

Wildlife Management Institute, Inc.
North Atlantic LLC
Grant 2011-2012
Final Report

Richard R. Veit
Biology Department
CSI/CUNY 2800 Victory Boulevard
Staten Island, NY 10314

Abstract: We scanned the entire 270,000 record USGS database on marine birds to search for inconsistencies in the use of four-letter species codes for birds. On the whole, we found the database to be clean, and species codes were generally consistent. There were some exceptions. Potentially the most influential confusion was in the use of codes for Roseate and Royal Terns. Unambiguous and correct codes for these species are “ROST” and “ROYT” respectively, but it is quite certain that the ambiguous code “ROTE” was used for both of these species on the order of a few hundred times in the past. We have made corrections or suggestions within a separate column in the database about how to interpret these ambiguous codes. Another general observation based on our scan of the database is the interpretation of large numbers of birds unidentified to species, for example “UNTE” for unidentified tern, “UNAL” for unidentified alcid or “UNSH” for unidentified shearwater. It is possible, for example, that the majority of Roseate Terns observed during the entire 35 year period covered by the database were entered originally in the field as “UNTE”, because the Roseates were in mixed flocks with Common Terns and perhaps some Arctics. In one sense there is no way we can know how many of these were Roseates, but there are many ways to estimate this quantity based on other data sources. We recommend this be done so that distributional models accurately reflect the entire content of the data on these birds collected at sea.

Introduction/Background

Standardized surveys of seabirds off the east coast of the United States began in earnest during the 1970s and have continued, with increasing sophistication, to the present (Powers 1983, Powers and Brown 1987, Nisbet et al. 2013). A recent surge in interest in these data derives from exploration for offshore wind “farms” and other possible offshore development .

Data on birds collected from ships and aircraft were originally recorded by hand on clipboards and subsequently entered into computer databases (Powers 1983). More recent efforts use a variety of data entry programs run on dataloggers with built in GPS recorders. All recording systems have used codes for species names, roughly similar to those used by the USGS/USFWS Bird Banding lab. These codes are made of four letters, often the first two letters of the species name and the first two letters of the family or other group, e.g. "SOSH" for SOoty SHearwater. For the most part these codes provide unambiguous separation, but there are some sources of confusion, for example both Royal Terns and Roseate Tern could be (and have been) recorded as "ROTE".

During our work with the modeling team, we noticed some surprising results in some of the model outputs, and wondered whether any of these unusual results could have been explained by errors or unambiguities in the species coding within the database. Perhaps the most important such potential error involved that for Roseate Tern, since it is the only federally endangered species occurring within the area covered by the database. But other ambiguities may have arisen with other species and species groups, so our rationale for perusing the database was to check for such errors.

Methods

We scanned through the entire database manually, looking for mistaken species codes within the species fields. Such would jump out fairly clearly as the database is sorted alphabetically. We selected what we thought would be the most likely sources of confusion to begin with – the terns mentioned above, Razorbills (for which we expected confusion in the pre-1990 data, which we did not find) and some unidentified groups ("UNTE", "UNSH", "UNAL"). For the unidentified groups, we did not make a suggestion in the database, because interpretation will have to be done in collaboration with modelers later, but it is our plan to consult on this issue. For other ambiguities, especially the Royal/Roseate Tern pair, we have made suggested changes with the column allocated for this purpose and returned the annotated database to Andrew Gilbert, Mark Wimer and Allison Sussman.

Results

We reviewed the entire database and made suggested changes in a file sent to USGS personnel. We feel that the database is "clean" and free of errors. We recommend that all analyses be

checked with knowledgeable seabird ecologist for inclusion of unidentified birds and for checking of potential “hotspots” that appear in modeled data.

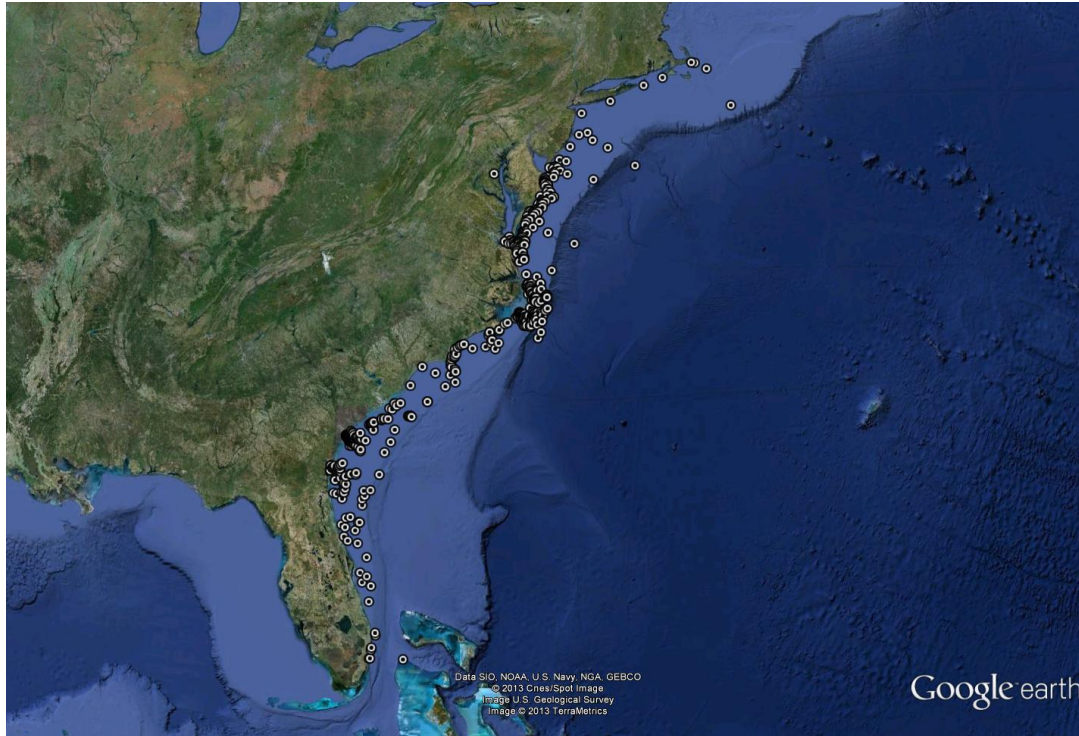


Figure 1.) All Royal Tern records in database.

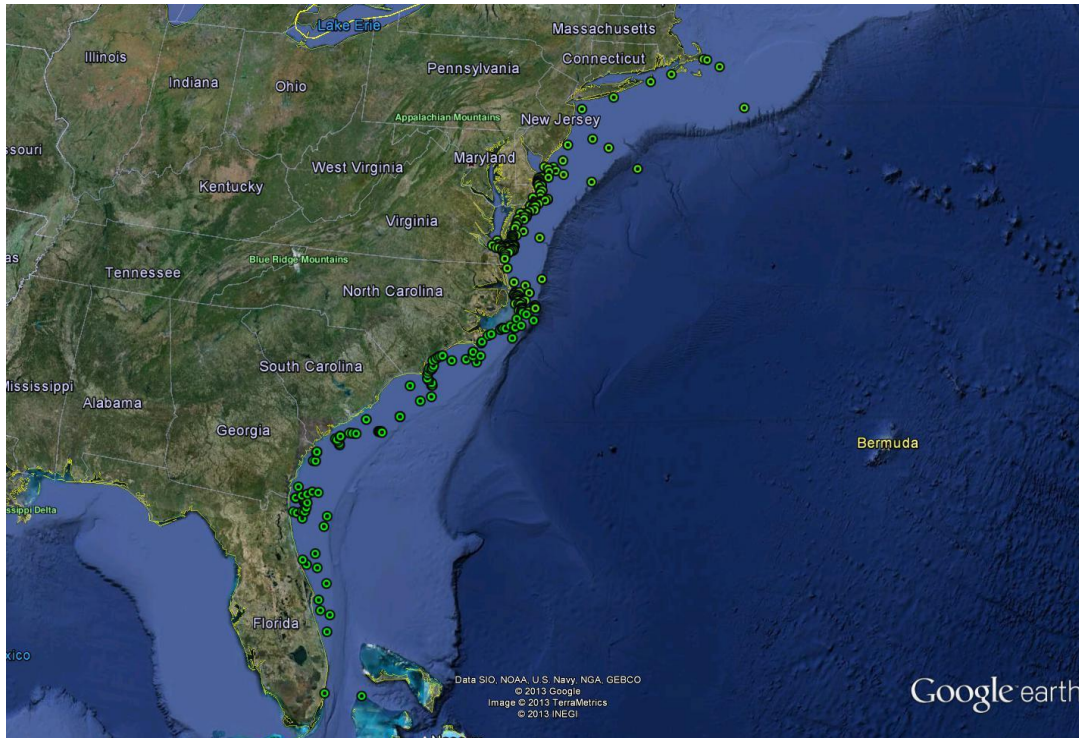


Figure 2.) Records originally coded as ROTE and designated as Royal Tern (248 records).

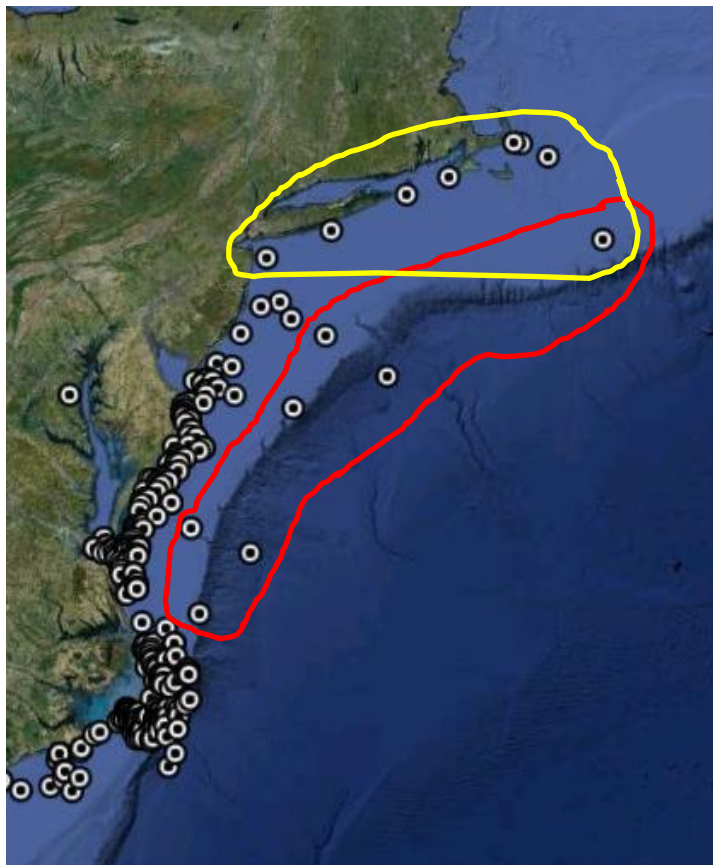


Figure 3.) Suspect Royal Tern records flagged in database as “way offshore” or “ North of 40.” These have been listed by us as “UNTE” in the database; models could be run with that designation, or with birds listed as either Roseate or Royal Terns.

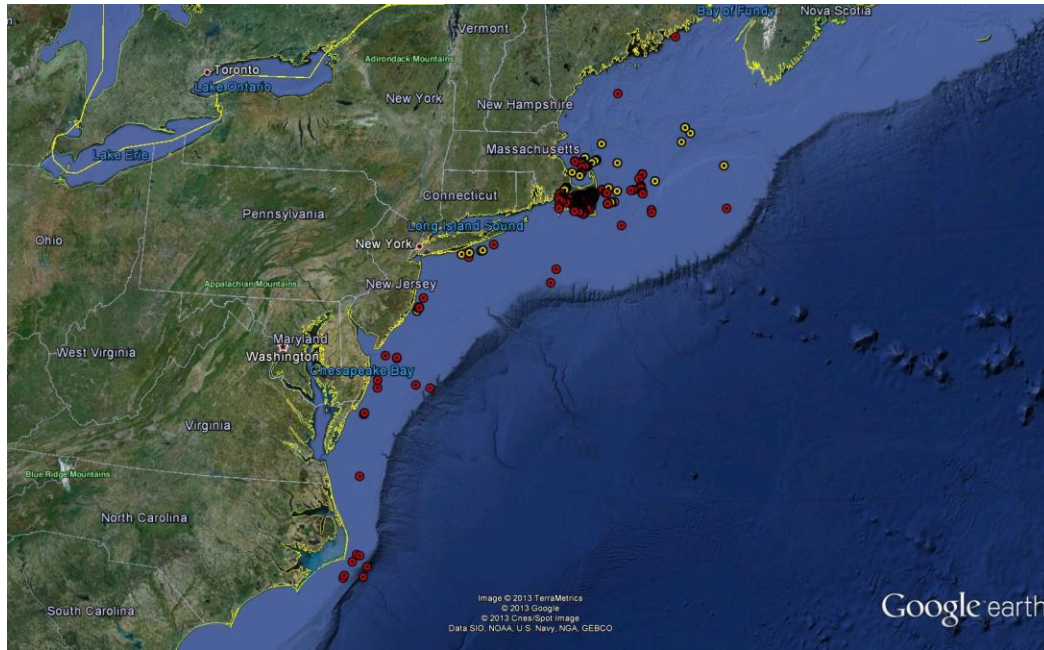


Figure 4. Roseate Tern records showing those originally coded as ROTE in yellow (56), others in red. Southernmost birds (off Hatteras) are in fact Roseates; but they are rare that far south and inshore and inclusion of those records may bias modeling output.

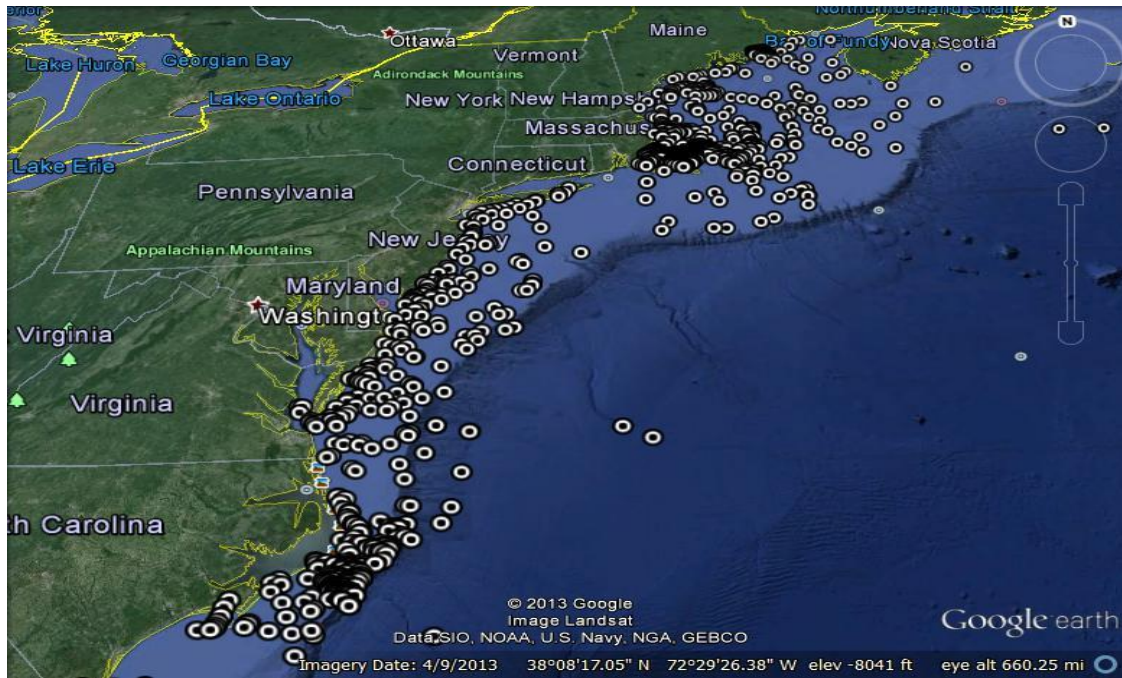


Figure 5.) Records of unidentified terns in database. The point here is to show the wide distribution, and to suggest that a large proportion of (endangered) Roseates Terns may be “hidden” here.

Discussion

The USGS database is a remarkable achievement, especially considering the disparate sources of information contained in it. There are still some issues of interpretation of the data, but we believe there are no further mistakes within the four letter species codes.

We recommend that models of abundance make use of the birds listed as “unidentified” to species, with perhaps 2-3 different versions of output, with, respectively, all unidentifieds included, all excluded, and some fraction included. The fraction to include can be determined through examination of other sources of data (e.g. Nisbet et al. 2013, state bird books, the journal *North American Birds*).

Acknowledgements

I thank Deborah Jaques for her work on the review of the database and the production of figures.

References

Powers, K.D. 1983. Pelagic distributions of marine birds off the Northeastern United States. NOAA Technical Memorandum NMFS-F/NEC-27. Woods Hole: US Department of Commerce, National Oceanographic and Atmospheric Administration. 201 pp.

Powers, K.D. and R.G. B. Brown. 1987. Seabirds. Pages 372-384 in Georges Bank (Backus, R.H. and D.W. Bourne, eds.) Cambridge, MA: MIT Press. 593 pp.

Nisbet, I.C. T., R.R. Veit, S.A. Auer and T.P. White. 2013. Marine Birds of the Eastern United States and the Bay of Fundy: Distribution, numbers, trends, threats and management. Nuttall Ornithological Monographs No. 29. 188 pp.