

NGDA Dataset Report

Official NGDA Title: U.S. Geological Survey Gap Analysis Program Species Distribution Models

Metadata Record Title: U.S. Geological Survey Gap Analysis Program Species Distribution Models

A-16 NGDA Theme: Biota

Executive NGDA Theme Champion(s):

Name: Jerry Johnston

Agency: DOI

Email: jerry_johnston@ios.doi.gov

Theme Lead(s):

Name: Marcia McNiff

Agency: USGS

Email: mmcniff@usgs.gov

Dataset Manager(s):

Name: Julie Prior-Magee

Agency: USGS

Email: jpmagee@usgs.gov

Metadata:

Registration Status: Complete

Registered on 12/10/2014

GeoPlatform Link*: <https://www.geoplatform.gov/node/243/ad3ad373-1a0c-4515-82e7-d10e2b05d326>

Data.gov Metadata Link*: <http://catalog.data.gov/harvest/object/af53cb32-c02b-49fe-8d20-6ffc7af5f46f/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 responses include dataset activities from 2008 to 2015. GAP began its focus on national modeling in 2008.

LMA Submission:

Status: Complete

Date: 10/22/2015

Extension Requested: Yes

LMA Reviewer(s):

Supervisor: Sky Bristol

Theme Lead: Did not review

Executive Champion: Did not review

SAOGI*: Did not review

Other: Did not review

LMA Verifier:

Name: Sky Bristol

Email: sbristol@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:

Managed; Predictable

General Questions: 72%

Mature; Consistent

Stage 4 - Access: 75%

Mature; Consistent

Stage 1 - Define/Plan: 78%

Mature; Consistent

Stage 5 - Maintain: 28%

Planned; Initial Development

Stage 2 - Inventory/Evaluate: 100%

Optimized; Established

Stage 6 - Use/Evaluate: 77%

Mature; Consistent

Stage 3 - Obtain: 60%

Managed; Predictable

Stage 7 - Archive: 33%

Transition; Transformation

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 exists but is not adequate to meet known requirements, most lifecycle stages are supported.

Justification Comment:

Attachment(s): 0

Funding for the development of the Gap Analysis program species distribution models comes through the Core Science Systems Mission Area of USGS. This is one of four key datasets necessary for the gap analysis. Species models for all taxa being analyzed are nearly complete, but we are working on improving the maintenance, use (access), and archiving aspects of the lifecycle. Currently the Core Science Analytics, Synthesis and Libraries (CSAS&L) program (under which the Gap Analysis Program species modeling activities is located) is undertaking several program planning activities that will help CSAS&L more adequately meet the requirements of the Maintain, Use/Evaluate, and Archive stages. These planning activities will address known needs for a proper data repository solution, as well as the need for Data Management Plans that will address several of the required activities in the Maintain through Archive lifecycle stages.

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 established, significant portions of the documentation is complete.

Justification Comment:

Attachment(s): 0

The species distribution models are being shared openly with the public through an online viewer, data downloads, and webservices (<http://gapanalysis.usgs.gov/species/viewer/>). We are in the process of changing the archiving of the completed species models which will require some modification of the access point for the data. The methods used to create the models are well documented in species reports that can be downloaded for any completed model. That report describes the variables used in the creation of the spatial model. We are taking steps to publish the full wildlife habitat relationship database. The species models metadata describes the general modeling process as well <http://gapanalysis.usgs.gov/wp-content/uploads/2013/09/GAPSpeciesDistributionModelmetadata.pdf>.

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 species distribution modeling effort has been a long-term effort with sufficient institutional knowledge and structure to successfully weather several staffing transitions. The central database design, ancillary data, and common spatial framework across models provides for continuity. We are now in the process of developing a more stable archive for the completed species models, which will further enhance continuity.

STAGE 1 - Define/Plan

4) Are user and business requirements defined and formalized?

Answer: A recurring process exists for gathering partners/ stakeholders requirements is in place and is in the beginning stages of implementation.

Justification Comment:

Attachment(s): 0

The Gap Analysis Program is a mature program with a highly focused stakeholder community. The requirements were defined by the conservation and academic community in the early stages with individual state and regional efforts. A 2008 program review recommended getting the national

datasets and analysis should be the priority. Since that time we have focused on that task and species distribution models for the native terrestrial vertebrate species are nearly complete. Upon completion of the national analyses, additional analyses and required enhancements to species distribution models will be defined by stakeholders.

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

Answer: A recurring process exists for gathering partners/ stakeholders requirements is in place and is in the beginning stages of implementation.

Justification Comment:

Attachment(s): 0

The Gap Analysis Program is a mature program with a highly focused stakeholder community. The requirements were defined by the conservation and academic community in the early stages with individual state and regional efforts. Those stakeholders defined the requirements through innovation and research and development. The GAP state projects showed the limits of the technology and a series of annual meetings were held to showcase the methods and approaches being tested and successfully applied. Successes from the early projects were used to identify the most meaningful approaches for the regional and national effort. Upon completion of the national analyses, additional analyses and required enhancements to species ranges will be defined by stakeholders.

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

Answer: Process established, significant portions of the documentation is complete.

Justification Comment:

Attachment(s): 0

Each species distribution model is reviewed by two wildlife biologists as a part of the modeling process. The models are based on literature review and descriptions of habitat affinities by species. In the metadata, data constraints and appropriate uses of the data are summarized. The species distribution data product will be going through the USGS data review and release process as described in the USGS Instructional Memorandum IM OSQI 2015-03 <http://www.usgs.gov/usgs-manual/im/IM-OSQI-2015-03.html> .

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 species distribution modeling does not involve any sensitive data with respect to personal information.

There are individual animal species in the database for which over-harvest in the wild is an issue. In those cases the species distribution models are not published online.

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

Answer: Standards being implemented.

Justification Comment:

Attachment(s): 0

The methods for the species distribution modeling and the standards are described in the Gap Analysis Handbook from 2007 (will send via email). While most of the process is the same, the move to a national extent means that handbook needs to be revised to reflect changes made to accommodate the new extent and changing technology. The species lists used are based on the taxonomic authorities - Amphibians and Reptiles - Crother 2008 (https://ssarherps.org/wp-content/uploads/2014/07/HC_37_6thEd.pdf), Mammals - Wilson and Reeder 2005 (<http://vertebrates.si.edu/msw/mswCFApp/msw/index.cfm>) , Birds - American Ornithological Union's 2008 checklist (<http://www.aoucospubs.org/doi/pdf/10.1525/auk.2008.9708>) and linked to the Intergrated Taxonomic Information System codes.

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

While there are individual species models available through other programs, there is no other comprehensive source of national species distribution models based on a thorough literature review. The comprehensiveness and consistent modeling approach are necessary to support a meaningful national gap analysis.

STAGE 3 - Obtain

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

Answer: Process is being implemented.

Justification Comment:

Attachment(s): 0

Specific data sets were identified that were needed in order to develop species distribution models. All of the input data used and the output models created in the development of the species distribution modeling are available for download from the National Gap Analysis program's website. <http://gapanalysis.usgs.gov/species/>. Currently there are performance issues with the species viewer that we are working to make the access to the data more direct, but all of the data are available for download.

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

Answer: Metadata is available in a format endorsed by the FGDC but does not fully describe the information needed to make the dataset discoverable, accessible, and usable.

Justification Comment:

Attachment(s): 0

The species distribution models are described in a common FGDC Metadata record. We are working to embed species specific metadata in each species model data download in the future. Metadata from Data.gov: <http://catalog.data.gov/dataset/u-s-geological-survey-gap-analysis-program-species-distribution-modelsaea06>

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

Part 1 Answer: Business requirement targets are being attained, cyclic updates being assessed.

Part 2 Answer: Dataset presently about 75% complete per current requirement.

Justification Comment:

Attachment(s): 0

Part 1 = Species distribution models for the full U.S. including Alaska, Hawaii, Puerto Rico are near completion. Updates to the species distribution modeling will be evaluated based on the availability updated data sources (e.g. land cover, new range information, high quality occurrence records to drive models) and stakeholders needs.

Part 2 = Species distribution models 86% (1485/1735) of the species distribution models for the nation have been completed. Native terrestrial birds (648) and reptiles (327) are complete. 50 of the 459 native terrestrial mammals and 200 of the 301 amphibians need to be modeled.

STAGE 4 - Access

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

Answer: Process is developed, documented, and beginning to be implemented.

Justification Comment:

Attachment(s): 0

The species distribution models are being provided as webservice (eg. http://gis1.usgs.gov/arcgis/rest/services/NAT_Species_Amphibians/aagtox/MapServer/). We are currently changing the process for generating and serving the webservice to make them more openly accessible through sciencebase.gov.

STAGE 5 - Maintain

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

Answer: Dataset maintenance process is under development.

Justification Comment:

Attachment(s): 0

We are in the process of changing the maintenance process for the species distribution models. There had been a working data workflow but changes in organizational structure are making updates to the process necessary.

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

Answer: Error correction process under development.

Justification Comment:

Attachment(s): 0

Because the species distribution models are still being developed the primary focus on errors is based on internal review. Some subsets of models have been formally reviewed by external partners – specifically the state biologists for the Western Governor’s Association. Those comments were compiled and evaluated by the modeling team. Where the proposed changes were well documented and consistent across the range of the species they were made. In the future the plan would be to compile comments from internal and external reviews and develop a reconciliation document explaining which changes have been made, and if no change was made the logic behind that choice.

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

The gap analysis program had a very specific focus and mandate, to assess the conservation status (representation) of terrestrial vertebrate species in the conservation network. In 2008 there was an independent programmatic review and the recommendations from that review were to complete the national species distribution models with the approach being applied and to test inductive modeling approaches. We have had two pilot projects to test the alternative methods, and found the current approach to be most appropriate for nationally consistent models.

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

On the National GAP Analysis program’s website there are multiple ways for the public to access the species models (through an interactive viewer, through download, or webservice). In the metadata there is a description about limitations of the data and appropriate uses of the data. We are working on changing the access point and soon the models will be hosted in ScienceBase.

Current access options:

<http://gapanalysis.usgs.gov/species/data/>

[http://gis1.usgs.gov/csas/gap/viewer/species/Map.aspx?__utma=105344476.1015654843.1440011370.1445266637.1445271389.12&__utmb=105344476.6.10.1445271389&__utmc=105344476&__utmz=105344476.1443799297.1.1.utmcsr=\(direct\)|utmccn=\(direct\)|utmcmd=\(none\)&__utmv=-&__utmj=30604479](http://gis1.usgs.gov/csas/gap/viewer/species/Map.aspx?__utma=105344476.1015654843.1440011370.1445266637.1445271389.12&__utmb=105344476.6.10.1445271389&__utmc=105344476&__utmz=105344476.1443799297.1.1.utmcsr=(direct)|utmccn=(direct)|utmcmd=(none)&__utmv=-&__utmj=30604479)

<http://gapanalysis.usgs.gov/species/data/web-map-services/>

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

Answer: Assessment process implementation started for taking advantage of changing technology.

Justification Comment:

Attachment(s): 0

The National Gap Analysis Program's website is well established and the data formats have been working well to date. When webservice became a standard way to share geospatial data those were created in addition to the more traditional gis formats. We are in the process of updating the data formats for the spatial models, getting the Wildlife Habitat Relationship Database formally published and made available on line, currently the information from the database is provided through species reports. We are also moving the models into the ScienceBase Catalog.

STAGE 7 - Archive

19) Is there an archiving process for the dataset?

Answer: Archival and/or disposition processes are in development.

Justification Comment:

Attachment(s): 0

The National Gap Analysis Program's website has been the central location for acquiring the data. Currently we are in the process of moving the data and creating webservices through the ScienceBase catalog. <https://www.sciencebase.gov/about/content/about-sciencebase>

The CSAS&L program is currently addressing data archival and disposition processes for all program datasets. The development of formal data management plans for CSAS&L products will address some of the needed actions under the Archive stage.