Fisheries Research 208 (2018) 167–178

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Evaluation of GLM and GAM for estimating population indices from fishery independent surveys

Stephen E. Potts[a,*], Kenneth A. Rose[b]

[a] Department of Biological Science, Louisiana State University, Baton Rouge, LA 70803 USA
[b] University of Maryland Center for Environmental Science, Horn Point Laboratory, PO Box 775, Cambridge, MD 21613 USA

ARTICLE INFO

Handled by A.E. Punt.

Keywords:
Stock assessment
Spatial heterogeneity
Fishery independent surveys
Generalized linear and additive models
Indices of abundance
Parsimony

ABSTRACT

We evaluate the performance of generalized linear (GLM) and generalized additive (GAM) models for deriving population indices from fishery independent survey data. Six model types (3 GLMs and 3 GAMs) were formulated that differed in how spatial covariates were represented, with each using one of three alternative ways to include temporal covariates. The models were applied to summer and fall survey data on 127 species from fisheries-independent bottom-trawl surveys conducted by the Southeast Monitoring and Assessment Program (SEAMAP) in the northwest Gulf of Mexico. Three response variables were analyzed: occurrence, density, and abundance. The best model (from the alternative temporal representations) were identified for each response variable for each of the six model types and their performance analyzed in more detail. Model performance was based on residual autocorrelation (Moran's I), prediction error (AIC weights), and predication variance (based on simulated sampling). We examined for patterns in these metrics based on the magnitude of the response variables (i.e., by quartiles). Results suggest that sample size (indexed here by response value quartiles) could be useful for *a priori* consideration when choosing among GLM and GAM models. GAM models that use geoposition with smoothing as the spatial covariate performed comparable to some of the other models at low abundances and densities (lower quartiles), and significantly outperformed all of the other models at higher densities and abundances (quartiles 3 and 4). We discuss how our results provide guidance on selecting GLM and GAM models for deriving population indices from survey data.

1. Introduction

A major challenge for species abundance modeling is to account for the high complexity inherent in the survey data in terms of spatial and temporal variation (Cao et al., 2017; Orio et al., 2017; Thorson and Barnett, 2017). Deriving indices of abundance that reflect the spatial distribution of a species and its dynamics over time is fundamental to the study of population and community ecology and plays a critical role in management and conservation (Beale and Lennon, 2012; Berger et al., 2017). For example, abundance indices are used in fisheries stock assessment and management as the basis for documenting fish population trends (Maunder and Punt, 2004), for the fitting of population dynamics models (Campbell, 2014), and for other management-related analysis, such as indices to tune multi-species and food web models (Storch et al., 2017) and management strategy evaluation (Ono et al., 2017).

The selection of the response variable to represent the population index and how the statistical model applied to the survey data to estimate the index is formulated can strongly influence the resulting inferences about the temporal and spatial dynamics of the population (Bučas et al., 2013). Generalized linear models (GLMs; Nelder and Wedderburn, 1972) and generalized additive models (GAMs; Hastie and Tibshirani, 1986) are two statistical approaches commonly applied to survey and monitoring data to generate population indices. GLMs that stratify by depth and by longshore gradients are very commonly used as part of marine fisheries stock assessment (Hart, 2012; Hoyle et al., 2014b; SEDAR31, 2013). A typical example is the use of GLMs with a polynomial representation of depth effects, in combination with stratified regional zones, to describe the population trends of skates and sharks in the North Sea (Sguotti et al., 2016).

While GLMs that fit spatially-stratified factor models are commonly reported (Brodziak and Walsh, 2013; Campbell, 2015; Lynch et al., 2012; Tascheri et al., 2010), there has also been analyses (albeit difficult to generalize) that demonstrate the efficacy of using GAMs (Berg et al., 2014; Bigelow et al., 1999; Minami et al., 2007; Walsh and Kleiber, 2001). GAMs have been shown to fit inherently nonlinear

* Corresponding author.
E-mail addresses: spotts2@lsu.edu (S.E. Potts), krose@umces.edu (K.A. Rose).

https://doi.org/10.1016/j.fishres.2018.07.016
Received 30 April 2018; Received in revised form 24 July 2018; Accepted 25 July 2018
0165-7836/ Published by Elsevier B.V.

relationships with lower prediction error than GLMs (Austin et al., 2006; Li et al., 2011; Moisen and Frescino, 2002; Murase et al., 2009). GAMs often also have less residual autocorrelation than GLMs, particularly when the model itself includes covariates designed to quantify spatial correlation effects; in such cases, GAMs can have greater predictive power (Polansky et al., 2018; Segurado and Araujo, 2004; Segurado et al., 2006). In particular, GAMs specifying 2-dimensional (2D) isotropic functions of geoposition, rather than product interactions of latitude and longitude may be especially useful for dealing with autocorrelation (Augustin et al., 2013; Wood and Augustin, 2002). For fisheries-dependent data, bias can be introduced to abundance estimates when spatial aggregation is too course such that sampling is no longer random (Campbell, 2004; Carruthers et al., 2010), and fine-scale resolution may be much better able to capture underlying dynamics (Carruthers et al., 2011). For fisheries-independent data, however, it remains unclear to what extent resolution or specifying similar covariates in GLMs affects how well autocorrelation is represented.

Survey data generally have a high prevalence of zeroes and skewed distributions. Such data can be problematic for fitting typical distributions (Shono, 2008; Zuur et al., 2012). Zero-inflated Poisson (Lambert, 1992) and negative binomial (Welsh et al., 1996) are common in ecological research; they are used to model population dynamics (Lyashevska et al., 2016; Zipkin et al., 2017) and species distributions for common and rare species (Cunningham and Lindenmayer, 2005; Welsh et al., 1996). In fisheries, zero-inflated negative binomials have been applied to bycatch estimation (Kuhnert et al., 2011; Minami et al., 2007). Alternatively, delta-distributions, though less common, are used in spatial models for populations and for community ecology (Arcuti et al., 2014; Lynch et al., 2014). Delta-GLMs and delta-GAMs in particular are two-part models that have been applied to fisheries indices of abundance. Delta-GLMs (Stefánsson, 1996), especially delta-lognormal distributions (Walter and Ortiz, 2012), are the most commonly used models in fisheries stock assessment abundance standardization (Hoyle et al., 2014a), though various forms of delta-GAMs (Swartzman et al., 1992) have also been proposed (Arcuti et al., 2014; Bacheler and Ballenger, 2018; Li et al., 2011; Orio et al., 2017). Systematic comparisons of GLMs and GAMs when used to derive population indices from survey data are rare (though see Venables and Dichmont, 2004; Yu et al., 2013).

The purpose of this analysis is to explore whether criteria *a priori* to model selection can be applied to better formulate statistical models used to generate population indices. Few direct comparisons among alternative models are available to allow for general guidance on how one should select which approach (GLM or GAM) to use. We evaluate GLMs and GAMs based their application to the same survey data on their: 1) residual autocorrelation, 2) prediction error, and 3) prediction variance. Both GLMs and GAMs have been extensively described (Hastie and Tibshirani, 1990; McCullagh and Nelder, 1989; Wood, 2017); our focus is on the performance of GLMs and GAMs when both are applied in a consistent manner to the same widely used fisheries survey dataset and whether the results lead to general guidelines in how the two approaches should be used in different situations. Model choices, such GLM or GAM, are often made *a priori* to analysis; yet, explicit criteria for *a priori* selection are often lacking and clear evidence which modeling approach should be used is equivocal and often undocumented. Regardless of choice, abundance models should reduce spatial correlation, account for important temporal patterns, and reduce survey bias.

Consistent estimators of performance are needed to quantify differences among modeling approaches and among different formulations (e.g., covariates) within an approach. Difficulties in comparing models often arise due to lack of sufficient quantity and quality of data to allow for rigorous comparisons and because the subsets of the many possible evaluation criteria vary among studies (Rocchini et al., 2011). Residual autocorrelation, prediction error, and prediction variance, as used here, provided three contrasting criteria of model performance. Model residuals are compared between samples and weighted by proximity; spatially significant clusters compared to a null sampling distribution denote autocorrelation. Relative prediction error is estimated using Information-Theoretic methods, where increasing prediction error denotes decreasing information contained in model likelihood functions (Arndt, 2012). Models should fit data well, but still be flexible enough to make reliable predictions from new data; therefore, we compare annual abundance indices from simulated replicate surveys to field survey data to compare prediction variation.

We examine the performance of GLMs and GAMs for numerous species and three commonly used response variables (occurrence, abundance, density). Our evaluation of model performance is based on a common set of criteria applied across modeling approaches and across alternative models with each approach. We then attempt to relate performance to the response sizes (e.g., low versus high average abundance indices) to derive information to guide decisions in model selection. Our test dataset was the summer and fall survey data on 127 species from fisheries-independent bottom-trawl surveys conducted by the Southeast Monitoring and Assessment Program (SEAMAP) in the northwest Gulf of Mexico. The large number of species ensures that models are compared across a wide range of abundances and densities for species that exhibit a diverse suite of spatial and seasonal patterns. Only spatio-temporal covariates (e.g., depth, latitude) are considered, as these are commonly used to derive population indices. Further analyses could investigate environmental variables (e.g., temperature) as spatial and temporal covariates. For each species, trawl-collected data were fit to binomial, lognormal, and negative binomial response distributions to generate response variables of occurrence, density, and abundance. We then fit a progression of GLM and GAM models with spatial covariates (depth, latitude-longitude, location), and for each spatial version, we select the best among three different ways to represent time (year, season, day). The performance of the best models for each species is then quantified using their autocorrelation, prediction error, and prediction variance. These results are accumulated over species to allow for comparisons of overall performance among GLMs and GAMs for the three response variables. We conclude with how our results lead to several recommendations on model selection for the surveys that resemble our SEAMAP test dataset.

2. Methods

We outline the step-by-step flow of methods for model fitting (Fig. 2) to response variables (Fig. 3), and for model variance estimation (Fig. 4). Fig. 2 outlines the steps for fitting and comparing the 3 GLMs and 3 GAMs at differing latitudinal and longitudinal spatial resolutions using observed survey covariate values. Fig. 4 outlines the steps used to simulate new covariate values for input into the 6 models to generate new model estimates that are then used to quantify prediction variance.

2.1. SEAMAP data

We used the bottom-trawl surveys for the Gulf of Mexico as collected by the Southeast Monitoring and Assessment Program (SEAMAP; http://www.gsmfc.org). SEAMAP is a long-term scientific survey conducted in the summer and fall (Fig. 1). Surveys were based on a stratified random design, where transects perpendicular to depth gradients were randomly selected within NOAA statistical zones (Fig. 1A). Trawls were then taken at approximately 5 fm isobath intervals on each transect.

We used a subset of the SEAMAP data in this analysis that we screened for sampling consistency (A in Fig. 2). We only included trawls that met the following criteria: summer and fall surveys during 1985 to 2013; depth of trawls was less 110 fm; statistical zones 10 to 21; only samples collected with a 12.8 m otter trawl fitted with a 1.63 cm mesh; trawl distances limited to 0.5 to 10 km; and where meta-data comments