NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) @ 30-m for South America: Cropland Extent Product (GFSAD30SACE)

User Guide
## Document History

<table>
<thead>
<tr>
<th>Document Version</th>
<th>Publication Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>September, 2017</td>
<td>Original</td>
</tr>
<tr>
<td>1.1</td>
<td></td>
<td>Modification made according to USGS reviewer comments</td>
</tr>
</tbody>
</table>
Contents

Document History ............................................................................................................. 2
1.0 Dataset Overview ...................................................................................................... 4
  1.1 Background .............................................................................................................. 4
2.0 Dataset Characteristics ............................................................................................ 5
  2.1 Global Food Security Support Analysis Data (GFSAD) 30-m V001 ...................... 5
    2.1.1 Collection Level ............................................................................................... 5
    2.1.2 Granule Level .................................................................................................. 5
    2.1.3 Data Layer Characteristics ............................................................................. 5
    2.1.4 Data Layers Classification ............................................................................ 6
    2.1.5 Filename Convention .................................................................................... 6
3.0 Dataset Knowledge .................................................................................................... 6
  3.1 Frequently Asked Questions .................................................................................. 6
4.0 Dataset Access (Applicable Data Tools) .................................................................. 8
5.0 Contact Information .................................................................................................. 8
6.0 Citations ................................................................................................................... 9
  6.1 GFSAD30SACE ....................................................................................................... 9
7.0 Publications .............................................................................................................. 9
  7.1 Peer-reviewed publications within GFSAD project ............................................. 9
  7.2 Web sites and Data portals: .................................................................................. 10
  7.3 Other relevant past publications prior to GFSAD project .................................. 10
  7.4 Books and Book Chapters .................................................................................... 11
1.0 Dataset Overview

The goal of the Global Food Security-support Analysis Data @ 30-m (GFSAD30) project is to provide the highest resolution, objective cropland datasets to assist and address global food and water security issues in the twenty-first century. The project proposed developing cropland products using time-series Landsat and Sentinel satellite sensor data, machine learning algorithms, and cloud-based computing. The project is funded by the National Aeronautics and Space Administration (NASA) with supplemental funding from the United States Geological Survey (USGS). The project is led by USGS and carried out in collaboration with NASA AMES, University of New Hampshire (UNH), California State University Monterey Bay (CSUMB), University of Wisconsin (UW), NASA GSFC, and Northern Arizona University (NAU). There are a number of International partners, including The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

This user’s guide provides information about the GFSAD30 cropland extent product for the South American continent (GFSAD30SACE) at nominal 30m resolution for 2015. The Coordinate Reference System (CRS) used for the GFSAD30SACE is a geographic coordinate system (GCS) based on the World Geodetic System 84 (WGS84) reference ellipsoid. The legend is presented in Section 2. Datasets are provided as 1° x 1° tiles in GeoTIFF format. The year, resolution, tiling, and file name convention details are provided in section 2.0 of this document.

1.1 Background

Monitoring global croplands is imperative for ensuring sustainable water and food security for people of the world in the twenty-first century. However, the currently available cropland products suffer from major limitations such as: (1) the absence of precise spatial location of the cropped areas; (2) The coarse resolution nature of the map products with significant uncertainties in areas, locations, and detail; (3) The uncertainties in differentiating irrigated areas from rainfed areas; (4) The absence of crop types and cropping intensities; and/or (5) The absence of a dedicated Internet data portal for the dissemination of these cropland products. This project aims to address all of these knowledge gaps.

Satellite-derived cropland extent maps at high spatial resolution are necessary for food and water security analysis. Therefore, the GFSAD30SACE cropland extent product was produced at a resolution of 30-m for the entire continent of South America for the nominal year 2015 using Landsat-5, 7, and 8 time-series data. These data are part of a global data release, whereby seven different regions are made publically available. Global cropland extent data, indicating cropland and non-cropland areas, provide a working baseline dataset to develop higher-level products, such as crop watering method (irrigated or rainfed), cropping intensities (e.g., single, double, or continuous cropping), crop type mapping, cropland fallow, as well as, the assessment of cropland productivity (productivity per unit of land), and crop water productivity (productivity per unit of water or “crop per drop”). Uncertainties associated with cropland extent data have a cascading effect on all these higher-level cropland datasets.

Cloud-based geo-spatial computing platforms and satellite imagery offer opportunities for producing precise and accurate data of cropland extent and area that meet the spatial and temporal requirements for a broad set of applications. Such data can be a significant improvement compared to existing products, which tend to be coarser resolution, are often not representative of highly dynamic regions, and have a fixed set of cover classes. Cloud-based computing platforms such as Google Earth Engine and new earth-observing satellites like those in the Landsat constellation have brought significant improvements to land use/land cover (LULC) mapping and agriculture monitoring. Specifically, the production of standard static maps of the past will be shifted to dynamic creation of maps from massively large volumes of big data, crowd-sourcing of training and validation samples, and implementing machine learning algorithms on these computing clouds to better serve specific applications.
For a very detailed description of the satellite and reference data, processing schemes, approaches, methods, results, and conclusions of this project, please refer to the algorithm theoretical basis document (ATBD) of GFSAD30SACE.

## 2.0 Dataset Characteristics

Global food security-support analysis data @ 30-m cropland extent for the South American Continent (GFSAD30SACE) data product and characteristics are described below.

### 2.1 Global Food Security Support Analysis Data (GFSAD) 30-m V001

#### 2.1.1 Collection Level

<table>
<thead>
<tr>
<th>Short name</th>
<th>GFSAD30SACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Granularity</td>
<td>Static</td>
</tr>
<tr>
<td>Temporal Extent</td>
<td>2015, nominal</td>
</tr>
<tr>
<td>Spatial Extent</td>
<td>South America</td>
</tr>
<tr>
<td>File size</td>
<td>~1.6 GB</td>
</tr>
<tr>
<td>Coordinate System</td>
<td>Geographic</td>
</tr>
<tr>
<td>Datum</td>
<td>WGS84</td>
</tr>
<tr>
<td>File Format</td>
<td>GeoTIFF</td>
</tr>
</tbody>
</table>

#### 2.1.2 Granule Level

<table>
<thead>
<tr>
<th>Number of Layers</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns/Rows</td>
<td>180029 x 258858</td>
</tr>
<tr>
<td>Pixel Size</td>
<td>~30 ~m</td>
</tr>
</tbody>
</table>

#### 2.1.3 Data Layer Characteristics

<table>
<thead>
<tr>
<th>SDS Layer Name</th>
<th>Description</th>
<th>Units</th>
<th>Data Type</th>
<th>Fill Value</th>
<th>Valid Range</th>
<th>Scale Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1</td>
<td>Crop Extent for South American Continent</td>
<td>N/A</td>
<td>8-bit unsigned integer</td>
<td>N/A</td>
<td>0, 1, 2</td>
<td>N/A</td>
</tr>
</tbody>
</table>
2.1.4 Data Layers Classification

<table>
<thead>
<tr>
<th>Class Label</th>
<th>Class Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Non-Cropland</td>
<td>Non-Cropland areas</td>
</tr>
<tr>
<td>1</td>
<td>Cropland</td>
<td>Cropland and Fallow-land</td>
</tr>
<tr>
<td>2</td>
<td>Water</td>
<td>Water bodies</td>
</tr>
</tbody>
</table>

2.1.5 Filename Convention

GFSAD30SACE_2015_S10W60_001_2017261200520.tif = File name
GFSAD30SACE = Product Short name
30 = 30 m Spatial Resolution
SA = South America
CE = Crop Extent
2015 = Nominal Year
S10W60 = 10 x 10 degree grid, starting at (S10, W60)
001 = Version
2017261200520= Processing Date in YYYYJJJHHMMSS

3.0 Dataset Knowledge

The following questions address the user information regarding the GFSAD30SACE collection.

3.1 Frequently Asked Questions

What does the GFSAD30SACE product contain?

This product provides cropland extent for continental South America at nominal 30-m. It covers all 55 South American countries.

What is the definition of cropland extent?

For the entire Global Food Security-Support Analysis Data project at 30-m (GFSAD30) project, cropland extent was defined as: “lands cultivated with plants harvested for food, feed, and fiber, include both seasonal crops (e.g., wheat, rice, corn, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, oil palms). Cropland fallows are lands uncultivated during a season or a year but are farmlands and are equipped for cultivation, including plantations (e.g., orchards, vineyards, coffee, tea, rubber)” (Teluguntla et al., 2015). Cropland extent includes all planted crops and fallow lands. Non-croplands include all other land cover classes other than croplands and cropland fallows.
How can the dataset be obtained?

All the GFSAD30 products are downloadable through the Land Processes Distributed Active Archive Center (LP DAAC). GFSAD30SACE, divided into 10x10 grids, is among them. You can also visualize these data at: croplands.org by going to the “products” drop-down menu.

What is the accuracy of the GFSAD30SACE product?

For the entire continent, the weighted overall accuracy was 93.2% with producer’s accuracy of 82.6% (errors of omissions of 17.4%) and user’s accuracy of 76.7% (errors of commissions of 23.3%) (Table 1). When considering all 5 zones (Figure 1), the overall accuracies ranged between 92.4-96.8%, producer’s accuracies ranged between 79.6-90.9%, and user’s accuracies ranged between 76.7-81.3% (Table 1).

Table 1. Independent Accuracy Assessment of 30-m Cropland Extent Map for South America. Accuracies were assessed for each of the 5 zones as well as for the entire continent.

<table>
<thead>
<tr>
<th>Zone 1</th>
<th>Reference Data</th>
<th>% TCASA= 1.8%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>37.5%</td>
<td>50.0%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>6</td>
<td>244</td>
<td>250</td>
<td>98.8%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zone 2</th>
<th>Reference Data</th>
<th>% TCASA= 17.6%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>25</td>
<td>8</td>
<td>33</td>
<td>75.8%</td>
<td>83.3%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>30</td>
<td>220</td>
<td>250</td>
<td>97.7%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zone 3</th>
<th>Reference Data</th>
<th>% TCASA= 29.7%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>39</td>
<td>9</td>
<td>48</td>
<td>81.3%</td>
<td>79.6%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>49</td>
<td>200</td>
<td>250</td>
<td>95.1%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zone 4</th>
<th>Reference Data</th>
<th>% TCASA= 40.6%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>40</td>
<td>11</td>
<td>51</td>
<td>78.4%</td>
<td>90.9%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>44</td>
<td>206</td>
<td>250</td>
<td>98.0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zone 5</th>
<th>Reference Data</th>
<th>% TCASA= 10.3%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>21</td>
<td>6</td>
<td>27</td>
<td>77.8%</td>
<td>80.8%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>26</td>
<td>224</td>
<td>250</td>
<td>97.8%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>All Zones</th>
<th>Reference Data</th>
<th>% TCASA= 100%</th>
<th>Crop</th>
<th>No-Crop</th>
<th>Total</th>
<th>User Accuracy</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Data</td>
<td></td>
<td></td>
<td>128</td>
<td>39</td>
<td>167</td>
<td>76.7%</td>
<td>82.6%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>155</td>
<td>1,095</td>
<td>1,250</td>
<td>97.5%</td>
<td></td>
</tr>
</tbody>
</table>

Note: Total net cropland area of SA (TCASA) = 282.60Mha  

Area weighted accuracy: 93.2%
Figure 1. Stratification of the South American continent into seven distinct refined FAO agro-ecological broad zones. The figure also shows the distribution of the reference training and validation data used in the machine learning algorithms.

Can I obtain the dataset through Google Earth Engine (GEE)?

No. At this time we are not releasing the data on GEE. All data are released through LP DAAC. In the future, such a release will be considered by the PI (Prasad S. Thenkabail).

4.0 Dataset Access (Applicable Data Tools)

The GFSAD30SACE dataset is available through the LP DAAC Data Pool and NASA Earthdata Search. GFSAD data visualization and information can also be found at https://croplands.org.

5.0 Contact Information

LP DAAC User Services  
U.S. Geological Survey (USGS)  
Center for Earth Resources Observation and Science (EROS)  
47914 252nd Street  
Sioux Falls, SD 57198-0001

Phone Number: 605-594-6116
For the Principal Investigators, feel free to write to:

Prasad S. Thenkabail at pthenkabail@usgs.gov

For 30-m cropland extent product of South America, please contact:
Pardhasaradhi Teluguntla at pteluguntla@usgs.gov
Prasad S. Thenkabail at pthenkabail@usgs.gov
Ying Zhong at ying.zhong105@gmail.com
Chandra Giri at Giri.Chandra@epa.gov

More details about the GFSAD project and products can be found at: globalcroplands.org

6.0 Citations

6.1 GFSAD30SACE

7.0 Publications

7.1 Peer-reviewed publications within GFSAD project


7.2 Web sites and Data portals:
http://croplands.org (30-m global croplands visualization tool)
http://geography.wr.usgs.gov/science/croplands/products.html#LPDAAC (dissemination on LP DAAC)
croplands.org (crowdsourcing global croplands data)

7.3 Other relevant past publications prior to GFSAD project


### 7.4 Books and Book Chapters


