IMOS ANIMAL TRACKING DATABASE REVIEW AND QUALITY CONTROL ANALYSIS REPORT

October 2015

Andre Steckenreuter, Colin Simpfendorfer, Katherine Tattersall, Xavier Hoenner

Table of Contents

1. INTRODUCTION

The Animal Tracking facility's national Database is a repository and data storage facility built through collaboration between the Integrated Marine Observing System's (IMOS) facilities, IMOS Animal Tracking, e-Marine Information Infrastructure (eMII), and the IMOS Animal Tracking Data Committee representing the IMOS Animal Tracking community. The Database is the primary storage facility for acoustic tag and receiver data collected by the IMOS Animal Tracking facility and members of the IMOS Animal Tracking community.

The Database is a key tool for aggregating data from a wide network of researchers and holds a diverse array of information, including information about projects, array locations, receiver deployments, tag releases, and individual animal detections. The Database is accessed online via a unique user account through a publicly available web application (https://aatams.emii.org.au/aatams), and through the IMOS Ocean Portal (https://imos.aodn.org.au).

2. OBJECTIVES OF THE REVIEW

The IMOS Animal Tracking Database holds more than 66 million detections from 12,137 individual tag IDs (Database back-up October 2015). The main objectives of the Quality Control Analysis and database review are to:

a) Identify false tag detections by applying an advanced False Detection Analysing Algorithm (FDAA);
b) Develop codes to flag false detections in the Database;
c) Identify owners of unregistered tags;
d) Identify registered tags with no detections;
e) Identify registered tags with incomplete metadata.

3. ADVANCED FALSE DETECTION ANALYSING ALGORITHM

As a general rule, false detections caused by signal collision are infrequent, but are positively correlated with the intervals between detections. A large number of tag detections recorded on acoustic receivers may not be attributable to existing tag IDs or detections of released tags although they were not in the vicinity of the receiver. Such invalid detections are due to a phenomenon known as 'false detections' (Pincock 2012). False detections are generated by collision of acoustic signals from multiple tags at one location, with a positive correlation of collision probability with an increasing number of tags and/or short delays between transmission

signals. In some less common instances, false detections can also be due to biological noise or the collision between real tags and biological noise.

Thus far, false detection analysing tools have been applied to discrete datasets of acoustic receiver arrays under controlled conditions (see Pincock 2012; Ocean Tracking Network 2015). The aim of our analysis was to develop an advanced FDAA that can be applied to large datasets including multiple acoustic receiver arrays, such as the IMOS Animal Tracking Database. Applying a process that flags tag detections that are likely to be false detections is an improvement over current practice, as researchers will gain access to datasets that are quality controlled as well as to information crucial for dataset interpretation. Our advanced FDAA method was based on an acoustic telemetry manufacturer's recommendation for identifying false detections in datasets (Pincock 2012). The latter is integrated as the FDA Analyzer Tool as part of the VUE software (VEMCO 2015). The current analysis has refined this method further by integrating two new parameters. All validity parameters are at the receiver level:

(1) *Number of Detections*: Detections are valid if detected more than once;

(2) *Interval Ratio*: Tag detections at a given station/receiver pair are valid if more short than long intervals between detections are observed;

(3) *Velocity Test*: Individual detections are valid if the velocity between neighbouring detections is less than a specified threshold;

(4) *Distance Test*: Individual detections are classified valid if the distance between neighbouring detections is less than a specified threshold.

These assumptions are conservative in that there are conceivable circumstances where real detections might be flagged in this process. However, the detections are not deleted from the database, merely flagged, leaving the final decision on how data is to be utilised to the discretion of users.

3.1 Number of Detections and Interval Ratio

Firstly, we applied the same conservative approach as Pincock (2012) and considered all tags with one detection only at a given station/receiver pair as likely to be invalid. Importantly, we excluded range test and sentinel tags and their associated detections from all analyses, which accounted for ca. 13.5 million detections.

We then considered the ratio of short intervals to long intervals between detections. Pincock's (2012) algorithm used a multiplication factor of 24 for short and long intervals, suggesting a default short interval of 30 minutes and a long interval of 720 minutes. For our analysis, we tested a range of short and long intervals to determine the combination of short/long interval durations that maximised the proportion of valid detections from

registered tags, with short intervals ranging from five minutes to 16 hours and long intervals from 45 minutes to 64 days (Fig. 1).

As no maximum was obtained for the proportion of registered and unregistered tags that passed the FDAA (Fig. 1), we decided to follow VEMCO's conservative approach by setting the short interval time to two hours and the long interval to 24 hours (D. Webber, pers. comm.) as they both yielded an acceptance rate of 99.8% of detections.

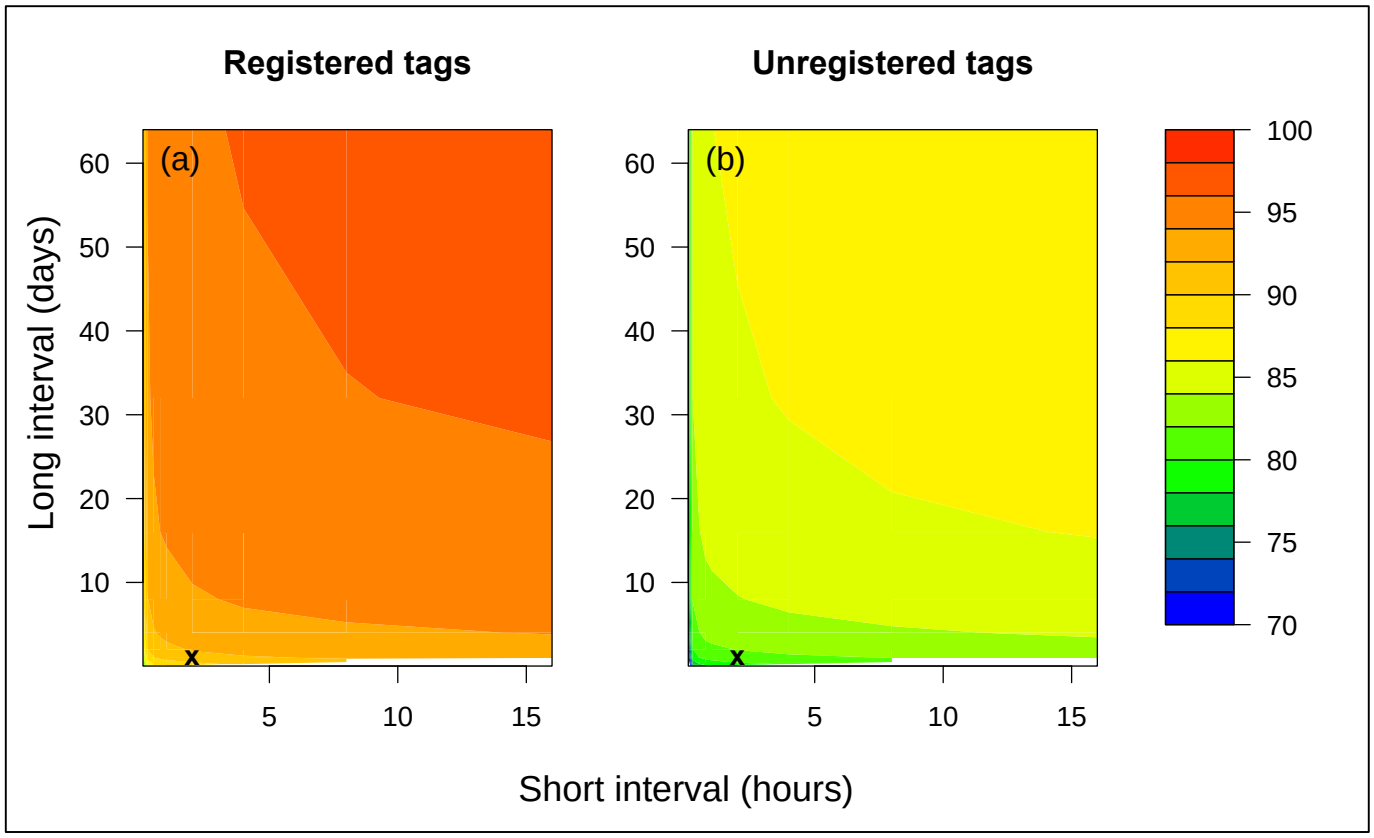

Fig. 1: Proportion of registered (a) and unregistered (b) tags for which detections at a given station/receiver pair were flagged as valid by the advanced False Detection Analysing Algorithm (FDAA), *i.e.* having more than one detection and more short than long intervals. The black cross on each plot indicates the short and long interval times selected for the FDAA (two and 24 hours, respectively). The legend represents the colour code used to display the variations in percentage for various short/long interval combinations.

3.2 Velocity and Distance Tests

3.2.1 Velocity Test

The two fastest species in the IMOS Animal Tracking Database are the Southern Bluefin Tuna (*Thunnus maccoyii*) and Mako Shark (*Isurus oxyrinchus*). In the literature, speed is usually expressed in body lengths (BL) per second.

For Southern Bluefin Tuna, the maximum recorded speed was 2.6 $BL.s^{-1}$ (Fitzgibbon *et al.* 2007). Based on the maximum size of approximately 2.5 m for this species, the corresponding maximum swimming speed was calculated to be 6.5 $m.s^{-1}$. For Mako Sharks, the maximum recorded speed was 1.5 $BL.s^{-1}$ (Sepulveda *et al.* 2004). Based on the maximum size of approximately 4 m for this species, the corresponding maximum swimming speed was calculated to be 6 $m.s^{-1}$. Given these swimming speed estimates, a value of 10 $m.s^{-1}$ between consecutive detections was chosen allowing for potential detection range discrepancies resulting from animals appearing to move faster than expected due to the relative location of receivers to the animal when receivers are close together.

3.2.2 Distance Test

The distance between subsequent tag detections was based on the average distance between acoustic receiver arrays in the IMOS Animal Tracking Database and the potential of tag re-deployment. A conservative value of 100 km was chosen, flagging detections for which the distance to the previous and next detections exceeded this specified limit.

3.3 Stepwise Application of Parameters

The number of registered and unregistered tags in the IMOS Animal Tracking Database and their associated number of detections were tested for validity using the four parameters described above. Since we implemented the advanced FDAA at the receiver level, we assessed the validity of detections for each tag at each station/receiver pair where it was detected.

3.3.1 Number of Detections and Interval Ratio – Validation Process

Of the 5,998 registered tags in the Database, 3,484 (58.1%) have been detected more than once (39.9% of registered tags have never been detected, 2.0% have been detected only once). Of the 6,139 unregistered tags in the Database, 3,224 (52.5%) have been detected more than once while 2,915 (47.5%) have been detected only once (Table 1).

Table 1: Number of registered and unregistered tags in the IMOS Animal Tracking Database and their associated number of detections.

	Tags detected (N)	Tags (N)	Tags (%)	All tags (%)	Detections (N)	Detections (%)	All detections (%)
Registered tags	0	2,393	39.9	19.7	0	0	0
	1	121	2.0	1.0	121	<0.1	<0.1
	>1	3,484	58.1	28.7	51,649,433	100	77.6
Subtotal		5,998	100	49.4	51,649,554	100	77.6
Unregistered tags	0	0	0	0	0	0	0
	1	2,915	47.5	24.0	2,915	<0.1	<0.1
	>1	3,224	52.5	26.6	14,923,335	100	22.4
Subtotal		6,139	100	50.6	14,926,250	100	22.4
Total		12,137		100	66,575,804		100

Of the 3,484 remaining registered tags available for the next step in the validation process, 3,317 (95.2%) encompassing a total number of 51,571,850 detections (99.8%) passed the Interval Ratio Test. Of the 3,224 unregistered tags, 1,655 (51.3%) passed the Interval Ratio Test, encompassing a total number of 14,891,214 detections (99.8%).

3.3.2 Velocity Test – Validation Process

Of the 3,317 registered transmitters that passed the previous test, all had at least one detection that passed the velocity test accounting for a total of 51,227,892 detections (99.2%). 2,163 tags (65.2%) had all their detections pass the velocity test, which accounted for 36,698,068 detections (71.1%).

Of the 1,655 unregistered transmitters that passed the previous test, 1,654 (99.9%) had at least one detection that passed the velocity test accounting for a total of 14,496,293 detections (97.2%). No unregistered transmitter had all their detections pass the velocity test.

3.3.3 Distance Test – Validation Process

Of the 3,317 registered transmitters that passed the previous test, all had at least one detection that passed the distance test accounting for a total of 51,575,080 detections (99.9%). 2,146 tags (64.7%) had all their detections pass the velocity test, which accounted for 38,952,606 detections (75.4%).

Of the 1,655 unregistered transmitters that passed the previous test, all had at least one detection that passed the distance test accounting for a total of 14,692,444 detections (98.5%). No unregistered transmitter had all their detections pass the velocity test.

4. FLAGGING PROCESS

We propose to flag both tag detections that were identified as invalid during the advanced FDAA process (section 3) and tag detections with incorrect or insufficient metadata (Table 2), in the AATAMS Database. The flags have the following values:

0. Test not performed;
1. Passes;
2. Questionable;
3. Missing information.

These flags do not match the IMOS Quality Control (QC) Flag Convention (see Mancini *et al.* 2012) and may be altered to match other IMOS datasets before implementation. Hereafter, we consider a detection as valid when it has a QC flag of '1' for the Number of Detections/Interval Ratio Test or when a detection passed at least two of the three FDAA results described in Table 2.

Using this approach, the total number of registered tags with at least one valid detection was 3,520, *i.e.* 97.6% of all tags with one or more than one detection, accounting for 51,648,205 valid detections (99.9%). 3,063 registered tags (85.0%) had all their detections classified as valid, accounting for 45,898,088 detections (88.9%). The total number of unregistered tags with at least one valid detection was 3,515, *i.e.* 51.4% of all tags with one or more than one detection, accounting for 14,921,130 valid detections (99.9%). 1,354 unregistered tags (22.1%) had all their detections classified as valid, accounting for 13,240,536 detections (88.7%). Of the 3,520 valid registered tags, 959 tags had either the release location, or the release date, or the species name missing or invalid, while 715 had no release information at all (no tag release created in the IMOS Animal Tracking web applocation).

Table 2: Detection validity flags after application of advanced False Detection Analysing Algorithm (FDAA) and assessment of tag/detection metadata quality in the IMOS Animal Tracking Database.

Test type	Parameter	Flag	Meaning	Comment
FDAA results	Number of Detections and Interval Ratio	0 1 2	Test not performed Passes Questionable	Valid if more than one detection and more short than long intervals
	Velocity	0 1 2	Test not performed Passes Questionable	First detection flagged as '3' if metadata missing (unable to calculate velocity)
	Distance	0 1 2	Test not performed Passes Questionable	First detection flagged as '3' if metadata missing (unable to calculate distance)
Tag metadata	Tag release date	0 1 2 3	Test not performed Release date before detection Release date after detection No release date	
	Species information	0 1 3	Test not performed Species information complete No species information	
	Tag ID	0 1 2 3	Test not performed Registered tag with tag release Registered tag without tag release Unregistered tag	
	Release location	0 1 3	Test not performed Metadata complete No latitude/longitude information	
	Tag code	0 1 2 3	Test not performed Tag code correct Tag code incorrect No tag code	Flag A180 tag codes

5. ONGOING INVESTIGATIONS

5.1 Contacting Owners of Unregistered Tags

During the IMOS Animal Tracking Database review, 6,139 unregistered tags were identified. These are partly the result of false detections due to tag collision (see section 3) but possibly include real tags (3,520 registered tags had at least one valid detection *vs.* 3,155 unregistered tags) that were not registered in the Database. Vemco contacted all owners of unregistered tags detected in the IMOS Animal Tracking Database — identified as valid by

the False Detection Analysing Algorithm — on our behalf. From a list of 1,655 unregistered valid tags that passed the first two steps of the validation process (Number of Detections and Interval Ratio Tests; see 3.3.2) supplied to Vemco, about 46% were sold in Australia and 39% were sold overseas and the owners were identified. Approximately 15% of tag owners were not identified yet. Tag owners were invited by Vemco to contact IMOS Animal Tracking and register their tags and to access their tag detection data. Since April 2015, the first few Australian tag owners have contacted IMOS Animal Tracking to register their tags.

5.2 Identifying Registered Tags with no Detections

The IMOS Animal Tracking Database currently contains 2,393 (39.9%) registered tags with no associated detections recorded (Table 1). The following may be potential causes for registered tags without detections:

a) Tags were not released yet;
b) Detection files were not uploaded from arrays where those tags were released;
c) Tags were released but not detected yet.

5.3 Contacting Tag Owners of Incomplete Metadata

The IMOS Animal Tracking Database currently contains incomplete tag metadata of varying degrees, e.g. no tag release coordinates, no tag release date, no species information, etc. Of the 3,484 registered transmitters with more than one detection, 2,529 (72.6%) transmitters had partially missing metadata such as tag release date, location, and species information. Similarly, 646 (27.0%) of registered tags with no detection had no species information recorded in the tag metadata. The following may be potential causes for registered tags with incomplete metadata:

a) Tag not released yet (no animal metadata available);
b) Tag owner has not entered some or all of the data;
c) Never detected (released but never detected by receivers);
d) Detected (released and detected but detection data was not uploaded).

The Principal Investigators of associated projects were contacted to provide further information.

6. REFERENCES

Fitzgibbon, Q. P., Seymour, R. S., Ellis, D. and Buchanana, J., 2007. The energetic consequence of specific dynamic action in the southern Bluefin tuna *Thunnus maccoyii*. *The Journal of Experimental Biology* 210: 290-298.

Mancini, S., Tattersall, K., Proctor, R. and Galibert, G., 2012. IMOS NetCDF User's Manual: NetCDF Conventions and Reference Tables. Integrated Marine Observing System - eMarine Information Infrastructure, Hobart, Australia.

Ocean Tracking Network, 2015. OTN Data Portal – OTN Tool Box. Accessed 22/06/2015 at http://members.oceantrack.org/data/otn-tool-box.

Pincock, G., 2012. False Detections: What They Are and How to Remove Them from Detection Data. Vemco Whitepaper Document #DOC-004691 v03, Amirix Systems Inc., Halifax, Canada.

Sepulveda C. A., Kohin, S., Chan, C., Vetter, R. and Graham, J. B., 2004. Movement patterns, depth preferences, and stomach temperatures of free-swimming juvenile mako sharks, *Isurus oxyrinchus*, in the Southern California Bight. *Marine Biology* 145: 191-199.

VEMCO, 2015. VUE Software Manual. DOC-4399-23, available at http://Vemco.com/products/vue-software.