

2010

FINDING THE SMOOTHEST PATH TO
SUCCESS: MODEL COMPLEXITY AND THE
CONSIDERATION OF NONLINEAR
PATTERNS IN NEST-SURVIVAL DATA

Encontrando el Camino Mas Facil Hacia el Exito:
Complejidad de los Modelos y Consideración de
Patrones No Lineales en Datos de Supervivencia de
Nidos

Max Post van der Burg

University of Nebraska-Lincoln, maxpostvanderburg@usgs.gov

Larkin A. Powell

University of Nebraska-Lincoln, lpowell3@unl.edu

Andrew J. Tyre

University of Nebraska at Lincoln, atyre2@unl.edu

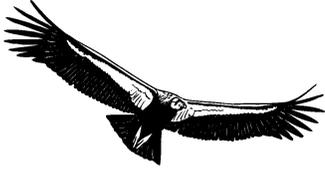
Follow this and additional works at: <http://digitalcommons.unl.edu/natrespapers>



Part of the [Natural Resources and Conservation Commons](#), [Natural Resources Management and Policy Commons](#), and the [Other Environmental Sciences Commons](#)

Post van der Burg, Max; Powell, Larkin A.; and Tyre, Andrew J., "FINDING THE SMOOTHEST PATH TO SUCCESS: MODEL COMPLEXITY AND THE CONSIDERATION OF NONLINEAR PATTERNS IN NEST-SURVIVAL DATA Encontrando el Camino Mas Facil Hacia el Exito: Complejidad de los Modelos y Consideración de Patrones No Lineales en Datos de Supervivencia de Nidos" (2010). *Papers in Natural Resources*. 424.

<http://digitalcommons.unl.edu/natrespapers/424>



RESEARCH PAPERS

The Condor 112(3):421–431
© The Cooper Ornithological Society 2010

FINDING THE SMOOTHEST PATH TO SUCCESS: MODEL COMPLEXITY AND THE CONSIDERATION OF NONLINEAR PATTERNS IN NEST-SURVIVAL DATA

MAX POST VAN DER BURG¹, LARKIN A. POWELL, AND ANDREW J. TYRE

School of Natural Resources, University of Nebraska-Lincoln, Hardin Hall, 3310 Holdrege Street, Lincoln, NE 68583

Abstract. Quantifying patterns of nest survival is a first step toward understanding why birds decide when and where to breed. Most studies of nest survival have relied on generalized linear models (GLM) to explore these patterns. However, GLMs require assumptions about the models' structure that might preclude finding nonlinear patterns in survival data. Generalized additive models (GAM) provide a flexible alternative to GLMs for estimating linear and nonlinear patterns in data. Here we present a comparison of GLMs and GAMs for explaining variation in nest-survival data. We used two different model-selection criteria, the Bayes (BIC) and Akaike (AIC) information criteria, to select among simple and complex models. Our study was focused on the analysis of Red-winged Blackbird (*Agelaius phoeniceus*) nests in the Rainwater Basin wetlands of south-central Nebraska. Under BIC, our quadratic model of nest age had the most support, and the model predicted a concave pattern of daily nest survival. We found more model-selection uncertainty under AIC and found support for additive models with ordinal effects of both day and age. These models predicted much more temporal variation than did the linear models. Following our analysis, we discuss some of the advantages and disadvantages of GAMs. Despite the possible limitations of GAMs, our results suggest that they provide an efficient and flexible way to demonstrate nonlinear patterns in nest-survival data.

Key words: *daily nest survival, logistic exposure, generalized linear models (GLM), generalized additive models (GAM).*

Encontrando el Camino Más Fácil Hacia el Éxito: Complejidad de los Modelos y Consideración de Patrones No Lineales en Datos de Supervivencia de Nidos

Resumen. Cuantificar los patrones de supervivencia de los nidos es un primer paso para entender por qué las aves deciden cuándo y dónde reproducirse. La mayoría de los estudios sobre la supervivencia de nidos se han basado en modelos lineales generales (MLG) para explorar esos patrones. Sin embargo, los MLG presentan supuestos sobre la estructura de los modelos que podrían impedir encontrar patrones no lineales en los datos de supervivencia. Los modelos aditivos generales (MAG) brindan una alternativa flexible distinta a los MLG para estimar patrones lineales y no lineales en los datos. En este estudio presentamos una comparación entre los MLG y los MAG para explicar la variación en los datos de supervivencia de nidos. Utilizamos dos criterios diferentes de selección de modelos, los criterios de información de Bayes (BIC) y de Akaike (AIC), para escoger entre modelos simples y complejos. Nuestro estudio se enfocó en el análisis de nidos de *Agelaius phoeniceus* encontrados en humedales de la cuenca del Rainwater, sur-centro de Nebraska. De acuerdo al BIC, nuestro modelo lineal cuadrático de la edad de los nidos tuvo el máximo respaldo y el modelo predijo un patrón cóncavo en la supervivencia diaria de los nidos. Encontramos que hubo mayor incertidumbre en la selección de modelos de acuerdo al AIC y encontramos respaldo para modelos aditivos con efectos ordinales tanto del día como de la edad. Estos modelos predijeron mucha más variación temporal que los modelos lineales. Siguiendo nuestro análisis, discutimos algunas de las ventajas y desventajas de los MAG. A pesar de las posibles limitaciones de los MAG, nuestros resultados sugieren que éstos representan una manera eficiente y flexible para demostrar la existencia de patrones no lineales en los datos de supervivencia de nidos.

Manuscript received 27 March 2009; accepted 20 February 2010.

¹Current address: School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36849. E-mail: mzp0022@auburn.edu

INTRODUCTION

Quantifying patterns in nest survival is a common analytical approach used in studies of avian population ecology (e.g., Brown and Collopy 2008, Bulluck and Buehler 2008). However, quantifying these patterns is only a first step toward understanding the mechanisms of nest survival or the consequences of individuals' decisions about breeding on fitness. We expect individual birds to make choices that minimize the likelihood of brood parasitism and nest predation, hence maximizing reproductive potential. Given the complexity of most ecosystems, there is no reason to expect brood parasitism or nest survival to respond to these choices in a linear fashion. If there are times or conditions that are particularly good, nest survival may even have an intermediate optimum.

Factors that influence nest survival range from the abiotic (e.g., microclimate) to the biotic (e.g., predator behavior). Abiotic factors associated with climate are likely to have nonlinear influences on nest survival not only because of the intrinsic variability of weather phenomena but also because of the behavioral responses of nesting birds and their predators to climate (Roberts and Porter 1998, Dinsmore et al. 2002). For example, increased temperature and precipitation may lead to an increase in olfactory cues from nesting adults (i.e., bacterial growth in plumage; Roberts and Porter 1998). Likewise, the selective force of nest predation is likely to interact with other factors, leading to effects that suggest the apparent optimal placement of the nest (Kruger 2004). The timing of nest initiation and nest development is also likely to influence nest survival non-linearly if the energetic value of a nest as a prey item increases as the nest ages or if predators' efficiency of searching increases during particular nest stages or with nest abundance (Schmidt 1999, Grant et al. 2005). The most popular methods for estimating nest survival with abiotic and biotic covariates, thus far, are based on linear models (Dinsmore et al. 2002, Hazler 2004, Shaffer 2004). While these analytical approaches are extremely useful, we have no reason to assume a priori that nest survival is best represented with a particular parametric form such as quadratic or cubic terms in a generalized linear model (GLM) (e.g., Grant et al. 2005). In fact, linear models may miss multimodal or intermediate optima in patterns of nest success. Generalized additive models (GAM) are a flexible extension of GLMs that can combine parametric forms along with nonparametric smoothers that are more sensitive to nonlinear patterns (Wood 2006). To be clear, both GLMs and GAMs smooth patterns in data. But the way that each approach to modeling accomplishes this is quite different. GLMs smooth by assuming the process can be approximated by linear terms, thus averaging over variation in the data. GAMs also smooth data, but they accomplish this by allowing the smoothing function to be broken into pieces. This procedure provides an approximated function of the

process, but variation in the data is allowed to affect this function more than in the linear case. The result is a model that is much more flexible in terms of capturing nonlinear variation in certain processes. For example, Post van der Burg et al. (2009) detected nonlinear temporal trends in the daily parasitism rate of nests of the Red-winged Blackbird (*Agelaius phoeniceus*) by using GAMs.

Here we compare generalized linear modeling and generalized additive modeling for nest-survival analysis. We analyzed daily survival of Red-winged Blackbird nests with regard to covariates of time, metrics of colonial nesting, and weather. Our goal was to fit both GLMs and GAMs to our data in an effort to show the differences between them and provide some guidance on when these different models may be appropriate. We assessed the ability of each of these model types to explain variation in our data by using an information-theoretic approach, and we further show how the choice of information criterion can influence whether simpler linear models or more complex additive models are chosen.

METHODS

STUDY SPECIES

The Red-winged Blackbird is abundant and found throughout North America (Yasukawa and Searcy 1995). It is a habitat generalist and nests in old fields as well as in wetlands (Yasukawa and Searcy 1995). Aspects of nest placement, such as nest height, are thought to be a response to terrestrial predators (Beletsky 1996, Beletsky and Orians 1996). Red-winged Blackbirds may nest colonially and engage in defensive behavior following hatching of their clutches (Knight and Temple 1988). There is conflicting evidence as to whether colonial nesting provides any benefit in terms of nesting success (e.g., Westneat 1992, Clotfelter and Yasukawa 1999). The role of mammals in nest predation in our study area was established with cameras trained on nests, which documented that least weasels (*Mustela nivalis*), deer mice (*Peromyscus maniculatus*), and striped skunks (*Mephitis mephitis*) depredate Red-winged Blackbird nests (L. Powell, unpublished data). We also made two isolated observations of Northern Harriers (*Circus cyaneus*) feeding Red-winged Blackbird nestlings to their young. Red-winged Blackbirds may nest up to two times per season (Yasukawa and Searcy 1995). For blackbirds nesting in wetlands, weather such as precipitation may affect their reproductive success (Fletcher and Koford 2004).

STUDY AREA AND FIELD METHODS

Within the Rainwater Basin wetlands in Clay Center, Nebraska, we sampled 10 wetlands that varied in size (range: 0.5–92 ha) and ownership (private, state, and federal). We chose our sample of wetlands on the basis of whether we could gain access to the wetland in repeated years. The plants dominating these wetlands were cattail (*Typha* spp.), river bulrush (*Scirpus*

fluvaitalis), reed canary grass (*Phalaris arundinacea*) and smartweed (*Polygonum* spp.). We began searching for nests in the second or third week of May and terminated searches at the end of July each year. At the beginning of the 2002 season we established random transects in wetlands exceeding 1 ha and covered both wetland and upland habitat. In 2003 we adjusted these transects to cover only the wetland; we used these same transects in 2004. We searched the entire wetland if it was smaller than 1 ha. Otherwise, we searched for nests within 100 m on either side of the transects. We identified each nest with a number by writing on vegetation near the nest with a felt-tip marker. We tied plastic flagging to vegetation >10 m from the nest and wrote the distance and direction to the nest on the flag. We also recorded the UTM coordinates for each nest with a GPS unit. We revisited nests every 3–5 days until the young fledged or the nest failed.

We measured nest height (from the ground to the top edge of the nest) to the nearest 0.5 dm. We also overlaid the nest's coordinates on a modified National Wetlands Inventory coverage (U.S. Fish and Wildlife Service 2005) and used the ruler tool in ESRI's ArcGIS (version 8.x) to measure the shortest distance from each nest to the wetland's edge in meters. We considered the edge of each wetland to be in the zone of transition between upland and wetland vegetation, and we verified the location of these transition zones in the field. Using the Hawth's Tools (Beyer 2004) extension for ArcGIS, we also used the nest coordinates to calculate the distance between simultaneously active nests. We obtained daily weather data (daily high temperature and daily precipitation) for Clay Center from the High Plains Climate Research Center database (University of Nebraska-Lincoln, Lincoln, Nebraska). All study sites were within 19 km of the weather station.

STATISTICAL ANALYSIS

We used the statistical program R (R Development Core Team 2006) to estimate daily nest survival with the logistic-exposure method outlined by Shaffer (2004). We fit both GLMs and GAMs to our binomial survival data (0 = failure, 1 = success). We used a modified logit-link function developed by Shaffer (2004): $g(\theta) = \ln[\theta^{1/t}(1 - \theta)^{t-1}]$, where θ is the survival estimate for the interval monitored and t is the interval length in days.

Generalized linear models. Most ornithologists are familiar with the structure of a GLM, $g(y) = \beta_0 + \beta_i x_i$, where the function g is a link function, β_0 is an intercept, and $\beta_i x_i$ is a collection of parameters plus covariate data. We fit GLMs by using the R package *glm*, which uses an iterative least-squares approach to fit each model to the data. This method uses a solution common in linear algebra to find maximum-likelihood estimates of the model's parameters. The iterative portion of the algorithm typically involves calculating weights, which are functions of the residuals. After each iteration, the weights are updated and the linear equation is solved again. As a result of this weighting, the estimates are more robust to outliers in

the data. Because these models are fairly simple we can present both parameter values and the models' predictions easily.

Generalized additive models. Now consider the structure of a GAM: $g(y) = \beta_0 + s(x_i)$, where s is known as a smooth function (Hastie and Tibirishani 1990, Wood 2006, Crawley 2009, Zuur et al. 2009). We used the package *mgcv* to fit GAMs to our data. This package also uses an iterative least-squares approach to fit each model and treats the number of smoothing terms as something to be estimated in the context of fitting the model (Wood and Augustin 2002). In *mgcv*, these smooth functions are represented as spline functions. A spline function is a piecewise polynomial function that is often used to represent smoothed and nonlinear relationships. The problem with modeling GAMs with spline functions is in the choice of the order of polynomial. If you choose too high an order the model may be overfitted. If you choose too small an order the model will not approximate the patterns in the data. The penalized-regression spline algorithm used by *mgcv* estimates the order of the smoothing function for each covariate on the basis of how it balances the model's fit with overall smoothness (Wood and Augustin 2002). The result is a model that has more "wiggleness" than a GLM but is not overfitted. If a model parameter fits better as a smooth term, then the degrees of freedom will increase. The number of parameters in each model is approximated by the effective degrees of freedom (edf); each linear term contributes only one degree of freedom. In the *mgcv* package, the estimation of smoothing parameters is accomplished by minimizing either the generalized cross validation criterion or the unbiased risk estimator. In our exercise we are using binomial data, so *mgcv* relies on the unbiased risk estimator, $D/n + 2\phi(\text{edf})/n - \phi$, where D is the model's deviance, n is the number of data points, and ϕ is a scale parameter. In the case of the unbiased risk estimator, the scale parameter, assumed to be the variance of the process, must be known. If the data were normally distributed, we could use generalized cross validation and estimate the scale term. For discrete Poisson and binomial data the score of the unbiased risk estimator is more appropriate (Crawley 2007). Notice that this score is really just a rescaled score of the Akaike information criterion (AIC). To get that score, we assume that the scale parameter is equal to one, which is the default for binomial and Poisson data in this package. The smooth portion of a GAM is represented as a matrix of basis functions (Wood and Augustin 2002), so it is often not practical to present those in lieu of parameter estimates. Therefore, we will focus on presenting the predictions of our models. However, the package does provide approximate P -values for assessing the significance of smooth terms. These values are calculated by comparing a statistic derived from the parameter vector and its covariance matrix against a chi-squared distribution with the degrees of freedom equal to the effective degrees of freedom (see *mgcv* help files for more information).

Candidate models. We compared multiple working hypotheses that were expressed as both generalized linear and additive models. We considered linear models composed of covariates related to the nest site, interval-specific weather conditions, colonial nesting, and time. To determine which of these possible models to include in our analysis, we conducted a simple exploratory analysis. We used Microsoft Excel to summarize the nests' fates and number of exposure days for various levels of our covariates. We then calculated simple Mayfield estimators (*sensu* Mayfield 1961) for the data in each level of our covariates, and we plotted these results against the covariates' values. On the basis of these results, we built models using parameters that looked as if they had an effect on nest survival. Likewise, if they looked as if they had a nonlinear effect on nest survival we included those parameters in our analysis of the additive models. On the basis of our exploratory analysis, neither of our nest-site variables, nest height (HT) and distance from wetland edge (EDGE), had an effect on daily nest survival, so we excluded them from further analysis.

We built a weather model that included average daily precipitation (PRECIP) and average daily high temperature (TEMP). Our exploratory plots suggested that increasing temperature and precipitation could have a nonlinear effect on daily nest survival. Daily nest survival was higher at the low end of the range of temperatures, lower at the seasonal average temperature, and then it seemed to increase again at higher temperatures. It also seemed to increase as precipitation increased, but the rate of increase seemed much faster for smaller amounts of precipitation. Precipitation and temperature could lead to lower nest survival by influencing olfactory cues used by predators to locate nests or by influencing activity of other predators like snakes (Roberts and Porter 1998, Dinsmore et al. 2002). For both covariates, we averaged daily measurements over each interval of monitoring. Our model for colonial nesting accounted for the effects of the number of simultaneously active nests (SIM) and the shortest distance between those nests (DIST). Our plots suggested that the number of simultaneous nests influenced nest survival nonlinearly. But in general this pattern suggested that nest survival was higher for lower and intermediate numbers of nests than for higher numbers of nests. On the basis of our analysis, as the distance between these nests increased, nest success also apparently increased, but the rate of increase was much faster when the distance between nests was small. Because the Red-winged Blackbird is colonial, it could reduce predation risk by choosing sites with greater numbers of individuals, though Clotfelter and Yasukawa (1999) failed to find this effect. There may be a trade-off between nesting in a colony with too many closely packed individuals (increased prey density) in comparison to more intermediate nest densities and distances between nests. We also found evidence that as nests aged, their survival rates declined, but survival increased again beyond a threshold age. The success of nests earlier in the season also

appeared higher than the success of those later in the season. We found that plots of both parameters had quite a lot of scatter in the Mayfield estimates, so we cannot rule out nonlinear dynamics in these parameters. Our time-specific model included an age effect (AGE) and an effect of day of the nesting season (ordinal day; DAY). For both age and ordinal day we assigned the average value of each interval (*i.e.*, the midpoint of the interval) as the covariate's value. Previous work suggests that blackbird nests are less likely to survive as they age (Young 1963, Robertson 1972, Caccamise 1976). These estimates, however, were not based on an analysis of daily nest survival. We expected that temporal patterns of daily nest survival to be nonlinear because the blackbirds' behavior might influence the risk of nest survival at different stages of development (Eichholz and Koenig 1992, Davis 2005, Grant et al. 2005). Likewise, we do not have any reason to expect nest survival to be linear throughout the season because predators may respond to temporal fluctuations in nesting densities or alternative prey (Jones et al. 2002).

Up to this point we have described our three "base" models, which were our weather model, colonial-nesting model, and temporal model. Because we found evidence in our exploratory analysis that the models' parameters may vary nonlinearly, we included an additional three models with quadratic effects on the parameters. We then composed pairwise combinations of the models, included a null model (intercept only), main-effect models, and two "global" models for a total of 27 GLMs (Table 1). We then composed the GAM equivalent of these models, which added an additional 13 models (see Table 1 for complete list).

Model-selection criteria. We fit a total of 40 models. We used information-theoretic methods to rank candidate models and to account for the models' uncertainty. We used two criteria to choose models, one known to select simple models and one known to select more complex models. We used the Bayes information criterion (BIC) to select simpler models. Link and Barker (2006) have shown that BIC scores can be used to approximate Bayes factors for comparing models i and j : $\exp[(\text{BIC}_i - \text{BIC}_j)/2]$. These approximate Bayes factors (BF) can then be used to calculate the posterior probabilities of models in a set in much the same way that AIC weights are calculated: $\text{pr}(\text{model}_i) = \text{BF}_{i,1} \pi_i / \sum \text{BF}_{j,1} \pi_j$, where the numerator is the Bayes factor of model i compared with model 1 in the set times the prior probability of that model (π). The denominator is the sum of all the j models in the set compared to model 1 times their respective priors. The advantage of this method for calculating posterior probabilities rests in the fact that the researcher can explicitly state prior belief in a particular model or set of models. Thus one is not confined to a model-selection criterion that has implicit prior assumptions. In our case, we chose to state our prior belief as a uniform distribution [*i.e.*, $\text{pr}(\text{model}) = 1/40$], which means that we did not favor any one model or set of models. We considered our best models

TABLE 1. Comparison of generalized linear models (GLM) and generalized additive models (GAM) of Red-winged Blackbird (*Agelaius phoeniceus*) daily nest survival in the Rainwater Basin region of Nebraska (2002–2004). Model parameters are average high temperature (TEMP), average precipitation (PRECIP), number of simultaneously active nests (SIM), the minimum distance between those nests (DIST), average nest age (AGE), average ordinal day in the nesting season (DAY), and a null, or intercept only, model (NULL). GLMs were fit with a maximum-likelihood approach. We present log likelihoods and the number of model parameters (K) for GLMs only. Smooth terms for GAMs, denoted by s (PARAMETER), were estimated for all of the parameters with a penalized spline regression algorithm. The number of smooth terms for each parameter was estimated by minimizing the unbiased risk estimator (UBRE). Models that include smooth terms have higher estimated degrees of freedom (edf) because nonsmooth terms only contribute one degree of freedom. All models were compared by Bayes (BIC) and Akaike (AIC) information criteria. BIC scores were used to approximate Bayes factors to estimate posterior model probabilities. Using differences in AIC scores, we also estimated the models’ posterior probabilities. All models with posterior probabilities greater than zero are presented in **bold**. For comparison, we also include a measure of the amount of deviance explained by a particular model.

GLM	Deviance explained (%)	K	BIC	Model probability (BIC)	AIC	Model probability (AIC)
Single-effect models						
Null ^a	5.81	1	1396.42	0.00	1390.79	0.00
AGE	5.94	2	1402.05	0.00	1390.79	0.00
AGE^{2a}	8.10	3	1377.87	0.99	1360.98	0.00
DAY	5.86	2	1403.26	0.00	1392.00	0.00
DAY ^{2a}	6.30	3	1404.39	0.00	1387.50	0.00
TEMP	5.90	2	1402.63	0.00	1391.37	0.00
TEMP ²	5.98	3	1409.09	0.00	1392.20	0.00
PRECIP	5.81	2	1403.98	0.00	1392.72	0.00
PRECIP ²	5.84	3	1411.24	0.00	1394.35	0.00
SIM	5.81	2	1404.05	0.00	1392.79	0.00
SIM ²	5.92	3	1410.05	0.00	1393.16	0.00
DIST	5.93	2	1402.29	0.00	1391.03	0.00
DIST ²	6.08	3	1407.65	0.00	1390.76	0.00
Temporal models						
AGE + DAY	5.95	3	1409.53	0.00	1392.64	0.00
AGE ^{2a} + DAY ^{2a}	8.35	5	1389.52	0.00	1361.37	0.00
Weather models						
TEMP + PRECIP	5.93	3	1409.92	0.00	1393.03	0.00
TEMP ² + PRECIP ²	6.04	5	1423.45	0.00	1395.30	0.00
Colonial-nesting models						
SIM + DIST	5.94	3	1409.78	0.00	1392.89	0.00
SIM ² + DIST ²	6.17	5	1421.61	0.00	1393.46	0.00
Temporal and weather models						
AGE + DAY + TEMP + PRECIP	6.03	5	1423.75	0.00	1395.60	0.00
AGE ^{2a} + DAY ^{2a} + TEMP ² + PRECIP ²	8.38	9	1419.48	0.00	1368.81	0.00
Temporal and colonial-nesting models						
AGE + DAY + SIM + DIST	6.07	5	1423.11	0.00	1394.96	0.00
AGE ^{2a} + DAY ^{2a} + SIM ² + DIST ²	8.72	9	1414.57	0.00	1363.90	0.00
Weather and colonial-nesting models						
TEMP + PRECIP + SIM + DIST	6.06	5	1423.29	0.00	1395.14	0.00
TEMP ² + PRECIP ² + SIM ² + DIST ²	6.34	9	1442.05	0.00	1397.01	0.00
Global models						
AGE + DAY + TEMP + PRECIP + SIM + DIST	6.14	7	1437.29	0.00	1397.88	0.00
AGE ^{2a} + DAY ^{2a} + TEMP ² + PRECIP ² + SIM ² + DIST ²	8.76	13	1444.52	0.00	1371.33	0.00

(continued)

TABLE 1. (Continued)

GAM	UBRE	Deviance explained (%)	edf	BIC	Model probability (BIC)	AIC	Model probability (AIC)
Single-effect models							
$s(\text{AGE})^a$	-0.35	9.76	7.36	1386.74	0.01	1345.31	0.00
$s(\text{DAY})^a$	-0.33	8.24	8.69	1419.25	0.00	1370.34	0.00
$s(\text{TEMP})$	-0.32	5.91	2.00	1402.64	0.00	1391.37	0.00
$s(\text{PRECIP})$	-0.32	5.81	2.00	1403.98	0.00	1392.72	0.00
$s(\text{SIM})$	-0.32	6.17	3.78	1412.34	0.00	1391.05	0.00
$s(\text{DIST})$	-0.32	6.01	2.48	1404.80	0.00	1390.82	0.00
Temporal model							
$s(\text{AGE})^a + s(\text{DAY})^a$	-0.35	11.8	15.19	1416.89	0.00	1331.36	0.61
Weather model							
$s(\text{TEMP}) + s(\text{PRECIP})$	-0.32	5.93	3.00	1409.92	0.00	1393.03	0.00
Colonial-nesting model							
$s(\text{SIM}) + s(\text{DIST})$	-0.32	6.31	4.99	1419.44	0.00	1391.37	0.00
Temporal and weather model							
$s(\text{AGE})^a + s(\text{DAY})^a + s(\text{TEMP}) + s(\text{PRECIP})$	-0.35	11.8	17.09	1431.11	0.00	1334.86	0.11
Temporal and colonial-nesting model							
$s(\text{AGE})^a + s(\text{DAY})^a + s(\text{SIM}) + s(\text{DIST})$	-0.35	12.2	19.23	1441.52	0.00	1333.22	0.24
Weather and colonial-nesting model:							
$s(\text{TEMP}) + s(\text{PRECIP}) + s(\text{SIM}) + s(\text{DIST})$	-0.32	6.46	7.13	1433.58	0.00	1393.44	0.00
Global model							
$s(\text{AGE})^a + s(\text{DAY})^a + s(\text{TEMP}) + s(\text{PRECIP}) + s(\text{SIM}) + s(\text{DIST})$	-0.35	12.2	21.08	1455.54	0.00	1336.86	0.04

^aSignificant at $\alpha = 0.05$.

to be the ones with the highest posterior probabilities (Link and Barker 2006).

We used AIC to select more complicated models. Link and Barker (2006) also showed that the calculation of AIC weights includes an implicit prior (implicit in the AIC score) that favors complex models. They also showed that AIC weights are, in fact, posterior probabilities. Therefore, we refer only to posterior probabilities when assessing the support for models under each criterion.

After ranking the models in our set on the basis of posterior probabilities, we then made model-averaged predictions, which we present as means with associated 95% confidence intervals. Parameters for GLMs are presented as means ± 1 SE, and means calculated directly from observed data are presented as means ± 1 SD. We further estimated nest success by bootstrapping the predictions made from our set of models, which entailed taking the predicted means and standard errors on the logit scale and randomly drawing values from a normal distribution. These values were then back transformed and multiplied together. We performed this routine 1000 times and calculated the mean and confidence intervals from the distribution of nest-success values.

RESULTS

Over the 3 years of our study we monitored 592 active Red-winged Blackbird nests. We found that clutch size varied from zero (for some parasitized nests) to five eggs and averaged 2.9 ± 1.2 eggs. Over all three years, the number of active nests generally increased from mid-May to early June, then slowly declined until the first week of August. For the analysis, our sample was reduced to 418 nests because we were unable to age all of our nests or because field workers did not record certain covariates. This did not seem to have an appreciable qualitative impact on our parameter estimates. Covariate data are summarized in Table 2.

As expected, the logistic-exposure model that best explained variation in daily nest survival depended on the type of information criterion we used. Assuming simple models should receive more emphasis (BIC), we found that the linear model with a quadratic effect of nest age best explained our data (Table 1). Because the probability for this model was so high, we did not model-average our predictions under this criterion (Burnham and Anderson 2002). The parameters of our estimated linear model were 3.58 ± 0.12 for the intercept,

TABLE 2. Data recorded from Red-winged Blackbird (*Agelaius phoeniceus*) nests in the Rainwater Basin region of Nebraska, 2002–2004. The number of nests refers to the size of the total sample of nests analyzed in this study. The number of intervals refers to the total number of observations of nest fate, with Mayfield estimates of daily nest survival in parentheses. All ordinal data (nest age, ordinal day, simultaneously active nests) are presented as means with the range of the observed data in parentheses. Continuous data are presented as means \pm 1 SD. Data are summarized for each of the three years of the study as well as for all three years aggregated.

	2002	2003	2007	Overall
Number of nests	79	133	206	418
Number of observed intervals	361 (0.96)	677 (0.97)	1021 (0.96)	2059 (0.96)
Nest age	10.55 (1–28)	12.58 (1–28)	12.20 (1–28)	12.05 (1–28)
Ordinal day	24.00 (3–79)	40.64 (6–72)	35.41 (1–73)	33.80 (1–79)
Average high temperature (°C)	28.62 \pm 5.01	29.92 \pm 3.48	26.94 (3.29)	28.22 (3.93)
Average precipitation (cm)	0.15 \pm 0.31	0.20 \pm 0.33	0.40 (0.96)	0.29 (0.73)
Simultaneously active nests	14.12 (1–41)	13.28 (1–37)	15.45 (1–46)	14.51 (1–46)
Distance to nearest simultaneously active nest	28.37 \pm 39.54	43.87 \pm 66.44	41.85 (60.01)	40.28 (59.66)

-0.29 ± 0.05 for the main age parameter, and 0.01 ± 0.001 for the quadratic effect of age. All were significantly different from zero ($P < 0.05$). We plotted the effect of age on the prediction scale in Fig. 1, which shows a clearly concave relationship between daily nest survival and nest age.

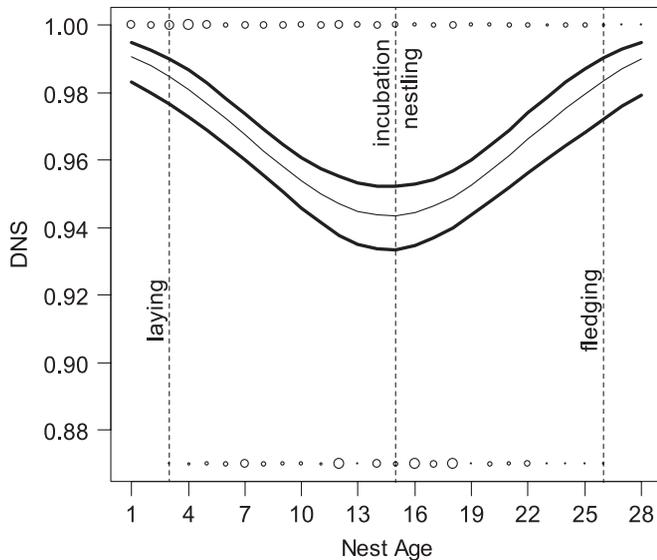


FIGURE 1. Model-predicted daily survival (DNS) of Red-winged Blackbird (*Agelaius phoeniceus*) nests in central Nebraska (2002–2004) as a function of nest age (day 1 = first day of nest building). These predictions were based on a generalized linear model (GLM) with a quadratic age term (age^2). The vertical dotted lines represent the average transition periods between different nest stages (laying, incubation, nestling, and fledging). The thin horizontal line represents the average survival prediction, and the thicker lines represent 95% confidence intervals. Interval data used to estimate survival are plotted as circles; those at the top of the figure represent intervals that the nests survived, those at the bottom represent intervals in which the nests failed. The size of each circle is scaled to the number of nests observed for each value of nest age (large = many observations, small = few observations).

The approach emphasizing more complicated models (AIC) implied more selection uncertainty than did BIC. The model with the highest probability was the additive temporal model with the effects of nest age and ordinal day (Table 1). In general, all the additive models that had a temporal component had some support under AIC posteriors, suggesting that both nest age and ordinal day had a good deal of support. We model-averaged our predictions across the five GAMs with nonzero probabilities. Only the age of the nest and ordinal day had a clear pattern, and the effects of other parameters were largely flat or were so uncertain as to be uninterpretable. Because the smoothed part of the model is nonparametric, we did not have parameter estimates for our smooth terms. Our models predicted that daily nest survival had a bimodal pattern with a decline following clutch completion and just after hatching (Fig. 2A). Much as in our exploratory analysis, daily nest survival generally appears to be lower during the middle of the breeding season than at the beginning and end of the season (Fig. 2B). However, the uncertainty at the end of the season makes it difficult to interpret the general pattern of DNS late in the breeding season.

Using the best-supported GLM and best-supported GAM, we estimated nest success. The average seasonal estimate based on the GLM was 0.37 (0.35–0.38), that based on the GAM was 0.36 (0.27–0.46). Assuming that nest success varied throughout the breeding season, we also estimated nest success by our bootstrapping method, but as a function of initiation date for the model-averaged GAM. These results suggested a nonlinear pattern of nest success through the season, with nest success being quite a bit higher earlier in the season than later (Fig. 3).

DISCUSSION

Our results show that modeling nest survival is a bit more complicated than simply developing biologically realistic hypotheses and using an objective criterion to sort among the

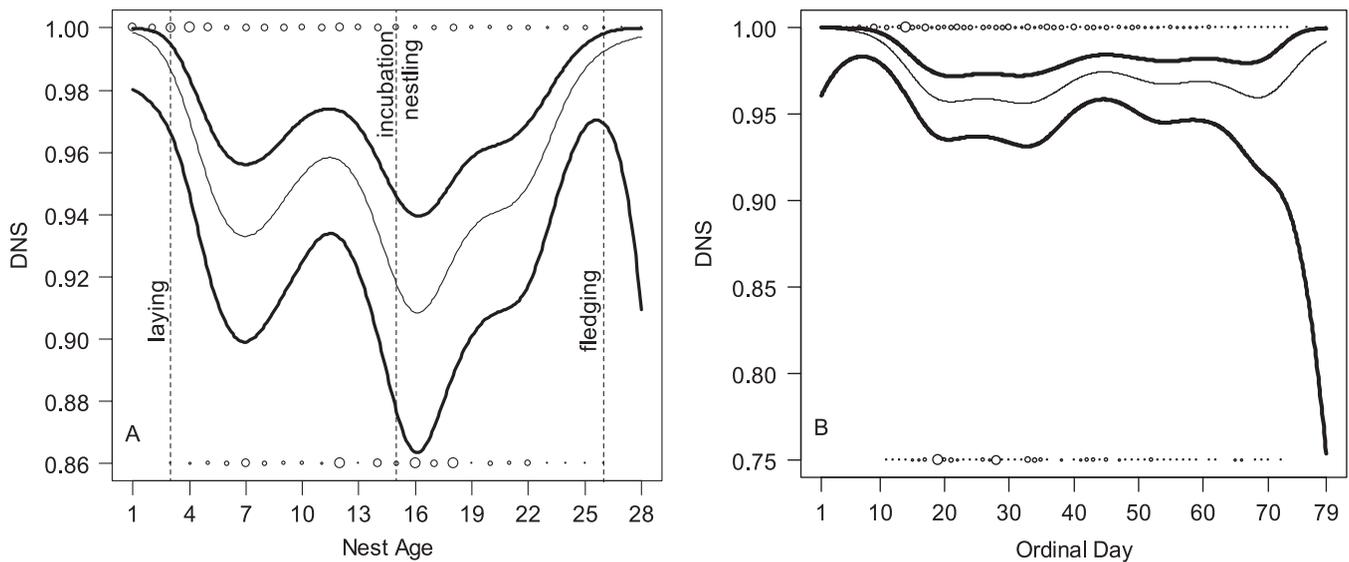


FIGURE 2. Daily survival (DNS) of Red-winged Blackbird (*Agelaius phoeniceus*) nests in central Nebraska (2002–2004) as a function of nest age (A; day 1 = first day of nest building) and ordinal day in the breeding season (B; day 1 = average date of 18 May), based on model-averaged predictions from a set of generalized additive models (GAMs). In (A), the vertical dotted lines represent the average transition periods between different nest stages (laying, incubation, nestling, and fledging). In both (A) and (B), the thin horizontal line represents the average predicted survival, the thicker lines represent 95% confidence intervals. Interval data used to estimate survival are plotted as circles; those at the top of each panel represent intervals that the nests survived, those at the bottom represent intervals in which the nests failed. The size of each circle is scaled to the number of nests observed for each value of nest age (large = many observations, small = few observations).

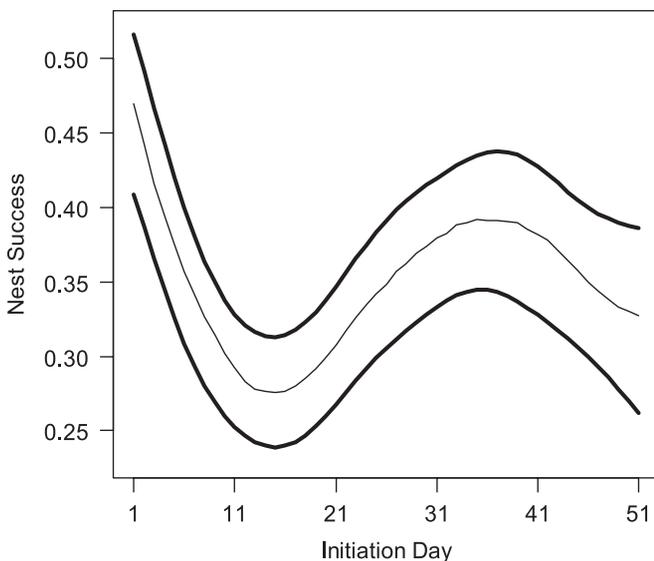


FIGURE 3. Bootstrapped predictions of success of Red-winged Blackbird (*Agelaius phoeniceus*) nests in central Nebraska (2002–2004), model-averaged from a set of generalized additive models. These predictions were made by conditioning each nest on a different initiation day within the season and taking the product of age-specific predictions. The thin line represents the mean of 1000 iterations, and the thick lines represent the 2.5% and 97.5% quantiles of those 1000 iterations.

models. In fact, choosing the criterion and the models’ form (i.e., linear or nonlinear) seems as important in modeling nest survival as making defensible biological inferences. Our results suggest that the choice of model-selection criterion plays a large role in distinguishing between simple models or more complex models in terms of their ability to explain variation in the data. Our results might appear to some readers as minor technical issues in terms of selection and development of models. However, the end goal of modeling nest survival should be to provide information for predicting avian fecundity. It appears that the choice of pattern used to represent nest survival and, ultimately, nest success in productivity models is likely to have an effect on the types of inferences that one draws. The effect of these models is evident in the patterns we discerned in the Red-winged Blackbird’s nest success in our study area. Nest success estimated on the basis of our linear model would likely give an answer different from a productivity estimate based on our seasonal GAM-based success estimates.

BIOLOGY OF BLACKBIRD NESTING

Both sets of nest-survival models suggest that temporal factors (i.e., age and day effects) account for more of the variation in daily nest survival than do other factors, such as nest-site characteristics. Similar results have been found for other passerines (e.g., Grant et al. 2005, Bonnot et al. 2008), but the

effect of time is difficult to generalize across species as it appears to be species specific. The general finding that time often explains variation in nest survival better could result from temporal variables often being recorded at a resolution that more closely matches the scale at which interval data are recorded. Habitat variables are typically not recorded at the same resolution as data on survival through an interval. As a result, perhaps temporal variables are better at explaining the finer-scale variation in this data.

The predictions of the more complex additive model are not all that different qualitatively from those of the equivalent linear model. But these more complex results do suggest that there are likely two periods of rapid decline in survival: one for the incubation period and one for the nestling period. Numerous studies have found that nest survival tends to vary temporally depending on the development of the nest (e.g., Sockman 1997, Peak et al. 2004, Burhans et al. 2002). Our results indicate that nest survival is high during the period when eggs are being laid, which could suggest that predators do not forage preferentially on newly initiated nests. Interestingly, our predictions suggested that the daily survival rate declined following clutch completion, followed by an increase in survival toward the end of the incubation period. One potential explanation for this pattern could be related to predators' foraging behavior (Schmidt 1999): predators could gain more energy by feeding on nests with completed, and possibly larger, clutches. However, predators must search for these nests; as nests that are easy to find are lost to predation, the value of remaining (more difficult to find) nests decreases, so predators should forage elsewhere. Thus the survival rate of the difficult-to-find nests increases during that period.

Just prior to the eggs' hatching, daily nest survival drops off and hits its lowest point after hatching. At this stage, predators may increase their foraging as nests become easier to find because of behavioral cues at the nest (increased movement and sound of both adult female and nestlings; Martin et al. 2000). Following this decrease, daily nest survival then increases. After the young hatch, Red-winged Blackbirds defend their nests, which might explain the increase in survival we observed at the nestling stage (Knight and Temple 1988).

GENERALIZED LINEAR VS. ADDITIVE MODELS

Regardless of what effects one is able to correlate with nest survival, the ultimate goal of estimating such patterns should be to help predict avian productivity. The choice of model form, however, could affect the nature of the pattern of nest survival and ultimately nest success used in productivity models. In turn, this choice is affected by whether one uses a model-selection criterion that has explicit (or implicit) preferences for simple or complex models. Our comparison was not meant to demonstrate the superiority of one approach over another but rather to suggest that there might be advantages to considering types of model forms other than GLMs (i.e., GAMs). We

suggest that the choice of one approach over another should be motivated by the types of questions that one would like to ask. These questions could be of a general sort. For example, theoretical predictions could lead a researcher to expect that nest survival should be highest at intermediate levels of certain nest-site characteristics (Kruger 2004). Likewise, and perhaps more practically, research questions can be honed by exploratory analyses. From our dataset we made predictions based on theory, previous observations, and initial exploratory plots of our data. These plots can be constructed in readily available spreadsheet software prior to a formal multi-model inference analysis. Then one could make a choice about model type (linear or additive) on the basis of consideration of the data.

We argue that the choice should also be based on whether the type of question has more to do with variation specific to a dataset or with making predictions that can be generalized. For example, GLMs have the characteristic of being built very simply and so are easily generalized both spatially and temporally. This approach might be of use in a comparison of the patterns of nest survival by habitats or by species. Likewise, if one is interested in making inferences about nest success in general, as in a comparison of regions, our results suggest that one could examine the patterns of nest survival with linear models and still make accurate assessments of nest success overall. Additionally, the output of a GLM is easy to interpret, and the parameter estimates are easily transferrable to other exercises in modeling.

GAMs, on the other hand, provide an efficient method for fitting nonlinear models to specific nest-survival datasets. If one's goal is to examine fine-scale nonlinearities particular to a given dataset, GAMs might be more useful than GLMs. GAMs could be especially important in analysis of temporal patterns in daily nest survival, which appear to be specific to a species (Grant et al. 2005) and perhaps even to a region. The complex results of an additive model may appear to be more difficult to apply to models of avian productivity (e.g., Powell and Knutson 2006, Pease and Grzybowski 1995). While some productivity models may require further development in order to accommodate an increase in temporal complexity, others that can incorporate the predictions of linear models can just as easily use the predictions from GAMs. Our focus on the specificity of GAMs is not to say that they preclude finding a general pattern in, say, estimates of nest success. In fact, our study shows that GLMs and GAMs give very similar estimates of average nest success. Using the additive set of models, however, we found a different temporal pattern of nest success, suggesting that ignoring the resulting differences in nest-success estimates affect our ability to estimate productivity because the mean success estimated from the linear model may underpredict the number of offspring produced.

We found two additional advantages of GAMs over GLMs. First, with a GLM, including quadratic and cubic effects entails an increase the set of possible models. Tenets of the multimodel

inference approach suggest that the set of potential models should be kept as small as possible (Burnham and Anderson 2002). Second, the *mgcv* package allowed us to estimate the number of smoothing terms as part of the procedure of fitting the model. Therefore, we did not have to make a priori statements about the structure of this variation and we were able to capture more variability without increasing the number of additional models in our set over that of linear models.

That said, trade-offs between GLMs and GAMs remain, stemming from differences in model complexity. In any model, there is an inverse relationship between interpretability and complexity. GAMs are very flexible and can be influenced by idiosyncrasies in the data. Fortunately, advancements such as *mgcv* (Wood and Augustin 2002) should minimize the risk of such influences. Still, the flexibility, and thus complexity, of GAMs may make their results difficult to interpret, particularly when one would like parameter estimates. Likewise, our finding nonlinearities in our dataset doesn't mean that these same patterns are likely to be found in others. Small data sets, for instance, may not provide enough information for the smoothing process to show as much variation in daily nest survival. In such a case, the GAM will collapse into a linear model (edf = 1). In this case, little would be lost in terms of considering these processes as linear models because the *mgcv* algorithm smooths this covariate as a linear parameter. The next step in this situation, to get the estimated linear parameter, would be to rerun the analysis with the smoothing term removed from the model statement.

The tendency of the *mgcv* algorithm to smooth when possible is useful because it suggests that the approach does not arbitrarily impose a nonlinear model on processes that are in fact linear. Put another way, it does seem to distinguish between linear and nonlinear dynamics in data. This is important because some ecological phenomena are, indeed, nonlinear, and imposing linear models on those processes does not increase our understanding, partially because linear models, even with polynomial terms, have limited flexibility when fit to complex nonlinear data (Austin et al. 2006). Therefore, simply making assumptions with a model for the purpose of tractability does little to help us understand the underlying biology of a system (Wood 2001).

ACKNOWLEDGMENTS

Funding was provided by the Layman Trust in the Institute of Agriculture and Natural Resources, as well as the School of Natural Resources at the University of Nebraska-Lincoln. MP was supported by a GAANN graduate training grant from the Initiative for Ecology and Evolutionary Analysis at the University of Nebraska-Lincoln. Logistical and financial support was also provided by the U.S. Fish and Wildlife Service and the Nebraska Game and Parks Commission. T. Kocer, Z. Cunningham, M. Lelevier, and N. J. Gonzalez assisted with field work. B. Tenhumberg, K. Christiansen, P. Unitt, and two anonymous reviewers provided constructive comments on the manuscript. R code for the GAM application is available from MP. This research was supported by Hatch Act funds through the University of Nebraska Agricultural Research Division, Lincoln, Nebraska.

LITERATURE CITED

- AUSTIN, M., L. BELBIN, J. MEYERS, M. DOHERTY, AND M. LUOTO. 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecological Modeling* 199:197–216.
- BELETSKY, L. 1996. The Red-winged Blackbird: the biology of a strongly polygynous songbird. Academic Press, San Diego.
- BELETSKY, L., AND G. ORIANI. 1996. Red-winged Blackbirds: decision-making and reproductive success. University of Chicago Press, Chicago.
- BEYER, H. L. [ONLINE]. 2004. Hawth's analysis tools for ArcGIS. <<http://www.spatial ecology.com/htools>> (December 2008).
- BONNOT, T., M. RUMBLE, AND J. MILLSPAUGH. 2008. Nest success of Black-backed Woodpeckers in forests with mountain pine beetle outbreaks in the Black Hills, South Dakota. *Condor* 110:450–457.
- BROWN, J. L., AND M. W. COLLOPY. 2008. Nest-site characteristics affect daily nest-survival rates of Northern Aplomado Falcons (*Falco femoralis septentrionalis*). *Auk* 125:105–112.
- BULLUCK, L. P., AND D. A. BUEHLER. 2008. Factors influencing Golden-winged Warbler (*Vermivora chrysoptera*) nest-site selection and nest survival in the Cumberland Mountains of Tennessee. *Auk* 125:551–559.
- BURHANS, D. E., D. DEARBORN, I. F. R. THOMPSON, AND J. FAABORG. 2002. Factors affecting predation at songbird nests in old fields. *Journal of Wildlife Management* 66:240–249.
- BURNHAM, K. P., AND D. R. ANDERSON. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York.
- CACCAMISE, D. F. 1976. Nesting mortality in the Red-winged Blackbird. *Auk* 93:517–534.
- CLOTFELTER, E. D., AND K. YASUKAWA. 1999. The effect of aggregated nesting on Red-winged Blackbird nest success and brood parasitism by Brown-headed Cowbirds. *Condor* 101:729–736.
- CRAWLEY, M. J. THE R BOOK. 2007. Wiley, Chichester, England.
- DAVIS, S. K. 2005. Nest-site selection patterns and the influence of vegetation on nest survival of mixed-grass prairie passerines. *Condor* 107:605–616.
- DINSMORE, S. J., G. C. WHITE, AND F. L. KNOPF. 2002. Advanced techniques for modeling avian nest survival. *Ecology* 83:3476–3488.
- EICHHOLZ, M. W., AND W. D. KOENIG. 1992. Gopher snake attraction to birds' nests. *Southwestern Naturalist* 37:293–298.
- FLETCHER, R., AND R. KOFORD. 2004. Consequences of rainfall variation for breeding wetland blackbirds. *Canadian Journal of Zoology* 82:1316–1325.
- GRANT, T. A., T. L. SHAFFER, E. M. MADDEN, AND P. J. PIETZ. 2005. Time-specific variation in passerine nest survival: new insights into old questions. *Auk* 122:661–672.
- HASTIE, T., AND R. TIBSHIRANI. 1990. Generalized additive models. Chapman and Hall, London.
- HAZLER, K. R. 2004. Mayfield logistic regression: a practical approach for analysis of nest survival. *Auk* 121:707–716.
- JONES, D. D., L. M. CONNER, R. J. WARREN, AND G. O. WARE. 2002. The effect of supplemental prey and prescribed fire on success of artificial nests. *Journal of Wildlife Management* 66:1112–1117.
- KNIGHT, R. L., AND S. A. TEMPLE. 1988. Nest-defense behavior in the Red-winged Blackbird. *Condor* 90:193–200.
- KRUGER, O. 2004. Breeding biology of the Cape Bulbul *Pycnonotus capensis*: a 40-year comparison. *Ostrich* 75:211–216.
- LINK, W. A., AND R. J. BARKER. 2006. Model weights and the foundations of multimodel inference. *Ecology* 87:2626–2635.
- MARTIN, T. E., J. SCOTT, AND C. MENGE. 2000. Nest predation increases with parental activity: separating nest site and parental activity effects. *Proceedings of the Royal Society of London B* 267:2287–2293.

- MAYFIELD, H. 1961. Nesting success calculated from exposure. *Wilson Bulletin* 73:255–261.
- PEAK, R. G., F. R. THOMPSON AND T. L. SHAFFER. 2004. Factors affecting songbird nest survival in riparian forests in a midwestern agricultural landscape. *Auk* 121:726–737.
- PEASE, C. M., AND J. A. GRZYBOWSKI. 1995. Assessing the consequences of brood parasitism and nest predation on seasonal fecundity in passerine birds. *Auk* 112:343–363.
- POST VAN DER BURG, M., L. A. POWELL AND A. J. TYRE. 2009. Modeling parasitism rate and parasitism risk: an example using a colonially nesting songbird, the Red-winged Blackbird *Agelaius phoeniceus*. *Journal of Avian Biology* 40:263–270.
- POWELL, L. A., AND M. L. KNUTSON. 2006. A productivity model for parasitized, multi-brooded songbirds. *Condor* 108:292–300.
- R DEVELOPMENT CORE TEAM. 2006. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- ROBERTS, S. D., AND W. F. PORTER. 1998. Influence of temperature and precipitation on survival of Wild Turkey poults. *Journal of Wildlife Management* 62:1499–1505.
- ROBERTSON, R. J. 1972. Optimal niche space of the Red-winged Blackbird (*Agelaius phoeniceus*). I. Nesting success in marsh and upland habitat. *Canadian Journal of Zoology* 50:247–263.
- SCHMIDT, K. A. 1999. Foraging theory as a conceptual framework for studying nest predation. *Oikos* 85:151–160.
- SHAFFER, T. L. 2004. A unified approach to analyzing nest success. *Auk* 121:526–540.
- SOCKMAN, K. W. 1997. Variation in life-history traits and nest-site selection affects risk of nest predation in the California Gnatcatcher. *Auk* 114:324–332.
- U. S. FISH AND WILDLIFE SERVICE [ONLINE]. 2005. National Wetlands Inventory. <<http://www.fws.gov/wetlands/>> (May 2005).
- WESTNEAT, D. 1992. Nesting synchrony by female Red-winged Blackbirds—effects on predation and breeding success. *Ecology* 73:2284–2294.
- WOOD, S. 2001. Partially specified ecological models. *Ecological Monographs* 71:1–25.
- WOOD, S. N. 2006. Generalized additive models: an introduction with R. Chapman and Hall, London.
- WOOD, S. N., AND N. H. AUGUSTIN. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling* 157:157–177.
- YASUKAWA, K., AND W. SEARCY. 1995. Red-winged Blackbird (*Agelaius phoeniceus*), no. 184. In A. Poole and F. Gill [EDS.], *The birds of North America*. Academy of Natural Sciences, Philadelphia.
- YOUNG, H. 1963. Age-specific mortality in the eggs and nestlings of blackbirds. *Auk* 80:145–155.
- ZUUR, A., E. IENO, N. WALKER, A. SAVELIEV, AND G. SMITH. 2007. *Mixed-effects models and extensions in ecology with R*. Springer, New York.