EGU24-2477, updated on 16 Apr 2024
https://doi.org/10.5194/egusphere-egu24-2477
EGU General Assembly 2024

Combining citizen science and Earth observation data to produce global maps of 31 plant traits

Daniel Lusk[1], Sophie Wolf[7], Álvaro Moreno Martínez[2], Jens Kattge[3,4], Francesco Maria Sabatini[3,5,6], and Teja Kattenborn[1]
[1]University of Freiburg, Earth and Environmental Sciences, Sensor-based Geoinformatics, Germany (daniel.lusk@geosense.uni-freiburg.de)
[2]Image Processing Laboratory, Universitat de Valéncia, Valéncia, Spain
[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
[4]Max Planck Institute for Biogeochemistry, Jena, Germany
[5]BIOME Lab, Department of Biological, Geological and Environmental Sciences (BiGeA), Alma Mater Studiorum University of Bologna, Bologna, Italy
[6]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle, Germany
[7]Leipzig University, Remote Sensing Center for Earth System Research, Earth System Data Science, Leipzig, Germany

The acceleration of global environmental change underscores the pressing need for a comprehensive understanding of how the biosphere interacts with its environment. To reliably examine these connections across diverse ecosystems, having extensive and spatially comprehensive data on plant functional traits is imperative. The TRY database boasts an extensive repository of plant trait measurements for thousands of species, and, while previous approaches have attempted to spatially extrapolate these traits using environmental predictors and remote sensing data, the scarcity of the original data leads to significant uncertainty in the extrapolations. Meanwhile, citizen scientists have been actively gathering dense observations of species occurrences worldwide, and when matched with trait data, can adequately represent global trait patterns. Here, we explore the use of citizen science and Earth observation data to generate global maps of 31 ecologically relevant plant functional traits. The study utilizes sparse spatial grids created by linking species occurrences from the Global Biodiversity Information Facility (GBIF) with the TRY gap-filled database to generate continuous global trait maps as a function of climate, soil, and remote sensing data. We first evaluated model performance using spatial cross-validation, and they demonstrated up to $R^2 = 0.53$ with a normalized RMSE = 0.21. We then compared mean trait values from the GBIF-based extrapolations to community-weighted mean traits from sPlotOpen, a global, environmentally balanced dataset of vegetation plot data. Our results show correlations between the two datasets of up to $r = 0.73$ with particular resilience to decreasing map resolution. When compared to similar extrapolations based on sPlotOpen alone, we found that GBIF-based extrapolations increased global spatial applicability for all maps by up to 12%. Additionally, we show that GBIF-based extrapolations have higher correlations to sPlotOpen-derived maps than the majority of previously published trait maps. Despite the inherent noise and biases of their crowd-sourced origins, GBIF-based models are remarkably capable of producing

even closer approximations of the trait distributions of scientifically controlled vegetation plots than their own sparse reference data. Considering the rapid growth and availability of crowd-sourced data, the capacity of models to overcome their noisy and opportunistic nature further affirms the potential of databases such as GBIF to complement more scientifically rigorous data collections.