# DS2500 Final Project: Biodiversity Decline and Species Extinction

Samara Shankar, Maygala Selvisudhakar, and Sydney Howard

shankar.sa@northeastern.edu, selvisudhakar.m@northeastern.edu, howard.syd@northeastern.edu

*Problem Statement and Background*

Biodiversity is the measure of the immense variety of life on earth. It is comprised of the total number of global species, genetic variation, ecosystem interactions, and regional population changes encompassing all living organisms. Biodiversity is essential to maintaining healthy ecosystems that provide critical services for human survival, including clean air and water, food sources, pollination, and climate regulation, ultimately contributing to human well-being and a stable environment. Without species diversity, many vital ecosystem functions would be dismantled. Biodiversity loss is one of the many pressing issues emerging with the current climate crisis, primarily due to human activity. We investigated the alarming decline in global biodiversity, as highlighted by the Living Planet Index (LPI), which tracks biodiversity trends in species of vertebrate populations. This project explores the changes in species variety and population sizes over the last few decades by comparing each region's LPI, rate of extinction risk, and the protected terrestrial biodiversity shares to better understand the current efforts to stabilize ecosystems. By examining the LPI since 1970, we will quantify and assess the current rates of biodiversity decline, identify areas that require the most conservation efforts, and understand the ways these declines impact the global ecosystem.

*Introduction to the Data*

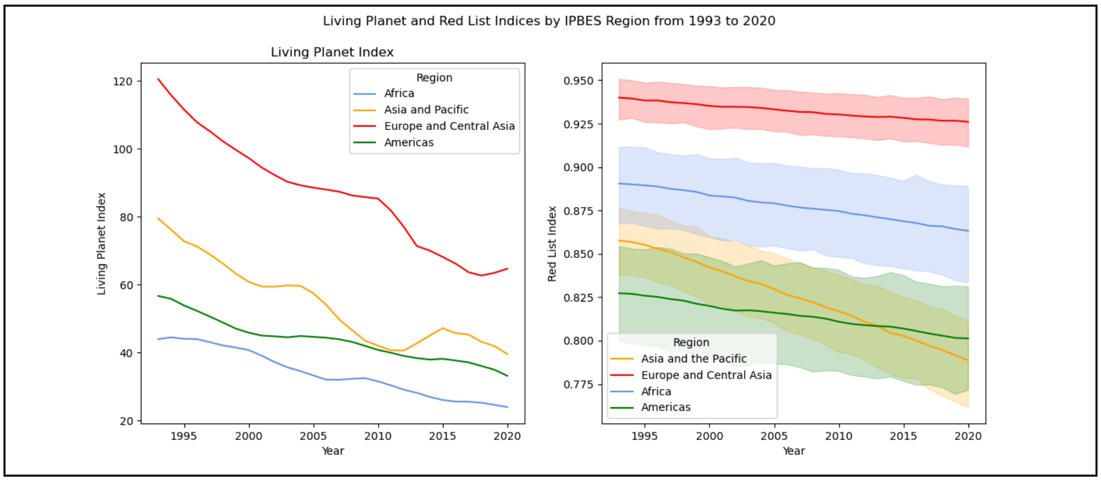
As mentioned, the Living Planet Index is a collection of the average change in population sizes for extensive vertebrate animal populations. The index value measures the abundance change in 34,836 populations across 5,495 native species relative to the first year of collection, where 1970 is 100%. It compiles its data from various sources from 1970 to 2020, including published scientific articles, online databases, and government reports. This collection process guarantees a broad scope of data, though it still has a handful of shortcomings. The index includes only vertebrates (mammals, birds, fish, reptiles, and amphibians) and excludes other critical groups like insects, corals, fungi, and plants. Its focus on vertebrates also limits its representativeness of global ecosystem biodiversity, as it fails to include vital non-vertebrate species critical to ecosystems. The dataset has few privacy or ethical concerns as it relies on publicly available and published ecological data. However, there are some ethical challenges in ensuring the inclusion of underrepresented regions and species groups to avoid skewing global biodiversity trends. Also, though the LPI contains an extensive sample of data globally, some regions must be represented more. Areas located around the tropics have less comprehensive data than more well-studied regions. The tropics also have substantially higher rates of species richness due to the warmer climate sustaining more life, which could further skew the data. Furthermore, Our World In Data highlights that the dataset is sensitive to outliers, meaning a few extreme values pulled from specific regions or species can alter the accuracy of the results. Despite these drawbacks, the LPI is the most comprehensive measurement currently possible and is available to encompass the areas of biodiversity loss that have been studied.

Using the Living Planet Index as a baseline, we compared a significant amount of the LPI data to other biodiversity measurements to understand correlation and identify areas of interest. The first of these datasets is the Red List Index, also pulled from Our World In Data. With entries from 1993 to 2024, the RLI demonstrates trends in overall extinction risk for species groups. It is an index between 0 and 1. A value of 1 indicates no current extinction risk to any included species, while a value of 0 would mean that all included species are extinct. This data estimates extinction risk estimates for mammals, birds, cycads, amphibians, and corals, unlike the vertebrates only LPI, which will be kept in mind upon comparison. The national and regional Red List Indices are calculated by weighting the fraction of each species' distribution occurring within them. It is important to note that the species collected here are likely different from the populations of species studied in the LPI, so our main comparison will be to see how extinction rates correlate to regions with high or low biodiversity loss. The second dataset used is the Share of terrestrial Key Biodiversity Areas that are protected. This data is from 2000 to 2023 and illustrates the proportion of terrestrial Key Biodiversity Areas (KBAs) that are covered by designated protected areas. A KBA is a site that significantly contributes to the long-term persistence of biodiversity in terrestrial, freshwater, and marine environments. The higher percentage means that more of the area is protected.

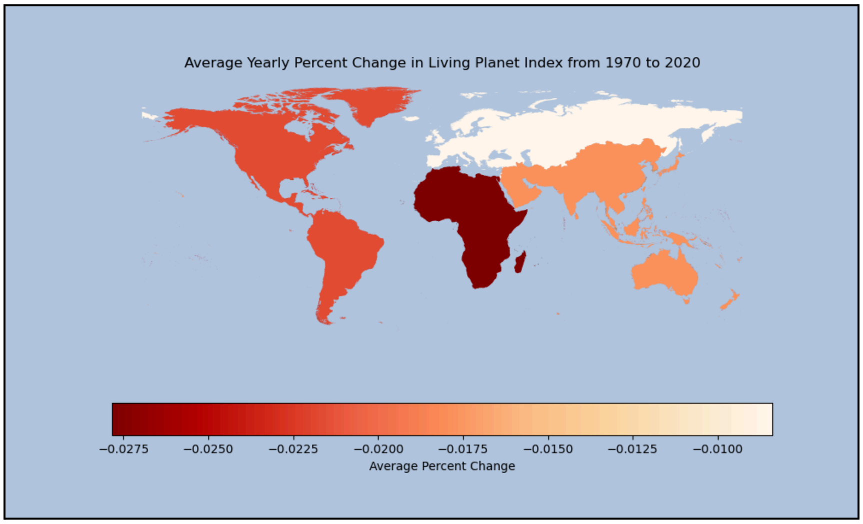
Rather than by region like the LPI, the Red List Index and the Share of terrestrial Key Biodiversity Areas that are protected data are organized by country, which is an issue addressed during the code's cleaning phase. The region boundaries used in the LPI are defined by the IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services), an intergovernmental organization established by the United Nations. To add a variable to the Red List and KBA datasets representing the IPBES region, we sourced a dataset of countries and their regions and subregions from the IPBES website. This allowed us to add a new variable to both data frames that mapped the IPBES region based on country code.

*Data Science Approaches*

The differences in their dataset's structures had to be dealt with to compare the Red List and Living Planet indices. The Living Planet index was aggregated across world regions, while the Red List was broken down by country. To standardize these metrics, country-level Red List values were averaged within IPBES regional boundaries for each year so that their respective line plots would have consistent visuals and the content could be compared well. Data was further filtered to ensure regional line colors stayed consistent and that only years that had data for both indices were displayed. By plotting these side-by-side, we could observe how trends in each index mirrored or diverged from the other. As both indices refer to wildlife, looking at wildlife conservation from two different angles simultaneously provides the most accurate, well-rounded understanding of how each region handles biodiversity.

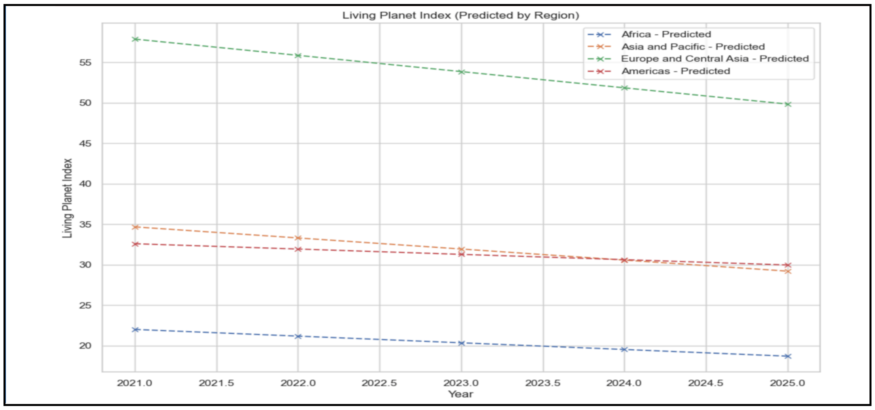



To further investigate the Living Planet Index, the change in value was investigated to demonstrate, over the past 50 years of LPI's use, which regions have done the best at keeping the index high and which have been dropping rapidly. To perform this analysis, we extracted each year, starting in 1971, and found the difference between that year and the one previous. Then, we averaged those values to understand what a typical year's change in LPI looks like. A geospatial dataset was introduced to plot these on a choropleth map, creating a visual illustration of the yearly changes. I was unable to find a shapefile of IPBES regions, so a lot of exploration had to be done with a generic shapefile of countries found off the internet. A similar merging technique from the original region-country-code dataset was employed to add the IPBES region of each country to the dataset. Country borders within each region were then dissolved to create a dataset of IPBES regions and their geometry. This geometry was mapped to the living planet index data, which allowed for a geographic display of the calculated average change in LPI.

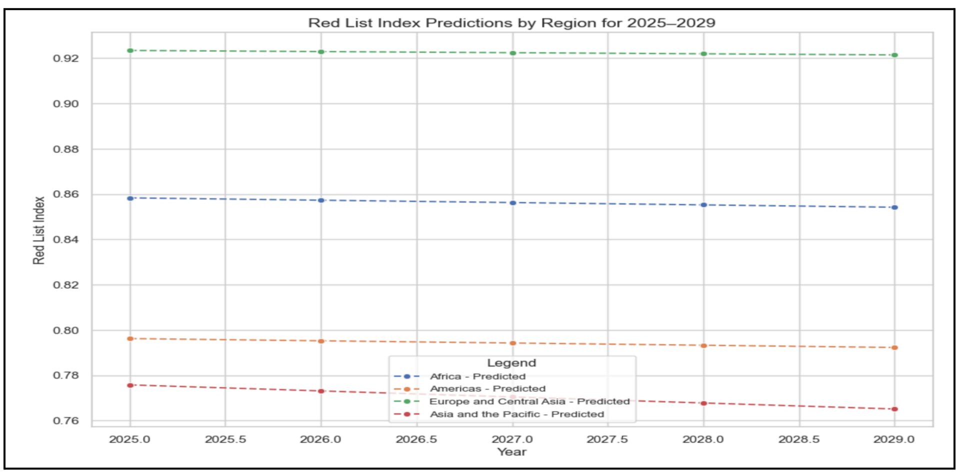



Dissolving the geometries across IPBES regions was the most challenging thing that had to be done regarding spatial data, as there were quite a few challenges that had to be troubleshot with help, and the solutions are not very intuitive. The Red List Index was displayed by country with choropleth color grading determined by the aggregated mean Red List Index across years in the dataset. The plotting of the proportion of protected KBAs across the world by country was similarly simple, but the aggregation in this case was performed by calculating the difference between the first year this was tracked (2000) and the most recent (2023), to highlight which countries have been focusing on improvement over the few decades.

Another way the Living Planet Index was investigated was using a linear regression model. The linear regression model examines the population trends from each of the given regions from 1995 to 2020. Using a linear regression model, the average decline in population size for recorded vertebrate each region is predicted for the next five years. The model given displays the normalized Living Planet Index prediction for each future year with no recorded data.




A linear regression model was also used as another way to analyze the red list index. Since the data for the red list index goes up to 2024 unlike the living planet which only goes up to 2020, the predictions were made for the years 2025 go 2029. The data shows predictions for how close species in a given region are to going extinct.




*Results and Conclusions*

Overall, biodiversity has been decreasing globally since 1970. According to the Living Planet Index, the size of animal populations for which data is available has declined by 73%, on average. The decline for some populations is much larger; for some, it's

Living Planet Index Average Percent Change From 1970-2020

| Region/Group | Relative Change % (1970 = 100%) |
|---|---|
| Africa | -71% |
| Asia and Pacific | -57% |
| Europe and Central Asia | -33% |
| North America | -37% |
| World | -73% |

much smaller. The average of the central, upper, and lower LPI estimates shows Africa has seen the most significant biodiversity loss. At the same time, Europe and Central Asia have seen the least.

We are aware that data for the linear regression model utilizes extrapolation, which is not always the best indicator for future data. However, we chose to use it as a preliminary way of making predictions. All regions are projected to experience a steady decline in LPI through 2025. Europe & Central Asia consistently have the highest LPI.  This general decline proves the significant loss of biodiversity happening on average across the globe as the population size and species richness of the studied regions decrease. The declining LPI is likely attributed to a variety of factors, such as climate change and habitat destruction. This project further establishes the importance of interdisciplinary collaboration to analyze biodiversity trends and preserve global ecosystems.

Based on another linear regression model, the predicted Red List Index values for each The data above shows that all regions are predicted to have a decreasing RLI over time. This means the rate at which extinction is taking place in these regions is worsening along with biodiversity health. According to the predictions, Asia and the Pacific are experiencing the greatest decline in the Red List Index.

The change in proportion of protected KBAs can be used as a metric to understand how each country has introduced policies to maintain biodiversity. We observed that America and Africa have consistently had the two lowest red list and living planet indices over the past few years, but their differing commitment to change informs us of where policy reform is most important. Despite being lower overall in Living Planet Index, and recently dropping to lowest in Red List as well, several countries within Africa have significantly increased their proportion of

KBAs protected over the past few years, indicating that biodiversity is a priority for these areas. In the Americas, on the other hand, very few countries (particularly in the Caribbean) have had any substantial increase in their amount of protected KBAs. These findings highlight the varying levels of commitment to biodiversity protection across regions, emphasizing the need for targeted policy reforms in the Americas. By understanding these disparities, efforts can be better directed to support regions like the Americas in increasing their protected KBAs and addressing biodiversity loss more effectively.

### *Future Work*

To build on the current scope of our project, future work could concentrate on specific regions or species to identify those with the highest risk of biodiversity loss. We explore in detail the correlations between extinction rate and loss of biodiversity but do not identify the majority of the causes or reasons for why these rates occur. New prediction models could be used to integrate climate projections, land-use changes, and further conservation strategies to identify actionable areas. Habitat destruction, deforestation, and fossil fuel emissions are all fascinating areas to explore further regarding their links to the environment. Additional research could also examine current programs in place to conserve biodiversity to understand the areas of improvement and regions with successful programs. Biodiversity loss, including socioeconomic impacts, impacts the ecosystem and the greater global society. Examining sustainable practices and communities could also provide more insight into the ways to prevent these extreme biodiversity losses.

With the series of predictors and maps we made exploring the past data, we can classify which countries and regions are doing the best in terms of environmental sustainability and examine their policies. By exploring the data we already have, we can see what works for protecting our wildlife and environment and let that inform future policy-making decisions.

*Sources*

Our World in Data. (2024). *Data page: Share of terrestrial Key Biodiversity Areas that are protected*. Data adapted from BirdLife International, IUCN, and UNEP-WCMC. Retrieved November 27, 2024, from https://ourworldindata.org/grapher/protected-terrestrial-biodiversity-sites

Our World in Data. (2024). *World region map definitions*. Retrieved November 27, 2024, from https://ourworldindata.org/world-region-map-definitions

Ritchie, H., & Roser, M. (2024). *Biodiversity*. Our World in Data. Retrieved November 27, 2024, from https://ourworldindata.org/biodiversity

Ritchie, H., & Roser, M. (2024). *Red List Index*. Our World in Data. Retrieved November 27, 2024, from https://ourworldindata.org/grapher/red-list-index?tab=table

IPBES Regions and Sub-Regions. (2021). *Zenodo (CERN European Organization for Nuclear Research)*. Retrieved November 27, 2024, from https://doi.org/10.5281/zenodo.5719431

World Administrative Boundaries - Countries and Territories. (2019). Retrieved December 3, 2024, from Opendatasoft.com website: https://public.opendatasoft.com/explore/dataset/world-administrative-boundaries/export/?flg=en-us