IMPUTING STOCK-OF-ORIGIN FOR ELECTRONIC TAGS USING STOCK-SPECIFIC MOVEMENT

Tom Carruthers[1]

SUMMARY

A method is described that uses stock-specific movement to impute stock-of-origin for electronic tags. A simple example of the approach was tested using cross-validation of electronic tags of known stock-of-origin. The imputation model achieved a failure rate of 3% and a success rate of 88%. In 9% of cases the imputation model did not classify stock-of-origin. Similar approaches may increase the number of tags available for multi-stock spatial population dynamics modelling.

KEYWORDS

Population modelling, fishery statistics, tagging

[1] IOF, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4. t.carruthers@oceans.ubc.ca

1 Introduction

Electronic tagging experiments have provided invaluable information regarding the ecology, migration and stock composition of Atlantic bluefin tuna (Block et al. 2005, Lutcavage et al. 2012, Cermeno et al. 2015). Electronic tagging data are a principal source of data for identifying plausible hypotheses for stock mixing and movement that may be included in Management Strategy Evaluation (MSE, Butterworth 1999) to identify robust management approaches (SCRS 2013). Electronic tagging data also provides additional information about the relative likelihood of encountering fish of each stock in a particular time and area, which is ultimately used by multi-stock population models to divide up observed catches by stock thereby informing the relative size of each stock (i.e. the unfished biomass of the Eastern stock spawning in the Mediterranean and the Western stock spawning in the Gulf of Mexico and Slope Sea).

A principal obstruction to the use of electronic tagging in multi-stock population dynamics models is that a large fraction of electronic tag tracks cannot be reliably assigned to a particular stock-of-origin (e.g. Carruthers et al. 2015). It may be possible assign stock-of-origin internally in the spatial population dynamics model by using the model estimated movement matrices. However the discrete assignment of stock-of-origin to tags using the relative likelihood of stock-of-origin, introduces instability in movement estimation. It may be preferable to assign stock-of-origin *apriori* using the movements of tags of known stock-of-origin. In this paper I describe a general approach and provide a simplistic proof of concept using cross validation of the coarsely aggregated data available to the bluefin tuna MSE process.

2 Methods

2.1 Theory

Let us assume that movement of animals does not vary substantially among years. If a Markov movement matrix *M* is available for all stocks *s* in a within-year time period *t* then the asymptotic fraction of fish inhabiting an area *k* can be found by recursively multiplying a vector *v* by the seasonal movement from area *r* to area *k*:

(1) $$v_{s,t,k} = \begin{cases} \sum_r v_{s,t-1,k} M_{s,t,r,k} & t > 1 \\ \sum_r v_{s,n_t,k} M_{s,t,r,k} & t = 1 \end{cases}$$

where n_t is the total number of within-year time steps (e.g. $n_t = 12$ for monthly model) and *v* is a vector representing stock fractions (i.e. $\sum_k v_{s,t,k} = 1$, and prior to iterative multiplication by the movement matrices, initial values for *v* can be set to $1/n_r$, the reciprocal of the number of areas).

In computer coding terms, this means looping over a number of iterative ‘years’ and subyears until *v* stabilizes in each within-year (typically *v* varies by less than 0.1% within 100 ‘years’). Once this iterative process has converged the fraction *N*, of each stock passing to and from areas can be calculated:

(2) $$N_{s,t,r,k} = \begin{cases} v_{s,t-1,k} M_{s,t,r,k} & t > 1 \\ v_{s,n_t,k} M_{s,t,r,k} & t = 1 \end{cases}$$

It follows that the probability of a tag transition *T*, of tag *i*, from area *r* to area *k* in subyear *t*, originating from a fish of stock *s* can be calculated by:

(3) $$P(T_{i,t,r,k} \in s|M,R) = P_{i,s} = \frac{R_s N_{s,t,r,k}}{\sum_s R_s N_{s,t,r,k}}$$

where *R* is the relative magnitude of each stock.

In instances where a tag *i* can be observed over many time steps (there are n_o track observations), the total probability *L* of that tag belonging to stock-of-origin *s* can be calculated as the product of the *P* terms:

(4) $$L_{i,s} = \prod_o P_{i,s,o}$$

The ratio of the probabilities for each stock may be used to impute stock-of-origin for tag tracks that do not have this covariate data.

2.2 Simplistic proof of concept using cross-validation

The aggregated electronic tagging data (provided by M. Lauretta) used in the bluefin tuna MSE process are aggregated to quarter (January-March, April-June, July-September, October-December) and by 10 areas including areas for the natal spawning grounds of Eastern and Western fish in the Mediterranean and Gulf of Mexico, respectively (Figure 1). It was assumed any tags that entered the Mediterranean were Eastern fish and any that entered the Gulf of Mexico were Western fish. This was the base dataset of known stock-of-origin and it constituted 222 tags and 459 seasonal transitions.

'Leave one out' cross-validation was used to test the theory laid out in section 2.1. Each tag of known stock-of-origin was removed from the base dataset one at a time. For each tag, only those transitions that did not include visitation to a known natal spawning ground could be used; natal grounds were used to assign stock-of-origin in the first place and consequently stock assignment would be 100% correct if these transitions were included. Of the 222 tags of known stock-of-origin just 67 had one or more transitions that did not include the natal spawning grounds and therefore just 67 tags could be used in the cross-validation.

For each of the 67 tags that were removed, an empirical seasonal movement matrix was calculated for each stock using the remaining tag tracks (the remaining 221 tags and approximately 430 quarterly transitions). Using equations 1-4 above and an assumed relative stock size R of 6 for the East and 1 for the West (the East assumed to be approximately 6 times larger), the relative likelihood $L_{i,s}$ of the removed tag *i* belonging to a given stock *s* was calculated.

For each tag removed in the leave-one-out cross validation, the probability ratio in favour of stock 1 (the eastern stock) was calculated: $LR_{\mathrm{i}} = L_{\mathrm{i},1}/L_{\mathrm{i},2}$. Arbitrarily, when the probability ratio exceeded 2 the origin of the tag was imputed as stock 1, the Eastern stock. When the probability ratio was less than 0.5, the origin of the tag was imputed as stock 2, the Western stock.

3 Results

The imputation model incorrectly classified stock-of-origin in 3% of cases (2 tags), correctly classified stock-of-origin in 88% of cases (59 tags) and failed to classify stock-of-origin in 9% of cases (6 tags)(Table 1).

4 Discussion

The results of this simplistic cross validation test appear encouraging. This is particularly the case as: (1) the tag track data in the example were very coarsely aggregated by quarter and 10 large ocean areas, (2) the algorithm for imputing stock-of-origin was a first attempt and not conditioned or adjusted in any way to improve the success rate of stock classification and (3) the tag tracks were generally quite short as they removed the majority of transitions that involved natal spawning areas.

(1) It would be highly desirable to use much finer resolution data derived empirically from the geolocation error models used to assign tracks to tags. Others (e.g. Arregui et al. 2015) have derived defensible movement matrices that might provide a better basis for imputation.

(2) The cross validation shows a tendency to mis-assign eastern fish with a western stock-of-origin, which may be counteracted by adjusting the relative stock sizes, *R*. A more sophisticated alternative would be to undertake the imputation using the estimated stock sizes from the stock assessment models thereby using a time-varying relative stock size, *R* for each stock. The imputation model appears to have a higher propensity to classify unknown origin (likelihood ratios *LR* between 0.5 and 2) than incorrectly classify stock-of-origin. Depending on priorities, the ratio of error to non-classification rates (perhaps not assigning stock-of-origin is more acceptable than erroneously assigning stock-of-origin) could be adjusted by altering the bounds of the probability ratios used to assign stock-of-origin.

(3) It would be desirable to test the imputation algorithm using electronic tag tracks that have stock-of-origin assigned by covariate data (e.g. otolith microchemistry). These data may include tags of with a greater number of transition observations (longer tag track) since at least one transition to a known natal spawning area does not have to be removed for each tag for the purposes of cross validation.

There are however various reasons why the results of the cross-validation above may overstate the stock classification success rate of the imputation algorithm. To assume similar classification success rates for other tags of unknown stock-of-origin is to assume that tags which include a transition to a natal spawning area (used in the cross-validation) are representative of all tags (for which stock of origin is not known). This is a dubious assumption. For one, the tags used in cross-validation include seasonal movements among areas adjacent to natal spawning areas, a behaviour likely to be highly indicative of stock-of-origin. Secondly, fish moving to and from spawning areas are likely to be mature fish whose behaviour is not representative of the foraging behaviour of intermediate age classes (oceanic juvenile fish mixing in the Atlantic for example).

Another limitation of the proposed approach is that the use of discrete movement matrices throws away a great deal of fine-scale spatio-temporal data. It may be worth investigating multivariate clustering techniques tailored to the fine-scale (e.g. daily by longitude and latitude) tag track data of fish of known stock-of-origin. Alternatively the method could be used probabilistically in conjunction with other assignment methodologies such as otolith microchemistry and otolith shape analysis.

Regardless of the limitations, the simple approach here provides a promising basis for unlocking a large amount of electronic tagging data for use in multi-stock modelling and the exploration of alternative hypotheses for stock structure and mixing (Arrizabalaga et al. 2014).

5 Acknowledgements

This work was carried out by TC under the provision of the ICCAT Atlantic Wide Research Programme for Bluefin Tuna (GBYP), funded by the European Union, several ICCAT CPCs, the ICCAT Secretariat and by other entities (see: http://www.iccat.int/GBYP/en/Budget.htm). The contents of this paper do not necessarily reflect the point of view of ICCAT or other funders and in no ways anticipate ICCAT future policy in this area.

6 References

Arrizabalaga, H., Fromentin, J.M., Kell, L.T., Kerr, L.A., Kimoto, A., Porch, C., Secor, D.H. 2014. Population hypotheses and stock assumptions for North Atlantic Bluefin tuna. SCRS/2014/xxx.

Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985-998.

Carruthers, T.R., Kimoto, A., Powers, J., Kell, L., Butterworth, D., Lauretta, M. and Kitakado, T. 2015. Structure and estimation framework for Atlantic bluefin tuna operating models. ICCAT SCRS/2015/179.

Cermeno, P., Quilez-Badia, G., Ospina-Alvarez, A., Sainz-Trapaga, S., Boustany, A., Seitz, A.C., Tudela, S., Block, B.A. 2015. Electronic tagging of Atlantic bluefin tuna (*Thunnus thynnus,* L.) reveals habitat use and behaviors in the Mediterranean Sea. PLoS ONE 10 (2): e0116638. doi: 0.1371/journal.pone.0116638.

Arregui I., Galuardi B., Goñi N., Arrizabalaga H., Lam C.H., Fraile I., Santiago J. and Lutcavage M. 2015.Movements and geographic distribution of juvenile bluefin tunas in the North Atlantic, described through electronic tags. SCRS/2015/044.

Lutcavage, M.E., Galuardi, B., Lam, T.C.H. 2012. Predicting potential Atlantic spawning grounds of western Atlantic bluefin tuna based on electronic tagging results, 2002-2011. Col. Vol. Sci. Pap. ICCAT. 69(2): 955-961.

Block, B., A., Teo, S.L.H., Walli, A., Boustany, A., Stokesbury, M.J.W., Farwell, C.J., Weng, K.C., Heidi, D., Williams, T.D. 2005. Electronic tagging and population structure of Atlantic bluefin tuna. Nature. 434: 1121-1127.

Table 1. Results of the 'leave one out' cross validation

		Imputed stock		
		East	West	Not imputed
Known stock	East	12	2	0
	West	0	47	6

n =	67
Success rate:	88.1%
Non assigment rate:	9.0%
Failure rate:	3.0%

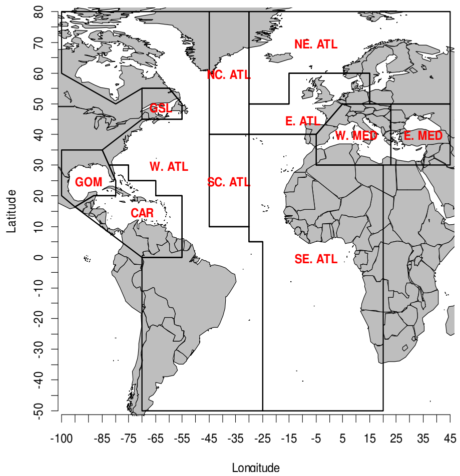

Figure 1. The 11-area spatial definitions of latest electronic tagging (Lauretta. pers. comm., right). In this analysis the two Mediterranean areas were combined into a single area, creating a total of 10 ocean areas.