

Global Land Use and Land Cover Data for Integrated Assessment Modeling

Navin Ramankutty^a, Tom Hertel^b, Huey-Lin Lee^{b,c}

^aCenter for Sustainability and the Global Environment (SAGE)
Institute for Environmental Studies
University of Wisconsin
Madison, WI – 53726
Email: nramanku@wisc.edu

^bCenter for Global Trade Analysis
Purdue University
1145 Krannert Building
West Lafayette IN 47907-1145
Email: hertel@purdue.edu

^c Email: hlee1@purdue.edu

Abstract

Changes in land use and land cover have been a major driver of global change over the last three centuries. Detailed spatially-explicit data sets characterizing global land cover changes are now emerging. By synthesizing remotely-sensed data with land use inventory data, the Center for Sustainability and the Global Environment (SAGE) is developing databases of contemporary and historical agricultural land cover change. These datasets are of immense value to the integrated assessment (IA) modeling community. One way to further enhance their value is by integrating them with global economic data in order to facilitate improved analysis of the costs of climate change mitigation. Towards this end, we are currently incorporating the SAGE land use data sets into the Global Trade and Analysis Project (GTAP) database. The GTAP database is currently the premier data set for economy-wide analysis of global economic issues and it has been widely used by the IA community. It includes detailed bilateral trade, protection, and energy data characterizing economic linkages among regions, as well as individual country input-output tables that account for inter-sectoral linkages within regions. Unfortunately, the current specification of land use in GTAP is rudimentary at best and does not distinguish between types of land. To incorporate the SAGE land use database into GTAP, we have divided each country's land endowment into eighteen Agro-Ecological Zones (AEZs). This paper describes the procedure developed to integrate SAGE land use data, as well as some crop yield data from the Food and Agriculture Organization (FAO), to derive global national-level statistics on land cover and agricultural production by AEZ. The resulting database permits economists interested in IA to better assess the competition for land among agricultural activities and between agriculture and forestry.

1. Introduction

Human land use activities, while extracting natural resources such as food, fresh water, and fiber, have significantly transformed the face of the planet. Recently, Hannah *et al.* [1994] estimated that roughly 50% of the planet's land surface (75% of the habitable area) has been either moderately or severely disturbed; only the core of the tropical rainforests and Boreal forests, deserts, and ice-covered surfaces are still relatively untouched by humans. Moreover, Vitousek *et al.* [1997] estimated that around 40% of the global net primary productivity is being co-opted by humans, while Postel *et al.* [1996] estimated that over 50% of the available renewable fresh water supply is being co-opted.

The major mode of human land transformation has been through agriculture. Since the invention of agriculture, ~10,000 years ago, humans have modified or transformed the land surface; today, roughly a third of the planet's land surface is being used for growing crops or grazing animals [Foley *et al.*, 2003; *National Geographic Maps*, 2002]. The pace of land clearing for cultivation has been particularly rapid in the last 300 years. Richards [1990] estimated that over the last 300 years, we lost 20% of forests and woodlands, 1% of grasslands and pastures (although most grasslands were converted to pastures), and croplands expanded by 466%.

Such large-scale changes in global land use and land cover can have significant consequences for food production, freshwater supply, forest resources, biodiversity, regional and global climates, the cycling of carbon, nitrogen, phosphorus, etc. In particular, it has been estimated that roughly a third of the total emissions of carbon into the atmosphere since 1850 has resulted from land use change (and the remainder from fossil-fuel emissions) [Houghton, 2003]. In the 1990s, while 6.4 Gt-C yr⁻¹ was emitted to the atmosphere from industrial activities, about 2.2 Gt-C yr⁻¹ was emitted from tropical deforestation.

While the clearing of tropical forests today results in a source of carbon to the atmosphere, natural ecosystems are also sinks of carbon. It is known, from attempts to close the global carbon budget, that roughly 3 Gt-C yr⁻¹ is taken up by terrestrial ecosystems, and all evidence points to a significant uptake in the Northern Hemisphere high latitude regions [Prentice, 2001; Tans *et al.*, 1990]. Many hypotheses for the uptake have been proposed including CO₂ fertilization, nitrogen fertilization, climate driven uptake, and carbon uptake by vegetation recovering from past disturbances [Prentice, 2001]. It is becoming increasingly clear that terrestrial ecosystems can be a source or sink of carbon depending on how they are managed. The notion of terrestrial *carbon sequestration* emerged from this understanding, wherein terrestrial ecosystems are actively manipulated, through afforestation, soil management, etc., to withdraw carbon dioxide from the atmosphere [IPCC, 2000]. This notion has become so widespread that in the Kyoto Protocol, terrestrial carbon sequestration was included as a possible means for nations to reduce their net emissions of carbon.

In addition to affecting carbon dioxide fluxes, land use changes can also mitigate other greenhouse gas emissions. In particular, appropriate management of forests and agricultural lands can reduce emissions of CO₂, CH₄, and N₂O, can further sequester CO₂, and biomass substitutes for fuels can reduce emissions that would otherwise result from burning fossil fuels. Numerous estimates have been made by the integrated assessment community of the cost of

abating greenhouse gas emissions through land use change [Richards and Stokes, 2004]. Early estimates suggest that mitigating emissions from land use management could be less expensive than through mitigation of traditional industrial activities [US-EPA/GTAP, 2004]. Therefore, numerous partial and general equilibrium models are being developed to study climate change policy and the role of land use change in abating GHG emissions.

1.1. The GTAP database

The GTAP database includes detailed bilateral trade, protection, and energy data characterizing economic linkages among regions, as well as individual country input-output tables that account for inter-sectoral linkages within regions. GTAP began as a database and modeling framework to assess the global implications of trade policies. However, over the past decade, through a series of grants from the US Department of Energy and the US Environmental Protection Agency, GTAP has become increasingly central to analyses of the global economic consequences of attempts to mitigate greenhouse gas emissions. The first step in this direction involved integrating the International Energy Agency's database on fossil fuel consumption into GTAP. When coupled with CO₂ emissions coefficients, this permitted researchers to more accurately estimate changes in economic activity and fossil-fuel-based emissions in the wake of policies aimed at curbing CO₂ emissions. More recent work has extended the GTAP database to include data on non-CO₂ greenhouse gas emissions. Most importantly, these emissions are linked to the underlying economic drivers of emissions, which are faithfully represented in the core GTAP database. However, the treatment of land use in the standard GTAP model has been rudimentary at best, with no differentiation across land types, and no link to the global land use databases.

In response to the increasing emphasis on land use change as a factor in greenhouse gas abatement, Global Analysis and Trade Project (GTAP) at Purdue University has been actively working to develop a new land use database which has the capability to support the associated land use modeling activities – particularly in agriculture and forestry. This has involved close collaboration with researchers at SAGE, the Center for Sustainability and the Global Environment, based at the University of Wisconsin, as well as at The Ohio State University. This paper documents the methodology used to derive land use information for the GTAP database.

2. Global Land Use Data for GTAP

There are several groups around the world that are developing global land use data sets. These groups are focusing different issues, and using different methodologies. GTAP organized a workshop at the Massachusetts Institute of Technology (MIT) to invite these groups to present their land use data and their approaches, and to discuss the feasibility of including them in GTAP [GTAP/MIT, 2002]. Following this workshop, the groups were invited to contribute their data for the U.S. and China. The data were comprehensively compared against each other (figures 1 and 2; [Lee, 2002]), and SAGE was invited to contribute land use data for GTAP, based on both the outcome of the intercomparison, as well as the ease of working within the GTAP framework.

2.1. SAGE Global Land Use Data

The Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin has been developing global databases of contemporary and historical agricultural land use and land cover. SAGE has chosen to focus on agriculture because it is clearly the predominant land use activity on the planet today, and provides a vital service – food – for human societies.

SAGE has developed a “data fusion” technique to integrate remotely-sensed data on the world’s land cover with administrative-unit-level inventory data on land use [Ramankutty and Foley, 1998; Ramankutty and Foley, 1999]. The advent of remote sensing data has been revolutionary in providing consistent, global, estimates of the patterns of global land cover. However, remote sensing data are limited in their ability to resolve the details of agricultural land cover from space. Therein lies the strength of the ground-based inventory data, which provide detailed estimates of agricultural land use practices. However, inventory data are limited in not being spatially explicit, and are plagued by problems of inconsistency across administrative units. The “data fusion” technique developed by SAGE exploits the strengths of both the remotely-sensed data as well as the inventory data.

Using SAGE’s methodology, Ramankutty and Foley [1998] [RF98 hereafter] developed a global data set of the world’s cropland distribution for the early 1990s (Fig. 3). This was accomplished by integrating the Global Land Cover Characteristics (GLCC; Loveland *et al.* [2000]) database at 1 km resolution (derived from the Advanced Very High Resolution Radiometer (AVHRR) instrument), with a comprehensive global inventory data (at national and subnational levels) of cropland area. The data set, at a spatial resolution of 5 min (~10 km) in latitude by longitude, describes the percentage of each 5 min grid cell that is occupied by croplands. Leff *et al.* [2004] further disaggregated the RF98 dataset to derive the spatial distribution of 18 major crop types of the world (plus one “other crop” type; maps of individual crops not shown – see Leff *et al.* [2004] for detailed maps). Ramankutty and Foley [1999] [RF99 hereafter] compiled historical inventory data on cropland areas to extend the global croplands data set back to 1700 (figures not shown). RF99 also derived a global data set of potential natural vegetation (PNV) types; this data set describes the spatial distribution of 15 natural vegetation types that would be present in the absence of human activities (Fig. 4). Furthermore, global data sets of the world’s grazing lands (Fig. 5) and built-up areas (not shown), representative of the early 1990 period, were also developed recently [Foley *et al.*, 2003; *National Geographic Maps*, 2002].

The SAGE data sets described above are being used for a wide array of purposes, including global carbon cycle modeling [McGuire *et al.*, 2001], analysis of regional food security [Ramankutty *et al.*, 2002b], global climate modeling [Bonan, 1999; Bonan, 2001; Brovkin *et al.*, 1999; Myhre and Myhre, 2003], estimation of global soil erosion [Yang *et al.*, 2003], OTHERS. etc. They also formed part of the BIOME300 effort, initiated by two core projects – LUCC (Land Use and Land Cover Change) and PAGES (Past Global Changes) of the International Geosphere-Biosphere Programme (IGBP). In other words, they are a widely recognized, and widely used data set of global agricultural land use.

The SAGE agricultural land use data formed the core of the GTAP land use database. In

addition to the SAGE data, to derive information on crop yields and irrigation, some ancillary data were obtained from the Food and Agriculture Organization (FAO) as well as the University of Kassel in Germany. In the next section, we describe the procedure used to adapt the SAGE data and the ancillary data to derive land use information for GTAP.

3. Incorporating SAGE data into GTAP

3.1. Global Agro-Ecological Zones (AEZs)

In constructing the GTAP land use database, we adopt the FAO/IIASA convention of agro-ecological zones (AEZs) The FAO was one of the first organizations to develop the concept of AEZs. Recently, FAO, in collaboration with the International Institute for Applied Systems Analysis (IIASA), developed a global database of AEZs [Fischer *et al.*, 2002; Fischer *et al.*, 2000]. For global AEZ data, this product is considered State of the Art. For our purposes, we adopted the “Length of Growing Period” (LGP) data from the IIASA/FAO GAEZ database. Fischer *et al.* [2000] derived the length of growing period by combining climate, soil, and topography data with a water balance model and knowledge of crop requirements. LGP describes the number of days during a growing season with adequate temperature and moisture to grow crops.

Table 1 shows the format of the GTAP land use data.¹ We identify land located in each agro-ecological zone and this land is allocated across the rows in Figure 1. The land uses (sectors or activities) are in turn allocated across the columns in Figure 1. In the GTAP database, each country has its endowment of agricultural land, used in the crops, livestock and forestry. This total land endowment is represented by the sum over all of the cells in Figure 1. We hypothesize that any activity that presently uses land in a given AEZ represents a candidate for use of other land within the same AEZ. In economic terms, *land is assumed to be mobile across uses within any given AEZ*. The extent of this mobility will determine the rate at which land use change responds to changes in the relative per hectare return to different activities. And this, in turn, is governed by an estimated parameter in the model.

While land is mobile across uses within an AEZ, it is *immobile* across AEZs. Therefore, if a given activity (e.g., irrigated rice production) is not observed in AEZ3, then it will never appear in that AEZ. Furthermore, if there is another activity, say forestry, which is only observed in AEZ3, then it will never compete directly for land with irrigated rice production. This assumption has some limitations. However, we believe it is much preferable to the alternative in which any type of land-using activity can compete with any other activity for the same land. The fact is that most crops can only grow on lands that is under certain temperature, moisture, soil type, land form, etc. The same concern arises for land use by the livestock and the forestry sectors. Lands that are suitable for growing wheat may not be good for rice cultivation. The introduction of agro-ecological zoning in GTAP helps to provide a sounder, more empirically-based presentation of sectoral competition for land.

3.2. Key Assumptions and Procedures

¹ This format was originally proposed at the 2002 MIT workshop [GTAP/MIT, 2002]

In order to supply the necessary data for this specification of GTAP, the spatially-explicit land use data sets from SAGE needed to be aggregated along the lines suggested in Figure 1. Furthermore, additional developments are required: 1) Development of global Agro-Ecological Zones for deriving subnational information on land endowments; 2) Mapping data to match GTAP crop sectors; 3) Deriving yield (and production) data for the crop sectors; and 4) Further disaggregating the data into irrigated versus rainfed areas and yields of major crops. These developments are described in detail below.

3.2.1. Definition of AEZs

We first derived 6 global LGPs by aggregating the IIASA/FAO GAEZ data into 6 categories of ~60 LGP days each: 1) LGP1: 0-59 days; 2) LGP2: 60-119 days; 3) LGP3: 120-179 days; 4) LGP4: 180-239 days; 5) LGP5: 240-299 days; 6) LGP6: > 300 days. We also subdivided the world into three climatic zones -- Tropical, Temperate, and Boreal -- using criteria based on absolute minimum temperature and Growing Degree Days, as described in Ramankutty and Foley [1999]. By overlaying the 6 categories of LGPs with the 3 climatic zones, we developed a global map of 18 AEZs (Figure 6).

3.2.2. Global land cover

A map of global land cover, representative of present-day conditions, was first derived by overlaying the SAGE global data set of potential natural vegetation, over the present-day global maps of croplands, grazing lands, and urban areas. The resulting map was overlain with the global AEZ map, to calculate land cover by country, for each of the 18 AEZs (a summary chart with global total numbers as a function of the 6 LGPs is shown in figure 7, with the distinction between tropical, temperate, and boreal zones removed for clarity of the figure).

3.2.3. Crop harvested area

The SAGE data provides information on crop areas [Leff *et al.*, 2004]. They were derived by compiling crop harvested area statistics from national and subnational sources, estimating the proportions of harvested area of each to total harvested area, and then redistributing it using the RF98 croplands map.

Because the RF98 croplands map represents the “physical” cultivated area on the ground, the redistribution of crop harvested area essentially resulted in a conversion to physical area on the ground. Here, we first recalibrate the SAGE data to the national crop harvested area statistics from FAOSTAT [FAO, 2004].

Let $A_{SAGE}(i, mc)$ be the SAGE crop area for pixel i and major crop mc . Note that the original SAGE data sets are gridded, at 0.5 degree resolution in latitude by longitude. We recalibrate the SAGE data to the FAOSTAT harvested area data as follows:

$$A'_{SAGE}(i, mc) = A_{SAGE}(i, mc) \times \frac{A_{FAO}(l, mc, t_{ref})}{\sum_{i \in l} A_{SAGE}(i, mc)},$$

where l = countries in FAOSTAT, $i \in l$, and t_{ref} is the reference time period = 2001.

The recalibrated SAGE major crops data are then overlain with: 1) Global AEZ map; and 2) Political boundaries, and aggregated to derive harvested areas of 19 crops for all nations of the world, for 18 AEZs within each nation.

Let this aggregated data be represented by $A_{SAGE}(l, mc, z)$, where,

l is the country, mc is one of 19 SAGE major crops, and $z =$ one of 18 AEZs.

This can then be mapped onto GTAP's 8 crop sectors (see table 4),

$$A_{SAGE}(l, mc, z) \rightarrow A_{GTAP}(l, s, z).$$

3.2.4. Estimating crop yields from FAO data

The FAO provided GTAP with estimates of harvested area, yield, and production, for 94 developing countries, for several agro-ecological zones (including an AEZ labeled "irrigated"). This data was developed based on primary data obtained in the 1970s, and has been periodically updated since then based on observed aggregates (Jelle Bruinsma, personal communication, 2003). So while the FAO data are not strictly reliable for direct use, they are the only available data on yields by AEZ within countries. We have therefore chosen to use the FAO data as a provisional measure, until improved estimates become available in the future. Here we describe how we adopted the FAO data for our purpose.

3.2.4.1. Derive yields from FAO data for 94 developing countries.

The FAO data were provided for 6 different agro-ecological zones (see Alexandratos [1995]), defined slightly differently from our AEZs, and for 34 different crops. We therefore had to match the FAO AEZs and crops with SAGE's 18 AEZs and crops.

We first map from the 34 FAO crops to the 18 SAGE major crops, based on the mapping given in Table 2.

Let $Y_{FAO,RF}(n, mc, fz)$ be the rainfed yield for nation 'n', crop 'mc', and FAO AEZ 'fz', where,
 $mc =$ one of 19 SAGE major crops,
 $fz =$ FAO's AEZs AT1, AT2, AT3, AT4+AT5 (see Table 3)

FAO also reported yield for a zone AT6+AT7 (Fluvisols/Gleysols), and for a zone called "Irrigated Land".

Let $Y_{FAO}(n, mc, tot)$ be the total yield from FAO (harvested area weighted average of all 6 zones)

The FAO separates out and reports yields for all irrigated areas and fluvisol/gleysols under one category each. Here we choose to think of the AEZs as a *climate only* constraint (or an endowment), and that irrigation or fluvisols/gleysols can occur within each AEZ. Therefore, we have to estimate the total yields for each rainfed AEZ (i.e., we essentially have to repartition the irrigated and AT6+AT7 yields back into the rainfed zones).

To do this, we first calculate *national total rainfed* yields for each crop,

$$Y_{FAO}(n, mc, RF) = \frac{\sum_{fz=AT1}^{AT4+AT5} Y_{FAO}(n, mc, fz) \times A_{FAO}(n, mc, fz)}{\sum_{fz=AT1}^{AT4+AT5} A_{FAO}(n, mc, fz)}, \text{ where}$$

A_{FAO} = harvested area data from FAO, corresponding to the yield data 'RF' refers to rainfed.

Then, calculate total yield for each of the FAO AEZs, AT1, AT2, AT3, & AT4+AT5

$$Y_{FAO,tot}(n, mc, fz) = \begin{cases} Y_{FAO,RF}(n, mc, fz) \times \frac{Y_{FAO}(n, mc, tot)}{Y_{FAO}(n, mc, RF)} \\ \bar{Y}_{FAO}(n, mc) \times \frac{1}{94} \sum_{n=1}^{94} \frac{Y_{FAO,tot}(n, mc, fz)}{Y_{FAO}(n, mc, RF)}, \text{ if } Y_{FAO,RF}(n, mc, fz) = 0 \end{cases}, \text{ where}$$

If total rainfed yield is zero (i.e., FAO reports that for a particular crop and country, the crop is entirely irrigated or found in the gleysol/fluvisol AEZ), then use estimates based on global average, as described in greater detail below in 3.2.4.2 (Note that the estimation of global-average yields in section 3.2.4.2 is executed prior to the calculation above for zero total rainfed yields).

3.2.4.2. Estimate yields for countries without FAO data.

As FAO data were available for only 94 countries, we estimated information for the remaining countries using averages calculated over the 94 countries and applying them to the national statistics for the remaining countries (Fig. 9). Note that we did not average the yields themselves, but rather the proportion of yield in each AEZ to national total yields.

For each country 'm', without FAO data by AEZ,

$$Y_{FAO,tot}(m, mc, fz) = \bar{Y}_{FAO}(m, mc) \times \frac{1}{94} \sum_{n=1}^{94} \frac{Y_{FAO,tot}(n, mc, fz)}{Y_{FAO}(n, mc, tot)}, \text{ where}$$

\bar{Y}_{FAO} = FAOSTAT national statistics on crop yields, and

3.2.4.3. Merge the data and adjust for consistency with SAGE harvested area.

The yields from the 94 countries are merged with the estimated yields for the remaining countries,

$$Y_{FAO,tot}(l, mc, fz) = Y_{FAO,tot}(n, mc, fz) \cup Y_{FAO,tot}(m, mc, fz)$$

FAO does not report yields for the three SAGE AEZs that have 0-60 day LGP (see Table 3). Furthermore, often the FAO yield data and SAGE harvested area data are inconsistent, with FAO

reporting non-zero yields even though SAGE reports zero harvested areas, and conversely, FAO reporting zero yields while SAGE reports non-zero harvested area. In all of these cases, we adjusted the FAO yield data to match the SAGE harvested area data.

We first map the FAO yield data from FAO's AEZs to SAGE AEZs based on Table 3. To fill in gaps in FAO yield data (i.e., zero reported yields when SAGE harvested area is non-zero), we estimate yields using a regression across all countries and all crops of yields in each rainfed AEZ to total rainfed yields (Fig. 8). For SAGE AEZ1, AEZ7, and AEZ13 (0-60 day LGP, with no data reported by FAO), we assumed that yield is one-tenth of the total rainfed yield for the corresponding crop and country. In other words,

$$Y_{FAO,tot}(l, mc, z) \begin{cases} \leftarrow Y_{FAO,tot}(l, mc, fz), \text{ based on table 3} \\ = 0, \text{ if } A_{SAGE}(l, mc, z) = 0 \end{cases}$$

If ($Y_{FAO,tot}(l, mc, z) = 0$ & $A_{SAGE}(l, mc, z) \neq 0$), $Y_{FAO,tot}(l, mc, z) = \alpha_z Y_{FAO}(l, mc, RF)$, where

$$\begin{aligned} \alpha_z = \alpha_{fc} &= 0.10 \text{ for AEZ1, AEZ7, and AEZ13;} \\ &= 0.50 \text{ for AT1;} \\ &= 0.87 \text{ for AT2;} \\ &= 1.30 \text{ for AT3; and} \\ &= 0.82 \text{ for AT4+AT5 (based on figure 8, and table 3).} \end{aligned}$$

3.2.4.4. Recalibrate the yield data to Year 2001 and map to GTAP crop sectors

Finally, because the harvested area data by AEZ are taken from SAGE, and the yield data are taken from FAO, the re-calculated national total yields will change and needs to be re-calibrated. Also, the yields need to be calibrated to the reference period of 2001. We do this as follows:

$$Y(l, mc, z) = Y_{FAO,tot}(l, mc, z) * \frac{P_{FAO}(l, mc, t_{ref})}{\sum_z Y_{FAO,tot}(l, mc, z) * A_{SAGE}(l, mc, z)}, \text{ where}$$

$P_{FAO}(n, c, t_{ref})$ = the FAOSTAT national production for $t_{ref} = 2001$.

This data can then be mapped onto GTAP's 8 crop sectors.

$$Y(l, mc, z) \rightarrow Y_{GTAP}(l, s, z) \text{ (see table 4 for mapping)}$$

4. Conclusions

In this study, we have combined several global land use data sets to derive land use information at the national level, by 18 different AEZs (Table 5) for use in the GTAP database. In particular, we have utilized the following global land use/land cover data sets developed by SAGE: spatially-explicit maps of croplands, pastures, built-up areas, potential natural vegetation, 18 major crops (plus Other crops), agro-ecological zones, national boundaries; and developed by FAO: harvested area, yield, and production for 94 developing nations, for 34 crops, by 6 AEZs within each nation, and FAOSTAT national statistical data for all crops and countries for 2001. These data sets were synthesized, adjusted for consistency, and calibrated to the year 2001 (the latter only for the major crops).

The results from our synthesis seem reasonable. For example, the land cover data set (figure 7) shows that forests dominate in LGP3 (120-179 days) and LGP6 (> 300 days), corresponding to boreal forests and tropical rainforests respectively. Shrub lands and pastures dominate in LGP1 (the driest AEZ) and their areas decrease as the AEZs get more humid. Savanna/grasslands are distributed fairly uniformly across all AEZs. Croplands are distributed with slightly higher proportions in LGP3 and LGP4 (i.e., in areas that are not too dry, but are not heavily forested) (see [Ramankutty *et al.*, 2002a] for a study on climatic constraints on cropland distribution). Built-up lands predominate in LGP4 and LGP5 (also the temperate regions of the world; [Small, 2003], but their total area is very small. The “otherland” category, which includes tundra, desert, and polar desert/rock/ice, is dominant in LGP1 (with some additional area in LGP2), as would be expected. The major crops data set (figure 8; only a few crops shown) shows that while rice and soy dominate in humid climates, cassava is grown in intermediate-to-humid climates, maize and pulses are mostly grown in intermediate climates, while millet dominates in semi-arid climates.

The SAGE land use data sets, integrated with ancillary information from FAO, have been customized for the GTAP database. The database provides several land cover and land use variables (table 5), for 160 nations of the world, and for 18 agro-ecological zones within each nation. The data will be valuable for use by global integrated assessment models examining the role played by land use in GHG mitigation strategies.

References

- Alexandratos, N., *World Agriculture Towards 2010*, 488 pp., Food and Agric. Organ. of the United Nations, Rome, 1995.
- Bonan, G.B., Frost followed the plow: impacts of deforestation on the climate of the United States, *Ecological Applications*, 9 (4), 1305-1315, 1999.
- Bonan, G.B., Observational evidence for reduction of daily maximum temperature by croplands in the Midwest United States, *Journal of Climate*, 14 (11), 2430-2442, 2001.
- Brovkin, V., A. Ganopolski, M. Claussen, C. Kubatzki, and V. Petoukhov, Modelling climate response to historical land cover change, *Global Ecology and Biogeography*, 8 (6), 509-517, 1999.
- FAO, FAOSTAT data, Food and Agriculture Organization of the United Nations, (Available at <http://apps.fao.org>), 2004.
- Fischer, G., H. van Velthuis, S. Medow, and F. Nachtergaele, Global Agro-Ecological Assessment for Agriculture in the 21st century [CD-ROM and web site <http://www.iiasa.ac.at/Research/LUC/SAEZ/index.html>], Food and Agricultural Organization/International Institute for Applied Systems Analysis (FAO/IIASA), (Available at <http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm>), 2002.
- Fischer, G., H. van Velthuis, F. Nachtergaele, and S. Medow, Global Agro-Ecological Zones (Global - AEZ) [CD-ROM and web site <http://www.fao.org/ag/AGL/agll/gaez/index.htm>], Food and Agricultural Organization/International Institute for Applied Systems Analysis (FAO/IIASA), (Available at <http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm>), 2000.
- Foley, J.A., M.H. Costa, C. Delire, N. Ramankutty, and P. Snyder, Green Surprise? How terrestrial ecosystems could affect earth's climate, *Frontiers in Ecology and the Environment*, 1 (1), 38-44, 2003.
- GTAP/MIT, Incorporation of land use and greenhouse gas emissions into the GTAP data base, Co-organized by the Massachusetts Institute of Technology and Center for Global Trade Analysis (GTAP), Purdue University, (Available at http://www.gtap.agecon.purdue.edu/databases/projects/Land_Use_GHG/MIT_Workshop/default.asp), 2002.
- Hannah, L., D. Lohse, C. Hutchinson, J.L. Carr, and A. Lankerani, A preliminary inventory of human disturbance of world ecosystems, *Ambio*, 23 (4-5), 246-250, 1994.
- Houghton, R.A., Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000, *Tellus Series B-Chemical and Physical Meteorology*, 55 (2), 378-390, 2003.
- IPCC, IPCC Special Report on Land Use, Land-Use Change and Forestry, edited by R.T. Watson, I.R. Noble, B. Bolin, N.H. Ravindranath, D.J. Verardo, and D.J. Dokken, Cambridge University Press, Cambridge, 2000.
- Lee, H.-L., Comparison of Land Cover Data from Various Specialist Groups, pp. 44, Purdue University, Lafayette, 2002.
- Leff, B., N. Ramankutty, and J. Foley, Geographic distribution of major crops across the world, *Global Biogeochemical Cycles*, 18, GB1009, doi:10.1029/2003GB002108, 2004.
- Loveland, T.R., B.C. Reed, J.F. Brown, D.O. Ohlen, J. Zhu, L. Yang, and J.W. Merchant, Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data, *International Journal of Remote Sensing*, 21 (no. 6/7), 1303-1330, 2000.
- 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.M. III, I.C. Prentice, N. Ramankutty,

- T. Reichenau, A. Schloss, H. Tian, L.J. Williams, and U. Wittenberg, 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, 2001.
- Myhre, G., and A. Myhre, Uncertainties in radiative forcing due to surface albedo changes caused by land-use changes, *Journal of Climate*, 16 (10), 1511-1524, 2003.
- National Geographic Maps, A World Transformed, Supplement to National Geographic September 2002, National Geographic Society, Washington, D.C., 2002.
- Postel, S.L., G.C. Daily, and P.R. Ehrlich, Human appropriation of renewable fresh water, *Science*, 271, 785-788, 1996.
- Prentice, I.C., G. Farquhar, M. Fashm, M. Goulden, M. Heimann, V. Jaramillo, H. Khesghi, C. Le Quéré, and R.J. Scholes, The carbon cycle and atmospheric carbon dioxide, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson, pp. 183-237, Cambridge University Press, Cambridge, 2001.
- Ramankutty, N., and J.A. Foley, Characterizing patterns of global land use: An analysis of global croplands data, *Global Biogeochemical Cycles*, 12, 667-685, 1998.
- Ramankutty, N., and J.A. Foley, Estimating historical changes in global land cover: Croplands from 1700 to 1992, *Global Biogeochemical Cycles*, 13, 997-1027, 1999.
- Ramankutty, N., J.A. Foley, J. Norman, and K. McSweeney, The global distribution of cultivable lands: current patterns and sensitivity to possible climate change, *Global Ecology and Biogeography*, 11 (5), 377-392, 2002a.
- Ramankutty, N., J.A. Foley, and N.J. Olejniczak, People on the land: Changes in Population and Global Croplands During the 20th Century, *Ambio*, 31 (3), 251-257, 2002b.
- Richards, J.F., Land transformation, in *The Earth as Transformed by Human Action*, edited by B.L. Turner, W.C. Clark, R.W. Kates, J.F. Richards, J.T. Mathews, and W.B. Meyer, pp. 163-178, Cambridge Univ. Press, New York, 1990.
- Richards, K., and C. Stokes, A Review of Forest Carbon Sequestration Cost Studies: A Dozen Years of Research, *Climatic Change*, 63 (1-2), 1-48, 2004.
- Small, C. (2003) Global Population Distribution and Urban Land Use in Geophysical Parameter Space. *Earth Interactions*, 8(Paper 8), doi: 10.1175/1087-3562(2004)008<0001:GPDAUL>2.0.CO;2.
- Tans, P.P., I.Y. Fung, and T. Takahashi, Observational constraints on the global atmospheric CO₂ budget, *Science*, 247, 1431-1438, 1990.
- US-EPA/GTAP, Land Use and Integrated Assessment Workshop, Co-organized by the U.S. Environmental Protection Agency (US-EPA) and Center for Global Trade Analysis (GTAP), Purdue University, (Available at http://www.gtap.agecon.purdue.edu/databases/projects/Land_Use_GHG/EPA_Workshop_2004/EPA_Workshop_May2004.asp), 2004.
- Vitousek, P.M., H.A. Mooney, J. Lubchenco, and J.M. Melillo, Human domination of Earth's ecosystems, *Science*, 277, 494-499, 1997.
- Yang, D.W., S. Kanae, T. Oki, T. Koike, and K. Musiak, Global potential soil erosion with reference to land use and climate changes, *Hydrological Processes*, 17 (14), 2913-2928, 2003.

Tables

Table 1. GTAP land use matrix

| AEZs | Land use types in region r | | | | | | | | |
|------------------|----------------------------|------|-------------------|------------------------|------|------------------------|---------------------|------|---------------------|
| | Crop ₁ | | Crop _N | Livestock ₁ | | Livestock _H | Forest ₁ | | Forest _v |
| AEZ ₁ | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| AEZ _M | | | | | | | | | |
| Total | | | | | | | | | |

Table 2. Mapping from FAO crops to SAGE crops

| No. | FAO crops | No. | SAGE crops |
|-----|-----------|-----|------------|
| 1 | WHEA | 19 | Wheat |
| 2 | RICE | 12 | Rice |
| 3 | MAIZ | 5 | Maize |
| 4 | BARL | 1 | Barley |
| 5 | MILL | 6 | Millet |
| 6 | SORG | 14 | Sorghum |
| 7 | OTHC | 13 | Rye |
| 8 | POTA | 9 | Potato |
| 9 | SPOT | 8 | Others |
| 10 | CASS | 2 | Cassava |
| 11 | OTHR | 8 | Others |
| 12 | BEET | 16 | Sugar beet |
| 13 | CANE | 17 | Sugar cane |
| 14 | PULS | 10 | Pulses |
| 15 | VEGE | 8 | Others |
| 16 | BANA | 8 | Others |
| 17 | CITR | 8 | Others |
| 18 | FRUI | 8 | Others |
| 19 | OILC | 8 | Others |
| 20 | RAPE | 11 | Rape |
| 21 | PALM | 7 | Oilpalm |
| 22 | SOYB | 15 | Soy |
| 23 | GROU | 4 | Groundnuts |
| 24 | SUNF | 18 | Sunflower |
| 25 | SESA | 8 | Others |
| 26 | COCN | 8 | Others |
| 27 | COFF | 8 | Others |
| 28 | TEAS | 8 | Others |
| 29 | TOBA | 8 | Others |
| 30 | COTT | 3 | Cotton |
| 31 | FIBR | 8 | Others |
| 32 | RUBB | 8 | Others |

Table 3. Mapping from FAO AEZs to SAGE AEZs

| SAGE AEZs | FAO AEZs | Description |
|----------------------------|----------------------|--------------------|
| AEZ1, AEZ7, AEZ13 | Estimated (see text) | |
| AEZ2, AEZ8, AEZ14 | AT1 | Dry Semi-arid |
| AEZ3, AEZ9, AEZ15 | AT2 | Moist Semi-arid |
| AEZ4, AEZ10, AEZ16 | AT3 | Sub-humid |
| AEZ5, AEZ11, AEZ17 | AT4+AT5 | Humid |
| AEZ6, AEZ12, AEZ18 | AT4+AT5 | Humid |
| No separate AEZ (see text) | AT6+AT7 | Fluvisols/Gleysols |
| No separate AEZ (see text) | | Irrigated Land |

Table 4. Mapping from SAGE crops to GTAP crop sectors

| SAGE crop name | GTAP sector No. | GTAP code | Description |
|-----------------|-----------------|-----------|-------------------------|
| Barley | 3 | gro | Cereals grain n.e.c. |
| Cassava | 4 | v_f | Vegetables, fruit, nuts |
| Cotton | 7 | pfb | Plant-based fibres |
| Groundnuts | 5 | osd | Oil seeds |
| Maize | 3 | gro | Cereals grain n.e.c. |
| Millet | 3 | gro | Cereals grain n.e.c. |
| Oilpalm | 5 | osd | Oil seeds |
| Others | 8 | ocr | Crops n.e.c. |
| Potato | 4 | v_f | Vegetables, fruit, nuts |
| Pulses | 4 | v_f | Vegetables, fruit, nuts |
| Rape | 5 | osd | Oil seeds |
| Rice | 1 | pdr | Paddy rice |
| Rye | 3 | gro | Cereals grain n.e.c. |
| Sorghum | 3 | gro | Cereals grain n.e.c. |
| Soy | 5 | osd | Oil seeds |
| Sugar beet | 6 | c_b | Sugar cane, sugar beet |
| Sugar cane | 6 | c_b | Sugar cane, sugar beet |
| Sunflower seeds | 5 | osd | Oil seeds |
| Wheat | 2 | wht | Wheat |

Table 5. Summary of the SAGE land use data set provided to GTAP

| Category | Variables | Specifications | Reference Period |
|-------------|--|--|------------------|
| Land Cover | Forest, Savanna/Grassland, Shrubland, Cropland, Pasture, Built-up land, Other land | 160 countries 18 AEZs within each country | ca. 1992 |
| Major Crops | 18 SAGE crops (+ Other crops) & 8 GTAP crop sectors | | 2001 |

Figures

Figure 1

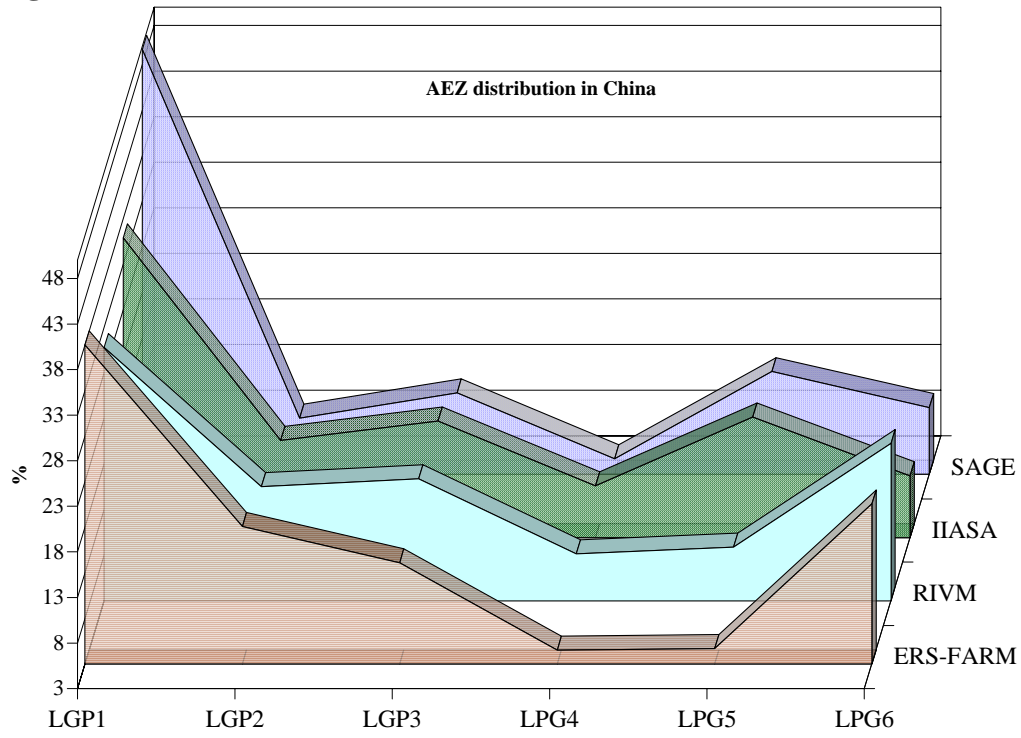


Figure 2

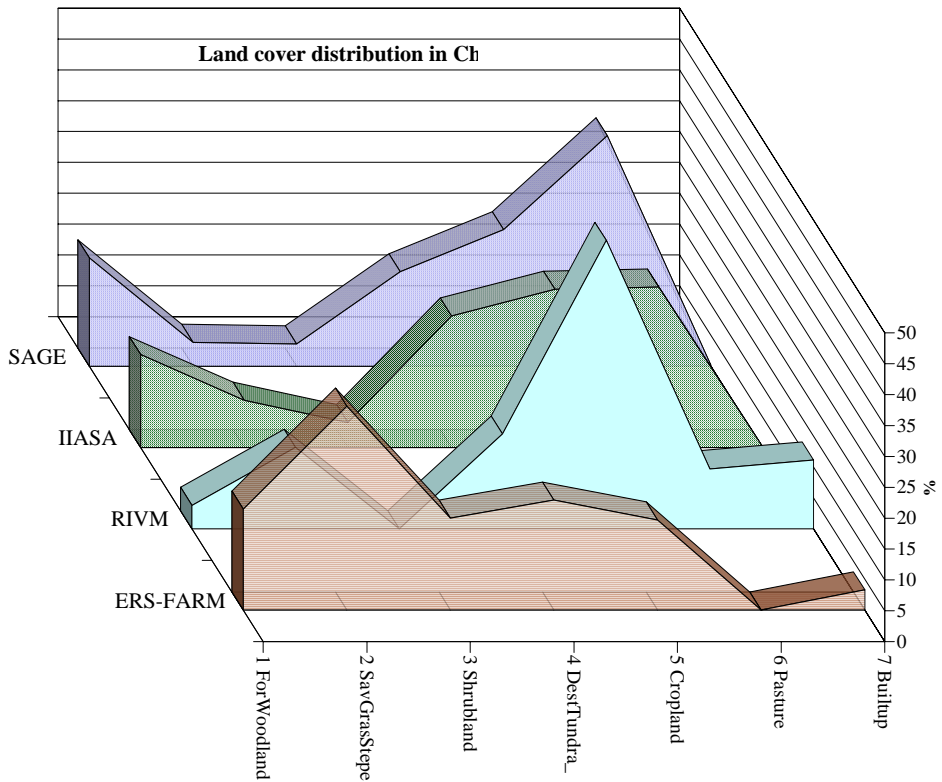


Figure 3



Figure 4

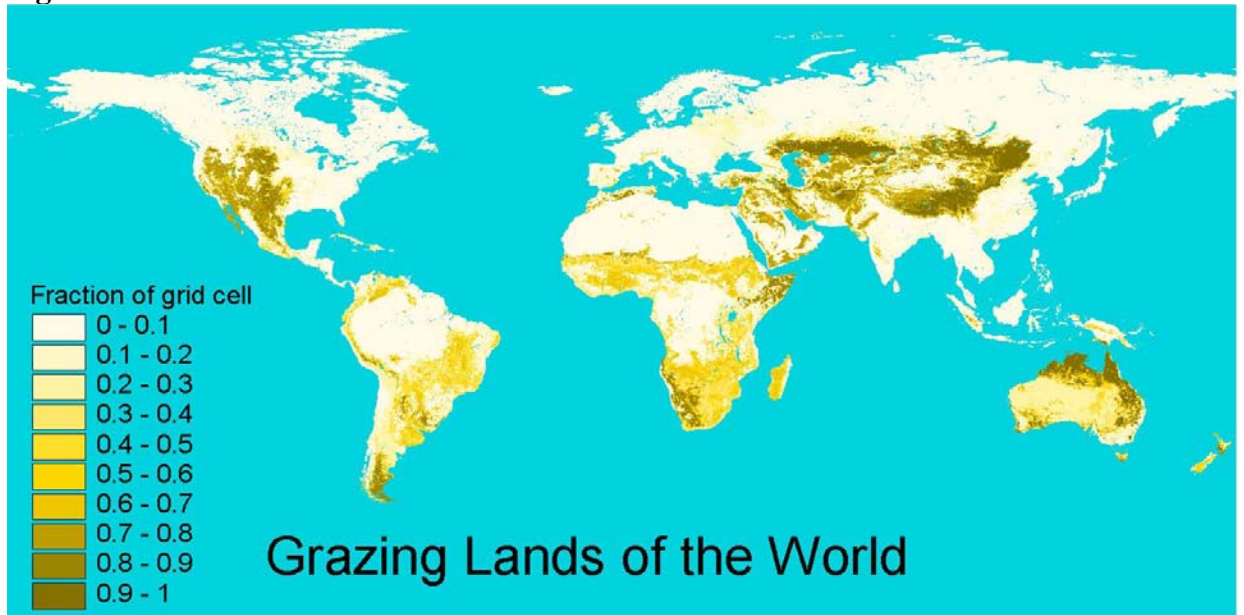


Figure 5

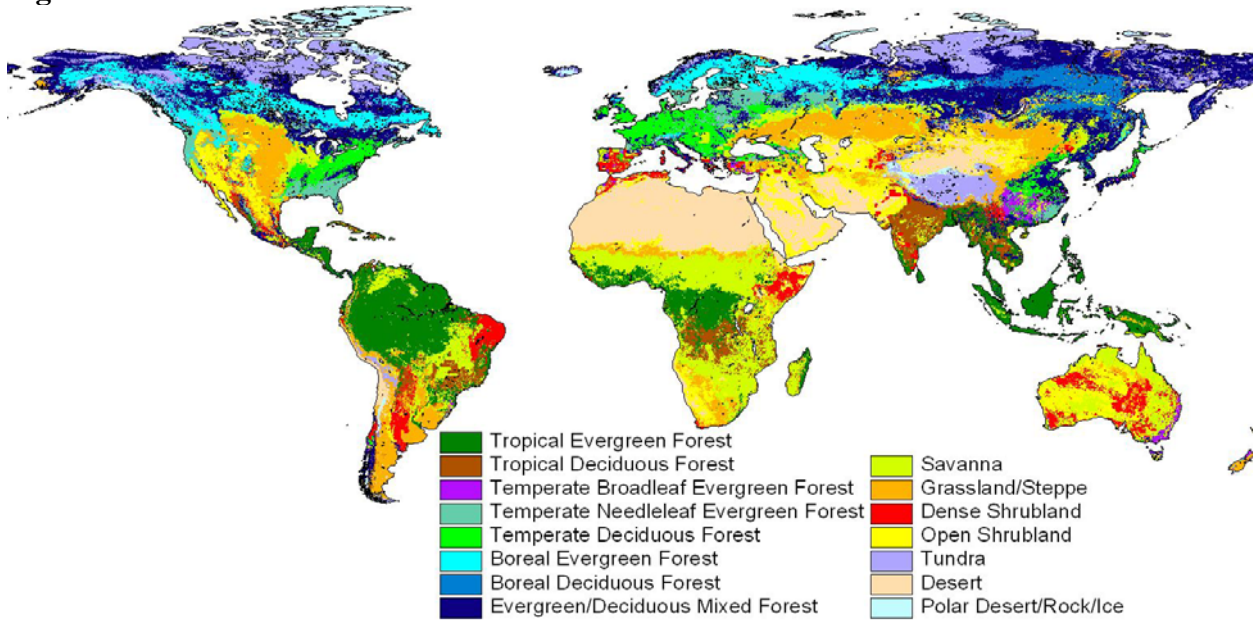


Figure 6

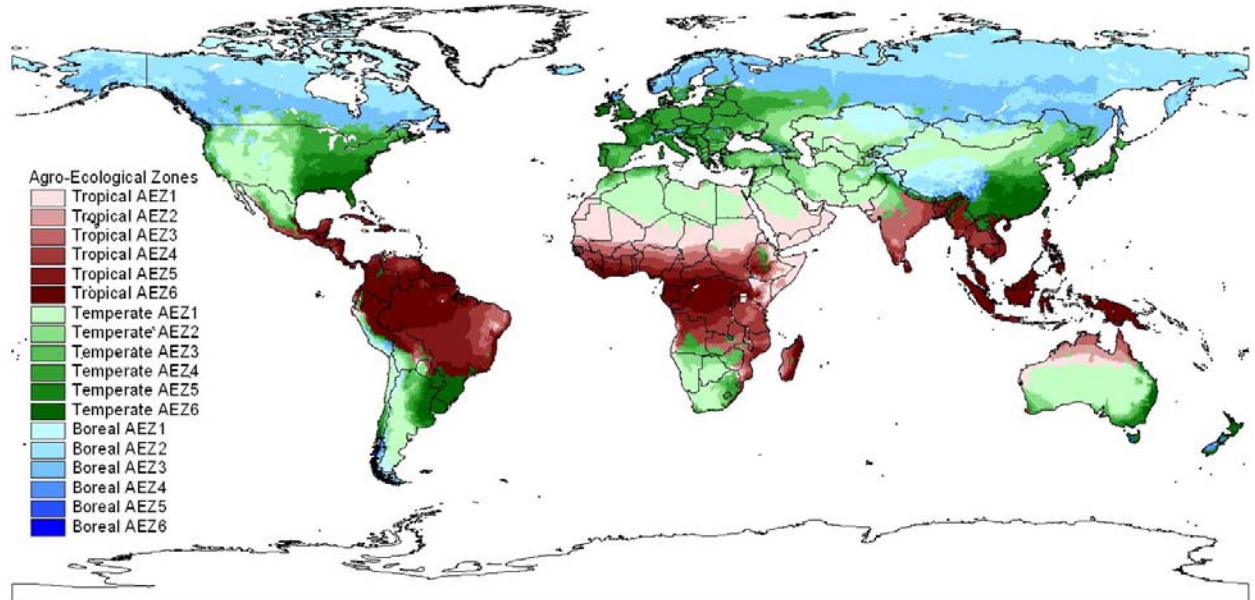


Figure 7

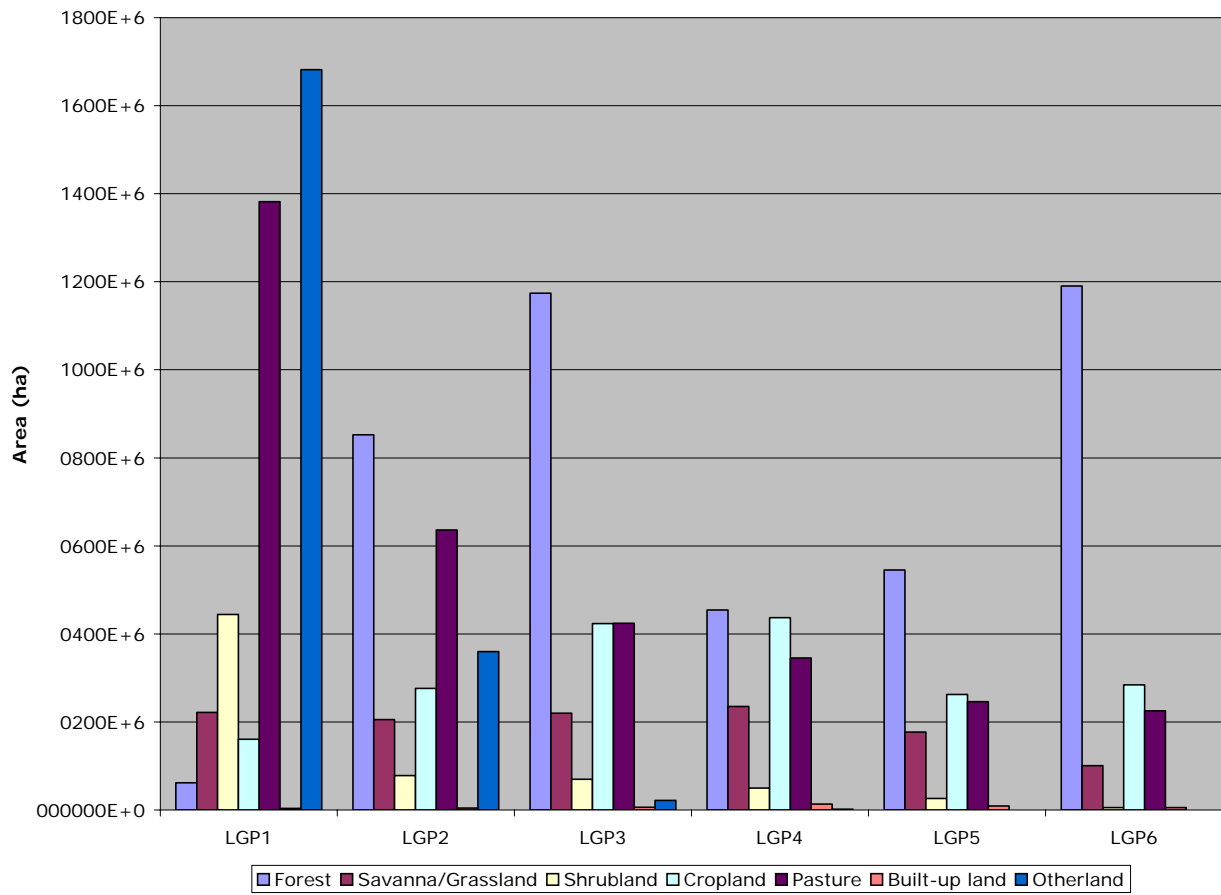


Figure 8

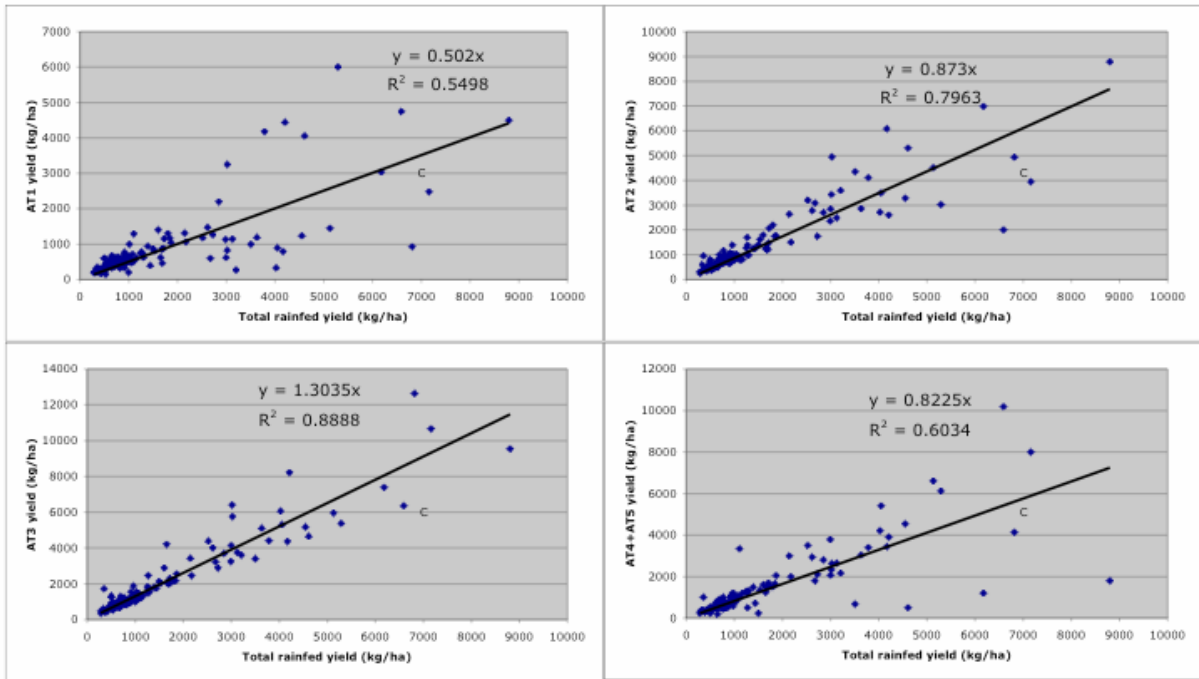


Figure 9

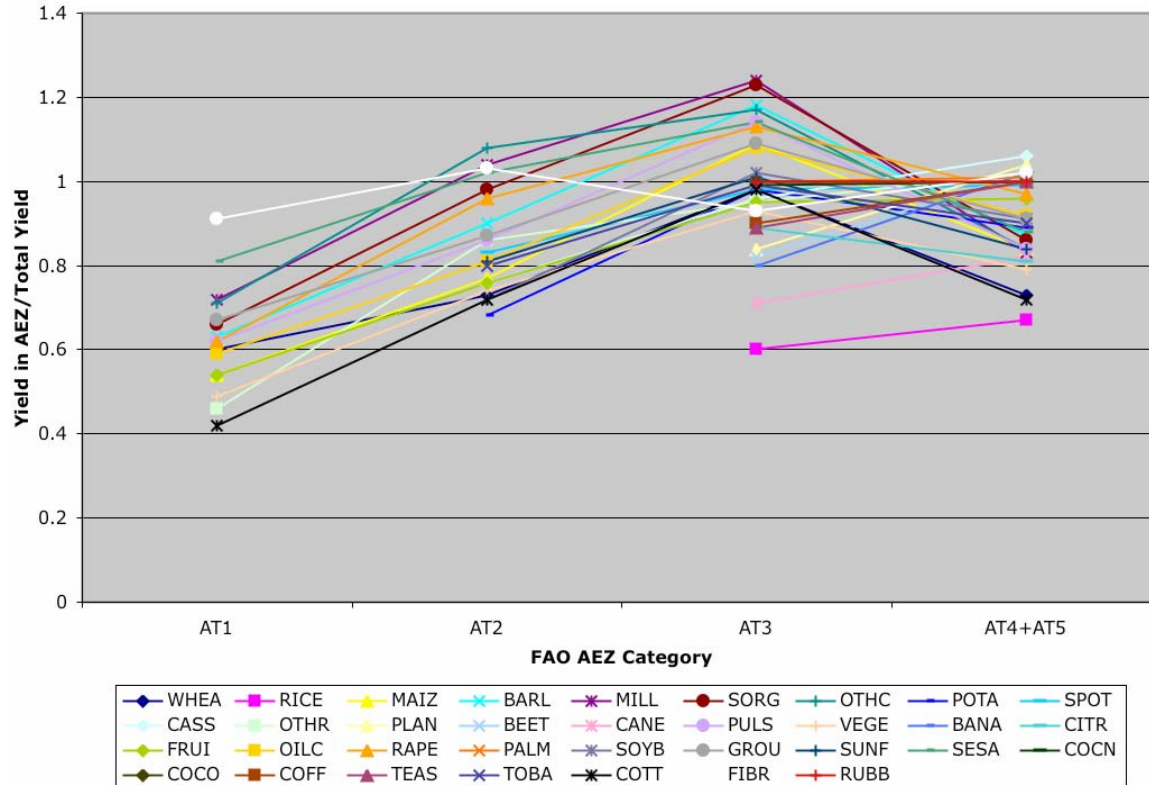


Figure 10

