1. TITLE

1.1 Data Set Identification

ISLSCP II Historical Croplands Cover (1700-1992).

1.2 Database Table Name(s)

Not applicable to this data set.

1.3 File Name(s)

The files in this data set are named **historic_cropland_XX_YYYY.asc**, where XX can be either hd or 1d, meaning a spatial resolution of 0.5 and 1.0 degree in both latitude and longitude, respectively and YYYY is the year from 1700 to 1992. For each spatial resolution, 24 files with global croplands fraction cover from 0 to 1 are provided for every 50 years from 1700 to 1850, every 10 years from 1850 to 1980, and every year from 1986 to 1992. An additional file called **historic_cropland_hd_chngmp.asc** is also provided that shows the differences between the original land/water mask and the land/water mask used in this collection. Note that the land/water mask difference file is only available at a 0.5 degree spatial resolution.

1.4 Revision Date of this Document

February 8, 2010

2. INVESTIGATOR(S)

2.1 Investigator(s) Name and Title

Dr. Navin Ramankutty
Dept of Geography & Earth System Science Program
McGill University

Dr. Jonathan A. Foley
Institute on the Environment
University of Minnesota

University of Wisconsin-Madison

2.2 Title of Investigation

Global land use data reconstruction.

2.3 Contacts (For Data Production Information)

	Contact 1	Contact2
2.3.1 Name	Dr. Navin Ramankutty	Dr. Jonathan A. Foley
2.3.2 Address	Dept of Geography & Earth System Science Program McGill University 627 Burnside Hall 805 Sherbrooke Strt W.	Institute on the Environment 1954 Buford Ave 325 Vo Tech Bldg St Paul, MN
City/St.	Montreal, QC,	55108
Zip Code	H3A 2K6	U.S.A.
Country	Canada	
2.3.3 Tel. No.	(514) 398-8428	(612) 626-0553
Fax No.	(514) 398-7437	
2.3.4 E-mail	navin.ramankutty@mcgill.ca	Jfoley@umn.edu

	Contact 3
2.3.1 Name	Dr. Eric Brown de Colstoun
2.3.2 Address	NASA/GSFC Code 614.4
City/St.	Greenbelt, MD
Zip Code	20771
Country	USA
2.3.3 Tel. No.	(301) 614-6597
Fax No.	(301) 614-6695
2.3.4 E-mail	ericbdc@ltpmail.gsfc.nasa.gov

2.4 Data Set Citation

Ramankutty, N. and J.A. Foley. 2010. ISLSCP II Historical Croplands Cover, 1700-1992. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. Available on-line [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.doi:10.3334/ORNLDAAC/966

2.5 Requested Form of Acknowledgment

Users of the International Satellite Land Surface Climatology (ISLSCP) Initiative II data collection are requested to cite the collection as a whole (Hall et al. 2006) as well as the individual data sets. Please cite the following publications when these data are used:

Hall, F.G., E. Brown de Colstoun, G. J. Collatz, D. Landis, P. Dirmeyer, A. Betts, G. Huffman, L. Bounoua, and B. Meeson, The ISLSCP Initiative II Global Data sets: Surface Boundary Conditions and Atmospheric Forcings for Land-Atmosphere Studies, *J. Geophys. Res.*, 111, doi:10.1029/2006JD007366, 2006.

Ramankutty, N., and J.A. Foley (1999). Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13(4), 997-1027.

3. INTRODUCTION

3.1 Objective/Purpose

This data set was developed to understand the consequences of historical changes in land use and land cover for ecosystem goods and services. In particular, this data set can be used to study how global changes in cultivated areas have influenced climate, biogeochemical cycles, biodiversity, etc. This data set can be used directly within spatially-explicit climate and biogeochemical models.

3.2 Summary of Parameters

This is a gridded data set describing the fraction of each grid cell in the globe that is occupied by cultivated land from 1700 to 1992. Data layers are provided for every 50 years from 1700 to 1850, every 10 years from 1850 to 1980, and every year from 1986 to 1992.

3.3 Discussion

There are two sources of global land cover/land use data. The most recent estimates are derived from *satellite* measurements, and are available in a spatially-explicit fashion for roughly the last 30 years. The other estimate is based on ground-based sources such as census statistics, land surveys, estimates by historical geographers, etc. These *land inventory* data are only available at the scale of political units, but have the advantage of being historical.

Ramankutty and Foley (1998) derived a spatially-explicit data set of croplands in 1992 by synthesizing remotely-sensed land cover data with contemporary land inventory data. Furthermore, Ramankutty and Foley (1999) extended this data set into the past (back to 1700) using historical land inventory data.

The data set should only be used for continental-to-global scale analysis and modeling. The data set captures the broad patterns of cropland change over history, but not necessarily the fine details at local to regional scales – please check the data quality before using it at fine spatial scales. The quality of historical data for the Russian Federation is poor. The quality of data prior to 1850 is poor -- only continental-scale historical data were used for that period.

The original 0.5 degree data set submitted by the Principal Investigators is available from the Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin-Madison (http://www.sage.wisc.edu/mapsdatamodels.html). This original data set has been modified by the staff of the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection in order to match the land/water boundaries used in the collection. A separate file that shows the differences between the original 0.5 degree land/water boundaries and

the ISLSCP II land/water mask is provided. Finally, the ISLSCP II staff has created a 1.0 degree version of the data set by aggregating the 0.5 degree data set to the coarser resolution. This 1.0 degree version was produced so that all data sets within the ISLSCP II collection contained 1.0 degree versions of all data sets and should be considered a 'browse' product. Both these data sets are consistent with the global potential vegetation types data sets submitted by the Principal Investigators.

4. THEORY OF ALGORITHM/MEASUREMENTS

In the Seasonal Land Cover Regions (SLCR) classification scheme of the Loveland *et al*. (2000) global land cover data set, there are many classes that have some degree of crop cover. For example, of the 205 SLCR classes for North America, 64 include some form of croplands. In this study, we regrouped the SLCR legend into six categories (or labels) according to their degree of crop cover: (0) other vegetation, (1) other vegetation with crops, (2) other vegetation/crop mosaic, (3) crop/other vegetation mosaic, (4) crops with other vegetation, and (5) crops. Here "other vegetation" denotes all types of land cover other than croplands, including natural vegetation and other types of land use such as pastures and shifting cultivation. These groups arose naturally from the SLCR legend for all continents except the Australia-Pacific region; for the latter we had to use some subjective judgment.

Reclassifying the SLCR legend in this way gave us a 1 km resolution global map with six categories indicating the relative density of crop cover. However, this map is still qualitative. For instance, "crop/other vegetation mosaic" indicates that there are roughly equal amounts of crops and other vegetation in that pixel, but no information is available about the actual fractional extent of crop cover. To determine this information, we used statistical inventory data on global land cover. Land cover inventory data are available at the national and subnational level (or political unit level) from various international and national organizations, such as the Food and Agricultural Organization (FAO) and the United States Department of Agriculture (USDA). In this analysis, we use cropland inventory data at the national level for various countries available from the FAO, supplemented with data at the state/province/region level for the United States of America, Canada, Mexico, Brazil, Argentina, India, China, and Australia. A more complete collection of agricultural inventory data, including subnational data for all the large countries of the world, would be ideal. However, to our knowledge, no such data set has been compiled at this time.

To combine the inventory data sets with the 1 km land cover maps, we calibrated the amount of fractional crop cover implied by the six labels to the fractional cover implied by the inventory data for each political unit. In this procedure, we assumed that each of the six labels could have any fractional crop cover value ranging from 0.05 to 1.0, except for the label "crops" which was assigned a value of 1.0 and the label "other vegetation" which was assigned a value of 0.0. Although we allowed labels to imply the same fractional cover values, we did not allow vegetation-dominant labels to have higher fractions of crops than crop-dominated labels. For instance, "other vegetation with crops" could not have a higher fractional crop cover than "crop/other vegetation mosaic."

To perform the calibration of fractional crop cover values for the six labels, we evaluated all possible combinations of fractional cover (in intervals of 0.05) and calculated the aggregate area of croplands over each political unit for which inventory data are available. The total amount of cropland for each political unit implied by the satellite classification was compared to

agricultural inventory data using a simple linear regression. We selected the set of fractional crop cover values that yielded the best correlation coefficient when restricting the slopes to range from 0.9 to 1.1. (Nigeria was a significant outlier in Africa, and we left it out of the calibration. In Eurasia, the Russian Federation, which has a crop area an order of magnitude larger than the other countries, dominated the solution. Hence we left out Russia during the calibration.) Table 2 and Figure 1 in Ramankutty and Foley (1998) show the results of the calibration procedure. The calibration procedure was performed separately for each continent because the original satellite-based land cover data was produced separately for each continent. Finally, we aggregated the 1 km resolution fractional cover maps into a 5 minute resolution grid using a simple area-weighted averaging procedure.

To estimate historical crop cover change, we first compiled an extensive database of historical croplands at the national and subnational (state, province, etc.) level (henceforth referred to as "political unit" level) (see Appendix A of Ramankutty and Foley (1999) for details). The data were collected for 339 countries at the national level and for 8 countries at the subnational level, consistent with present-day political boundaries. With the exception of the Russian Federation, we have subnational information for most of the large countries, or countries with extensive croplands. The inventory data are obtained at 5-10 year intervals in the best situation, and often at much wider time intervals. We linearly interpolate in between data to obtain annual values. Often the data need adjustments for consistency, and these are described in Appendix A of Ramankutty and Foley (1999).

The 1992 croplands data set was then used as an initial condition for a simple land cover change model, which runs backward in time-generating historical land cover maps, using the historical crop inventory data as a constraint for each political unit (see figure in section 9.4). In other words, the land cover change model is merely a simple algorithm for spatially distributing the historical cropland inventory data within each political unit. However, there is insufficient global scale data of historical land use and land cover change to calibrate an elaborate model, much less to validate it. Hence, in our judgment, the simplest possible approach is most appropriate (employing the principle of Occam's razor). Our basic assumption was that within each political unit, the cropland pattern of 1992 represents the historical spatial patterns. The historical inventory data provides the temporal information needed to describe the differences in cropland area among the political units. Within each political unit, for each year in the past, we adjusted the spatial crop cover pattern of 1992 so that the cropland total for that unit matches the historical inventory data. This assumption will cause problems in large countries with no subnational information, but for several large countries (with the exception of the Russian Federation), we have subnational cropland inventory. The model simulations begin with the initial conditions for 1992 and simulate the crop cover backward in time, annually, until 1700. The reconstruction procedure is outlined in more detail in Appendix B of Ramankutty and Foley (1999).

5. EQUIPMENT

5.1 Instrument Description

The data set used the 1 km resolution land cover classification data developed at the EROS data center (Loveland *et al*., http://eros.usgs.gov/). The base satellite data for the 1 km land cover product were monthly composites from the Advanced Very High Resolution

Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) measurements for the April 1992-March 1993 period.

5.1.1 Platform (Satellite, Aircraft, Ground, Person)

Not applicable to this data set.

5.1.2 Mission Objectives

Not applicable to this data set.

5.1.3 Key Variables

Not applicable to this data set.

5.1.4 Principles of Operation

Not applicable to this data set.

5.1.5 Instrument Measurement Geometry

Not applicable to this data set.

5.1.6 Manufacturer of Instrument

Not applicable to this data set.

5.2 Calibration

5.2.1 Specifications

5.2.1.1 Tolerance

Not applicable to this data set.

5.2.2 Frequency of Calibration

Not applicable to this data set.

5.2.3 Other Calibration Information

Not applicable to this data set.

6. PROCEDURE

6.1 Data Acquisition Methods

The satellite-derived land cover classification data was obtained from the following web site: http://eros.usgs.gov/#Find_Data/Products_and_Data_Available/NDVI

The historical census data were obtained from various sources – world wide web, published census reports, publications by historical geographers, etc. The various sources are listed in Appendix A of the publication, Ramankutty and Foley (1999).

6.2 Spatial Characteristics

6.2.1 Spatial Coverage

The data coverage is global, except for Greenland, Antarctica and some islands.

6.2.2 Spatial Resolution

The data are provided in two equal-angle latitude/longitude Earth grids with spatial resolutions of 0.5 by 0.5 and 1.0 by 1.0 degree in both latitude and longitude.

6.3 Temporal Characteristics

6.3.1 Temporal Coverage

From 1700 to 1992.

6.3.2 Temporal Resolution

The temporal resolution of the data set is yearly. Data are provided for every 50 years from 1700 to 1850, every 10 years from 1850 to 1980, and every year from 1986 to 1992. The original data set described in Ramankutty and Foley (1999) has data for every year from 1700 to 1992. The original data set which contains data for every year from 1700 to 1992 available from SAGE at http://www.sage.wisc.edu/mapsdatamodels.html

7. OBSERVATIONS

7.1 Field Notes

Not applicable to this data set.

8. DATA DESCRIPTION

8.1 Table Definition with Comments

Not applicable to this data set.

8.2 Type of Data

8.2.1 Parameter/ Variable Name	8.2.2 Parameter/ Variable Description	8.2.3 Data Range	8.2.4 Units of Measurement	8.2.5 Data Source
Historical Croplands Fractional Cover	Fraction of each grid cell in croplands from 1700 to 1992.	0.0 to 1.0 Water=-9.999 No data over land=-8.888	Unitless	Ramankutty and Foley (1999)
Point Changed	Differences between the ISLSCP II land/water mask and the original data: -1 = ISLSCP II mask is water and original data is land (data removed) 0 = Data sets agree over land or water (data unchanged) 1 = ISLSCP II mask is land or water and original	-1 to 1	See 8.2.2	Original data and ISLSCP II land/water mask

	data is missing (fill value used).			

8.3 Sample Data Record

Not applicable to this data set.

8.4 Data Format

All of the files in the ISLSCP Initiative II data collection are in the Arc GIS ASCII Grid format. The file format consists of numerical fields of varying length, which are delimited by a single space and arranged in columns and rows. The files at different spatial resolutions each contain the following numbers of column and rows:

0.5 degree: 720 columns by 360 rows
1.0 degree: 360 columns and 180 rows

All files are gridded to a common equal-angle lat/long grid, where the coordinates of the upper left corner of the files are located at 180 degrees W, 90 degrees N and the lower right corner coordinates are located at 180 degrees E, 90 degrees S. Data in the map files are ordered from North to South and from West to East beginning at 180 degrees West and 90 degrees North. The files have all had the ISLSCP II land/water mask applied to them.

8.5 Related Data Sets

A historical land cover/land use data set has been developed by Klein Goldewijk (2000) and is provided in this collection. The Principal Investigators have also produced a global potential vegetation types data set (this collection) that is consistent with this data set and can be used together with this one. Several land cover data sets are provided in the ISLSCP II data collection.

9. DATA MANIPULATIONS

9.1 Formulas

9.1.1 Derivation Techniques/Algorithms

Ramankutty and Foley (1998) derived a global crop cover map for 1992 by calibrating the cropland categories of the 1km-resolution land cover data set of Loveland *et al*. (2000) against cropland inventory data for 1992. Ramankutty and Foley (1999) derived a historical croplands data set from 1700 to 1992, by extrapolating the 1992 croplands data into the past using historical cropland inventory (See Section 4 above and appendices of Ramankutty and Foley (1999) for more details).

9.2 Data Processing Sequence

9.2.1 Processing Steps and Data Sets

See Section 4 above and appendices of Ramankutty and Foley (1999).

9.2.2 Processing Changes

None.

9.2.3 Additional Processing by the ISLSCP II Staff

The original 0.5 degree files submitted contained data layers for every year from 1700 to 1992. The ISLSCP II staff has extracted layers for every 50 years from 1700 to 1850, every 10 years from 1850 to 1980, and every year from 1986 to 1992. In these original files both water bodies and missing data were assigned the same value of –9.999. The ISLSCP II staff has created two categories from this single category, assigning a value of –9.999 to all water bodies, and a value of –8.888 to missing data over land (e.g. Greenland, Antarctica). The ISLSCP II staff has also made the data set consistent with the 0.5 degree land/water mask used in the collection. Points where the original data showed water and the ISLSCP II mask showed land have been filled in, where possible, from an average of the cropland fraction of all surrounding cells in a 3 by 3 window, not including water or missing data. Points that did not have any values within this surrounding window have been assigned a value of –8.888 (i.e. No data over land). ISLSCP II water points have also been forced over original land points where needed. A separate file that shows the differences between the original 0.5 degree land/water boundaries and the 0.5 degree ISLSCP II land/water mask was produced. Finally, the ISLSCP II staff has created a 1.0 degree version of the data set by aggregating the 0.5 degree files to the coarser resolution. This was done by averaging the cropland fraction for the 4 0.5 degree cells contained within each 1.0 degree cell, ignoring any water or missing data cells. This 1.0 degree data set was also adjusted to match the water/missing data/land cells of the 1.0 degree potential vegetation data set submitted by the Principal Investigators. During this process, it was not possible to match 11 land cells from the potential vegetation data set with any actual cropland fraction cells so these were assigned the value of –8.888, or missing data over land.

9.3 Calculations

9.3.1 Special Corrections/Adjustments

See Ramankutty and Foley (1999) for more details.

9.4 Graphs and Plots

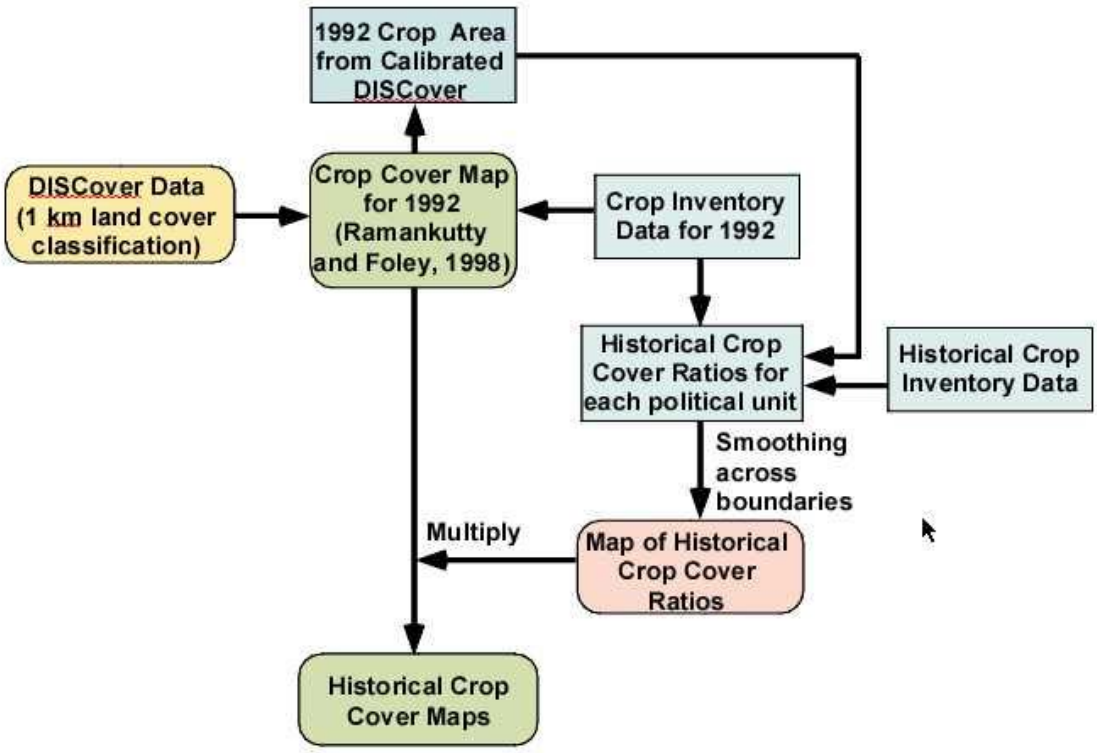

Figure 1. Algorithm for reconstructing historical crop cover maps. Boxes with sharp corners indicate data at the level of political units, while boxes with rounded corners indicate spatially-explicit maps. A crop cover map for 1992 is first derived by calibrating the DISCover data set against crop inventory data for 1992 (Ramankutty and Foley, 1998). Then the ratio of crop cover in the past to the crop cover in 1992 is derived for each political unit, and then further converted to a spatial map and smoothed across the boundaries of political units. The resulting map is multiplied by the crop cover map for 1992 to derive historical crop cover maps.

10. ERRORS

10.1 Sources of Error

Errors in both the satellite-derived land cover classification data set, and errors/biases in the cropland inventory data will affect the quality of this historical croplands data set. See Ramankutty and Foley (1999) for more details.

10.2 Quality Assessment

10.2.1 Data Validation by Source

A systematic quality assessment of the data set described here has not been performed. The original 1km DIScover global land cover data set has been validated from higher spatial resolution data (Scepan 1999), but only for its International Geosphere Biosphere Programme classification legend.

10.2.2 Confidence Level/Accuracy Judgment

Not available at this revision.

10.2.3 Measurement Error for Parameters and Variables

Not available at this revision.

10.2.4 Additional Quality Assessment Applied

None.

11. NOTES

11.1 Known Problems with the Data

In particular, we are aware of errors in the inventory data for China, Nigeria, and the Former Soviet Union that are yet to be resolved. Lack of subnational inventory data for the Russian Federation affects the quality of this data set in that region. The presence of extensive cultivation in Patagonia, and strips of cultivation in the Sahara likely reflects misclassifications in the satellite data.

11.2 Usage Guidance

The data are recommended for use only for continental-to-global scale studies over the timescale of decades to centuries. We note that the 1.0 degree version is a simple aggregation of the 0.5 degree data and is not produced according to the same methodology described in Section 4. We consider these 1.0 degree data a 'browse' product that is more suitable for visual assessment of the data and not necessarily scientific purposes. The user should always refer to the 0.5 degree data or the original data set at http://www.sage.wisc.edu/mapsdatamodels.html

11.3 Other Relevant Information

None.

12. REFERENCES

12.1 Satellite/Instrument/Data Processing Documentation

None.

12.2 Journal Articles and Study Reports

Klein Goldewijk, K., Estimating global land use change over the past 300 years: the HYDE database, *Global Biogeochemical Cycles*, Vol 15 (2), 417-434, 2000.

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.

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.

Scepan, J., 1999. Thematic Validation of High-Resolution Global Land-Cover Data Sets, *Photogrammetric Engineering and Remote Sensing*, v. 65, no. 9, p. 1051-1060.

13. DATA ACCESS

13.1 Contacts for Archive/Data Access Information

The ISLSCP Initiative II data are available are archived and distributed through the Oak Ridge National Laboratory (ORNL) DAAC for Biogeochemical Dynamics at http://daac.ornl.gov.

13.2 Contacts for Archive

E-mail: uso@daac.ornl
Telephone: +1 (865) 241-3952

13.3 Archive/Status/Plans

The ISLSCP Initiative II data are archived at the ORNL DAAC. There are no plans to update these data.

14. GLOSSARY OF ACRONYMS

AVHRR	Advanced Very High Resolution Radiometer
DAAC	Distributed Active Archive Center
FAO	Food and Agriculture Organization (United Nations)
GSFC	Goddard Space Flight Center
ISLSCP	International Satellite Land Surface Climatology Project
NDVI	Normalized Difference Vegetation Index
ORNL	Oak Ridge National Laboratory
SAGE	Center for Sustainability and the Global Environment
SLCR	Seasonal Land Cover Regions
USDA	United States Department of Agriculture