

基于数据驱动的野生生物保护的应用研究

庞树霞, 韩 印, 马若铭, 刘 博, 丁龙远

上海理工大学管理学院, 上海

收稿日期: 2024年12月15日; 录用日期: 2025年1月8日; 发布日期: 2025年1月16日

摘 要

野生生物是地球环境的重要组成部分, 然而受到非法野生生物贸易的影响, 野生生物的种类和数量正面临着严重的威胁和衰退。为了显著减少非法野生动物贸易, 本研究对美国野生生物保护法规体系进行了分析, 通过多元岭回归模型, 对当地野生生物保护种类数、公众对非法野生动物贸易的认知度、公众对加强执法的支持率、公众每年对野生动物保护组织的捐款数、野生生物保护机构年投资额、野生生物保护政策数量、野生生物保护地数量及面积、当地猎人证数量这9个重要指标进行分析, 评估他们对非法野生动物年贸易量、年度GDP、当地犯罪统计数、当地人口数等4个政府重点关注因素的影响。并采用灰色预测构建了数据驱动下, 野生生物保护的预测模型。最终得到针对性的政策和投资可以在促进社会经济发展的同时, 大大降低非法野生生物贸易规模及当地犯罪率。

关键词

岭回归模型, 数据可视化, 灰色预测, 野生生物保护

Applied Research Based on Data-Driven Wildlife Conservation

Shuxia Pang, Yin Han, Ruoming Ma, Bo Liu, Longyuan Ding

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Dec. 15th, 2024; accepted: Jan. 8th, 2025; published: Jan. 16th, 2025

Abstract

Wildlife is an important part of the Earth's environment, yet wildlife species and populations are facing serious threats and declines due to the impact of illegal wildlife trade. In order to significantly reduce illegal wildlife trade, this study analyzed the US wildlife conservation regulatory system by using multivariate ridge regression modeling for the number of local wildlife conservation species,

public awareness of illegal wildlife trade, public support for increased law enforcement, number of public donations to wildlife conservation organizations per year, annual investment in wildlife conservation organizations, number of wildlife conservation policies, nine important indicators, namely the number and area of wildlife reserves and the number of local hunter permits, were analyzed to assess their impacts on four government-focused factors, including the annual trade in illegal wildlife, the annual GDP, the number of local crime statistics, and the number of local population. And a data-driven, predictive model for wildlife conservation was constructed using gray prediction. The final result is that targeted policies and investments can make significant progress in reducing illegal wildlife trade and crime rates while promoting socioeconomic development.

Keywords

Ridge Regression Model, Data Visualization, Gray Prediction, Wildlife Conservation

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

自《濒危物种法案》(ESA)在美国实施以来,已有超过 2300 种动植物被列为受威胁或濒危物种。虽然 ESA 成功防止了物种灭绝,但只有 2%的物种被移除出名单[1]。Wilcove [2]等人通过量化濒危物种面临的威胁发现,栖息地丧失和外来物种是它们面临的主要威胁。他使用了“管理债务”一词来强调这一问题,表明这是由于政府缺乏主动的栖息地管理投资。因此,本项目将重点关注州野生动物保护区内的野生动物栖息地保护相关数据。

在美国,濒危物种和受威胁物种的保护以及狩猎活动主要发生在私有土地上,导致政府对私有土地的投资和关注更多[3]。统计数据显示,超过 75%的濒危物种依赖私有土地作为其栖息地。在许多方面,由于国家土地私有化制度[4],野生动物保护政策往往难以广泛实施。针对濒危物种的土地使用限制和潜在的民事罚款并没有通过减税和其他市场驱动的激励措施得到缓解[5]。因此,在开发项目时,不同州的野生动物保护政策数量也是一个重要的考虑因素。

对于大多数州野生动物保护机构来说,猎人仍然是野生动物的主要威胁之一。全国范围内,平均 85%的猎人在私有土地上进行狩猎,其中 64%的人仅在私有土地上狩猎[6]。因此,考虑到对受威胁和濒危物种的保护,猎人的数量也将是数据驱动项目开发的一个关键关注点。

2. 数据收集及可视化

2.1. 可视化表达

在构建数据驱动模型前,我们首先收集了众多数据,如表 1 所示。我们选择了权威的官方网站作为数据来源,以确保模型构建和数据分析结果的准确性。

我们对表 1 中的数据进行了可视化的表达,如图 1 所示。

由图 1 我们可以看出,该地区从 2005 年开始,人口数量及当地犯罪数均在缓慢减少,人口数从 4100 万降低至 3600 万左右,犯罪数随着人口数的减少从 22.5 万降低至 17.5 万左右。尽管人口数量减少,该地区居民对非法野生生物贸易的关注比例,及支持打击非法野生生物贸易的比例逐年上升,从 2005 年不足 80%的比例,上升到了 2022 年 85%左右的关注贸易活动比例和支持打击非法野生生物贸易比例。

Table 1. Data sources
表 1. 数据来源

数据描述	数据来源
野生生物保护物种数量	https://www.traffic.org/cn
公众对非法野生生物贸易的认知度	https://www.macrorends.net/states/new-york/population
公众对加强执法的支持率	https://www.criminaljustice.ny.gov/crimnet/ojsa/tableau_index_crime_by_agency.htm
公众对野生生物保护机构捐款总额	https://www.neighborhoodscout.com/about-the-data/crime-rates
野生生物保护机构投资额	
野生生物保护政策数量	
野生生物保护区数量	https://www.statista.com/topics/9382/new-york-city
野生生物保护区域的面积	
当地猎人证数量	

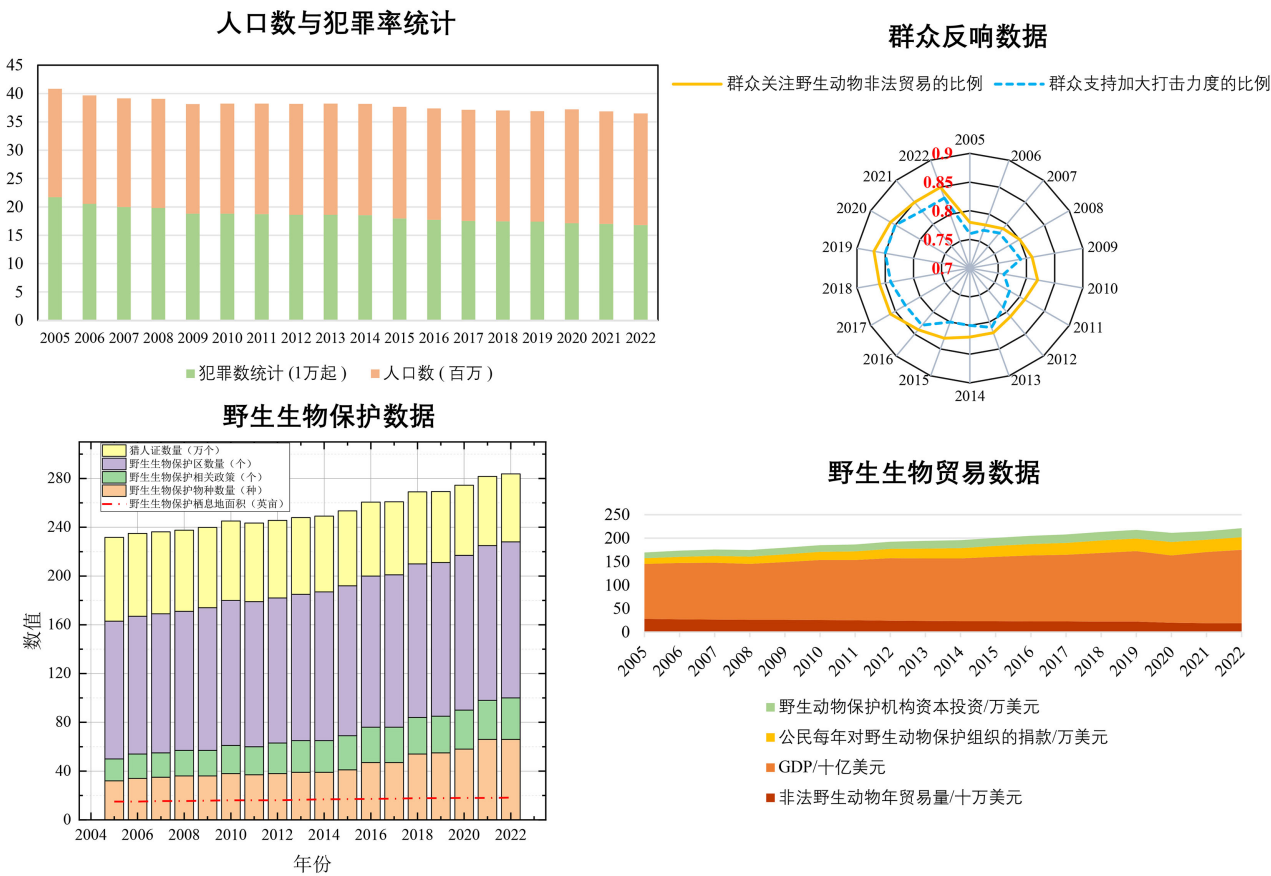


Figure 1. Data visualization
图 1. 数据可视化

当地野生生物保护机构及相关政府单位对野生生物保护的重视程度也在逐年上升，从 2005 年起，规定的野生生物保护物种个数从 38 种增加到了 73 种，相关政府单位颁布的野生生物保护相关政策从 50 条

增加到了 100 条左右, 野生生物保护区数量从 160 个增加到了 240 个左右。然而, 野生生物保护栖息地的总面积并未看出有明显的上升, 这可能与当地自然条件及规划用地区域相对紧缺有关, 由于野生生物出现的区域较为集中, 且不能影响城市周边重要交通枢纽的规划与建设, 因此只能在原有的自然保护区周围, 新增一些面积较小的保护区域, 尽可能的扩大野生生物的活动范围保护区。但是, 我们可以看到, 相关政府机构正在逐年增加发放的猎人证数量, 从 237 万增加到 280 万左右, 这对野生生物的生存环境造成了极大的威胁, 不利于野生生物的保护。

受当地犯罪数降低, 群众关注及支持打击非法野生生物的比例增加, 及野生生物保护机构及政府相关机构的努力, 该地区从 2005 年至 2022 年, 非法野生生物贸易的年贸易量从 281 万美元下降至 188 万美元。此外, 当地居民向野生生物保护机构捐款金额及资本投资金额, 也分别从 12.3 万美元增加到 27.1 万美元及 12 万美元增加到 19 万美元。随着野生生物贸易量的逐年降低, 该地区的非法经济活动受到极大的限制, 该地区的经济状况也呈现向好的趋势, 该地区年生产总值(GDP)从 1168.8 亿美元增加到了 1563 亿美元, 这体现出地区经济发展的繁荣与该区域的社会稳定性息息相关。

2.2. 数据预处理

在数据预处理前, 首先对变量名称进行定义, 表 2 展示了本研究所使用的变量及符号。

Table 2. Variable annotation

表 2. 变量注释

符号	变量描述	单位
V_b	非法野生动物年贸易量	十万美元
GDP	年度 GDP 产出	十亿美元
N_c	犯罪统计数据	百
P	人口数	百万
N_w	保护物种数	每种
R_t	公众对非法野生动物贸易的认知百分比	%
R_a	公众对加强执法的支持百分比	%
V_d	公民每年对野生动物保护组织的捐款	一万美元
V_i	野生动植物保护机构年投资额	一万美元
N_p	野生动植物相关保护政策	每条
N_a	野生生物保护地数量	每个
S_a	野生生物保护地面积	一万英亩
N_h	年发放猎人证数量	一万

接着通过如图 2 所示的流程对收集到的数据进行预处理。

处理数据集中的异常值和噪声值是数据预处理中的关键步骤。首先基于“ 3σ 原则”, 我们考虑数据落在区间 $(-3, +3)$ 内的概率为 0.9974, 而落在这一区间之外的概率小于千分之三。接下来, 我们使用加权平均方法去除数据中存在的噪声并进行平滑处理。由于不同变量具有不同的量纲和显著的数量级差异, 为了解决数据之间的可比性问题, 我们使用标准化方法消除独立变量上的量纲影响。图 3 展示了经过预处理后的数据分布, 显示数据预处理效果显著。

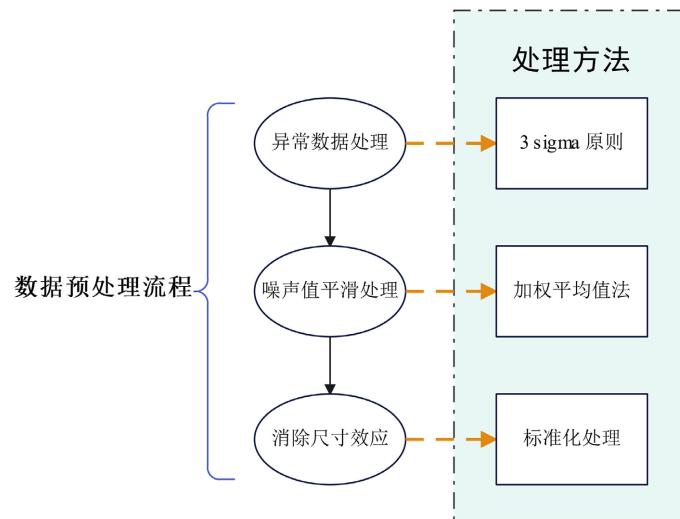


Figure 2. Data preprocessing process
图 2. 数据预处理流程

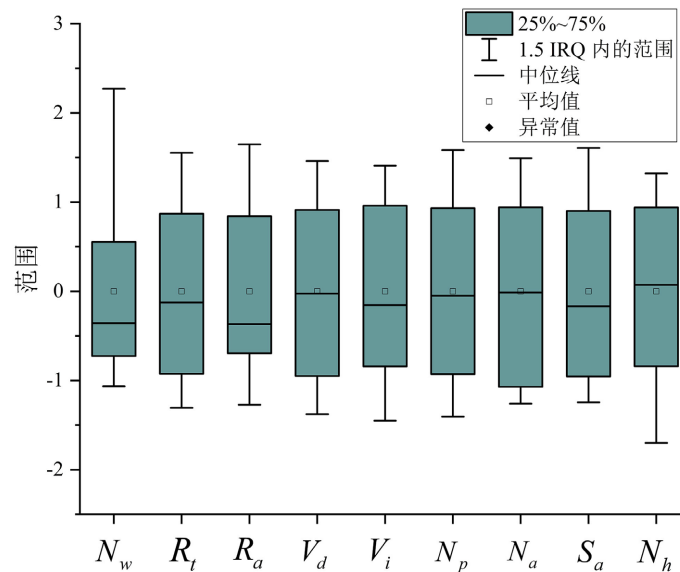


Figure 3. Data box plot distribution
图 3. 数据箱型图分布

3. 多元岭回归模型构建

非法野生生物贸易活动的规模与当地人口数量、生产总值(GDP)及犯罪率都有着重要的关联。因此,本研究构建了四个岭回归模型,以 N_w 、 R_t 、 R_a 、 V_d 、 V_i 、 N_p 、 N_a 、 S_a 和 N_h 作为自变量,评估其对 V_b 、 GDP 、 N_e 和 P 的影响。

岭回归模型是一种专门用于分析共线数据的回归方法。本质上,它是最小二乘估计方法的改进版本。通过牺牲最小二乘法的无偏性,得到更符合实际且更可靠的回归系数。这是以丢失一些信息和降低精度为代价实现的。首先,该模型通过最小化真实值(y_i)与预测值($\hat{y}_i$)之间的残差平方和(RSS)进行训练。

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

其中, y_i 表示实际值, $\hat{y}_i$ 表示预测值。在 RSS 之后, 成本函数加入了一个惩罚项, 用于控制模型系数的 L2 范数, 以控制模型的复杂度:

$$J(\beta) = \text{RSS} + \lambda \sum_{j=0}^n \beta_j^2 = \text{RSS} + \lambda \|\beta\|^2 \quad (2)$$

通过最小化目标函数, 我们可以求解岭回归模型的系数 β_{ridge} 。这种方法与普通最小二乘法的关键区别在于增加了正则化项 λI_n , 其中 I_n 是单位矩阵, λ 是正则化参数:

$$\beta_{\text{ridge}} = (X^T X + \lambda I_n)^{-1} X^T y \quad (3)$$

与普通最小二乘法的预测结果类似, 岭回归的预测结果也是由自变量乘以各自的系数得到的。

$$\hat{y}_i = X \beta_{\text{ridge}} \quad (4)$$

我们使用 RMSE 和 R^2 来评估岭回归的性能。表 3 列出了 MSE 和 R^2 的计算结果。

Table 3. MSE and R^2 calculations

表 3. MSE 和 R^2 的计算结果

模型(用自变量符号表示)	RMSE	R^2
V_b	0.226	0.94722
GDP	9.3	0.958
N_c	62.47	0.926
P	0.028	0.9444

我们发现, 在我们建立的所有四个模型中, 预测误差都在可接受的范围内, R^2 值均超过 0.9。这表明这些模型与数据集的拟合度非常高, 预测能力出色。接近于 1 的 R^2 值强调了模型在解释数据变化方面的有效性, 为决策提供了可靠的支持。

4. 灰色预测模型构建

灰色预测模型由于自身的简洁性、适应性、灵活性等优势[7], 在不同领域的预测问题中都得到了广泛的应用, 因此本研究采用了灰色预测模型来预测所选自变量在未来 5 年内的变化。灰色预测模型(GM(1, 1))是一种处理有限和不完整信息的方法。它是将原始数据累积一次(1-AGO), 生成一个新的数据序列, 减少数据的随机性。随后, 建立一阶微分方程模型进行预测。

首先, 我们假设原始数据序列为:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (5)$$

使用一阶累积生成(1-AGO)生成新序列:

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (6)$$

其中 $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$ 。

接下来, 我们建立灰色微分方程模型, 其一般形式为:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (7)$$

其中, a 和 b 是需要估算的参数, 分别代表发展系数和灰色作用量。

用最小二乘法估计参数 a 和 b 。构建数据矩阵 B 和矢量 Y ，其中 $B=[-Z \ 1]$ ， Z 是相邻均值序列， $Y=X^{(0)}(2:n)$ 代表原始数据序列的最后 $n-1$ 项。根据参数估计公式，我们可以得到 $\hat{a}$ 和 $\hat{b}$ ：

$$\hat{a} = (B^T B)^{-1} B^T Y \tag{8}$$

根据估计参数，使用以下预测公式预测未来 n 年的数值：

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} \tag{9}$$

预测的差分序列可以用下面的公式计算出来：

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{10}$$

通过计算误差与真实值的比率，我们评估了灰色预测的精确度，精确度结果如表 4 所示。利用这种方法进行精度测试，我们发现大多数自变量的预测精度都在 95% 以上。这表明所采用的预测方法有效地反映了自变量的自然趋势，为后续分析和决策提供了可靠的数据支持。

Table 4. Gray prediction accuracy
表 4. 灰色预测精度

预测变量	精度(%)
N_w	95.18
R_t	97.1
R_a	97.09
V_d	95.24
V_i	94.25
N_p	94.49
N_a	96.68
S_a	95.68
N_h	90.65

5. 野生生物保护应用评价

根据灰色预测模型的结果，我们将预测的自变量数据输入岭回归模型，计算出未来五年四个因变量的详细数值。随后，我们设计了一些变量的变化，并制定了一个分为两个阶段的 10 年计划。表 5 列出了该计划的详细变量设置，旨在通过预先确定的干预措施，有效指导和优化长期规划和决策过程。

公众对非法野生动物贸易的关注比例、公众对加大执法力度的支持比例、捐款总额等变量对政府部门控制具有挑战性。因此，我们仍采用灰色预测法模拟的自然变化结果。

对于其他变量，如受保护物种的数量、野生动物保护项目的投资、相关政策的数量、野生动物保护区的数量和规模等，我们设计了主动规划和干预的两阶段计划。政府在这些方面具有更直接的控制和干预能力，可以通过具体的政策措施和项目投资来影响这些变量的发展方向和结果。

该战略旨在通过有计划的干预措施提高野生动物保护的有效性，促进保护区的建立和管理，提高相关政策的数量和质量，并加强对野生动物保护项目的投资。这将加速减少非法野生动植物贸易，同时实现降低犯罪率、增加 GDP 和人口增长的目标。

Table 5. Variable settings
表 5. 变量设置

年	N_w	R_t	R_a	V_d	V_i	N_p	N_a	S_a	N_h
2023	-0.5	0.004614	0.004552	2.721098	25	1	2	2	-1.5
2024	-0.5	0.004584	0.004521	2.588827	25	1	2	2	-1.5
2025	-0.5	0.004553	0.004491	2.462985	25	1	2	2	-1.5
2026	-0.5	0.004523	0.004461	2.34326	25	1	2	2	-1.5
2027	-0.5	0.004493	0.00443	2.229355	25	1	2	2	-1.5
2028	0	0.004463	0.004401	2.120987	20	0.5	1	1.5	-1.8
2029	0	0.004433	0.004371	2.017887	20	0.5	1	1.5	-1.8
2030	0	0.004404	0.004341	1.919799	20	0.5	1	1.5	-1.8
2031	0	0.004374	0.004312	1.826478	20	0.5	1	1.5	-1.8
2032	0	0.004345	0.004283	1.737694	20	0.5	1	1.5	-1.8

从前 5 年人工干预前后对比图(图 4)和整个十年的结果(图 5)显示了实施相关项目计划对主要指标的影响。从图 4 中我们可以看出,在最初的五年中,在保护物种数量、增加野生动物保护项目投资、实施更多政策和管理野生动物保护区等方面的人工干预,与干预前的情况相比,产生了显著的积极变化。

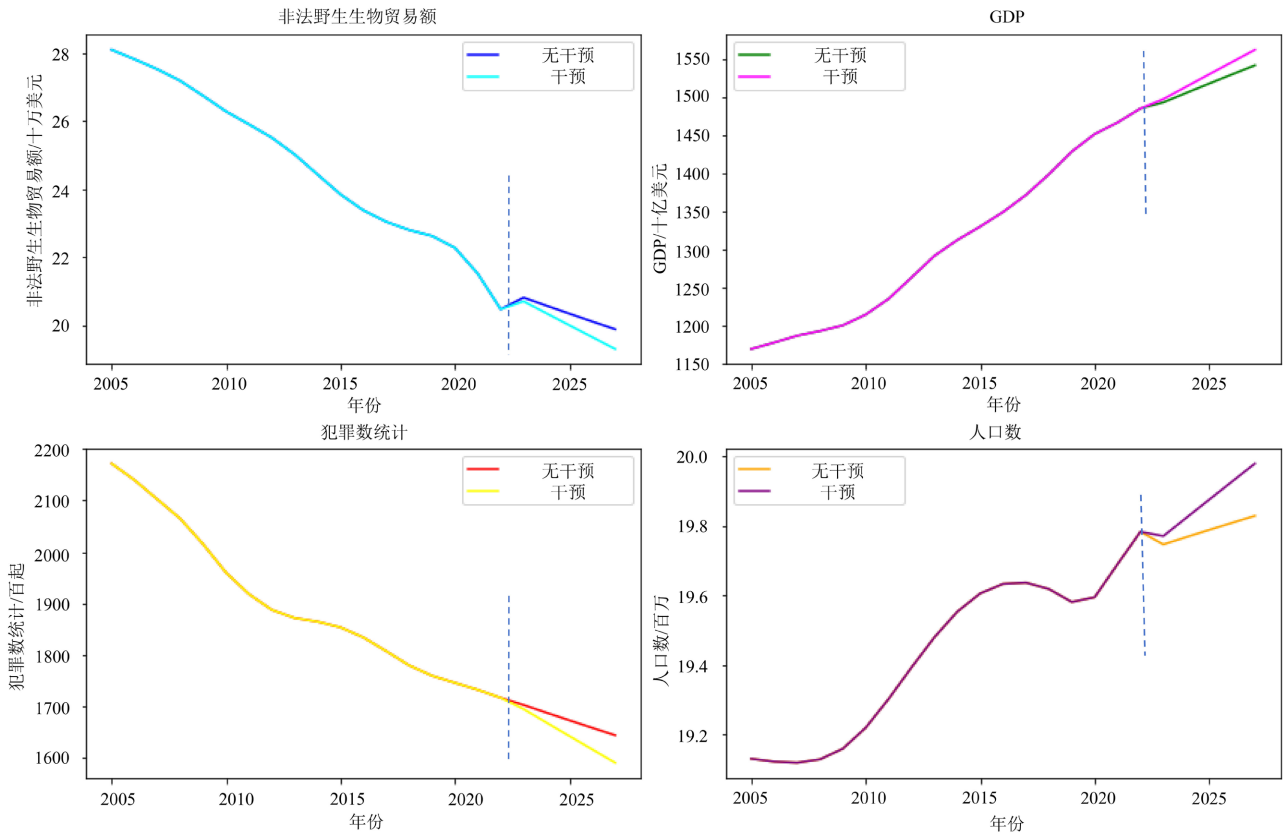


Figure 4. Comparison of project implementation expectations over 5 years
图 4. 5 年项目实施预期对比

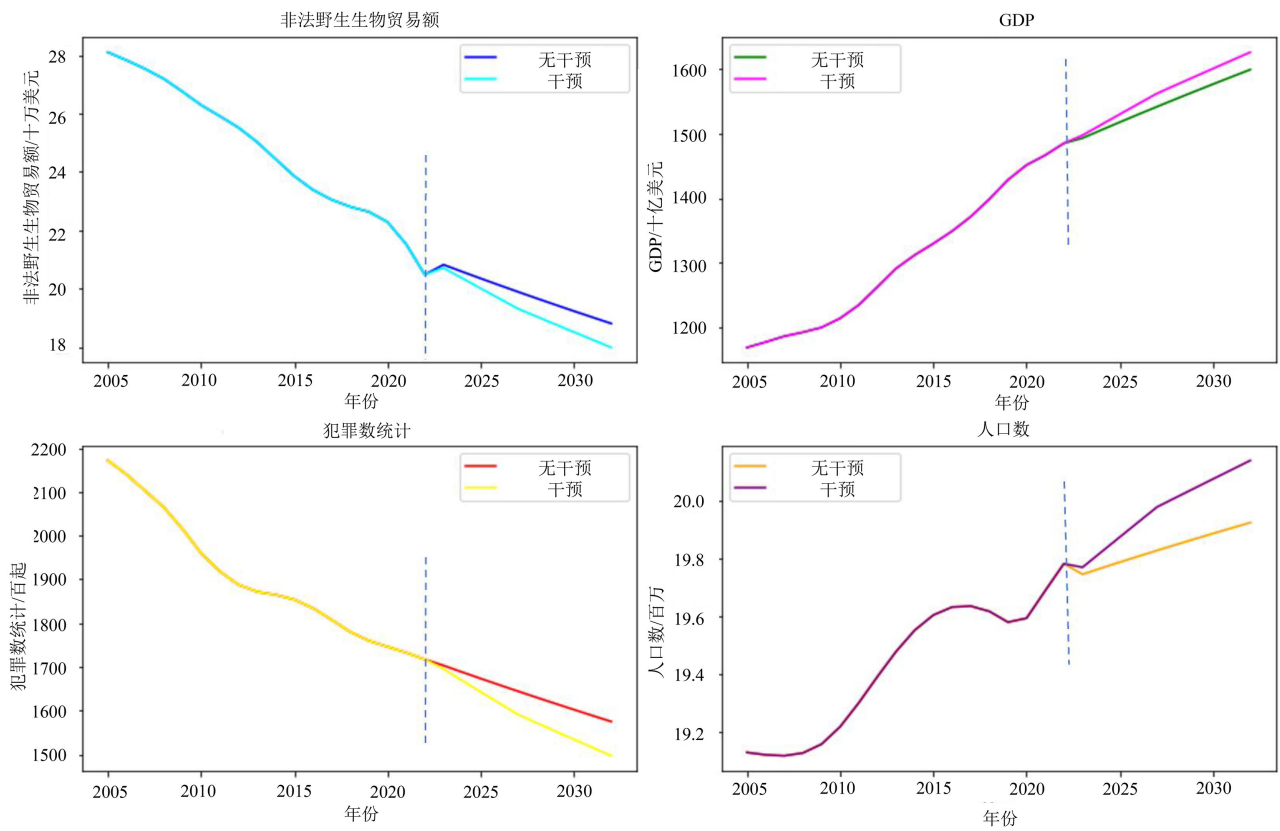


Figure 5. Comparison of project implementation expectations over 10 years
图 5. 10 年项目实施预期对比

图 5 进一步表明，在整个十年中，通过持续实施这些计划，不仅国内生产总值和居住人口大幅增加，而且非法贸易额和犯罪率也显著下降。这表明，有针对性的政策和项目投资可以在促进社会经济发展的同时，在降低非法贸易和犯罪率方面取得重大进展。

这些成果强调了采取像本项目这样的综合战略来解决非法野生生物贸易和社会可持续发展的重要性。通过精心设计和实施项目计划，可以在多个方面产生积极影响，包括促进经济增长、提高居民生活质量、保护野生动物及其栖息地，以及有效减少非法活动和加强社会安全。项目优化效果如表 6 所示。

Table 6. Project optimization effectiveness statistics
表 6. 项目优化效果统计

项目优化效果	V_b	GDP	N_c	P
项目实施第五年与未实施的差值	-0.58	20.45	-53.28	0.15
项目实施后第五年的相关参数变化	2.88%	1.34%	3.21%	0.76%
项目实施五年内的平均变化	0.34	12.29	-30.59	0.087

6. 结论

为了有效减少非法野生生物贸易的规模，我们提出了“多维数据驱动减少非法野生生物贸易策略”项目，考虑了社会发展、生物多样性保护和经济水平等多个维度。

项目基于广泛的历史数据，针对影响非法野生生物贸易的重要因素进行分析。接着根据分析结果，

制定非法野生动物贸易改进策略。通过这些措施，我们制定了一个全面的项目计划，旨在有效减少非法野生动物贸易的规模，保护濒危物种和生态平衡。我们基于岭回归建立了变量间的回归模型，并使用准确率超过 95% 的灰色预测方法预测了未来 5 年的影响。通过对预测结果进行岭回归分析，本研究分析了项目制定的 5 年和 10 年计划的优势。

研究表明，通过适当减少受保护物种的数量，减少狩猎许可证的发放，增加用于野生动物保护项目的投资金额，增加公众捐赠的总金额、增加相关政策的数量、增加野生动物保护区的数量、扩展野生动物保护区的面积，可以实现非法野生动物贸易减少 2.88%，GDP 增长 1.34%，犯罪率降低 3.21%，人口增加 0.76% 的结果。其中，本研究提出适当减少受保护的野生生物物种数量，是一种有利于野生生物保护的措施，这是因为大部分野生生物的价值受到其珍稀度和可获得性的影响，本研究认为当珍稀野生生物的栖息地和种群数量已经得到改善时，将该物种从移出保护物种名单，可以帮助其免于遭受犯罪分子的捕猎，可以更加有效地保护该物种的数量。本研究通过多元岭回归模型和灰色预测模型的分析，也验证了这一观点。

参考文献

- [1] Wikipedia (2013) Quantum Entanglement. https://en.wikipedia.org/wiki/Quantum_entanglement
- [2] The Wildlife Society [TWS] (2017) Endangered Species Act Policy Brief. https://wildlife.org/wp-content/uploads/2014/11/Policy-Brief_ESA_FINAL.pdf
- [3] Wilcove, D.S., Rothstein, D., Dubow, J., Phillips, A. and Losos, E. (1998) Quantifying Threats to Imperiled Species in the United States. *BioScience*, **48**, 607-615. <https://doi.org/10.2307/1313420>
- [4] Morgan, J.J., Rhoden, C.M., White, B. and Riley, S.P. (2019) A State Assessment of Private Lands Wildlife Conservation in the United States. *Wildlife Society Bulletin*, **43**, 328-337. <https://doi.org/10.1002/wsb.997>
- [5] Lockwood, J.A. (1998) The Intent and Implementation of the Endangered Species Act: A Matter of Scale. In: Shogren, J.F., Ed., *Private Property and the Endangered Species Act*, University of Texas Press, 70-91.
- [6] Innes, R., Polasky, S. and Tschirhart, J. (1998) Takings, Compensation and Endangered Species Protection on Private Lands. *Journal of Economic Perspectives*, **12**, 35-52. <https://doi.org/10.1257/jep.12.3.35>
- [7] 何金泽, 熊婧秋, 黎欣茹, 黄婷, 余佳蔓. 灰色预测模型在碳排放预测中的应用[J]. 应用数学进展, 2024, 13(1): 84-90.