Documentation for the India Annual Winter Cropped Area, 2001-2016

November 2017

Meha Jain[1]*, Pinki Mondal[2], Gillian L. Galford[3], Greg Fiske[4], Ruth S. DeFries[5]

[1] School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109, USA
[2] Center for International Earth Science Information Network, Columbia University, Palisades, NY 10964, USA
[3] Gund Institute of Environment and Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT 05405, USA
[4] Woods Hole Research Center, Falmouth, MA 02540, USA
[5] Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY 10027, USA
* Correspondence: mehajain@umich.edu; Tel.: +1-734-764-5707

Abstract

This India Annual Winter Cropped Area, 2001 - 2016 consists of annual winter cropped areas for most of India (except the Northeastern states) from 2000-2001 to 2015-2016. This data set utilizes NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) Enhanced Vegetation Index (EVI; spatial resolution: 250m) for the winter growing season (October-March). The methodology uses an automated algorithm identifying the EVI peak in each pixel for each year and linearly scales the EVI value between 0% and 100% cropped area within that particular pixel. Maps were then resampled to 1 km and were validated using high-resolution QuickBird, RapidEye, SkySat, and WorldView-2 images spanning 2008 to 2016 across 11 different agricultural regions of India. The spatial resolution of the data set is 1 km, resampled from 250m. The data are distributed as GeoTIFF and NetCDF files and are in WGS 84 projection.

Data set citation:

Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. India Annual Winter Cropped Area, 2001-2016. Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H47D2S3W. Accessed DAY MONTH YEAR.

Suggested citation for this document:

Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. Documentation for the India Annual Winter Cropped Area, 2001-2016. Palisades NY: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/H43R0QSG. Accessed DAY MONTH YEAR.

We appreciate feedback regarding this data set, such as suggestions, discovery of errors, difficulties in using the data, and format preferences. Please contact:

NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network (CIESIN)
Columbia University
Phone: 1 (845) 365-8920
Email: ciesin.info@ciesin.columbia.edu

Contents

I. Introduction

The India Annual Winter Cropped Area, 2001- 2016 data set utilizes NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) Enhanced Vegetation Index (EVI) time-series to produce a collection of 16 gridded data files, one for each winter season during 2001-2016, at a spatial resolution of 1 km (aggregated from 250m). The grids were developed following an automated scaling technique (see Jain et al., 2017). The data files were validated using high-resolution satellite images and were compared with Indian agricultural census data. The maps cover most of India (except the Northeastern states).

Sustainable agricultural intensification is one of the ways to meet increasing food demand worldwide without further environmental degradation. Winter crops (from the secondary growing season) play an important role in ensuring food security, especially in developing countries. This new data set can be combined with socio-economic or meteorological data sets to better understand how, where, and why Indian farmers are switching between single and double crops. Thus, this data set will provide a useful social science application of remote sensing.

II. Data and Methodology

A detailed description of the methodology is provided in Jain et al., 2017.

Input data

MODIS Terra imagery-derived 16-day composites of 250m X 250m mosaicked EVI data were downloaded for all of India from June 2000 to May 2016 from the Columbia University International Research Institute for Climate and Society/Lamont-Doherty Earth Observatory (IRI/LDEO) Climate Data Library (http://iridl.ldeo.columbia.edu/). The MODIS Version 5 EVI products are compiled by the United States Geological Survey (USGS) at the Land Processes Distributed Active Archive Center (LP-DAAC).

Methods

MODIS EVI time-series are first smoothed using a cubic spline function to remove any high-frequency noise. Cropped pixels are then identified by examining EVI phenology during the winter growing season (October 1 – March 31). A pixel is identified as cropped if there is (i) an inflection point in the EVI time-series indicating green up after sowing, and (ii) a period of peak greenness within the growing season window (see Fig. 2 in Jain et al., 2017). This method captures only annual crops planted during winter (*rabi*) season, and not perennial crops or crops planted before winter (i.e. in rainy/monsoon season) that continue to grow throughout winter. These cropped pixels are then scaled between 0% and 100% cropped by identifying appropriate 0% and 100% cropped end-member pixels. The assumption here is that the highest and lowest EVI values in a given agro-ecological sub-region represent a pixel with 100% and 0% cropped area, respectively. See Jain et al., 2017 for details on the end-member selection procedure.

These cropped area maps are validated using single-date high-resolution QuickBird, RapidEye, SkySat, and WorldView-2 images (available for 2008 to 2016) across 11 agricultural regions of India representing a wide range in crop type, irrigation access, integration with market, and soil type. Cropped and uncropped pixels (50 each) are selected using visual inspection from the high-resolution imagery, and Normalized Difference Vegetation Index (NDVI) values are extracted for each of these 100 pixels. A regression tree is then used to identify the NDVI threshold that best differentiates between the cropped and uncropped pixels. The high-resolution images are then classified as cropped vs. non-cropped using this NDVI threshold value in a decision tree. These classified high-resolution images are aggregated to match the spatial resolution of the MODIS cropped area data products (1km X 1km), and the proportion of cropped area within each of these aggregated pixels are quantified. Five-hundred random points are then selected (1km X 1km) for comparing cropped area estimates derived from MODIS data products and high-resolution images. Comparison results are reported in R^2 and Root Mean Square Error (RMSE) – see Jain et al., 2017 for details. All analyses were conducted in R Project software (Vienna, Austria) using the raster, rgdal, rpart, and sp packages.

The accuracy of the data product varies based on location with R^2 values ranging from 0.19 (West Bengal) to 0.89 (Rajasthan) with an overall R2 of 0.71. RMSE ranged from 7.75 (Punjab) to 34.38 (Andhra Pradesh) with an overall RMSE of 18.47. Please see Jain et al., 2017 for an explanation for these varied R2 and RMSE values.

III. Data Set Description(s)

Data set description:
The winter cropped area grids consist of percent of area cropped per 1 km grid cells for each of the sixteen years between 2001 and 2016.

Data set web page:
http://sedac.ciesin.columbia.edu/data/set/india-india-annual-winter-cropped-area-2001-2016

Data set format:
The data are available in GeoTIFF format as downloadable zip files. Each downloadable is a compressed zip file containing: 1) GeoTIFF for the year, 2) PDF documentation, and 3) PDF and plain text R script source code. Two additional zip files are also available that provide the raster data in GeoTIFF and NetCDF format for all 16 years stacked together.

Data set downloads:
india-india-annual-winter-cropped-area-2001-2016-1km-XXXX-geotiff.zip

where XXXX are years 2001 through 2016

The 16-year stacked files:

- india-india-annual-winter-cropped-area-2001-2016-1km-2001-2016-geotiff.zip
- india-india-annual-winter-cropped-area-2001-2016-1km-2001-2016-netcdf.zip

IV. How to Use the Data

The raster data are in GeoTIFF format and can be used directly in mapping and geospatial analysis.

V. Potential Use Cases

The annual winter cropped area maps can be used to answer various biophysical and socio-economic research questions. For example, this data set can be used in conjunction with other remote sensing derived data sets to understand if agricultural burning before the winter season in preparation for winter planting correlates with air pollution during

winter. This data set can also be used to understand if surface irrigation (as opposed to groundwater irrigation) can sustainably support winter crop production in India.

VI. Limitations

This method does not capture perennial crops or crops planted before the winter season (or before October 1st of each year); these continue to grow throughout the winter. The methods described above may not be as accurate in regions with sparsely cropped fields, where approximately 10% or less of the MODIS pixel appears to be cropped based on high-resolution imagery. It is also likely that the cropped area estimates have increased error in regions with disparate crop types, intercropped fields, and/or fields with very different yields across space.

VII. Acknowledgments

Funding for development of this data set was provided by the U.S. National Aeronautics and Space Administration (NASA) Land Use and Land Cover Change (LCLUC) grant (NNX11AH98G) awarded to R. S. DeFries, an NSF Graduate Research Fellowship and a NASA New Investigator Award (NNH15ZDA001N) to M. Jain, and a Google Earth Engine Research Award to G. Galford. Validation data were provided by Jaime Nickeson and colleagues at NASA's NGA Commercial Data Archive Site (https://cad4nasa.gsfc.nasa.gov) and Planet and their Ambassador's program (https://www.planet.com/). Funding for dissemination of this data set was provided under the NASA contract NNG13HQ04C for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University.

VIII. Disclaimer

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at ciesin.info@ciesin.columbia.edu. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

IX. Use Constraints

X. Recommended Citation(s)

Data set(s):

Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. India Annual Winter Cropped Area, 2001-2016. Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H47D2S3W. Accessed DAY MONTH YEAR.

Scientific publication:

Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. An Automated Approach to Map Winter Cropped Area of Smallholder Farms across Large Scales Using MODIS Imagery. *Remote Sensing* 9 (6): 566. https://doi.org/10.3390/rs9060566.

XI. Source Code

The R codes used to develop this data set are being disseminated at https://doi.org/10.7927/H47D2S3W and are available as open source software.

Recommended citation for the Source Code:

Jain, M., P. Mondal, G. L. Galford, G. Fiske, and R. S. DeFries. 2017. Source Code for the India Annual Winter Cropped Area, 2001-2016. Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H47D2S3W. Accessed DAY MONTH YEAR.

XII. References

Columbia University International Research Institute for Climate and Society/Lamont-Doherty Earth Observatory (IRI/LDEO) Climate Data Library. http://iridl.ldeo.columbia.edu/.

XIII. Documentation Copyright and License

Appendix 1. Data Revision History

No revisions have been made to this data set.

Appendix 2. Contributing Authors & Documentation Revision History

Revision Date	Contributors	Revisions
November 8, 2017	Pinki Mondal	This document is the 1st instance of documentation