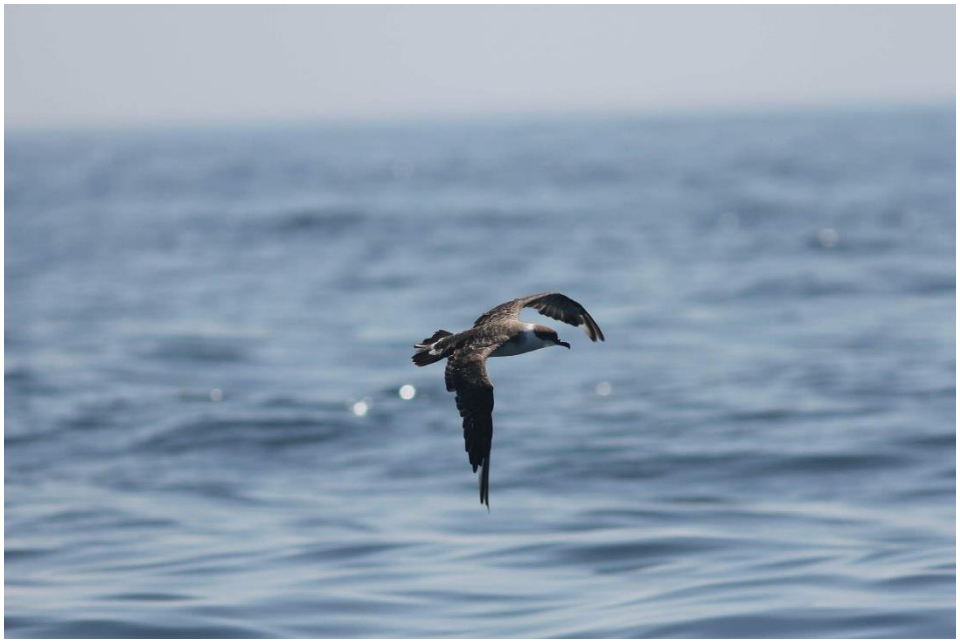
RWSC

Regional Wildlife Science Collaborative

Northwest Atlantic Seabird Catalog Metadata and Field Definitions



RWSC Marine Bird Distribution Working Group

September 2026

Contents

About the Northwest Atlantic Seabird Catalog 2
Versioning 3
Overview of Data Structure 3
Data Standardization 4
Field Definitions 5
Survey Metadata 5
Observations 7
Effort 9
Appendix A. Guide to Changes in Fields 11
Appendix B. Lookup Tables and Codes 12

About the Northwest Atlantic Seabird Catalog

The Northwest Atlantic Seabird Catalog (NWASC) compiles at-sea observational survey data from the Northwest Atlantic region from southern Atlantic Canada to Florida. Datasets are collected by various partners primarily during systematic aerial or vessel-based seabirds surveys at sea that includes both observations and survey effort, though the database was designed to be flexible to accommodate additional coastal and at sea observations of birds and other taxa. While the primary focus of the database is seabirds, data an a range of other at-sea observations (other birds, marine mammals, fishes, sea turtles, anthropogenic objects such as vessels and debris) are also included in many datasets. The NWASC is a powerful tool for examining the abundance and distribution of seabirds in the region to inform coastal and offshore planning activities and supports research and decision-making requiring large, compiled datasets across the region. The data in the NWASC have been used to develop modeled monthly predictions of marine bird species relative densities across the U.S. Atlantic which have been integrated into regional ocean data portals, for oil spill response and planning, listed species assessment, and species inventory for a Marine National Monument management plan, among other uses.

The NWASC was created in 2005 and has been supported by the U.S. Geological Survey, U.S. Fish and Wildlife Service, the U.S. Bureau of Ocean Energy Management, and the U.S. National Oceanic and Atmospheric Administration. In 2025-2026 the Regional Wildlife Science Collaborative, Northeast Regional Ocean Council, Mid-Atlantic Regional Council on the Ocean, New York State Energy Research and Development Authority, and Maine Department of Energy Resources supported the integration of NWASC into OBIS-SEAMAP.

Versioning

Prior to integration into OBIS-SEAMAP, versions of the NWASC were tracked in the Survey Metadata table in the “version_nb” column, indicating the first version of the Northwest Atlantic Seabird Catalog that included a particular dataset. With the shift to self-submission via OBIS-SEAMAP, the definition of a “version” of the Catalog has shifted and a new version will only be released when the structure of the database has changed. Thus, the addition of new datasets will be tracked in the change history directly on OBIS-SEAMAP. We recommend following standard OBIS citation guidance for individual datasets with DOIs using the below citation when citing the full Catalog.

Example Citation for Individual Dataset
Lockhart, G.G., DiGiovanni Jr., R.A., DePerte, A.M. 2014. Virginia and Maryland Sea Turtle Research and Conservation Initiative Aerial Survey Sightings, May 2011 through July 2013.

Downloaded from OBIS-SEAMAP (http://seamap.env.duke.edu/dataset/1201) on xxxx-xx-xx. DOI

Suggested Citation for Full Catalog

NWASC. Year. The Northwest Atlantic Seabird Catalog VXX. Downloaded from OBIS-SEAMAP (http://seamap.env.duke.edu/ on xxxx-xx-xx.

Overview of Data Structure

This document defines field definitions for both data submission and use of existing data, with two goals: 1) to reflect the structure of the NWASC prior to integration with OBIS-SEAMAP, and 2) to work towards additional data standardization where possible. There are three primary components to datasets in the NWASC:

1. **Survey Metadata** – High-level metadata about the survey effort, including key methodological information such as survey and platform type, survey strip width, project information, etc.
2. **Observations** – Observations of individuals or groups of birds with associated spatial data (latitude, longitude), along with additional information about the observed birds (for example, behavior) and information about the observer and observation conditions.
3. **Effort** – Standardized measures of the locations in which a survey was conducted, which provides key information on where birds were not observed and allows for the modeling of density. This may be provided as 1) track points and/or transect lines plus survey strip width, which can be combined into area covered or 2) survey polygons that already combine this information. As with observations, effort should include information on conditions (e.g., visibility, sea state) and methodology (e.g., observer) that may vary within a survey.

During submission of data, IDs are assigned across components ensuring that these values are unique across the full database (e.g., Catalog/OBIS). These include:

- **Dataset ID –** Included in metadata, observations, and effort to link these components of a dataset together
- **Transect ID –** Unique identifier for individual transects, used to link to Tracks and Observations; Tracks and Observations are linked based on IDs supplied by the data contributor between **source_transect_ID** and **source_track_ID** and **source_obs_ID**
- **Track ID** – Unique identifier for individual points along transects
- **Observation ID –** Unique identifier for individual observations

In addition to the IDs created during data integration into the Catalog, there are typically versions of Transect ID, Track ID, and Observation ID within each dataset that are provided during data submission (source_transect_ID, source_track_ID, source_obs_ID).

Data Standardization

Column names in the sections below represent the current structure of the database. In some cases, column names have been changed to align with the OBIS-SEAMAP broader infrastructure or for clarity (see Appendix A for changes from previous nomenclature).

In an effort to standardize particular components of observations with the integration of the NWASC into OBIS-SEAMAP, we have developed new value-restricted fields, indicated by standardized language that utilizes lookup tables and data restrictions to ensure data is standardized where possible moving forward (fields ending in "cd" or "nb"). As full standardization of historical data was not possible, there may also be a unstandardized version of a column containing historical data (fields ending in "tx"). Moving forward data should be submitted in the standardized format or with a custom lookup table if needed. Please refer to Appendix B for standardized options for fields. Categories and definitions were standardized to the degree possible with existing survey protocols and data collections tools, including Seascribe, Eastern Canadian Seabirds at Sea, European Seabirds at Sea, and Gulf of America protocols. In some cases, standardized columns have been implemented from the inception of the NWASC (e.g., species codes, survey type) and in other cases were implemented during the OBIS-SEAMAP implementation (e.g., visibility). Finally, during the integration process, additional columns were added to improve clarity in methodology and consistency in reporting. In some cases it was feasible to fill these columns with the information provided, but in other cases these should be viewed as information only provided consistently from the date of integration forward.

Field Definitions

Survey Metadata

Metadata is included as a table representing broad information about the overall survey design and methodology (Table 1). This table in the NWASC is called "**Dataset,**" with "**Dataset ID**" as the unique overall identifier linking to observations and effort. Dataset IDs are automatically assigned upon data submission. Definitions are included below the table with the exception of lookup tables, which can be found in Appendix B. **Metadata should be provided as a CSV or Excel file.**

*Table 1. Metadata table column names and definitions. * indicates suggested required fields. Green fields have been added to existing structure during the integration with OBIS-SEAMAP; grey indicate fields that are created during submission (not by the data provider). Bolded indicate fields with lookup tables (see Appendix B for definitions).*

Column name	Definition
dataset_id	Unique numeric identifier for dataset across full database – created during submission (same as OBIS-SEAMAP Dataset ID)
dataset_name*	Survey/Dataset name. To ensure this is unique, if long-term project, include Month, Year information.
platform_type_cd*	The type of survey which includes information on a combination of platform and method: a=boat; b=aircraft; c= camera; f=fixed point ground survey; g=area-wide ground survey; v=video.
survey_method_cd*	The method of how the survey was conducted: cbc= Christmas bird count; byc=bycatch; cts=continuous time strip; dth=discrete time horizon; dts=discrete time strip; go=general observation; tss= targeted species survey.
distance_sampling_cd*	Indicates whether distance sampling was used: DA=measured distance and angle"; PD=perpendicular distance; NA=none.
platform_name*	Platform name (for vessels) or type/model (aircraft)
dataset_type_cd*	The design of the survey: de=derived effort; og=original general observation; ot=original transect.
whole_survey_width_m*	Planned width of transects in meters (this may include a single observer on one side of a platform or multiple observers on different sides of a platform). See "individual_observer_survey_width_m" for information on how these two fields interact.
individual_observer_survey_width_m	Planned width for single observers. If value matches whole_survey_width_m, this assumes there were not multiple simultaneous observers. If different values are entered for 'whole_survey_width_m' and "individual_observer_survey_width_m" then assume there are multiple simultaneous observers.
pooled_observations	Refers to whether effort and observations are separated by observer or pooled across multiple observers. Set pooled to yes when there were confirmed multiple simultaneous observers looking at different areas whose effort was additive and their data are not separated by observer. Set to no when there were confirmed multiple simultaneous observers at a time with effort and observations supplied separately. Leave blank if single observer.
dataset_summary*	Brief description of the dataset, including project information, geographic area, timing, and goals.
planned_speed_knots	Planned speed of platform in knots.
effort_type*	Indicates whether effort is provided as points, lines, polygons, or none.
flight_height_method	Describes the method used for determining flight height (e.g., human estimate, laser range finder, etc.).
sponsors	Sponsors/Funders
parent_project	If survey is part of a larger project, this name should be consistent across all surveys submitted.

version_nb	Corresponds to the first version of the Northwest Atlantic Seabird Catalog that included this dataset.
publications	Information on publications resulting from this dataset, including full citation and DOI
additional_info	Any additional information related to the dataset that would be useful; in the case of targeted species surveys, this should include information on what species were included/excluded from observations.
share_policy*	OBIS-SEAMAP defines the following data sharing policy options: 1) Public domain, 2) CC BY, 3) CC BY-NC, 4) Permission required. More information: https://seamap.env.duke.edu/content/user_faq
citation*	Dataset citation following OBIS-SEAMAP suggested citation format. More information: https://seamap.env.duke.edu/content/content/termsofuse
updated	Date the dataset was last updated in OBIS-SEAMAP. In the case of historical datasets, dates were added based on version number updates from previous iterations of the NWASC.
catalog_id	Dataset ID in the Catalog ver. 0.10.0.

Observations

Observations represent the primary data for surveys (Table 2). While the NWASC focuses on seabirds, sightings of other species groups and at-sea objects may also be recorded and included in observations. Each observation (individual or group of individuals) must have associated time and location information and ideally be linked to effort (via "source_transect_id" and "whole_transect"). In addition to information on the individuals observed and associated behavior and characteristics, each observation should also have associated information on the observer (name, position) and sea conditions that affect detectability. **Observations should be provided as CSV or ESRI shapefile.**

*Table 2. Observation table column names and definitions. * indicates suggested required fields. Green fields have been added to existing structure during the integration with OBIS-SEAMAP; grey indicate fields that are created during submission (not by the data provider). Bolded indicate fields with lookup tables (see Appendix B for definitions).*

Column Name	**Definition**
dataset_id	Unique numeric ID for dataset across whole database.
observation_id	Unique numeric IDs for observation records across whole database.
transect_id	Unique numeric IDs for transects across whole database.
source_observation_id*	Unique, numeric ID of observation record.
source_transect_id*	Identifier linking to transect effort.
whole_transect*	Name corresponding with the "whole_transect" in transect dataset.
longitude*	Longitude of the sighting in decimal degrees (WGS84 datum not UTM).
latitude*	Latitude of the sighting in decimal degrees (WGS84 datum not UTM).
obs_dt*	Date of the sighting in YYYY-MM-DD format.

obs_tm*	Time of the sighting in local time zone in HH:MM:SS in 24-hour clock.
spp_cd*	Four letter species code (standardized).
original_species_tx	How you defined the species in your survey (common name, scientific name, code, ITIS number).
obs_count*	The number of individuals of a species observed (i.e., see at one location and time. Whether the observation occurred during the standardized effort or outside/off transect are indicated in the "in_transect" column.
in_transect	Indicates whether observations indicated in the Obs_count are within or outside of the transect. If blank, assumed to be in transect.
obs_count_intrans_nb*	The number of individuals of a species included in a single observation (i.e. seen at one location and time) within the transect standardized survey effort. Any individuals observed off transect should be indicated in "obs_count_general_nb" instead.
obs_count_general_nb	Indicates only observations recorded outside of standardized survey effort. All observations during the standardized effort should be indicated in "obs_count_intrans_nb".
life_stage_tx	Animal life stage/age (unstandardized) .
life_stage_cd	Animal life stage/age category (standardized). If multiple ages within a group, ideally separate observations are provided for unique individual types linked via "flock_id".
sex_tx	Animal sex (unstandardized).
sex_cd	Animal sex (standardized). If multiple sexes, ideally separate observations are provided for unique individual types linked via "flock_id".
plumage_tx	Plumage of animal (unstandardized) used to indicate plumage characteristics that may be helpful in age/sex determination or indicating morphotype. Suggested categories included in lookup tables.
behavior_tx	Behavior of animal (unstandardized).
behavior_cd	Behavior of animal (standardized).
travel_direction_tx	Cardinal direction the animal is traveling (unstandardized).
travel_direction	Cardinal direction the animal is traveling (standardized). Accepted values include N, NNE, NE, ENE, E, ESE, SE, SSE, S, SSW, SW, WSW, W, WNW, NW, NNW.
flight_height_tx	Flight height including number and units (unstandardized) from historic data.
flight_height_nb	Flight height in meters (used only in cases where flight heights are not binned).
flight_height_bin_min	If using binned heights, list minimum flight height for the selected bin.
flight_height_bin_max	If using binned heights, list maximum flight height for the selected bin.
association_tx	Association (unstandardized). If an individual/group is observed with another species or with a boat, each observation should be a record of it's own but linked via association.
association_cd	Association (standardized). If an individual/group is observed with another species or with a boat, each observation should be a record of it's own but linked via association.
flock_id	If multiple observations correspond to the same group of animals (e.g., they were separated into distinct observations based on different morphology/behavior), they can be linked via a unique ID for the group.
distance_to_animal_tx	Direct distance to animal (unstandardized). Distance between the observer and the animal (list units in text).
distance_to_animal_nb	Direct distance to animal in meters; a combination of angle_from_observer and distance_to_animal can be used to calculate perpendicular distance.

distance_bin_min	If using distance bins, list the minimum distance for the selected bin.
distance_bin_max	If using distance bins, list the maximum distance for the selected bin.
perpendicular_distance_nb	Perpendicular distance to animal in meters. If the provider provides perpendicular distance, fill in this field instead of distance_to_animal_nb. If the provider provides distance_to_animal_tx/nb with blank angle_from_observer_tx/nb, the value in distance_to_animal_tx/nb is condidered as a perpendicular distance and it is copied to this field.
angle_from_observer_tx	Angle from observer (unstandardized) in relation to the direction of the transect .
angle_from_observer_nb	Radial angle from observer (standardized); numeric values only in relation to the direction of the transect (not true direction); for example, for observations on the port side values should range from 270-359; starboard 0-90; acceptable values range from 0 to 360.
observer_confidence_tx	Observer confidence in ID (definite, possible, probable).
camera_reel	ID for imagery (digital aerial camera reel or if photos were taken as part species ID documentation for other data types).
seastate_beaufort_nb*	Beaufort sea state- A scale relative to observation conditions from 0 (calm) to 6 (strong breeze) to 12 (hurricane force winds).
wave_height_tx	Wave height (list units in text).
whitecaps_tx	Whitecaps - describing the state of the sea and degree of windiness – alternatively provided instead of beaufort in historic data (Beaufort preferred and required moving forward).
visibility_tx	Visibility provided as a number plus units. Can be supplied as a range; complete visibility is indicated as “unlimited”.
visibility_cd	Visibility category (standardized).
glare_tx	Describing the amount of glare off the water (unstandardized).
glare_cd	Describing the amount of glare off the water (standardized).
wind_speed_tx	Wind speed (list units in text).
wind_dir_tx	The direction from which the wind is coming from.
cloud_cover_tx	Cloud cover indicated as a percent.
weather_tx	Description of the weather conditions (unstandardized).
weather_cd	Description of the weather conditions (standardized).
seasurface_tempc_nb	Sea surface temperature in Celsius.
salinity_ppt_nb	Salinity in ppt.
observer_tx*	Observer name.
obs_position_cd*	Describing the observer position in the craft (standardized).
observer_comments_tx	Comments.
geom	Geometry added by OBIS-SEAMAP
catalog_id	Dataset ID in the Catalog ver. 0.10.0.

Group Observations

There may be times when a group of birds is observed that includes multiple species, sexes, ages, etc. Ideally, these are entered as separate observations with associated categories of information and linked via “flock_id”.

Effort

There are two types of effort provided historically in the NWASC, including "**Transects**" and "**Tracks**". It is possible that these were interpreted differently throughout the life of the Catalog, but "Transects" generally represent lines of overall effort, while "Tracks" represent points along those lines. All historical data have "Transects", some of which were derived by data managers during incorporation into the Catalog (see Dataset Type above), but only a portion of datasets have "Tracks". **Data can be provided as CSV, Excel, ESRI shapefile or geodatabase.** In cases where effort data is provided points, OBIS-SEAMAP will convert these to tracklines for viewing purposes in the data portal. Polygon effort can be turned on by going to the "Layers" subpanel.

We recommend that data providers include both Transects and Tracks for new data submissions (as relevant), with tracks represented as points during the survey and transects as line data (planned or actual survey effort). We also recommend the use of Tracks when available for analysis of historic data as a more accurate spatial representation of effort, though effort-related metadata may be more complete in the transect component. There is overlap in fields between these two effort types.

*Table 3. Effort table column names and definitions. * indicates suggested required fields. Green fields have been added to existing structure during the integration with OBIS-SEAMAP; grey indicate fields that are created during submission (not by the data provider). Bolded indicate fields with lookup tables (see Appendix B for definitions). Effort type indicates whether this information is provided for tracks, transects, or both.*

Column Name	Definition	Effort Type
dataset_id	Unique numeric ID for dataset across whole database.	Both
track_id	Unique numeric IDs for points along track across whole database.	Track
transect_id	Unique numeric IDs for transects across whole database.	Both
source_track_id*	Unique numeric track ID.	Track
source_transect_id*	Unique transect ID (links effort to observations; also links between effort types).	Both
whole_transect*	Name of transect.	Both
traversal_speed_nb	Speed of platform during transect.	Both
Platform name (previously conveyance_name_tx)	Name of platform – If this differs across transects within a survey indicate platform name separately for observations/effort. If blank, assume platform indicated in metadata.	Both
altitude_nb	Altitude of observer in meters.	Both

transect_distance_nb	Total length of transect (km).	Both[1]
start_dt* / track_dt*	Start date of transect line or instantaneous time for track points.	Both
start_tm* / track_tm*	Start time of transect or instantaneous time for track points.	Both
end_dt*	End date of transect.	Transect
end_tm*	End time of transect.	Transect
longitude*	Instantaneous Longitude (WGS84, decimal degrees).	Track
latitude*	Instantaneous Latitude (WGS84, decimal degrees).	Track
start_lon*	Starting Longitude (WGS84, decimal degrees).	Transect
start_lat*	Starting Latitude (WGS84, decimal degrees).	Transect
end_lon*	Ending Longitude (WGS84, decimal degrees).	Transect
end_lat*	Ending Latitude (WGS84, decimal degrees).	Transect
seastate_beaufort_nb*	Beaufort sea state represents a scale relative to observation conditions from 0 (calm) to 6 (strong breeze) to 12 (hurricane force winds) based on lookup table	Both
glare_tx*	Describing the glare off the water (unstandardized) .	Both[2]
glare_cd	Describing the glare off the water (standardized).	Both[2]
visibility_tx*	Distance of adequate visibility from observation location provided as a number plus units (unstandardized). Can be supplied as a range; complete visibility indicated as “unlimited”.	Both[1]
visibility_cd	Distance of adequate visibility from observation location (standardized).	Both[1]
wind_speed_tx	Wind speed (list units in text).	Both[1]
wind_dir_tx	Indicates direction from which the wind is coming.	Both[1]
wave_height_tx	Wave height (list units in text).	Both[2]
whitecaps_tx	Describing the state of the sea and degree of windiness – alternatively provided instead of Beaufort (Beaufort preferred).	Both
cloud_cover_tx	Amount of cloud cover as a percent.	Both[2]
weather_tx	Description of weather conditions (unstandardized).	Both
weather_cd	Description of weather conditions (standardized).	
observers_tx	Observer name.	Both
obs_position_cd	Describing the observer position in the craft (standardized).	Both
comments_tx	Comments on transect/track.	Both
catalog_id	Dataset ID in the Catalog ver. 0.10.0.	Both

[1] Previously only in Track, but now included in both

[2] Previously only in observation table but now included effort tables as well

Appendix A. Guide to Changes in Fields

General changes

- All the columns that expect standardized codes have a suffix "_cd" (e.g. age_cd renamed from age_id). They have corresponding free-text columns (e.g. age_cd has age_tx).
- Columns that expect text have a suffix "_tx" (e.g. glare_tx).

Metadata

- distance_sampling_cd is added to indicate if the survey followed the distance sampling method.
- platform_name_id is replaced with platform_name. platform_name accepts free text of the platform name and is no longer tied to a lookup table.
- share_level_id is replaced with share_policy. share_policy accepts Creative Common's share policies such as CC-BY, CC-BY-NC.

Observations

- Obs_count_intrans_nb and obs_count_general_nb have been combined into Obs_count; whether observations occurred on transect or outside of transect are now indicated in the "in_transect" column
- age_tx & age_cd are renamed to life_stage_tx & life_stage_cd to better represent the expected values (e.g. "Adult", "Juvenile")
- perpendicular_distance_nb is added. This field should be used if the distance to the animal is a perpendicular distance (already calculated when submitting the data).
- distance_bin_min and distance_bin_max are added in case the distance to the animal is a range.
- obs_position is replaced with obs_position_cd which accepts standardized codes (e.g. "LF", "LR" etc.).

Effort

- stop_dt & stop_tm are renamed to end_dt & end_tm.
- For tracks, start_dt & start_tm are renamed to track_dt & track_tm as track data expect points.
- temp_start_lon, temp_stop_lon, temp_start_lat & temp_stop_lat in the tracks table are removed. Use longitude & latidude.
- temp_start_lon, temp_stop_lon, temp_start_lat & temp_stop_lat in the tracks table are renamed to start_lon, end_lon, start_lat & end_lat.

- seconds_from_midnight_start is removed.

Appendix B. Lookup Tables and Codes

- association
- beaufort
- behavior
- dataset_type
- distance_sampling
- glare
- life_stage
- observer_position
- platform_type
- sex
- species
- survey_method
- visibility
- weather