

NGDA Dataset Report

Official NGDA Title: National Land Cover Database (NLCD) Land Cover Collection

Metadata Record Title: National Land Cover Database (NLCD) Land Cover Collection

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: Collin Homer

Agency: USGS

Email: homer@usgs.gov

Metadata:

Registration Status: Complete

Registered on 4/10/2015

GeoPlatform Link*: <https://www.geoplatform.gov/node/243/ec7f7cc1-fb25-4f69-9804-fe64c411c69a>

Data.gov Metadata Link*: <http://catalog.data.gov/harvest/object/735fee49-68b4-4f0c-8a0b-a02d1dee9727/html>

*If the metadata has been updated and reharvested after publication of this report, the link may no longer be valid. The dataset may be searched for manually in Data.gov or GeoPlatform.gov.

NGDA Lifecycle Maturity Assessment (LMA) Report

Time Frame:

Baseline assessment covers NLCD 1992, 2001, 2006, and 2011.

LMA Submission:

Status: Complete

Date: 9/25/2015

Extension Requested: No

LMA Reviewer(s):

Supervisor: Collin Homer <homer@usgs.gov>

Theme Lead: Jonathan Smith <jhsmith@usgs.gov>

Executive Champion: Did not review

SAOGI*: Did not review

Other: Did not review

LMA Verifier:

Name: Collin Homer

Email: homer@usgs.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: 100%

Optimized; Established

Stage 4 - Access: 100%

Optimized; Established

Stage 1 - Define/Plan: 100%

Optimized; Established

Stage 5 - Maintain: 100%

Optimized; Established

Stage 2 - Inventory/Evaluate: 100%

Optimized; Established

Stage 6 - Use/Evaluate: 100%

Optimized; Established

Stage 3 - Obtain: 100%

Optimized; Established

Stage 7 - Archive: 100%

Optimized; Established

NGDA Dataset Maturity Definitions:

How To Calculate Maturity: 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): 1

Work on the National Land Cover Database (NLCD) is funded through the USGS Land Change Science Program, along with reimbursable funding and in kind contributions from the MRLC (Multi-Resolution Land Characterization) Consortium members. The USGS has funded this work for more than 20 years. The MRLC Consortium has grown from the original 4 member agencies (USGS, EPA, NOAA, and the Forest Service) to, currently, the 10 member agencies listed below:

U.S. Geological Survey (<http://www.usgs.gov>)

Environmental Protection Agency (<http://www.epa.gov>; EPA also maintains a Web page on MRLC at <http://www2.epa.gov/eco-research/multiresolution-land-characteristics-mrlc-consortium>)

National Oceanic and Atmospheric Administration (in particular the NOAA Office for Coastal Management at <http://www.coasat.noaa.gov>)

U.S. Forest Service (<http://www.fs.fed.us/>)

Bureau of Land Management (<http://www.blm.gov>)

National Agricultural Statistics Service (<http://www.nass.usda.gov/>)

National Park Service (<http://www.nps.gov>)

National Aeronautics and Space Administration (<http://www.nasa.gov/>)

U.S. Fish and Wildlife Service (<http://www.fws.gov/>)

U.S. Army Corps of Engineers (<http://www.usace.army.mil/>)

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

An assessment of NLCD practices against the 8 principles of open government data (opengovdata.org) shows that NLCD follows open government and transparency guidelines:

8 Principles of Open Government Data

1. Complete. NLCD datasets are made available to the public in their entirety. Data downloads at <http://www.mrlc.gov> contain entire datasets that are split only by 4 major geographies, the conterminous United States, Alaska, Hawaii, and Puerto Rico and U.S. Virgin Islands. This split is done so that each of the 4 pieces is as compact as possible, without empty fill, and can be downloaded as quickly and efficiently as possible.

2. Primary. NLCD datasets are derived from 30 meter resolution Landsat data, and are delivered to the public at the same 30 meter resolution, with no aggregation or subjective generalization.
3. Timely. NLCD datasets are developed and made available in a timely fashion. Datasets are generally compiled, completed, error checked, and published within 5 years of the nominal date of the source imagery.
4. Accessible. NLCD datasets are available for download from <http://www.mrlc.gov>, free of charge, in industry-standard file formats.
5. Machine processable. NLCD dataset bundles include raster data, attribute tables in standard formats, color lookup tables, and associated metadata and other files to enable the public to perform geographic analysis of these datasets.
6. Non discriminatory. NLCD datasets are available for download to anyone.
7. Non proprietary. NLCD datasets are made available in 2 raster formats. One format, IMG, was developed by ERDAS Imagine but is not exclusively controlled by that company. The other format, GeoTIFF, is an open source standard.
8. License free. NLCD datasets are not subject to any copyright, patent, trademark, or trade secret regulation, and are in the public domain.

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 in place and implemented for all lifecycle stages.

Justification Comment:

Attachment(s): 0

The NLCD project (<http://www.mrlc.gov>) rigorously documents the strategic model for NLCD products and ensures, through quarterly and annual meetings, that MRLC Consortium members thoroughly review, comment, and propose changes to the overall strategy. The operations manual that guides the NLCD production flow is a detailed guide to the datasets, models, functions and all other processing considerations required to produce the NLCD. The operations manual is rigorously maintained and updated by the project's scientific and technical staff, so that NLCD production is thoroughly documented to maintain continuity in case of staffing changes, and to help new staff learn the process quickly. The operations manual is especially critical to ensuring a consistent processing approach is used by production staff spread among multiple locations.

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

The quarterly and annual meetings of the MRLC Consortium, along with frequent communication in between, provide a venue for stakeholders to bring business and technical needs to the strategic planning process for NLCD products. MRLC Consortium members are both partners in the production of NLCD and users of the resulting datasets. The periodic meetings are an important way for stakeholders to define new requirements and needs for the planning process.

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 quarterly and annual meetings of the MRLC Consortium (<http://www.mrlc.gov>), along with frequent communication in between, provide a venue for stakeholders to bring business and technical needs into the strategic planning process for NLCD products. The (currently) 10 member agencies of the MRLC Consortium (Web sites are listed under question 1) are both partners in the production of NLCD and users of the resulting datasets. As a result, the periodic meetings are an important way for these stakeholders to bring new requirements and needs to the planning process. Requirements and future work on NLCD development are discussed and a future course agreed to as the primary activity and goal of MRLC meetings. In addition, the potential scope of future NLCD work is generally greater than the resources available to perform it. The MRLC Consortium meetings are where business requirements – funding and in-kind resources available, acceptable timelines for dataset delivery – are assessed against desired characteristics and capabilities of upcoming NLCD work, and adjustments are made to balance the two.

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

For the major updates in 2001 and 2011, the NLCD project has performed an extensive accuracy assessment, based on large samples of random points. Results of the 2001 assessment for the 48 conterminous states were published in: Wickham, J. D., S. V. Stehman, J.A. Fry, J.H. Smith, and C.G. Homer, 2010, Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. Remote Sensing of Environment 114 (6):1286 to 1296 (<http://www.sciencedirect.com/science/article/pii/S0034425710000490>). An accuracy assessment for NLCD 2011 is underway, and will be published when completed.

In addition, the NLCD project employs a Quality Supervisor (QS), who is responsible for the overall accuracy and validity of the NLCD datasets. The QS ensures that the mapping teams employ consistent approaches and methodology in their land cover mapping, and also ensures that the data received from the mapping teams meshes into one consistent, final land cover mosaic.

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 NLCD does not contain sensitive information of any kind. Datasets produced by the NLCD project are created from freely available public data sources, including Landsat data, the National Elevation Dataset, and other ancillary data, and NLCD final products are freely available to the public.

NLCD products are published at 30 meter resolution. Each 30 meter pixel represents an area of about .22 acres on the ground, suitable for mapping land cover at a medium level of detail, but not for seeing fine detail on the ground, or for locating any but the most massive manmade features. Instead, NLCD is designed to support analysis of broad, landscape patterns of land cover, imperviousness, and tree canopy, and their change over time.

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

Because the NLCD project employs an approach of constant innovation in the development of each

successive update of the NLCD, it's important to use standard, repeatable data sources (primarily Landsat images) that have well documented characteristics and defined spectral and geometric properties, so that NLCD data derived from them are consistent and verifiable. Processing flows are based on commercial image processing software, such as ERDAS Imagine, eCognition, See5 (and others), and standard scripting languages, such as Python and PERL. Rigorous, detailed documentation is maintained of the exact steps and parameters applied to produce the NLCD.

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

Because the NLCD project is the only multiagency Federal Government effort to map land cover on a regular, consistent basis, there really are no alternative data sources that could be substituted for NLCD. All source data, such as Landsat, are obtained at no cost to the government. Perhaps more fundamentally, the whole point of generating land cover updates is not to reuse old datasets, but to generate new ones that are derived from the most recent satellite imagery. Working through the MRLC, the NLCD project coordinates with other Federal agencies to share costs and incorporate land cover mapping being done by other, non USGS government agencies, especially the GAP Analysis Program, LandFire, and NOAA's CCAP Program, all of which create land cover for more specific purposes, and apply these more specialized datasets to create NLCD products with attributes and characteristics of greatest use to the widest possible user community.

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 NLCD is distributed for user access in 2 primary ways: 1. The Website <http://www.mrlc.gov/> is the primary distribution point for NLCD datasets. This Website is maintained at USGS EROS and is a portal for all distributed NLCD primary datasets and associated supporting datasets (for example, change pixels). All datasets are available for download in standard raster formats that preserve geospatial information. In addition, mrlc.gov provides links to publications, including journal articles and USGS publications, that explain the content of NLCD. All downloads of data and other materials are freely available to the public. 2. NLCD datasets are also distributed in 1-degree-by-1-degree tiles by The National Map, <http://nationalmap.gov/> and viewable in The National Map viewer.

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

All NLCD datasets are distributed with accompanying metadata in CSDGM format. Metadata describe in detail the source datasets and processing steps used to create NLCD products. All metadata records include NGDA keywords to aid discoverability. All metadata records have been loaded into geoplatform.gov and data.gov.

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 NLCD was originally developed in 1992, then updated in 2001, 2006, and 2011. The next 5 year update, in 2016, will rebaseline the NLCD, under a development approach that has been vetted and approved by the member agencies of the MRLC. All 5 epochs of NLCD cover the conterminous United States. The 2001 and 2011 updates additionally cover Alaska, and 2001 additionally covers Hawaii and PRVI. Percent Developed Imperviousness was developed in 2001, 2006, and 2011. Percent Tree Canopy was developed in 2001 and 2011. Resolution of all datasets is 30 meters. Geographic coverage is considered complete, as judged by MRLC consortium members.

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 NLCD is distributed for user access in 2 primary ways, the Website <http://www.mrlc.gov/> (the primary distribution point for NLCD datasets) and as 1-degree-by-1-degree tiles by The National Map, <http://nationalmap.gov/>. Both delivery methods use one of two open file formats, either GeoTIFF (<http://www.digitalpreservation.gov/formats/fdd/fdd000279.shtml>) or ERDAS Imagine file format (IMG) with, for images greater than 2 GB in size, an accompanying Large Raster Spill File (IGE) (<http://www.digitalpreservation.gov/formats/fdd/fdd000420.shtml>). USGS makes use of both GeoTIFF and IMG/IGE for data distribution.

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

With 4 epochs of data already released 1992 (<http://www.mrlc.gov/nlcd1992.php>), 2001 (<http://www.mrlc.gov/nlcd2001.php>), 2006 (<http://www.mrlc.gov/nlcd2006.php>), and 2011 (<http://www.mrlc.gov/nlcd2011.php>) and a 5th (2016) currently in development, the NLCD project considers dataset maintenance and update to be a core mission. Changing requirements are identified in the quarterly and annual MRLC meetings and incorporated into NLCD products in the following update cycle. The MRLC consortium has been an effective vehicle for capturing the changing needs of its member agencies.

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

Any error found in NLCD products is examined by the project's quality assurance supervisor, and evaluated as to whether it is a major, systematic error that warrants a rerelease of the full dataset, or if minor, corrected during the next regular dataset update. Corrected datasets, with updated metadata that details the repair(s) made, are posted to <http://www.mrlc.gov/> along with an explanatory notice on the Web site.

STAGE 6 - Use/Evaluate

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

Answer: Process is fully implemented and repeated on a recurring basis.

Justification Comment:

Attachment(s): 0

The NLCD project receives feedback from the user community in 3 ways: 1. The MRLC consortium, the producer of NLCD, is itself also a consumer of the datasets produced. A part of each quarterly and annual meeting of the consortium is devoted to discussions of NLCD product quality and utility, and how changing requirements should be reflected in NLCD planning strategy. 2. The NLCD receives questions and inquiries directly from users through an established process at the USGS EROS Center's Customer Services section. Customer Service (CS) staff log calls from a wide range of users on a wide range of topics. CS staff answer basic questions about dataset properties, file formats, and access, but pass along to NLCD staff questions of a more technical or scientific nature. 3. NLCD staff have published in peer reviewed journals and presented at conferences as a way to receive feedback from the science community on both the theoretical underpinnings of NLCD's development approach, and on the practical experiences of users of NLCD data.

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

The NLCD project has a mature, 2 prong approach to informing users on how to access and use NLCD products: 1. The MRLC consortium's flagship Web site, <http://www.mrlc.gov/>, is the primary data portal to find, access, and download all NLCD datasets. In addition, [mrlc.gov](http://www.mrlc.gov/) has an extensive FAQ section that answers typical user questions, sometimes in great technical detail, including step by step instructions to solve particular problems. In particular, one section of the FAQ deals exclusively with data access and use issues (http://www.mrlc.gov/faq_dau.php). 2. The USGS Land Cover Institute's Web site, <http://landcover.usgs.gov/>, is an information source for USGS land cover in general, both the current NLCD product and other efforts, such as global land cover and historic USGS land cover. As an information site, landcover.usgs.gov has extensive news and descriptions of datasets, its own FAQ, and a references list, as well as links to the [mrlc.gov](http://www.mrlc.gov/) data portal. Between these 2 Web sites, users of NLCD products have access to comprehensive information about dataset properties, formats, and use.

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

NLCD project staff regularly review new hardware and software, and have incorporated both into the NLCD production process. For example, the 1992 NLCD process was based primarily on clustering software followed by manual labeling. The next update, NLCD 2001, was primarily a product of decision tree software that had been evaluated and adopted by the project team. The 2016 effort will employ new software to generate land cover change trajectories. In addition, the NLCD project is actively engaged in centerwide initiatives at USGS EROS, such as the LCMAP project, and will take advantage of Landsat Science Team developments that make geometrically and radiometrically corrected Landsat data more readily available to users, potentially saving several steps in the NLCD processing flow.

STAGE 7 - Archive

19) Is there an archiving process for the dataset?

Answer: Archival and disposition processes are fully implemented.

Justification Comment:

Attachment(s): 0

NLCD relies on the Long Term Archive (LTA) project at USGS EROS to perform critical services for both interim and longterm archiving of NLCD products. LTA performs these services for many projects at USGS EROS. By relying on LTA processes and technology, the NLCD project is assured of following best practices, making use of best technology, and assuring that longterm archive with the National Archives and Records Administration (NARA) is done by established practices.