Agricultural Resilience and Biodiversity in Southern Saskatchewan

Author: Jocelyn Pender

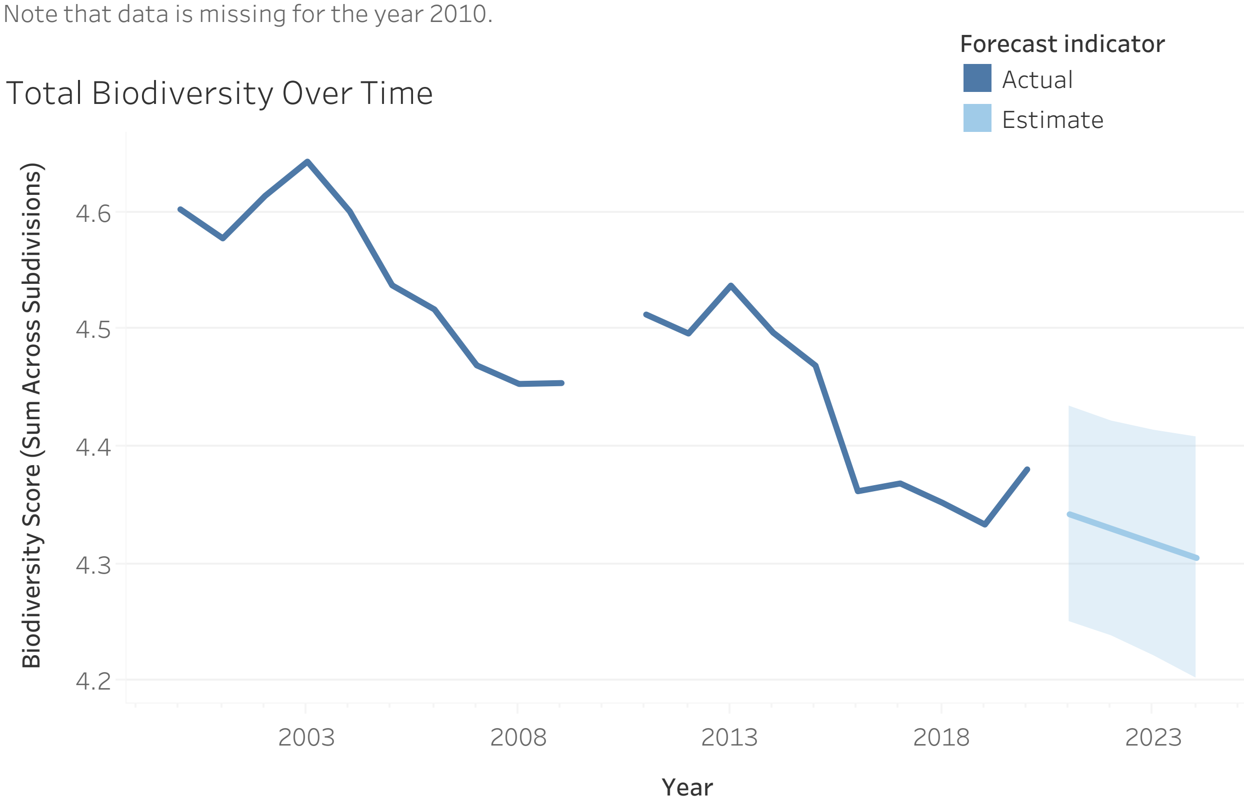
FCC wishes to understand biodiversity and its relationship to land use in southern Saskatchewan. To support this goal, I analyzed a land use dataset containing grassland, wetland, forest and cropland area estimates and biodiversity scores from 10 subdivisions measured over 21 years (2020-2020 inclusive).
To view an interactive version of this report, click here or navigate to https://public.tableau.com/app/profile/jocelyn.pender/viz/PENDERJOCELYN-DataScientist1/IntroductionandKeyFinding in your browser.

Note that dataset descriptive statistics can be found in the associated Jupyter notebook stored on GitHub here:
https://github.com/jocelynpender/technical-assessment/blob/main/notebooks/clean_data.ipynb

Results

Key Finding

My analysis shows that **biodiversity is declining** across the 10 agricultural subdivisions assessed. Additionally, a forecast of future biodiversity scores (exponential smoothing, multiplicative model) demonstrates that if the dataset's patterns reflect future trends, biodiversity will likely continue to decline (-0.9%, 4 years forecasted). However, the model quality is poor, indicating that future data collection is required to bolster the trend forecast.
Note that data is missing for the year 2010.

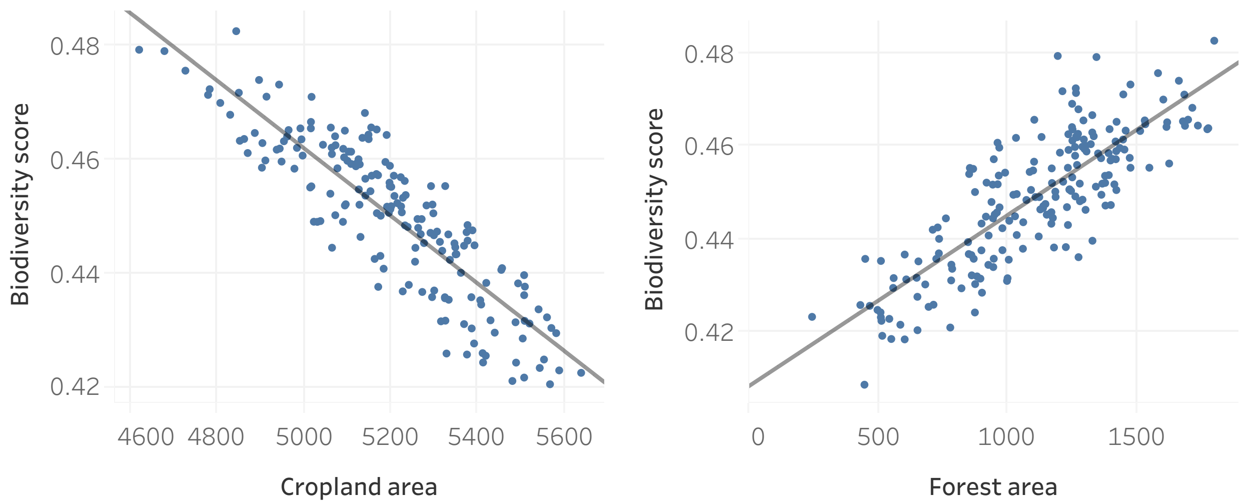


Possible Cause

As the area of cropland increases within a subdivision, the biodiversity score of the subdivision decreases. This is demonstrated by a strong, <u>statistically significant</u> **negative linear relationship** (R squared: 0.76, P-value <0.0001, Pearson's correlation coefficient: -0.87) between biodiversity and cropland area in the 10 subdivisions assessed.

Additionally, as forest cover within a subdivision increases, the biodiversity score of the subdivision increases. This is seen by a strong, <u>statistically significant</u> **positive linear relationship** between biodiversity and forest area (R squared: 0.67, P-value <0.0001, Pearson's correlation coefficient: 0.82).

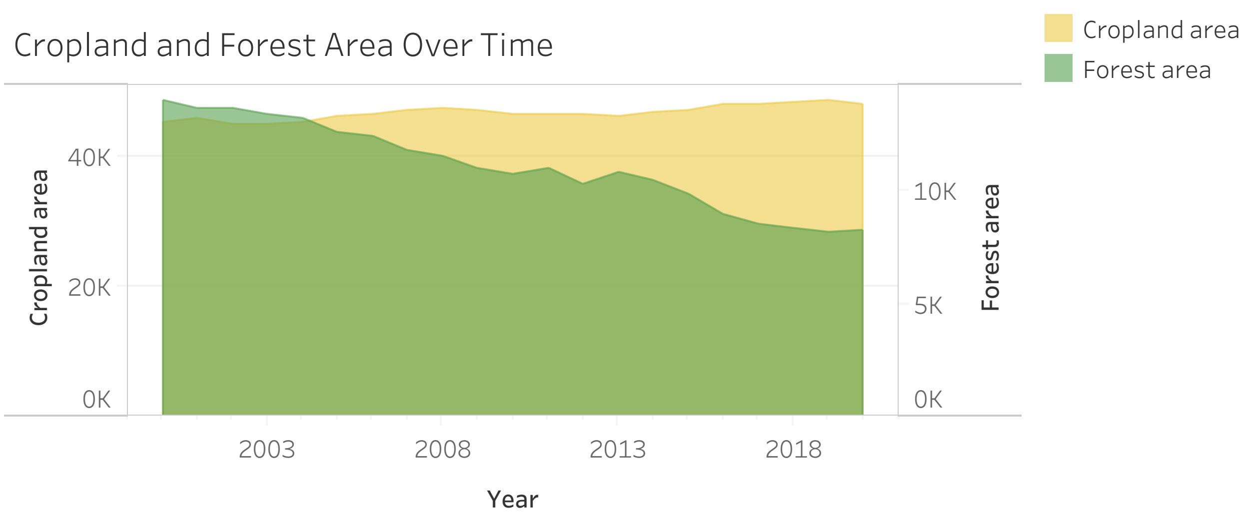
Biodiversity by Cropland Area



Importantly, I found that across the 10 subdivisions in southern Saskatchewan, the area of cropland is increasing whereas the area of forest is decreasing over time. This suggests that land use is changing and specifically, that forests are being converted into cropland (2000-2020).

<u>Given the relationship between biodiversity and cropland & forest cover seen above, the trend of decreasing forest cover may have negatively impacted biodiversity over time.</u>

Cropland and Forest Area Over Time



Recommendation

By prioritizing the conservation of **forest area** across the 10 subdivisions assessed, biodiversity scores would be most effectively stabilized or increased in southern Saskatchewan.

By conserving forest areas, Saskatchewan can conserve biodiversity and ensure agricultural resilience for years to come.

Saskatchewan land planners should consider incentivizing the conversion of other land types (except forest) to cropland, *if* it is required.

Limitations and Assumptions

My analysis assumes that biodiversity score data effectively captures and quantifies biodiversity (species richness, species diversity) and the associated resilience it supplies to agriculture in southern Saskatchewan.
I assume that cropland (%), grassland (%), forest (%) and wetland (%) cover sum to 100%.
Additionally, I assume that missing data (see next page) does not significantly impact the findings and recommendations provided in this report.
Lastly, biodiversity, cropland cover and forest cover changes over time differ between subdivisions. However, I assume that variability across subdivisions does not significantly impact the key findings and recommendations presented. See subdivisional differences visualized below:

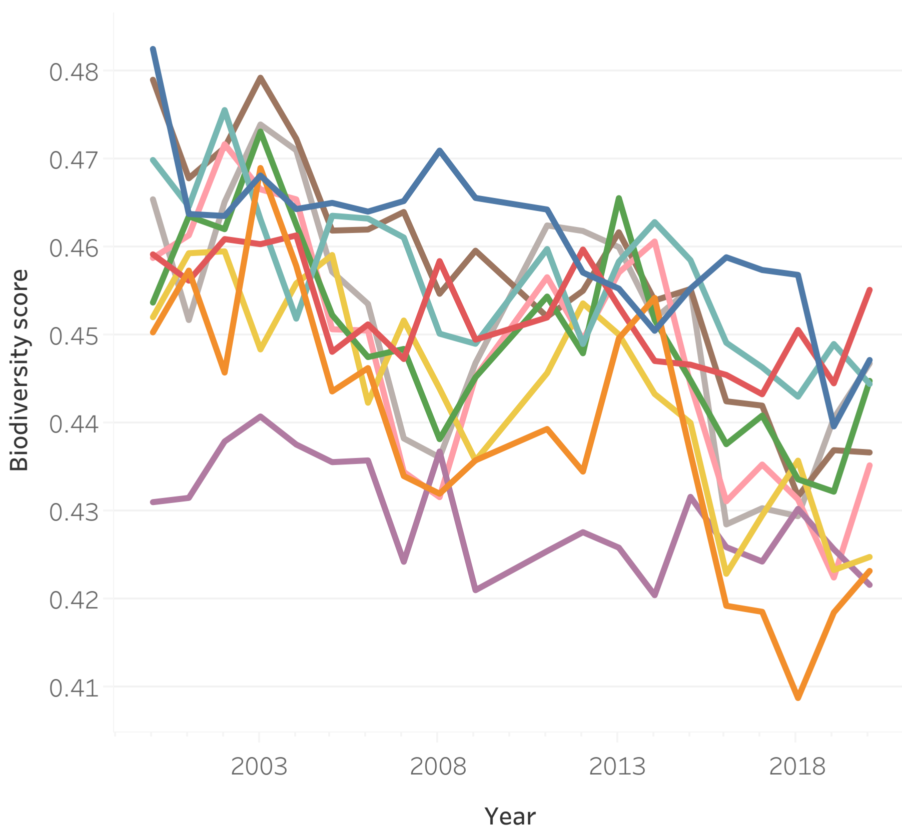
Subdivisions By Total Biodiversity

Subdivision	
1	9.2164
2	8.7758
3	9.0513
4	9.1337
5	9.0013
6	8.8782
7	8.5925
8	8.9612
9	9.1406
10	9.0274

Biodiversity score

8.5925 9.2164

Biodiversity by Subdivision Over Time



Next Steps

This analysis advocates for the protection of forest areas. I recommend a future analysis to determine which forest areas provide the most habitat to the most number and diversity of species at risk.

Deeper analysis of forest areas by location and forest type (e.g., proximity to water, connectedness, forest maturity, stand density) and biological significance (e.g., habitat is provided to how many and which species) should help to determine which forest areas should be prioritized for conservation.

I recommend further biodiversity monitoring and data collection to determine if biodiversity scores adequately capture biodiversity value for building agricultural resilience.

Lastly, I recommend continuing to collect biodiversity scores over time, to enable a robust biodiversity future forecasting.

Methodology

All data and code associated with this report can be found on GitHub: https://github.com/jocelynpender/technical-assessment/.

Data Cleaning

I trimmed text data (subdivision) and modified data types (year).
I calculated missing wetland (%) data by subtracting the sum of other land cover types (%) from 100. I calculated missing wetland area data by multiplying by 100.
I filled missing values in the cropland (%) column for the year 2015 and subdivision 5 by dividing the cropland area by 100.

Missing Data

Some missing data could not be populated. This includes:
1. Biodiversity score for the year 2010 (all subdivisions, 10 values missing)
2. Cropland (%) and cropland (area) for Subdivision 2 (all years, 21 values missing per column)
3. Wetland (%) and wetland (area) for Subdivision 2 (all years, 21 values missing per column)

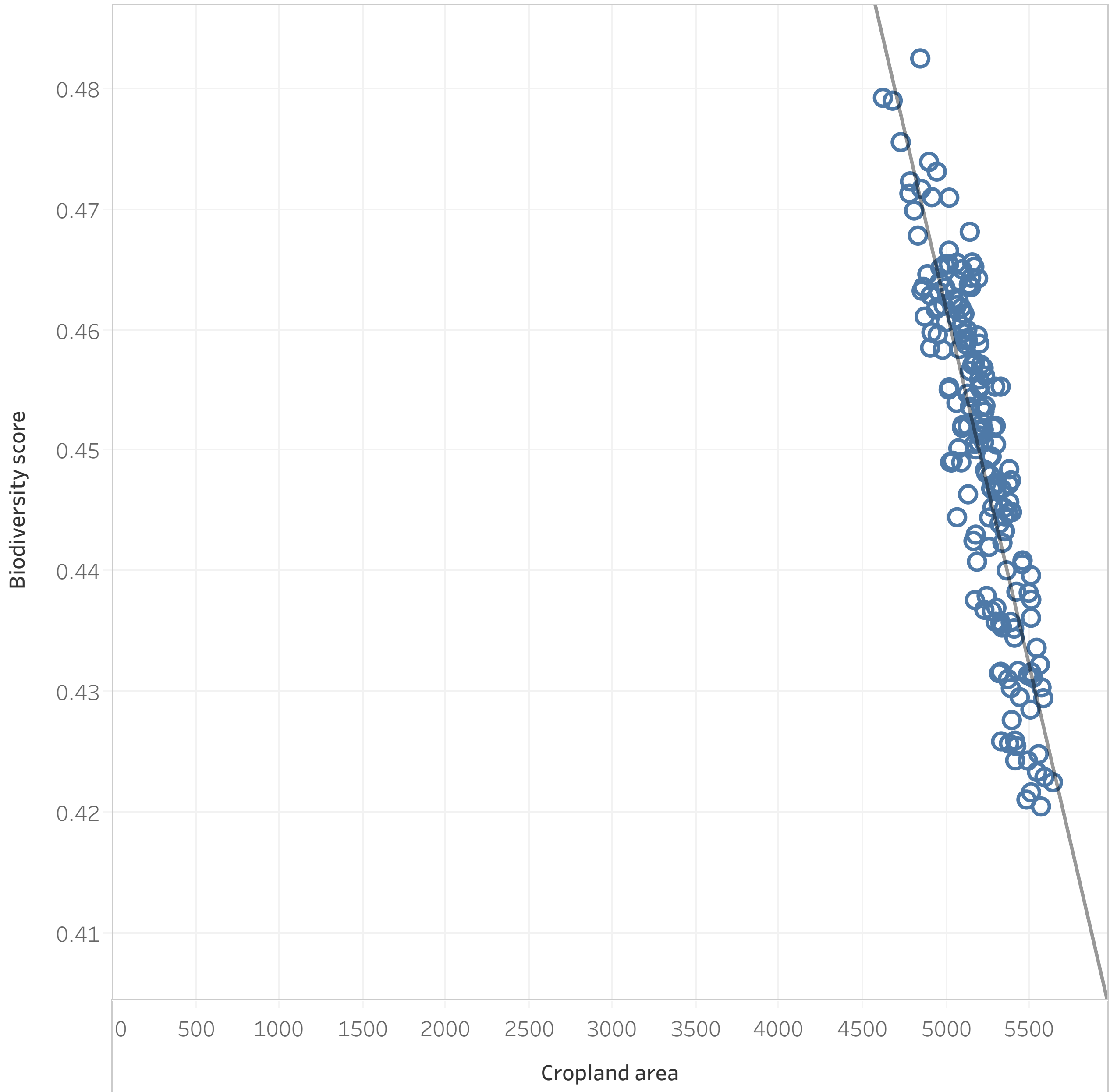
Statistical Analysis

I leveraged exponential smoothing to forecast future trends in biodiversity scores.

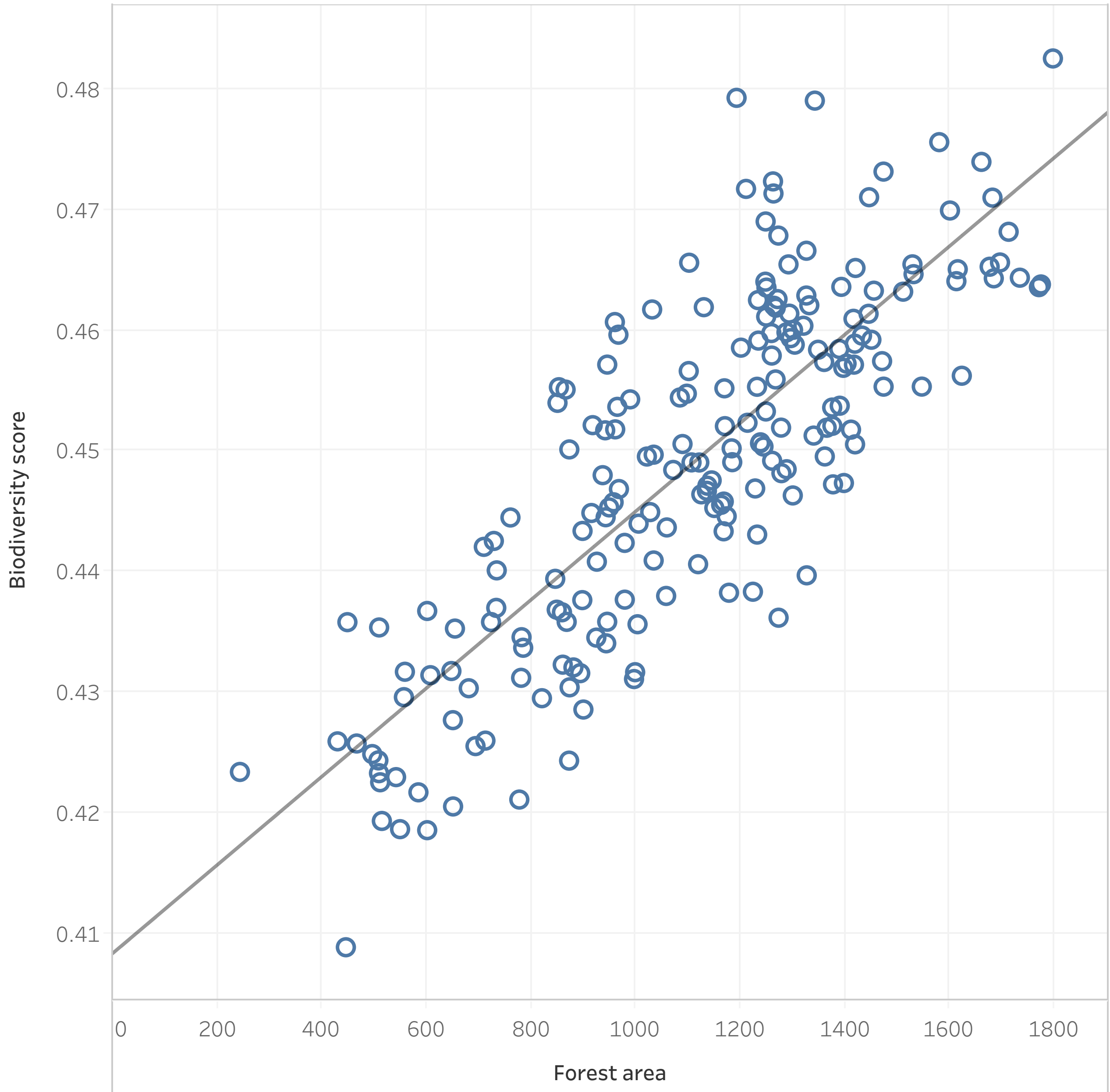
I leveraged linear regression to elucidate trends between biodiversity score and land use types. The linear regressions shown in this report assume that errors are randomly distributed around the trend line, errors are independent and observations are independent. Only statistically significant trends have been included here. See the appendix (next page) for all regressions performed.

Statistical analyses were performed using Tableau Analytics.

Appendix - Biodiversity Score and Land Area - All Land Use Types



Appendix - Biodiversity Score and Land Area - All Land Use Types



Appendix - Biodiversity Score and Land Area - All Land Use Types

Appendix - Biodiversity Score and Land Area - All Land Use Types