



## Modelling Insect Natality Using Splines

Brian S. Yandell; David B. Hogg

*Biometrics*, Vol. 44, No. 2. (Jun., 1988), pp. 385-395.

Stable URL:

<http://links.jstor.org/sici?sici=0006-341X%28198806%2944%3A2%3C385%3AMINUS%3E2.0.CO%3B2-E>

*Biometrics* is currently published by International Biometric Society.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ibs.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## Modelling Insect Natality Using Splines

Brian S. Yandell

Departments of Horticulture and of Statistics, 1210 West Dayton St.,  
University of Wisconsin, Madison, Wisconsin 53706, U.S.A.

and

David B. Hogg

Department of Entomology, Russell Laboratories, University of Wisconsin,  
Madison, Wisconsin 53706, U.S.A.

### SUMMARY

We present methodology for analyzing age-specific birth rates, with particular emphasis on describing variability within a cohort of like-aged individuals. We use a model structure appropriate for natality in insects and other arthropod species. Our approach separates natality into a baseline rate representative of the effects of the environment for the population and a factor that represents the unique effect of each individual in the cohort. We employ smoothing splines with the penalized likelihood methodology for estimation. We illustrate the methodology by applying it to data on leafhopper egg-laying rates.

### 1. Introduction

The age-specific natality schedule provides information of fundamental importance in animal population ecology. Conventional methods for studying population dynamics [e.g., the Leslie matrix model (Plant, 1986)] require information on age-specific natality for the population, but for the most part these approaches require only population averages and ignore any associated variability. However, there has been a growing interest in exploring the implications that variability in vital rates (birth and death) has for overall population behavior (Meyer et al., 1986; Slade and Levenson, 1982; Strong, 1986). Arthropods in general, and certain insect species in particular, exhibit considerable variation in age-specific birth rates and thus make good candidates for investigating individual variability and its effects on population dynamics.

A common approach in insect natality studies considers the cumulative number of births as a function of time for an individual as

$$N(t) = \int_0^t m(x) dx$$

with  $m(t)$  the observed birth rate. The expected birth rate,  $h(t) = Em(t)$ , is then modelled with some parametric form, such as a logistic, exponential, or Gompertz curve (cf. Madden, 1985). In contrast, we develop a nonparametric method for modelling insect natality that allows us to characterize an overall or "baseline" birth rate and to analyze individual contributions to that rate. Our starting point considers the number of eggs laid per unit time as a nonstationary Poisson process (Cinlar, 1975), and we propose to estimate the

age-specific Poisson rates by penalized maximum likelihood using spline-based generalized linear models.

The model structure we propose embodies the general characteristics of insect natality. In addition to high individual variability, these include strong dependency on age and sensitivity to environmental conditions (especially temperature). Let  $h^0(t)$  be the birth rate mediated by the common environment, the age-specific "baseline rate" common to all individuals subject to the same conditions. Furthermore, we suppose that variability introduced by individual differences over time (= age) affects the birth rate in a multiplicative fashion. Let  $a_i(t)$  be the "individual factor" that depends on unique characteristics during the life of individual  $i$ . We consider modelling the birth rate for the  $i$ th individual,  $h_i(t)$ , as

$$h_i(t) = h^0(t)a_i(t). \quad (1)$$

That is, individual-to-individual variation has a relative effect on the baseline rate. Other models, such as an additive one where individual factors have an absolute effect on the number of births, are possible but not considered here. The model structure (1) has been considered extensively for mortality rates (Cox, 1972). If individual birth rate curves  $h_i(t)$  follow the same shape as the baseline birth rate (on a log scale), then  $a_i(t)$  is constant. If not, changes in  $a_i(t)$  reflect time-dependent changes in individual natality (e.g., due to physiological changes). Taking logs in (1) and defining  $\theta_i = \log h$ ,  $\theta^0 = \log h^0$ , and  $\alpha_i = \log a_i$ , we have

$$\theta_i(t) = \theta^0(t) + \alpha_i(t). \quad (2)$$

Bartoszyński et al. (1981) and Clevenson and Zidek (1977) considered such a model for Poisson data with  $\alpha_i(t)$  assumed constant. This model is a special case of the generalized additive model (Hastie and Tibshirani, 1986).

We estimate the log baseline rate  $\theta^0(t)$  and individual factors  $\alpha_i(t)$  as smoothly changing functions of time. Our method requires the imposition of some measure of desired smoothness on the estimates through a "penalty" on overall curvature of the functions. We maximize the likelihood plus the penalty, striking a compromise between fidelity to the data (with rather bumpy estimates) and simplicity in curve shape (with straight-line estimates). This maximum penalized likelihood methodology has received growing attention recently for graphical examination of smoothly changing curves (cf. O'Sullivan, Yandell, and Raynor, 1986; Green and Yandell, 1985; Silverman, 1985; and Hastie and Tibshirani, 1986).

We outline the methodology in Section 2. We demonstrate the utility of our approach by applying it to leafhopper data sets: baseline rates are analyzed in Section 3, and individual variability is examined in Section 4. We make some conclusions and remarks in Section 5.

## 2. Splines and Penalized Likelihoods

We present here our estimation procedure for the baseline birth rate and the individual factors using the penalized likelihood method. Added detail is presented in the Appendix. We consider first a single female and for notational simplicity drop the individual subscript  $i$ . Thus, our goal is estimation of  $h(t)$  or, equivalently,  $\theta(t) = \log h(t)$ . Suppose the female is observed at  $n$  distinct times,  $0 < t_1 < \dots < t_n$ , with  $t_0 = 0$  being the time of adult emergence. During an interval between observation times, say  $t_{j-1}$  and  $t_j$ , one would expect to see  $H_j$  births, with

$$H_j = \int_{t_{j-1}}^{t_j} h(t) dt. \quad (3)$$

The number of births occurring during this interval,  $m_j$ , should have approximately a Poisson distribution with mean  $H_j$ . The best (maximum likelihood) estimate of  $H_j$  is  $m_j$ , the observed count, and the best estimate of  $h(t)$  for  $t$  between  $t_{j-1}$  and  $t_j$  is

$$\hat{h}(t) = m_j/d_j, \quad \text{with } d_j = t_j - t_{j-1}. \tag{4}$$

Unfortunately, if there are many short intervals between observations, this (nonsmoothed) estimate of birth rate  $\hat{h}(t)$  ends up being very rough.

The basic idea in our approach using smoothing splines is analogous to attaching a flexible rod to the crude estimates of  $\theta(t) = \log h(t)$  from (4) at the “nodes”  $t_j$ ,

$$\hat{\theta}(t_j) = \log(m_j/d_j), \tag{5}$$

using springs that move vertically. Strong springs force the rod (or curve) to adhere to the point estimates, whereas weak springs lead to a linear estimate. Springs with intermediate strength allow the rod to form a smooth curve. Our approach leads to a curve that approximately follows a cubic polynomial between nodes. We use maximum likelihood imposing a “penalty” on rough curves, with a large penalty corresponding to a weak spring.

For a single individual, the Poisson distribution

$$\Pr\{m_j | H_j\} = (H_j)^{m_j} \exp(-H_j) / m_j!$$

leads to the log-likelihood based on observed counts  $m_1, \dots, m_n$ , of

$$L = \log \prod_{j=1}^n \left( \frac{\Pr\{m_j | H_j\}}{\Pr\{m_j | m_j\}} \right) = \sum_{j=1}^n m_j (\log H_j - \log m_j) + \sum_{j=1}^n (m_j - H_j). \tag{6}$$

The second sum can always be made 0 by setting  $\sum H_j = \sum m_j$ , and is ignored below. If one assumes that  $h$  is “smooth,” one can allow the data to indicate what shape  $h$  has without specifying a parametric form. Approximating the integral  $H_j$  by  $h(t_j)d_j$  (see Appendix) and using  $\theta = \log h$ , we can write the log-likelihood as

$$L = \sum_{j=1}^n m_j [\theta(t_j) + \log(d_j) - \log(m_j)]. \tag{7}$$

Maximizing  $L$  directly leads to the estimates (5). Instead, we maximize the penalized likelihood,

$$L - \frac{1}{2} \lambda J(\theta), \tag{8}$$

with  $J(\theta)$  the roughness penalty and  $\lambda$  an appropriate “spring constant.” With a penalty  $J(\theta)$  on overall curvature,

$$J(\theta) = \int [\theta''(t)]^2 dt, \tag{9}$$

the maximum penalized likelihood estimate is cubic between observation times but is continuous with a continuous derivative over the domain of observation. Other choices of the penalty  $J$  are possible, e.g., to emphasize a more Bayesian interpretation of  $\theta(t)$  (Silverman, 1985; Wahba, 1985).

If the spring constant  $\lambda$  is near 0 (strong springs), the estimate reduces to the rough maximum likelihood estimate (4). A large value (weak springs) forces the estimate to be a straight line. An intermediate value of  $\lambda$  can be chosen by generalized cross validation (GCV) (Craven and Wahba, 1979), which compares each count ( $m_j$ ) with the curve estimate obtained by excluding that count. Since nonparametric curve estimates are biased, one must strike a balance between bias and variance reduction. Selecting that  $\lambda$  which minimizes the GCV criterion tends to achieve this in the sense of approximately minimizing the expected mean squared error.

Suppose one observes  $r$  individuals, each with its own birth rate according to the model (1), and that  $m_{ij}$  offspring are found for female  $i$  at time  $t_j$ . The overall penalized likelihood can be written, using (5), as

$$L = \sum_{i=1}^r \sum_{j=1}^n m_{ij} [\theta^0(t_j) + \alpha_i(t_j) + \log d_j - \log m_{ij}] - P$$

where  $P$  incorporates the penalty. This can be expressed as

$$L = L(\theta^0) + \sum_{i=1}^r L(\alpha_i) - P,$$

with

$$\begin{aligned} L(\theta^0) &= \sum_{j=1}^n m_{+j} [\theta^0(t_j) + \log(rd_j) - \log m_{+j}], \\ L(\alpha_i) &= \sum_{j=1}^n m_{ij} [\alpha_i(t_j) + \log(m_{+j}/r) - \log m_{ij}], \quad i = 1, \dots, r, \end{aligned} \tag{10}$$

and  $m_{+j} = \sum_i m_{ij}$ . Note that  $L(\theta^0)$  is a Poisson likelihood, and at time  $t_j$ ,  $\{m_{ij}\}_{i=1}^r$  is multinomial conditional on  $m_{+j}$ . Thus, one now has  $(r+1)$  log-likelihoods in (10).

Our earlier discussion of roughness focused on a single function,  $\theta(t)$ , for a single individual. We now have  $(r+1)$  functions,  $\theta^0(t)$  and  $\alpha(t) = \{\alpha_i(t)\}_{i=1}^r$ . We postulate a smooth model that allows for two distinct roughness penalties, one for the baseline rate  $\theta^0(t)$  and the other jointly for the individual factors  $\alpha(t)$ . Thus, we need to specify two spring constants,  $\lambda_1$  and  $\lambda_2$ , one for each penalty. The resulting penalized likelihoods are

$$L(\theta^0, \lambda_1) = L(\theta^0) - \frac{1}{2}\lambda_1 J(\theta^0), \tag{11}$$

$$L(\alpha, \lambda_2) = \sum_{i=1}^r L(\alpha_i) - \frac{1}{2}\lambda_2 \sum_{i=1}^r J(\alpha_i). \tag{12}$$

Recall that if  $\lambda_2$  is large (weak springs), the individual factors  $\alpha_i$  are forced to be nearly linear.

The penalized likelihood  $L(\theta^0, \lambda_1)$  can be solved for fixed  $\lambda_1$  using a Newton–Raphson iterative scheme with Fisher scoring typical of generalized linear models (McCullagh and Nelder, 1983; cf. Green, 1984). A similar argument holds for  $L(\alpha, \lambda_2)$ . At convergence, the deviances  $D(\hat{\theta}^0) = -2L(\hat{\theta}^0)$  and  $D(\hat{\alpha}) = -2 \sum_{i=1}^r L(\hat{\alpha}_i)$  have approximately  $\chi^2$  distributions (Yandell and Green, 1986). One can estimate the degrees of freedom for each statistic and use these for testing. For instance, comparing  $D(\hat{\alpha})$  with the deviance for constant  $\alpha_i$  tests whether birth rates can be considered parallel (on a log scale), with a constant relative difference. Details are presented in the Appendix.

Algorithms for solving this problem can be found in Green and Yandell (1985) based on the B-spline algorithm of Hutchinson and de Hoog (1985). This is currently being prepared for general distribution in FORTRAN77. A multidimensional algorithm based on GCVPACK (Bates et al., 1987) appears in Yandell (1988) and is available in FORTRAN77 from the author.

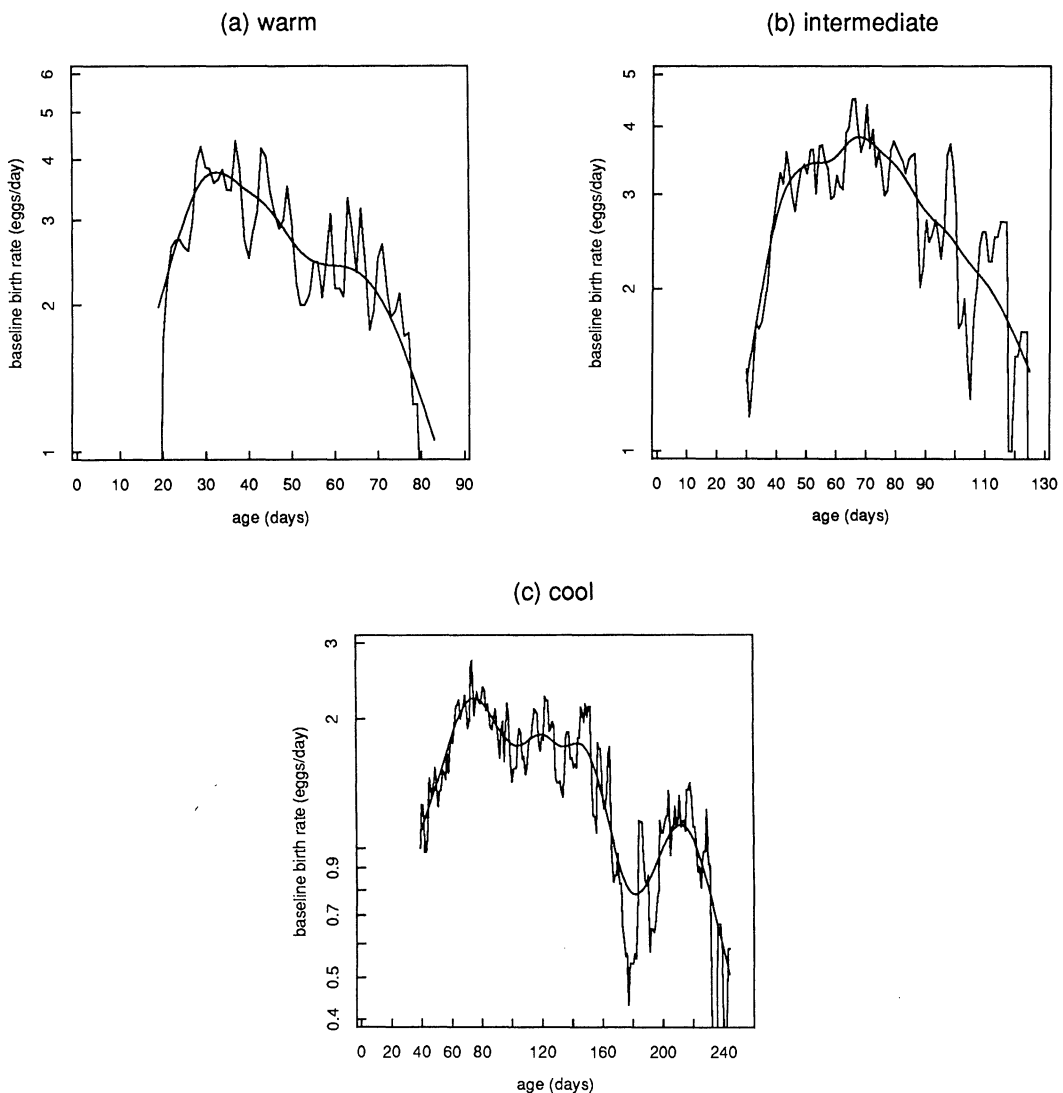
### 3. Potato Leafhoppers and Temperature

Hogg (1985) studied population dynamics of the potato leafhopper, an important crop pest, in controlled environments at the University of Wisconsin–Madison Biotron. Three temperature regimes were established with daily fluctuations designed to follow the cycle typically found under “cool,” “intermediate,” and “warm” summer field conditions in

southern Wisconsin. Analytical methods included age-specific life tables and Leslie matrix modelling, but did not incorporate individual variation in birth rates.

Here, we estimated the baseline oviposition rate  $h^0$  for leafhoppers at each of the three temperature regimes. Each regime had 25 or 26 females observed three times a week from adult emergence to death or removal from study. Of these, some never produced eggs and were not further considered, leaving 24, 21, and 23 individuals for cool, intermediate, and warm regimes, respectively. We based estimates on data from the age of first egg-laying through the last interval in which eggs were laid. As a consequence, the population of each group could change over the course of the study as individuals died or were lost from the study.

We estimated the baseline rates using (11). The smoothing parameters were chosen using GCV and gave fairly smooth estimates (Figure 1). Note that the cool regime baseline rate



**Figure 1.** Curve estimates for potato leafhopper baseline oviposition rates for (a) warm, (b) intermediate, and (c) cool regimes. Smooth lines are penalized estimates; rough lines are interpolated daily estimates.

estimate has more structure to it, reflecting the fact that there were two or three times as many records for this regime because leafhoppers tended to live longer at lower temperatures. The baseline rate curves all tend to exhibit an early peak in oviposition rate, followed by a gradual decline.

Potato leafhopper longevity is greatly influenced by the environment (note the different time scales in Figure 1). One might ask whether the intermediate and warm regime baseline rates are accelerated versions of the cool regime baseline rate. To examine this, we converted time to "physiological time" units of degree-days, which combine calendar time and temperature to reflect more realistically the leafhoppers' growth and maturation (Figure 2). It appears that the shapes of the baseline rate curves are roughly similar, supporting the notion of an accelerated time scale. However, there are differences in magnitude of oviposition rates among the three regimes. In particular, on a physiological time scale, the intermediate regime appears to be the most favorable for oviposition (cf. Hogg, 1985).

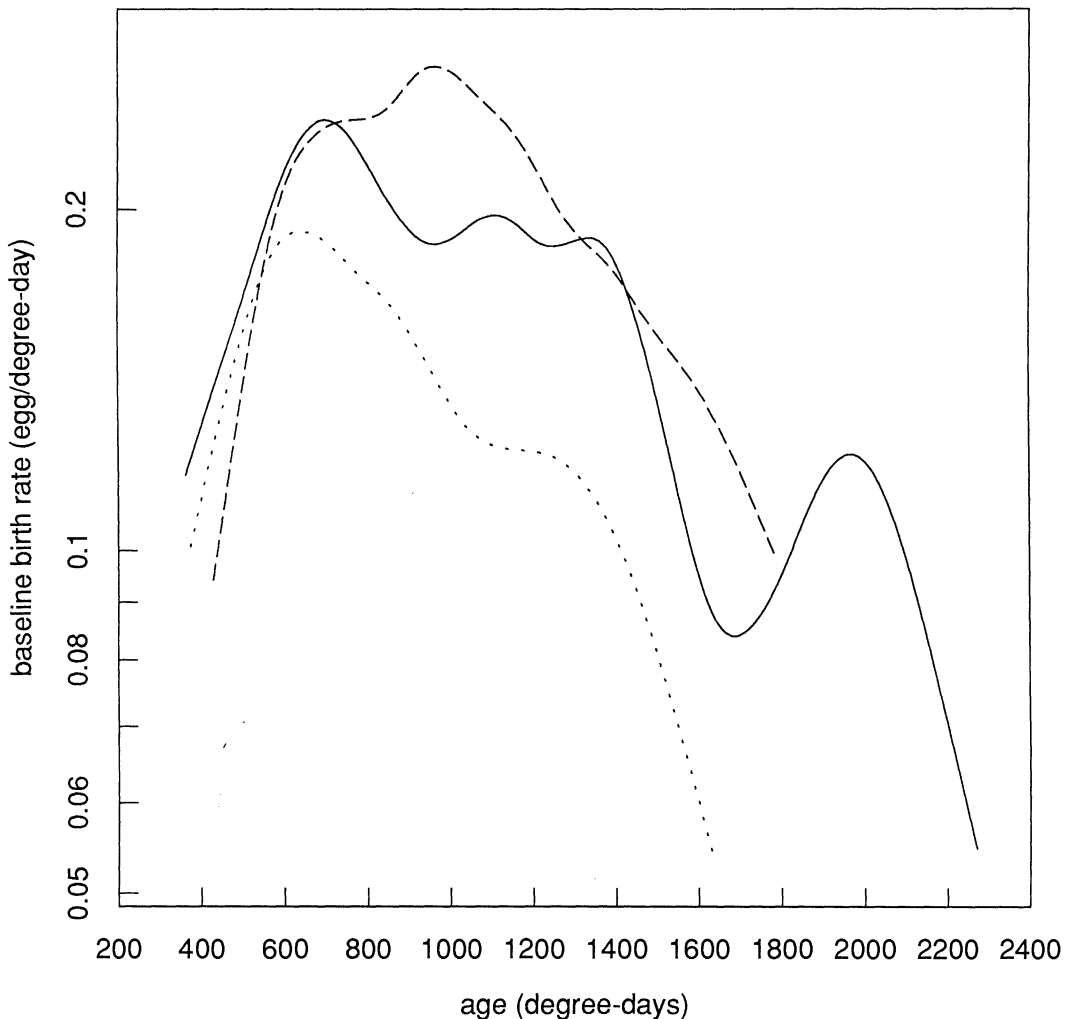


Figure 2. Baseline oviposition rate estimates on a common physiological time scale for warm (dot), intermediate (dash), and cool (solid) regimes.

4. Examining Individual Variability

We estimated individual factors  $a_i$  in each of the three temperature regimes using (12). We present here only results for the cool regime (Figure 3) since we have longer records for these individuals. Results for the other two regimes were similar.

Our main interest was in exhibiting the individual factors to look for interesting patterns not easily detected in the raw data or baseline rates. There were several individuals that had increased birth rates relative to the others early, but then died at a relatively early age. Note that some individuals had a much lower  $a_i$  in midlife (15, 20, 24), while some were just the opposite (18, 23).

We constructed deviance tests for constant individual factors to see whether the individual oviposition rates are parallel on a log scale, meaning that individual relative differences in egg-laying are constant over time. The tests were performed by comparing the difference in deviance between the constant  $a_i$  model and the smooth (spline)  $a_i$  model to a  $\chi^2$

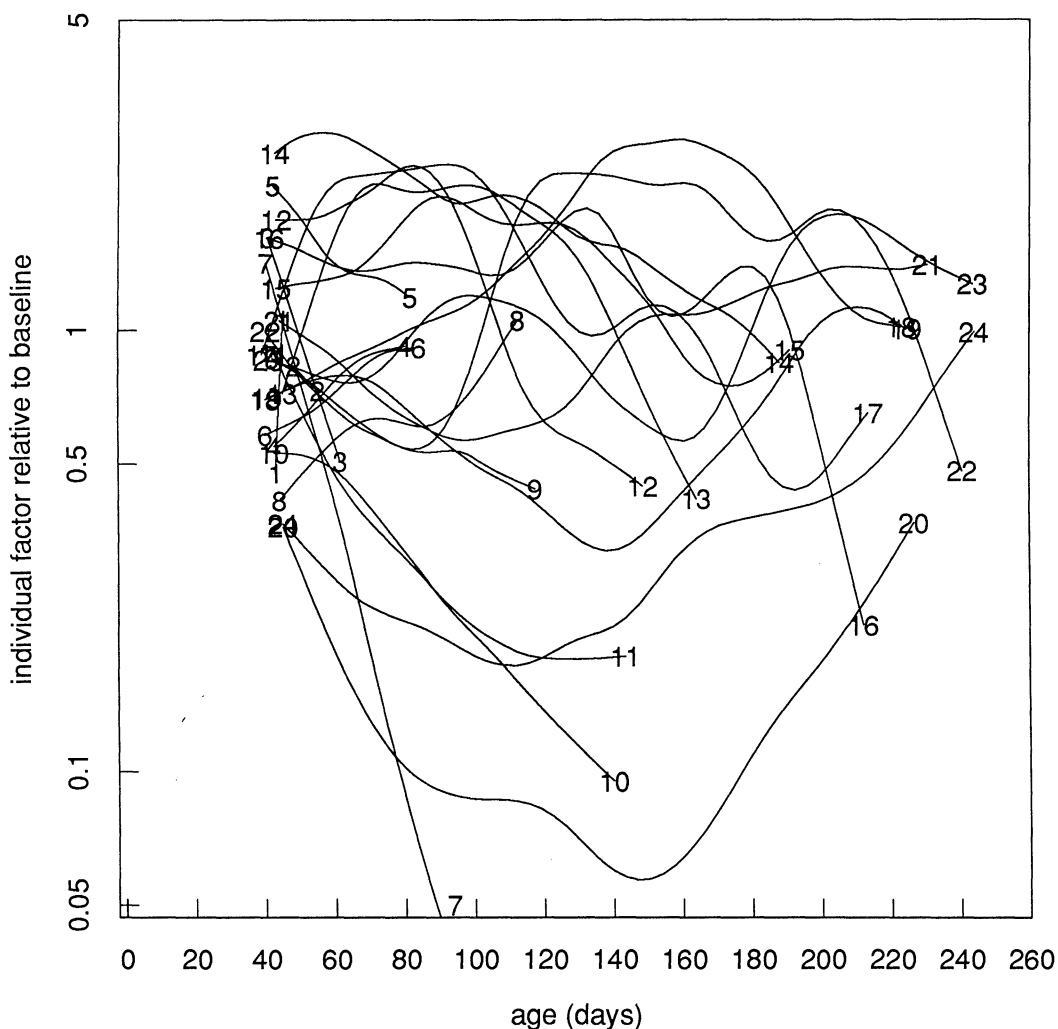


Figure 3. Variability among individual oviposition curves for 24 leafhoppers in the cold regime. Each curve is the individual factor, identified by survival rank.



distribution with degrees of freedom being the difference in degrees of freedom between the two models (Table 1). Note that the degrees of freedom are fractional and the test is approximate (Yandell and Green, 1986). These tests indicate that the individual factors are not constant; this result was consistent for all three regimes. Thus, the smoothly changing nature of the individual factors appears to be real, and marked differences exist in individual oviposition rates over time among leafhoppers in the same controlled environment.

**Table 1**  
*Deviance and degrees of freedom (in parentheses) for individual factors*

Regime	Model		Difference
	Constant	Spline	
Warm	265.90 (211)	145.70 (177.81)	120.20 (33.19)
Intermediate	856.55 (436)	614.31 (393.31)	242.24 (42.69)
Cool	1,727.07 (965)	974.87 (861.91)	752.20 (103.09)

Deviances are approximately  $\chi^2$  with d.f. which may be fractional.  
Difference is test of smooth spline against constant individual factor model.

## 5. Discussion

We have shown that it is possible to separate a baseline rate from individual factors that may modulate this baseline. We can examine pictures of the baseline rates and the individual factors, compare them visually, and even test for apparent patterns. While we have developed our model specifically for insect natality, the approach promises broader applicability for examining complicated counting data in other biological systems.

Different egg-laying patterns were exhibited by the potato leafhoppers within each regime, even among the longest-lived leafhoppers. However, we do not know if these represent distinct oviposition "strategies" within the leafhopper population, or whether they are simply a random sample from all possible patterns. These results suggest further work aimed at understanding the basis of individual variation in insect age-specific birth rates. It also suggests that population characteristics may not adequately reflect individual characteristics. For instance, the population net reproduction rate (cf. Pianka, 1978),

$$R^0 = \int_0^{\infty} l(t)h^0(t) dt,$$

with  $l(t)$  the survival probability, may differ markedly from the individual net reproduction rates

$$R_i = \int_0^{\infty} l(t)h_i(t) dt.$$

This net reproduction rate  $R^0$  plays a key role in evolutionary theory (Williams, 1966).

Some aspects of the theory need refinement, notably demonstrating conclusively the distributions of tests and properties of estimates for semiparametric GLMs. However, as it stands, the methods presented here can be used for exploratory investigation to examine temporal and individual variability, and to suggest particular parametric models that may be useful for subsequent experiments. New questions about temporal patterns and

individual variability will emerge as this methodology is incorporated into statistical packages accessible to biologists.

## ACKNOWLEDGEMENTS

We wish to thank Erik V. Nordheim and the referees for helpful suggestions on the manuscript. This work was supported in part by USDA-CSRS Grant 511-100, National Sciences Foundation Grant DMS-8404970 for Yandell, and the University of Wisconsin Graduate School and Regional Research Project NC-149 for Hogg. Computing was performed on the Statistics VAX 11/750 Research Computer at the University of Wisconsin.

## RÉSUMÉ

Nous présentons une méthodologie d'analyse des taux de natalité âge-spécifiques, en insistant particulièrement sur la description de la variabilité au sein d'une cohorte d'animaux d'âges voisins. Nous utilisons une structure de modèle appropriée à la natalité des insectes et d'autres arthropodes. Notre approche sépare la natalité en un niveau de base représentatif des effets de l'environnement sur la population, et en un facteur représentant l'unique effet de chaque individu de la cohorte. Nous employons conjointement les fonctions de lissage splines et une méthode de vraisemblance pénalisée pour l'estimation. Nous illustrons cette méthodologie en l'appliquant à des données sur le taux d'oviposition des cicadelles.

## REFERENCES

- Bartoszyński, R., Brown, B. W., McBride, C., and Thompson, J. R. (1981). Some nonparametric techniques for estimating the intensity function of a cancer-related nonstationary Poisson process. *Annals of Statistics* **9**, 1050–1060.
- Bates, D. M., Lindstrom, M. J., Wahba, G., and Yandell, B. S. (1987). GCVPACK—Routines for generalized cross validation. *Communications in Statistics, Series B—Simulation and Computing* **16**, 263–297.
- Cinlar, E. (1975). *Introduction to Stochastic Processes*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Clevenson, M. L. and Zidek, J. V. (1977). Bayes linear estimators of the intensity function of the nonstationary Poisson process. *Journal of the American Statistical Association* **72**, 112–120.
- Cox, D. R. (1972). Regression models and life tables (with Discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187–220.
- Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik* **31**, 377–403.
- Green, P. J. (1984). Iteratively reweighted least squares for maximum likelihood estimation and some robust and resistant alternatives (with Discussion). *Journal of the Royal Statistical Society, Series B* **46**, 149–192.
- Green, P. J. and Yandell, B. S. (1985). Semiparametric generalized linear models. In *GLIM85: Proceedings of the International Conference on Generalized Linear Models, September 1985*, R. Gilchrist (ed.). Lecture Notes in Statistics, Vol. 32. New York: Springer-Verlag.
- Hastie, T. J. and Tibshirani, R. J. (1986). Generalized additive models (with Discussion). *Statistical Science* **1**, 297–318.
- Hogg, D. B. (1985). Potato leafhopper (Homoptera: Cicadellidae) immature development, life tables, and population dynamics under fluctuating temperature regimes. *Environmental Entomology* **14**, 349–355.
- Hutchinson, M. F. and de Hoog, F. R. (1985). Smoothing noisy data with spline functions. *Numerische Mathematik* **47**, 99–106.
- Madden, L. V. (1985). Modeling the population dynamics of leafhoppers. In *The Leafhoppers and Planthoppers*, L. R. Nault and J. G. Rodriguez (eds), 235–258. New York: Wiley.
- McCullagh, P. and Nelder, J. A. (1983). *Generalized Linear Models*. New York: Chapman and Hall.
- Meyer, J. S., Ingersoll, C. G., McDonald, L. L., and Boyce, M. S. (1986). Estimating uncertainty in population growth rates: Jackknife vs bootstrap techniques. *Ecology* **67**, 1156–1166.
- O'Sullivan, F., Yandell, B. S., and Raynor, W. J., Jr. (1986). Automatic smoothing of regression functions in generalized linear models. *Journal of the American Statistical Association* **81**, 96–103.

- Pianka, E. (1978). *Evolutionary Ecology*, 2nd edition. New York: Harper & Row.
- Plant, R. E. (1986). A method for computing the elements of the Leslie matrix. *Biometrics* **42**, 933–939.
- Shiau, J.-J., Wahba, G., and Johnson, D. R. (1986). Partial spline models for the inclusion of tropopause and frontal boundary information in otherwise smooth two- and three-dimensional objective analysis. *Journal of Atmospheric and Ocean Technology* **3**, 714–725.
- Silverman, B. W. (1985). Some aspects of the spline smoothing approach to nonparametric regression curve fitting (with Discussion). *Journal of the Royal Statistical Society, Series B* **47**, 1–52.
- Slade, N. A. and Levenson, H. (1982). Estimating population growth rates from stochastic Leslie matrices. *Theoretical Population Biology* **22**, 299–308.
- Strong, D. R. (1986). Density-vague population change. *Trends in Ecology and Evolution* **1**, 39–42.
- Wahba, G. (1985). A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. *Annals of Statistics* **13**, 1378–1402.
- Williams, G. C. (1966). Natural selection, the costs of reproduction, and a refinement of Lack's principle. *American Naturalist* **100**, 687–690.
- Yandell, B. S. (1988). Algorithms for multidimensional semiparametric GLMs. *Communications in Statistics, Series B—Simulation and Computing* **17**, in press.
- Yandell, B. S. and Green, P. J. (1986). Semiparametric generalized linear model diagnostics. *Proceedings of the Statistical Computing Section, ASA, Chicago*, 48–53. Washington, D.C.: American Statistical Association.

Received April 1987; revised September 1987.

#### APPENDIX

We first justify our approximation of the integral (3) by the mean value theorem, and then present details of the maximum penalized likelihood estimation procedure for  $\theta^0$ , followed by a brief discussion of test statistics and the choice of spring constant  $\lambda_1$ . Brief mention is made of estimating  $\alpha$ .

The mean value theorem ensures that  $H_j = h(t_j^*)(t_j - t_{j-1}) = h(t_j^*)d_j$  for some  $t_{j-1} \leq t_j^* \leq t_j$ . In practice, we assume that  $t_j^* \approx (t_j + t_{j-1})/2$ , but we ignore this below and treat  $t_j^*$  as  $t_j$ . Note that this approximation introduces some bias that could be important if there are only a small number of observation times. Using  $\theta = \log h$  and the mean value theorem gives us (7).

Initial investigations showed that mortality in the population affects the baseline rates but has only a modest effect on the individual factors. One can adjust for mortality by modifying the population, and even allowing the baseline rate to have jumps (Shiau, Wahba, and Johnson, 1986). However, this has not been treated here.

The log-likelihood  $L(\theta^0)$  can be approximated near its maximum, using a second-order Taylor series expansion, by a weighted least squares. Given an initial estimate  $\theta_1^0$  of  $\theta^0$ ,

$$\begin{aligned} L(\theta^0) &\approx L(\theta_1^0) + \sum [\theta^0(t_j) - \theta_1^0(t_j)](m_{+j} - w_j) - \frac{1}{2} \sum w_j [\theta^0(t_j) - \theta_1^0(t_j)]^2 \\ &= L(\theta_1^0) + \frac{1}{2} \sum w_j (m_{+j} - w_j)^2 - \frac{1}{2} \sum w_j [y_j - \theta^0(t_j)]^2, \end{aligned} \quad (13)$$

with “working-weights” and “working-values” being, respectively,

$$w_j = rd_j \exp[\theta_1^0(t_j)] \quad \text{and} \quad y_j = \theta_1^0(t_j) + (m_{+j} - w_j)/w_j. \quad (14)$$

Note that given  $\theta_1^0$  only the last term of (13) depends on the unknown  $\theta^0$ .

The problem of finding  $\theta^0$  to maximize  $L(\theta^0, \lambda_1)$  is equivalent to finding  $\theta^0$  to minimize

$$\sum w_j [y_j - \theta^0(t_j)]^2 + \lambda_1 J(\theta^0) \quad (15)$$

and iterating to convergence. Typically one makes an initial guess at  $\theta^0$ , assigns  $w_j$  and  $y_j$  in (14), solves for a new  $\theta^0$  in (15), replaces  $\theta_1^0$  by the estimated  $\theta^0$ , and updates  $w_j$  and  $y_j$ , and so on, until the estimates converge adequately.

The estimate that minimizes (15) for penalty (9) can be written as (cf. Bates et al., 1987)

$$\theta^0(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \sum_{j=1}^n \delta_j |t - t_j|^3$$

with the condition that

$$\sum_{j=1}^n \delta_j t_j^k = 0, \quad k = 0, 1, 2.$$

To actually solve for this, we can write the penalty (9) in a quadratic form in the  $\delta_j$ 's.

$$J(\theta^0) = \sum_{j=1}^n \sum_{k=1}^n \delta_j \delta_k |t_j - t_k|^3,$$

allowing us to restate (15) as a quadratic form in the  $\delta_j$ 's and  $\beta_k$ 's, which can be minimized by standard methods.

At convergence, the estimate can be written as

$$\hat{\theta}^0 = \mathbf{A} \mathbf{y},$$

with  $\hat{\theta}^0$  and  $\mathbf{y}$  vectors of the parameter estimates and working values, respectively, and  $\mathbf{A}$  the matrix of coefficients. The "effective number of parameters" are  $\text{trace}(\mathbf{A})$  (Craven and Wahba, 1979). The deviance  $D^2(\hat{\theta}^0) = -2L(\hat{\theta}^0)$  has approximately a  $\chi^2$  distribution with  $\text{trace}(\mathbf{I} - \mathbf{A})$  degrees of freedom (often a fraction), which can be used for approximate tests (Yandell and Green, 1986).

The choice of  $\lambda_1$  can be made by minimizing the GCV function,

$$V(\lambda_1) = n[\text{trace}(\mathbf{I} - \mathbf{A})]^{-2} \sum_{j=1}^n w_j [y_j - \hat{\theta}^0(t_j)]^2. \tag{16}$$

Since  $w_j$  and  $y_j$  depend on the estimate  $\hat{\theta}^0(t_j)$ , one repeats the above minimization for several choices of  $\lambda$  and chooses that value which minimizes (16) by a grid search.

Similar arguments can be made for  $\alpha$  and  $\lambda_2$ . However, when data are not collected on the same days for all insects one must estimate the total number of offspring per each observation period on individual  $i$ . This can be done by linear interpolation, smoothing out the births for each individual, and replacing  $m_{+j}$  by the interpolated value in  $L(\alpha_i)$  of (10). The penalty is a sum of terms such as those given above for  $J(\theta^0)$ . One then proceeds to the weighted least squares iteration to convergence. The deviance under the null model of constant  $a_i$  is simply  $D^2(\bar{\alpha}) = -2 \sum_{i=1}^n L(\bar{\alpha}_i)$  from (10) with  $\alpha_i(t_j)$  replaced by  $\bar{\alpha}_i = \log(m_{i+}/m_{++})$ , the maximum likelihood estimator for constant individual factors. The difference in deviances reported in Table 2 is simply  $D^2(\bar{\alpha}) - D^2(\hat{\alpha})$ .