7-2-2007

# Incorporating Results of Avian Toxicity Tests into a Model of Annual Reproductive Success

Richard S. Bennett*

Matthew A. Etterson

Integrated Environmental Assessment and Management — Volume 3, Number 4—pp. 498–507


# Incorporating Results of Avian Toxicity Tests into a Model of Annual Reproductive Success

*Richard S Bennett* and Matthew A Etterson*

*US Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, Minnesota 55804*

*(Received 18 April 2007; Accepted 2 July 2007)*



Original Research

## ABSTRACT

Modeling the effects of pesticide exposure on avian populations requires knowledge of how the pesticide changes survival and fecundity rates for the population. Although avian reproduction tests are the primary source of information on reproductive effects in the pesticide risk assessment process, current tests cannot provide a direct estimate of the effects of a pesticide on fecundity rates. We present a mathematical model that integrates information on specific types of effects from reproduction tests with information on avian life history parameters, the timing of pesticide applications, and the temporal pattern of pesticide exposure levels to estimate pesticide effects on annual reproductive success. The model demonstration follows nesting success of females in no-pesticide or pesticide-exposed populations through a breeding season to estimate the mean number of successful broods per female. We demonstrate the model by simulating populations of a songbird exposed to 1 of 2 hypothetical pesticides during a breeding season. Finally, we discuss several issues for improving the quantitative estimation of annual reproductive success.

**Keywords:** Avian reproduction test  Annual reproductive success  Pesticide  Population-level assessment  Markov chain models

## INTRODUCTION

An ongoing challenge in the field of ecotoxicology is to improve methods for characterizing chemical risks to wildlife populations (Kendall 1994; Sample et al. 2000). One approach for understanding these risks is the use of population models that integrate toxicity information on reproduction and survival endpoints to estimate chemical effects on population growth rates or other population-level endpoints. This requires estimates of demographic parameters, such as fecundity and survival rates, and a mathematical model of the species' life history. Because such parameter estimates often are derived from previously published studies, it is important that they are evaluated to understand the quality of the data, including potential biases incorporated during the process of data collection and analysis and the precision with which they are estimated (Etterson and Bennett 2006a). Although there are many sources of bias in model parameters (e.g., Etterson and Bennett 2006a), our focus in this manuscript is on the currency mismatch between laboratory-estimated pesticide effects and field estimates of parameters pertaining to avian annual reproductive success (ARS). This is a particularly vexing problem in wildlife ecotoxicology (Sample et al. 2000; Bennett and Etterson 2006; Etterson and Bennett 2006a).

In a typical avian population model, fecundity rates usually are expressed as the number of female young produced per adult female per year, also known as the ARS rate. However, it is uncommon for ARS to be reported in the literature directly (Murray 2000). Usually it is estimated from a combination of intermediate parameters that are more easily estimated, such as clutch size, nest success rate, number of broods per season, and proportion of adult females in a breeding population (Nagy and Holmes 2004). This practice of compiling an estimate of ARS from component parameters makes such estimates vulnerable to error propagation because each component parameter will, in turn, be subject to both bias and sampling error (Etterson and Bennett 2006a).

Estimating the effects of a chemical stressor on a population requires understanding how exposure to that stressor quantitatively changes survival, reproduction, or both, either by incorporating stressor–response relationships directly into the model or by modifying the vital rates corresponding to a specific exposure estimate (Bartell et al. 2003). If the stressor is a pesticide, data from laboratory toxicity tests need to be translated into estimates of effects on survival or fecundity rates. To do this, we must 1st understand the information provided by the laboratory tests.

In the current US Environmental Protection Agency (USEPA) pesticide risk assessment process, a pair of laboratory avian reproduction tests with mallards (*Anas platyrhynchos*) and northern bobwhites (*Colinus virginianus*) are conducted to evaluate how dietary pesticide exposure affects a standard suite of reproduction endpoints (USEPA 1996). The results of these tests are used in calculating risk quotients (RQs) by comparing the lowest reported no-observed-effect concentration (NOEC) from any of the measured endpoints with estimates of the maximum dietary exposure expected for a given application rate. As a screening tool, RQs are compared with an established regulatory level of concern to categorize the potential for unacceptable risk. Risk quotients can be used to identify the environmental concentration above which adverse effects to avian reproduction could occur, but they cannot determine the probability or magnitude of potential reproductive effects.

Ideally, test endpoints could be used as estimators for how ARS changes as a function of pesticide exposure. For example, one of the endpoints in the existing avian reproduction test is

* To whom correspondence may be addressed: bennett.rick@epa.gov
Published on the Web 7/17/2007.

the mean number of 14-d-old chicks produced per pair, which seems to be expressed in a currency comparable to ARS. However, there are several reasons why none of the endpoints in the avian reproduction test is sufficient on its own to directly estimate changes in ARS (Bennett and Etterson 2006). First, several types of potential pesticide effects on ARS either cannot be evaluated directly in the avian reproduction test or are inadequately evaluated, such as effects on the parent's ability to defend territories, incubate, or provision young (Mineau et al. 1994; Mineau 2005). Second, although existing tests typically are conducted with 2 to 4 pesticide treatment groups plus a control group, they usually are not designed to quantitatively define dose–response relationships for important endpoints. Third, although the test simulates a single, extended laying period, ARS represents the cumulative productivity for an entire breeding season, which could include raising multiple broods and renesting after nest losses. Finally, the reproduction test simulates a continuous pesticide exposure over 20 weeks, whereas wild birds usually experience variation in pesticide exposure because pesticide applications can occur at various times throughout the breeding season. Consequently, avian reproduction test results do not reflect the types of events during a breeding season that ultimately affect ARS. For these reasons, the reproduction test alone does not provide information in a currency that is equivalent to ARS (Bennett and Etterson 2006), but it does provide data on specific pesticide-related effects that can be combined with information on life history and pesticide application timing to translate information from these multiple sources into an estimate of change in ARS.

An alternative conceptual framework for interpreting the results of avian reproduction tests was recently proposed by Bennett et al. (2005). Briefly, it involves linking the types of effects that can occur during each phase of a bird's reproductive cycle (e.g., pair formation, egg laying, incubation, brood rearing) to selected surrogate endpoints from avian toxicity tests and relate that to the estimated exposure during each phase under a given pesticide use scenario (Bennett et al. 2005). Because the great majority of avian reproduction tests do not provide dose–response information for surrogate variables, the approach is based on a series of phase-specific decision points for determining whether the nest fails or continues. In the framework proposed by Bennett et al. (2005), if the estimated exposure is less than the NOEC for surrogate endpoints at each phase, the nest continues without disruption. However, if exposure exceeds the surrogate endpoint NOEC, the nest is assumed to have failed, and the female might be able to renest if conditions permit. The simulated performance of females in relation to the timing of pesticide applications is then modeled over the course of a full breeding season (Bennett et al. 2005).

The framework described above identifies 3 categories of effects resulting from direct exposures that could occur: 1) Effects on adult behavior and reproductive performance from external exposure (e.g., dietary), 2) effects on juvenile growth and survival from external exposure, and 3) effects on juvenile growth and survival from in ovo exposure. Some potential effects have direct corollary measurements from the reproduction test (e.g., percent hatchability related to in ovo exposure), whereas other effects have more indirect surrogate measures (e.g., the use of change in adult body weight during the prelaying period as an indicator of overall parental well being and behavioral effects). Some effects, such as juvenile toxicity from external exposures, have no surrogate from the reproduction test because chicks are not exposed to treated diets. However, information from other toxicity tests might be useful.

To explore the potential for using this conceptual framework to estimate pesticide effects on avian populations, 3 case studies have been developed to estimate the risks to UK populations of skylarks (*Alauda arvensis*), an omnivorous songbird, exposed to a hypothetical pesticide application. Shore et al. (2005) calculated deterministic RQs specific to each reproductive phase on the basis of the timing of pesticide application relative to the timing of nest initiation. Roelofs et al. (2005) presented a method for calculating the probability of exceeding phase-specific risk thresholds. They also demonstrated a simulation model for estimating the number of new adults produced in a breeding population under various exposure scenarios. The simulation output can be used as an estimate of the reduction in ARS from a specific pesticide use scenario compared with a no-pesticide scenario. Topping et al. (2005) demonstrated how the phase-specific decision points from Bennett et al. (2005) can be integrated into a spatially explicit individual-based landscape model to assess pesticide exposure on the population size of skylarks.

The models used by Roelofs et al. (2005) and Topping et al. (2005) estimate pesticide-related declines in ARS with use of the extensive life history information available for skylarks. In fact, these papers integrated information on skylark life history into their models that went beyond the basic framework proposed by Bennett et al. (2005). For example, Roelofs et al. (2005) demonstrated how additional parameters could be incorporated to estimate partial brood success, rather than nest success versus failure. However, most risk assessments include other species of concern for which there is considerably less life history data available. In this paper, we propose a flexible mathematical model for implementing the conceptual framework of Bennett et al. (2005) to estimate pesticide effects on ARS that can be applied to a broad range of species and can be modified to incorporate either sparse or abundant life history data.

## A GENERAL MATHEMATICAL MODEL FOR ESTIMATING CHEMICAL EFFECTS ON ANNUAL REPRODUCTIVE SUCCESS

We propose a modeling approach for implementing the conceptual framework described in Bennett et al. (2005) that builds on more than 40 y of avian nest–survival modeling in the ornithological literature. Etterson and Bennett (2005) showed that a simple Markov chain model is equivalent to the well-known Mayfield (1961, 1975) model when similar assumptions are imposed and unifies many current formulations of nest survival estimation models (e.g., Johnson 1979; Hensler and Nichols 1981; Bart and Robson 1982; Dinsmore et al. 2002; Shaffer 2004). The ability to reformulate the problem as a Markov chain immediately suggested ways to improve our modeling and estimation of avian nesting success. Two important extensions to traditional nest survival estimation have already been made with the Markov chain formulation: The ability to incorporate heterogeneous ages at transition (e.g., hatching and fledging; Etterson and Bennett 2005, 2006b) and the ability to easily incorporate multiple classes of nest failure (Etterson, Nagy, et al. 2007; Etterson, Olsen, et al. 2007).

Our ARS model can be understood best as a compound Markov chain, in which each phase of a female's breeding season is represented with the use of a phase-specific Markov chain. Each phase begins and ends with a well-defined observable event (although these are not necessarily unique). Parameters within the phase-specific Markov chain describe the probabilities of transition among states within the phase and probabilities of exiting the phase. For each phase, the subject of the Markov chain is a breeding female who takes on different attributes (occupies different states) depending on the path she follows through the chain. Below, we illustrate with an example that describes the egg-laying phase for a female of a species that lays a clutch ranging from 2 to 5 eggs.

The initiation of the egg-laying phase is conditioned upon a female laying the 1st egg. Once that event occurs, many other events can occur. For example, the nest could now be subject to some nonzero background probability of failure for ecological reasons (e.g., nest predation, adverse weather). Depending on species ecology, a typical female might reinitiate pair formation, reinitiate the copulation phase with the same mate, or abandon her breeding season altogether. Similarly, the nest is susceptible to pesticide exposure, which could cause nest failure, again forcing the female to revert to an earlier phase (e.g., pair formation, copulation), or cessation of breeding altogether. Assuming none of these events occur, the female can lay additional eggs (1 egg/d in the example below) until she completes her clutch. Finally, the female will begin incubating the eggs, which defines the end of the laying phase and the beginning of the incubation phase. Thus, in this example, 3 well-defined events can terminate the egg-laying phase (ecological nest failure, pesticide exposure exceeding the NOEC, or the onset of incubation). In practice, there could be other terminating events too, such as death of the laying female.

The paragraph above informally describes a set of transitions that must be formalized in a mathematical model. Suppose, for example, that the number of eggs laid (clutch size) could range from 2 to 5 eggs. Then, a female could transition into 8 possible states during the laying phase, with 5 possible clutch sizes plus 3 termination states. For this example, we treat the latter 3 states as absorbing states, but more generally they represent separate Markov chains, each with its own transition matrix. In the matrix below, these states are represented in the following order (proceeding from top to bottom for rows and from left to right for columns): 1) Incubation begun, 2) 5 eggs, 3) 4 eggs, 4) 3 eggs, 5) 2 eggs, 6) 1 egg, 7) ecological failure, and 8) NOEC exceeded.

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ s & 0 & 0 & 0 & 0 & 0 & m_b & m_e \\ sg_i(4) & s[1-g_i(4)] & 0 & 0 & 0 & 0 & m_b & m_e \\ sg_i(3) & 0 & s[1-g_i(3)] & 0 & 0 & 0 & m_b & m_e \\ sg_i(2) & 0 & 0 & s[1-g_i(2)] & 0 & 0 & m_b & m_e \\ 0 & 0 & 0 & 0 & s & 0 & m_b & m_e \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The parameters of $\boldsymbol{L}$ are $m_b$, the background daily probability of nest failure; $m_e$, the daily probability of failure because of pesticide exposure; $s$, which is $1 - m_b - m_e$; and $g_i(a)$, the probability that the female enters the incubation phase with a clutch size of $a$ eggs.

To understand how the above model works, it will be useful to closely examine the rows of $\boldsymbol{L}$. First, the individual elements of $\boldsymbol{L}$ are the probabilities of transition from a given state (row) to any other state (column). Second, note that all rows sum to unity, indicating that each can be considered the conditional probability distribution of states for the female given her state the previous day. Consider, for example, the 3rd row, which has 4 nonzero entries. Thus, a female with 4 eggs in her nest can transition to any of 4 states: Incubating (column 1), 5 eggs (column 2), ecological failure (column 7), or failure because of pesticide exposure (column 8).

The 1st 2 transitions are conditional on not undergoing either of the latter 2 transitions (i.e., they are conditional on not failing, whether because of ecological reasons or pesticide exposure: $s = 1 - m_b - m_e$). They also depend on the probability of initiating incubation (or not) after laying only 4 eggs, $g_i(4)$. Note that the transition probability is assumed to be unity after the female lays her 5th egg (because this is the maximum clutch size for this species). More generally, we can think of the transition probability function, $g_i(a)$ as describing the cumulative probability of initiating incubation at a given clutch size, given that the female has not already initiated incubation. However, in empirical applications, the observed transition probabilities must also be adjusted for survival because only surviving nests are used to estimate the empirical distributions (Etterson and Bennett 2005, 2006b).

The laying cycle described in Equation 1 would be inappropriate for many species (e.g., any species that lays more than 5 eggs). Equation 1 also makes several assumptions about the life history of this hypothetical species that might not be obvious. First, it assumes that once the 1st egg is laid the female continues laying 1 egg/d until she initiates incubation. Second, it assumes that incubation begins with the final egg. However, in many species, incubation begins with the penultimate egg and, in others, with the 1st egg. Thus, even limiting our consideration to the egg-laying phase, we see that a general mathematical model for avian productivity must provide considerable flexibility to incorporate the diverse life histories of North American birds. The Markov chain provides this flexibility by adjusting the sizes of the phase-specific matrices, as well as the location and definition of parameters within the matrices. Furthermore, it does so in a way that is consistent with the manner in which daily survival parameters are estimated in field applications (Etterson and Bennett 2005).

The above flexibility extends to the way in which Equation 1 is incorporated into an overall Markov chain for a full breeding attempt (Equation 2).

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{N}_T & 0 & 0 & 0 & \boldsymbol{M}_N \\ \boldsymbol{G}_{IN} & \boldsymbol{I}_T & 0 & 0 & \boldsymbol{M}_I \\ 0 & \boldsymbol{G}_{LI} & \boldsymbol{L}_T & 0 & \boldsymbol{M}_L \\ 0 & 0 & \boldsymbol{G}_{CL} & \boldsymbol{C}_T & \boldsymbol{M}_C \\ 0 & 0 & 0 & \boldsymbol{G}_{PC} & \boldsymbol{P}_T \end{bmatrix} \quad (2)$$

The entries are now individual matrices, with the diagonal entries corresponding to the phases of nesting cycle ($\boldsymbol{P}$ = pair formation, $\boldsymbol{C}$ = copulation, $\boldsymbol{L}$ = egg laying, $\boldsymbol{I}$ = incubation, and $\boldsymbol{N}$ = nestling; after Bennett et al. 2005). The subscript T indicates that only the transient states are kept in the phase-specific submatrices. Failure parameters are removed to the matrices $\boldsymbol{M}_i$ and phase transition parameters are removed to the matrices $\boldsymbol{G}_i$. For example, if the matrix $\boldsymbol{L}_T$, were derived from a phase-specific Markov chain like that of Equation 1, then these matrices follow.

$$\boldsymbol{L}_{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ s[1-g_i(4)] & 0 & 0 & 0 & 0 \\ 0 & s[1-g_i(3)] & 0 & 0 & 0 \\ 0 & 0 & s[1-g_i(2)] & 0 & 0 \\ 0 & 0 & 0 & s & 0 \end{bmatrix},$$

$$\boldsymbol{M}_{\mathrm{L}} = \begin{bmatrix} 0 & \cdots & 0 & m_{\mathrm{p}} & m_{\mathrm{e}} \\ 0 & \cdots & 0 & m_{\mathrm{p}} & m_{\mathrm{e}} \\ 0 & \cdots & 0 & m_{\mathrm{p}} & m_{\mathrm{e}} \\ 0 & \cdots & 0 & m_{\mathrm{p}} & m_{\mathrm{e}} \\ 0 & \cdots & 0 & m_{\mathrm{p}} & m_{\mathrm{e}} \end{bmatrix}, \quad \boldsymbol{G}_{\mathrm{LI}} = \begin{bmatrix} 0 & \cdots & 0 & s \\ 0 & \cdots & 0 & sg_i(4) \\ 0 & \cdots & 0 & sg_i(3) \\ 0 & \cdots & 0 & sg_i(2) \\ 0 & \cdots & 0 & 0 \end{bmatrix}.$$

The number of columns in $\boldsymbol{M}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{LI}}$ will depend on the dimensions of $\boldsymbol{P}_{\mathrm{T}}$ and $\boldsymbol{I}_{\mathrm{T}}$, respectively.

Again, the overall structure of Equation 2 depends on the life history of the bird species to which the model is applied. In Equation 2, all females must undergo pair formation after nest failure. Other species might attempt to breed again with the same mate after a failed attempt, in which case they could (after an appropriate waiting period) transition directly to the copulation phase after nest failure (whether because of pesticides or background ecological causes). Clearly, much more complicated models could be constructed with this general template. However, we believe for most species available data will justify a simple model at best.

Equation 2 still would not be appropriate for a species in which some females have more than 1 successful nest attempt in a given season. Once again the added complexity can be handled by compounding the model so that the individual nest attempts ($\boldsymbol{B}$, from Eqn. 2) form the building blocks of yet another Markov chain. In the example below, a female may lay up to 3 successful clutches in a breeding season,

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{B}_3 & 0 & 0 \\ \boldsymbol{R}_{23} & \boldsymbol{B}_2 & 0 \\ 0 & \boldsymbol{R}_{12} & \boldsymbol{B}_1 \end{bmatrix} \tag{3}$$

where the submatrices along the diagonal (the $\boldsymbol{B}_i$) are the nest attempts and the subdiagonal submatrices ($\boldsymbol{R}_{ij}$) house the parameters governing the probabilities of initiating additional breeding attempt $j$ conditional on successful attempt $i$.

### *Benefits of the Markov chain approach*

We believe the Markov chain approach to modeling avian ARS has several important benefits. First, the matrix structure is very flexible, providing a common structure for describing diverse life histories because parameters, breeding phases, and breeding attempts can all be rearranged in a straightforward fashion to suit a given model species. Similarly, flexibility is great within a species to develop a model with any degree of complexity that our knowledge and data can justify. Second, the specific form of the matrices, as Markov chains, forces an explicit probabilistic structure for simulating ARS, for which the computer programming is easily verified. Third, the Markov chain approach treats avian nesting in a manner consistent with the way in which nesting parameters are typically estimated in the field, thus reducing the potential for currency mismatch between the model and available ecological data for the species of interest. Fourth, the adoption of the Markov chain model makes a large toolbox of analytical methods that have been developed for the analysis of Markov chains available to risk assessors.

### *Similarities and differences with other models of ARS*

On the whole, our model is similar in concept and in performance to that of Roelofs et al. (2005), although the mathematical model as described above is quite different. In its current form, our model for pesticide-induced nest failure includes only threshold responses, in keeping with the nature of RQs. It allows multiple nesting attempts per season by a given female.

The largest differences between the version of the model we demonstrate below and that of Roelofs et al. (2005) are related to our focus on the female as the subject of the model, rather than individual eggs or chicks. In our model, the number of eggs in a nest is an attribute of a given female, not a model subject that survives or fails independently. In Roelofs et al. (2005), both the performance of females and the survival of individual eggs or juveniles is tracked to estimate the number of juveniles produced per female per season. Our Markov chain model could incorporate variable clutch size and estimate partial brood reduction given appropriate parameters describing the individual probabilities of egg or juvenile survival. However, for many species–pesticide combinations, the paucity of life history information, the pesticide dose–response data on individual egg or juvenile survival, or both would mean that model estimates of the number of juveniles produced per female would be highly dependent on assumptions with limited empirical basis. Consequently, following Bennett et al. (2005), in our demonstration of the model we assume only that nest attempts either succeed or fail and estimate the number of successful broods per female per season.

Other potential model complexities that we have omitted from our demonstration include tracking juvenile survival postfledging and random variation in model parameters. We acknowledge that these are important processes, but again we doubt that sufficient information will be available for most species to accurately model such complexity. For species in which suitable information on random variation in model parameters exists, this information can be incorporated into the Markov chain model. Postfledging survival of juveniles might be modeled more appropriately as a separate process. Also, whereas fledging in altricial species might be marked by departure of juveniles from the nest, the parental responsibilities of females might or might not end at fledging. This becomes a factor in determining at what point the female is free to initiate an additional nesting attempt after a successful brood. For precocial species the definition of a fledgling might not be obvious and might need to be addressed on a species-by-species basis.

Finally, in the following demonstration, we do not address interspecies variability. Shore et al. (2005) demonstrated 1 approach by dividing no-observed-effect levels (NOELs) of phase-specific endpoints by interspecies extrapolation factors. Luttik et al. (2005) presented a proposal for estimating interspecies extrapolation factors for use in assessments of avian reproduction effects. Roelofs et al. (2005) demonstrated how extrapolation factors can be incorporated into estimates of ARS.

## DEMONSTRATION OF THE MARKOV CHAIN MODEL

To demonstrate the use of a Markov chain model to estimate ARS in species with sparse data, we have attempted to closely adhere to the conceptual framework laid out by Bennett et al. (2005). The Markov chain model used in the

demonstration below is simpler than the example matrix in Equation 1 because we did not model variation in clutch size, but rather assumed a fixed clutch size, and did not estimate the number of juveniles produced per nest, but rather estimated only the mean number of successful broods per female per season, where a successful brood is defined as a nesting attempt that produces fledglings. To estimate ARS, the number of successful broods is multiplied by the mean number of fledglings per successful nest, an endpoint that is often presented in the literature.

Below we use a Markov chain model to simulate ARS for a population of insectivorous songbirds exposed to 1 of 2 hypothetical pesticides. Pesticide X primarily affects hatchability of eggs and juvenile survival and is similar to bioaccumulative pesticides. Pesticide Y primarily affects egg production and is similar to an organophosphorous pesticide. We assume that this species has a breeding season of 45 d (from arrival of the 1st bird in the territory to initiation of the last nest attempt) and commonly attempts to raise 2 broods if time and resources permit. We also assume a fixed clutch size of 4 and that ova fully develop in 3 d.

We used the following parameters, adapted from skylark parameters used in Roelofs et al. (2005), to describe the avian breeding cycle. Birds were assumed to arrive on the breeding grounds according to a Poisson distribution with $\lambda = 5$. After arrival, birds proceeded through pair formation, copulation, egg laying, incubation, and nestling phases of fixed durations (Table 1). To these active phases we also added 3 inactive phases that described the minimum time required to initiate a new attempt after successfully fledging, failure because of background causes, or failure because of pesticide exposure (Table 1). After the 1st egg was laid, each nest was subject to a fixed daily background probability of failure ($m = 0.03$).

For this demonstration, we used the same phase-specific surrogate endpoints as described in Bennett et al. (2005), with the exception that our model did not consider the postfledging phase (Table 2). The NOELs for both hypothetical pesticides, expressed as daily ingested dose (i.e., mg/kg/d) rather than NOECs on the basis of dietary concentration, are given in Table 2. For each of the 2 hypothetical chemicals, we simulated 4 exposure scenarios, with application dates on day 0, 15, 30, or 45 after birds start to arrive on the breeding ground. In other words, the day 0 application occurred on the 1st day that nests would be initiated, whereas the day 45 application occurred on the last day that new nests could be initiated. For demonstration purposes, our model did not employ a realistic exposure model. Rather, we simulated exposure by assuming a maximum exposure dose of 150 mg/kg/d at application to both adults and juveniles consuming an invertebrate diet, after which exposure decayed according to a fixed half-life of 5 d. For both pesticides, the lowest NOEL was 50 mg/kg/d and the maximum exposure was 150 mg/kg/d. In the current risk assessment process, both pesticides would have overall RQs of 3. With the use of a 5-d half-life, pesticide exposure would exceed the 50 mg/kg/d NOEL for approximately 8 d.

Similar to Roelofs et al. (2005), as each simulated female in the breeding population progressed through various breeding phases, each phase-specific surrogate endpoint was compared with an estimate of pesticide exposure on each day appropriate for that endpoint (Figure 1). For this demonstration, we used the same exposure durations for estimating exposure doses as described in Bennett et al. (2005). Most surrogate endpoints are compared with the estimated single-day exposure on each day of a breeding phase (Figure 1). Endpoints affected by in ovo exposure (i.e., percent hatch and percent juvenile survival) are compared with a time-weighted average (TWA) exposure over the duration of ova development (in this demonstration, we assumed 3 d). Similarly, the endpoint from a 5-d juvenile dietary test is compared with a 5-d TWA.

Each day during the breeding season, phase-specific RQs (i.e., exposure estimate divided by NOEL for surrogate endpoint) were calculated and compared with a regulatory level of concern of 1. If any RQ was greater than 1, there was a presumption of unacceptable adverse effects, and the nesting attempt was assumed to fail (Figure 1). If no RQ was greater than 1 throughout the breeding phases, the nesting attempt succeeded or failed according to the background daily failure rate (0.03). At the end of either a successful or failed nesting attempt, a female can begin a new nesting attempt after a minimum recovery period (see Table 1) and if time remains in the breeding season (Figure 1). However, a new nesting attempt cannot be initiated as long as pesticide exposure levels exceeded the NOEL for changes in adult body weight (Figure 1).

For each exposure scenario, we simulated 100 populations of 10,000 females each and report 3 statistics pertaining to ARS: The expected number of successful broods per female per season, the expected number of nest attempts per female, and the expected nest survival rate. To compare these results to an unexposed population, we also simulated a population of 10,000 females using the same input parameters, but with no pesticide application. This simulation will be referred to as the no-pesticide population. For the no-pesticide population, we also simulated 100 populations of 1,000 and 100,000 females each as a guide to the effect of simulated population size on the precision of simulation results.

## SIMULATION RESULTS

### *Model outputs*

Each model run created several outputs for the simulated population. In addition to the 3 statistics pertaining to ARS mentioned above, the state of each female on each day of the breeding season was output in both tabular and graphical form to allow close examination of the timing of nest failures in relation to female nesting status. In particular, the graphical output was especially useful for visualizing the effect of a pesticide or other stressor on nesting success, when comparing across model scenarios (e.g., pesticide vs no-pesticide scenarios, or alternative application dates). A companion graph showed the pesticide exposure profile over the same time course.

### *No-pesticide population*

In the absence of pesticides, each simulated female would be expected to produce about 1.06 successful nests out of 2.07 attempts, resulting in an expected overall probability of nesting success of 0.51 (Table 3). As a whole, the population of females experienced 2 pulses of offspring production over the course of the breeding season (Figure 2). Thus, for a given female, on average only 1 of the 2 attempts was successful.

### *Pesticide X, with effects on hatchability*

The effects of application of pesticide X varied considerably with application date. Of the 4 application dates, the greatest

**Table 1.** Duration of phases of breeding cycle

| Phase | Duration (d) |
|---|---|
| Arrival | Poisson (5) |
| Pair formation | 3 |
| Copulation | 3 |
| Laying | 3 |
| Incubation | 11 |
| Nestling | 8 |
| Fledged | 4 |
| Failed for ecological reasons | 4 |
| Failed because of pesticide | 7 |

effect on the population was observed with application on day 45 (i.e., the last day for initiating new nests), whereas the least effect was observed with application on day 30 (Table 3). The initial nest attempt was unaffected by applications on days 30 or 45. The difference reflected that, for renesting attempts lost after an application on day 30, females might have had time remaining in the breeding season to renest, whereas if the application were on day 45, there was no further opportunity to renest. The results highlight the difficulty in interpretation of laboratory testing results without considering the interaction between the timing of exposure and the specific effects of a chemical on avian breeding. Another important result in Table 2 was the lack of any effect of exposure to Chemical X on the total number of nesting attempts made by a given female, which was constant regardless of application date. By contrast, nest survival rate, a typical ecological index of avian reproductive output, varied consistently and proportionately with numbers of successful broods produced per female.

### *Pesticide Y, with effects on egg production*

The effects of pesticide Y also varied with application date, with the worst case again on day 45 and best on day 30. Even though nest failures because of pesticide did occur after application on day 30 (Figure 2), there was virtually no reduction in the mean number of successful broods per female (Table 3). The timing of this application was such that the population compensated for nest failures because of pesticide by increases in the number of nest attempts. As above, number of attempts per female, although a better index for ARS than for Chemical X, was still a poor indicator, whereas nest survival rate was better.

### *Comparison between pesticides and among application dates*

Simulations with both pesticides showed that the timing of pesticide applications in relation to timing of nest initiations can have a significant influence on the mean number of successful broods (Table 3), primarily because the proportion of the population in any particular sensitive breeding phase varied daily. Also, the types of effects caused by a pesticide can affect the estimate of the number of successful broods. In the comparison of pesticide X (primarily affecting hatchability) and pesticide Y (primarily affecting egg production), the critical period of pesticide exposure for both pesticides is at, or just before, egg laying, but because effects of pesticide Y are expressed immediately in the egg-laying phase, the results differ from pesticide X, for which effects are not expressed until the end of an 11-d incubation period (Figure 2). To examine the effects of pesticide exposure on population growth rates, the relative difference between no-pesticide and pesticide scenarios can be used to modify the fecundity rates used in population models.

## DISCUSSION

We have presented a flexible mathematical model for estimating the effects of pesticide exposure on reproductive success in avian populations. In many respects our model is similar in performance to that of Roelofs et al. (2005), although the model structure is quite different. Our model combines data from toxicity test endpoints representing specific types of effects with information on avian life history and timing of pesticide applications. It results in an estimate of the overall effect on the number of successful broods produced per female in a simulated population.

To estimate the ARS rate for a particular model scenario expressed as the number of young produced per female per year, the mean number of successful broods per female can be multiplied by the mean number of fledglings per successful nest, a parameter that is often reported in the literature. Because of the limited nature of data from avian reproduction

**Table 2.** Phase-specific surrogate endpoints and NOELs for 2 hypothetical chemicals

| Surrogate endpoint (expressed as NOEL) | Phase[a] | NOEL (mg/kg/d) | |
|---|---|---|---|
| | | Pesticide X | Pesticide Y |
| Adult body wt prelaying | PF, IN, NE, FL, FE, FP | 400 | 200 |
| Eggs laid | CO, LA | 400 | 50 |
| Eggshell thickness | CO, LA | 400 | 400 |
| % Fertile eggs | CO, LA | 400 | 400 |
| % Hatch | IN | 50 | 400 |
| % Survival to 14 d | NE | 100 | 400 |
| Juvenile dietary toxicity | NE | 500 | 500 |

[a] CO = copulation; FE = failed for ecological reasons; FL = fledged; FP = failed because of pesticide; IN = incubation; LA = laying; NE = nestling; PF = pair formation.

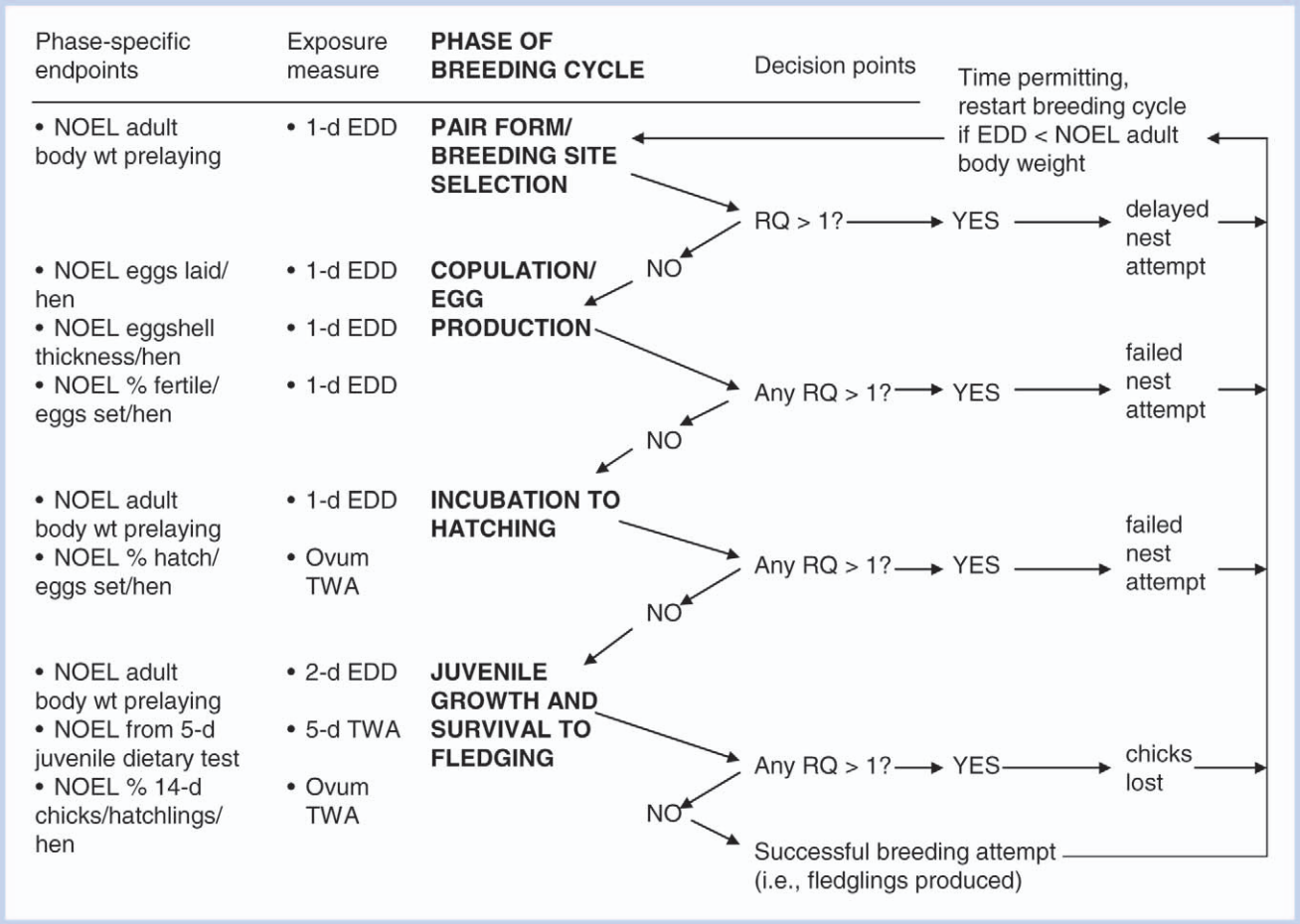


**Figure 1.** Four phases of avian breeding cycle with phase-specific toxicity endpoints and associated exposure estimates (i.e., estimated daily dose [EDD]; or time-weighted average dose [TWA]) used in risk quotients (RQs) at each decision point.

tests, phase-specific decision points are based on a success/failure determination with the use of NOELs for surrogate variables, instead of having dose–response relationships for estimating proportional responses (Bennett et al. 2005; Mineau 2005). Consequently, the estimated ARS rates for this application of the model, as well as those from Roelofs et al. (2005), should be interpreted as protective rather than predictive of an expected outcome. In other words, on the basis of the information available, we would not expect the true ARS rate to be lower. However, the true ARS rate may be higher because use of NOELs in phase-specific decisions could indicate a nest failure at exposure levels that, in reality, could produce less significant effects.

Although the approach for parsing the breeding season into a series of phase-specific decision points has limitations in the degree to which it can quantify ARS because of the nature of avian reproduction test results, we believe it is vastly superior to simply assuming that a dose–response relationship for an

**Table 3.** Simulation results

| Pesticide | Application date[a] | Successful broods per female | Nest attempts per female | Overnest success |
|---|---|---|---|---|
| None | NA | 1.06 | 2.07 | 0.51 |
| X | 0 | 0.84 | 2.07 | 0.40 |
| | 15 | 0.85 | 2.08 | 0.41 |
| | 30 | 0.90 | 2.07 | 0.44 |
| | 45 | 0.73 | 2.07 | 0.35 |
| Y | 0 | 0.97 | 2.16 | 0.45 |
| | 15 | 0.87 | 2.09 | 0.41 |
| | 30 | 1.03 | 2.28 | 0.45 |
| | 45 | 0.74 | 2.06 | 0.36 |

[a] Date is relative to arrival of 1st birds to breeding grounds. NA = not applicable.

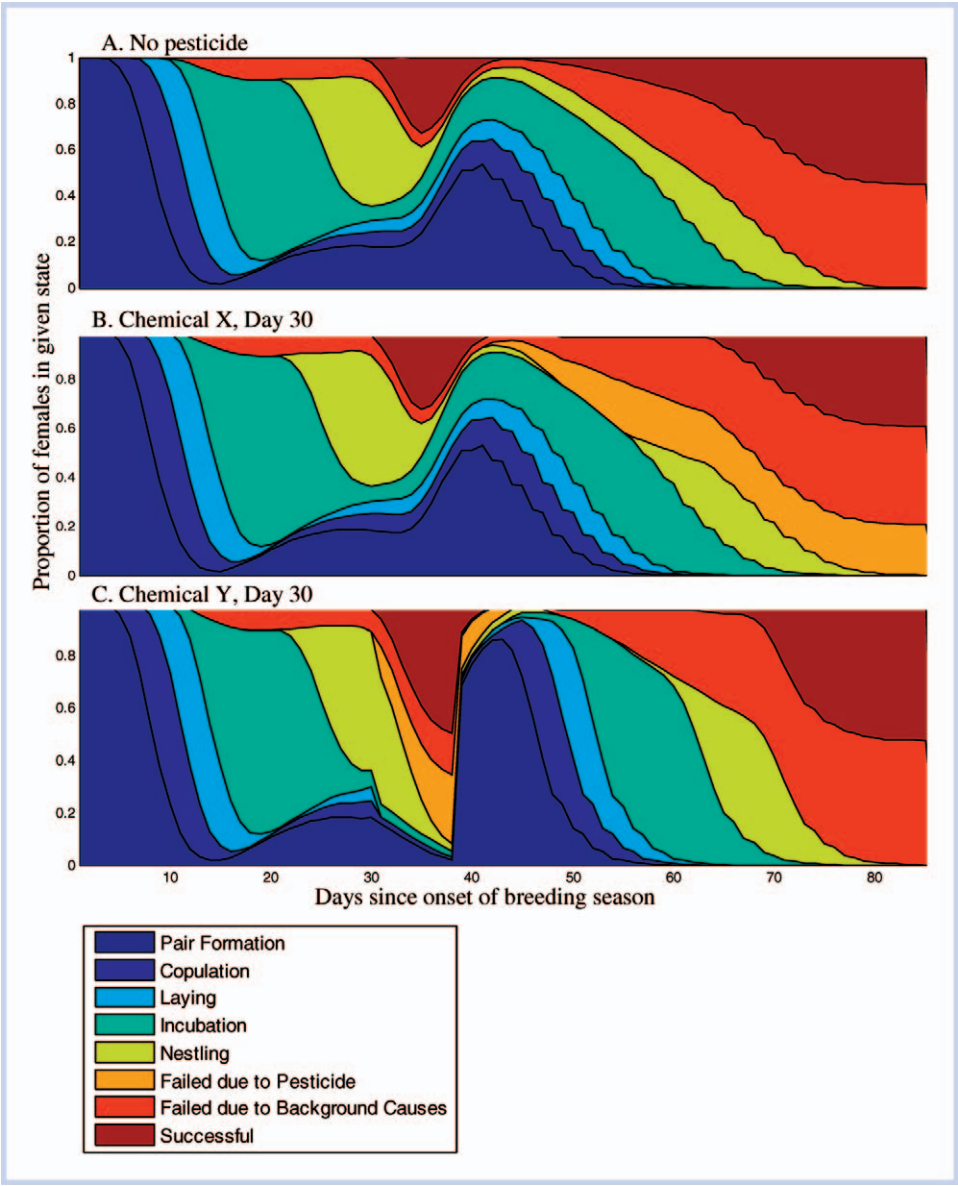


**Figure 2.** Temporal proportion of female population in each breeding phase throughout the breeding season.

endpoint (e.g., number of chicks per hen) taken directly from the test can serve as an estimate for ARS. In the model simulations above, 2 hypothetical pesticides, both considered to pose the same degree of risk by the current assessment process (i.e., RQ = 3), were observed to vary in their effects on the number of successful broods depending on the types of pesticide effects expected and the timing of applications over the course of the breeding season. Thus, this approach provided a method for dealing with a difficult source of error in current procedures for estimating pesticide effects on avian reproductive success.

However, many other sources of error exist, including the bias in the underlying ecological parameters because of violations of statistical assumptions, sample bias, and failure to account for important covariates (Etterson and Bennett 2006a). Furthermore, the simulations we presented above do not incorporate sampling error of estimated parameters into estimates of sampling error of ARS, which will have the effect of broadening the simulation distributions around the expected value. Ultimately, the structure of the modeled process (ARS) will also be subject to spatial and temporal variability, typically referred to as environmental stochasticity, which will further reduce the precision with which ARS can be predicted. Methods for handling all of these problems have been developed in the ecological modeling literature and can be incorporated into the model we describe (e.g., Morris and Doak 2002; Etterson and Bennett 2006a). The challenge will be to incorporate these methods in a way that informs us as to which management actions will be protective of the populations of interest.

In implementing the conceptual framework for estimating pesticide effects on ARS as laid out in Bennett et al. (2005), we present a generalized model that requires fewer life history parameter estimates than the models used in case studies by Roelofs et al. (2005) and Topping et al. (2005). It also relies on parameters that are typically estimated in field studies and reported in the literature. Our goal is a model that is applicable to a broad range of species with limited life history data. All 3 of these models used the same toxicity test endpoints. It is true that, in many cases, more is known about

the nature of possible effects of a pesticide on reproduction in avian species than is captured in a series of NOELs for surrogate test endpoints. When additional information exists, the Markov chain model approach is very flexible in the way it allows integration of information into a more complex model.

Adding model complexity introduces additional issues to address, however. For example, if an avian reproduction test is designed such that dose–response relationships can be quantified for important endpoints, users would want to know if that information can be used to estimate the proportion of eggs or juveniles that survive by replacing phase-specific RQs driven by NOELs, rather than assuming only nest success or failure. Although the Markov chain model can incorporate this type of information into the model, to do so can result in greatly expanded matrices with new issues emerging that require additional data or assumptions. In this example, modeling the survival of individual eggs or juveniles requires additional information that might or might not exist on the degree of correlation of survival probabilities of individuals (i.e., how does the death of 1 chick affect the survival probability of others?). The onus is on the model user to ensure that the additional complexity being designed into a model actually accomplishes the intended purpose of improving the quality of ARS estimates without compromising the level of protection afforded the population of interest.

In the above example on the use of dose–response relationships, it is important to remember that laboratory tests are providing endpoints that act as surrogates for effects occurring in wild birds. No matter how well we quantify dose–response relationships in the laboratory, these endpoints might or might not reflect field responses. To use dose–response relationships for surrogate endpoints effectively, we need to understand how the dose–response relationship of the measured surrogate endpoint quantitatively relates to the dose–response relationship for the field effect of concern (Bennett and Etterson 2006). If the concern relates to effects on hatching rate from in ovo exposure, we can be reasonably confident that the measure of hatchability in the laboratory is reflective of effects on hatchability in the field from in ovo exposure. However, with the use of prelaying adult body weight as a surrogate indicator of parental well being and behavioral changes, for example, it is quite unlikely that we will have information on how they are functionally related. Similarly, a change in the number of eggs laid in the laboratory might be used as a surrogate indicator for reduced clutch size in the field, but even though reduced egg production is commonly observed in laboratory tests, it is seldom seen in wild birds exposed to pesticides (Mineau 2005). Ultimately, to improve quantification of ARS by estimating proportional responses on nest success, additional specialized tests or field assays might be required to better understand important relationships between laboratory endpoints and field effects of concern.

The usefulness of our Markov chain model, as well as the model presented in Roelofs et al. (2005), is dependent on selecting appropriate surrogate endpoints to represent the full range of possible pesticide effects on reproduction. Bennett et al. (2005) discuss the rationale for selecting the surrogate endpoints used in this demonstration. However, it is important during an assessment of each chemical to evaluate the appropriateness of these surrogate endpoints on the basis of knowledge of the chemical and to consider additional or replacement surrogate endpoints from the avian reproduction test or other sources (e.g., field studies, nontraditional laboratory studies) where justified. For example, although a change in prelaying body weight might serve as a useful surrogate endpoint for parental well being when considering pesticides that affect food consumption rates and body weight (e.g., cholinesterase-inhibiting pesticides), it could be an inappropriate indicator for other modes of action that affect parental behavior without affecting weights (e.g., endocrine-mediated behaviors).

Additional types of effects could occur in the field for which adequate surrogate endpoints have not yet been established. The model simulations discussed above consider only effects on reproductive endpoints from direct pesticide exposure. For some pesticides, the maximum exposures after application might cause direct mortality of some females during the breeding season. This would further reduce ARS because females exposed to lethal exposures could not renest. The Markov chain model can incorporate additional decision points on the basis of estimates of pesticide-related mortality of females and modify the overall estimates of ARS to reflect the risk of breeding season mortalities. Also, current testing procedures provide little information on the potential for indirect pesticide effects (e.g., through altering food availability) to adversely affect avian reproduction, but specifically designed field studies could provide surrogate measures for including consideration of indirect effects in assessing the effects on ARS. By starting with a more complete list of the potential reproductive effects occurring in the field and examining the available evidence to determine whether, and to what extent, each of these potential effects might occur from a pesticide exposure scenario, we will have a stronger basis for characterizing risks in a population context.

## SUMMARY

At a conference on the effects of environmental contaminants on vertebrate populations and communities, Sample et al. (2000, p 241) concluded that, "perhaps the greatest obstacle preventing widespread use of population dynamics models for contaminant effects assessment is the incompatibility between commonly reported toxicological endpoints and population model inputs." Bennett et al. (2005) proposed a conceptual framework to bridge this incompatibility by translating data from the existing avian reproduction test into a currency compatible with population model inputs. We have discussed several modeling approaches developed to implement this framework. All can estimate the effect of pesticide exposure on overall reproductive success throughout the breeding season by integrating information on specific types of effects from avian toxicity tests with avian life history and timing of pesticide. Our Markov chain model was developed to provide a more generalized approach that can be applied to a broad range of species with limited life history data. In an ecological risk assessment involving a variety of avian species of concern, our modeling approach provides a means to estimate the effects of a given pesticide exposure scenario on the ARS for each species and to examine which species or life history characteristics could be at greatest risk.

*Acknowledgment*—This paper benefited greatly from review of an early draft by, and discussions with, Pierre Mineau. The authors also thank James Warren, Ed Odenkirchen, and Christopher Salice for comments on the paper and the model.

*Disclaimer*—The information in this document has been funded wholly by the USEPA. It has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

## REFERENCES


Bart J, Robson DS. 1982. Estimating survivorship when the subjects are visited periodically. *Ecology* 63:1078–1090.

Bartell SM, Pastorok RA, Akcakaya HR, Regan H, Ferson S, Mackay C. 2003. Realism and relevance of ecological models used in chemical risk assessments. *Human and Ecological Risk Assessment* 9:907–938.

Bennett RS, Dewhurst IC, Fairbrother A, Hart ADM, Hooper MJ, Leopold A, Mineau P, Mortensen SR, Shore RF, Springer TA. 2005. A new interpretation of avian and mammalian reproduction toxicity test data in ecological risk assessment. *Ecotoxicology* 14:801–815.

Bennett RS, Etterson MA. 2006. Estimating pesticide effects on fecundity rates of wild birds using current laboratory reproduction tests. *Human and Ecological Risk Assessment* 12:762–781.

Dinsmore SJ, White GC, Knopf FL. 2002. Advanced techniques for modeling avian nest survival. *Ecology* 83:3476–3488

Etterson M, Nagy L, Robinson T. 2007. Partitioning risk among different causes of nest failure. *Auk* 124:432–443.

Etterson M, Olsen B, Greenberg R. 2007. The analysis of covariates in multi-fate Markov chain nest-failure models. *Stud Avian Biol* 34:55–64.

Etterson MA, Bennett RS. 2005. Including transition probabilities in nest survival estimation: A Mayfield Markov chain. *Ecology* 86:1414–1421.

Etterson MA, Bennett RS. 2006a. On the use of published demographic data for population-level risk assessment in birds. *Human and Ecological Risk Assessment* 12:1074–1093.

Etterson MA, Bennett RS. 2006b. The effects of uncertainty about age at transition on bias in the Mayfield family of estimators. *Ecol Model* 199:253–260.

Hensler GL, Nichols JD. 1981. The Mayfield method of estimating nesting success: A model, estimators and simulation results. *Wilson Bull* 93:42–53.

Johnson DH. 1979. Estimating nest success: The Mayfield method and an alternative. *Auk* 96:651–661.

Kendall RJ. 1994. Using information derived from wildlife toxicology to model ecological effects of agricultural pesticides and other environmental contaminants on wildlife populations. In: Kendall RJ, Lacher TE Jr, editors. Wildlife toxicology and population modeling: Integrated studies of agro-ecosystems. Boca Raton (FL): Lewis. p 1–11.

Luttik R, Mineau P, Roelofs W. 2005. A review of interspecies toxicity extrapolation in birds and mammals and a proposal for long-term toxicity data. *Ecotoxicology* 14:817–832.

Mayfield HF. 1961. Nesting success calculated from exposure. *Wilson Bull* 73:255–261.

Mayfield HF. 1975. Suggestions for calculating nest success. *Wilson Bull* 87:456–466.

Mineau P. 2005. A review and analysis of study endpoints relevant to the assessment of "long term" pesticide toxicity in avian and mammalian wildlife. *Ecotoxicology* 14:775–799.

Mineau P, Boersma DC, Collins B. 1994. An analysis of avian reproduction studies submitted for pesticide registration. *Ecotoxicol Environ Saf* 29:304–329.

Morris WF, Doak DF. 2002. Quantitative conservation biology: Theory and practice of population viability analysis. Sunderland (MA): Sinauer Associates. 480 p.

Murray BG Jr. 2000. Measuring annual reproductive success in birds. *Condor* 102:470–473.

Nagy LR, Holmes RT. 2004. Factors influencing fecundity in migratory songbirds: Is nest predation the most important? *J Avian Biol* 35:487–491.

Roelofs W, Crocker DR, Shore RF, Moore DRJ, Smith GC, Akcakaya HR, Bennett RS, Chapman PF, Clook M, Crane M, Dewhurst IC, Edwards PJ, Fairbrother A, Ferson S, Fischer D, Hart ADM, Holmes M, Hooper MJ, Lavine M, Leopold A, Luttik R, Mineau P, Mortenson SR, Noble DG, O'Connor RJ, Sibly RM, Spendiff M, Springer TA, Thompson HM, Topping C. 2005. Case study part 2: Probabilistic modeling of long-term effects of pesticides on individual breeding success in birds and mammals. *Ecotoxicology* 14:895–923.

Sample BS, Rose KA, Suter GW II. 2000. Estimation of population-level effects on wildlife based on individual-level exposures: Influence of life-history strategies. In: Albers PH, Heinz GH, Ohlendorf HM, editors. Environmental contaminants and terrestrial vertebrates: Effects on populations, communities, and ecosystems. Pensacola (FL): Society of Environmental Toxicology and Chemistry. p 225–243.

Shaffer TL. 2004. A unified approach to analyzing nest success. *Auk* 121:526–540.

Shore RF, Crocker DR, Akcakaya HR, Bennett RS, Chapman PF, Clook M, Crane M, Dewhurst IC, Edwards PJ, Fairbrother A, Ferson S, Fischer D, Hart ADM, Holmes M, Hooper MJ, Lavine M, Leopold A, Luttik R, Mineau P, Moore DRJ, Mortenson SR, Noble DG, O'Connor RJ, Roelofs W, Sibly RM, Smith GC, Spendiff M, Springer TA, Thompson HM, Topping C. 2005. Case study 1. How to calculate appropriate deterministic long-term toxicity to exposure ratios (TERs) for birds and mammals. *Ecotoxicology* 14:877–893.

Topping CJ, Sibly RM, Akçakaya HR, Smith GC, Crocker DR. 2005. Risk assessment of UK skylark populations using life-history and individual-based landscape models. *Ecotoxicology* 14:925–936.

[USEPA] US Environmental Protection Agency. 1996. Avian reproduction test. ecological effects test guidelines OPPTS 850.2300. Washington DC: USEPA. EPA 712-C-96-141.