

NGDA Dataset Report

Official NGDA Title: CropScape - Cropland Data Layer

Metadata Record Title: CropScape - Cropland Data Layer

A-16 NGDA Theme: Land Use - Land Cover

Executive NGDA Theme Champion(s):

Name: Jerry Johnston

Agency: DOI

Email: jerry_johnston@ios.doi.gov

Name: Stephen Lowe

Agency: DOA

Email: stephen.lowe@ocio.usda.gov

Theme Lead(s):

Name: Jonathan Smith

Agency: USGS

Email: jhsmith@usgs.gov

Name: Greg Reams

Agency: USFS

Email: greams@fs.fed.us

Dataset Manager(s):

Name: Rick Mueller

Agency: USDA

Email: Rick.Mueller@nass.usda.gov

Metadata:

Registration Status: Complete

Registered on 2/2/2015

GeoPlatform Link*: <https://www.geoplatform.gov/node/243/a54f065c-17b8-4516-b309-618faa92fe18>

Data.gov Metadata Link*: [. The CDL product depicts agricultural land cover over the Continental United States at 30 meters resolution. CropScape includes historical data dating back to 1997 for some locales. The CDL is an annual raster, geo-referenced, crop-specific land cover da](#)

NGDA Lifecycle Maturity Assessment (LMA) Report

Time Frame:

CDL products dating back to 1997.

LMA Submission:

Status: Complete

Date: 10/8/2015

Extension Requested: Yes

LMA Reviewer(s):

Supervisor: Did not review

Theme Lead: Patrick Willis

Executive Champion: Did not review

SAOGI*: Did not review

Other: Did not review

LMA Verifier:

Name: Patrick Willis

Email: patrick.willis@nass.usda.gov

Attachments:

To get access to any attachments referenced in the report, email the LMA Help Desk at NGDA_LMA_help@fgdc.gov. Please use the subject "Dataset Report Attachment(s)" and indicate the associated official NGDA title.

*Senior Agency Official for Geospatial Information (SAOGI)

Lifecycle Maturity Assessment (LMA) Summary

Overall Maturity:

Optimized; Established

General Questions: 90%

Optimized; Established

Stage 1 - Define/Plan: 100%

Optimized; Established

Stage 2 - Inventory/Evaluate: 100%

Optimized; Established

Stage 3 - Obtain: 100%

Optimized; Established

Stage 4 - Access: 100%

Optimized; Established

Stage 5 - Maintain: 100%

Optimized; Established

Stage 6 - Use/Evaluate: 89%

Mature; Consistent

Stage 7 - Archive: 66%

Managed; Predictable

NGDA Dataset Maturity Definitions:

How To Calculate Maturity: [https://www.geoplatform.gov/sites/default/files/How to Calculate Maturity.pdf](https://www.geoplatform.gov/sites/default/files/How_to_Calculate_Maturity.pdf)

Maturity	Maturity Characteristics for All Lifecycle Stages
Optimized; Established Rank = 5	Dataset meets virtually all business needs of all users. The dataset is considered authoritative by owners and secondary users. It is curated across all stages of the approved lifecycle. Future needs are defined on a regular basis and resources for addressing both current and future business requirements are available.
Mature; Consistent Rank = 4	Dataset meets all the business needs of the primary owner and most of the secondary users. The dataset is curated and used as authoritative by the primary owner. Dataset is used widely by secondary users actively engaged in sustaining the dataset. Future needs are identified and steps are planned to address these. All stages are supported and reviewed on a recurring basis. The dataset is well managed in relation to the approved lifecycle.
Managed; Predictable Rank = 3	Dataset meets a significant number of the business needs of the primary owner and is widely used as an authoritative resource by secondary users. Benchmark activities are occurring in at least four of the approved lifecycle stages. Management practices in relation to the approved lifecycle is moderate but consistent. Dataset is integrating changing business requirements in lifecycle stages impacting overall maturity.
Transition; Transformation Rank = 2	Dataset meets business needs of the primary owner and has moderate use by secondary users. Benchmark activities are occurring in at least three stages. Efforts to integrate funding, include partners, and obtain data are not supported in a sustained manner. Management practices in relation to the stages of the approved lifecycle is limited.
Planned; Initial Development Rank = 1	Dataset limited in meeting business needs of the primary owner. Benchmark activities in the approved lifecycle are just starting to consider secondary uses, partnerships are forming to support additional dataset uses. Dataset development is in a very early stage. Minimal or limited management against the benchmarks in the approved lifecycle.
No Activity Rank = no activity	Dataset meets project or local business needs of the primary owner, secondary or additional uses or users were not considered, not recognized as an authoritative data or is part of a similar dataset. Not managed to any of the benchmarks in the approved lifecycle.

General Questions for All Stages

1) Is there a recurring process to obtain funding for all lifecycle stages of this dataset?

Answer: Funding support is part of agency budget on a recurring basis, funding is consistent and tied to business processes, and supports all lifecycle stages.

Justification Comment:

Attachment(s): 0

Funding for the Cropland Data Layer or CDL is provided by the Agency on a recurring basis. The CDL program is considered operational, producing in-season acreage estimates supporting the NASS Ag Statistics Board (ASB) estimation process. Funding/staffing is provided to continue program operations throughout all system lifecycle phases.

2) Is there a process in place to ensure that open government and transparency guidelines are followed in all lifecycle stages for this dataset?

Answer: Process is published as appropriate with respect to sensitivity requirements, process is transparent, published appropriately.

Justification Comment:

Attachment(s): 0

The CDL product is used to create internal (CUI) acreage estimates across the county in an operational environment. Acreage estimates from the CDL product are submitted to the NASS Ag

Statistics Board for setting official acreage estimates.

Upon completion of growing season, the CDL product is published in the public domain upon release of official county estimates. The CDL process uses CUI inputs or survey and administrative data and that data remains CUI and is never released.

The CDL is derived using published methods and processes and is considered open data, discoverable and queryable to the public upon publication.

3) Are there processes and tools in place so that staff are sufficiently knowledgeable to ensure a continuity of the dataset for all stages of the lifecycle, especially during staffing transitions?

Answer: Processes and tools to ensure dataset continuity are defined and beginning to be implemented.

Justification Comment:

Attachment(s): 0

The CDL uses standardized methods and processes for deriving the public product. The training manual provides instruction for analysts to follow standardized production work-flows. All CDL analysts are trained using these methods and they correspond/collaborate with one another on a daily and/or frequent basis throughout the production process. The analysts are provided the latest industry training opportunities and strive for continuous product improvement. Once finalized the CDL product is released on the CropScape web portal for public dissemination.

The CDL product and CropScape portals have not been assessed for records management requirements and long term value. All other terms of program meet standard.

STAGE 1 - Define/Plan

4) Are user and business requirements defined and formalized?

Answer: A recurring process is in place, including defining new partner and stakeholder business needs as they arise, and is fully implemented.

Justification Comment:

Attachment(s): 0

CDL production is operational and requested and funded by the Agency as an independent assessment of crop acreage throughout the growing season. There are cases when the Agency has requested special CDL tabulations in cases of extreme weather events as well. The Ag Statistics Board's needs and requirements are met by providing acreage estimates for stated commodities on a monthly basis throughout the growing season.

The CDL is published annually on a national scale with the intent of publishing the data during the first week of February coincident with the release of official county estimates. The CropScape portal is used as the official dissemination method of publication. Serving as redundant publication portals NASS releases the CDL product via the NASS website for ftp as well as the NRCS Geospatial Data Gateway. The entire inventory is published on CropScape and the Data Gateway.

5) How are partners/stakeholders involved in the requirements collection process?

Answer: A recurring process is in place, including defining new partner and stakeholder business needs as they arise, and is fully implemented.

Justification Comment:

Attachment(s): 0

The ASB is the major stakeholder of CDL estimates. The other stakeholders are the NASS Regional Field Offices stationed around the county. Since the CDL provides estimates to these stakeholders, methods are now operational to deliver these estimates using secure protocols. CDL program meets internal stakeholder needs with accurate and timely acreage estimates and serves the public with timely annual national public domain dataset updates characterizing the agricultural landscape.

6) Is there a quality assurance process for the dataset?

Answer: Quality assurance published as appropriate with respect sensitivity requirements.

Justification Comment:

Attachment(s): 0

The CDL product is a supervised land cover classification. The methods consume 70 percent of the available ground reference data for classification purposes, while the remaining 30 are used for accuracy assessment (AA). The AA is published in CDL metadata. The AA can also be used for correctional purposes to identify deficiencies in the land cover classification. A visual inspection is provided by the analysts throughout the production process and as the CDL products are prepared for publication an analyst compares prior CDL products to ensure mapping standards/compliance.

7) Is there a process to evaluate the sensitivity, privacy, and confidentiality of this dataset?

Answer: Sensitivity, privacy, and confidentiality evaluations fully implemented, reviewed and updated on a recurring basis.

Justification Comment:

Attachment(s): 0

The ground reference data comes from Agency survey and administrative sources. This data is CUI and not releasable. We sample from the reference data to build the land cover classification or CDL product. The CDL is released at 30 meters resolution and does not contain PII or sensitive data.

The ground reference data is published on our website at

http://www.nass.usda.gov/Research_and_Science/Cropland/metadata/meta.php The AA assessment uses standardized remote sensing metrics.

8) Are defined data standards used in collecting, processing, and/or rendering the data?

Answer: Standards fully implemented documented and published as appropriate.

Justification Comment:

Attachment(s): 0

The CDL product has been cleared by NASS for public release. There are no CUI data in-bedded within the product. This is a land cover classification, so there is inherent error, albeit quantified and published in this product.

Additionally the CDL product is released after the growing season when county estimates have been released and the financial markets have already determined the outcome.

STAGE 2 - Inventory/Evaluate

9) Is there a process for determining if data necessary to meet requirements already exist from other sources (either within or outside the agency) before collecting or acquiring new data?

Answer: Process for determining appropriate data is being reused fully implemented, reviewed, and updated on a regular basis.

Justification Comment:

Attachment(s): 0

NASS partners with USDA FAS (Foreign Ag Service) and FSA (Farm Service) on commercial satellite imagery and administrative data respectively. NASS has been a member of FAS' Satellite Image Archive (SIA) since 1997 and has been leveraging FSA's administrative data since the early 2000's. These are one of a kind data products/services with FSA's data being CUI and FAS's being member only to the SIA.

NASS CDL satellite program requirements are being met with satellite collections from the US government's Landsat program, but to provide an annual operational CDL product additional satellite collections provided by FAS are necessary to provide high temporal/frequent satellite coverage.

There are other datasets utilized for CDL production, namely the US Geological Survey's NLCD program. The following NLCD products are used for each CDL epoch, the National Land Cover

Database, imperviousness layer, forest canopy layer, and national elevation dataset. The latest NLCD epochs are used for deriving the CDL products.

STAGE 3 - Obtain

10) Is there a process for obtaining data in relation to this dataset?

Answer: Process is fully implemented, reviewed and updated on a regular basis.

Justification Comment:

Attachment(s): 0

The operational CDL program procures SIA services from FAS annually. This is a fixed cost derived from a long standing NASS/FAS MOU. NASS funds this agreement annually. NASS downloads the SIA satellite as well as the Landsat data streams. These data streams are downloaded and processed on a daily basis. As newer satellite systems are launched, these data streams are tested for needs/compliance in our work processes.

The FSA administrative data is downloaded on a monthly basis in preparation for build monthly in-season CDL products and estimates for the ASB. NASS has a liaison officer that handles communication between agencies regarding product acquisition/download and any potential formatting changes.

11) Is the metadata in a FGDC endorsed geospatial metadata standard?

Answer: Metadata is available in a format endorsed by the FGDC, it fully describes the dataset and provides all the information required to make the dataset discoverable, accessible, and usable.

Justification Comment:

Attachment(s): 0

As mentioned earlier or CDL product specific statewide metadata is published at http://www.nass.usda.gov/Research_and_Science/Cropland/metadata/meta.php The CDL and CropScape portal are published, tagged, and attributed in data.gov and the geoplatform websites. The NGDA is discoverable, accessible and understandable.

12) How complete is the geographic coverage as defined in the requirements for the dataset?

Part 1 Answer: Business requirements for cyclic updates identified and a process is in place.

Part 2 Answer: Dataset has presently attained the greatest geographic coverage as defined by the current requirements or roughly 100%.

Justification Comment:

Attachment(s): 0

The CDL is published in annual epochs upon completion of the growing season. Data is not published before, nor in-between and there are no intermediate products. The dataset is published at 30 meters contiguous US. The CDL has been published nationally/contiguous US since 2008.

STAGE 4 - Access

13) Do you have a process for providing users access to the data in an open digital machine readable format?

Answer: User access process is fully implemented, data is available, process is reviewed and updated on a recurring basis.

Justification Comment:

Attachment(s): 0

The CDL data is rasterized imagery. If the user has an image processor, imagery reader or GIS they can easily read/ingest the data. The CropScape portal facilitates user interaction with a visual interface allowing for user interaction, queries and .wms type services. CropScape is published using OGC standards.

STAGE 5 - Maintain

14) Is there a maintenance process for updating and storing the dataset?

Answer: Dataset maintenance process is fully implemented and processes are reviewed and periodically updated.

Justification Comment:

Attachment(s): 0

The CDL product is stored on numerous online portals. 1) CropScape @ <http://nassgeodata.gmu.edu/CropScape/>, 2) National FTP download @ http://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php, and 3) NRCS Data Gateway @ <https://gdg.sc.egov.usda.gov/>.

Original CDL source data/products and content are stored and backed up internally on our secured Agency IT infrastructure.

15) Is there an error correction process as part of dataset maintenance?

Answer: Error correction process includes user notification, process reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

We have issued correction notices on our Agency website. Users are notified in that manner or via email.

STAGE 6 - Use/Evaluate

16) Is there a process to determine if the dataset meets user needs?

Answer: Process is complete and being implemented on ad hoc basis.

Justification Comment:

Attachment(s): 0

Internally we meet with stakeholders to determine if our acreage estimates are accurate and of value to the Agency.

We do not have a great mechanism to receive external feedback on the CDL product. We regularly attend professional conferences and participate in outreach activities as budget allows, but feedback is on a restricted basis. If something is identified as incorrect, or just plain wrong, or a web service is not running, we hear about it rather quickly. If the product services the public well, then we generally don't hear about that.

17) Is there a process to provide users information on how to access and properly use the dataset?

Answer: Process is fully implemented supporting access and proper use, process is reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

We have online documentation on CropScape for users and developers as well as an video tutorial. We have a FAQ page published on the NASS webpage. The FAQ page generally describes user experiences with the data and recommended solutions to commonly asked questions. There is an announcement section on our main CDL page about upcoming releases or new papers/publications/presentations as well.

18) Are the business processes and management practices assessed to meet changing technology?

Answer: Assessment process is fully implemented for taking advantage of changing technology, process is reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

The methods, technology, and infrastructure are robust, transparent, and repeatable. CDL analysts regularly attend trade shows and industry conferences as well as correspond with other .gov entities who work on similar type products. We perform standardized testing as time/budget allows to determine if enhancements or upgrades are necessary and warranted.

STAGE 7 - Archive

19) Is there an archiving process for the dataset?

Answer: Archival and/or processes are in early implementation.

Justification Comment:

Attachment(s): 0

We store our geospatial data on NASS IT infrastructure as well as archive the data on blu-ray and large capacity usb drives. We do not have a robust implementation of long term handling and storage of geospatial materials such as NARA.