:** Here we convert it to coordinate information in the latitude and longitude grid in order to make the machine easy to deal with.

b. **Climate information:** Since Tanzania has a significant savannah climate, here we input the dry season as number 1 and the wet season as 2 in order to simplify the model and facilitate data processing.

c. **Village and town density information:** We also use the coordinate method of transformation to delineate the coordinate range, but here in order to facilitate the quantitative description of the density data according to the population density gradient we divided into large, medium and small. We use the number 1 for small, 2 for medium and 3 for large.

d. **Poaching alert information:** Through relevant reports and surveys of elephant carcass distribution, we identified red areas as areas of high poaching risk, and in the datasets we divided the poaching alert level into five segments, from number 1 to 5 in ascending order.

Following the method described above we collated the data required for model training, and only the data processing diagrams for the three most important paths are listed here (Fig. 4, Fig. 5, Fig. 6):

migrate	climate	populatio	alert	
37.43E,7.74S	2	1	5	5
35.81E,7.10S	2	2	5	5
37.55E,8.24S	2	1	4	4
36.63E,7.60S	2	2	4	4
33.88E,6.81S	2	1	4	4
31.18E,6.70S	1	3	3	3
37.30E,9.20S	2	1	2	2
34.85E,7.44S	2	1	2	2
34.33E,7.22S	2	1	2	2
32.00E,7.00S	1	2	2	2
36.80E,9.85S	2	1	1	1

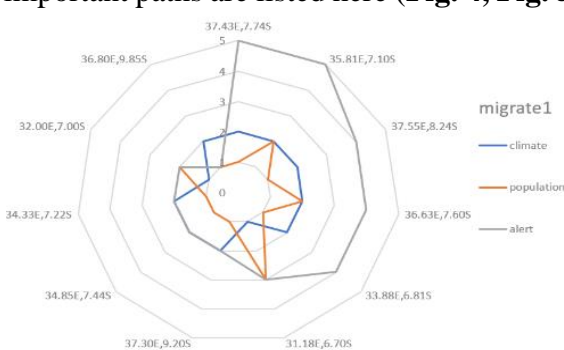


Fig. 4 First path

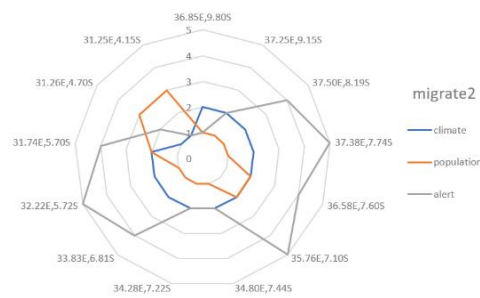
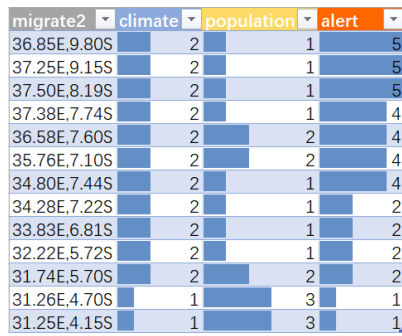


Fig. 5 Second path

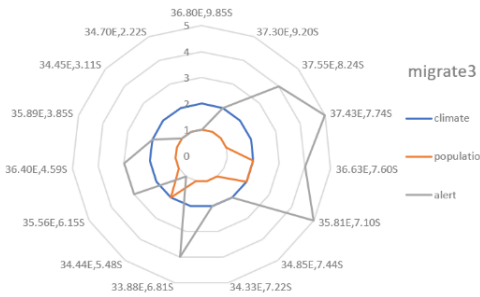
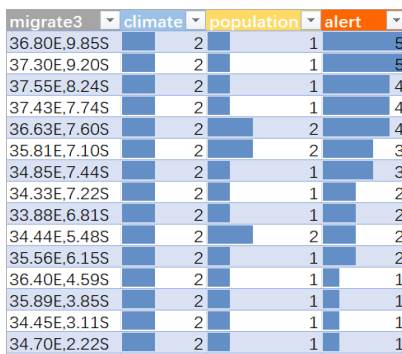


Fig. 6 Third path

3.2. Random Forest Algorithm

Random forest algorithm belongs to one of the extensions of Bagging algorithm, which is based on the decision tree model as the base learner, through the construction of a combination of multiple decision tree models. Random Forest algorithm is one of the extensions of Bagging algorithm.

In order to construct diversified decision trees, the random forest algorithm introduces random attributes in the model training process.

Randomness is divided into two randomness: the first one is data sampling randomness, we now number the collected data by year as well as by region, and generate random numbers according to python's random package to select them; the second one is the random extraction of features, according to the frequency of the potential features appearing in the process of data collection, we screen out the features that have the greatest correlation with the number of elephant herds, and carry out a feature 3.1 transformation.

The regression model of random forest can be interpreted as follows: assuming that the independent variable X of the input data and the dependent variable Y of the real results are distributed independently (X is the selected feature and Y is the location of elephant carcasses/poachers, which is both the key prevention area), randomly generating the training set θ in (X, Y), and predicting the results as $g(X)$, the mean square error of the model can be defined as:

$$E_{X,Y}[Y - g(X)]^2 \tag{1}$$

The model uses the average of the predicted values of h decision trees as the prediction result, if h tends to infinity, then there are:

$$E_{X,Y}[Y - g_h(X, \theta_h)]^2 \rightarrow E_{X,Y}[Y - E_\theta(X, \theta_h)]^2 \tag{2}$$

In equation (2), the left side represents the mean square error of the average prediction result of h decision trees, denoted as S_{E1} . When h is infinite, the average mean square error of a single decision tree is denoted as S_{E2} . S_{E2} satisfies:

$$S_{E2} = E_\theta E_{X,Y} [Y - g(X, \theta)]^2 \tag{3}$$

Assuming that the correlation coefficient between each residual is α , then θ in equation (3) is satisfied:

$$S_{E1} \leq S_{E2} \quad (4)$$

Then the final random forest model can be expressed as:

$$Y = E_{\theta}g(X, \theta) \quad (5)$$

In this paper, the mainstream migration path of elephants, climate information, village and town density information, poaching alert information and other data are treated as different variables of random forest. Through the decision tree adaptive determination of the weights of different variables, and then realize the high precision inversion of the poacher's location, so as to determine the key prevention area.

4. Summary

This paper describes the 5-year plan and the core algorithmic model of the ETIPS project, which aims to reduce the illegal international trade in ivory and establish a global wildlife monitoring network through modern technology. Based on the Random Forest algorithm, we use data collected on African elephant herds over the years and test the feasibility of the model. The project is targeted at specific clients and caters to the client's needs and capabilities. The core technology required for the project follows the general trend of the times and will effectively protect wild elephant herds after implementation. The core of the algorithm requires the accumulation of data in quantity and quality, and effective data can empower our model.

After the implementation of the project, ETIPS will not only be limited to the protection of elephants, but can also be extended to a variety of wildlife. In addition to predicting the path of poaching, ETIPS can also be used for field exploration to provide the necessary protection for wildlife living in nature, so that more people can gain a deeper understanding of wildlife [5].

References

- [1] Yufang Gao, Susan G. Clark, Elephant ivory trade in China: Trends and drivers, Biological Conservation, Volume 180, 2014,
- [2] Ahumada J A, Silva C E F, Gajapersad K, et al. Community structure and diversity of tropical forest mammals: data from a global camera trap network[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2011, 366(1578): 2703-2711.
- [3] Hauenstein, S., Kshatriya, M., Blanc, J. et al. African elephant poaching rates correlate with local poverty, national corruption and global ivory price. Nat Commun 10, 2242 (2019).
- [4] Symes, W.S., Edwards, D.P., Miettinen, J. et al. Combined impacts of deforestation and wildlife trade on tropical biodiversity are severely underestimated. Nat Commun 9, 4052 (2018).
- [5] Guergana Guerova, Tzvetan Simeonov, Chapter 1 - Global Navigation Satellite System (GNSS—GPS, GLONASS, Galileo), Editor(s): Guergana Guerova, Tzvetan Simeonov, Global Navigation Satellite System Monitoring of the Atmosphere, Elsevier, 2022.