

AERIAL SEABIRD SURVEY DATA MANAGEMENT PLAN

Gulf of Mexico Marine Assessment Program for Protect Species (GoMMAPPS)

ABOUT THE PROJECT

The Northern Gulf of Mexico is a vital resource for many large marine vertebrate (LMV) species and information on their abundance, distribution, habitat use, and behavior in the Gulf of Mexico (GoM) is needed to monitor and mitigate for potential impacts of human activities, including offshore energy development. Understanding of cumulative impacts on protected species in the Gulf from both natural and anthropogenic forcing is required to inform NEPA documents and consultations and rulemaking related to Endangered Species Act (ESA), Marine Mammal Protection Act (MMPA), and Migratory Bird Treaty Act (MBTA), as well as other statutes.

The Gulf of Mexico Marine Assessment Program for Protect Species (GoMMAPPS) aims to improve upon the scope of information on marine species in the Gulf of Mexico and provide a baseline assessment of LMV species' abundance and distributions within the Gulf. Biologists from the US Fish and Wildlife Service (USFWS), US Geological Survey (USGS), and the National Oceanic and Atmospheric Administration (NOAA), in partnership with the Bureau of Ocean Energy Management (BOEM), have collected data on marine mammals, sea turtles, and seabirds throughout the Gulf of Mexico. The results of this study will provide important information for regulatory needs, as well as for effective management and conservation of protected species.

The seabird team objectives are to 1) characterize the distribution, abundance, and diversity of seabirds in the GoM and 2) estimate and interpret the effects of natural and anthropogenic factors on bird distribution and abundance in the GoM. To achieve these goals, the seabird team conducted aerial surveys in the nearshore environment and offshore vessel surveys on ships of opportunity. This Data Management Plan (DMP) describes the *aerial seabird survey data only*, although the vessel survey data is included within the aerial survey database for completeness. *Please note* that this DMP is considered a living document and as such is subject to change. Project personnel, (including roles/responsibilities and contact information), database tables, QA/QC methods/software, and products are subject to change, particularly if there are future aerial surveys conducted that are not included within the scope of this initial phase of the GoMMAPPS.

The aerial surveys are low-level strip transect surveys, which extend from the USA-Mexico border near Brownsville, TX to Key West, FL, including the Dry Tortugas. The surveys were conducted during winter (2018, 2019, 2020) and summer (2018) following a pilot-effort off the Louisiana coast in July 2017. Using the U.S. EPA Environmental Monitoring & Assessment Program's 40-square kilometer hexagon sampling grid and a generalized random tessellation stratified sampling technique, we drew a random sample of 180 hexagons to survey. We then selected a random flight direction for each hexagon to include two additional, adjacent hexagons, and ran three parallel transects across the three hexagons.

Two flight crews each surveyed half of the Gulf of Mexico, i.e., 90 hexagons per crew. The data recorded include the species, count, time, latitude, longitude, and any additional notes. Other fields include date, aircraft, observer, seat location, survey unit (hexagon and transect), flight direction, and survey conditions. More information on these data fields can be found in the *Data Elements* section or on Confluence ([Data Collection](#) & [Transcription](#)). The surveys included a double observer protocol with three observers collecting data in each aircraft: the pilot-biologist and two biologists who rotated their seat position daily. This survey design allows for assessment of flock detection probability, count estimation bias, and species misidentification from two, independent observers.

Further information on the seabird surveys can be found on the seabird survey team space on USGS Confluence (<https://my.usgs.gov/confluence/display/GoMMAPPS>). The site includes survey information, standard operating procedures (SOPs), an archive of datafiles, survey summaries, data & analysis products and presentations, reports, and meeting information and notes.

Many people have collaborated on this project; the seabird team consists of those listed below. See the *Responsibilities* section for specific roles related to aerial survey data management.

Randy Wilson – aerial survey PI
Station Leader, Migratory Bird Field Office
US Fish and Wildlife Service
6578 Dogwood View Parkway, Suite C
Jackson, MS 39213
601-321-1121
randy_wilson@fws.gov

mark_koneff@fws.gov

Allison Sussman – data management and GIS support
Biologist
USGS Patuxent Wildlife Research Center
12100 Beech Forest Road
Laurel, MD 20708
301-497-5867
asussman@usgs.gov

Jim Lyons – aerial survey design & data management
Research Ecologist
USGS Patuxent Wildlife Research Center
12100 Beech Forest Road
Laurel, MD 20708
301-497-5567
jelyons@usgs.gov

Mark Koneff – aerial survey coordination and conduct
Chief, Migratory Bird Surveys Branch
US Fish and Wildlife Service - DMBM
York Village Building 6 - Suite C
69 Grove Street Extension
Orono, Maine 04469
301-980-0125

Jeff Gleason – vessel survey PI
 US Fish and Wildlife Service
 Gulf of Mexico Migratory Bird Coordinator
 Gulf Restoration Office
 Lower Suwannee National Wildlife Refuge
 16450 NW 31 Place
 Chiefland, FL 32626
 352-493-0238 x231
 jeffrey_gleason@fws.gov

Elise Zipkin – aerial survey design and analysis
 Assistant Professor
 Department of Integrative Biology
 Michigan State University
 288 Farm Lane RM 219A
 East Lansing, MI 48824
 517-884-8039
 ezipkin@msu.edu

Emily Silverman – aerial survey design & data management
 Statistician
 US Fish and Wildlife Service
 Division of Migratory Bird Management
 11510 American Holly Drive
 Laurel, MD 20708
 301-497-5801
 Emily_Silverman@fws.gov

Kayla Davis – aerial survey analysis
 PhD student
 Department of Integrative Biology
 Michigan State University
 288 Farm Lane
 East Lansing, MI 48824
 davisk93@msu.edu

The USFWS pilot biologists were Steve Earsom, Mark Koneff, and Jim Wortham. The aerial survey field biologists included: Dean Demarest (USFWS), Barret Fortier (USFWS), Wade Harrell (USFWS), Pat Stinson (USFWS), Allison Sussman (USGS), Robert Wheat (USFWS), Randy Wilson (USFWS; also the aerial survey PI), and Nick Wirwa (USFWS). Kayla Davis (MSU) flew part of the final survey (Feb 2020) and assisted in running GoPro cameras to collect and analyze video survey data.

Jeff Gleason (USFWS) was the vessel survey PI, and the vessel research team consisted of Pat Jodice (USGS & Clemson), J. Chris Haney (Terra Mar), Yvan Stage (Clemson), Pam Michael (Clemson), and Kathryn Hixson (Clemson). More information on the seabird surveys team for GoMMAPPS, can be found on Confluence.

DATA MANAGEMENT RESPONSIBILITIES AND RESOURCES

The seabird surveys are coordinated by the US Fish and Wildlife Service; Randy Wilson (aerial) and Jeff Gleason (vessel) are the Principle Investigators (PIs) for the seabird team.

Required Resources

The active project requires the support of the entire aerial survey team, a Project Coordinator for project management, and a fulltime Database Specialist. Prior to the project's inception and throughout the initial data collection period, additional support was required for designing the survey and sampling framework. Following active data collection, a PhD student will be funded to complete aerial survey data analysis and modeling.

The Database Specialist required physical office space and a computer. USFWS and/or USGS server space and access is needed for the Database Specialist to store and actively manage the database. The seabird team also requires internal access to the database throughout the development and duration of the project. To create and maintain the database, the Database

Specialist needs Microsoft SQL Server, ArcGIS and/or QuantumGIS, Google Earth (free version), Microsoft Office, and R/RStudio. Software that is not required but is useful to the Database Specialist includes the Adobe Suite, an advanced text editor (e.g., Notepad++), a metadata creator/editor tool (e.g., USGS Metadata Wizard), and a tool to convert spatial files (i.e., .shp, .kml, .gdb) to .gpx files for the pilots to use while conducting aerial surveys.

Each aircraft must be equipped with GPS and data recording equipment & software for each observer: the Hodges RECORD program (J. I. Hodges, USFWS, unpublished software) is run on a Sony Vaio computer in the aircraft and records observations during flight. Each observer also requires equipment and software to transcribe their data: a USB flash drive to transfer data files and a laptop with the TRANSCRIBE program (J. I. Hodges, USFWS, unpublished software) which translates observations to ASCII files. Members of the seabird team require access to the Confluence space; access is requested through the myUSGS Service Desk and requires a USGS sponsor.

Aerial Survey Data Management Roles

The responsible parties' duties for aerial survey data management will include, but are not strictly limited to, the following roles: SOP creation & modification, data collection, QA/QC process for data collection, data checking & correction, data synthesis & data products, data storage, metadata creation, documentation, corrections & versioning, access & data requests, data archives & backups, funding & resources.

The Database Specialist (Allison Sussman) is responsible for aerial survey preparation (e.g., file preparation for pilots and flight crews), database creation and development, data checking, correction, & QA/QC, data storage, archives & backups, metadata and documentation, data products, and access & data requests.

The Project Coordinator (Randy Wilson) is responsible for aerial survey coordination & conduct, funding & resources, SOP creation & modification, data collection, project reports & products, and overall project management.

The USFWS Migratory Bird Program is responsible for designing and conducting the aerial surveys, maintaining & providing project and data support, and server and file space for the project and data collected. The Branch of Migratory Bird Surveys (Branch Chief: Mark Koneff) will provide aerial survey support, airplanes to conduct surveys, pilot biologists, and necessary training tools and equipment for observers. The Branch of Monitoring and Data Management will be responsible for housing the data/database, data storage and space, and data management.

DATA ELEMENT DESCRIPTIONS

1. Sampling Framework & Survey Design

The base sampling unit for the aerial surveys is the EPA 40km hexagon. Surveys are conducted along transects across three hexagons. See Confluence ([Survey Design](#)) for more information

on the sampling framework, design, and methodology. The full 40km hexagon grid and associated metadata was compiled and released by USGS for public consumption (White, D., 1992, EPA 40km Hexagons for Conterminous United States: U.S. Environmental Protection Agency, <https://doi.org/10.5066/P9C56AY1>).

2. Base Layers

For the purposes of the GoMMAPPS project, the EPA Hexagons were limited to the study area, the Gulf of Mexico. As the layer did not extend into the southern Gulf, the data team working with the marine mammal group extended the hexagons to cover the entire Gulf of Mexico. That layer can be found here: `gommapps_root_folder/gis/hexagons/Hexagons_ExtendedGoM`.

Other important base layers include (but may not be limited to):

BOEM Gulf of Mexico Planning Areas (polygons)

Description: The BOEM Outer Continental Shelf (OCS) region contains three planning areas used for administrative purposes under the oil and gas leasing program (<https://www.boem.gov/regions/gulf-mexico-ocs-region>).

Location: `gommapps_root_folder/gis/boem_layers/Gulf_Planning_Areas`

BOEM Platforms (points)

Description: BOEM maintains a database of all oil and gas platforms within US waters, including whether the platforms are active or inactive. In addition to being stored within the GoMMAPPS aerial survey database and archive, the data are also accessible through the BOEM Data Center (<https://www.data.boem.gov/>).

Location: `gommapps_root_folder/gis/boem_layers/Gulf_Platforms`

Full Set Hexagons (polygons)

Description: Hexagons in the Gulf of Mexico which met the 'surveyable' criterion and made up the set of selectable hexagons used in the GRTS sampling technique that would select the survey sample and oversample hexagons.

Location: `gommapps_root_folder/gis/base_layers/FullHexagonSetGoM`

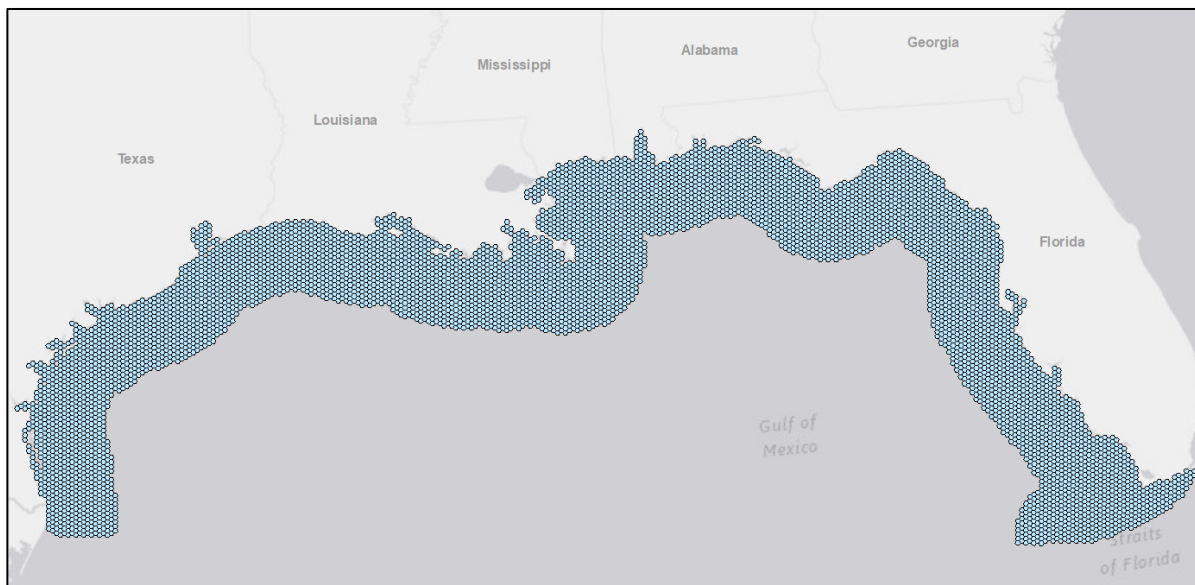


Figure 1. Full Set of hexagons for the Gulf of Mexico aerial surveys, based on the EPA 40km Hexagons for Conterminous United States, DOI: 10.5066/P9C56AY1.

Hexagons (polygons)

Description: Master hexagon polygon layer that includes all surveyed hexagon units, used for conducting aerial surveys and mapping data.

Location: gommapps_root_folder/gis/base_layers/Hexagons_Master

Transects (lines)

Description: Transect lines for the selected hexagons, the survey unit to be used for aerial surveys.

Location: gommapps_root_folder/gis/base_layers/Transects_Master

Transect End Points (points)

Description: End points for the surveyed transect lines, used by the pilots when conducting aerial surveys.

Location: gommapps_root_folder/gis/base_layers/EndPoints_Master

3. Survey Data

The GoMMAPPS aerial seabird surveys are conducted by the USFWS Migratory Bird Survey Branch (Branch Chief: Mark Koneff). For specific information on survey design and aircraft used see Confluence [Survey Design](#). All data is currently collected using the Hodges RECORD program (J. I. Hodges, USFWS, unpublished software) and the survey system installed in each aircraft. The Aerial Survey [Data Collection](#) and [Data Transcription](#) SOPs describe the process in more detail.

Two aircraft and flight crews are utilized for each survey, splitting the survey area (the Gulf of Mexico) into two sectors, east and west. Each crew consists of a pilot-biologist and two biologist observers. An unreconciled [double observer protocol](#) is used, whereby each observer records birds seen independently from each other and no attempt is made to reconcile between specific individual observations. See Confluence for more information about the double observer

protocol and observer rotation used during the survey to ensure equal and complete survey coverage.

Data Files

Each observer submits two flat files (.asc or .txt) per survey day (per seat) to the Database Specialist: an observation file and track file. If the observer(s) rotates seats within a single day, they have these files for each seat. Once received, the observation and track files will be checked for errors using R and ArcGIS. If any problems are found in the data files, the observer will be notified so that they may correct the problem for any remaining survey days (e.g., missing direction code or forgetting to change the seat during the transcription process).

Raw Data

Observations. The observations file is a text file (.asc or .txt) and should be named:

####_YEARMODA_**_obsinitials_birds.txt

Where #### = aircraft tail number (n7__); YEAR=four digit year, MO=two digit month, DA=two digit day; ** = lf (left front); rf (right front); rr (right rear); lr (left rear) seat position; and obsinitials = observers initials in lower case. For example: n723_20170710_rf_rrw_birds.txt is the observation file for Randy Wilson from the right front seat on 10 July 2017 on aircraft N723.

The elements of a transcribed seabird observation are: year, month, day, aircraft, seat position, observer, hexagon number, transect number, direction, file sequence, latitude, longitude, seconds after midnight, gps error, species code, count, and notes. For example: 2017,07,10,n723,rf,rrw,052,2,3,RRWRF001,44.767,-67.0213,34590.47,0.0,COLO,2. The species code field is also used to indicate the begin and end counts for each survey unit.

Once the observation data are checked, they are compiled into one master file for data QA/QC and processing. After the data have been QA/QC'd, the processed observations are imported into the main observations data table in the database. In addition, a flat CSV file is created for each survey and stored as an additional backup. Location: gommapps_root_folder/data/raw_data/

Tracks. The track data is also a text file (.asc or .txt) and should be named similarly to the observations file:

####_YEARMODA_**_obsinitials_hexagon_track.txt

Where #### = aircraft tail number (n7__); YEAR=four digit year, MO=two digit month, DA=two digit day; ** = lf (left front); rf (right front); rr (right rear); lr (left rear) seat position; and obsinitials = observers initials in lower case. For example: n723_20170710_rf_rrw_track.txt is the track file for Randy Wilson from the right front seat observer on 10 July 2017 on aircraft N723. The track file contains the flight path, collected as a string of XY coordinates, and is processed with the track files from the other observers in the same aircraft to create a single master track line for each survey. The processed tracks are then imported into the main transects data table in the database. Location: gommapps_root_folder/data/raw_data/

During every survey, each pilot maintains a record of [Survey Notes](#) where they record notes throughout the survey on items such as weather interruptions, data collection problems (e.g., computer failure), or other survey issues. The survey notes are documented within the database and used by the Database Specialist throughout the QA/QC process.

Data Processing

Once the Database Specialist receives and checks the raw data files, they are compiled together into one master data table/frame. A series of R scripts (location: `gommapps_root_folder/code/processing/`) are run to process the data, cleaning and correcting any errors found while documenting any changes. Once QA/QC'd, the cleaned data are imported into the database and prepared for analysis.

Processed Data*Observations*

Description: Observation data for aerial surveys

Location: `gommapps_root_folder/data/clean_data/`

Tracks

Description: Track data for aerial surveys

Location: `gommapps_root_folder/data/clean_data/`

Survey Effort

Description: Compiled survey effort for aerial surveys

Location: `gommapps_root_folder/data/effort/`

Surveyed Hexagons

Description: Compilation of all hexagons that were actually surveyed throughout the study.

Location: `gommapps_root_folder/data/effort/`

Surveyed Transects

Description: Compilation of all transects that were actually surveyed throughout the study.

Location: `gommapps_root_folder/data/effort/`

Database (version 1.0, August 1, 2021)

The data are maintained within a relational database and include a full metadata record, including the relational diagram. Figure 2 shows the relational diagram, a larger version can be found here: [Database Description](#).

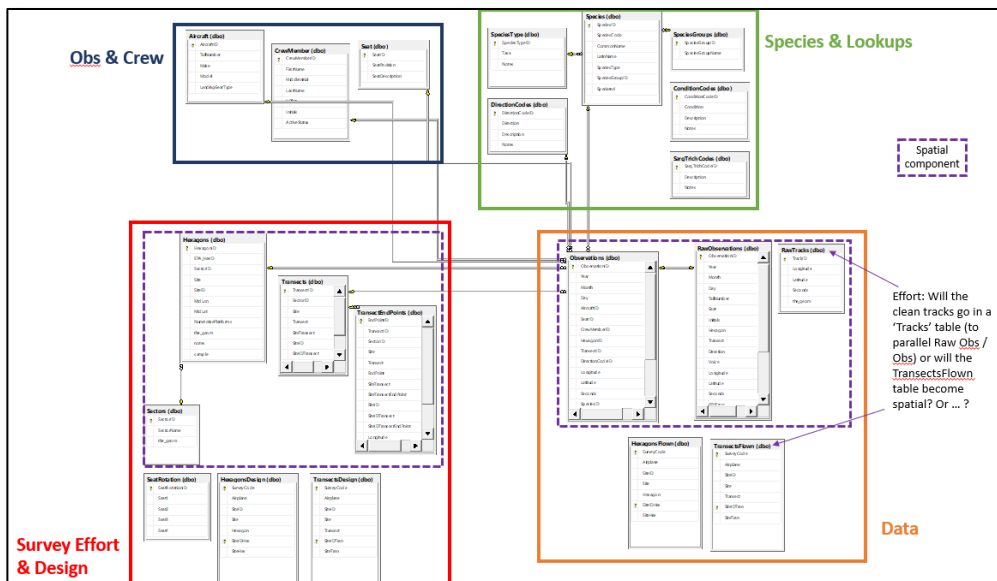


Figure 2. GoMMAPPS aerial seabird survey database diagram.

As the database is a relational & spatial database, a backup of the observation and track data is also being stored outside the database as spatial files (.qdb and .shp) and on Confluence.

DATA LOCATION, ACCESS & ARCHIVE

All aerial seabird survey data are of long-term value and should be retained and preserved. However, if storage space becomes an issue, the backup of the database (which includes survey files, processed observation and track data, and lookup tables) should be retained over raw data files. The documentation and metadata, SOPs, and project record should be preserved and kept with the database.

Storage

Raw and processed data files are stored on USFWS/USGS servers and on [Confluence](#):

GoMMAPPS Database: ifw-dbcsqlcl1.fws.doi.net/ifw9mb-gomapps

GoMMAPPS Folder & Files: \\ifw-hqfs1.fws.doi.net\MB SeaDuck\GOMMAPPS\

The processed data is imported into the database which resides on USFWS server ifw-dbcsqlcl1.fws.doi.net as ifw9mb-gomapps. The database is also backed up using USFWS server/cloud. See the Data Backup SOP on Confluence for instructions on manually backing up the data and database. More information on the database, including a database diagram, can be found on Confluence (see the *Database* section for more information).

Sharing & Access

As data are actively being collected, they will be shared internally with the seabird team through Confluence or through a request to the Database Specialist. Nonbird data (marine mammals and sea turtles) is to be shared with the cross-taxa partners to supplement their own data collection efforts.

Primary access to the database is limited to the Database Specialist. USFWS and USGS team members can obtain access to the database upon request to USFWS IT staff. USFWS IT employees have access to the database for creation, maintenance, and assistance. Non-USGS/USFWS seabird team members cannot access the database directly but have access to the raw and processed data shared on Confluence. Non-DOI employees cannot access the database directly, however they can request a copy of the data.

During active data collection and management, the data/base will be archived twice a year (February & September). Active archiving will ensure the latest version of the database will be backed up and retained, and will incorporate new data as well as updated data checks & processing code.

Restrictions

Data will not be shared publicly until the project is completed. Data can be requested using our [data sharing & use policy](#), and any requests will be reviewed. A record of all data requests, external of the seabird team, are documented for posterity. Any proprietary or sensitive information will be handled on a case-by-case basis.

Archive & Preservation

Upon completion of the project, data will be accessible online through NOAA's National Centers for Environmental Information (NCEI; formerly NODC) <https://www.ncei.noaa.gov/> data portal as well as through USGS Science Base. The GoMMAPPS team contact with NCEI is Errol Ronje - errol.ronje@noaa.gov.

The data archive will include the following: aerial survey data csv, aerial survey data dictionary, aerial survey data management plan, shapefiles of aerial survey design (hexagons, transects, endpoints), vessel survey data csv, and vessel survey data dictionary csv. The archive with NCEI will be a joint effort across all taxa partners: each group (marine mammal, sea turtle, seabird) will submit their archive separately and the archives will be grouped together as a collection for ease of access and dissemination. Once the data is archived with NCEI, they will also be submitted to USGS Science Base.

Products & Media

Throughout the project, pictures and videos were taken by various crew members and collected by the Database Specialist. The media files are stored within project folder system (location: `gommapps_root_folder/media/`) and photos are uploaded to the USFWS Pilot Biology Flickr page, in the GoMMAPPS folder. The Migratory Bird Surveys Branch Chief will either provide access to the Database Specialist to upload the photos or upload the photos for the project directly. The Flickr site is located: <https://www.flickr.com/photos/pilotbiologist/albums/72157669350461999>.

The project resulted in various products and deliverables. For example, at the end of each survey season, a one page summary report was compiled for distribution to all parties involved in the project, including funding agencies, supervisors, and other stakeholders. These reports,

as well as other products, such as presentations, are retained with the project files and database. Location: gommapps_root_folder/documentation/

This document is intended to be a living document to reflect the relationship between the parties and the addition of new potential signatories.