

## THE ANALYSIS OF COVARIATES IN MULTI-FATE MARKOV CHAIN NEST-FAILURE MODELS

MATTHEW A. ETTERTSON, BRIAN OLSEN, AND RUSSELL GREENBERG

*Abstract.* In this manuscript we show how covariates may be included in Markov chain nest-failure models and illustrate this method using nest-monitoring data for Coastal Plain Swamp Sparrows (*Melospiza georgiana nigrescens*) from Woodland Beach Wildlife Area, Delaware. First, we explore hypotheses for nest failure as a single event class, which is the converse of modeling covariates to survival. We then generalize to consider separate covariates to two classes of nest failure – predation and flooding. Temporal variability, both within and between years, was the most important factor for describing daily nest failure probabilities, though percent cover around the nest also received strong support. The Markov chain estimators for a single class of failure are likely to be similar to other generalizations of the original Mayfield estimator. The estimators for modeling two or more classes of failure should prove useful, but must be employed with caution. They are sensitive to nest-fate classification errors and they can lead to a proliferation of models, which could result in over-fitting.

*Key Words:* competing risks, covariates, Mayfield Markov chain, *Melospiza georgiana nigrescens*, nest survival.

### EL ANÁLISIS DE COVARIANTES EN MODELOS MULTI DESTINO MARKOV DE FRACASO DE NIDO EN CADENA

*Resumen.* En el presente manuscrito mostramos de qué manera las covariantes pueden ser incluidas en modelos Markov de fracaso de nido en cadena, y también ilustramos este método utilizando datos de monitoreo de nido para los Gorriónes Pantaneros (*Melospiza georgiana nigrescens*) del Área Silvestre Woodland Beach, en Delaware. Primero exploramos hipótesis para fracaso de nido, como clase de evento separado, el cual es inverso al modelaje de covariantes para la sobrevivencia. Por ello generalizamos para considerar separar covariantes en dos clases de fracaso de nido – depredación e inundamiento. La variabilidad temporal durante y entre los años, fue el factor más importante para describir las probabilidades de fracaso de nido diarias, sin embargo, el porcentaje de cobertura alrededor del nido también recibió soporte fuerte. Los estimadores de cadena Markov por una clase separada de fracaso suelen ser similares a otras generalizaciones del estimador original Mayfield. Los estimadores para modelar dos o más clases de fracaso deberían probar utilidad, sin embargo deben ser empleados con cautela. Son sensibles a errores de clasificación de destino de nido y pueden dirigir hacia la proliferación de modelos, lo cual podría resultar en un exceso en el ajuste.

Nest-survival analysis has developed beyond simple survival estimation. Current methodologies now allow scientists to hypothesize and model sources (e.g., ecological and natural history) of variation in nest survivorship (Natarajan and McCulloch 1999, Dinsmore et al. 2002, Rotella et al. 2004, Shaffer 2004a). Historically, studies of daily nest survival have sought to explain nest failure, focusing on predation as the major cause. Thus, modeling daily survival as functions of covariates identifies important correlates to the absence of the event(s) of interest. When failure is the simple complement of survival, then the approaches are conversely equivalent and the appropriate inference is easy to make. However, when more than one cause of nest failure is present, covariate models of survival may identify models that are difficult to interpret as to their importance for a given cause of failure. In such cases, if researchers are in a position to unambiguously determine the fate of nests, more insight may

be gained by modeling the different causes of failure separately.

Recently, Etterson and Bennett (2005) introduced a simple non-stationary Markov chain likelihood estimator for daily survival that allows incorporation of age-specific transition probabilities (hatching and fledging) in nest survival modeling. This Mayfield-Markov chain can be further generalized to incorporate multiple categories of nest failure while relaxing the requirement that nests are visited daily (Etterson et al., in press). This formulation is ideal for considering multiple simultaneous risks to nests because, when iterated, it correctly adjusts the probabilities of failure due to one cause conditional on failure due to another cause not occurring. The need for such discounting methods, typically referred to as competing risks, has long been recognized in human demography and actuarial science (Chiang 1968). In ecology, Royama (1981) and Carey (1989) have analyzed competing risks

in insect demography using multiple decrement life-table analyses. Below we show how Markov chain models of competing risks may be applied toward a greater understanding of cause-specific avian nest failure by incorporating covariates thought to influence risk of nest failure.

Coastal Plain Swamp Sparrows (*Melospiza georgiana nigrescens*) breed in tidal marshes and, along with other tidal-marsh breeding birds, face two major challenges to successful reproduction (Greenberg et al., 2006)—predation and inundation due to tidal and storm-caused flooding. Nest failures from these two causes account for >95% of the total nest loss in Coastal Plain Swamp Sparrows (Greenberg et al., 2006). Predation and loss to flooding are likely to select for different nest-placement strategies, and therefore a trade-off may exist between behaviors that help protect against one factor or the other. For example, nests placed higher in the vegetation may help reduce the chance of flood loss, but at the same time increase the vulnerability of the nest to aerial predators. Therefore, it is quite plausible that in this subspecies the two major causes of failure may be negatively correlated via important covariates, if, to extend the above example, construction of the nest lower in the vegetation or over areas of deeper water deters nest predators. If such a trade-off exists, then female sparrows are faced with an optimization problem in where they place their nests.

Based on extensive studies of nest location and phenology, the basic natural history of nest placement can be summarized. Coastal Plain Swamp Sparrows are most common in high marsh (at or above the mean high-tide line). In this zone they tend to anchor their nests on shrubs or reed-like grasses at a fairly consistent height (approximately 30 cm above the substrate) where they can be covered in tussocks of salt hay (*Spartina patens*). Nests are found disproportionately in areas of high surface heterogeneity where water wells up forming moats around the nest plant. The salt hay cover dies back in the winter and re-grows relatively slowly in the spring. The nesting season is relatively long, beginning in mid-May and ending in late July to mid-August. Nest cover increases between mid-May (when the first nests are constructed) and the summer months (R. Greenberg, unpubl. data).

In this manuscript we use nest-monitoring data for Coastal Plain Swamp Sparrows to show how the Markov chain models can be adapted to incorporate age-, time-, and nest-specific covariates when estimating daily failure probabilities.

#### MARKOV CHAIN NEST-SURVIVAL MODELS

Before describing how to incorporate covariates into the Markov chain model, we briefly review the previously published Markov chain formulations, emphasizing the known limitations of those models. As in previous publications, our development of the likelihood functions will provide the kernel of the likelihood for an arbitrary exposure interval bounded by two visits on which the state of the nest was determined. The likelihood for a sequence of observations on a single nest or a collection of nests is generated by taking the product of the likelihoods over all such intervals. The simplest Markov chain model is:

$$\mathbf{M} = \begin{bmatrix} s & 0 \\ 1-s & 1 \end{bmatrix} \quad (1)$$

where  $s$  = daily probability of survival. The likelihood of an observation beginning in state  $\mathbf{X}_n$ , lasting  $d_n$  days, and ending in state  $\mathbf{X}_{n+1}$  is:

$$l(s | \mathbf{X}_n, \mathbf{X}_{n+1}, d_n) \propto \mathbf{X}_{n+1}^T \mathbf{M}^{d_n} \mathbf{X}_n \quad (2)$$

where  $n$  indexes the sequential visits to the nest,  $T$  is the transpose operator, and  $\mathbf{X}_n$  and  $\mathbf{X}_{n+1}$  are column vectors describing the observed states of the nest (Etterson and Bennett 2005):

$$\mathbf{X}^T = [1 \ 0] \leftrightarrow \text{nest is still active}$$

$$\mathbf{X}^T = [0 \ 1] \leftrightarrow \text{nest has failed}$$

The estimator (2) is closely related to the original Mayfield (1961, 1975) estimator and formulations by Johnson (1979), Hensler and Nichols (1981), and Bart and Robson (1982).

Etterson and Bennett (2005) extended (1) to incorporate stage-specific survival (Stanley 2000, 2004a) and transition (hatching and fledging) probabilities:

$$\mathbf{M}(a, b) = \begin{bmatrix} \sim H(a)s_1 & 0 & 0 & 0 \\ H(a)s_1 & \sim F(b)s_2 & 0 & 0 \\ 0 & F(b)s_2 & 1 & 0 \\ 1-s_1[\sim H(a)+H(a)] & 1-s_2[\sim F(b)+F(b)] & 0 & 1 \end{bmatrix} \quad (3)$$

In (3),  $s_1$  = daily probability of survival during laying and incubation,  $s_2$  = daily probability of survival during the nestling phase,  $a$  = age (in days) since the first egg was laid,  $b$  = age (in days) since hatching,  $H(a)$  is the probability of hatching at age  $a$ , and  $F(b)$  is the probability of fledging at age  $b$  ( $\sim H(a)$  and  $\sim F(b)$  are the probabilities of not hatching or fledging at the respective ages). With equation (3) there are four corresponding state vectors:

$\mathbf{X}^T = [1 \ 0 \ 0 \ 0] \leftrightarrow$  eggs present, but not yet hatched,  
 $\mathbf{X}^T = [0 \ 1 \ 0 \ 0] \leftrightarrow$  nestlings have hatched, but not yet fledged,  
 $\mathbf{X}^T = [0 \ 0 \ 1 \ 0] \leftrightarrow$  nestlings fledged, and  
 $\mathbf{X}^T = [0 \ 0 \ 0 \ 1] \leftrightarrow$  nest failed.

While simple in theory, equation (3) is difficult to apply empirically for several reasons. First, the hatching and fledging probabilities are typically not known because the distributions observed in field data are joint probabilities of survival and hatching or fledging (Etterson and Bennett 2005). Second, the use of transition probabilities requires knowledge of age, and the likelihood based on (3) is sensitive to errors in age-estimation, especially when temporal heterogeneity occurs in survival probability (Etterson and Bennett 2006). Third, equation (3) presumes that the state of a nest is determined unambiguously at each visit. This assumption may not be true for states three and four, i.e., when the nest is scored as either fledged or failed if determination of fate was made upon finding the nest empty and the nestlings were sufficiently developed to have fledged during the interval. Stanley (2004b) recommended discarding such observations, in which case (3) could be simplified to:

$$\mathbf{M}(a) = \begin{bmatrix} \sim H(a)s_1 & 0 & 0 \\ H(a)s_1 & s_2 & 0 \\ 1 - s_1[\sim H(a) + H(a)] & 1 - s_2 & 1 \end{bmatrix} \quad (4)$$

Equation (1) can also be extended to incorporate multiple classes of failure:

$$\mathbf{M} = \begin{bmatrix} s & 0 & 0 \\ m_p & 1 & 0 \\ m_f & 0 & 1 \end{bmatrix} \quad (5)$$

subject to the constraints  $0 < s, m_p, m_f < 1$  and  $s + m_p + m_f = 1$ . Under this formulation, there are three state vectors:

$\mathbf{X}^T = [1 \ 0 \ 0] \leftrightarrow$  nest active,  
 $\mathbf{X}^T = [0 \ 1 \ 0] \leftrightarrow$  nest failed due to cause 'p' (predation in our example), and  
 $\mathbf{X}^T = [0 \ 0 \ 1] \leftrightarrow$  nest failed due to cause 'f' (flooding in our example).

The likelihood function incorporating (5) is:

$$l(m_p, m_f | \mathbf{X}_n, \mathbf{X}_{n+1}, d_n) \propto \mathbf{X}_{n+1}^T \mathbf{M}^{d_n} \mathbf{X}_n'$$

where  $m_p$  = daily probability of failure due to cause 'p',  $m_f$  = daily probability of failure due to cause 'f', and other terms are as defined above.

With this manuscript we combine (4) and (5) into a Markov chain with temporal heterogeneity and multiple causes of failure. We apply the resulting model to Coastal Plain Swamp Sparrow data with two main objectives. First, we develop and demonstrate methods for analysis of covariates in ecologically interesting models using the Markov chain formulation. To begin we demonstrate an application in which a single class of failure is modeled as a function of covariates to produce results similar to other current methods (Dinsmore et al. 2002, Shaffer 2004a). Next, we re-analyze the data in the first example, considering two classes of failure (predation and flooding) and use the results to discuss potential benefits and pitfalls of such analyses. Our second objective was to perform preliminary analyses of available Coastal Plain Swamp Sparrow data to help focus the allocation of ongoing field efforts for understanding the breeding ecology of this unique subspecies.

## METHODS

A Markov chain incorporating temporal heterogeneity and two states of failure can be formulated as equation (6). The failure probabilities ( $m_p(t)$ ,  $m_f(t)$ ,  $m_p(b,t)$ ,  $m_f(b,t)$ ) are expressed as functions of time ( $t$  = Julian date relative to 1 May in our example), and age ( $b$  = age of nestlings in days since hatching). Because eggs were neither floated nor candled, we did not have reliable knowledge of age of eggs for most nests unless they hatched, so we chose not to model age-specific failure probabilities for eggs. Thus equation (6) specifies age- and time-specific failure for nestlings, but only time-specific failure for eggs. In a more general formulation (Etterson and Bennett 2005) the treatment of age ( $b$ ) must handle cases in which age surpasses the maximum empirical fledging age, but this was not necessary here because the data were truncated prior to the minimum fledging age, after Stanley (2004a).

$$\mathbf{M}(a, b, t) = \begin{bmatrix} \sim H(a)(1 - m_p(t) - m_f(t)) & 0 & 0 & 0 \\ H(a)(1 - m_p(t) - m_f(t)) & (1 - m_p(b,t) - m_f(b,t)) & 0 & 0 \\ [\sim H(a) + H(a)]m_p(t) & m_p(b,t) & 1 & 0 \\ [\sim H(a) + H(a)]m_f(t) & m_f(b,t) & 0 & 1 \end{bmatrix} \quad (6)$$

To express the failure probabilities as functions of covariates, we used the multinomial logit:

$$m_i = \frac{\exp(\beta_i Y_{in})}{1 + \sum_i \exp(\beta_i Y_{in})}$$

where  $\beta_i$  is a row-vector of structural parameters and  $Y_{in}$  is a column-vector linking the covariates for observation  $n$  (on nest  $k$ , not subscripted) to failure due to fate  $i$ . In all models presented below, the first element of  $Y_{in}$  is reserved for a global intercept. Then, the likelihood of an arbitrary observation with initial state  $\mathbf{X}_n$  and final state  $\mathbf{X}_{n+1}$  and two states of failure can be written:

$$l(\beta_p, \beta_f | t_n, a_n, b_n, \mathbf{X}_n, \mathbf{X}_{n+1}, d_n, Y_{pn}, Y_{fn}) \propto \mathbf{X}_{n+1}^T \left[ \prod_{d=1}^{d_n} \mathbf{M}(a_n + d, b_n + d, t_n + d) \right] \mathbf{X}_n \quad (7)$$

As above, the likelihood over all observations on all nests is the product of the likelihoods of each observation on each nest.

For the Swamp Sparrow models considered below, we found the maximum likelihood estimates (MLEs) of the  $\beta_i$  by numerically maximizing equation (9) using Matlab 7.04 (Mathworks 2004). All continuous covariates were standardized to improve convergence. Using the value of the likelihood function at the MLEs and formulae for effective sample size provided by Rotella et al. (2004), we compared models using Akaike's information criterion corrected for small sample sizes (AIC<sub>c</sub>) and associated model weights (Burnham and Anderson 2002). Following the recommendation of Stanley (2004b) we censored all observations for which the nest was found empty and the nestlings in the nest could have been old enough to fledge ( $\geq 8$  days since hatching, a conservative estimate) to avoid misclassification of success versus failure. Classification of failed nests into failure classes is described below.

#### STUDY SITE AND FIELD METHODS

The data used in this manuscript are from an ongoing study of Coastal Plain Swamp Sparrows in two ~15 ha plots on the State of Delaware's Woodland Beach Wildlife Area. The plots are found on upland tidal salt marsh in a matrix of farmland and wetland forest along the Smyrna River, and they represent a wide range of mid-Atlantic marsh vegetation, Coastal Plain Swamp Sparrow densities, and flooding periodicity.

Nests were discovered primarily using nest-departure calls (Greenberg 2003). After discovery, nests were monitored every 2–3 d (2002) or almost daily (2003–2005) until failure or fledging. Most failed nests failed due to predation, with evidence ranging from observations of the predation event to broken eggshells, and torn up nests. However, some nests are known to have failed due to inundation during exceptionally high tides combined with storm surges in 2004 and 2005. These tides typically occur at night during a full or new moon and cause the synchronized failure of a subset of nests (or all active nests in 2004) with identical failure evidence. Flooded nest-sites show high water marks on the vegetation above the nest; eggs typically are floated out of the dish but otherwise unharmed; and chicks are killed without external evidence of injury. Abandonment was rare (six nests in the data analyzed below), and all such nests were combined in the class of depredated nests.

Covariates included in our models included a wide array of factors (Table 1) from vegetation characteristics, to spatial arrangement, to specific descriptions of nest placement. All represent possible influences on nest survival and are accompanied by specific hypotheses as to their effect. Some measures (e.g., nest height) attempt to explain flooding, while others attempt to explain predation (e.g., percent cover) and others may impact both (e.g., hummock index). Not all covariates were measured at all nests so the analyses presented here include the largest subset of nests for which all covariates were measured (Table 1).

#### SWAMP SPARROW NEST-FAILURE MODELS

We modeled Coastal Plain Swamp Sparrow nest failure as a single class of failure using equation (4) generalized to include temporal and age heterogeneity. Daily failure probability was linked to covariates using a binomial logit. We included an intercept-only model for historical reasons, and to see how traditional estimators (Bart and Robson 1982) would perform over the same data. The Markov chain for this model is equivalent to equation (1) above. For the intercept only model, we computed 95% confidence intervals around the failure rate on the logit scale and report them on the probability scale. More realistically, we suspected that temporal heterogeneity would be important and we modeled it in two ways. First, we modeled variation due to year as a classification variable. The Markov chain for this model is:

TABLE 1. COVARIATES USED FOR SWAMP SPARROW ANALYSES.

Abbreviation	Name	Definition	Hypothesized relationship with failure
dte	Julian date	Chronological date after 1 May in each respective year	Positive (typical relationship in open-nesting passerines).
yr	Year	—	Random.
cov	Percent cover	Percent of nest hidden by vegetation when viewed from above	Negative.
ht	Nest height	Distance from lip of nest dish to the substrate	Negative for flooding, positive for predation.
pat5	Patens	Percent of ground cover within 5 m occupied by <i>Spartina patens</i> , which is the primary nest material, and often used for cover	Negative.
iva5	Iva	Percent of ground cover within 5 m occupied by <i>Iva frutescens</i> , which is a common nest anchor and often used for cover	Negative.
salt	Salinity	Salinity (parts per thousand) within 1 m of nest: salinity tolerance in Swamp Sparrows is low compared with other salt marsh passerines	Positive.
hmk	Hummock index	Difference between the water depth below the nest and at the deepest point within 1 m of the nest at high tide	Negative for predation, positive for flooding.
edg	Distance to edge	Straight-line distance from each nest to the nearest terrestrial habitat, which is wetland forest in both plots.	Negative for predation, positive for flooding.
stg	Developmental stage	Phase of nesting cycle, divided into two stages: laying and incubation vs. nestling	No prediction.
age	Nestling age	Days post-hatch	No prediction.



$$\mathbf{M} = \begin{bmatrix} 1 - m\{\text{yr}\} & 0 \\ m\{\text{yr}\} & 1 \end{bmatrix} \quad (8)$$

with 2002 effects lumped within the global intercept and yr functioning as an indicator variable with three levels (corresponding to 2003, 2004, and 2005; Table 1). Second, we modeled variation due to year (again as a class variable), but with an additional slope parameter, constrained to be equal across all years, describing changes in failure rate with Julian date. The choice to use a single slope parameter across years was made both for ecological reasons (to determine whether a general trend of increasing versus decreasing probability of failure occurs as the season progresses) and for reasons of parsimony (separate slopes would require the estimation of three additional parameters). The Markov chain for this model is:

$$\mathbf{M} = \begin{bmatrix} 1 - m\{\text{yr} + \text{dte}\} & 0 \\ m\{\text{yr} + \text{dte}\} & 1 \end{bmatrix} \quad (9)$$

with yr defined as above, and dte indicating a single slope parameter constrained to be equal across years (Table 1).

All models except the intercept only model contained one of the two above versions of temporal effects. The two temporal effects models were considered individually and were also crossed with seven additional ecological covariates (above, Table 1), each added as a single additional main effect. An example of the Markov chain for one of these models (the percent cover model) could be written as:

$$\mathbf{M} = \begin{bmatrix} 1 - m\{\text{yr} + \text{dte} + \text{cov}\} & 0 \\ m\{\text{yr} + \text{dte} + \text{cov}\} & 1 \end{bmatrix} \quad (10)$$

The two temporal-effects models were also combined with each of two models of nest development: a stage-specific model and a model that included age of nestlings nested within the stage effect. The Markov chain for the model of failure as a function of age of nestlings nested

within the stage effect can be expressed in equation (11).

This resulted in 20 models plus the intercept only model, giving a total of 21 models. All models included a global-intercept parameter.

To model Coastal Plain Swamp Sparrow nest failure in two categories we used the above 20 models (excluding the intercept-only model) as models of nest predation. To save space, we do not present the Markov chains for each model, though the Markov chain for the most complicated model is presented at the end of this paragraph. Because the sample size for flooding was small, we did not include temporal effects within season. However, we did wish to explore whether the placement of flooded nests might place them at greater risk of failure to flooding. Thus we chose three covariates related to placement of the nests: distance to the edge of the marsh, nest height above ground, and the hummock index (Table 1). We combined the three placement variables with a year variable and an intercept-only model to produce five basic models of inundation probability—intercept only, year, distance to edge, height, and hummock index (Table 1). Each of these five flooding models was combined with the 20 predation models described above to give 100 models. Covariates were linked to the failure probabilities via structural parameters using the multinomial logit. The Markov chain for the temporal and age-specific predation model combined with the yearly flooding model is expressed in equation (12). Because of the short monitoring interval employed, we assumed that all nests that failed, having last been observed in incubation, had not hatched prior to failure, and, for nests that hatch, hatch date is inferred from the age of nestlings. In this case, the function  $H(a)$  is greatly simplified because it is set to 1 on hatch date and zero otherwise (Etterson and Bennett 2005). Again, for historical reasons, we include an intercept-only model (i.e., equation 5 above), this time containing two parameters: an intercept for each category of failure. Thus, we had 101 models for the two causes of failure.

$$\mathbf{M} = \begin{bmatrix} \sim H(a)[1 - m\{\text{yr} + \text{dte}\}] & 0 & 0 \\ H(a)[1 - m\{\text{yr} + \text{dte}\}] & 1 - m\{\text{yr} + \text{dte} + \text{age}(\text{stg})\} & 0 \\ m\{\text{yr} + \text{dte}\} & m\{\text{yr} + \text{dte} + \text{age}(\text{stg})\} & 1 \end{bmatrix} \quad (11)$$

$$\mathbf{M} = \begin{bmatrix} \sim H(a)(1 - m_p\{\text{yr} + \text{dte}\} - m_f\{\text{yr}\}) & 0 & 0 & 0 \\ H(a)(1 - m_p\{\text{yr} + \text{dte}\} - m_f\{\text{yr}\}) & (1 - m_p\{\text{yr} + \text{dte} + \text{age}(\text{stg})\} - m_f\{\text{yr}\}) & 0 & 0 \\ [\sim H(a) + H(a)]m_p\{\text{yr} + \text{dte}\} & m_p\{\text{yr} + \text{dte} + \text{age}(\text{stg})\} & 1 & 0 \\ [\sim H(a) + H(a)]m_f\{\text{yr}\} & m_f\{\text{yr}\} & 0 & 1 \end{bmatrix} \quad (12)$$

## RESULTS

Of 476 Swamp Sparrow nests discovered between 2002 and 2005, survival data and the full complement of seven ecological covariates were collected at 192. Of these, 110 were depredated, 63 fledged young, 13 failed due to flooding, and six were abandoned. The earliest observed active nest was on 12 May 2004 and the latest active nest was last seen active on 30 August 2005. For the 192 nests analyzed here, the mean interval between nest visits was  $1.12 \pm 0.054$  days and all observations together accounted for an effective sample size of 1,697 exposure days.

When Swamp Sparrow failure was modeled as a single fate, the best model included annual variation, seasonal variation, and percent cover around the nest (Table 2). All of the models containing seasonal variation in addition to annual variation scored higher than models with simple annual variation (Table 2). The best model containing simple annual variation (i.e., no seasonal effects) also contained percent cover (Table 2). Both models that included percent cover estimated an increasing probability of failure with increasing percent cover around the nest. For the intercept-only model, the estimated constant daily failure probability was 0.077, with upper and lower 95% confidence intervals estimated as (0.065–0.091). However, it was by far the worst of the 21 models, scoring over 17  $AIC_c$  units worse than the best model (Table 2).

When the same data were modeled using two classes of failure, the best model of predation was also the annual variation, seasonal variation, and percent cover model (Table 3). The best model of flooding failure was the intercept-only model, though model uncertainty with respect to flooding is large as can be seen in the relatively similar performance provided by the five best models (Table 3), all of which differ in their parameterization of flooding effects. As above, the two parameter intercept-only model was the worst model considered, scoring 16.8  $AIC_c$  units worse than the best model. It gave an estimated constant daily failure probability due to predation of 0.069 (0.057–0.082) and an estimated daily failure probability due to flooding of 0.009 (0.006–0.016). As above, all models containing percent cover as a covariate predicted increasing probability of failure with increasing cover.

All models of seasonal variation predicted increasing failure probabilities later in the season, with as much as a three-fold difference over the course of the season (Fig. 1). Removing the effects of flooding had little impact on the shape of the failure curve for the best models of failure (Fig. 1), with the single-fate curve being 10–15% higher, depending on year. None of the covariates aside from percent cover received much support either in the single-failure-class models or in the predation models in the dual-class models. Neither developmental stage, nor nestling age received much support in either the single- or dual-class estimators.

TABLE 2.  $AIC_c$  STATISTICS FOR 21 MODELS OF SWAMP SPARROW NEST SUCCESS, TREATING FAILURE AS A SINGLE CLASS.

Model <sup>a</sup>	Parameters	$AIC_c$	$\Delta AIC_c$	Weight
$m\{yr + dte + cov\}$	6	869.63	0.00	0.49
$m\{yr + dte\}$	5	873.02	3.40	0.09
$m\{yr + dte + edg\}$	6	873.47	3.84	0.07
$m\{yr + dte + hmk\}$	6	873.92	4.29	0.06
$m\{yr + dte + ht\}$	6	874.61	4.98	0.04
$m\{yr + dte + salt\}$	6	874.68	5.05	0.04
$m\{yr + dte + age(stg)\}$	6	874.76	5.13	0.04
$m\{yr + dte + pat5\}$	6	874.87	5.24	0.04
$m\{yr + dte + stg\}$	6	874.88	5.25	0.04
$m\{yr + dte + iva5\}$	6	874.92	5.30	0.03
$m\{yr + cov\}$	5	875.83	6.20	0.02
$m\{yr\}$	4	877.4	7.78	0.01
$m\{yr + edg\}$	5	877.44	7.82	0.01
$m\{yr + ht\}$	5	877.78	8.15	0.01
$m\{yr + salt\}$	5	878.61	8.99	0.01
$m\{yr + hmk\}$	5	878.77	9.15	0.01
$m\{yr + pat5\}$	5	879.13	9.50	0.00
$m\{yr + age(stg)\}$	5	879.28	9.66	0.00
$m\{yr + iva5\}$	5	879.31	9.69	0.00
$m\{yr + stg\}$	5	879.36	9.73	0.00
$m\{.\}$	1	886.88	17.25	0.00

<sup>a</sup> In single class of failure models  $m\{...\}$  indicates model of overall failure for which covariates are contained within brackets. Covariate abbreviations follow Table 1.

TABLE 3. TEN BEST DUAL-FAILURE-CLASS MODELS FOR COASTAL PLAIN SWAMP SPARROW NEST FAILURE.

Model <sup>a</sup>	Parameters	AIC <sub>c</sub>	ΔAIC <sub>c</sub>	Weight
$m_p\{yr + dte + cov\}m_f\{.\}$	7	965.55	0.00	0.21
$m_p\{yr + dte + cov\}m_f\{edg\}$	8	966.52	0.96	0.13
$m_p\{yr + dte + cov\}m_f\{yr\}$	8	966.74	1.19	0.12
$m_p\{yr + dte + cov\}m_f\{ht\}$	8	967.38	1.82	0.09
$m_p\{yr + dte + cov\}m_f\{hmk\}$	8	967.57	2.02	0.08
$m_p\{yr + dte\}m_f\{.\}$	6	969.91	4.36	0.02
$m_p\{yr + dte + edg\}m_f\{.\}$	7	970.51	4.95	0.02
$m_p\{yr + dte + hmk\}m_f\{.\}$	7	970.56	5.00	0.02
$m_p\{yr + dte\}m_f\{edg\}$	7	970.89	5.34	0.01
$m_p\{yr + dte\}m_f\{yr\}$	7	971.09	5.54	0.01

<sup>a</sup> In dual class of failure models  $m_p\{...\}$  indicates model of predation for which covariates are contained within brackets.  $m_f\{...\}$  indicates models of flooding probability.

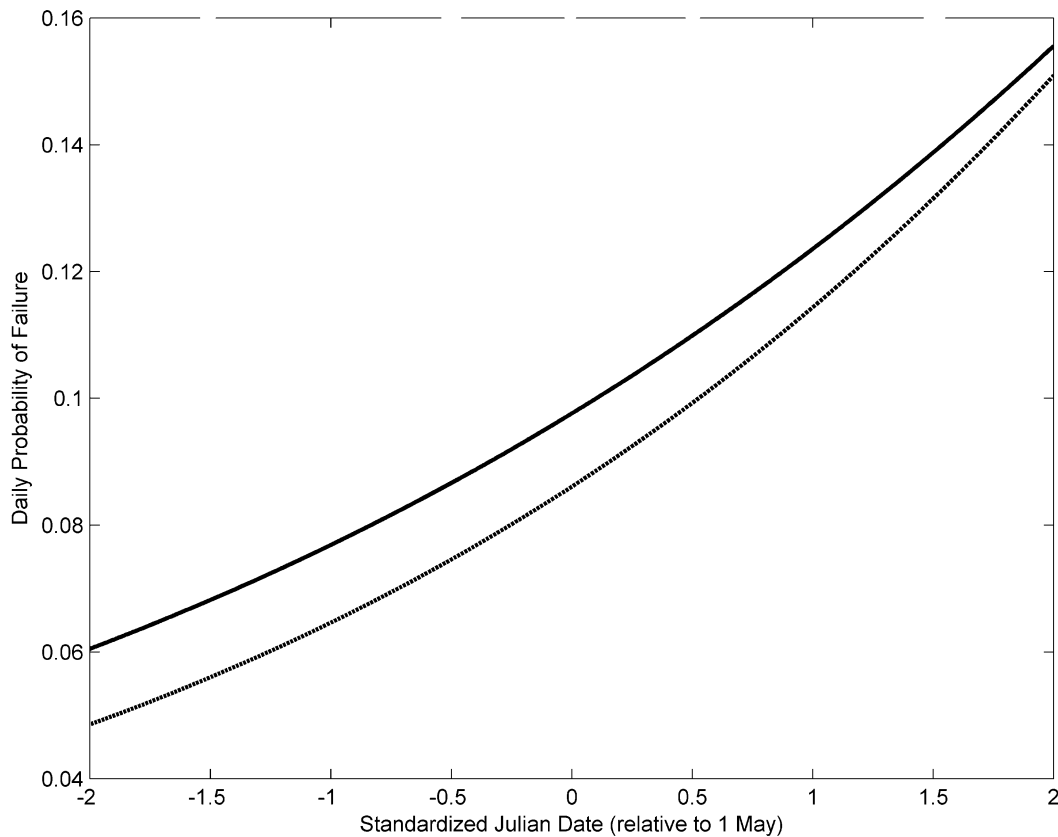


FIGURE 1. Representative patterns of within season nest failure showing change in daily failure rate with Julian date (standardized) in 2005 for best single-fate model and for predation component of best dual-fate model. For these analyses, modeling flooding failure separately simply reduces overall daily failure attributable to predation, without changing the seasonal pattern of failure.

## DISCUSSION

Clearly, temporal variation, both within and between years, is critical for understanding Coastal Plain Swamp Sparrow nest failure. This temporal variation makes it difficult for us to

decide conclusively the importance of the other ecological effects measured because of the relatively large number of estimated parameters that were required to describe temporal variation. With more data it may be possible to estimate such effects while controlling for temporal



variation, as for example, percent cover appears to be an important factor explaining swamp sparrow nest failure, albeit in a surprising direction.

The increased predation rate with percent cover may be partly confounded with the within-season effect because available cover in this habitat increases as the season progresses. However, the selection of both variables as important for describing overall fate, and predation in particular, suggests that each contributes unique information to understanding nest failure. The increase in failure rates with increasing cover may be attributable to the foraging habits of the small mammals that constitute the primary nest predators. Heavy predation on these small mammals from both nocturnal and diurnal raptors may prevent serious foraging bouts in open areas. Avian nest predation in the study plots is low. Neither Blue Jays (*Cyanocitta cristata*) nor Fish Crows (*Corvus ossifragus*) are seen on the plots and all observed avian nest-predation events occurred as holes in eggs or nestling necks on territories that neighbored those of Marsh Wrens (*Cistothorus palustris*). Territory location may therefore play a larger role than nest camouflage in these failure events. Relatively few nests, however, completely lack cover, and the question remains why sparrows would place such a large proportion of their nests in deep cover if it is a reliable predictor of failure. Thus, it seems likely that nest cover is correlated with an additional factor that we failed to identify.

Other analyses of age-specific nest failure have shown age to be an important predictor of failure probability. For example, Dinsmore et al. (2002) showed that the daily probability of failure decreased with age in Mountain Plovers (*Charadrius montanus*). Conversely, Natarajan and McCulloch (1999:558) showed increasing daily probability of failure with age for another wetland passerine, the Red-winged Blackbird (*Agelaius phoeniceus*). In our analyses, the fitted age-specific model suggested decreasing probability of failure with age, though the model was not competitive with the better models described above (Table 2). However, age of eggs was not included as a covariate because many nests were discovered after clutch completion and we did not attempt to determine age through candling or flotation. Thus, we were only able to assess age effects for nestlings. Stanley (2004b) cautioned that the right-censoring we performed here might result in failure to detect age-specific heterogeneity if failure probabilities change substantially just prior to fledging. We do not think such effects were present in our data for two reasons. First, the

terminal exposure intervals that were censored were usually only 1 d, and never more than 2 d, so little information would have been lost from censoring those observations. Second, in other analyses where we did not right-censor the data, the age-specific models also did not perform well. In fact, the minimal truncation had little effect on model performance for any of the models we considered.

The ability to include covariates in the Markov chain models will make this class of models more useful to ecologists and managers for determining causes of nest failure. The close relationship between the basic likelihood (equations 1 and 2) of the Markov chain model and that of Johnson (1979) and Bart and Robson (1982) suggests that the Markov chain model will give similar results to other generalizations of those estimators (Dinsmore et al. 2002, Shaffer 2004a), when similar covariate models are analyzed. However, there are some important differences. The Markov chain model has not yet been extended to allow the incorporation of random effects, as can be done in the SAS implementation of the logistic-exposure model (Rotella et al. 2004, Shaffer 2004a). In contrast, the incorporation of stage-transition probabilities (Etterson and Bennett 2005) cannot currently be done in MARK or in the logistic-exposure model, though Stanley's (2000, 2004a) model does provide this capability using SAS. Similarly, while multiple-fate nest-survival models can be implemented in SAS (Thompson and Burhans 2004), they require the assumption that the dates on which failure events occur are known precisely. The Markov chain model we present here relaxes that assumption. Other differences will occur due to the way in which time- and age-specific covariates are handled. In the analyses presented here, we allowed time and age to progress within exposure intervals, as did Dinsmore et al. (2002), whereas Shaffer (2004a) used the mean age of a nest during an exposure interval as a covariate to the entire interval. In practice, given the short monitoring interval for these data, the two methods would be virtually identical.

As suggested by Greenberg et al. (2006), flooding does not appear to be a major cause of nest failure for Coastal Plain Swamp Sparrows. Furthermore, little evidence exists in our data supporting hypotheses that flooding risk is related to any of the nest-placement variables we considered, though this result may be due to small sample size for flooding. The lack of correlates with ecological variables may also be due to the nature of flooding events, which, although rare, tend to destroy most or all active nests. Finally, the removal of flooding effects, in

these data, has little effect on the shape of the resulting model for predation effects. Thus, it would appear that fitting the more complex model offers little benefit for these analyses, but this may not always be the case. Furthermore, even in this case, we have confirmed our previous belief that flooding effects are of relatively minor importance.

For now, the choice to use a multi-fate model such as the one we present here must ultimately be a subjective one, depending on the goals of the study, the interpretation one wishes to place in the failure parameters, and the confidence with which nests are classified. If the goal of research is to assess the importance of nest predation and to explore the ecological conditions that result in increased predation pressure, then it is sensible to model fates separately, assuming they are known with confidence. Our analyses suggest that the risk of failure due to predation is 10–15% lower than the overall risk of failure for Coastal Plain Swamp Sparrows. In fact, this difference is actually >10–15% because some nests that failed due to abandonment were also classified with depredated nests.

The latter observation highlights the need for the development of methods for controlling and reducing probabilities of classification error. To date, we know very little about the effects of such error because they cannot be estimated under typical monitoring protocols. In our opinion, this remains the largest obstacle to the use of competing risks nest-survival analyses. Some authors have begun to consider the importance and interpretation of ancillary evidence at nests (Manolis et al. 2000), and we think this is a promising direction for further research.

Another potential drawback to multiple-fate nest-failure modeling is the proliferation of models that can occur when a small set of models for one fate is combined with a small set of models for another fate. In our example, we had relatively simple sets of models for each fate, yet ended up with 101 models! If we had considered more main effects and interactions

between ecological variables we could easily have conceived >1,000 plausible models describing the two fates. Thus, careful a priori consideration of models and objectives will be absolutely necessary to avoid over fitting.

The data we used for this demonstration were ideal. The short monitoring interval resulted in very well-characterized hatching dates; the unequivocal evidence available for determining whether a nest was destroyed due to flooding allowed us to apply the dual-fate model without much risk of classification error (but note that we still confounded abandonment and nest predation). In most cases data will be somewhat less ideal and the decision of appropriate modeling framework will require judgment on the part of the researcher. Our models also currently require a greater degree of attention to programming details than, for example, the nest-survival module in MARK (Dinsmore et al. 2002) or the logistic-exposure method (Shaffer 2004a) implemented in SAS. Nevertheless we believe the Markov chain framework will continue to prove itself a flexible template for the development of sophisticated nest failure models and for testing interesting ecological hypotheses about avian nest failure.

#### ACKNOWLEDGMENTS

This manuscript benefited greatly from comments by J. D. Nichols and J. F. Bromaghin. Field work conducted as part of this study was funded by grants from the Delaware Division of Fish and Wildlife, the Maryland Ornithological Society, the Smithsonian Institution's Abbot Fund, the Washington Biologists' Field Club, the Explorer's Club: Washington D.C. Group, the Eastern Bird Banding Association, Virginia Tech, and a Smithsonian Institution Graduate Fellowship. Field work was conducted in part by the following technicians: J. Wang, M. Powell, K. Murabito, J. Kolts, K. Callaway, A. Wessel, J. Adamson, B. Augustine, B. Beas, K. Heyden, and J. Felch.