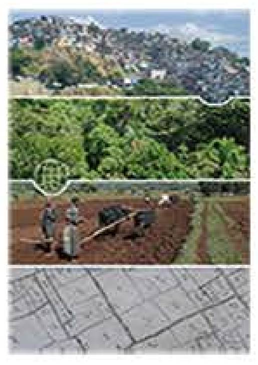
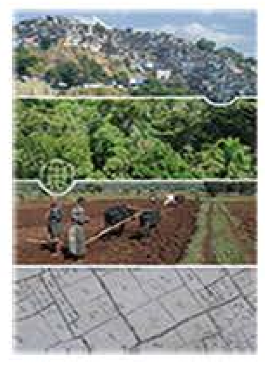


Responsible Land Governance: Towards an Evidence Based Approach

ANNUAL WORLD BANK CONFERENCE ON LAND AND POVERTY
WASHINGTON DC, MARCH 20-24, 2017



Downscaling Land Use Data: Pixel Level Cropland Allocation and the Impacts of Biophysical Factors

Jingyu Song*, Michael S. Delgado*, Paul V. Preckel*, and Nelson B. Villoria**

***Department of Agricultural Economics, Purdue University**

****Department of Agricultural Economics, Kansas State University**

Corresponding author: Jingyu Song song173@purdue.edu

Introduction

Land allocation has long been a focus for policymakers across the world. It is also an important research focus in the academic community. Researchers treat land use patterns as a means to assess the impact of various human activities on ecological-environmental systems and to inform public policy. Observational cropland allocation data is widely available only at an aggregate level, which is typically defined as national or subnational (i.e., state/provincial, and in some cases district/county).

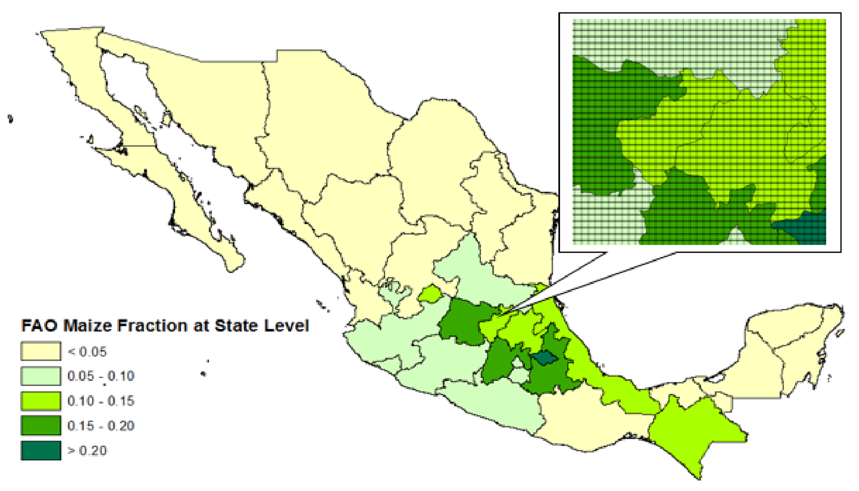


Figure 1. Observed FAO maize fraction data for Mexico at state level and zoomed in pixelated land area.

Objective: to fill the gap between the need for generally unavailable fine-scale land allocation data that is based on observational measurements, and to contribute to the scientific understanding of cropland allocation across a large geographic area at a fine-scale.

Theoretical Framework

We are interested in estimating the parameters β in the conditional mean for pixel level share z_{ijk}:

$$E[z_{ijk} \mid x_{ij}] = G_{ijk}(x_{ij}, \beta_k) \qquad (1)$$

where j denotes administrative units, k represents crops. y_{jk} is the observed fraction of land area in administrative unit j that is in crop k, $0 \le y_{jk} \le 1$. z_{ijk} is the (unobserved) fraction of cropped land in pixel i in administrative unit j that is cropped in crop k, where $i = \{1, 2, \ldots, I_j\}$ is the pixel index that allows the total number of pixels in each administrative unit to vary. x_{ij} is an M-dimensional vector of observable land attributes for pixel i in administrative unit j. We parameterize $G(\cdot)$ using a logistic function, and the predicted fraction of land in crop k in pixel i in administrative unit j becomes:

$$G_{ijk}(x_{ij}, \beta_k) = \frac{\exp(x_{ij}\beta_k)}{\sum_{l=1}^{K} \exp(x_{ij}\beta_l)} \quad \text{where } \beta_1 \equiv 0. \qquad (2)$$

The predicted fraction of land in crop k in administrative unit j is:

$$H_{jk} = \frac{\sum_{i \in I_j} G_{ijk}(x_{ij}, \beta_k) A_{ij}}{\sum_{i \in I_j} A_{ij}} \qquad (3)$$

where A_{ij} is the area of pixel i in state j. Given H_{jk} the log-likelihood function to be maximized with respect to the parameters β_k is:

$$\sum_{j=1}^{J} \sum_{k=1}^{K} y_{jk} \ln(H_{jk}). \qquad (4)$$

Acknowledgements

The authors are grateful for comments and suggestions from Timothy Baker, Thomas Hertel, David Laborde, Navin Ramankutty, Juan Sesmero, AAEA 2015 participants, and seminar participants at IFPRI, Washington, DC. This research is supported in part by USDA grant Agreement #58300010058, Purdue grant #105651.

Data and Results

We consider two empirical applications:

- Single-crop for maize
- Multi-crop for maize, soybeans, and wheat.

Both focus on harvested crop area throughout North, Central and South America.

The total areas of land harvested in maize, soybeans, and wheat at Administrative Unit Level 1 are collected from FAO Agro-MAPS.

Independent variables include average temperature over the growing season (New et al., 1999), average annual total precipitation (New et al., 1999), elevation (United States National Geophysical Data Center TerrainBase global model of terrain and bathymetry (1995)), soil pH (IGBP-DIS, 1998), soil carbon content (IGBP-DIS, 1998), slope (IIASA/FAO, 2012), and latitude.

Table 1. Parameter estimates for the maize model

Variable	Coefficient	Standard Error
Intercept	-7.1552***	1.9488
Temperature	0.5818***	0.1575
Temperature Squared	-0.0212***	0.0041
Precipitation	-0.5780	0.7133
Precipitation Squared	-0.5841***	0.1407
Temperature·Precipitation	0.1142***	0.0278
Elevation	-0.0623	0.2810
Max(pH6.5-pH,0)	-0.9227***	0.2031
Max(pH-pH6.5,0)	-1.6218***	0.3832
Soil Carbon	0.1566***	0.0467
Slope	-6.1423**	2.2885
Latitude	0.0442**	0.0148
Latitude Squared	0.0004	0.0005

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Table 2. Implied marginal effects and odds ratios for the multi-crop model

Variable	Maize		Soybeans		Wheat	
	Marginal Effect	Odds Ratio	Marginal Effect	Odds Ratio	Marginal Effect	Odds Ratio
Temperature	-0.0119	0.7330	-0.0248	0.7585	-0.0051	1.1831
Precipitation	0.2041	3.9022	-0.2257	32.8382	-0.0134	2.7792
Elevation	0.2219	10.0657	-0.4400	0.0382	-0.0024	70.6299
Max(pH6.5-pH, 0)	0.0066	0.8911	-0.0788	0.5189	-0.0046	3.0646
Max(pH-pH6.5, 0)	-0.0968	0.3468	-0.0061	1.1989	0.0050	9.5469
Soil Carbon	0.0079	1.0947	-0.0049	0.9820	0.0018	1.0646
Slope	0.5607	18.7818	-1.2198	0.0002	-0.3749	0.0001
Latitude	0.0163	1.3702	0.0107	1.3750	0.0021	1.1277

Model Validation and Conclusion

We compare the crop fraction predicted by our model to the actual crop fraction reported by FAO Agro-MAPS at both Administrative Levels 1 and 2, considering both in-sample and out-of-sample validation. The points generally cluster around the 45 degree line at both levels for both models, indicating that our models generally predict well.

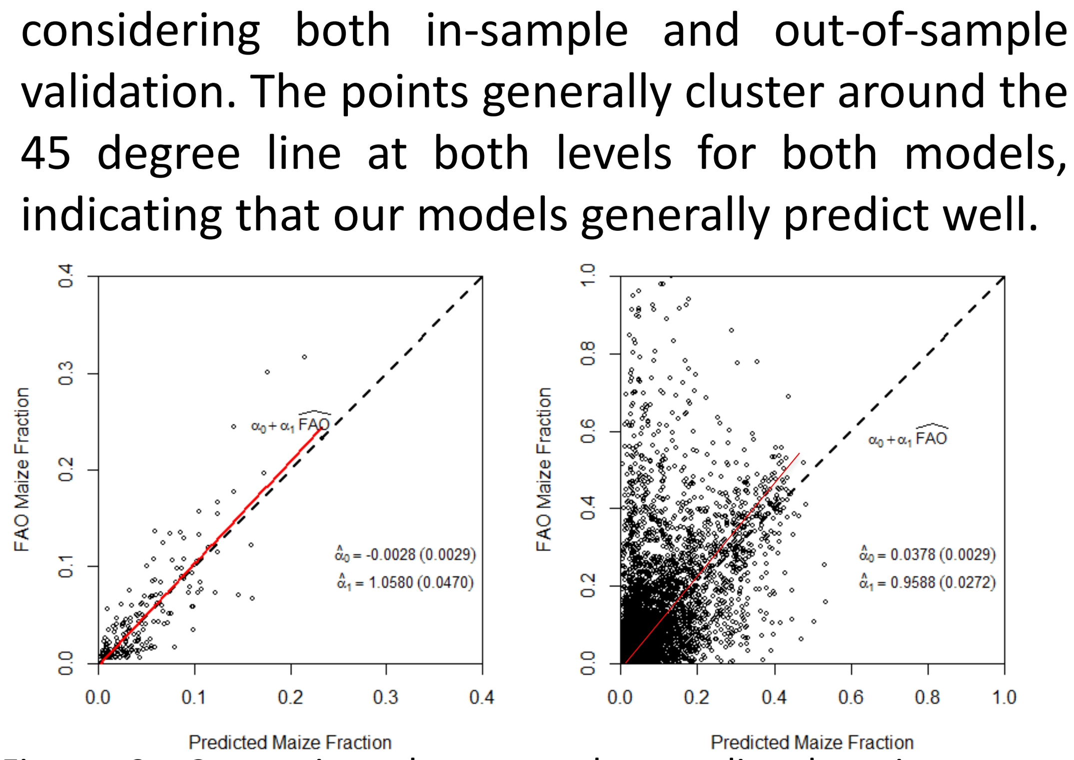


Figure 2. Comparison between the predicted maize versus observed FAO maize area fraction at Administrative Levels 1 (left) and 2 (right) for the maize model

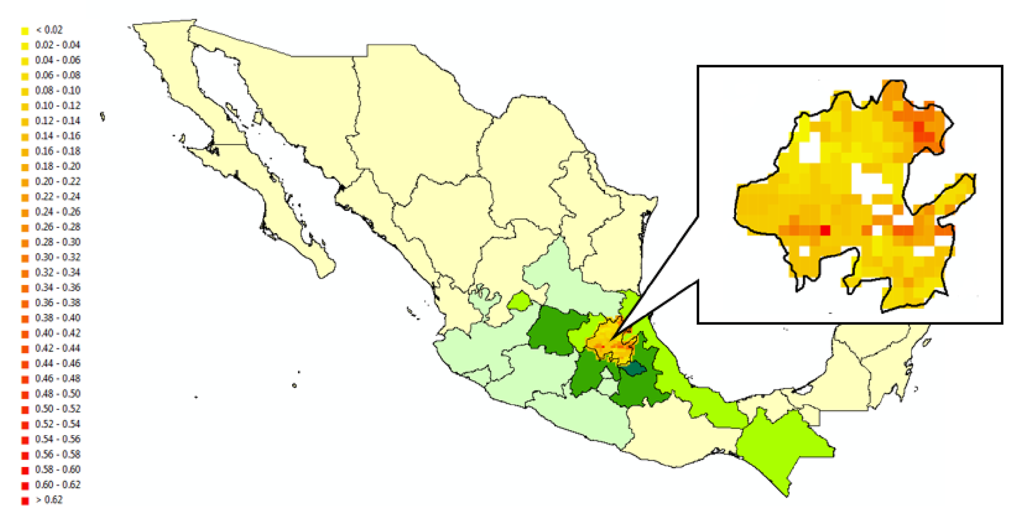


Figure 3(a). Down-scaled maize fraction predictions for Hidalgo, Mexico using constant FAO Level 2 shares for all pixels

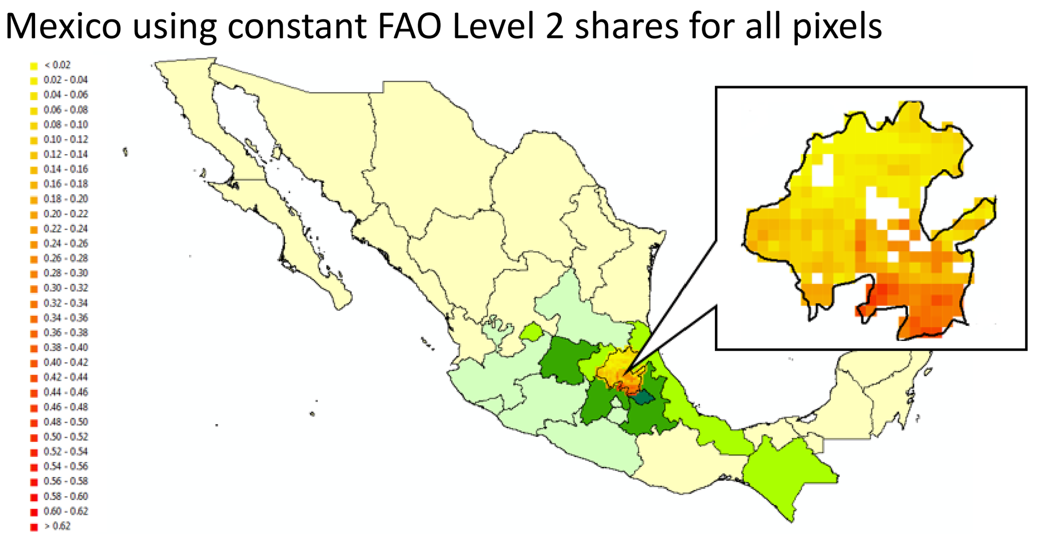


Figure 3(b). Down-scaled maize fraction predictions for Hidalgo, Mexico using the estimated pixel-specific maize shares

Validation Illustration: the south-eastern part of Hidalgo is home to the Valley of Mexico, a largely agricultural area. Our model correctly identifies that.

References

1. New, M., Hulme, M., and Jones, P. (1999). Representing twentieth-century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, 12(3), 829-856.
2. National Geophysical Data Center/NESDIS/NOAA/U.S. Department of Commerce. 1995. TerrainBase, global 5 arc-minute ocean depth and land elevation from the U.S. NationalGeophysical Data Center (NGDC). Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.
3. IGBP-DIS. SoilData (V.0) a program for creating global soil-property databases, IGBP global soils data task, France, 1998.
4. IIASA/FAO. Global Agro-ecological Zones (GAEZ v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2012.