Measurement of Black Rhino Conservation Methods

Project Report

Spencer Will, Karnav Patel, Dineth Gunawardena, Vaishnavi Annabhemoju

May 4, 2022

1 Introduction

Animals are endangered because of many human activities and biological activities. This is important due to species diversification, ecosystem impact, etc. Some of the causes for animal endangerment are poaching, invasive species, climate change, and many others. Currently, conservation efforts are being enforced such as: training rangers to deal with poachers, wildlife trade monitoring network (TRAFFIC), and criminal investigation with DNA analysis on animal item trade. Our project is focusing on the black rhino population. We will implement numerical analysis approaches such as interpolation and extrapolation to approximate the population for black rhino.

2 Related Work

There is a substantial amount of research going on in implementing mathematical models to forecast the population of Mathematical and statistical models have become essential tools for predicting how species demographic rates will react to their environment, and thus how they will cope with ever-changing conditions. However, their ability to allow us to see into the future is dependent on the accuracy of the information we use to build them.

According to the researchers at University of California, Riverside [3], they investigated how extinction probability varies over time as a function of species age, time of observation, current geographical range, change in geographical range, climate state, and change in climate state and they claimed that their models have a 70–80 percent probability of correctly forecasting the rank order of extinction risk for a random out-of-sample species pair, implying that determinants of extinction risk have only changed modestly.

Researchers applied a discrete-time survival modeling framework to see how well they can anticipate the likelihood of extinction. The core of this approach was multilevel logistic regression. In order to measure the performance of the model, they have used two distinct contexts: in-sample performance, and out-of-sample predictive performance. In-sample forecasting is a posterior predictive check in which they estimated their model's ability to categorize the data to which it was fitted correctly.However they stated that, in-sample forecasting measures are not the same as understanding their model's ability to forecast future extinctions. To quantify their ability to forecast species' extinction risk, they estimated average out-of-sample forecasting performance using five-fold time-series cross-validation. Cross-validation is a procedure for estimating a model's expected out-of-sample error.

They concluded that past extinction patterns can provide valuable information about which extant species are most threatened with extinction in the near geological future.

3 Problem Description

We are focusing on the black rhino population for this project. Some examples of why the black rhino is an endangered species are increasing poaching of black rhino horns due to poverty, wars hampering conservation, loss in habitat due to human settlement, and many other causes. Also, the current conservation efforts include sub-population field protection, limiting sports and hunting quotas, and improved biological management. We need to understand whether these efforts are enough for conserving the black rhino population. We have chosen the numerical analysis approaches including least squares approximation and linear regression to approximate the future population of the black rhino using the existing population data. Using the different models, we retrieve different accuracy which can provide a metric for how current black rhino conservation efforts are performing.

4 Methodology

To find the percentage population decline for current years and future years, we first needed to find an adequate data source. The best data found came from *Our World in Data*, who collected data from many different reliable sources.

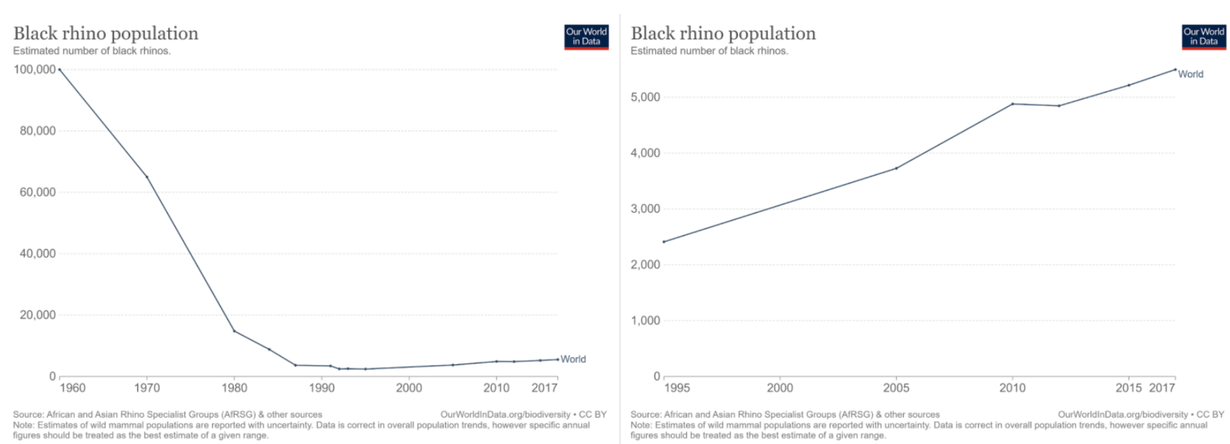


Figure 1: *Black Rhino Population from Our World In Data* Source: [2]

The full set of data can be seen on the left of Fig.1, but we only used the data from after 1995. This is because the data from before 1995 does not have any relation to the conservation measurements we were looking for, so we only included the data points once the population of the Black Rhino was rising. This new data set can be seen on the right of Fig.1.

With this new data set, another problem came up in that there was very little data to work with and there were often large gaps between data points. To ease this, we used Cubic Spline interpolation to interpolate intermittent values to try and create a more accurate model. Cubic Spline interpolation was chosen because it was the most accurate form of interpolation covered in class.

Using this data set, we used three different methods to approximate future values: Linear Least Squares, Polynomial Least Squares, and Linear Regression with Gradient Descent. For each of these methods, we found the approximating function and used this function to calculate the population values from 2023 to 2027. Then, we used these values, as well as ones from before 1995, to calculate the percentage population decline in a 44-year period.

We also calculated two different measures of accuracy on each method. First, we found the Mean Squared Error for each model and then we found the Coefficient of Determination for each model to find their percentage accuracy. The formula for the Coefficient of Determination (R^2) is:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}.$$

4.1 Least Squares

In Linear Least Squares, coefficients to the function $P(x) = a_1 x - a_0$ are found using the equations:

$$a_0 = \frac{\sum_{i=1}^m x_i^2 \sum_{i=1}^m x_i - \sum_{i=1}^m x_i y_i \sum_{i=1}^m x_i}{m(\sum_{i=1}^m x_i^2) - (\sum_{i=1}^m x_i)^2}, a_1 = \frac{m \sum_{i=1}^m x_i y_i - \sum_{i=1}^m x_i \sum_{i=1}^m x_i}{m(\sum_{i=1}^m x_i^2) - (\sum_{i=1}^m x_i)^2}.$$

More generally, Polynomial Least Squares can be used to find the best fit polynomials of higher order than the linear case. For a polynomial with n degree, Polynomial Least Squares gives us $n+1$ equations and $n+1$ unknown coefficients to solve for. For each $j = 0, 1, ..., n$, we have:

$$\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j.$$

Using the coefficients, we are given a best fit polynomial of $P_n(x) = a_n x^n + \cdots a_1 x + a_0$ for the given set of data. We can use this polynomial to find the error for already existing points and interpolating points between them. We can also use the trend of the polynomial to predict future values with reasonable accuracy. For this project, we used two cases for Least Squares: Linear and Quadratic.

Figure 2: *Linear and Polynomial Least Square Plots*

The results for each method can be found in the next section.

4.2 Linear Regression with Gradient Descent

In linear regression, the model calculates the cost function which measures the root Mean Squared error between the predicted value and true value. The model targets to minimize the cost function.

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

To minimize the cost function, the model tries to have the best value of θ_0 and θ_1, which are the slope and intercepts of the line equation. Initially model selects θ_0 and θ_1 values randomly and then iteratively update these values in order to minimize the cost function until it reaches the minimum. A graph is drawn for the cost function as a function of parameter estimates. Taking the derivative of the cost function, moves downward towards the pits in the graph to find the minimum value.

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

Gradient Descent step-downs the cost function in the direction of the steepest descent. The size of each step is determined by parameter α known as Learning Rate.

$$\theta_j = \theta_j - \frac{\alpha}{m}\sum_{i=1}^{m}[(h_\theta(x_i) - y_i)x_i]$$

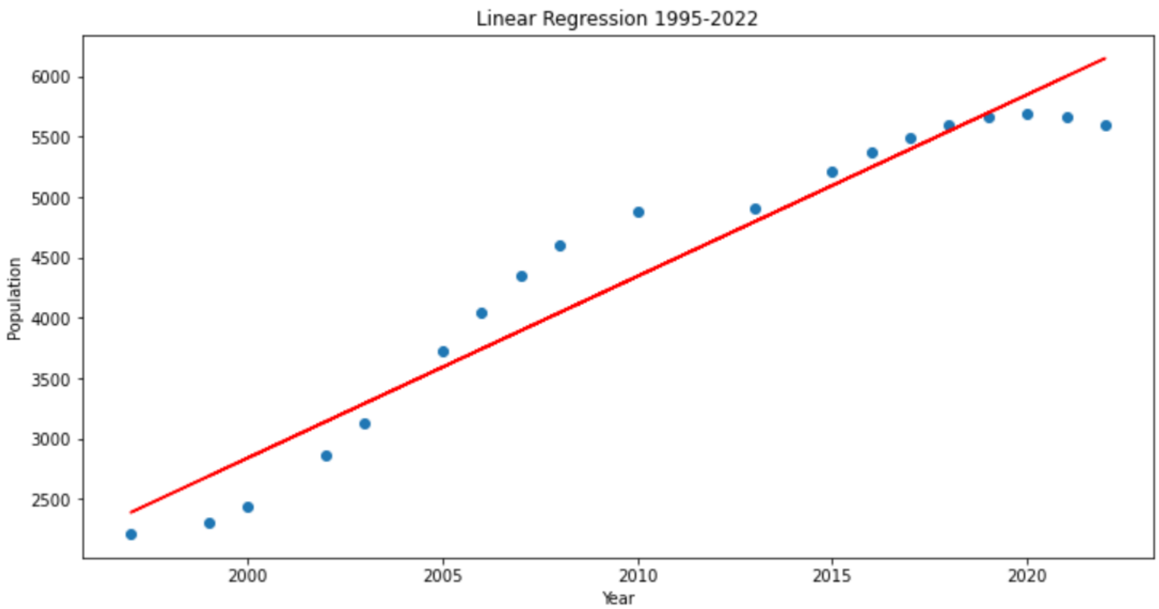


Figure 3: *Linear Regression 1995-2022*

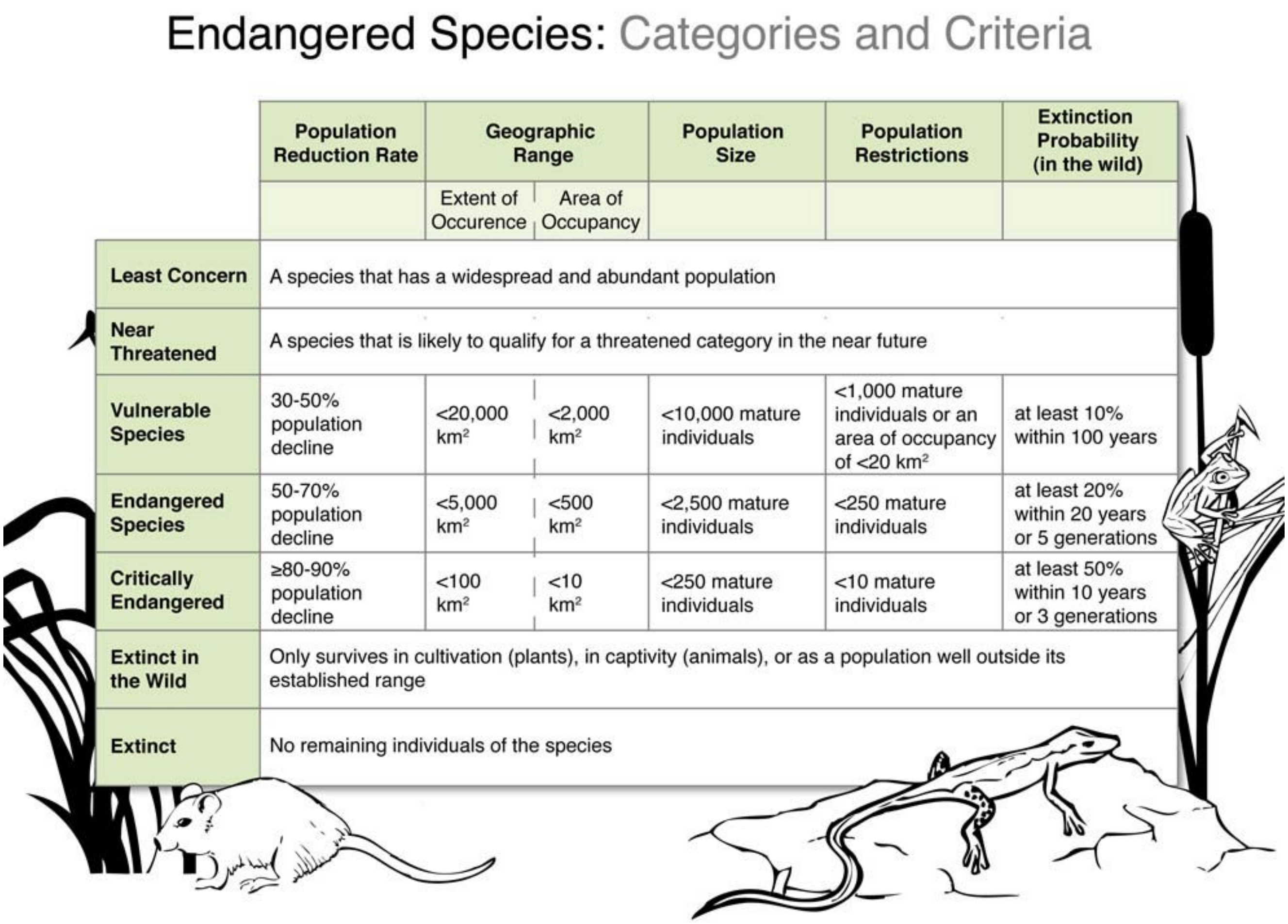
Using each method, we found the error and accuracy to be as follows:

Model	Mean Squared Error	Accuracy
Linear	95902.58	94.06%
Quadratic	63538.61	96.06%
Linear Regression	97352.49	93.105%

Table 1: *Results of each model*

5 Result Analysis

Endangered species are classified in the following criteria described in [1, Fig. 4]. Many of the criteria shown in the figure/table isn't monitored and available to the public over the past 40 years. This includes geographic range of rhinos, population size of mature individuals, population restrictions and the extinction probability. It is difficult to monitor data over a period of time for these criteria and some of this information may be kept from the public due to safety purposes. The data that is most available on black rhinos is their overall population which our project uses to monitor the first criteria (Population Reduction Rate) in [1, Fig 4]. Additional criteria from this figure would also incorporate heavy computational load in our numerical analysis models and were left out due to the availability of data and time constraint of the project.



Endangered Species: Categories and Criteria

	Population Reduction Rate	Geographic Range: Extent of Occurence	Geographic Range: Area of Occupancy	Population Size	Population Restrictions	Extinction Probability (in the wild)
Least Concern	A species that has a widespread and abundant population					
Near Threatened	A species that is likely to qualify for a threatened category in the near future					
Vulnerable Species	30-50% population decline	<20,000 km²	<2,000 km²	<10,000 mature individuals	<1,000 mature individuals or an area of occupancy of <20 km²	at least 10% within 100 years
Endangered Species	50-70% population decline	<5,000 km²	<500 km²	<2,500 mature individuals	<250 mature individuals	at least 20% within 20 years or 5 generations
Critically Endangered	≥80-90% population decline	<100 km²	<10 km²	<250 mature individuals	<10 mature individuals	at least 50% within 10 years or 3 generations
Extinct in the Wild	Only survives in cultivation (plants), in captivity (animals), or as a population well outside its established range					
Extinct	No remaining individuals of the species					

Figure 4: *Species extinction categories and criteria.* Source: [1]

Currently, the black rhino has a population decline of about 72.16% over the past 3 generations (44 years) of rhinos. The population reduction rate criteria alone would currently classify black rhinos between the endangered and critically endangered species categories based on [1, Fig. 4]. Although the black rhino is currently listed as critically endangered by the International Union for Conservation of Nature (IUCN), this classification takes into account of the other criteria in [1, Fig. 4] which we aren't able to quantify or have access to.

As the recent trend of black rhino population in the past 20 years has been increasing, the black rhino population reduction rate has been decreasing. With extrapolated population values, our models predict the year at which black rhinos can be classified from endangered/-critically endangered to a vulnerable species. This calculation was done by using the following equation:

$$VulnerableYear \rightarrow \frac{44thYearPopulation - CurrentYearPopulation}{44thYearPopulation} \leq 50\%$$

44thYearPopulation is the population of black rhinos from 44 years ago and CurrentYearPopulaiton is the population of black rhinos using the current year. This calculation is reiterated for the extrapolated population values within the next 5 years to determine at which year the black rhino population reduction over 3 generations (44 years) is less than 50%. At this year (VulnerableYear), the population is classified as vulnerable as indicated in [1, Fig. 4].

Each model's approximated VulnerableYear and accuracy are summarized in Table 2

	VulnerableYear	Reduction Rate	MSE	Accuracy
Linear Least Squares	2025	48.2%	95902.58	94.06%
Polynomial Least Squares	2026	47.2%	63538.61	96.06%
Linear Regression	2025	48.8%	97352.49	93.11%

Table 2: *Summary of each model's results including approximation and accuracy*

Table 2 shows that Polynomial Least Squares had the highest accuracy and MSE with a reduction rate of 47.2% by the year 2026. By 2025, the other two models approximated the black rhino as a vulnerable species with lower accuracy. Higher weight is given to the Polynomial Least Squares model due to its higher accuracy so our models collectively conclude the black rhino as a vulnerable species around the year 2026.

6 Conclusion

Our project's approximation of 2026 is within the next 5 years of the current year 2022. Given that the past 62 years of black rhino population data shows the black rhino as critically endangered and almost extinct in the wild, our project measures the current conservation efforts for the black rhino as doing very well. Although our project doesn't consider some of the many factors that determine species extinction classification (geographic range, mature population size, etc, this is due to the availability of data to quantify the other criteria and use them in numerical analysis models. Furthermore, the overall black rhino population data can provide a general measure of the current conservation efforts and its models can be later adapted based on the actual future black rhino populations.

Additional analysis can be conducted on black rhino population data such as numerical differentiation which takes into account of the black rhino population growth (slope) and how that changes over time.

More work can also be done to include the other factors in determining a species extinction class. This could include multiple linear regression or multivariate Lagrange interpolation, which quantify additional characteristics of black rhino extinction (invasive species, habitat, human disturbance) and creates an approximation model based on all of the considered characteristics. These characteristics would incorporate more of the complex causes and effects in ecological systems which factor into a species extinction.

The mentioned future work could effectively monitor/measure the efforts of species extinction classification accurately. Having an accurate metric for animal conservation efforts allows practitioners to focus their needs of conservation for a specific species which can save a lot of time, funding, and lives of animals truly in danger. There needs to be more data available for these models to predict extinction classification, which can lead to more trade monitoring, trained rangers, and community engagement for a species in danger of extinction.

7 Contributions

Spencer: Created Least Squares models: Linear and Quadratic, computed the errors and accuracy of these models, completed the Methodology section for the slides and the report.

Vaishnavi: Created Linear Regression model, interpolated values for the data set using cubic spline method, related work section, completed slides and report.

Karnav: Research on project topic, introduction and problem description sections (presentation and paper)

Dineth: Research on project topic, result analysis section and conclusion section (presentation and paper), created slides for introduction and problem description sections.

References

[1] National Geographic Society, "Endangered species categories and criteria," National Geographic Society, 09-Nov-2012. [Online]. Available: https://www.nationalgeographic.org/media/endangered/#: :text=The%20four%20categories%20of%20endangered,and%20extinct%20in%20the%20wild.

[2] "Black Rhino Population," Our World in Data. [Online]. Available: https://ourworldindata.org/grapher/black-rhinos. [Accessed: 03-May-2022].

[3] Smits Peter and Finnegan Seth 2019How predictable is extinction? Forecasting species survival at million-year timescalesPhil. Trans. R. Soc. B3742019039220190392 http://doi.org/10.1098/rstb.2019.0392