Farming the Planet. Part 1: The Geographic Distribution of Global Agricultural Lands in the Year 2000

Navin Ramankutty[1], Amato T. Evan[2], Chad Monfreda[3], Jonathan A. Foley[3]

Corresponding Author: Navin Ramankutty; Email: navin.ramankutty@mcgill.ca

[1]Department of Geography and Earth System Science Program; McGill University; 805 Sherbrooke St. W.; Montreal, QC H3A 2K6; Canada.

[2]Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison, 1225 W. Dayton St., Madison, WI, 53706, USA

[3]Center for Sustainability and the Global Environment (SAGE), Nelson Institute for Environmental Studies, University of Wisconsin - Madison, 1710 University Avenue, Madison WI, 53726, USA

Submitted to *Global Biogeochemical Cycles*, Feb 2007

Running Title: global agricultural lands in 2000

Keywords: land-use, land-cover, cropland, pasture, agriculture, global, data

ABSTRACT

Agricultural activities have dramatically altered our planet's land cover. To understand the extent and spatial distribution of these changes, we have developed a new global data set of croplands and pastures *ca*. 2000 by combining national and sub-national agricultural inventory data and satellite–derived land cover data. The agricultural inventory data, with much greater spatial detail than previously available, is used to train a land cover classification data set obtained by merging two different satellite-derived products. By utilizing the agreement and disagreement between Boston University's MODIS global land cover product and the GLC2000 data set, we are able to predict the spatial pattern of agricultural land better than by using either data set alone. We present a new global 5 min (~10 km) resolution cropland and pasture dataset for the Year 2000 that is of greater accuracy than previously available, and for the first time, statistical confidence intervals on these estimates.

1. Introduction

Human land use activities are a force of global significance [*Foley, et al.*, 2005]. Humans have extensively modified the Earth's land surface, altering ecosystem structure and functioning, and diminishing the ability of ecosystems to continue providing valuable resources such as food, freshwater and forest resources, and services such as regulation of climate, air quality, water quality, soil resources.

Agricultural activities, in particular, have been responsible for a vast majority of these land-use related ecosystem consequences [*Richards*, 1990; *Tilman, et al.*, 2001; *Green, et al.*, 2005]. Nearly 40% of the planet's ice-free land surface is now being used for agriculture, and much of this land has replaced forests, savannas, and grasslands [*Foley, et al.*, 2005]. Clearing of tropical forests for cultivation or grazing is responsible for ~12-26% of the total emissions of carbon dioxide to the atmosphere [*DeFries and Achard*, 2002; *Houghton*, 2003], and land use changes can significantly modify regional and global climate [*Pitman, et al.*, 1999; *Pielke, et al.*, 2002]. Furthermore, ~20-30% of the total available surface water on the planet is withdrawn for irrigation [*Cassman and Wood*, 2005], and nitrogen fixation through fertilizer production and crop cultivation currently equals or even exceeds natural biotic fixation [*Galloway, et al.*, 1995; *Smil*, 1999].

As such, agriculture is partly or wholly responsible for environmental concerns such as tropical deforestation and biodiversity loss, fragmentation and loss of habitats, emissions of important greenhouse gases, losses of soil quality through erosion and salinization, decreases in quantity and quality of water resources, alteration of regional climates, reduction in air quality, and increases in infectious diseases [*Foley, et al.*, 2005]. On the other hand, agricultural expansion and intensification has provided a crucial service to humanity by meeting the food demands of a rapidly growing population [*Cassman and Wood*, 2005], and thereby involves a trade-off between food production and environmental deterioration [*DeFries, et al.*, 2004][*Foley et al.*, 2005].

In order to assess the Earth system consequences of agriculture – both the positive social and economic benefits and the often negative environmental consequences -- it is essential to develop global data sets of the geographic distribution of agricultural land use and land cover change [e.g., *Wood, et al.*, 2000; *Bauer, et al.*, 2003; *Donner and Kucharik*, 2003; *Cassman and Wood*, 2005]. Recent advances have led to the emergence of new continental-to-global scale data sets of agricultural land cover, developed by merging satellite-derived land cover data sets and ground-based agricultural inventory data sets [*Ramankutty and Foley*, 1998; *Frolking, et al.*, 1999; *Ramankutty and Foley*, 1999; *Hurtt, et al.*, 2001; *Klein Goldewijk*, 2001; *Cardille, et al.*, 2002; *Frolking, et al.*, 2002; *Cardille and Foley*, 2003; *Donner*, 2003; *Leff, et al.*, 2004; *Ramankutty*, 2004].

Our earlier work, in particular, pioneered the development of a statistical “data-fusion” technique to merge a satellite-derived, global, 1-km resolution land-cover data set, with ground-based national and sub-national cropland inventory statistics, to develop global maps of the world’s croplands in the early 1990s [*Ramankutty and Foley*, 1998], and their historical changes since the year 1700 [*Ramankutty and Foley*, 1999]. These data sets have been widely used by the global change community and have been employed in various analysis and assessments, including analysis of regional food security [*Ramankutty, et al.*, 2002b], an assessment of the regions of the world undergoing the most rapid land-cover changes over the last decade [*Lepers, et al.*, 2005], global carbon cycle modeling [*McGuire, et al.*, 2001], global climate modeling [*Bonan*, 1999; *Brovkin, et al.*, 1999; *Bonan*, 2001; *Myhre and Myhre*, 2003; *Brovkin, et al.*, 2006], estimation of global soil erosion [*Yang, et al.*, 2003], and as input to global economic models [*Lee*, 2005; *Ramankutty, et al.*, in press]. These data have also provided the essential information on historical croplands for other global land use/cover data sets [*Hurtt, et al.*, 2006; *Wang, et al.*, 2006].

In this paper, we present a critical update to our global agricultural land cover data sets. In particular, we present new global data sets for the year 2000, developed using an order-of-magnitude enrichment of our agricultural inventory data, a combination of two different satellite-derived global land cover data sets for year 2000, and improved methods to merge the satellite data and inventory data. We also present, in addition to an updated map of global croplands for the year 2000, a new map of global pastures, as well as estimated confidence intervals for both of these data sets. These new data sets will form valuable products for the global environmental change community.

2. Data Sets

In this section, we describe how we compiled the two different sources of information used in this study: 1) Satellite-based global land cover classification data sets; and 2) Ground-based agricultural census/inventory data sets.

2.1. Satellite data sets

We used two different high resolution (1-km) satellite-based, global land cover classification data sets that are available for the year 2000: Boston University's MODIS (MODerate resolution Imaging Spectrometer) -based global land cover product [*Friedl, et al.*, 2002] [BU-MODIS hereafter], and the Satellite Pour l'Observation de la Terre (SPOT) VEGETATION based Global Land Cover 2000 (GLC2000) data set [*Bartholome and Belward*, 2005]. The BU-MODIS land cover product used data acquired from 15 October 2000 to 15 October 2001 to derive 17 land cover classes using a supervised classification scheme (see Table 1 for legend). The GLC2000 data set utilized data acquired from 1 November 1999 to 31 December 2000 to derive 22 global land cover classes based on a flexible classification system that is determined by regional institutions (Table 1).

We applied a simple set of climatic parameters to mask obviously non-agriculture areas within the satellite data sets, else we obtain some spurious results in the Northern Hemisphere high latitudes. This mask included all regions north of 50°N with Growing Degree Day (GDD; base 5°C) less than 1000 days. GDD data were calculated according to Ramankutty *et al*. [2002a], and interpolated to 1km resolution. We also masked out protected / minimal use areas in the central part of the Australian continent [*Bureau of Rural Sciences*, 2000], which otherwise gets classified entire as pasture.

2.2. Agricultural inventory data

We extensively compiled cropland and pasture inventory data for the globe at the national and sub-national level (Figs. 1a-c, Table 2, more details in online supplement) for *ca*. year 2000. We compiled data for 15,990 different administrative units of the world -- ranging from poltical units like countries, states and counties -- which represents a 46-fold improvement in the richness of our inventory data compared to our previous effort [348 units in *Ramankutty and Foley*, 1998]. For 57 countries, we compiled census data at the sub-national level (e.g., "Level 1" indicating states in the U.S. or India, provinces in Canada or Argentina, departments in Bolivia or Columbia, etc., and "Level 2" indicating smaller units like U.S. counties, Brazilian municipios, or Indian districts). For 159 countries, we used national-level statistics from the Food and Agriculture Organization's (FAO) FAOSTAT database (http://faostat.fao.org); for these 159 countries, we calculated an average around the year 2000 using data from 1998 to 2002. For another 19 countries in our database, no FAOSTAT data was available, and we set the data to be missing.

We compiled the cropland and pasture data to be consistent with the FAO definition of "Arable lands and permanent crops" and "Permanent pastures" respectively. Arable land is defined by FAO (http://faostat.fao.org/site/375/default.aspx) as including "land under temporary crops (double-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market and kitchen gardens and land temporarily fallow (less than five years). The abandoned land resulting from shifting cultivation is not included in this category. Data for arable land are not meant to indicate the amount of land that is potentially cultivable."

Permanent crops are defined as “land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest, such as cocoa, coffee and rubber; this category includes land under flowering shrubs, fruit trees, nut trees and vines, but excludes land under trees grown for wood or timber.” Permanent pastures are defined as “land used permanently (five years or more) for herbaceous forage crops, either cultivated or growing wild (wild prairie or grazing land). The dividing line between this category and the category ‘Forests and woodland’; is rather indefinite, especially in the case of shrubs, savannah, etc., which may have been reported under either of these two categories.”

The agricultural inventory data were seldom available exactly for the year 2000 because agricultural censuses are only taken every 5-10 years in most industrialized nations of the world, and less frequently in other countries. We collected inventory data between the years 1998 to 2002 where possible, but in several instances we relied on older data (see Table 2). For nations where the inventory data did not fall within the 1998-2002 period, or where cropland or pasture data were unavailable but some proxies were available (such as harvested area of individual crops or heads of livestock), we estimated cropland and pasture data for *ca.* 2000 by calibrating the available information to national totals from FAOSTAT (see online supplement for details).

The quality of our census data is varied. Some regions of the world are not well represented in terms of the resolution of inventory data, with the African continent and the Former Soviet Union being the most underrepresented. Sometimes, the national-level census statistics were inconsistent with the FAOSTAT data; in such cases we mostly relied on the national statistics (as recommended by FAO), except for a few cases where we believed that FAOSTAT data was more reliable (see online supplement for details). Inconsistencies were a result of either unclear definitions of the category, or sometimes poor reporting by the national statistics agency. For example, the cropland census statistics for China has been noted by various studies to be particularly problematic [*Crook*, 1993; *Frolking, et al.*, 1999; *Heilig*, 1999; *Seto, et al.*, 2000]. Here we have used the data for China from the study of Verburg and Chen [2000], which seems reliable. The definition of permanent pasture is particularly problematic, as acknowledged by the FAO (http://faostat.fao.org/site/375/default.aspx). Some countries (e.g., the USA) clearly distinguish between grassland pasture and range and forest-use land grazed, while most countries do not. So it is not clear whether grazed forest land or semi-arid grazing is included under the definition of pastures. One egregious example is Saudi Arabia for which FAOSTAT reports 1.7 million km^2 of permanent pasture in 2000, which is 80% of its total land area. However, most of Saudi Arabia is arid land and it is clear that much of the nomadic grazing areas are included under pasture. The Saudi Arabian sub-national census data we obtained reports 486 km^2 of pasture, which is 3500 times smaller; we have chosen to rely on this lower value. Similarly, FAOSTAT reports 4 million km^2 of pasture in China. However, Verburg *et al*. [2000] report a total of only 2.6 million km^2 of grassland in China. In this case, we have used the FAOSTAT value for the total amount of pasture in the country (but our final estimates for China are lower, as we shall discuss later). Similar inconsistencies exist in other countries and are reported in the online supplement.

In summary, our agricultural inventory database yields a global total of 15 million km^2 of cropland and 31.5 million km^2 of pasture. This compares to 15.3 million km^2 of cropland and 34.4 million km^2 of pasture reported by FAOSTAT; significantly, our census data compilation and interpretation yields about 8% less pasture than FAOSTAT. The difference between FAOSTAT and our own inventory for pasture can mainly be explained by differences in Saudi Arabia (as described earlier), Australia, Nigeria, Brazil, Mexico, Indonesia, Argentina, Colombia, Russia, and Spain. There were very few countries where the national census pasture area was greater than FAOSTAT; the one exception is Iran, for which FAOSTAT reports 0.4 million km^2 of pasture while the national census reported 0.9 million km^2. For croplands, while the global total areas are comparable between FAOSTAT and our inventory, there are significant national differences. For example, the national inventories of Australia, Brazil, Canada, China and Turkey report lower cropland area compared to FAOSTAT, while Iran, Argentina, Nigeria, Mexico, and Indonesia report greater cropland area.

3. Methodology

The basic methodology for creating the new cropland and pasture data sets originated from our earlier work [*Ramankutty and Foley*, 1998], where we calibrated a high-resolution satellite-derived land-cover data set against agricultural inventory data to derive a global map of croplands for 1992. In this paper, however, we updated the methodology in three important ways: (1) Instead of a single satellite-derived land cover data set, here we used a merger of two different satellite data sets; (2) Instead of calibrating only the *a priori* identified agricultural land-cover classes against inventory data [as in *Ramankutty and Foley*, 1998], we utilized *all* the land-cover classes in our training procedure [as in *Hurtt, et al.*, 2001; *Cardille, et al.*, 2002]; and (3) Because we had much higher-resolution census data compared to our previous efforts, we considered the census data sets to represent an approximate "truth", and used the 2-step method developed by Ramankutty [2004] whereby the satellite data is used to spatially locate agricultural lands within an administrative unit, but the total area of agricultural land in the administrative unit is derived from the census data (with some exceptions as described later). The following section provides a detailed description of the steps taken to create the final data set (see figure 2 for a flowchart of our algorithm).

3.1. Combining the two satellite-derived land-cover classification data sets

Giri *et al*. [2005] compared the GLC2000 and BU-MODIS land cover data sets and found some consistency at the aggregate class level, but widespread disparities in the details. The GLC2000 data set was developed using SPOT vegetation data with the assistance of regional experts and used a flexible classification scheme. The BU-MODIS data set, on the other hand, was developed using a globally-consistent procedure with a fixed classification scheme, but without regional expertise. The two data sets therefore bring different kinds of expertise and information that are potentially complimentary. Here we decided to capitalize on both by combining them into a single land cover data set. To do so, we overlaid the BU-MODIS and GLC2000 data sets and developed new land-cover classes that contain all possible combinations of the individual land-cover classes (see Table 3 for an example). The resulting combined land-cover data set consists of 391 possible land cover types and is shown later to provide more accurate results when calibrated against the inventory data, as opposed to using either data set individually.

3.2. Step 1: Calibrating the satellite data sets against the agricultural inventory data

We followed methods developed in previous efforts [*Ramankutty and Foley*, 1998; *Cardille, et al.*, 2002; *Ramankutty*, 2004] to merge the satellite data sets and the agricultural inventory data to develop global maps of croplands and pastures for the year 2000. As mentioned earlier, our statistical data fusion procedure is different from Ramankutty and Foley [1998] in a couple of major ways. First, to allow for potential misclassifications in the satellite-derived land-cover data sets [e.g, *Hurtt, et al.*, 2001; *Cardille, et al.*, 2002], and also because pasture is not explicitly identified as a land-cover class by the satellite-derived data sets, we utilized all 391 land-cover classes in our training procedure, as described below. Second, we used the two-step procedure of Ramankutty [2004] which assumes that the inventory data is the "truth" (except for identified outliers), and uses the satellite data to spatially disaggregate the census data within each administrative unit.

For each administrative unit, i, the proportion of cropland and pasture area from the inventory data, cf_i and pf_i respectively, was calculated by dividing the inventory cropland and pasture areas by the total land area (A_i) for the administrative unit. Then, we determined $\lambda_{j,i}$, which is the proportion of each of the 391 satellite-derived land-cover classes, j, within administrative unit i.

We formulated a linear model relating the satellite-derived data sets to the agricultural inventory data, as follows:

$$cf_i = \sum_{j=1}^{n_\lambda} \left(\alpha_j \times \lambda_{j,i} \right) + \varepsilon cf_i , \qquad (1)$$

and

$$pf_i = \sum_{j=1}^{n_\lambda} \left(\beta_j \times \lambda_{j,i}\right) + \varepsilon pf_i, \tag{2}$$

where α and β are unknown parameters associated with each land-cover category j, n_λ is the number of land-cover categories, and εcf_i and εpf_i are error terms that represent the residual difference between the inventory and linear-model predicted cropland and pasture proportions.

Additionally, (1) and (2) were subject to the following constraints,

$$\begin{aligned} &0 \le \alpha_i \le 1, \text{ and} \\ &0 \le \beta_i \le 1, \text{ and} \\ &\alpha_i + \beta_i \le 1. \end{aligned} \tag{3}$$

These constraints ensured that the cropland or pasture proportions in any pixel (when the model is later applied at pixel level) will be between 0 and 100%, and that the sum of cropland and pasture proportions will be less than 100%.

We used a least-squares minimization method to solve for the parameters α and β. In particular, we specified the weighted least squares error (*LSE*) to be minimized as:

$$LSE = \sum_{i=1}^{n_i} \omega_i \left[(\varepsilon cf_i)^2 + (\varepsilon pf_i)^2 \right], \tag{4}$$

where n_i is the number of administrative units, and ω_i is a term that weights the residuals εcf_i and εpf_i by the land area, A_i, of each administrative unit, normalized by the maximum value, i.e.,

$$\omega_i = \frac{A_i}{\max(A_i)}. \quad (5)$$

3.2.1. Implementation

To estimate the parameters in equations 1 and 2, we used a multiple linear regression model from the STARPAC package (http://www.cisl.ucar.edu/softlib/STARPAC.html). We developed three separate models – first using the BU-MODIS and GLC2000 data sets separately, and then using the combined land cover data set. In each case, we started with a complete model specifying all the land cover classes as potentially being cropland or pasture; the following classes: BU13 (urban), BU15 (snow & ice), BU16 (barren), BU17 (no data), GLC19 (bare), GLC20 (water), GLC21 (snow & ice), GLC22 (artificial surfaces), and GLC23 (no data), and their combinations in our combined land cover data set, were left out of the model. We then used stepwise regression using backward selection to estimate the parameters. The details of this procedure are outlined in the online supplement.

We applied our optimization procedure separately to six different regions of the world (Fig. S1 in online supplement), similar to Ramankutty and Foley [1998]. These six regions were a compromise between selecting small enough regions with similar agricultural characteristics, but large enough regions to have enough observations within each to obtain robust parameter estimates. Clearly some of the regions extend across different types of agricultural land uses, but subdividing the world into smaller regions resulted in too few observations in some regions to get robust model estimates.

The estimated parameter values (not shown) were used to make global cropland and pasture maps at 5 min spatial resolution in latitude by longitude for the BU-MODIS and GLC2000 data sets individually, and for the combined land-cover data set (Fig. S2 in online supplement). For each 5 min grid cell, x, y (latitude by longitude), we determined $\lambda_{j,x,y}$, which is the proportion of each of the satellite-derived land-cover classes, j, within 5 min grid cell x, y. Using the estimates of α and β, we calculated cropland and pasture proportions in each 5 min grid cell, $cf_{x,y}$ and $pf_{x,y}$ respectively, using:

$$cf_{x,y} = \sum_{j=1}^{n_\lambda} \left(\alpha_j \times \lambda_{j,x,y} \right) \tag{6}$$

and

$$pf_{x,y} = \sum_{j=1}^{n_\lambda} \left(\beta_j \times \lambda_{j,x,y} \right). \tag{7}$$

3.4. Comparison of the performance of BU-MODIS, GLC2000, and combined data sets

We now present a comparison of the inventory agricultural land area to the predicted values for each data set from calibration Step 1 (from equations 6 and 7, aggregated from 5 min resolution to the administrative level) (Table 4). The cropland and pasture areas predicted using the combined satellite data set is better correlated to the inventory data compared to the models using either the BU-MODIS or GLC2000 data sets alone in every region except for a few exceptions (croplands in Africa, and pastures in Europe & Russia and Australia). Further, neither the BU-MODIS nor the GLC2000 data set always performs better than the other; for example, in *Africa & Middle East*, GLC2000 predicts cropland much better than BU-MODIS, but the reverse is true for pastures. We next present a regional comparison of the inventory and predicted agricultural land areas from the combined data set (Fig. S3 in online supplement). The most notable differences lie in the underestimation of pasture areas in Asia and Australia/New Zealand. We will discuss the significance of this underestimation in the next section.

We now consider the spatial patterns of predicted agricultural land from the combined data set versus the individual data sets, using South American pastures as an example (Fig. S4 in online supplement). In Figure S4a, the inventory data shows large extent of pasture along the arc of deforestation, along coastal and southern Brazil, in Uruguay, and in the Patagonia region of Argentina. Figures S4b and S4c are the BU-MODIS and GLC2000 based pasture maps. Both do a reasonable job of reproducing the inventory data in Figure S4a. The most glaring difference is in the Nordeste (Northeast) region of Brazil, where both data sets overestimate the distribution of pasture as compared to the inventory data in Figure S4a. The combined land-cover data was able to best reproduce the inventory data, especially in the Brazilian Nordeste region (Fig. S4d).

Thus, by using the combined data set we are able to capitalize on whichever satellite-based land cover data set is best suited to each region. If we were to use either the BU-MODIS or GLC2000 data sets alone, we would get reasonably good global results, but would lose accuracy in some regions. The use of the combined land-cover data set especially yields marked improvements over Asia and South America (Table 4). Therefore, we use the combined data set in the next step of this study.

3.4. Bootstrap procedure to estimate uncertainly

We further used a bootstrap technique in order to estimate the uncertainty in our parameter estimates. This procedure was applied at this stage for only the combined land-cover data set. We performed 1000 bootstrap runs, where the census data was sampled with replacement each time, and re-estimated our regression model using the combined satellite-based land-cover data set. This yielded a distribution of values for each parameter α and β, and we report the mean and 90% confidence intervals here (see Table S1 in online supplement).

Using the 1000 estimates of α and β, we calculated 1000 estimates of cropland and pasture proportions in each 5 min grid cell, $cf_{x,y}$ and $pf_{x,y}$ respectively, using:

$$cf_{x,y}(i) = \sum_{j=1}^{n_\lambda}\left(\alpha_j(i) \times \lambda_{j,x,y}\right),\ i = 1,\ 1000 \qquad (8)$$

and

$$pf_{x,y}(i) = \sum_{j=1}^{n_\lambda}\left(\beta_j(i) \times \lambda_{j,x,y}\right),\ i = 1,\ 1000. \qquad (9)$$

From these 1000 estimates, we calculated the mean, 5th percentile and 95th percentile values of the cropland and pasture values, represented as $cf_{x,y}^{mean}$, $cf_{x,y}^{5th\%ile}$, and $cf_{x,y}^{95th\%ile}$, and as $pf_{x,y}^{mean}$, $pf_{x,y}^{5th\%ile}$, and $pf_{x,y}^{95th\%ile}$ respectively (figures not shown from this stage of analysis).

3.5. Step 2: Adjusting the predicted cropland and pasture data to match inventory data

In this final step, we followed the methods of Ramankutty [2004], to adjust our spatially-explicit predictions from Step 1 (the bootstrap model estimates using the combined data set) to match the inventory data at the administrative level where available. To do so, we first aggregated our 5 min resolution cropland and pasture data sets to the administrative level. We then compared them to the inventory data to derive a correction factor for each administrative unit. The correction factors were, however, constrained to be within a factor of 5 (i.e., to lie between 0.2 and 5.0) for administrative units that were considered outliers in the regression, thereby trusting the satellite-based land cover data more than the inventory data in those cases. Outliers were determined to be those administrative units with residuals (predicted cropland area from the Step 1 calibration procedure minus inventory cropland area) that were greater than 2 standard deviations from the mean. Correction factors were set to 1.0 for administrative units with missing data, thereby relying on the satellite-estimated spatial patterns from Step 1 in these units. We then applied Pycnophylactic Interpolation [*Tobler*, 1979] to the administrative-level correction factors to obtain a smooth surface of correction factors at 5 min resolution (without this smoothing, artificial boundaries between administrative units might appear in the final product; note, however, that only the correction factors were smoothed, so any real boundaries in the original satellite data will remain). The spatial correction factors were then applied to our results from Step 1 to derive our final maps of cropland and pastures at 5 min resolution (Fig. 3) and respective confidence intervals (Fig. 4). The final equations for cropland and pasture proportions can be represented as:

$$\begin{aligned} Cropland_{x,y}^{mean} &= \mu cf_{x,y}\ x\ cf_{x,y}^{mean} \\ Cropland_{x,y}^{5th\%ile} &= \mu cf_{x,y}\ x\ cf_{x,y}^{5th\%ile}, \\ Cropland_{x,y}^{95th\%ile} &= \mu cf_{x,y}\ x\ cf_{x,y}^{95th\%ile} \end{aligned} \tag{10}$$

and,

$$\begin{aligned} Pasture_{x,y}^{mean} &= \mu pf_{x,y}\ x\ pf_{x,y}^{mean} \\ Pasture_{x,y}^{5th\%ile} &= \mu pf_{x,y}\ x\ pf_{x,y}^{5th\%ile}, \\ Pasture_{x,y}^{95th\%ile} &= \mu pf_{x,y}\ x\ pf_{x,y}^{95th\%ile} \end{aligned} \tag{11}$$

where $\mu cf_{x,y}$ and $\mu pf_{x,y}$ are the spatially-explicit correction factors for cropland and pasture respectively.

4. Results

4.1. Total global area of croplands and pastures in 2000

Our final results (Fig. 3) indicate that there were 15.1 million km^2 of cropland and 28.3 million km^2 of pasture in the world in the Year 2000. This compares to 15.3 million km^2 of cropland and 34.4 million km^2 of pasture reported by the FAOSTAT database. Thus we predict significantly lower extent of pasture (by 6.1 million km^2 or ~18% lower) than reported by FAO.

Our own inventory data reports 15.0 million km^2 of cropland and 31.5 million km^2 of pasture. Thus our inventory data for pasture is already lower than FAO statistics; this difference was explained earlier in section 2.2.

Our final predicted pasture area is even lower than our inventory data, especially in Asia and Australia/New Zealand (as already evidenced in Fig. S3 in online supplement). We anticipated that this problem would be overcome in Step 2, when we scaled our spatial cropland and pasture data to match our inventory data, but despite this our final predicted extent of pasture differs significantly from our inventory data. This is because we did not allow pixels with predicted agriculture proportions of 0% in Step 1 to be scaled, and each pixel was limited to have a maximum agriculture proportion of 100%. The biggest difference is in China, where the inventory reports 4 million km^2 of pasture, but we estimate a final area of only 2.9 million km^2. Our estimate, however, is similar to the 2.6 million km^2 of grassland reported by Verburg *et al.* [2000]. It is likely that grazed forestland and semi-arid land are included under the pasture category in the Chinese inventory. Other big differences in pasture area arise in Australia (inventory area of 3.2 million km^2 versus our prediction of 2.7 million km^2), Mongolia (inventory of 1.3 million km^2 versus our estimate of 0.9 million km^2), Mauritania (inventory of 0.4 million km^2 versus our 0.1 million km^2), Iran (inventory of 0.9 million km^2 versus our 0.6 million km^2), and the U.S.A. (inventory of 2.3 million km^2 versus our 2.1 million km^2). It is interesting to note that all of these countries have significant amount of semi-arid land. For croplands, while the global totals agree, there are compensating national-level differences but these differences are less significant compared to the differences in pasture areas.

Our final predicted cropland extent of 15.1 million km^2 in Year 2000 amounts to roughly 12% of the global land area (excluding Greenland and Antarctica), and pasture extent of 28.3 million km^2 amounts to 22% of global land area. Thus, humans are using 34% of the global land area for their agricultural needs.

4.2. The geographic distribution of croplands and pastures in Year 2000

We analyzed the distribution of agricultural regions by 14 different regions of the world (Fig. 5a). The greatest proportion of croplands in the world are found in South Asia (39%), Europe (27%) and USA-East of the Mississippi (23%), while the greatest proportion of pastures are found in Argentina, Uruguay & Chile (33%), Pacific Developed countries (33%), China (33%), Mexico & Central America (31%), USA-West of Mississippi (31%), and Tropical Africa (30%). The smallest proportion of croplands are found in Canada, the Pacific Developed Countries, and northern South America, while the smallest proportion of pastures are found in Southeast Asia, Canada, and USA-East of the Mississippi.

We also examined which potential natural vegetation types of the world have been most affected by agriculture (Fig. 5b). We overlaid our agricultural maps over the global map of potential natural vegetation developed by Ramankutty and Foley [1999]. We find that croplands have mostly replaced temperate deciduous forests (in Europe and eastern US), and tropical deciduous forests (in South Asia), while, pastures have mostly replaced grasslands, savannas, and shrubland. Roughly 30% of temperate deciduous forests have been converted to cropland, while 50% of grasslands have been converted to pasture. However, this global picture varies regionally (Fig. 6). While forests have been cleared for croplands predominantly in Asia, in

North America, Africa, and the Former Soviet Union, a substantial amount of savanna/grasslands have been converted to croplands (Fig. 6a). Also, a significant amount of forests in South America have been cleared for pastures, even though most pastures have replaced savanna/grasslands (Fig. 6b).

Next, we examined the amount of spatial overlap between croplands and pastures (Fig. 7). Our analysis shows that cropland and pastures are distinct biomes over much of our planet's land surface. The regions of the world with significant overlap lie along the western edge of cultivation in the Midwestern U.S. and in Texas, northeast Brazil, in parts of West Africa and eastern China, and scattered regions elsewhere. This is not to say that grazing does not occur in cultivated regions of the world – many regions of the world are characterized by multi-functional agricultural lands, subject to different uses during different parts of the year (e.g., grazing occurs following the harvest of a crop) [*Reenberg and Fog*, 1995]. Therefore, this lack of overlap partly reflects the inability of global monitoring systems, including satellite data and agriculture inventory data, to characterize multiple uses, and land is often classified as a single category. Our final estimates likely underestimate the real overlap between cultivation and grazing, especially the multiple uses that occur within a year.

We also investigated the frequency distribution of croplands and pastures globally, and across the 14 different regions of the world. We calculated probability distributions using all pixels with non-zero cropland or pasture values, and estimated what proportion of total cropland or pasture area can be attributed to different categories of cropland or pasture proportions (Fig. 8).

Globally, we find that a large proportion of the total cropland area comes from land that is between 60% and 80% cultivated. This pattern is observed in most regions of the world, except in northern South America, Tropical Africa, and North Africa and Middle East (where larger number of pixels with low proportion of cultivation contribute most to the total cropland area, although in the latter two regions, there are a few pockets of greater than 90% cropland). With pastures, the total area is dominated by pixels with greater than 90% pasture. This pattern is true over most regions of the world, but is reversed in the eastern USA, Europe, and Southeast Asia (and South Asia, to a lesser extent). However, it is to be noted that the proportion of pasture in a grid cell provides no information on the grazing intensity in that grid cell; an area of grassland with very low stocking density of livestock would have the same proportion of pasture as one with very high stocking densities.

4.3. Evaluation against other independent data sets

While there are no consistent, global spatial data sets of agriculture in the Year 2000 to evaluate our products against, there are numerous regional products against which we can compare our global products (see online supplement for figures). While many of these regional data are not for the Year 2000 (and are sometimes a decade older), they nevertheless provide an independent measure of the large-scale spatial patterns of agriculture in these regions. These data were not always available in a consistent digital format -- they were often in vector format that is difficult to quantitatively compare to our raster data, and sometimes only available as images in publications. Therefore, rather than making quantitative comparisons, we present detailed regional maps from our data set compared to these independent regional maps, as online supplements.

Our regional comparisons are mostly for croplands because there are few spatial data sets depicting pastures (likely because it is difficult to distinguish between natural and grazed grasslands). Visual comparison (online supplement) suggests that our distribution of croplands is reasonable in North America (USA, Canada, Mexico), and so is the distribution of pastures in the USA (except for noticeable differences in northeastern Texas, and eastern Oklahoma). In Brazil, our distribution of both croplands and pastures shows similar geographic patterns but greater intensity compared to Cardille *et al*. [2002]. This may be a result of our data being representative of 2000, while Cardille *et al*. [2002] data is for the mid 1990s – changes in the southern Amazon are rapid -- or because of differences in statistical methods with the regression tree method of Cardille *et al*. [2002] not able to deal well with extremes. In China also, our geographic patterns of croplands matches well with that of Liu *et al*. [2005], but our intensity is lower. In West Africa, our cropland distribution from this study is a significant improvement over our earlier work [Ramankutty *et al*., 1998], but consistent with our more recent effort [*Ramankutty*, 2004]. Our distribution of croplands compares reasonably well with Africover data in East Africa, but is not a significant improvement over our earlier effort, while in South Africa, our estimated patterns are an improvement over our previous effort but of much lower intensity. In Europe, our data set of croplands compares well to the CORINE land cover database. We seem to underestimate croplands in eastern Australia, while our distribution of pastures is significantly underestimated in the Northern Territory. The problem with pastures in Australia actually emerges from Step 2, with the census data in the Northern Territory suggesting very little pastures; our predicted pattern from Step 1 compares better to the data from the Australian Natural Resources Atlas.

5. Discussion and Conclusions

We have merged two different satellite-based land cover classification data sets with an extensive compilation of national and subnational-level agricultural census statistics to develop global maps of croplands and pastures for the Year 2000. These maps form the first comprehensive characterization of the distribution of global agricultural lands in the Year 2000, describing the spatial extent of croplands and pastures within 5 min resolution (~10km) grid cells in latitude by longitude; in addition, 90% confident intervals on the mean estimates are being presented for the first time. In creating these new data sets, we have built on our earlier work in 1998 [*Ramankutty and Foley*, 1998], where we developed a statistical "data fusion" method to merge satellite data and census data to map global croplands in 1992. Here we have improved our statistical methods, brought in two new satellite data sets, and enriched our agricultural census data to update our global croplands map to the Year 2000, and to create a new global pasture data set for 2000. We would like to strongly caution that our two croplands maps from 1992 and 2000 cannot be directly compared to detect changes over that time period – changes in methodology and data sources preclude such a direct comparison. We plan to develop a consistent historical time series of croplands and pastures in the future.

The global area of cropland from this study for the Year 2000 of 15.1 million km^2 is smaller than the area of 18 million km^2 for 1992 estimated by our earlier study [*Ramankutty and Foley*, 1998]. This is not a real decrease, but rather an artifact of change in methodology between the two studies. In this study, we used our step 2 to scale the calibrated cropland patterns to identically match the census data. In our earlier study, our final result was directly out of the calibration and we did not do any scaling. We have changed our philosophy here to trust the total area of agricultural land reported by the census data (unless we have strong reason to suspect their quality as discussed earlier, or unless they are outliers in the step 1 calibration procedure), and use the satellite data sets for information on the spatial distribution within each census administrative unit.

The global area of pasture of 28.3 million km^2 is 18% lower than the standard FAOSTAT estimate of 34.4 million km^2. The major differences are found in Saudi Arabia, Australia, China, and Mongolia. This is likely because the census data on pasture reported to FAOSTAT include grazed forestland and semi-arid land. The definition of pasture has always been problematic, as acknowledged by FAOSTAT (http://faostat.fao.org/site/375/default.aspx), and one way to improve the situation in the future may be to develop global maps of livestock density [e.g., *Kruska, et al.*, 2003; *Wint*, 2005], and overlay that data with an independent global estimate of herbaceous vegetation.

Although we now have new global estimates of the distribution of cropland and pastures, several caveats need to be noted. First, there is much misunderstanding and confusion regarding the definitions of croplands and pastures. In this study, we have followed the FAO definition, as described earlier. For croplands, this includes temporary fallow lands (less than 5 years), which are not cultivated. It is not clear how strictly this restriction of less than 5 years was applied when accounting for fallow land. For example, the U.S. data on croplands used by FAOSTAT includes idled cropland, which includes land under the Conservation Reserve Program that amounts to roughly 9% of the total cropland area, and is often idled for longer than 5 years [*Lubowski, et al.*, 2002]. Secondly, in tropical nations characterized by extensive fallow cropping systems, such a definition may include much land that is not currently cultivated, and therefore portray a misleading picture of what may be commonly thought of as cropland.

The definition of pasture is subject to even greater uncertainty. Where is the dividing line between herding and grazing? For example, is reindeer herding reported under pastureland? Does the data represent both planted pastures and natural pastures? Is grazing underneath a forest cover or in semi-arid areas included in the pasture data? In other words, pasture is only a subset of the land (on herbaceous vegetative cover) that is used for grazing. The data on extent of pasture also says nothing about the intensity of grazing -- an acre of land with one cow and another acre with 10 cows would both be considered 1 acre of pastureland.

An additional concern related to the definition of croplands and pastures arises from the existence of multifunctional landscapes, as discussed earlier [*Reenberg and Fog*, 1995]. In some countries, especially in Asia and Africa, land is cropped for a while, and then after harvest, is grazed for the remainder of the year. Thus, during the year, the land is put to multiple uses, and it is not clear how to classify these lands, and how these lands were accounted for in the census statistics. Mixed-use classifications need to be used to characterize such landscapes rather than discrete classes such as cropland and pasture. It is not clear how much of the global agricultural land area is influenced by such multifunctional land use practices.

Finally, our synthesized data sets have uncertainty related to the fact that we are trying to merge two different observation systems – remote-sensing based and ground based. Remote sensing satellites can only observe land cover, i.e., the top of the vegetative canopy, and have little information on what happens below the canopy. The ground-based land use data, on the other hand, may include different information. For example, the cropland census data include permanent crops such as tree crops. It is not clear whether the remote sensing observations consider tree crops as tree cover or whether they classify them as cropland. Similarly, as discussed earlier, census data may not distinguish between grazing on grasslands, forests, and bare ground.

There is great demand by the global environmental change community to understand how global agricultural lands are changing and evaluate their implications for a sustainable future [e.g., *Tilman, et al.*, 2001; *Foley, et al.*, 2005]. Therefore, despite large uncertainties, we need to make progress toward developing new methods to characterize the spatial patterns of global agricultural lands. Here we have obtained the best available global data on agricultural lands and synthesized them to create a single homogeneous database of the world's croplands and pastures. We believe that these data sets would be enormously useful to at least two different communities of scientists/practitioners: (1) Global change scientists, interested in the consequences of global agriculture for climate, carbon cycle, water resources, etc., and would use these data sets for global scale analysis or as inputs to climate and ecosystem models [e.g., *McGuire, et al.*, 2001; *Myhre and Myhre*, 2003; *Jain and Yang*, 2005]; and (2) ecologists and conservation practitioners, interested in the role of agriculture in modifying natural ecosystems and habitats, in reducing biodiversity, and enhancing species extinction [e.g., *Green, et al.*, 2005; *Vandermeer, et al.*, 2005].

Acknowledgments

We are grateful to Peter Verburg from Wageningen University for sharing his China data. We would like to thank Prof. Mark Friedl from Boston University for providing feedback on early versions of this work. This work was accomplished through support from a NASA Interdisciplinary Science (IDS) grant, for which we are thankful.

References

Bartholome, E., and A. S. Belward (2005), GLC2000: a new approach to global land cover mapping from Earth Observation data, Int. J. of Remote Sens., 26, 1959-1977.

Bauer, E., M. Claussen, and V. Brovkin (2003), Assessing climate forcings of the Earth system for the past millennium, Geophys. Res. Lett., 30, 1276, doi:1210.1029/2002GL016639.

Bonan, G. B. (1999), Frost followed the plow: impacts of deforestation on the climate of the United States, Ecol. Appl., 9, 1305-1315.

Bonan, G. B. (2001), Observational evidence for reduction of daily maximum temperature by croplands in the Midwest United States, J. Clim., 14, 2430-2442.

Brovkin, V., M. Claussen, E. Driesschaert, T. Fichefet, D. Kicklighter, M. F. Loutre, H. D. Matthews, N. Ramankutty, M. Schaeffer, and A. Sokolov (2006), Biogeophysical effects of historical land cover changes simulated by six Earth system models of intermediate complexity, Clim. Dyn., 1-14; DOI: 10.1007/s00382-00005-00092-00386.

Brovkin, V., A. Ganopolski, M. Claussen, C. Kubatzki, and V. Petoukhov (1999), Modelling climate response to historical land cover change, Glob. Ecol. Biogeogr., 8, 509-517.

Bureau of Rural Sciences (2000), 1996/97 Land use of Australia (version 1), National Land & Water Resources Audit, Department of Agriculture, Fisheries and Forestry-Australia, http://www.nlwra.gov.au/, Canberra.

Cardille, J. A., and J. A. Foley (2003), Agricultural land-use change in Brazilian Amazonia between 1980 and 1995: Evidence from integrated satellite and census data, Remote Sensing of Environment, 87, 551–562.

Cardille, J. A., J. A. Foley, and M. H. Costa (2002), Characterizing patterns of agricultural land use in Amazonia by merging satellite classifications and census data, 16, 1045, doi:1010.1029/2000GB001386.

Cassman, K. G., and S. Wood (2005), Cultivated Systems, in Ecosystems and Human Well-being: Current State and Trends, Volume 1, edited by R. Hassan, et al., pp. 745-794, Island Press, Washington D.C.

Crook, F. (1993), Underreporting of China's cultivated land area: Implications for world agricultural trade, in China International Agriculture and Trade Report, edited, pp. 33-39, Situation and Outlook Series RS-93-4, U.S. Department of Agriculture, Washington, D.C.

DeFries, R., and F. Achard (2002), New estimates of tropical deforestation and terrestrial carbon fluxes: Results of two complementary studies, LUCC Newsletter, 8, 7-9.

DeFries, R. S., J. A. Foley, and G. P. Asner (2004), Land-use choices: balancing human needs and ecosystem function, Frontiers in Ecology and the Environment, 2, 249-257.

Donner, S. D. (2003), The impact of cropland cover on river nutrient levels in the Mississippi River Basin, Glob. Ecol. Biogeogr., 12, 341-355.

Donner, S. D., and C. J. Kucharik (2003), Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin, Glob. Biogeochem. Cycles, 17, 1085, doi:1010.1029/2001GB001808.

Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily, H. K. Gibbs, J. H. Helkowski, T. Holloway, E. A. Howard, C. J. Kucharik, C. Monfreda, J. A. Patz, I. C. Prentice, N. Ramankutty, and P. K. Snyder (2005), Global Consequences of Land Use, Science, 309, 570-574.

Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. Schaaf (2002), Global land cover mapping from MODIS: algorithms and early results, Remote Sens. Environ., 83, 287-302.

Frolking, S., J. J. Qiu, S. Boles, X. M. Xiao, J. Y. Liu, Y. H. Zhuang, C. S. Li, and X. G. Qin (2002), Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China, Glob. Biogeochem. Cycle, 16, art. no.-1091.

Frolking, S., X. M. Xiao, Y. H. Zhuang, W. Salas, and C. S. Li (1999), Agricultural land-use in China: a comparison of area estimates from ground-based census and satellite-borne remote sensing, Glob. Ecol. Biogeogr., 8, 407-416.

Galloway, J. N., W. H. Schlesinger, H. Levy, II, A. Michaels, and J. L. Schnoor (1995), Nitrogen fixation: Anthropogenic enhancement-environmental response, Glob. Biogeochem. Cycles, 9, 235-252.

Giri, C., Z. Zhu, and B. Reed (2005), A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets, Remote Sens. Environ., 94, 123-132.

Green, R. E., S. J. Cornell, J. P. W. Scharlemann, and A. Balmford (2005), Farming and the Fate of Wild Nature, Science, 307, 550, DOI: 510.1126/science.1106049.

Heilig, G. K. (1999), Can China feed itself? -- A system for evaluation of policy options, edited, International Institute for Applied Systems Analysis.

Houghton, R. A. (2003), Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000, Tellus B, 55, 378-390.

Hurtt, G. C., S. Frolking, M. G. Fearon, B. Moore, E. Shevliakova, S. Malyshev, S. W. Pacala, and R. A. Houghton (2006), The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands, Glob. Change Biol., 12, 1-22; doi: 10.1111/j.1365-2486.2006.01150.x.

Hurtt, G. C., L. Rosentrater, S. Frolking, and B. Moore (2001), Linking remote-sensing estimates of land cover and census statistics on land use to produce maps of land use of the conterminous United States, Global Biogeochemical Cycles, 15, 673-685.

Jain, A. K., and X. Yang (2005), Modeling the Effects of Two Different Land Cover Change Data Sets on the Carbon Stocks of Plants and Soils in Concert With CO2 and Climate Change, Glob. Biogeochem. Cycles, 19, GB2015, doi:2010.1029/2004GB002349.

Klein Goldewijk, K. (2001), Estimating global land use change over the past 300 years: The HYDE Database, Global Biogeochemical Cycles, 15, 417-433.

Kruska, R. L., R. S. Reid, P. K. Thornton, N. Henninger, and P. M. Kristjanson (2003), Mapping livestock-oriented agricultural production systems for the developing world, Agricultural Systems, 77, 39-63.

Lee, H.-L. (2005), The GTAP Land Use Data Base and the GTAPE-AEZ Model: Incorporating Agro-Ecologically Zoned Land Use Data and Land-based Greenhouse Gases Emissions into the GTAP Framework, edited.

Leff, B., N. Ramankutty, and J. Foley (2004), Geographic distribution of major crops across the world, Global Biogeochemical Cycles, 18, GB1009, doi:1010.1029/2003GB002108.

Lepers, E., E. F. Lambin, A. C. Janetos, R. DeFries, F. Achard, N. Ramankutty, and R. J. Scholes (2005), A synthesis of information on rapid land-cover change for the period 1981-2000, Bioscience, 55, 115-124.

Liu, J., M. Liu, H. Q. Tian, D. Zhuang, Z. Zhang, W. Zhang, X. Tang, and X. Deng (2005), Spatial and temporal patterns of China's cropland during 1990-2000: An analysis based on Landsat TM data, Remote Sensing of Environment, 98, 442-456.

Lubowski, R. N., M. Vesterby, S. Bucholtz, A. Baez, and M. J. Roberts (2002), Major Land Uses in the United States, 2002, Economic Information Bulletin No. (EIB-14), 54 pp, U. S. Department of Agriculture, Econ. Res. Serv., Washington, D.C.

McGuire, A. D., S. Sitch, J. S. Clein, R. Dargaville, G. Esser, J. Foley, M. Heimann, F. Joos, J. Kaplan, D. W. Kicklighter, R. A. Meier, J. M. Melillo, B. Moore, I. C. Prentice, N. Ramankutty, T. Reichenau, A. Schloss, H. Tian, L. J. Williams, and U. Wittenberg (2001), Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO2, climate and land use effects with four process-based ecosystem models, Global Biogeochemical Cycles, 15, 183-206.

Myhre, G., and A. Myhre (2003), Uncertainties in radiative forcing due to surface albedo changes caused by land-use changes, J. Clim., 16, 1511-1524.

Pielke, R. A., G. Marland, R. A. Betts, T. N. Chase, J. L. Eastman, J. O. Niles, D. D. S. Niyogi, and S. W. Running (2002), The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases, Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences, 360, 1705-1719.

Pitman, A., R. Pielke Sr., R. Avissar, M. Claussen, J. Gash, and H. Dolman (1999), The role of the land surface in weather and climate: does the land surface matter?, IGBP Newsletter No. 39, 4-9.

Ramankutty, N. (2004), Croplands in West Africa: A geographically explicit dataset for use in models, Earth Interactions, 8, 1-22.

Ramankutty, N., and J. A. Foley (1998), Characterizing patterns of global land use: An analysis of global croplands data, Global Biogeochem. Cycles, 12, 667-685.

Ramankutty, N., and J. A. Foley (1999), Estimating historical changes in global land cover: Croplands from 1700 to 1992, Global Biogeochemical Cycles, 13, 997-1027.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002a), The global distribution of cultivable lands: current patterns and sensitivity to possible climate change, Glob. Ecol. Biogeogr., 11, 377-392.

Ramankutty, N., J. A. Foley, and N. J. Olejniczak (2002b), People on the land: Changes in Population and Global Croplands During the 20th Century, Ambio, 31 (3), 251-257.

Ramankutty, N., T. Hertel, H.-L. Lee, and S. K. Rose (in press), Global Agricultural Land Use Data for Integrated Assessment Modeling, in Human-Induced Climate Change: An Interdisciplinary Assessment, edited by M. Schlesinger, Cambridge University Press, New York, NY.

Reenberg, A., and B. Fog (1995), The spatial pattern and dynamics of a Sahelian agro-ecosystem, GeoJournal, 37, 489-499, DOI 410.1007/BF00806938.

Richards, J. F. (1990), Land transformation, in The Earth as Transformed by Human Action, edited by B. L. Turner, et al., pp. 163-178, Cambridge Univ. Press, New York.

Seto, K. C., R. K. Kaufmann, and C. E. Woodcock (2000), Landsat reveals China's farmland reserves, but they're vanishing fast, Nature, 406, 121-121.

Smil, V. (1999), Nitrogen in crop production: An account of global flows, Global Biogeochemical Cycles, 13, 647-662.

Tilman, D., J. Fargione, B. Wolff, C. D'Antonio, A. Dobson, R. Howarth, D. Schindler, W. Schlesinger, D. Simberloff, and D. Swackhamer (2001), Forecasting Agriculturally Driven Global Environmental Change, Science, 292, 281-284.

Tobler, W. R. (1979), Smooth pycnophylactic interpolation for geographical regions, Journal of the American Statistical Association, 74, 519-530.

Vandermeer, J., I. Perfecto, R. E. Green, S. J. Cornell, J. P. W. Scharlemann, and A. Balmford (2005), The Future of Farming and Conservation, Science, 308, 1257b-1258b, DOI: 1210.1126/science.1308.5726.1257b.

Verburg, P. H., and Y. Chen (2000), Multiscale characterization of land-use patterns in China, Ecosystems, 3, 369-385, DOI: 310.1007/s100210000033.

Verburg, P. H., Y. Chen, and T. A. Veldkamp (2000), Spatial explorations of land use change and grain production in China, Agriculture, Ecosystems & Environment, 82, 333-354.

Wang, A., D. T. Price, and V. Arora (2006), Estimating changes in global vegetation cover (1850 – 2100) for use in climate models, Global Biogeochemical Cycles, 20, GB3028, doi:3010.1029/2005GB002514.

Wint, W. (2005), Global Livestock Distributions: Data archive prepared by Environmental Research Group Oxford Limited for the Pro-poor Livestock Policy Initiative of the Animal Production and Health Division of the Food and Agriculture Organisation of the United Nations [http://ergodd.zoo.ox.ac.uk/agaagdat/index.htm], Rome, Italy.

Wood, S., K. Sebastian, and S. J. Scherr (2000), Pilot Analysis of Global Ecosystems: Agroecosystems, 110 pp, International Food Policy Research Institute and World Resources Institute, Washington D.C.

Yang, D. W., S. Kanae, T. Oki, T. Koike, and K. Musiake (2003), Global potential soil erosion with reference to land use and climate changes, Hydrological Processes, 17, 2913-2928.

Figure Captions

Figure 1. (a) Administrative units for which we calculated agricultural inventory data, (b) Percentage cropland within each administrative unit from the agriculture inventory data; (c) Percentage pasture within each administrative unit from the agriculture inventory data.

Figure 2. Flowchart depicting our methodology for combining satellite data and agricultural inventory data to derive global data sets of croplands and pastures.

Figure 3. Final estimates of croplands and pastures from this study. This is the final result obtained by calibrating the combined land-cover data set against the agricultural inventory data (Step 1), using 1000 bootstrap estimates for the parameters, and then further adjusting the predictions to match the inventory data at the administrative-unit level (Step 2).

Figure 4. 90% confidence intervals on our final estimates; here we show the difference between the 5^{th} and 95^{th} percentiles from the mean estimate. Note that the scales indicate absolute differences in the percentage of grid cells occupied by agriculture (e.g., if cropland mean = 50% and cropland 5^{th} percentile = 30%, then the difference is as 20%).

Figure 5. (a) Cropland and pasture areas estimated by this study aggregated over 14 regions of the world. The data are represented as percentage of total land area of each region. (b) Cropland and pasture areas estimated by this study aggregated over the 15 different potential natural vegetation types of the world of Ramankutty and Foley [1999]. The data are represented as percentage of total land area within each potential vegetation type.

Figure 6. (a) Cropland and (b) pasture areas estimated by this study aggregated over 6 regions of the world and by 4 vegetation types within each region.

Figure 7. Map showing the amount of overlap between croplands and pastures. Croplands or pastures were considered to be dominant when they were a factor of 3 greater in magnitude than the other category, and mixed otherwise.

Figure 8. Probability distribution of cropland and pasture areas as a function of the different proportions (or "intensities") of cropland and pastures for the different regions of the world. The figures indicate the percentage of the total area of cropland or pasture that is contributed by different intensities of cropland and pasture values (e.g., globally, croplands in grid cells that have between 60% and 70% cropland contribute to 13% of the total area of cropland).

Tables

Table 1. Land-cover classification schemes used by the two global satellite data sets

BU - MODIS legend[1]	GLC2000 Legend
0: Water 1: Evergreen Needleleaf Forest 2: Evergreen Broadleaf Forest 3: Deciduous Needleleaf Forest 4: Deciduous Broadleaf Forest 5: Mixed Forests 6: Closed Shrublands 7: Open Shrublands 8: Woody Savannas 9: Savannas 10: Grasslands 11: Permanent Wetlands 12: Croplands 13: Urban and Built-Up 14: Cropland/Natural Vegetation Mosaic 15: Snow and Ice 16: Barren or Sparsely Vegetated 254: Unclassified (recoded to 17) 255: Fill Value (recoded to 17)	1: Tree Cover, broadleaved, evergreen 2: Tree Cover, broadleaved, deciduous, closed 3: Tree Cover, broadleaved, deciduous, open (15-40% tree cover) 4: Tree Cover, needle-leaved, evergreen 5: Tree Cover, needle-leaved, deciduous 6: Tree Cover, mixed leaf type 7: Tree Cover, regularly flooded, fresh water (& brackish) 8: Tree Cover, regularly flooded, saline water, 9: Mosaic: Tree cover / Other natural vegetation 10: Tree Cover, burnt 11: Shrub Cover, closed-open, evergreen 12: Shrub Cover, closed-open, deciduous 13: Herbaceous Cover, closed-open 14: Sparse Herbaceous or sparse Shrub Cover 15: Regularly flooded Shrub and/or Herbaceous Cover 16: Cultivated and managed areas 17: Mosaic: Cropland / Tree Cover / Other natural vegetation 18: Mosaic: Cropland / Shrub or Grass Cover 19: Bare Areas 20: Water Bodies (natural & artificial) 21: Snow and Ice (natural & artificial) 22: Artificial surfaces and associated areas 23: No Data

[1]We obtained product 2000289 V003, SDS 01 Land_Cover_Type_1 with IGBP land cover classification scheme.

Table 2. Source of Census Data

Country	No. of admin. units	Source	Year of data
Argentina	499	Instituto Nacional de Estadistica y Censos, 2002	2001-02
Australia	59	Australian Bureau of Statistics, 2005	2002-03
Austria	9	Eurostat, 2005	2000
Belarus	6	Ministry of Statistics & Analysis, 1994; FAO, 2005a	1993; 2000
Belgium	11	Eurostat, 2005	2000
Bolivia	9	Instituto Nacional de Estadistica, 2005; FAO, 2005b	2000
Brazil	5510	Instituto Brasileiro de Geographia e Estatistica;http://www.ibge.gov.br;1995-1996 Census of Agriculture	1996
Bulgaria	28	Eurostat, 2005	2000
Canada	273	Statistics Canada;http://www.statcan.ca;Census of Agriculture 2001	2001
Chile	13	Instituto Nacional de Estadísticas;http://www.ine.cl;Censo Agropecuario 1997	1997
China	2400	Verburg and Chen, 2000; Liu et al. (2005);Multiscale Characterization of Land-Use Patterns in China	1998
Colombia	32	Departamento Administrativo Nacional de Estadística;http://www.dane.gov.co;Encuesta Nacional Agropecuaria: Resultados 2001	2001
Czech Republic	8	Eurostat, 2005	2000
Ecuador	22	Instituto Nacional de Estadistica y Censos;http://www.inec.gov.ec;III Censo Nacional Agropecuario	2003
Finland	6	Eurostat, 2005	1998
France	22	Eurostat, 2005	2000
Germany	40	Eurostat, 2005	1999
Greece	13	Eurostat, 2005	2000
Hungary	20	Eurostat, 2005	2000
India	552	IndiaAgristat.com; http://indiaagristat.com	1991–2002
Indonesia	26	BPS - Statistics Indonesia;http://www.bps.go.id;Land Utilization by Province 2003	2002
Iran, Islamic Rep of	24	Statistical Centre of Iran; Iran Statistical Year book 1382 [2003–2004]; http://www.sci.org.ir.	2004
Ireland	2	Eurostat, 2005	1999
Italy	20	Eurostat, 2005	1999
Japan	9	Statistics Bureau;http://www.stat.go.jp;Historical Statistics of Japan: Chapter 7.8	2000

Kazakhstan	19	Agricultural statistics of the Republic of Kazakhstan, Almaty, 1993	1993
Korea, Republic of	14	Korea National Statistical Office;http://www.nso.go.kr;Statistical DataBase (KOSIS)	2000
Lithuania	10	Eurostat, 2005	2000
Mexico	2402	INEGI;http://www.inegi.gob.mx/;VII Censo Agrícola-Ganadero 1991/Unidad de producción rural/Según uso actual del suelo/	1991
Mongolia	20	National Statistical Office of Mongolia;Mongolia in a Market System Statistical Yearbook: 1989-2002, 328pp., National Statistical Office, Ulaanbaatar, 2004	2000
Nepal	14	Central Bureau of Statistics;Statistical Year Book of Nepal 2001, 8th ed., 447 pp., Central Bureau of Statistics, Khatmandu, 2001	1991/92
Netherlands	12	Eurostat, 2005	1994
New Zealand	14	Statistics New Zealand;http://www.stats.govt.nz;Agriculture statistics (2002) - reference report	2002
Nigeria	31	Forestry Management, Environmental and Coordinating Unit (FORMECU);The Assessment of Vegetation and Land Use Changes in Nigeria Between 1976/78 and 1993/95, Draft report, pp 75., Geomatics International Inc., 1996	1993-95
Norway	19	Statistics Norway;http://www.ssb.no;Table: 04414: Agricultural area per 31 July, by use	2004
Pakistan	5	Government of Pakistan Statistics Division;http://www.statpak.gov.pk;Agricultural Census 2000	2000
Paraguay	19	Ministerio de Agricultura y Ganaderia;http://www.mag.gov.py;Produccion Agropecuaria Ano Agricola 2000/2001	2000
Peru	26	Instituto Nacional de Estadistica y Informatica;http://www.inei.gob.pe;III Censo Nacional Agropecuario 1994 (CENAGRO)	1994
Philippines	11	U.S. Department of Commerce monograph; National Statistics Office, Philippines;http://www.census.gov.ph;1991 Census of Agriculture of the Philippines; 2002 scenario of the agriculture sector in the Philippines	1991/2002
Poland	16	Eurostat, 2005	2000
Portugal	7	Eurostat, 2005	2000
Romania	41	Eurostat, 2005	2000

Russian Federation	75	State Commitee of the Russian Federation on statistics;Rossiiskoi Federatsii po Statistike. 1995. Sel´skoe khoziaistvo Rossii: Statisticheskii sbornik.Moscow (Russia): Gos. komt Rossiiskoi Federatsii po Statistike.	1993
Saudi Arabia	14	Agro-MAPS (cropland), Ministry of Economy & Planning (pasture);Pasture -- personal communication (George Allez with embassy);	2000, 1999
Slovakia	8	Eurostat, 2005	2000
South Africa	11	Central Statistical Service;http://www.statssa.gov.za;Natural resource accounts: Land accounts, 1994/1995	1995
Spain	17	Eurostat, 2005	1999
Sri Lanka	24	Department of Census and Statistics;http://www.statistics.gov.lk;Census of Agriculture 2002	2002
Sweden	8	Eurostat, 2005	1998
Thailand	72	National Statistical Office;http://web.nso.go.th;1993 Agricultural Census	1993
Turkey	73	Turkey's statistical yearbook, State Institute of Statistics, Prime Ministry, Republic of Turkey, 2004. [http://www.die.gov.tr/ENGLISH/index.html]	2001
Ukraine	25	Center for Agriculture and Rural Development;Bouzaher, Aziz, The Structure of Ukrainian Agriculture: Comparative Efficiency and Implications for Policy Reform, 115 pp., Iowa State University, Ames, IA, 1994	1991
United Kingdom	12	Eurostat, 2005	2000
United States of America	3077	2002 Census of Agriculture, National Agricultural Statistics Service, U.S. Department of Agriculture. [http://www.nass.usda.gov/Census_of_Agriculture/]	2002
Uruguay	19	Ministerio de Ganaderia Agricultura y Pesca;http://www.mgap.gub.uy;Censo General Agropecuario 2000	2000
Venezuela	24	Infoagro Zulia;http://www.zulia.infoagro.info.ve;VI Censo Agrícola Nacional - Datos Preliminar	1997/98
Vietnam	61	General Statistical Office of Vietnam;http://www.gso.gov.vn;Land use in 2003 & Number of Livestock	2003

1

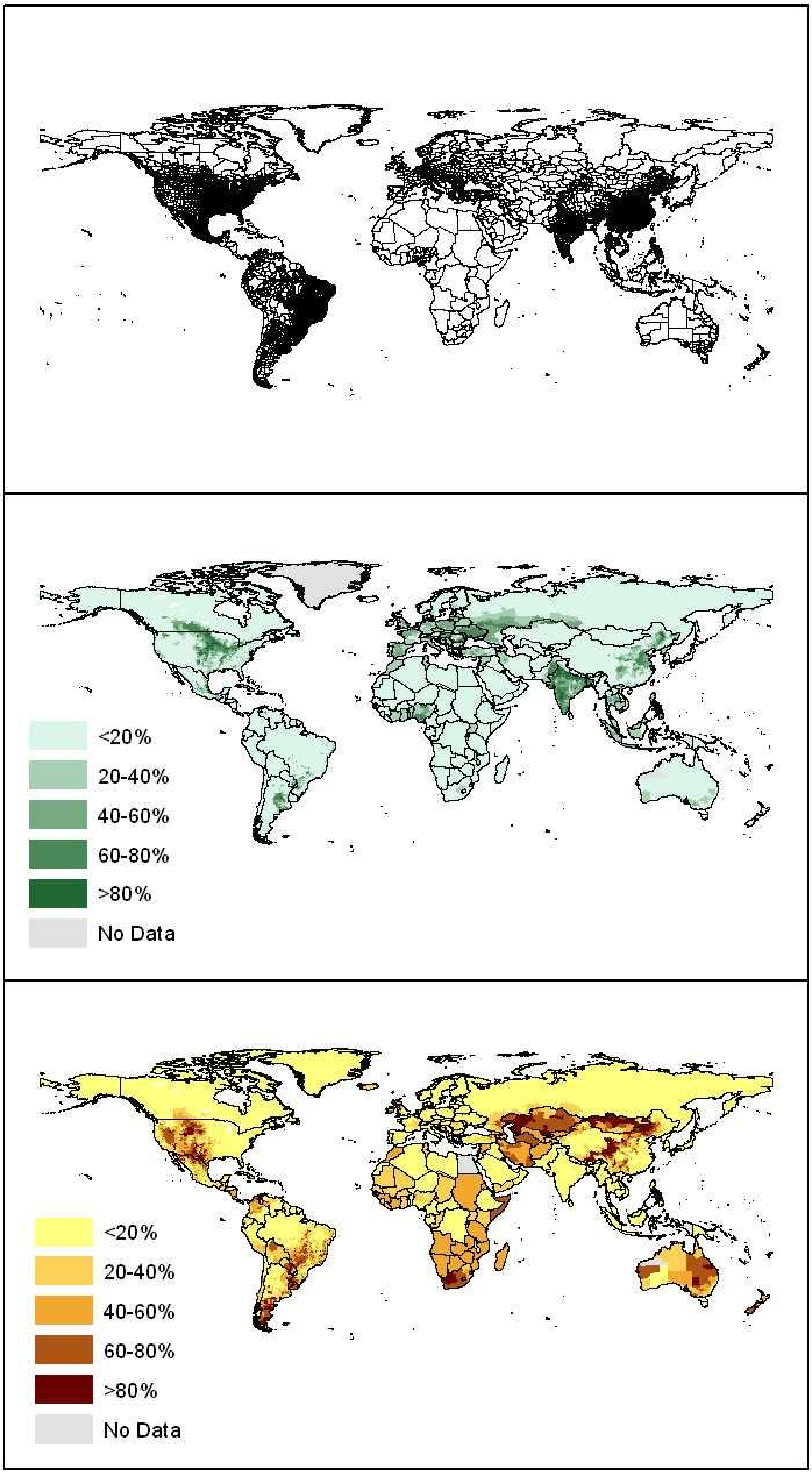
Table 3. Examples of combined land cover categories

BU-MODIS	GLC2000	Combined category
1 (Evergreen Needleleaf Forest)	1 (Tree Cover, broadleaved, evergreen)	1 (BU1GLC1)
1 (Evergreen Needleleaf Forest)	2 (Tree Cover, broadleaved, deciduous, closed)	2 (BU1GLC2)
1 (Evergreen Needleleaf Forest)	3 (Tree Cover, broadleaved, deciduous, open)	3 (BU1GLC3)
1 (Evergreen Needleleaf Forest)	4 (Tree Cover, needle-leaved, evergreen)	4 (BU1GLC4)
…	…	…
10 (Grasslands)	11 (Shrub Cover, closed-open, evergreen)	218 (BU10GLC11)
10 (Grasslands)	12 (Shrub Cover, closed-open, deciduous)	219 (BU10GLC12)
10 (Grasslands)	13 (Herbaceous Cover, closed-open)	220 (BU10GLC13)
10 (Grasslands)	14 (Sparse Herbaceous or sparse Shrub Cover)	221 (BU10GLC14)

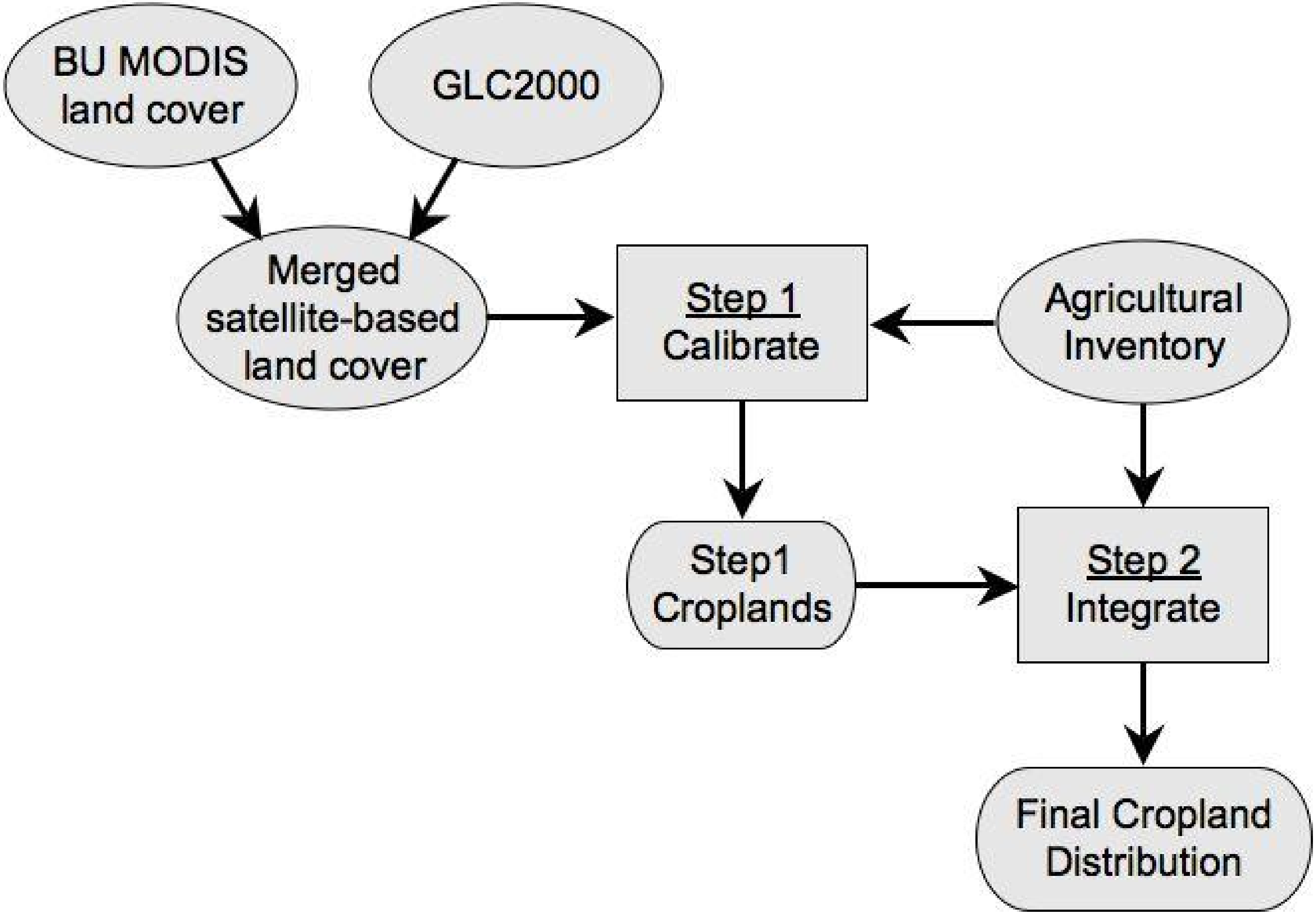
Table 4. Weighted* correlation coefficient between inventory data and model predictions from Step 1

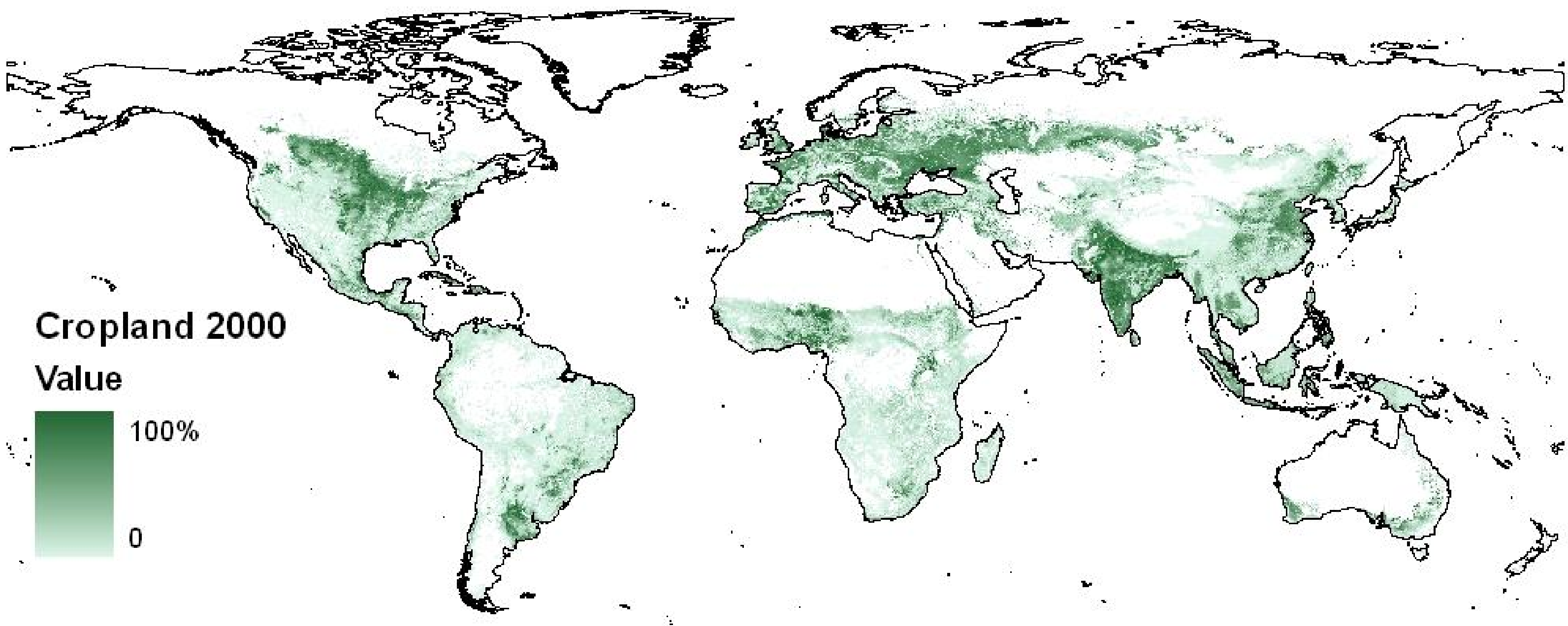
Region	Number of admin. units	BU-MODIS		GLC2000		COMBINED	
		Cropland	Pasture	Cropland	Pasture	Cropland	Pasture
Africa and Middle-East	242	0.54	0.70	0.65	0.55	0.41	0.73
Europe and Russia	448	0.93	0.87	0.95	0.85	0.95	0.84
Asia	3190	0.83	0.79	0.87	0.66	0.91	0.85
North America	5752	0.90	0.84	0.89	0.84	0.93	0.89
South America	6201	0.77	0.67	0.59	0.70	0.83	0.79
Australia/NZ	71	0.90	0.90	0.81	0.86	0.91	0.74

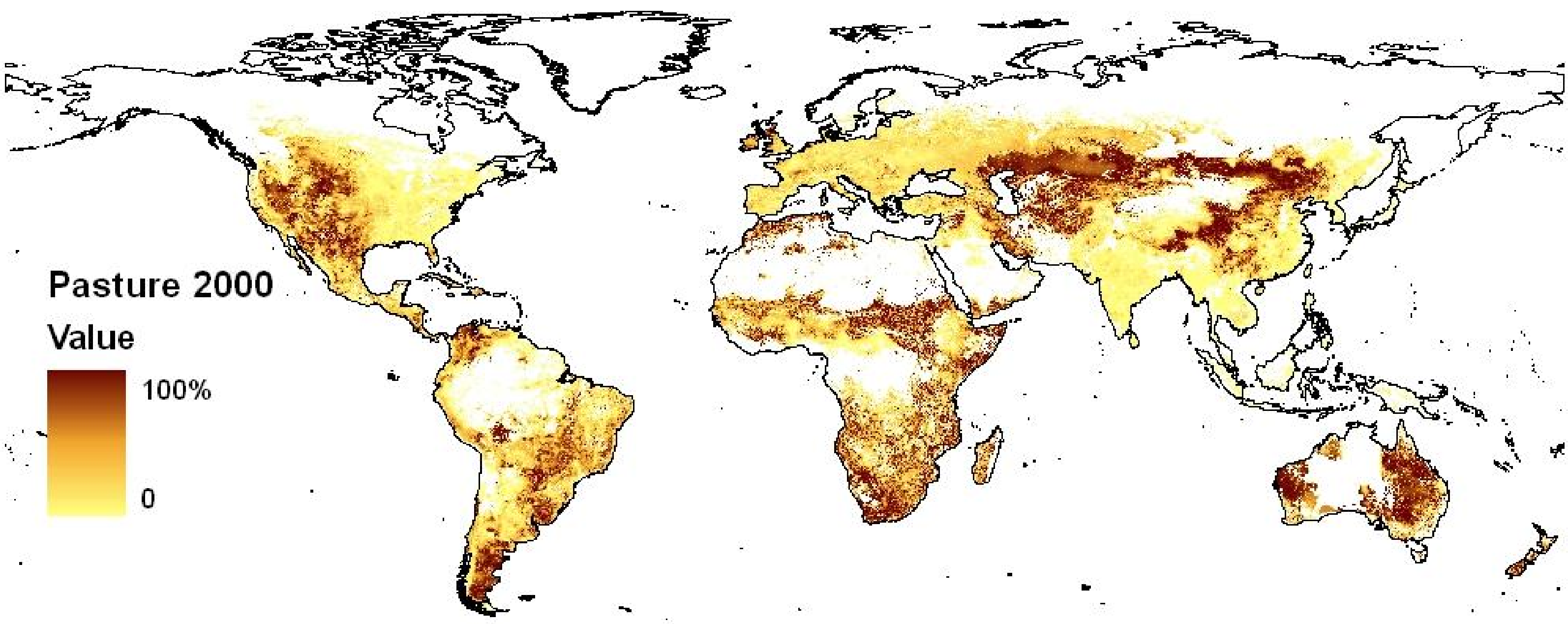
*Each data point is weighted by the area of the administrative unit.

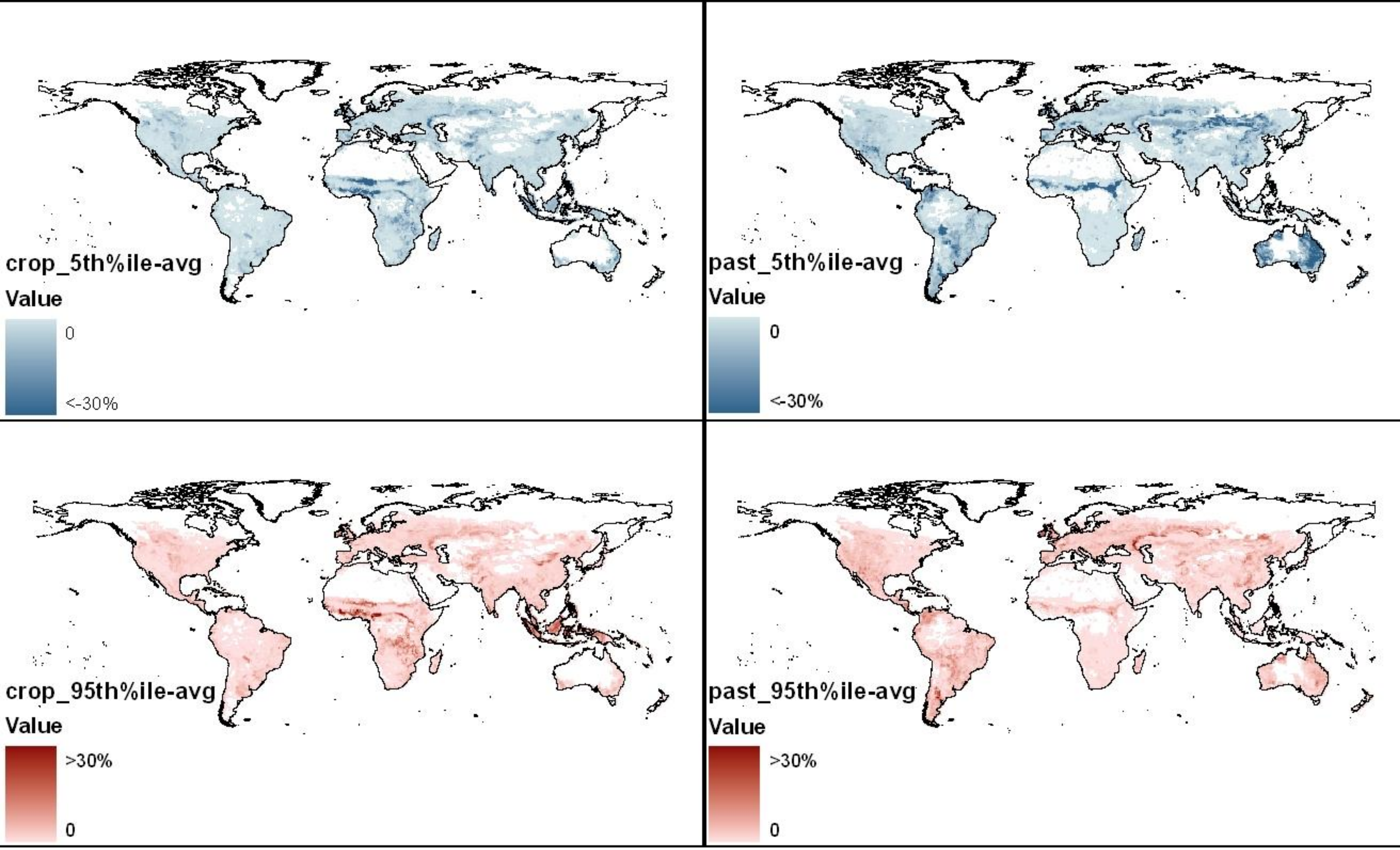


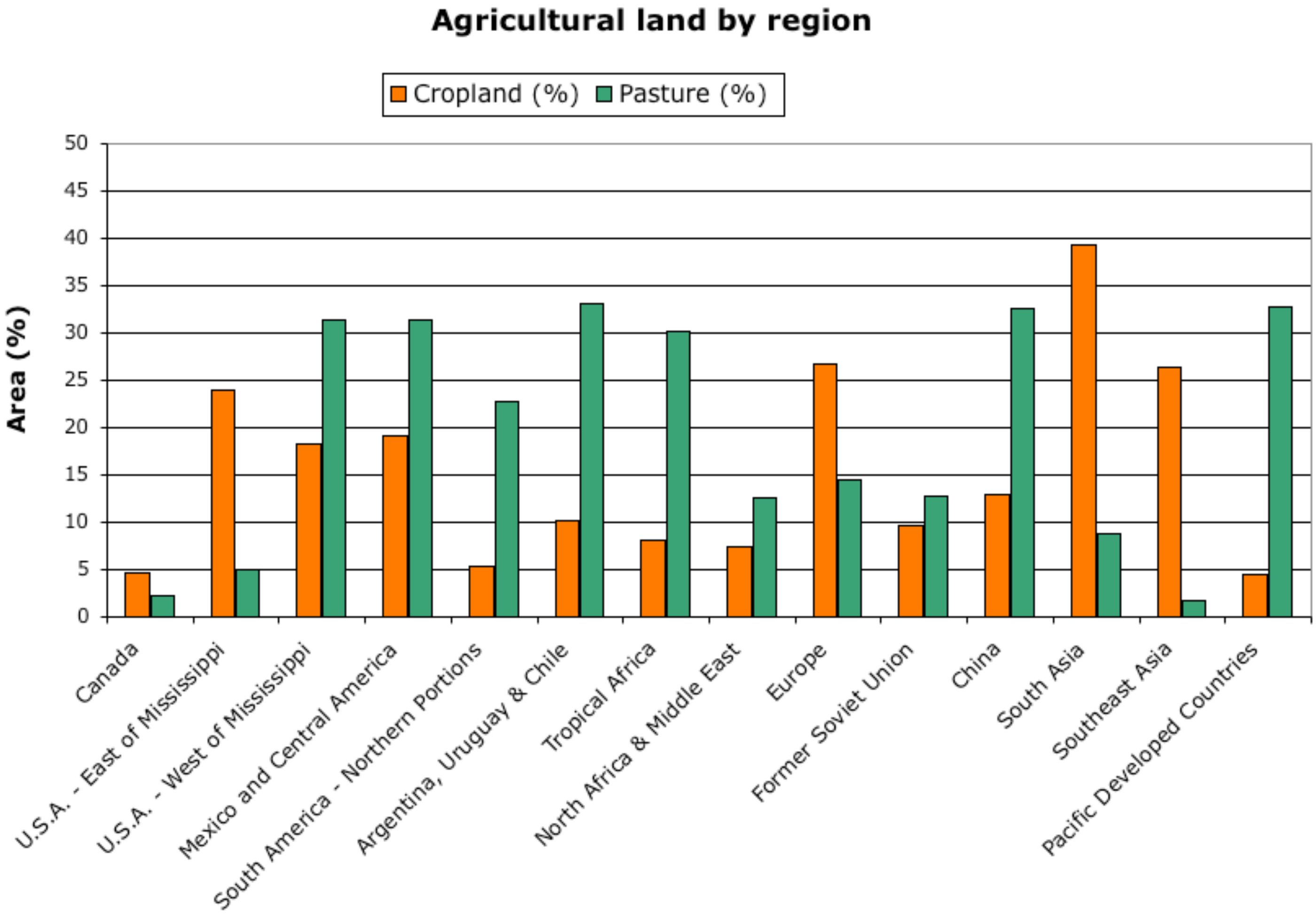
<20%
20-40%
40-60%
60-80%
>80%
No Data
<20%
20-40%
40-60%
60-80%
>80%
No Data

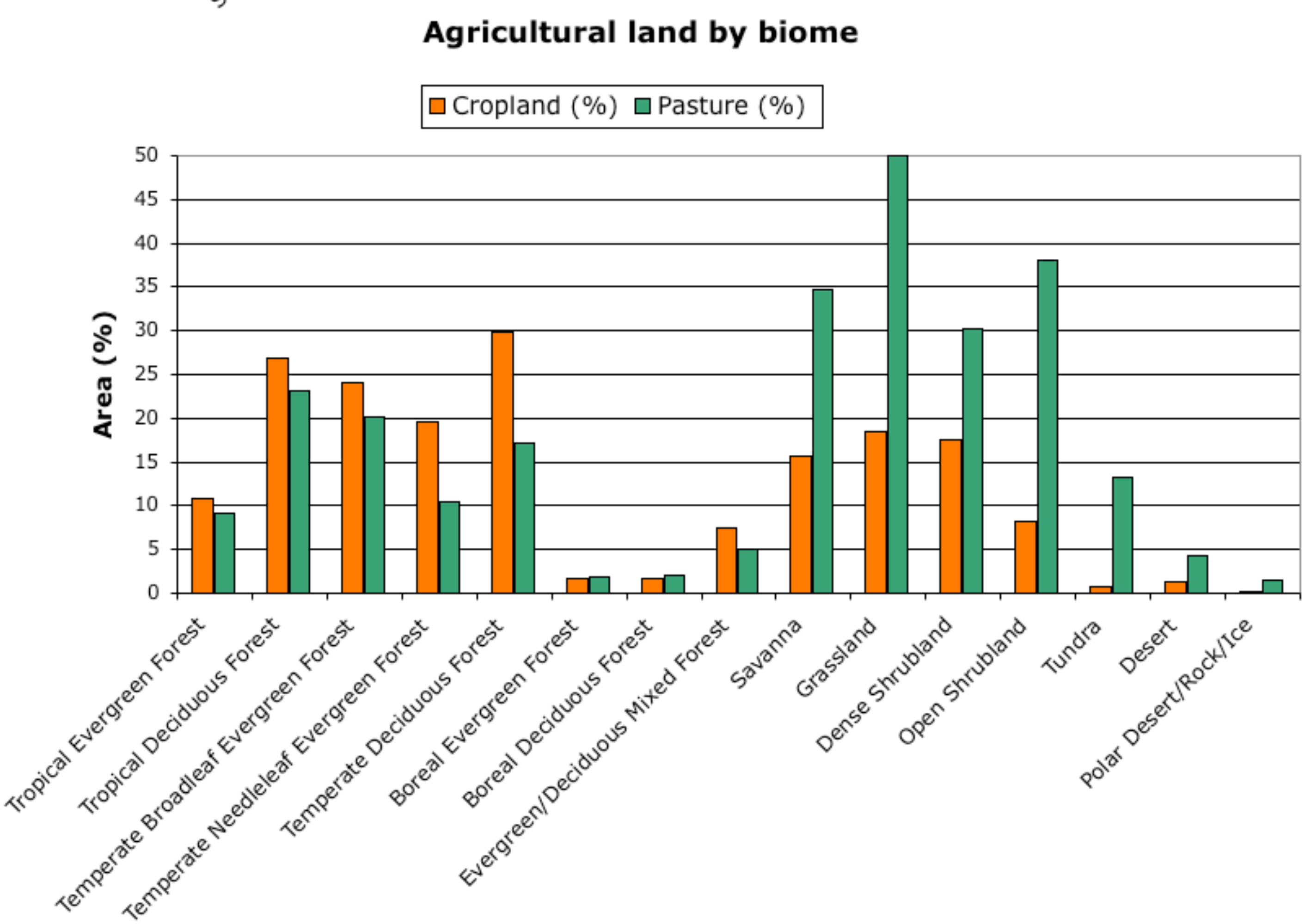

BU MODIS land cover
GLC2000
Merged satellite-based land cover
Step 1
Calibrate
Agricultural Inventory
Step1 Croplands
Step 2
Integrate
Final Cropland Distribution

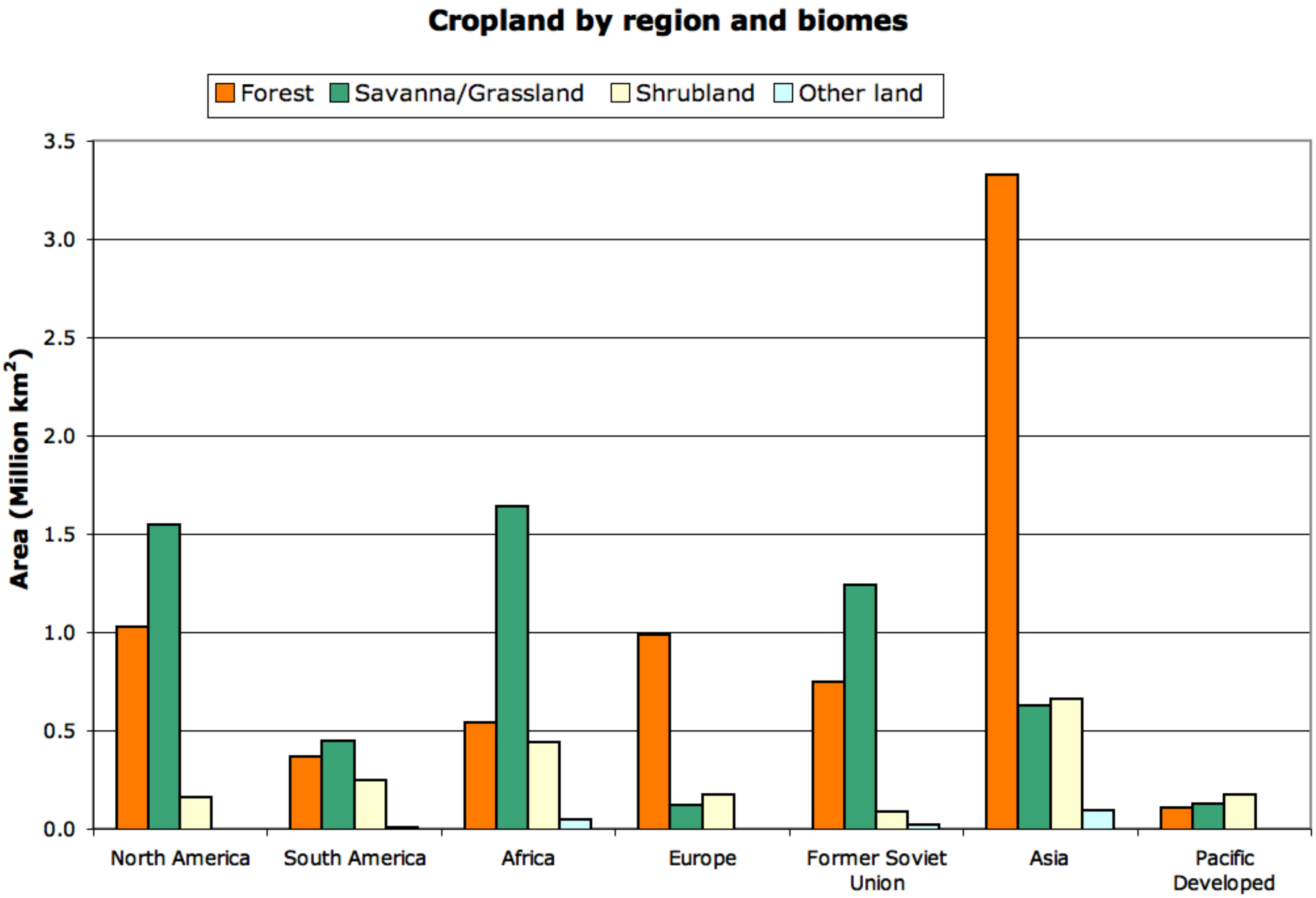

Cropland 2000
Value
100%
0


Pasture 2000
Value
100%
0

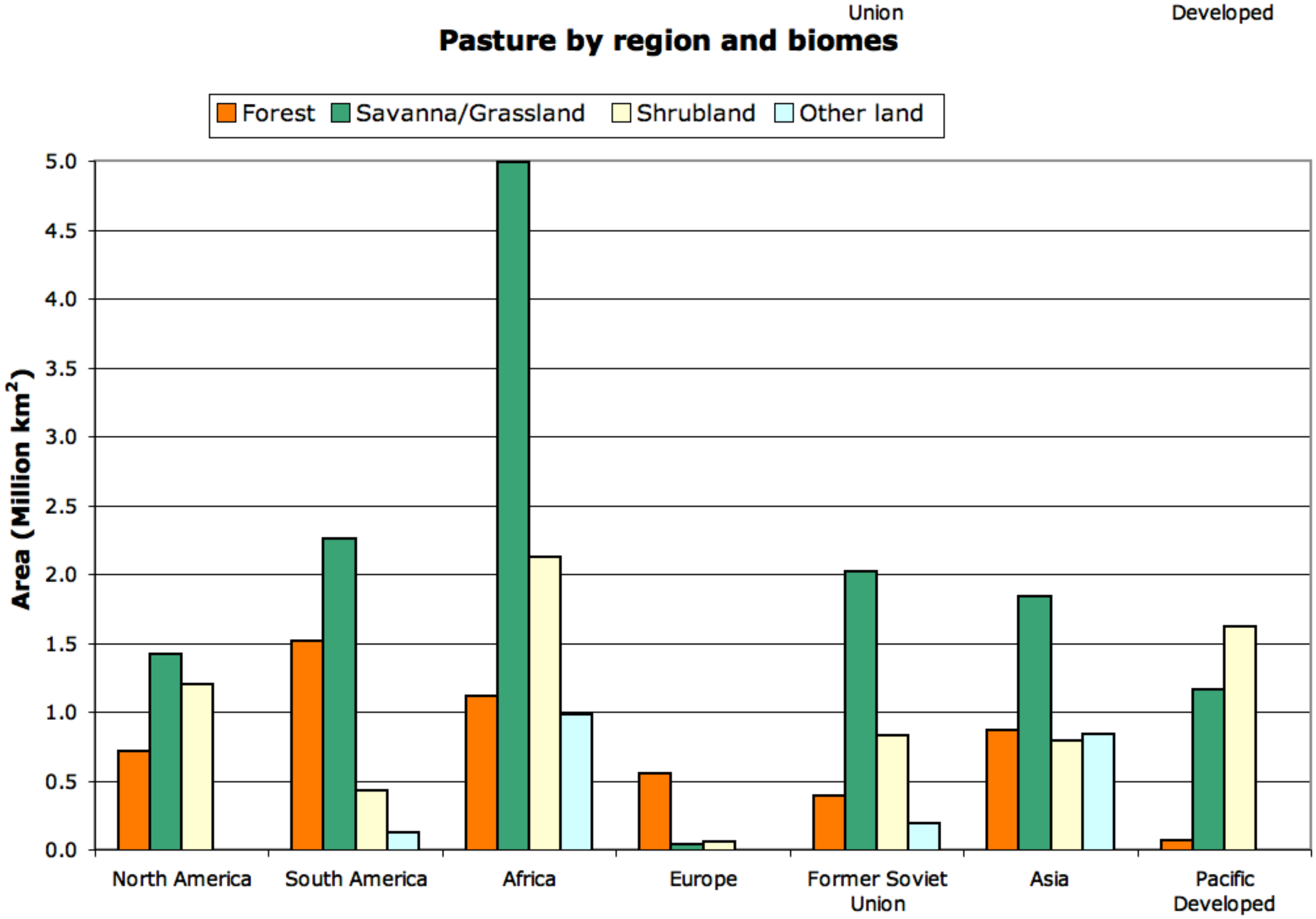

crop_5th%ile-avg
Value
0
<-30%
past_5th%ile-avg
Value
0
<-30%
crop_95th%ile-avg
Value
>30%
0
past_95th%ile-avg
Value
>30%
0


Agricultural land by region
Cropland (%)
Pasture (%)
Area (%)
50
45
40
35
30
25
20
15
10
5
0
Canada
U.S.A. - East of Mississippi
U.S.A. - West of Mississippi
Mexico and Central America
South America - Northern Portions
Argentina, Uruguay & Chile
Tropical Africa
North Africa & Middle East
Europe
Former Soviet Union
China
South Asia
Southeast Asia
Pacific Developed Countries

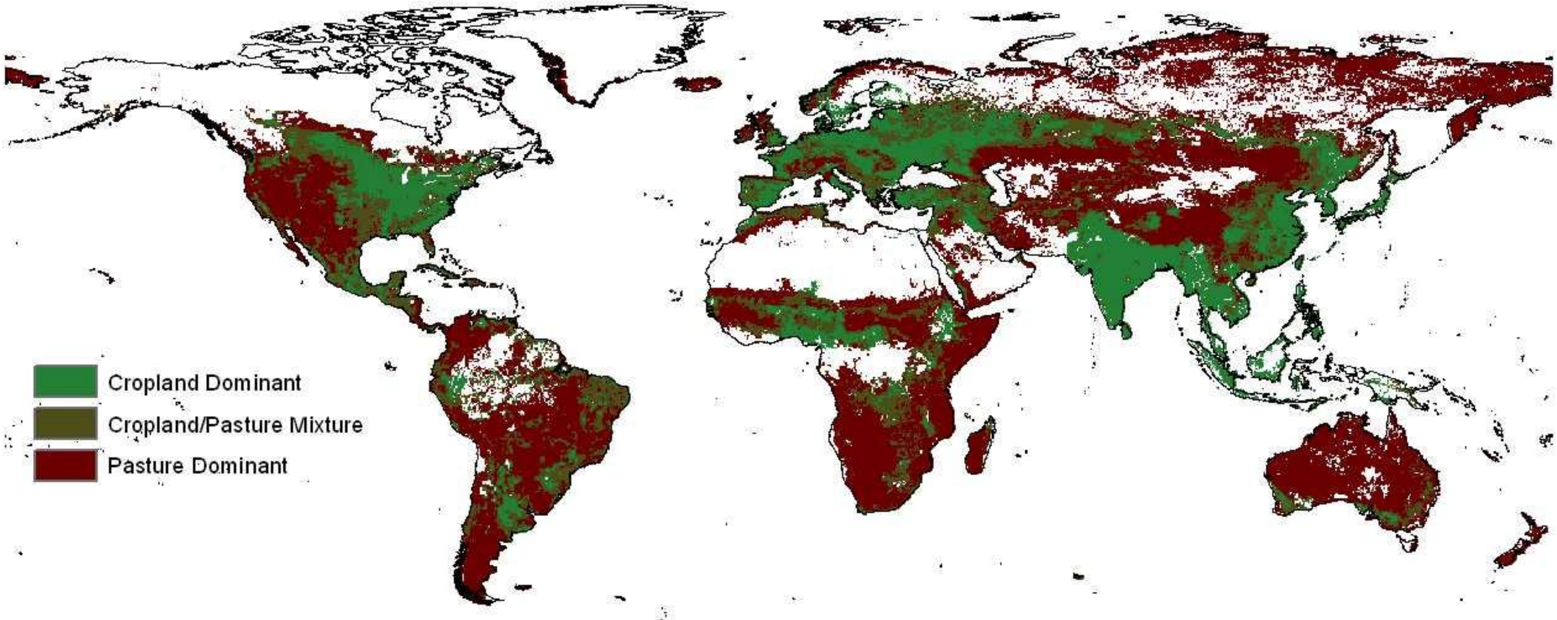

Agricultural land by biome
Cropland (%)
Pasture (%)
Area (%)
50
45
40
35
30
25
20
15
10
5
0
Tropical Evergreen Forest
Tropical Deciduous Forest
Temperate Broadleaf Evergreen Forest
Temperate Needleleaf Evergreen Forest
Temperate Deciduous Forest
Boreal Evergreen Forest
Boreal Deciduous Forest
Evergreen/Deciduous Mixed Forest
Savanna
Grassland
Dense Shrubland
Open Shrubland
Tundra
Desert
Polar Desert/Rock/Ice

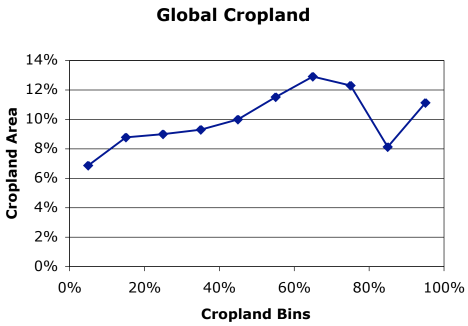


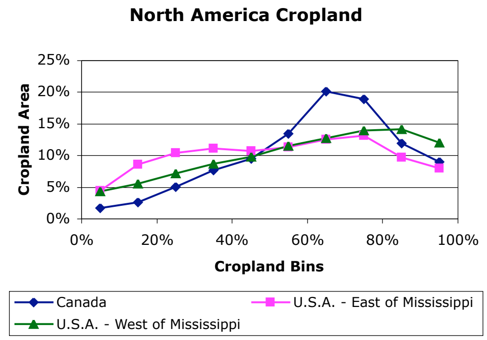
Cropland by region and biomes
Forest
Savanna/Grassland
Shrubland
Other land
Area (Million km²)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
North America
South America
Africa
Europe
Former Soviet Union
Asia
Pacific Developed

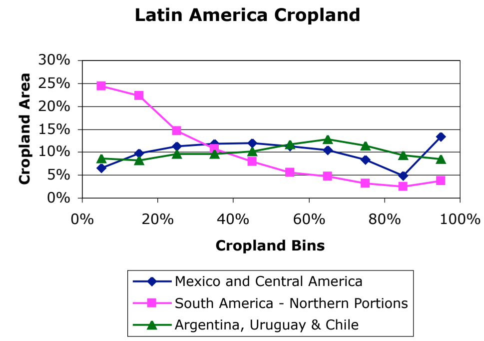


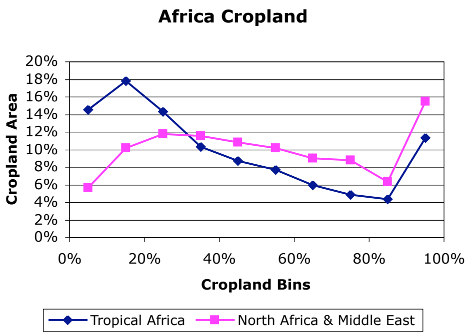
Pasture by region and biomes
Forest
Savanna/Grassland
Shrubland
Other land
Area (Million km²)
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
North America
South America
Africa
Europe
Former Soviet Union
Asia
Pacific Developed

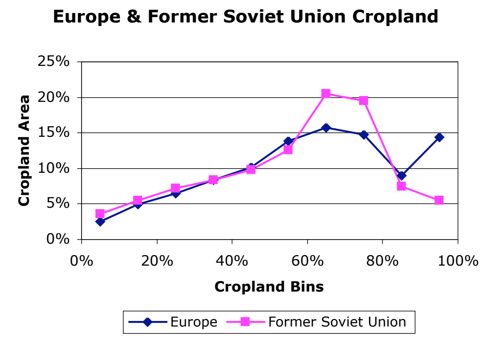

Cropland Dominant
Cropland/Pasture Mixture
Pasture Dominant

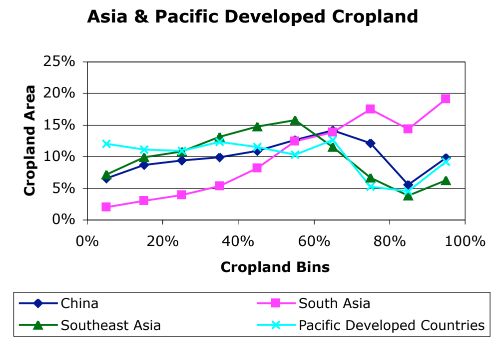

Global Cropland
Cropland Area
14%
12%
10%
8%
6%
4%
2%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins

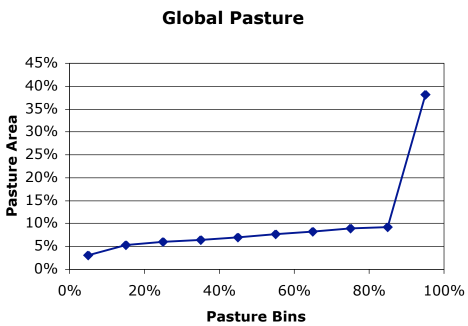

North America Cropland
Cropland Area
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins
Canada
U.S.A. - East of Mississippi
U.S.A. - West of Mississippi

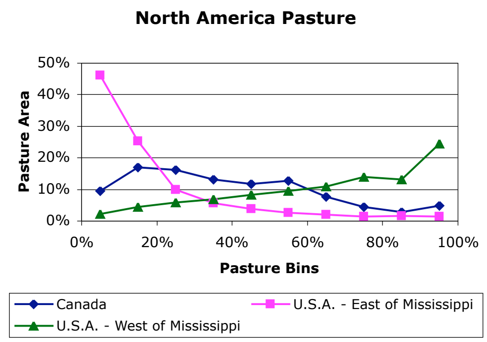

Latin America Cropland
Cropland Area
30%
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins
Mexico and Central America
South America - Northern Portions
Argentina, Uruguay & Chile

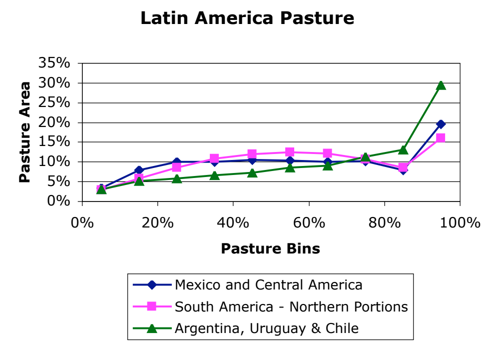

Africa Cropland
Cropland Area
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins
Tropical Africa
North Africa & Middle East

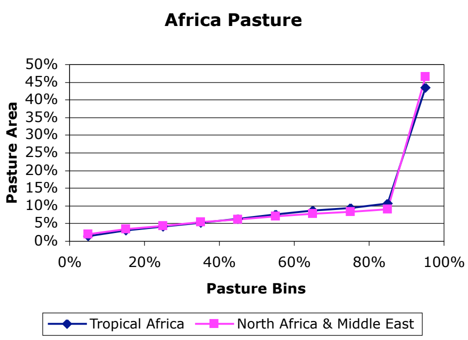

Europe & Former Soviet Union Cropland
Cropland Area
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins
Europe
Former Soviet Union

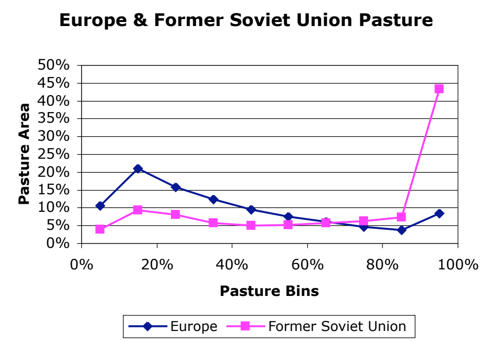

Asia & Pacific Developed Cropland
Cropland Area
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Cropland Bins
China
South Asia
Southeast Asia
Pacific Developed Countries

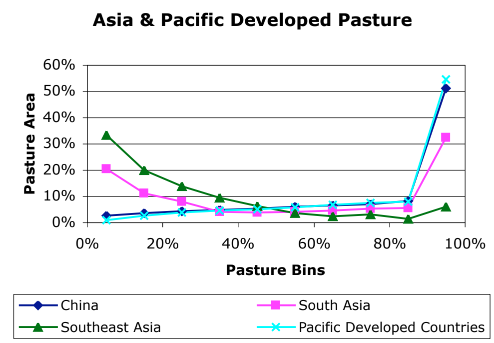

Global Pasture
Pasture Area
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins


North America Pasture
Pasture Area
50%
40%
30%
20%
10%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins
Canada
U.S.A. - East of Mississippi
U.S.A. - West of Mississippi


Latin America Pasture
Pasture Area
35%
30%
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins
Mexico and Central America
South America - Northern Portions
Argentina, Uruguay & Chile


Africa Pasture
Pasture Area
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins
Tropical Africa
North Africa & Middle East


Europe & Former Soviet Union Pasture
Pasture Area
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins
Europe
Former Soviet Union


Asia & Pacific Developed Pasture
Pasture Area
60%
50%
40%
30%
20%
10%
0%
0% 20% 40% 60% 80% 100%
Pasture Bins
China
South Asia
Southeast Asia
Pacific Developed Countries