Biomathematics and Statistics Scotland

Cross-Validation of tern *Sterna sp.* tracking data modelling (Phase 1) (Under Agreement C10-0206-0387)

CONTRACT No: C10-0206-0387

Report submitted to:

Joint Nature Conservation Committee

December 2013

Revised: January 2014

BioSS

Authors:

Jackie M Potts, Mark J Brewer

CONTENTS PAGE

In addition to this report, there are ancillary files associated with this project:

(i) R code for cross-validation

1. Introduction

The Joint Nature Conservation Committee (JNCC) is currently working with the four Statutory Nature Conservation Bodies (SNCBs) to identify important marine areas around the UK that are used by terns *Sterna sp.* during the breeding season. This is to inform the identification of areas that may be suitable for designation as marine Special Protection Areas (SPAs) under the EC Birds Directive (2009/147/EC).

Previous work undertaken developed a weighted logistic regression modelling approach to analyse JNCC tern tracking data for the four larger species of terns (Arctic, common, Sandwich and roseate terns). This was developed in two phases: Phase 1 developed individual site-specific models (Brewer *et al*. 2012a) while Phase 2 developed generic models which, in theory, could be applied to any colony which had the requisite covariate data (Brewer *et al*. 2012b). Subsequent to this, some refinements were made in relation to the candidate covariate data set for the Phase 1 models (Brewer *et al*. 2012c; Potts *et al*. 2013a) which addressed some issues previously identified in the earlier work and final models which will be used to underpin any boundary delineation of any possible SPAs were then chosen (Potts *et al*. 2013b).

This report describes cross-validation work carried out on the final Phase 1 models to assess their predictive ability.

2. Methods

Models were fitted using the methodology described in earlier reports for Phase 1 (Brewer *et al*., 2012a, Potts *et al*., 2012). Two approaches to cross-validation were taken. First, we assessed the predictive ability of the final models by using a sub-set of data to fit the model and testing how well it predicted the remaining data. Second, we assessed the how well data from one year were able to predict another, using sub-sets of the data to select a model as well as to fit it.

The first analysis was carried out on colony/species combinations with at least 50 tracks in total, using the final models chosen in Potts et al. (2013b) as shown in Table 1

Table 1, Colonies and species with at least 50 tracks and the terms chosen in the final model

Colony	**Species**	**Terms in final model**
Coquet	Arctic	distance to colony, bathymetry, June chlorophyll
Coquet	Common	distance to colony, bathymetry, June chlorophyll, April SST
Coquet	Sandwich	distance to colony, distance to shore
Leith	Common	distance to colony, distance to shore, bathymetry, spring salinity, slope
North Norfolk	Sandwich	distance to colony, distance to shore, bathymetry, shear stress wave
Cemlyn	Sandwich	distance to colony, distance to shore, spring salinity
Forvie	Sandwich	bathymetry, temperature stratification

For each colony and species we formed 100 bootstrap samples for each of 10, 20 or 30 tracks in the training set. In each case we held back 30 tracks sampled randomly without replacement for use as a test set and then sampled 10, 20, or 30 tracks randomly with replacement from the training set, together with the corresponding control tracks. Three performance scores were used (the likelihood score, the MSE score and the AUC) as described in Potts et al. (2013c).

Initially we considered holding back just 10 tracks for use as a test set, but we found that the performance measures were then considerably more variable between test sets than they were between training sets.

For the second analysis, we used colonies for which multiple years of data were available with more than 10 tracks in at least one year. For each year with more than 10 tracks we found a minimum AIC model by stepwise selection and fitted the model to the other years.

3. Results and Discussion

The average performance scores for the bootstrap samples are shown in Table 2.

Table 2. Average performance scores for bootstrap sample sizes of 10, 20 or 30 tracks.

Colony	Species	Bootstrap Sample Size	Likelihood Score	MSE Score	AUC
Coquet	Arctic	10	-0.225	0.056	0.791
		20	-0.213	0.055	0.801
		30	-0.213	0.056	0.795
Coquet	Common	10	-0.232	0.059	0.838
		20	-0.197	0.056	0.848
		30	-0.193	0.056	0.849
Coquet	Sandwich	10	-0.213	0.059	0.915
		20	-0.192	0.059	0.913
		30	-0.193	0.059	0.917
Leith	Common	10	-0.305	0.086	0.734
		20	-0.294	0.084	0.744
		30	-0.291	0.084	0.744
North Norfolk	Sandwich	10	-0.215	0.053	0.883
		20	-0.199	0.052	0.886
		30	-0.201	0.053	0.884
Cemlyn	Sandwich	10	-0.205	0.055	0.934
		20	-0.193	0.053	0.940
		30	-0.176	0.051	0.943
Forvie	Sandwich	10	-0.104	0.030	0.989
		20	-0.085	0.027	0.990
		30	-0.082	0.026	0.991

All of the AUC scores exceeded 0.7, which indicates good performance. There was only a slight improvement in the scores with increasing sample size. However, this was using a model containing covariates that had already been selected on the basis of the full sample of tracks. Poorer performance might be expected with smaller sample sizes if these samples had been used to select (as well as fit) the model. The results from the inter-annual cross-validation are shown in Table 3.

Table 3. Performance of model fitted to one year's data when tested on data for the same species and colony from other years.

Colony	Species	Training Year	Test Years	Minimum AIC Model	Likelihood Score	MSE Score	AUC
Coquet	Arctic	2009	2010, 2011	dist_col, chl_june, sst_april, sal_spring, ss_wave, bathy_1sec	-0.238	0.058	0.760
		2010	2009, 2011	dist_col, sal_spring	-0.197	0.047	0.741
		2011	2009, 2010	dist_col, dist_shore, sst_june, sal_spring, ss_wave	-0.307	0.065	0.760
Coquet	Common	2009	2010, 2011	dist_col, chl_june, bathy_1sec	-0.173	0.051	0.860
		2010	2009, 2011	summ_front, strat_temp, sal_summ, sal_spring, ss_current	-0.240	0.064	0.677
		2011	2009, 2010	dist_col, sst_april, sst_june, summ_front, ss_wave	-0.268	0.067	0.783
Coquet	Sandwich	2009	2010, 2011	dist_col, dist_shore, chl_apr, chl_may, spring_front, sal_summ, sal_spring	-0.404	0.047	0.871

		2010	2009, 2011	dist_col, chl_apr, chl_june, sal_summ, ss_current	-0.143	0.042	0.880
		2011	2009, 2010	dist_col, chl_june, sst_june, ss_wave, ss_current	-0.180	0.048	0.869
Leith	Common	2009	2010	dist_col, dist_shore, chl_apr, sst_april, spring_front, ss_wave. bathy_1sec	-0.327	0.075	0.590
		2010	2009	dist_col, sst_april, sst_june, spring_front	-0.190	0.053	0.805
Larne Lough	Common	2009	2010, 2011	dist_col, ss_wave, bathy_1sec	-0.339	0.034	0.591
		2011	2009, 2010	dist_col, dist_shore, sst_june	-0.243	0.060	0.897
Larne Lough	Sandwich	2011	2009, 2010	dist_col, dist_shore, sal_spring	-0.109	0.007	0.998
Outer Ards	Arctic	2011	2009, 2010	dist_col, dist_shore, chl_apr, chl_may, chl_june, sal_summ, sal_spring, bathy_1sec	-0.484	0.096	0.627
North Norfolk	Sandwich	2006	2007, 2008	dist_col, dist_shore, chl_apr, chl_june, sst_april, bathy_1sec	-0.160	0.041	0.822
		2007	2006, 2008	dist_col, sst_april, bathy_1sec	-0.187	0.052	0.853

		2008	2006, 2007	dist_col, dist_shore, bathy_1sec	-0.178	0.044	0.857

The majority of the AUC scores exceeded 0.7, indicating good performance, but some were only around 0.6.

References

Brewer M.J., Potts J.M., Duff E.I. & Elston D.A. (2012a). To carry out tern modelling under the Framework Agreement C10-0206-0387. Contract Report to JNCC, March 2012.

Brewer, M.J., Potts, J. M., Duff, E.I. & Elston, D. A. (2012b). *Prediction of new colonies – seabird tracking data (under Agreement C10-0206-0387)*. Contract Report to JNCC, November 2012.

Potts J.M., Brewer M.J., Duff E.I. & Elston D.A. (2012). Additional Work Coquet Colony - Seabird Tracking Data (Under Agreement C10-0206-0387). Contract Report to JNCC, December 2012.

Potts, J. M. Brewer, M.J., Duff, E. I. & Elston, D. A. (2013a). *Additional Work Other Colonies – Seabird Tracking Data (Under Agreement C10-0206-0387).* Contract Report to JNCC, March 2013

Potts, J. M. Brewer, M.J. & Duff, E. I. (2013a). *Refinements of tern Sterna sp. tracking data modelling (Phase 1) – Seabird Tracking Data (Under Agreement C10-0206-0387).* Contract Report to JNCC, September 2013

Potts, J. M. Brewer, M.J. & Duff, E. I. (2013b). *Refinements of tern Sterna sp. tracking data modelling (Phase 1) – Seabird Tracking Data (Under Agreement C10-0206-0387).* Contract Report to JNCC, September 2013

Potts, J. M. Brewer, M.J. & Duff, E. I. (2013c). *Refinements of tern Sterna sp. tracking data modelling (Phase 1) – Seabird Tracking Data (Under Agreement C10-0206-0387).* Contract Report to JNCC, October 2013