Dugong aerial survey database

USER MANUAL

Updated 02.06.2015

1. Dugong aerial survey database

This dugong aerial survey database has been compiled as part of a project funded by the Australian Marine Mammal Centre (AMMC). It is MySQL based and currently contains data from 54 aerial surveys for dugongs in nine regions along the Australian coast since 1984 including: Shark Bay (WA), Exmouth (WA), Pilbara (WA), Gulf of Carpentaria (NT/QLD), Torres Strait, northern Great Barrier Reef (QLD), southern Great Barrier Reef (QLD), Hervey Bay (QLD), and Moreton Bay (QLD).

While the database is available for any user, you are required to contact the data custodian to obtain access details.

Please familiarize yourself with the methodology and the limitations of the database before using the data for your own projects. Given that the data span more than 30 years, and data collection and analyses have been continuously refined over that period, inevitable differences between individual surveys exist which need to be considered when working with the dataset. We highly recommend reading the corresponding publications (reports and papers), comments in the database and examining the provided shapefiles to fully understand each survey.

2. Survey description

All surveys used the strip transect aerial survey technique detailed in Marsh and Sinclair (1989) and later refined by Pollock et al. (2006). Surveys were usually conducted at a height of 450 or 500 feet (134 or 152 m, respectively) above sea level. Individual parameters for each survey are noted in the database. Transects on the water surface on each side of the aircraft were demarcated using fiberglass rods attached to artificial wing struts on the aircraft. Usually, two tandem teams of observers on each side of the aircraft scanned their respective transect and recorded sightings onto separate tracks of an audio recorder. The two members of each tandem team operated independently and could neither see nor hear each other when on transect. Dugongs were the main focus of these surveys, followed by dolphins, marine turtles and other marine megafauna, such as sharks, rays and seasnakes. The database currently only holds dugong data.

The data from each survey area were analysed to determine estimates of relative dugong abundance, following the methodologies developed by Marsh and Sinclair (1989) and later modified by Pollock et al. (2006). These methods attempt to correct for availability bias (animals not available to observers because of water visibility) and perception bias (animals visible in the survey transect but missed by observers (*sensu* Marsh and Sinclair, 1989).

3. Caveats

Limitations of the data collection:

- It is important to note that dugong sightings during aerial surveys (and consequently population size estimates) are an underestimate of the true number of animals in the study area. Observers were instructed to only record animals that they could clearly identify as dugongs, therefore minimizing the risk of false positives. In instances, where observers were unsure about their observation (i.e. could not confidently identify the object as a dugong), the sighting was disregarded. False negatives were therefore unavoidable in the dataset.

- Recent work of Hagihara et al. (2014) indicates that improved methodologies (e.g., the use of Availability Correction Factors that are based on water depth as well as Environmental Conditions inferred from Sea State and Water Turbidity only to estimate dugong availability bias during aerial surveys) can improve dugong population abundance estimates from surveys from 2000 (from when the Sea State and Water Turbidity were collected for each sighting).

- Early surveys (before the mid-90s) did not use Global Positioning Systems (GPS) and sighting locations were inferred from the average flight speed and the direction of flight.

- Each survey used trained observers, however, observer specific differences and potential observer error was unavoidable. Some observer error can be corrected for using the tandem observer data and Mark-Recapture models.

- Dugongs can move considerable distances (100s of km) in a few days and due to the large spatial extend of each survey, it cannot be ruled out that animals moved in and/or out of the study area. While transects were spaced far enough apart to avoid double-counting of animals during a single flight, breaks between flights (e.g. overnight or due to poor weather conditions) may have influenced the animal counts.

- Conducting dugong aerial surveys in large and often remote regions present several logistical challenges. In addition, financial limitations have to be considered and often, weather can be a critical factor impacting on the progress of the survey. All these issues have to be weighed against the scientific rigor of the study and a balance needs to be found. Consequently, surveys will be slightly different from each other (e.g. different transect lengths or number of transects flown), which needs to be considered when working with the dataset.

Limitation of the data analyses:

- Data analyses following the Marsh and Sinclair (1989) method have been conducted for most surveys. The refined methodology of Pollock et al. (2006) is only available for selected surveys conducted past 2005.

- Some study areas have a very low detection frequency of dugongs, which can pose difficulties for population size estimations. In order to make statistical analysis more robust, blocks with less than five dugong sightings were usually excluded from the analyses. This restriction results in an underestimate of the actual population size.

Limitations of the database:

- Combining data from over 30 years of aerial surveys cannot be achieved without making compromises with regards to the database structure and layout. The database was compiled using all available data (published reports, shapefiles, and raw data). On some occasions, errors in the published reports were detected and these were pointed out in the database. Please make sure you read the comments linked to the entries.

- Like other population survey relational databases, the dugong aerial survey database has to differentiate between '0' (i.e. surveyed, but no dugongs seen) and 'non-observations' (i.e. not surveyed) with a type and/or value. Varchar (Variable Character Field) was the type chosen and the value is an empty string. This approach differs from some other MySQL implementations that use 'NULL' instead. Using an empty string aims to make it easier for end users to read, paste and copy data to perform analysis with a minimum concession for the specific use case.

- Therefore, we recommend copying the database to CSV to be used in Excel or another tool for detailed data analysis. The example SQL queries section documents what queries and use cases the database itself is optimal for.

4. Database design

Information in the database is organized under the following headings.

Region

The database encompasses data from nine survey regions across Australia:

- Moreton Bay (MB)
- Hervey Bay (HB)
- southern Great Barrier Reef (sGBR)
- northern Great Barrier Reef (nGBR)
- Torres Strait (TS)
- Gulf of Carpentaria (GoC)
- Pilbara (Pil)
- Exmouth (Ex)
- Shark Bay (SB)

The spatial extend of individual regions can be seen in the accompanying shapefiles (links provided in this table). Regions are primarily based on logistical limitations for aerial surveys and do not necessarily represent habitat borders for individual dugong populations.

Survey

This tab provides an overview of individual surveys in each survey region. Information includes:

- Survey ID (region-year-month)
- Region ID
- Reference ID (published and unpublished reports and papers)

Survey Background Note

Here the user can find important information linked to individual survey IDs.

Survey Methodology

Survey specific methodology is listed here. Information includes:

- Survey ID (region-year-month)
- Survey start (year-month-day)

- Survey end (year-month-day)
- Flight height (in feet)
- Transect width (in km)
- Number of observers (e.g. 2*2, which means two rows of two observers each)
- Comments (explanatory notes)

Block

Each survey region is sub-divided into individual blocks. Block names follow an alphanumeric system used in each respective survey, with the letters usually symbolizing the region (e.g. MB for Moreton Bay) and numbers the block number. Large regions (i.e. sGBR and nGBR) used a slightly different conversion, with 'southern' blocks (S), 'central' blocks (C), and 'northern' blocks (N). Block locations did not vary between years, but block sizes may differ slightly over time. The location of blocks can be seen in the provided shapefiles. Dugong population estimates are usually calculated on the block level.

Survey Block Summary

For each survey ID, this tab provides information about the size (in km^2) of a survey block (block ID) and its sampling intensity (in %). In addition, information is included for some surveys about the maximum perpendicular (to the transect direction) length (in km) of each block, and the maximum number of possible transects per block. Some additional comments are provided.

Transect

This tab provides information about individual transects. Transect IDs (trans_id) are unique identifiers, while transect names during individual surveys (tf_num_in_surv) may vary between surveys (e.g. transect ID 1003 is named transect 3 in the Moreton Bay survey in November 1999, but called transect 1003 in the Moreton Bay survey in November 2011). Coordinates for each transect start and end point are provided and may vary slightly between surveys due to variations in the actual flight paths. Information is provided about transect length (in km), average flight height (in feet), transect area (in km^2), start and end times and flight direction. Some additional comments are provided.

Survey weather summary

The weather conditions during each survey (survey ID) are provided as:

- Maximum wind speed (in knots)
- Minimum and maximum cloud cover (in octas)
- Minimum and maximum cloud height (in feet)

- Beaufort sea state (minimum, maximum and the overall mean of the individual modes for each transect)
- Glare in the north, south, and overall (minimum, maximum and the overall mean of the individual modes for each transect). Glare is recorded on a scale from 0 to three with 0 = no glare; 1 = less than 25% of the field of view affected; 2 = between 25-50% of the field of view affected; 3 = more than 50% of the field of view affected.
- Minimum horizontal air visibility (in km)

Transect weather

The weather conditions for each transect (identified with its unique ID (trans_id) and the identifier during a survey (tf_num_in_survey) are provided as:

- Beaufort sea state (minimum, maximum and mode)
- Glare in the north and south (minimum, maximum and mode). Glare is recorded on a scale from 0 to three with 0 = no glare; 1 = less than 25% of the field of view affected; 2 = between 25-50% of the field of view affected; 3 = more than 50% of the field of view affected.

Zone

During aerial surveys, the transect strip was usually sub-divided into three to four zones of equal width: upper, middle and lower zones were used in earlier surveys; and Very high (V), High (H), Medium (M), and Low (L) zones were used in more recent surveys. In addition to sightings that occurred on the transect strip, some surveys also provide information on animals sighted outside (O –away from the plane) or inside (I – closer to the plane) the transect strip. Sightings that were not on the transect strip will need to be excluded for dugong population size estimations. Sometimes, groups of animals were sighted across (A) several zones.

Sighting

Dugong sightings during an aerial survey are provided with the following information:

- Sighting ID (unique identifier)
- Survey ID
- Sighting type (dugong or herd)
- Transect number in survey
- Transect ID
- Block ID
- Observation time
- Observer (PF – port front; PR – port rear; PB – port both; SF – starboard front; SR – starboard rear; SB – starboard both)
- Number in group (minimum count)
- Number at the surface

- Number breaking the surface (where some part of the animal breaks the water surface) or deep (which one of those two descriptor is valid is specified in the next column)
- Break_or_deep (whether the previous column refers to numbers recorded breaking the surface or deep)
- Zone (see above)
- Number of calves (calves are defined as less than 2/3 the size of the cow and swimming in close proximity to her)
- Coordinates of the sighting
- Water turbidity (recorded on a scale from 1 to 4 with 1 = clear, shallow water with the seafloor being clearly visible; 2 = variable depth and visibility with the seafloor being visible but unclear; 3 = clear, deep water with the seafloor not visible; 4 = turbid water with variable depth with the seafloor not visible.
- Comments

Sighting Calves Summary

For each survey (survey ID), the total number of dugongs (num_total), total number of calves (num_calves) and the calf percentages (perc_calves) are summarized. Some additional comments are provided.

Survey Group Size

The average dugong group sizes for individual sections in a survey with the coefficient of variation and/or the standard error are provided for each survey.

Survey Pop Summary Marsh & Sinclair Block

Dugong population size estimates ($\hat{N}$ ± standard error) after the methodology developed by Marsh and Sinclair (1989) are provided for individual blocks for each survey. In some cases, estimates are not available because (1) the block was not surveyed (ns), (2) there were too few estimates (tfe, <5 dugongs sighted), (3) the block was surveyed using a zig zag transect pattern (zzt) across the depth gradient instead of the standard parallel transect pattern, or (4) a different survey design (dd) was used. Estimates in the database originate from published references and have not been verified against the raw data. If reports were subsequently corrected, the database displays the corrected estimate plus details of the original estimate in the comments column.

Survey Pop Summary Marsh & Sinclair Total

Dugong population size estimates ($\hat{N}$ ± standard error) after the methodology developed by Marsh and Sinclair (1989) are provided for the entire survey region for each survey. In some cases, the estimate for the entire region differs slightly to the sum of the estimates for the individual blocks due to rounding.

Survey Pop Summary Pollock Block

Dugong population size estimates ($\hat{N}$ ± standard error) after the methodology developed by Pollock *et al.* (2006) are provided for individual blocks for each survey In some cases, estimates are not available because (1) the block was not surveyed (ns), (2) there were too few estimates (tfe, <5 dugongs sighted), (3) the block was surveyed using a zig zag transect pattern (zzt) across the depth gradient instead of the standard parallel transect pattern, or (4) a different survey design (dd) was used. Estimates in the database originate from published references and have not been verified against the raw data. If reports were subsequently corrected, the database displays the corrected estimate plus details of the original estimate in the comments column.

Survey Pop Summary Pollock Total

Dugong population size estimates ($\hat{N}$ ± standard error) after the methodology developed by Pollock *et al.* (2006) are provided for the entire survey region for each survey. In some cases, the estimate for the entire region differs slightly to the sum of the estimates for the individual blocks due to rounding.

Survey Correction Factors

This tab provides information about the correction factors used for dugong population size estimations:

- Perception Correction Factors (for port and starboard side; with coefficient of variation, plus the number of animals sighted by each observer (abbreviations as above) that were used for calculations.)
- Availability Correction Factors (with coefficient of variation, plus the number of animals in groups less than 10 sighted at the surface and below the surface that were used for the calculations)

Section lists the particular blocks/transects that the correction factors apply too, if it's not the entire surveyed region. Some additional comments are provided.

Reference

A list of references (e.g. papers or reports) and their URLs is provided for each aerial survey.

5. Example queries

The database is coded in MySQL, the standard language for relational database management systems. For users unfamiliar with this computer language, the following list provides commonly used queries which may help creating your own queries.

Query: Show all dugong sightings (replace 'dugong' with 'herd' for herds)

```
SELECT *
FROM `Sighting`
WHERE `sighting_type` = 'dugong';
```

Query: Show all dugong sightings within a particular region (replace 'MB' with region of interest)

```
SELECT *
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
AND `surv_id` IN (SELECT `surv_id` FROM `Survey` WHERE `reg_id` = 'MB');
```

Query: Show all dugong sightings from a particular survey (replace 'MB-1999-11' with survey of interest)

```
SELECT *
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
AND `surv_id` = 'MB-1999-11';
```

Query: Show all dugong sightings with at least one calf from a particular survey (replace 'TS-2013-11' with survey of interest)

```
SELECT *
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
AND CAST(`calves` AS UNSIGNED) >= '1'
AND `surv_id` = 'TS-2013-11';
```

Query: Show a count of all dugong observation events per survey, ordered descending

```
SELECT `surv_id`, COUNT(*) AS `total observations`
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
GROUP BY `surv_id`
ORDER BY `total observations` DESC;
```

Query: Show a count of all dugong observation events per region, ordered descending

```
SELECT `reg_id`, COUNT(*) AS `total observations`, COUNT(DISTINCT `surv_id`) AS `total surveys`
FROM `Sighting` NATURAL JOIN `Survey`
WHERE `sighting_type` = 'dugong'
GROUP BY `reg_id`
ORDER BY `total observations` DESC;
```

Query: Show a count of all dugong observation events per survey and block

```
SELECT `surv_id`, `block_id`, COUNT(*) AS `total observations`
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
GROUP BY `surv_id`, `block_id`
ORDER BY `surv_id`, `block_id`;
```

Query: Show a count of all dugong observation events by zones (change `zones` to `observers` for observer)

```
SELECT `zones`, COUNT(*) AS `total observations`
FROM `Sighting`
WHERE `sighting_type` = 'dugong'
GROUP BY `zones`
ORDER BY `total observations` DESC;
```

Query: Show the longest and shortest transect lengths per survey

```
SELECT `surv_id`, MAX(CAST(`length` AS DECIMAL(11, 2))) AS `longest transect length`,
MIN(CAST(`length` AS DECIMAL(11, 2))) AS `shortest transect length`
FROM `Transect_Flight`
GROUP BY `surv_id`;
```

Query: Show block summaries ordered by area descending

```
SELECT *
FROM `Survey_Block_Summary`
ORDER BY CAST(`area` AS DECIMAL(10, 2)) DESC;
```

Query: Show a count of all dugong observation events, split by observer

```
SELECT 'Port - front/rear/either/both' AS `observer`, `total observations`
FROM (SELECT COUNT(*) AS `total observations` FROM `Sighting` WHERE `Sighting_Type` = 'dugong'
AND `observer` LIKE '%P%') AS `port_either`
UNION
SELECT 'Starboard - front/rear/either/both' AS `observer`, `total observations`
FROM (SELECT COUNT(*) AS `total observations` FROM `Sighting` WHERE `Sighting_Type` = 'dugong'
AND `observer` LIKE '%S%') AS `starboard_either`
UNION
SELECT 'Front - port/starboard/either/both' AS `observer`, `total observations`
FROM (SELECT COUNT(*) AS `total observations` FROM `Sighting` WHERE `Sighting_Type` = 'dugong'
AND `observer` LIKE '%F%') AS `front_either`
UNION
SELECT 'Rear - port/starboard/either/both' AS `observer`, `total observations`
```

```
FROM (SELECT COUNT(*) AS `total observations` FROM `Sighting` WHERE `Sighting_Type` = 'dugong'
AND `observer` LIKE '%R%') AS `front_either`
UNION
SELECT `observer`, COUNT(*) AS `total observations`
FROM `Sighting`
GROUP BY `observer`;
```

Query: Show all surveys sorted by availability correction factor

```
SELECT `cf_id`, `surv_id`, `avail_cf`, `section`
FROM `Survey_Correction_Factors`
ORDER BY CAST(`avail_cf` AS DECIMAL(11,2)) DESC;
```

Query: Show all surveys sorted by windiness

```
SELECT *
FROM `Survey_Weather_Summary`
ORDER BY CAST(`max_wind_speed` AS UNSIGNED) DESC;
```

Query: Show all transects that had a maximum Beaufort Sea State of 3.

Code: `SELECT * FROM `Transect_Weather` WHERE `beaufort_max` = '3';`

Query: Using the Marsh and Sinclair population estimation data, which survey region (across all surveys) had the largest fluctuation in population estimates (difference between MAX(`nhat`) and MIN(`nhat`) ?

```
SELECT `reg_id`, MAX(CAST(`nhat` AS UNSIGNED)) AS `max nhat`, MIN(CAST(`nhat` AS UNSIGNED))
AS `min nhat`,  MAX(CAST(`nhat` AS UNSIGNED)) - MIN(CAST(`nhat` AS UNSIGNED)) AS `nhat
fluctuation`
FROM `Survey_Pop_Summary_Marsh_Sinclair_Total` NATURAL JOIN `Survey`
GROUP BY `reg_id`
ORDER BY `nhat fluctuation` DESC;
```

6. Acknowledgements

This project was supported through funding from the Australian Government's Australian Marine Mammal Centre, Australian Antarctic Division, Department of the Environment. The database was designed by Steven van der Valk, eResearch Centre, James Cook University.

7. References

Hagihara, R., Jones, R.,E., Grech, A., Lanyon, J.M., Sheppard, J.K., and Marsh, H. (2014) Improving availability estimates by quantifying dugong diving and surfacing patterns. *Marine Mammal Science* **30** (1): 348-366 DOI: 10.1111/mms.12041

Marsh, H. & Sinclair, D.F. (1989). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* **53**: 1017-1024.

Pollock, K., Marsh, H., Lawler, I., & Alldredge, M. (2006). Estimating animal abundance in heterogeneous environments: an application to aerial surveys for Dugongs. *Journal of Wildlife Management* **70**: 255-262.