ANALYZING LAND COVER DATA USING SURROUNDING DATA

YUVRAJ SAHU - WESTLAKE HIGH SCHOOL

ABSTRACT

Land cover can be predicted where it is not feasible to make observations. Using the GLOBE Observer app, I collected land cover data in a 6 by 6 grid where adjacent points were 500 meters apart. I categorized the data and determined the percent coverage for each category. Since the observations only included the 100 meter by 100 meter space centered around a point, there were gaps where land cover was not classified. I selected 4 factors to test - deciduous trees, evergreen trees, grass, and urban. I wrote a program to test every point within all possible rectangles in the grid using a weighted average. For all 4 corners, it used the 2 directions facing the point being predicted. As the distance or angle difference increased, the weight decreased. Then, I calculated the mean absolute error, in percentage points, for the factors. The overall error was 16.6 for deciduous trees, 13.4 for evergreen trees, 17.5 for grass, and 38.5 for urban. In general, the small rectangles had lower error than the large rectangles. This confirms that, for certain factors, surrounding data can be used to predict land cover at a certain point.

RESEARCH QUESTION

How can surrounding observations be used to predict land cover at any point in a grid?

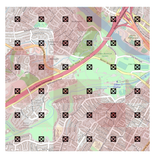
IMPLICATIONS

Land cover predictions made using this algorithm, can be combined with research from other data sets to find new correlations.

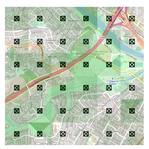
DATA COLLECTION

GLOBE Observer is an app that allows users to collect data about the Earth. Using this app, I collected and classified land cover data in a 6 by 6 grid, where the points were spaced 500 meters apart.

Area of Interest (AOI) Points

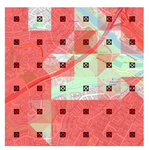


Deciduous Trees



Evergreen Trees

Grass



Urban Area

RESEARCH METHOD

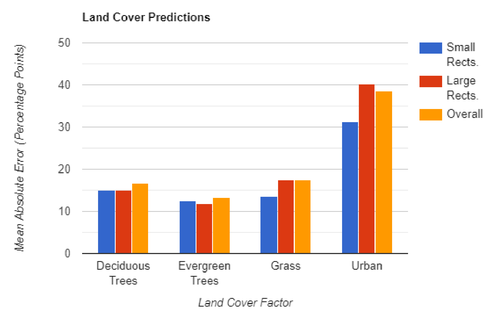
For every possible rectangle that can be made from the grid, and for every point within each rectangle, it predicts the land cover based on the four corners in order to simulate predicting the areas between each of the points.

To do this, it uses a weighted average. For each of the corners, only the two directions that are facing the point to be predicted are used. This means that 8 factors are used in the weighted average.

Distance and weight have an inverse relationship - as the distance increases, the weight in the weighted average decreases. Also, angle difference and weight have an inverse relationship - as the difference in angle increases, the weight decreases.

RESULTS

For each prediction, I calculated the absolute error - the positive difference between the predicted value and the actual value. By taking the average of these errors, I arrived at the mean absolute error. The units for the errors are percentage points.



CONCLUSION

I learned that surrounding data can be used to predict land cover at a point in a grid. For future research, I could use new data (other AOI grids, satellite data, or assorted GLOBE data), test new models, or apply this model to new contexts.

MENTORS

Scientist Mentors

Rusty Low Cassie Soeffing Peder Nelson
Becky Boger Surya Khudan

Peer Mentors

Matteo Kimura Pratham Barbaria Faguni Gupta
Kavita Kar

BIBLIOGRAPHY

Global Learning and Observations to Benefit the Environment (GLOBE) Program, 6/19/2021, globe.gov