

Study on Illegal Trade in Wildlife Based on AHP and ARMI-Intervention-SVR Series Models

Yingping Tang*, Xinle Wang and Mingle Chen

International Business School, Beijing Foreign Studies University, Beijing, China

*Corresponding author: 202220102163@bfsu.edu.cn

Abstract. This study looks at the environmental impact of the illegal trade in the world's fourth largest wildlife and aims to combat it. By setting two project objectives, namely strengthening law enforcement and enhancing community economic capacity, with the government as the focus client. In the data processing stage, the government of Vietnam was identified as the target client by assigning weights using the Analytical Hierarchy Processing (AHP) and entropy weighting methods and combining the weights through the Minimum Information Entropy (MIE) and Lagrange Multiplier (LM) methods. In terms of project impact quantification, an intervention analysis model is introduced to establish the ARMI-Intervention-SVR series of models, and the PSO algorithm is used to find the best fitting parameters of the SVR model to improve the model performance. After combining the intervention prediction models, it was concluded that the impact of the Vietnam project on the five-year transaction volume of illegal wildlife trade was 31.22 per cent, 28.87 per cent, 18.25 per cent, 13.77 per cent and 12.31 per cent, respectively. These results provide an important reference for effectively combating illegal wildlife trade and are of practical significance.

Keywords: AHP; entropy-weighted; ARMI-Intervention-SVR.

1. Introduction

The amount of money involved in the world's fourth largest illegal trade in wildlife and its impact on the environment have attracted international attention. As the scale of the global illegal wildlife trade continues to grow, the harm caused to ecosystems and species diversity is becoming increasingly evident. In order to effectively combat this phenomenon and protect wildlife resources, this study aims to explore ways to strengthen law enforcement measures, enhance community economic capacity, and focus on the government to address this challenge [1].

In the current environment, Vietnam faces serious challenges as a major hub for illegal wildlife trade. Therefore, in this study, Vietnam was selected as the focus and the data was processed through methods such as hierarchical analysis and entropy weighting in order to identify the target customers of the project. Meanwhile, by introducing the intervention analysis model and SVR series model, we quantify the impact of the project on illegal wildlife trade to provide a scientific basis for future policy formulation and implementation.

This study aims to provide empirical support to better understand and respond to the problem of illegal wildlife trade and contribute to the protection of the environment and biodiversity. Through in-depth analysis of the implementation effects of the Vietnam project, we will provide lessons and insights for combating illegal wildlife trade globally.

2. Analysis of Compatibility between Customer and Program

2.1. Data Acquisition and Pre-processing

This paper collects 11 categories of indicators from 267 countries and territories around the world, covering a wide range of aspects of social development and human well-being. We then pre-processed the data before applying it to construct the model.

2.2. Customer Selection: using EWM-AHP Model

3.2.1 Data Standardization

Consistent treatment of indicators The scale and distribution of the different indicators vary considerably. Therefore, we should standardize the data according to the type of indicator. Since all our indicators are very large and non-positive, the formula for converting to positive indicators is as follows.

Dimensionless processing of indicators Then we should transform the data to the same interval to remove the effects caused by the scale. For subsequent use of the entropy weighting method, we processed the data using the following.

To facilitate subsequent processing, we first processed the data using the polar transform method.

$$a_{ij}^* = \frac{a_{ij} - \min\{a_{1j}, a_{2j}, \dots, a_{nj}\}}{\max\{a_{1j}, a_{2j}, \dots, a_{nj}\} - \min\{a_{1j}, a_{2j}, \dots, a_{nj}\}} \quad (1)$$

In which, a_{ij} denoting the data for the j th country under the i th indicator. $i = 1, 2, \dots, 11; j = 1, 2, \dots, m$.

The data from the NCA were non-negative using the efficacy coefficient method in addition to the polar transform method:

$$a_{ij}^{**} = C a_{ij}^* + D \quad (2)$$

Where the scaling factor $C = 0$ and the translation factor $D = -\min a_{ij} = 0.042$, ($j = 1, 2, \dots, m$).

3.2.2 Model Establishment: based on EWM-AHP-Topsis

1) The Entropy Weight Model

According to the interpretation of the basic principles of information theory, information is a measure of the degree of order of the system, and entropy is a measure of the degree of disorder of the system. The smaller the information entropy value, the greater the degree of dispersion of the index, and the greater the weight of the index to the comprehensive evaluation. Therefore, information entropy can be used as a tool to calculate the weight of each indicator and provide a basis for the comprehensive evaluation of multiple indicators [2].

The specific steps are as follows. *i.* determine the evaluation object, *ii.* establish the evaluation index system, *iii.* construct the level matrix, *iv.* standardize the evaluation matrix to obtain the matrix R .

Then calculate the entropy value of each index, where f_{ij} is the weight of the index value of the i_{th} item under the j_{th} index.

$$H_j = -k \sum_{i=1} f_{ij} \cdot \ln f_{ij} \quad (3)$$

$$f_{ij} = \frac{r_{ij}}{\sum_{i=1}^m r_{ij}}, k = \frac{1}{\ln m} \quad (4)$$

After calculating the entropy weight of the j_{th} index w_j , calculate the comprehensive weight of the index λ_j , and modify the weight obtained by the entropy weight method.

$$w_j = \frac{1-H_j}{\sum_{j=1}^n (1-H_j)} = \frac{1-H_j}{n-\sum_{j=1}^n H_j} \quad (5)$$

$$\lambda_j = \frac{\lambda'_j w_j}{\sum_{j=1}^n \lambda'_j w_j} \quad (6)$$

2) Analytic Hierarchy Process (AHP)

A decision analysis method that combines qualitative and quantitative methods, which was proposed by the American operations researcher, Prof. T. L. Saaty of the University of Pittsburgh, in the early 1970s. Analytic Hierarchy Process(AHP) is a method to decompose a decision problem into

different hierarchical structures through in-depth study of the nature of the decision problem, the influencing factors and their internal relationships, then use mathematical methods to calculate the weights reflecting the relative importance of the elements in each level [3].

We usually use the following formula to test whether the judgment matrix has satisfactory consistency:

$$CR = \frac{CI}{RI} = \frac{\lambda_{\max} - n}{(n-1)RI} \quad (7)$$

Where: CR is the test coefficient, RI is the random consistency indicator, n is the number of indicators, $\lambda_{\max}$ is the maximum characteristic root of the judgment matrix.

3) TOPSIS Model

Technique for Order Preference by Similarity to an Ideal Solution is a method of ranking according to the proximity between a limited number of evaluation objects and the idealized target, which is to evaluate the relative advantages and disadvantages of the existing objects.

In order to simplify the analysis and make our model convincing, we should first perform the standardization of data according to the methods mentioned above. Here we transform all the indicators to the NEGATIVE type to make it more intuitive. Then we determine the best and the worst set of data to evaluate the target (the maximum and the minimum). We then calculate the Euler distance between each sample and these two vectors separately.

$$D_+(j) = |X - X_{\text{best}}| = \sqrt{\sum_{i=1}^m (x_{ij} - x_{\text{best}}(i))^2} \quad (8)$$

$$D_-(j) = |X - X_{\text{worst}}| = \sqrt{\sum_{i=1}^m (x_{ij} - x_{\text{worst}}(i))^2} \quad (9)$$

In the end, we would determine the determination coefficient. The larger c_j is, the more important the j^{th} indicator is to the target to be evaluated.

$$C_j = \frac{D_-(j)}{D_-(j) + D_+(j)} \quad (10)$$

3.2.3 Model Solving: Using Minimum Relative Information Entropy

We obtained 11 different indicators of client-program fit. For clarity, we categorized these metrics into three groups: power (P), resources (R), and interest (I).

First, for each of the Level 1 indicators, we apply the AHP to determine the weights, and then use the EWM to determine the weights of the Level 2 indicators below them.

The calculated CR value is 0.003185, which is within the range of (0.1) and meets the consistency requirements, which in turn leads to the first-level indicator weights:

$$w_j^1 = [0.10945229, 0.30899564, 0.58155207]^T \quad (11)$$

Then, in order to make up for the lack of a single assignment, we use a combination of subjective empowerment method (AHP) and objective empowerment method (EWM) assignment method with the principle of Minimum Relative Information Entropy, we establish an optimization model to minimize the relative deviation of the results under the two decision methods.

$$\min \sum_{j=1}^n w_j (\ln w_j - \ln \alpha_j) + \sum_{j=1}^n w_j (\ln w_j - \ln \beta_j) \quad (12)$$

$$\text{s.t.} \begin{cases} \sum w_j = 1 \\ w_j > 0 \\ j = 1, 2, \dots, n \end{cases} \quad (13)$$

Lagrange Multiplier Method is used to solve the above optimization problem [4], and we got the final weights W_j :

$$W_j = \frac{\sqrt{w_j^1 w_j^2}}{\sum_{j=1}^n \sqrt{w_j^1 w_j^2}} \quad (14)$$

We integrate the weights in this way to obtain the weights for each secondary indicator as shown in Fig. 1.

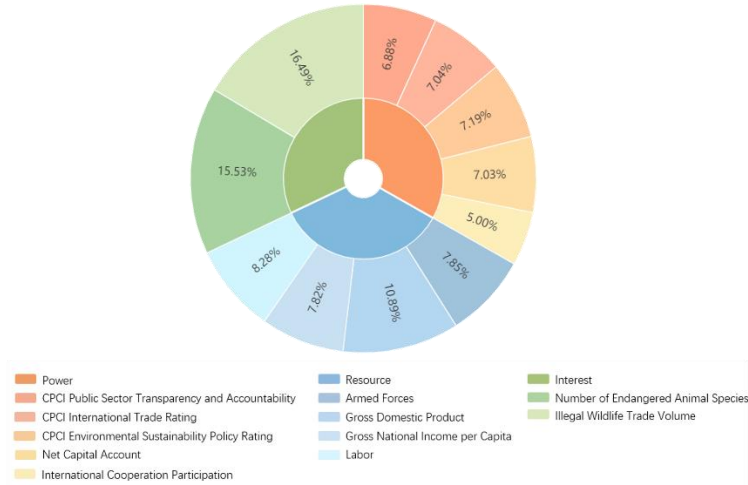


Fig. 1: The Weights of Indicators

Finally, we perform a TOPSIS analysis to determine the relative proximity of each country to the Ideal Sister as a measure of client fit. The larger the value corresponding to the country, the more suitable the client is for the program. The top ten clients with more ideal conditions are shown in Fig. 2.

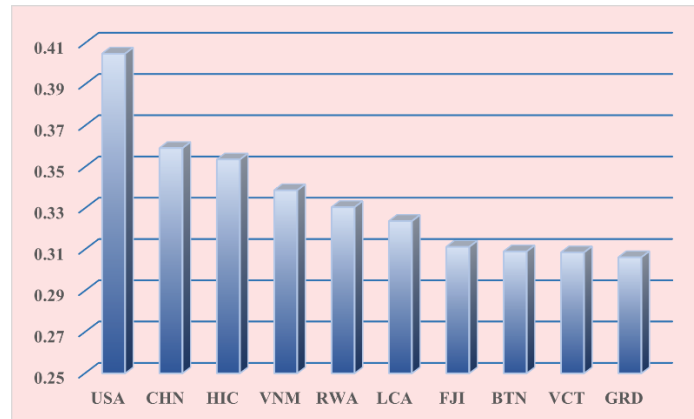


Fig. 2 Top Ten Customers

3. Analysis of the Impact of Our Model on IWF

3.1. Establishment of ARMI-Intervention-SVR Series Model

Autoregressive Integrated Moving Average Model (ARIMA) combining the autoregressive model (AR) [5], the moving average model (MA) and the difference method:

$$X = \alpha_1 X_{t-1} + \alpha_2 X_{t-1} + \dots + \alpha_p X_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (15)$$

We obtain the differential autoregressive moving average model ARIMA (p, d, q), where d is the order in which the data need to be differenced.

Intervention Analysis Model (IAM) is a model proposed by American statisticians Box and Tiao to quantify the specific impact of an intervention event on an economic environment, process or

outcome. The Intervention Analysis Model is widely used to measure the impact of a policy or an event on the subject of study.

Combined forecasting model is a model proposed by Bates and Granger scholars to use different forecasting methods for a problem and combine them appropriately to make full use of the useful information provided by each method and improve the forecasting accuracy. It overcomes the limitations of a single model.

Aiming at the linear and nonlinear characteristics of the data of illegal wildlife trade in reality, as well as its susceptibility to certain policy changes or unexpected events and outliers, we propose the ARIMA-Intervention-SVR tandem model to deal with this type of complex time-series data to improve the prediction accuracy compared to a single model when intervened by unusual events, and to predict the development trend effectively. In this way, we are able to test the results of the project implemented by our client.

Here, the time series data are divided into three parts: the linear part, the intervention event and the nonlinear residuals, i.e.,

$$Y_t = L_t + I_t + N_t \quad (16)$$

The construction of the ARIMA-Intervention-SVR tandem model is divided into the following steps:

3.2. ARIMA Modeling

The data from 2000 to 2012 before being affected by several policies are the preintervention data, for which a SARIMA model is built, firstly, the data are subjected to the S-W test, and $p = 0.045$ does not obey the normal distribution. Then unit root test (ADF) was performed to test the smoothness of the series. The results show that the series is a smooth time series when the difference is of first order, followed by the Ljung-Box test, which shows that the series has significant autocorrelation. Based on the above results, we can determine that the series is a smooth non-white noise series, so the ARIMA model is used. Then the partial autocorrelation test is carried out and then the optimal ARIMA model is established based on the Akaike Information Criterion (AIC), and the final ARIMA (1,1,1) is used to represent the unintervened process, and then the parameters of the model are estimated by the method of Great Likelihood (MLE) [6], and the results are shown as follows:

$$(1 - 0.4690B)(1 - B)Y_t = (1 - 0.1932B)\epsilon_t \quad (17)$$

B denotes the backward shift operator, $(1 - B)$ denotes the difference operation, is the predicted value of the time series at time t , and ϵ_t is the error term.

3.3. ARIMA-Intervention modeling

After determining the model of the basic no-intervention process, the next step is to establish a model for the data of the intervention process, first test the outliers, identifiable outliers are divided into two kinds, addable outliers (AO) and new interest outliers (IO), identified outliers are added to SARIMAX in python, and the impact of the intervention event on the sequence is calculated. Due to the implementation of several strong policies to combat illegal wildlife trade from 2012 to 2014, the three of 2013, 2014, and 2015 are outliers, and by incorporating these three outliers into the model, it is found that the AIC becomes smaller, which indicates that the model fit is enhanced. Construction of the m_t model:

$$m_t = \omega_0 P_t^{(T)} + \frac{\omega_1}{1 - \omega_2 B} P_t^{(T)} \quad (18)$$

In the above equation,

$\omega_0 + \omega_1$ tenor represents an immediate effect produced by a contingency.

$\omega_1(\omega_2)^k$ represents the effect caused by the effect on the subsequent k months.

$$m_t = 0.0104 P_t^{(T)} - \frac{0.4209}{1 - 0.7395B} P_t^{(T)} \quad (19)$$

Impulse function

$$P_t^{(T)} \begin{cases} 1, t = T \\ 0, \text{others} \end{cases} \quad (20)$$

According to the above equation, it can be seen that the effect of relevant policy intervention has reduced the volume of illegal wildlife trade by $33.7\% = 1 - \exp(0.0104 \times 0.4209) \times 100\%$, and the transaction volume has decreased by $1 - \exp(0.4209 \times 0.7395k) \times 100\%$ after k years. Fig. 3 10 gives an estimate of the impact of national policies on transaction volume, and shows that the impact of national policies is gradually decreasing.

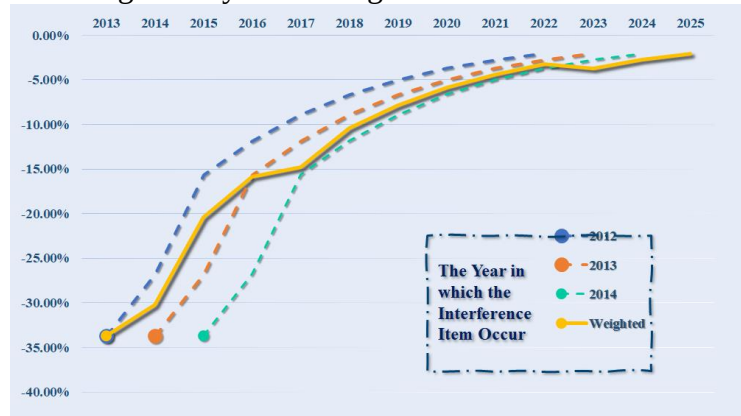


Fig. 3: Comparison Chart of Policy Interference Effects

AO means that the underlying process is superimposably perturbed at the moment of work, which can be viewed as an intervention with an impulse response at the moment of work, and the sequence is affected by AO only at the moment of work. If the time series has an additive outlier at the moment of work, then the sequence is only affected by AO at the moment of work. If an additive outlier occurs at the moment of work, then the sequence becomes:

$$Y'_t = Y_t + \omega_A P_t^{(T)} \quad (21)$$

Bringing $P_t^{(T)}$ into equation 23, the above expression becomes:

$$Y'_t \begin{cases} Y_t + \omega_A, t = T \\ Y_t, t \neq T \end{cases} \quad (22)$$

IO is said to occur if the error at time work T is perturbed, and the perturbed error can be expressed as:

$$e'_t = e_t + \omega_I P_t^{(T)} \quad (23)$$

ω_I usually represents the coefficient of the Innovation Outlier, which is a parameter used to capture a momentary shock at a specific point in time in a time series.

Bringing $P_t^{(T)}$ into the above equation, the expression becomes:

$$e'_t = \begin{cases} e_t + \omega_I, t = T \\ e_t, t \neq T \end{cases} \quad (24)$$

Assume that the undisturbed process is smooth and can be expressed in the form of $MA(\infty)$: $Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots$.

Then, the process subjected to perturbation can then be written as:

$$Y'_t = e'_t + \psi'_1 e'_{t-1} + \psi'_2 e'_{t-2} + \dots \quad (25)$$

The above equation is equivalent to:

$$Y'_t = Y_t + \psi_{t-T} \omega_I \quad (26)$$

Where $\psi_0 = 1$, and $\psi_j = 0$ when j is negative. IO has an effect on the observations at moment T and beyond, which diminishes over time but does not disappear completely. Combining the above process, the intervention model was added to the ARIMA model to get:

$$Y'_t = ARIMA(Y_t) - 0.3922BY_t - 0.2879B^SY_t + 0.2116P_t^{(2012)} \quad (27)$$

The superscript S usually indicates the number of lags of the seasonal component.

3.4. ARIMA-InTERVENTION-SVR modeling

The fitted values as well as the predicted values obtained from ARIMA-Intervention above are subtracted from the true values of the sequences to obtain the fitted residual sequences, and SVR training is performed. First, 80% of the residuals are randomly selected as the training set, and the rest is used as the validation set to adjust the parameters of the model and evaluate the model capability.

Due to the small amount of data in this problem and in order to avoid overfitting, the SVR model applies a sliding window approach to determine the training set as well as the test set.

Although the theory of SVR model has been perfected, there are still difficulties in parameter selection. In order to find the optimal parameter pairs of penalty coefficients C and kernel function coefficients g , a swarm intelligence optimization algorithm is used to optimize the parameters of SVR model. The swarm intelligence algorithm is characterized by flexibility, simplicity and robustness, and the particle swarm algorithm is chosen for parameter searching because of its high accuracy and difficulty in falling into local optimality.

3.5. Particle Swarm Algorithm (PSO)

An algorithm proposed by Eberhart and Kennedy in 1995 through the study of the foraging behavior of natural bird group collaboration. The basic principle of the algorithm is to mimic the process in which birds continuously cooperate and exchange information during foraging, so that the whole foraging action can be changed from disordered to ordered, thus obtaining better food locations faster.

The initial range of C and g is set to be 0.0001-100000, and then the model performance is evaluated by the ten-fold cross-validation method, and it is verified that, for this problem, c and g are both 0.1 is the best prediction of SVR, and then the optimal parameter is used to fit the SVR model, and then the residual value of the prediction and the prediction of the ARIMA-Intervention model are used to get the final prediction result.

3.6. Model testing

In order to effectively evaluate the prediction results of the models, three error indicators are chosen to judge the prediction results of each model:

Root Mean Squared Error (RMSE)

Mean Absolute Error (MAE)

Mean Absolute Percentage Error (MAPE)

Next, we constructed other models and compared them with the final model, and the results are shown in Table 1:

Table 1. Comparison Chart of Policy Interference Effects

Model	RMSE	MAE	MAPE (%)
ARIMA	44.305	33.35	5.7323
SVR	99.577	79.45	13.0296
ARIMA-SVR	23.46	18.72	3.0558
ARIMA-Intervention	28.617	22.83	4.1540
ARIMA-Intervention-SVR	16.510	13.17	2.3350

Therefore, based on the above process, we chose the ARIMA-Intervention-SVR model to determine the impact of our project. Finally, it can be concluded that after the relevant anti-illegal wildlife trade policies were introduced, the trade volume decreased by 31.22%, 28.87%, 18.25%, 13.77% and 12.31% with the passage of years. When the client chooses our project and fully implements it, it will effectively combat illegal wildlife trade.

4. Summary

Through this study, we have successfully quantified the impact of the project on the illegal wildlife trade and identified the Government of Vietnam as an important target client. By introducing the intervention analysis model and the SVR series of models, we have produced the results of an impact analysis of the project's five-year trade volume in Vietnam, which provides an important reference for future policy development and implementation. This research provides empirical support for us to better understand and respond to the problem of illegal wildlife trade.

However, we also realise that many challenges and opportunities remain in this area. In order to combat illegal wildlife trade more effectively, we suggest that future research and actions could further explore community participation and co-operation mechanisms, strengthen international co-operation, and enhance law enforcement power and technology. In addition, the differences between different countries and regions can be studied in depth to explore more personalised and effective solutions.

In summary, this study provides important empirical support and insights for us to understand and respond to the problem of illegal wildlife trade. It is hoped that we can continue to conduct in-depth research in the future and make joint efforts to protect valuable biological resources, maintain ecological balance and achieve the goal of sustainable development.

References

[1] Ma Zhongfa, Xu Ziyun. Study on the Supervision Mechanism of Trade in Endangered Species of Wild Fauna and Flora[J]. Journal of Capital Normal University (Social Science Edition),2021,(03):47-58.

[2] CHEN Xiaoguo,DING Pengkun,LI Zhenyu. Evaluation of power emergency plan indexes based on grey great entropy weights[J]. Electrical Automation,2020,42(02):37-39+63.

[3] HE Yinshui, XIAO He, LUO Canghai, et al. Autonomous decision-making of welding position for arc welding of thick plate T-joints based on hierarchical analysis process[J]. Journal of Jilin University (Engineering Edition),2024,54(03):657-662.DOI:10.13229/j.cnki.jdxbgxb.20220518.

[4] Lili Zhao,Lei Wang. Lagrange multiplier method and its promotion[J]. Journal of Henan College of Finance and Economics (Natural Science Edition),2024,33(01):50-54.

[5] SUN Jiale, MEI Yuan, XU Baichao, et al. Application of ARIMA seasonal model in the prediction of the incidence of out-of-hospital cardiac arrest in Haikou City[J/OL]. Journal of Preventive Medicine Intelligence,1-7[2024-05-30].<https://doi.org/10.19971/j.cnki.1006-4028.230595>.

- [6] Han Xianping. Research on high-precision positioning method of target range photoelectric longitude and latitude instrument based on MLE[J/OL]. Measurement and Control Technology,1-6[2024-05-30]. <https://doi.org/10.19708/j.ckjs.2024.01.203>.