

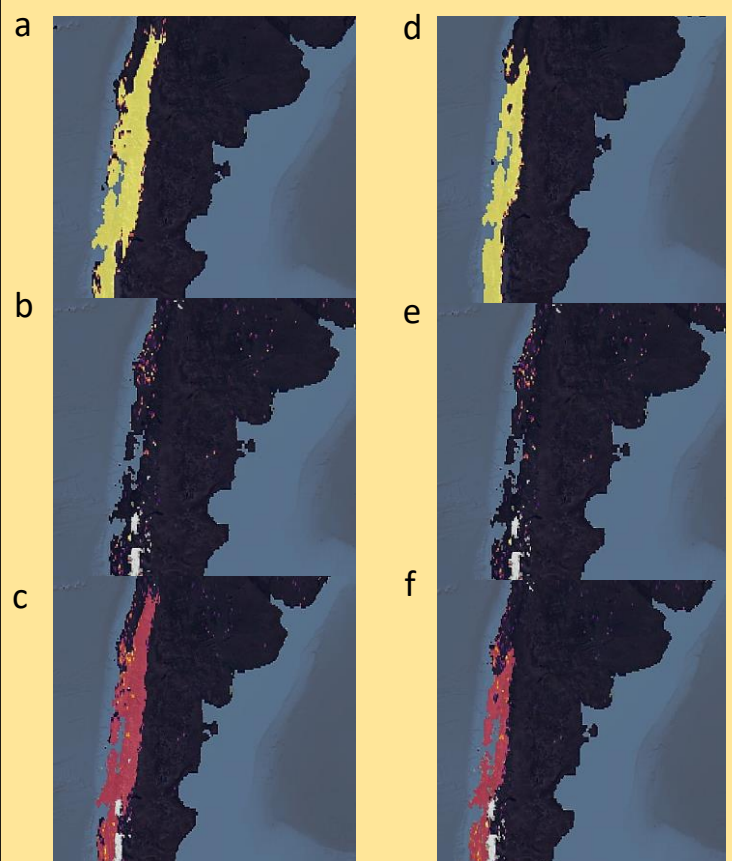
Earth observation, open data and machine learning for near real time threat monitoring of vulnerable plant species

Burra DD (d.burra@cgiar.org), Barua MA, Palao LK and Reymondin L,
International Center for Tropical Agriculture – Asia



The IUCN database lists several plant species whose existence is currently threatened by human activities and climatic extremes. Here we report a methodology that monitors threat status of these species in near real time, by deriving data from multiple open data sources, by linking them via a machine learning analytical framework, with interpretations facilitated by a web based geospatial visualization framework.

1. Red listed species list extraction (as on May 2018) using batch query system on <http://www.iucnredlist.org/> resulted in > 500 species.
2. Use of rGBIF R package interface (R package that interacts with the GBIF repository) to batch extract geotagged occurrence records of the red listed species, resulted in 140 species with at least one geotagged occurrence record.
3. The top 5 species with most number of occurrence records (Pilgerodendron uviferum had the most number of occurrence records – 128 in total) were used to develop the methodology and build the current version of the tool.
4. Used bioclimatic data (Worldclim, 2.5 minutes) to model, predict and map species distribution using random forests and Xtreme gradient boosting algorithms across continents, wherein the species were originally identified. Only models above 0.8 AUC value were used for prediction.
5. Transformed the deforestation events captured by terra-i (<http://www.terra-i.org/terra-i.html>) as on 15 August 2018 into presence and absence points of deforestation.
6. Developed a web based geospatial visualization framework built using GeoServer and Leaflet to visualize raster data outputs from steps four and five.



Figures a and d are species distribution model outputs for *Prumnopitys andina* and *Pilgerodendron uviferum*, b and e are deforestation point detections from terra-i for the area where the above species presence has been predicted, while c and f overlay deforestation detection with detected species presence (yellow dots in c and f correspond to locations where the species is predicted and a deforestation event was detected) for *P. andina* and *P. uviferum* respectively