A New Five-Minute Land Use Data Set for Amazonia Produced from Satellite- and Agricultural Census-Based Data

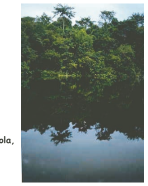
Center for Sustainability and the Global Environment,
Institute for Environmental Studies,
University of Wisconsin-Madison
http://sage.aos.wisc.edu

Large-Scale Biosphere-Atmosphere Experiment in Amazonia
http://lba-ecology.gsfc.nasa.gov/ lbaeco/

Jeff Cardille (1), Jonathan Foley (1), Marcos Costa (2)
(1)Center for Sustainability and the Global Environment, University of Wisconsin-Madison
(2) Federal University of Vicosa, Vicosa, Brazil
contact: Jeff Cardille, cardille@students.wisc.edu
Phone: 608-262-4775

University of Wisconsin-Madison
Madison, WI, USA

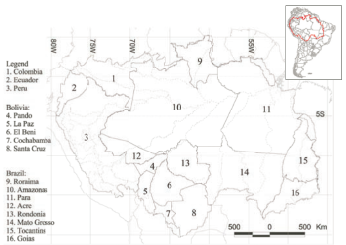
Departamento de Engenharia Agricola,
Universidade Federal de Viçosa,
Viçosa, MG, Brazil



1- INTRODUCTION

Although there are several existing land cover classifications of Amazonia for the mid-1990s, they neither agree on the location of cropland nor explicitly include pasture in the classification schemes. As a result, this vast (6.7 x 10^^8 ha) region covering most of northern South America is under great threat of development, yet lacks a definitive map of human land use. Most agricultural development in Amazonia is the conversion of land to rangeland and planted pasture, two land uses which are spectrally similar to other ecosystems in the region and are thus not well captured in satellite classifications.

Here we present a new data set of cropland and pasture density for Amazonia at five-minute (~ 10 km) spatial resolution for the mid-1990s. Produced from a unique fusion of existing land cover classifications and sub-national agricultural census data, this data set retains the attributes of the census while gaining the spatial detail inherent in satellite imagery.



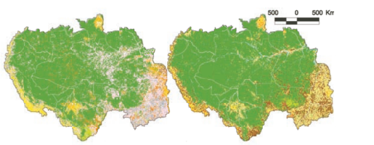
Major administrative boundaries in the Amazon and Tocantins river drainage basins.

Total Basin Area:

68.9 x 10^^5 km^^2

- Brazil:
  - 45.4 x 10^^5 km^^2
- Peru:
  - 9.4 x 10^^5 km^^2
- Bolivia:
  - 6.9 x 10^^5 km^^2
- Colombia:
  - 3.3 x 10^^5 km^^2
- Ecuador:
  - 1.3 x 10^^5 km^^2

2- LAND USE AND LAND COVER DATA SETS



IGBP DISCover land cover classification (Belward and Loveland, 1996)

Belward, A.S., and T.R. Loveland, The DIS 1 km land cover data set, in *Global Change Newsletter*, pp. 7-8, The International Geosphere-Biosphere Programme: A study of global change (IGBP) of the International Council of Scientific Unions, 1996.

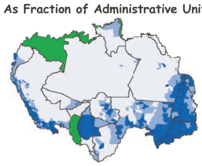
University of Maryland Global Land Cover Facility classification (Hansen et al., 2000)

Hansen, M.C., R.S. Defries, J.R.G. Townshend, and R. Sohlberg, Global land cover classification at 1km spatial resolution using a classification tree approach, *International Journal of Remote Sensing*, *21* (6-7), 1331-1364, 2000.

Per-pixel agreement (black= agree, white = disagree) indicates that these land cover classifications, both built using 1992-93 1 km AVHRR data, disagree in most areas that are not Evergreen Broadleaf Forest (in green at left).

Since we could not directly extract land use information from the land cover data sets above, we created a mid-1990s map of cropland and pasture density using all available agricultural census data from countries of Amazonia (right). This map presents, for the first time, the fraction of each administrative unit that was reported in the census as Cultivated Area, Natural Pasture, or Planted Pasture. For this new map, we used the finest detail available: 2 levels of organization below the country level for Brazil; and at the third level of organization in Peru (far right).

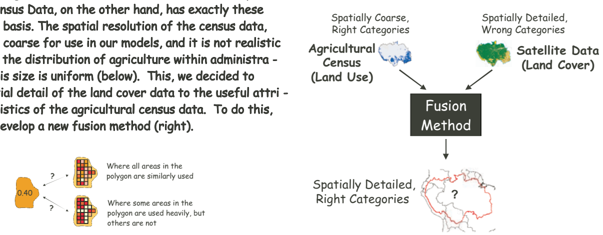
Mid-90s Agricultural Area, As Fraction of Administrative Unit



- Brazil:
  - 1995 Agricultural Census (Municipio level)
  - '80 (Muni), '75, '70 (Muni), '60 (State)
- Bolivia
  - 1984 Ag Census (County)
  - 1990-1998 (State), 1950
- Peru
  - 1994 (Distrito)
  - 1972 (Prov.), 1961(State)
- Ecuador
  - 1995 Encuesta Nacional (State)
- Colombia
  - No data found within the basin

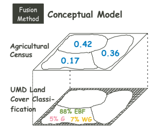
3- FUSION METHOD

Existing land cover data sets do not contain categories for Pasture, and do not agree on the location and abundance of Cropland. Agricultural Census Data, on the other hand, has exactly these variables as its basis. The spatial resolution of the census data, however, is too coarse for use in our models, and it is not realistic to assume that the distribution of agriculture within administrative units of this size is uniform (below). This, we decided to merge the spatial detail of the land cover data to the useful attribute characteristics of the agricultural census data. To do this, we needed to develop a new fusion method (right).



Our conceptual model is based on the fact that the agricultural censuses and satellite classifications looked at the same place at nearly the same time. We hypothesized that although the land cover classifications did not look for agriculture to this level of detail, there should be a systematic relationship between these classifications and agricultural censuses. In particular, we hypothesized that there is a set of categories in the land cover classifications (such as, for example, "Wooded Grassland"), which were likely "confused" in agricultural areas. Our fusion method seeks to find and use this relationship, in effect "pulling" land use data from the land cover information.

Fusion Method — Conceptual Model

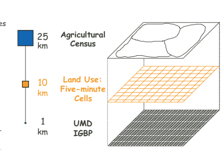


In the above illustration, we use the county boundaries from the census data to reinterpret the land cover classification by computing, in each county, the proportions of each land cover. In the figure above, according to the UMD classification the highlighted county has 88% Evergreen Broadleaf Forest, 5% Grassland, 7% Wooded Grassland, and 17% agriculture according to the census.

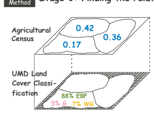
This method fuses data sources of quite different scales (right). Satellite data, derived from AVHRR imagery, was at 1 km spatial resolution. In this region, the median size of an administrative unit in the agricultural censuses was about 625 km2. We can think of this as a 25 km x 25 km square.

What is an appropriate scale for the merged data? The five-minute, or about 10 km x 10 km, scale is roughly halfway between the spatial scale of the two data sources. We have developed our fusion method to determine land use density at this scale intermediate between the two inputs.

Fusion Method — Scale Questions



Fusion Method — Stage 1: Finding the relationship

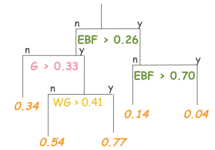


First, we compute the land cover fractions in each administrative unit.

Unit	Ag Frac	EBF	WG	G
A	0.17	0.88	0.07	0.05
B	0.42	0.30	0.27	0.43
C	0.36	0.52	0.21	0.27

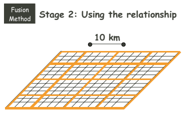
Dependent Variable — Independent Variables

Next, we assemble them into a table to be analyzed.



A regression tree, built using half the data, captures the relationship. This tree stores how land cover conditions in a county relate to observed agricultural fraction.

Fusion Method — Stage 2: Using the relationship



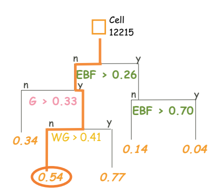
To use the relationship between land cover and agricultural fraction, we then compute the land cover fractions in each five-minute (~10 km) cell.

Recalling that these land cover fractions were the predictors in the regression tree, we assemble the fractions into a table with one line for each cell.

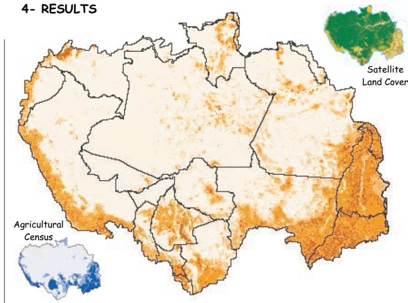
Cell	EBF	WG	G	Ag Frac
12215	0.21	0.14	0.46	?
12216	0.54	0.11	0.19	?
12217	0.44	0.21	0.17	?

The goal is to determine the most appropriate agricultural fraction for each cell given the cell's land cover conditions.

Using the tree's ability to "key out" each cell, we assemble the agricultural fractions for each cell in the map. After 50,000 extrapolations, we have created the map.

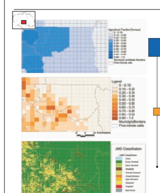


4- RESULTS



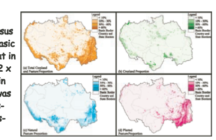
These resulting map of agricultural activity (left) represents the basin-wide relationship between census data and interpreted land cover as extrapolated even into countries without censuses. The five-minute spatial scale is suitable for regional ecosystem modeling and is midway between that of the coarse census data and fine satellite imagery. The method has the effect of "scaling down" the census, distributing the data in each administrative unit into those cells most likely to have been converted for human use.

The resulting map (in orange) shows the density of agricultural activity for the mid-1990s, expressed as a fraction of each five-minute cell. This map clearly has characteristics of both the census data and satellite data (insets) used to form it.



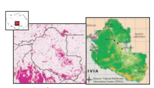
At left, the relative scales of the census data (top), the satellite data (bottom), and the resultant map (middle) are evident. The fusion method appears to create a map that successfully distributes agricultural census data into those parts of the administrative unit most likely to contain agricultural land use, according to the land cover characteristics.

After creating the total-agriculture maps, we used census data to apportion the three basic land uses. Results indicate that in the mid-1990s, there were 8.2 x 10^^7 ha of agricultural land in Amazonia; of this total, 21% was cropland, 40% was natural pasture, and 40% was planted pasture.



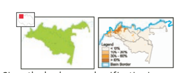
This new image of agricultural activity has several distinct advantages over using only satellite imagery or census data:

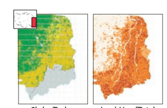
The maps compare well to existing maps of deforestation. Here, in Rondonia, Brazil (right), we can see good agreement between the new land use map and the 1992 deforestation map, developed using 30 m Landsat data.



Land Use (Planted Pasture) — Skole-Tucker Deforestation Map

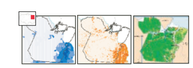
In areas where there is no census data like Colombia (right), we can use the basin-wide relationship between agriculture and land cover to estimate the actual amount of agriculture. Since the land cover classification is globally available and is used to derive predictor variables, we can estimate agriculture where there is no census data to guide us.





Skole-Tucker (Cerrado is beige) — Land Use (Total Agriculture)

Most deforestation maps do not specify the land use of areas of converted Cerrado. Cerrado, a shrubland formerly found in much of the eastern Brazilian Amazon, has been very highly converted to cropland and pasture. This map allows us to estimate the amount of agriculture in areas typically classified as merely "Cerrado".



Agriculture Census — Land Use — Skole-Tucker Deforestation

In areas where census data exists but is weak, the relationship built using data from across Amazonia can fill in deficiencies. In Para, for example, the Transamazon Highway is not seen in the census data, though it is evident on the new land use map as well as the deforestation imagery.

5- CONCLUSION

This map represents a new blend of ground-based and satellite-based spatially explicit data. Although this method was derived for Amazonia, we believe it is also widely applicable for those who seek to merge these two usually disparate data sources. The results indicate that the map has spatial detail near to that provided by satellite information, yet includes key attribute data available only in agricultural censuses.

It is our hope that this data set will be widely adopted, and that the technique will prove useful to other researchers. We plan to make the data available at our web site at http://sage.aos.wisc.edu.