Farm Sites Features Exploration

-Mayaguez, Puerto Rico

Jingzhe Wang

3/6/2018

UNIVERSITY OF MINNESOTA
**Driven to Discover**®

Motivation

- Derived from a previous project on organic farm site selection
- Context about organic farming :
  - Enhance the resilient capacity facing extreme events such as Hurricane (Jouzi et al., 2016)
  - Help reduce the unemployment rate (Jouzi et al., 2016)
  - A great portion of production expenses goes to labors; competitive wages (USDA, 2012 & 2015)
  - Choose Area of Interest (AOI) in the forest region located in the mountain area of Mayaguez, western Puerto Rico based on the location of one known organic farm information

Data preparation

- Data search and data prepossessing were done

- Aim to explore the physical environment features for the existing organic farms, compare with the ones of non-farm areas or non-organic farms → very limited information

- Better remove the 'organic' limitation first

- One organic farm site on the AOI, two other farm sites found on the satellite image of the AOI

- Randomly specified three polygon areas on the satellite image which looked like residential area

- Extract points from three farm areas and three non-farm areas to form the observations
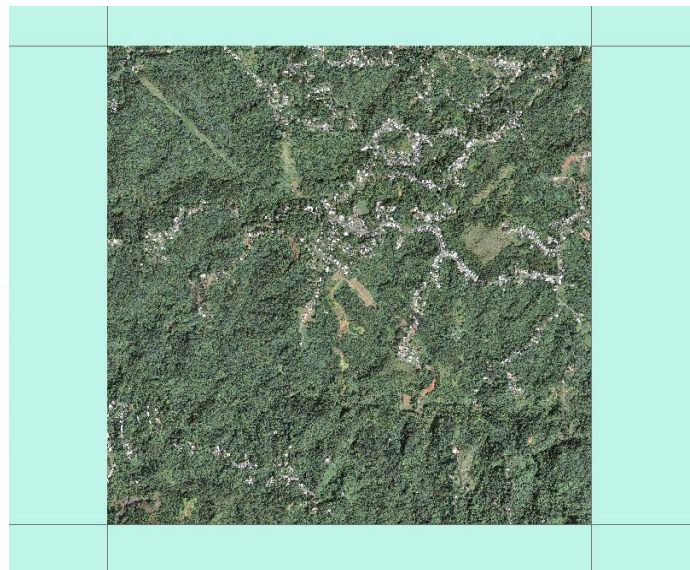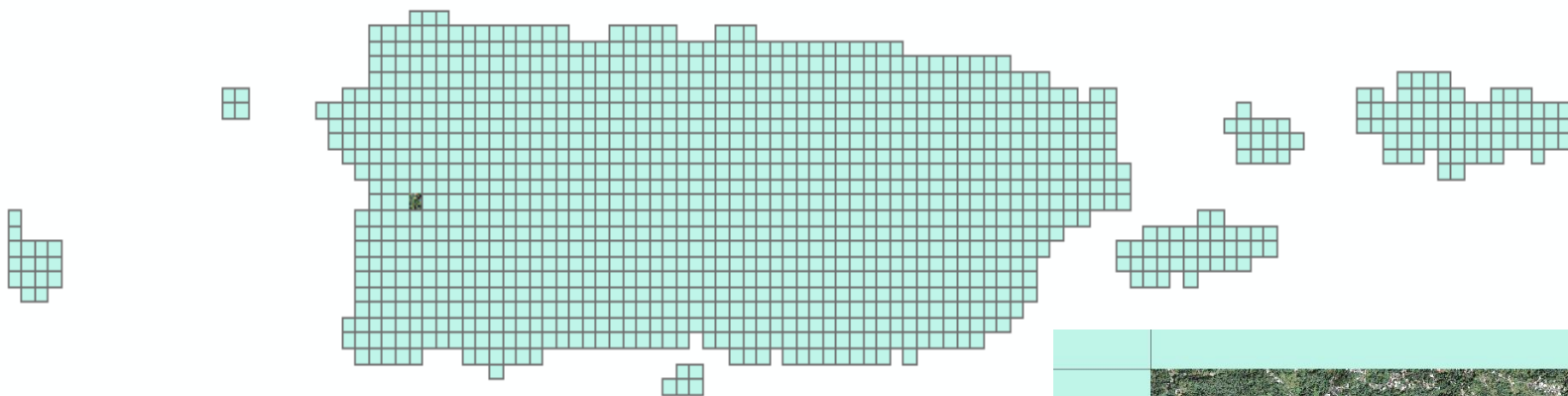
**Source:** U.S. Geological Survey
USGS, 2009

Data and Methods

- Response: farmland (binary, 'Yes' or 'No')
- Variables:
  - Factor variables: stype (soil type, 7 levels), landcover(4 levels),
  - Numeric or integer: organic_percent, avail_water, slope, canopy(tree canopy coverage rate)
  - XY-Coordinate
- Spatial autocorrelations as location points from the same location area (polygon) will have similar or same attribute values
- Common parametric regression models might not be good choice
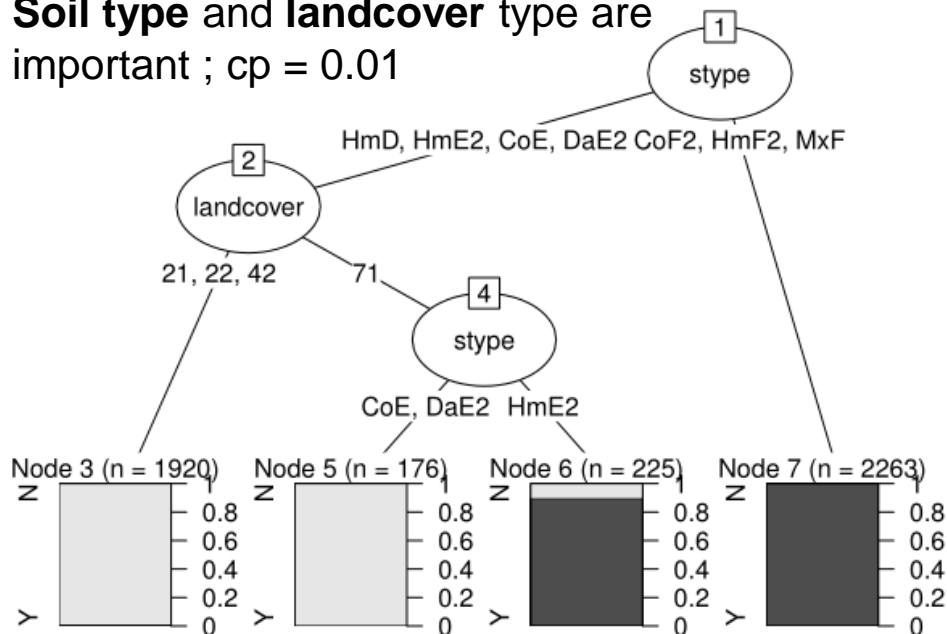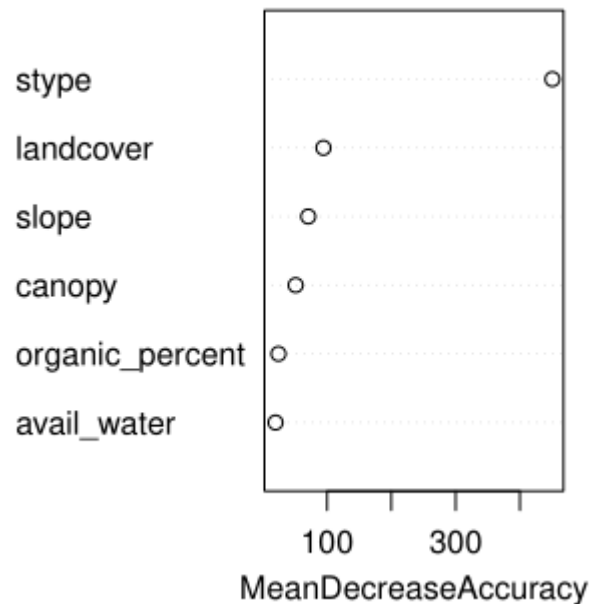- Focus on single decision tree and bagging methods

Scatterplots
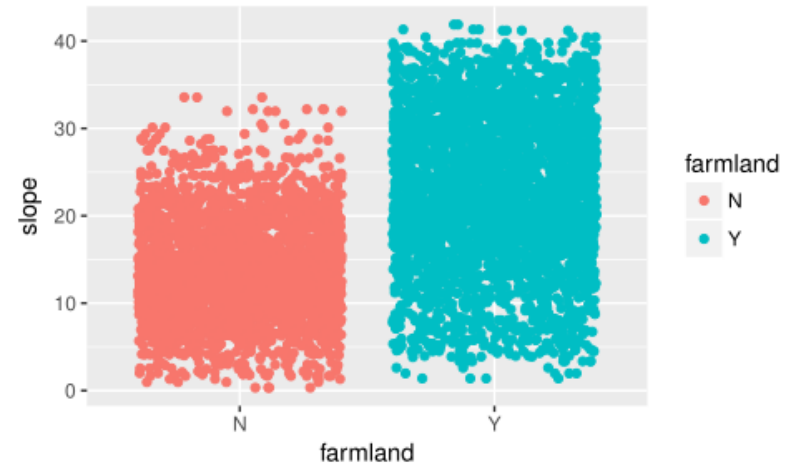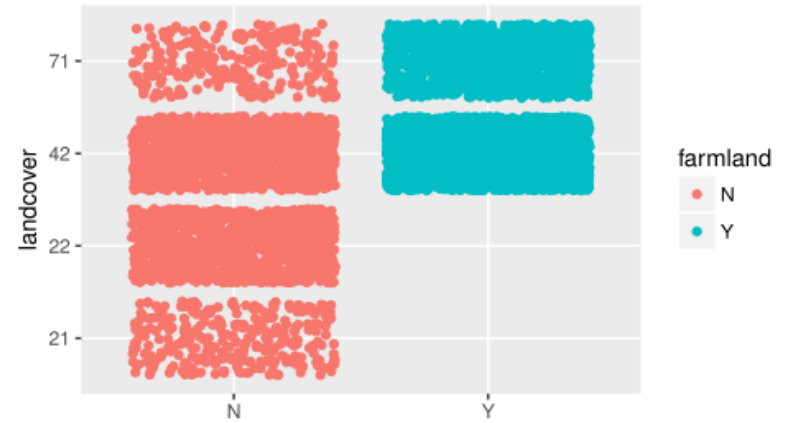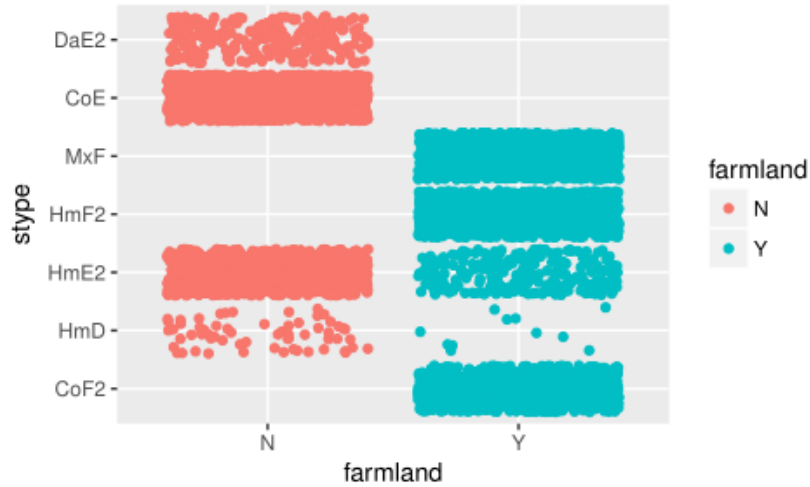
Decision Tree and Bagging

Soil type and **landcover** type are important ; cp = 0.01

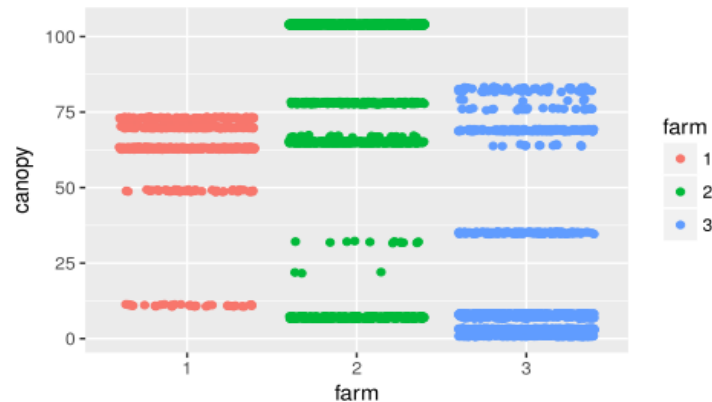Misclassification rate on test set: 0.92%

No misclassification on test set

Comparing Three Farm Sites

Comparing Three Farm Sites

Soil type is important
cp = 0.03

Canopy is important
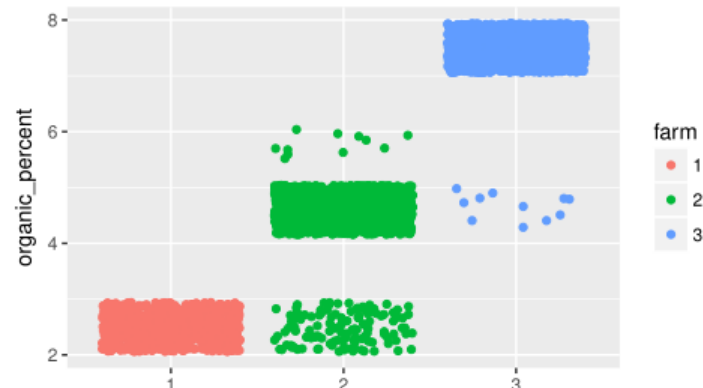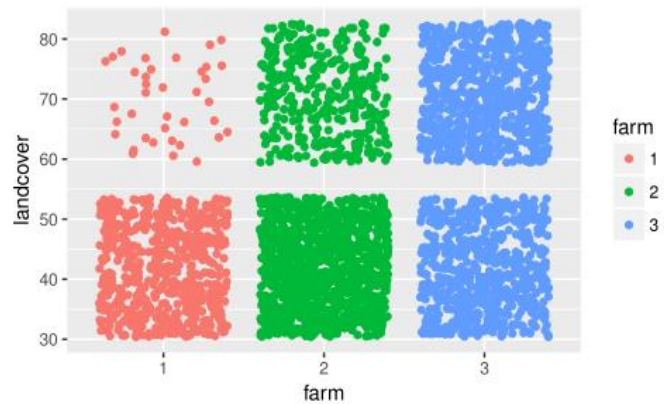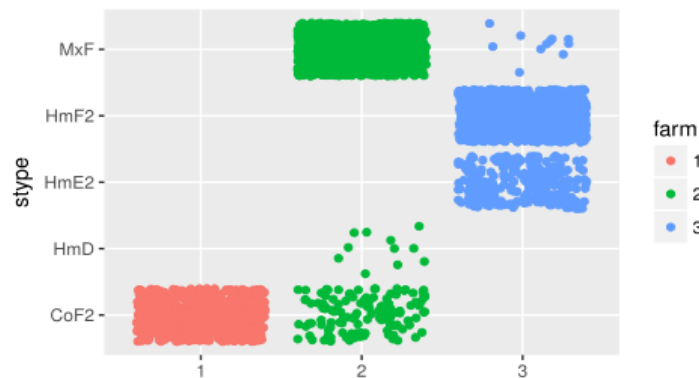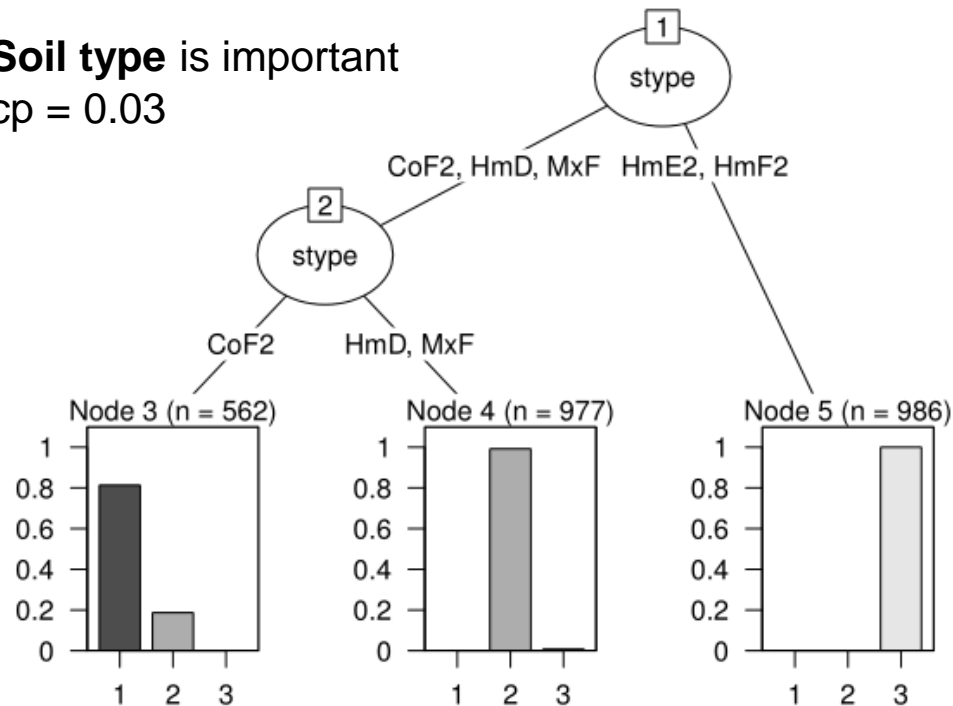

Misclassification rate on test set: 2.97%
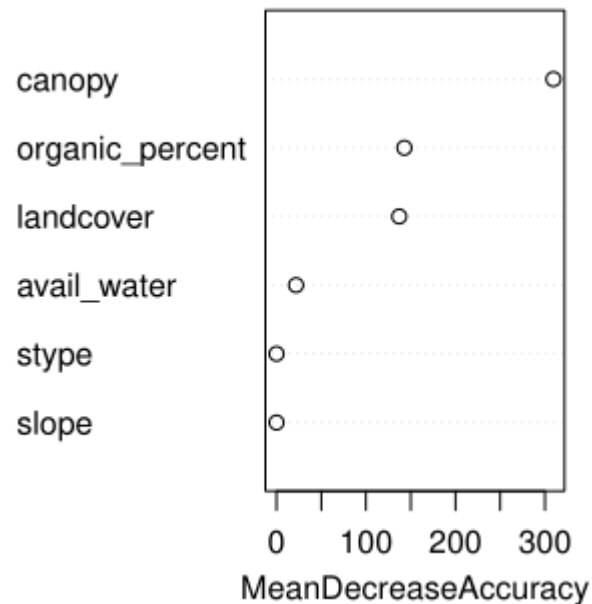
No misclassification on test set

Conclusion based on the sample

- To distinguish farmland from non-farmland:
  - Soil type: CoF2, HmF2, MxF, HmE2
  - Landcover type:
    avoid 21 (developed, open space), 22 (developed, low intensity), 42 (evergreen forest); 71 (grassland/herbaceous) may be neutral

- For three farms (farm3 is the organic farm):
  - Soil type: HmE2 and HmF2 for organic farm
  - Tree canopy coverage rate may be important (tried if not prune the tree)

Insights

```
farm   count
<fct>  <int>
1        457
2       1074
3        994
```

- Different results from single decision tree and bagging
  - More likely to pick location points from farm2 and farm3 rather than farm1 during the resampling process. The resulting decision trees will have fairly different important variables be selected then as the variance for each single decision tree is high.

- Non-farmland points all from residential area, non-farmland is not necessarily residential area; limited information about other non-farmland types

- Sample location points were from six polygon areas, they might not be representative for the whole AOI (e.g. some different kinds of soil are in other locations)

- Regression might be really helpful for assigning weights to different spatial attributes → weighted overlay methods

References

Farm Economics *Record high agriculture sales; income and expenses both up.*, Highlights, 2012 Census of Agriculture, United States Department of Agriculture, 2014. Retrieved from https://www.agcensus.usda.gov/Publications/2012/Online_Resources/Highlights/Farm_Economics/

How Weighted Overlay Works, Tool Reference, ArcGIS Desktop (2017). Retrieved from http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-weighted-overlay-works.htm

Jouzi, Z., Azadi, H., Taheri, F., Zarafshani , K., Gebrehiwot, K., Passel, S. V., & Lebailly, P. (2017). Organic Farming and Small-Scale Farmers: Main Opportunities and Challenges. *Ecological Economics, Elsevier*, *132*, 144–154. Retrieved from https://doi.org/10.1016/j.ecolecon.2016.10.016

Organic Farming *Results from 2014 Organic Survey*, Highlights, 2012 Census of Agriculture, United States Department of Agriculture, 2015. Retrieved from https://www.agcensus.usda.gov/Publications/2012/Online_Resources/Organics/index.php

Soil Survey Staff, Soil Survey Geographic (SSURGO) database for Mayaguez Area, Puerto Rico Western Part, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at the following link: https://websoilsurvey.sc.egov.usda.gov/ . Accessed [11/25/2017].

Soil Survey staff, Natural Resources Conservation Service ,United States Department of Agriculture. Retrieved from

https://www.nrcs.usda.gov/Internet/FSE_MANUSCRIPTS/puerto_rico/PR686/0/San_Juan.pdf

U.S. Geological Survey, 2009. High Resolution Orthoimagery: 2739264_PUERTORICO20100531. https://lta.cr.usgs.gov/high_res_ortho

U.S. Department of the Interior, U.S. Geological Survey, 2013. Land Cover 100 Meter Resolution-Puerto Rico and the U.S. Virgin Islands, Albers

Projection.

https://nationalmap.gov/small_scale/atlasftp.html?openChapters=chpbio#chpbio

U.S. Department of the Interior, U.S. Geological Survey, 2013. Tree Canopy-Puerto Rico and the U.S. Virgin Islands 100 Meter Resolution Albers

Projection.

https://nationalmap.gov/small_scale/atlasftp.html?openChapters=chpbio#chpbio

U.S. Geological Survey, 2017. USGS NED 1/3 arc-second n19w068 1 x 1 degree ArcGrid 2017
http://nationalmap.gov/viewer.html