

Regression of biodiversity indicators from satellite observations and environmental parameters

Vincenzo Madaghiele, Cosimo Chetta

MALIS Project, Fall Semester 2020

Motivation

The loss of biodiversity is one of the existential threats humanity has to face in today's world, and protecting habitats is one of the most important measures for reducing global warming. Unfortunately, measuring biodiversity accurately is a long and costly task, which requires observations on the field; it is therefore applicable to relatively small areas.

Luckily ecologists have recognized that some environmental parameters of a given area can be indicators of biodiversity, and can give a rough idea of biodiversity status. Over the years, as satellite technology evolved, ecologists have used some key variables from satellite observations to monitor the biodiversity status in important areas of the world, mainly based on how the vegetation reflects different wavelength signals sent from the satellites ([1], [2]).

Method

The purpose of this project will be to regress important biodiversity metrics from satellite data and environmental parameters (vegetation indices, climate, weather and water data, soil composition); allowing for a more accurate biodiversity estimation. In order to obtain biodiversity data the GBIF

dataset will be used ([4]), and to obtain satellite observations the geospatial environmental data offered by the Copernicus project ([3]) will be used. These two datasets will be combined using the common geographic coordinates for the chosen area, assigning a biodiversity index to each area of interest.

Intended experiments

- Regression of biodiversity index from satellite observations and environmental parameters for a given area. A comparison of performance will be made between the different regression methods studied in the course.
- Identification of the most important indicators of biodiversity among all of the variables considered in the dataset using regression trees and random forest.

Conclusions

This project does not aim to be universal: a suitable geographical area (or a set of similar areas) will be selected based on the availability of data, and the experiments will be conducted on data from the selected area and then tested also on similar habitats but in different geographic locations, to determine the generalization possibilities of the model. However, in order to obtain a better performing model, a collaboration with ecologists and expert in the field would be needed.

If the performance of the experiments prove to be successful these results could be used to have a rough constant monitoring of biodiversity in the chosen areas based on data which is relatively cheap and constantly collected. Of course these results would have to be accompanied by a more precise ecological monitoring on the field, but they could help to identify areas of major importance (biodiversity hotspots) and areas at risk, and help to efficiently allocate the necessary resources for environmental protection.

References

- [1] Zbigniew Bochenek. Monitoring forest biodiversity and the impact of climate on forest environment using high-resolution satellite images. *European Journal of Remote Sensing*, 51(1):166–181, 2017.
- [2] Woody Turner. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18(6):306–314, 2003.
- [3] www.copernicus.eu. Copernicus, 2020.
- [4] www.gbif.org. Gbif database, 2020.