

**Statistical Analysis of Population Dynamics in Space and Time Using  
Estimating Functions**



Subhash Lele; Mark L. Taper; Stuart Gage

*Ecology*, Vol. 79, No. 5 (Jul., 1998), 1489-1502.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9658%28199807%2979%3A5%3C1489%3ASAOPDI%3E2.0.CO%3B2-6>

*Ecology* is currently published by The Ecological Society of America.

---

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/esa.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).

## STATISTICAL ANALYSIS OF POPULATION DYNAMICS IN SPACE AND TIME USING ESTIMATING FUNCTIONS

SUBHASH LELE,<sup>1,4</sup> MARK L. TAPER,<sup>2</sup> AND STUART GAGE<sup>3</sup>

<sup>1</sup>*Department of Biostatistics, The Johns Hopkins University, 615 North Wolfe Street, Baltimore, Maryland 21205 USA*

<sup>2</sup>*Department of Biology, Montana State University, Bozeman, Montana 59715 USA*

<sup>3</sup>*Department of Entomology, Michigan State University, East Lansing, Michigan 48824 USA*

**Abstract.** The interplay of dispersal, disturbance, and local dynamics in spatial mosaics has profound effects on the stability and viability of populations. There are two main reasons to consider spatial models in population dynamics: (1) improved estimation of the parameters by utilizing spatial replications, and (2) ecologically realistic modeling. In this paper, we suggest models that are generalizations of the univariate population dynamics models (for example, Ricker or Gompertz) to space–time situations. We accommodate both spatially correlated environmental perturbations and dispersal. Moreover, we suggest computationally simple parameter estimation procedures based on estimating functions and provide an approach for obtaining approximate confidence intervals. The methodology is illustrated on the spatial time series of gypsy moths in the lower peninsula of Michigan.

**Key words:** density dependence; dispersal; estimating functions; gypsy moth; logistic growth; *Lymantria dispar* L.; spatially explicit models.

### INTRODUCTION

The interplay of dispersal, disturbance, and local dynamics in spatial mosaics has profound effects on the stability and viability of animal populations. Ecologists are increasingly appreciating the importance of spatial dimension and spatial heterogeneity in the study of natural populations. The importance of spatial aspects in ecology has been recently emphasized in special feature articles in *Ecology* (Goldwasser et al. 1994, Holmes and Lewis 1994, Molofsky 1994, Tilman 1994). In fact, in a paper title, Kareiva (1994) proclaims: “Space: the final frontier for ecological theory.” This paper is a foray to this “final frontier.”

There are two main reasons to consider spatial models in population dynamics.

#### *Improved estimation of model parameters*

Some biological populations exhibit density dependent growth (Andrewartha and Birch 1954, Bulmer 1975, Royama 1977, Dennis 1989, Kemp and Dennis 1993, Dennis and Taper 1994, Hooten 1995, Zeng et al. 1998). There are various models of population growth that incorporate density dependence (Dennis and Taper 1994, Hooten 1995). These models are mostly stochastic versions of deterministic differential equations models. Estimation of the parameters in these models is crucial for various applied problems and management decisions. Unfortunately, the commonly available length of the time series for parameter esti-

mation is short, in the range of 10–30 yr. Shorter time series lead to larger variability in parameter estimates.

However, in many ecological situations, spatial replications are available and should be utilized (Cigliano et al. 1995, Dennis et al. 1997) to improve parameter estimation. These spatial replications can be utilized to improve the estimators in three different ways. One can consider the spatial replications identical and independent replications of the time series and combine the information under that hypothesis. This, in effect, assumes that the phenomenon is spatially homogeneous, the environmental noise is spatially uncorrelated, and that there is no dispersal. Another approach is to assume that spatial replicates are related to each other only because they share the environment. This leads to the “joint density dependence” (JDD) model by Dennis et al. (1997), which assumes that the environmental noise is spatially correlated, but there is no immigration and emigration between the spatial locations. In this paper we explore a third approach which is to explicitly take into account the effects of dispersal in addition to the effects of shared environment.

#### *Ecologically realistic modeling*

There have been a number of theoretically interesting and biologically sensible models for population dynamics that explicitly accommodate spatial aspects. There have been two basic approaches in this modeling: reaction diffusion models involving partial differential equations (Hamilton and May 1977, Okubo 1980, Kot and Schaffer 1986, Okubo et al. 1989, Holmes and Lewis 1994), and cellular automata models (Molofsky 1994, Zhou and Liebhold 1995). The reaction diffusion

Manuscript received 7 July 1997; accepted 17 July 1997; final version received 6 October 1997.

<sup>4</sup> Address reprint requests to this author.

models consider populations living in a spatial continuum but are generally deterministic in nature. On the other hand, models based on cellular automata consider spatially discrete (or patch based) populations and, while often stochastic, assume some finite neighborhood for the spatial dependence.

Both of these approaches fail to capture important attributes of nature. First, the dynamics of all animal and plant populations are highly stochastic and typically affected strongly by variations, such as weather, in the environment. When viewed over a very small area, even well established populations may go extinct locally as a result of this stochastic variation. When the population is established over large areas, zones of local extinction are recolonized quickly from adjoining areas (Liebhold et al. 1993). Second, long range dispersal occurs in many animal and plant taxa (Harper 1977, Chepko-Sade and Halpin 1987, Goldsworthy and Wheeler 1989). It is clear that a model incorporating both environmental noise and an unspecified, possibly nonfinite, neighborhood would be highly desirable.

Sound theoretical models are the foundation for statistical analysis of real life phenomena. The next essential step is to couple empirical data to these theoretical models. This involves developing statistical inference procedures, such as point estimation and confidence intervals for the model parameters, and checking the validity of the models. In this paper we progress towards these goals. Our principal points and demonstration are:

1) We generalize the models developed by Dennis et al. (1997) to explicitly take into account the effects of immigration and emigration. These models correspond to stochastic versions of deterministic models suggested by Kot and Schaffer (1986). We call this class of models "dispersive joint density dependence" (DJDD) models.

2) The usual method of estimation, maximum likelihood estimation, is difficult to implement for the proposed model. We suggest an estimation procedure utilizing the newly developed statistical methodology of estimating functions (Godambe 1991). This method leads to computationally simpler and intuitively appealing estimation procedures. We use the parametric bootstrap (Monte Carlo) approach to obtain confidence intervals for the parameters. Prediction error estimates are calculated by a jackknife procedure.

3) Habitat heterogeneity can be represented in the model with site-specific parameter values. Both the model and the estimation procedure are flexible enough to take into account habitat covariates when they are available.

4) We illustrate the use of the model and the inference procedures on a space-time series of gypsy moth abundance in Michigan (1985–1994). Our conclusions about the average dispersal distance of range expansion seem to match empirical observations well. We also observe theoretical findings from the reaction-diffu-

sion models, such as the stabilizing influence of dispersal to be borne out in this data set.

#### MODELS FOR GROWTH AND DISPERSAL

The statistical models presented here accommodate both density-dependent growth and spatial dispersal. Our approach assumes that the data are available at each location over the whole period of study. If this is not the case, then such a data set must be constructed either by aggregation or smoothing of the original data.

We use the following notation throughout the paper. We assume that there are  $n$  locations and  $T$  time periods. The location of site  $i$  is denoted by  $s_i$  where  $i = 1, 2, \dots, n$ . The population size at time  $t$  at location  $i$  is denoted by  $N_{i,t}$ , where  $t = 1, 2, \dots, T$ . Thus, corresponding to each time period  $t$ , we have a vector of observations of length  $n$ .

Given present population sizes,  $N_{i,t}$ , population sizes in the next time period,  $N_{i,t+1}$  are obtained by using the following mechanism.

*Step 1.*—This involves immigration and emigration. Let  $k(s_i, s_j)$  denote the proportion of individuals of location  $i$  that emigrate from location  $i$  to location  $j$ . Similarly  $k(s_j, s_i)$  corresponds to the proportion of individuals immigrating into location  $i$  from location  $j$ .

This function  $k(s_i, s_j)$  need not be symmetric in  $i$  and  $j$ . Also notice that this is a proportion and not the actual numbers. Actual numbers would depend on the number of individuals present at the location from which the individuals are emigrating.

It is easy to see that the total proportion of individuals emigrating out of location  $i$  is given by  $\sum_{j \neq i} k(s_i, s_j)$ . Since the proportion of emigration can not exceed 1, a constraint on the function  $k(s_i, s_j)$  is that  $\sum_{j \neq i} k(s_i, s_j) \leq 1$  for any fixed  $i$ . We will describe a variety of possible  $k(s_i, s_j)$  functions subsequently.

*Step 2.*—After immigration and emigration, the population sizes,  $N_{i,t}^*$ , are

$$N_{i,t}^* = N_{i,t} - \sum_{j \neq i} k(s_i, s_j)N_{i,t} + \sum_{j \neq i} k(s_j, s_i)N_{j,t}.$$

The second term shows the total number of individuals leaving the location  $i$  and the third term signifies the total number of individuals immigrating into location  $i$  from all other locations.

*Step 3.*—Next there is population growth. The population at time  $(t + 1)$  at location  $i$  is given by

$$N_{i,t+1} = \exp[f(N_{i,t}^*, \boldsymbol{\theta} + \boldsymbol{\epsilon}_{i,t})]N_{i,t}^*$$

where  $\boldsymbol{\epsilon}_{i,t}$  denotes the environmental noise at location  $i$  and time  $t$ . The function  $f(N_{i,t}^*, \boldsymbol{\theta})$  describes population growth rate as a function of current population size and a parameter vector  $\boldsymbol{\theta}$ . We assume that the environmental noise is spatially correlated but temporally independent. We will describe the models for environmental noise subsequently.

*Models for local growth*

Following are some of the most commonly used population dynamic models.

*Exponential growth model.*—This model supposes that the populations grow at an exponential rate and that there is no upper bound to population size. This model, in fact, is a model for “density-independent” growth. For this model  $f(N; a, b) = a$ .

*Ricker growth model.*—This is a discrete-time logistic growth model (Ricker 1954) and is one of the most commonly used models for density dependent growth. See Hooten (1995) for various situations in which the Ricker growth model seems to fit well to real populations. For this model  $f(N; a, b) = a + bN$ . The parameter  $a$  corresponds to the intrinsic growth rate, and the parameter  $b$  corresponds to the impact of intraspecific competition on growth rate. If  $b$  is negative, this model has an asymptote corresponding to the carrying capacity given by  $K = -a/b$ .

*Gompertz growth model.*—This is another commonly used model for density-dependent growth (Gompertz 1825, Dennis and Taper 1994). For this model  $f(N; a, b) = a + b \ln(N)$ . The Gompertz model has an asymptote corresponding to the carrying capacity given by  $K = \exp(-a/b)$ , provided  $b < 0$ .

In this paper, we will use only the Ricker growth model. Application of Gompertz and other models will be illustrated elsewhere. We discuss a variety of models to indicate and emphasize the generality of the statistical procedures described in this paper. All models can be analyzed using the estimation procedures discussed in this paper with minor and straightforward modifications.

*Models for dispersal*

We will describe models for isotropic dispersal, where the dispersal depends only on the distance and not on the direction (Kot and Schaffer 1986). These models can be modified to accommodate nonisotropic dispersal based on covariates, such as prevalent wind direction, water bodies, traffic patterns, etc.

*Half exponential model.*—The most commonly used model, suggested by Fisher (1937) and further used by Skellam (1951), is the half exponential model. This model has proved useful in wind-aided local dispersal in insect populations (Kettle 1951, Liebhold et al. 1993). For this model,

$$k(s_i, s_j; \delta) = \delta^{\|s_i - s_j\|}$$

where  $\|s_i - s_j\|$  denotes the distance between locations  $i$  and  $j$ . The dispersal parameter  $\delta$  takes values in  $[0, 1)$  with the constraint that the sum over  $j$ , for any fixed  $i$ , is less than 1.

*Half Gaussian model.*—If the rate of dispersal is high for nearby locations but decreases rapidly thereafter, this might be an appropriate model for dispersal. For this model,

$$k(s_i, s_j; \delta) = \exp\left(\frac{-\|s_i - s_j\|^2}{2\delta}\right).$$

For this model, the dispersal parameter  $\delta$  takes values in  $(0, \infty)$  with the summation constraint. When  $\delta$  is zero,  $k(s_i, s_j; 0) \equiv 0$ .

*Half Cauchy model.*—If the rate of dispersal decreases slowly with increasing distance, this model may be useful. For this model,

$$k(s_i, s_j; \delta) = \frac{1}{1 + \delta\|s_i - s_j\|^2}.$$

For this model, the dispersal parameter  $\delta$  takes values in  $(0, \infty)$  with the summation constraint. When  $\delta$  is zero,  $k(s_i, s_j; 0) \equiv 0$ .

As mentioned previously, we discuss a variety of models for dispersion to show the generality possible. This paper uses only the exponential dispersion model because it is appropriate in modeling gypsy moth population dynamics. For other organisms, one may use other models for dispersion. The estimation procedures discussed in this paper can be modified easily to accommodate other dispersion models as well as covariate-dependent dispersion.

*Models for environmental noise*

The vector  $\epsilon_t$  denotes the environmental noise at time period  $t$ . This is an  $n \times 1$  vector. We assume that the environmental noise does not depend on time. We also assume that the environmental noise is independent from one time period to another. These assumptions are routinely made both in time series modeling as well as in population dynamics modeling. Mathematically, this implies that  $\epsilon_t$ 's are independent, identically distributed vector valued random variables. We also assume, without loss of generality, that the mean of this vector is zero.

We do, however, allow for spatial dependence in the environmental noise. We assume that the components of the vector  $\epsilon$  are correlated to each other. There are various ways to model covariance structure over space. The most extensive source for such models is the geostatistical literature, e.g., Cressie (1991). In this paper, we use a simple isotropic exponential covariance structure. According to this structure,

$$\text{cov}(\epsilon_i, \epsilon_j; \rho, \sigma^2) = \sigma^2 \rho^{\|s_i - s_j\|}.$$

Thus, the spatial covariance decreases exponentially with distance.

*Parameters of interest*

Given the above set of models, the parameters of interest are:

1) Density dependence parameters  $a$  and  $b$  (and the corresponding carrying capacity  $K$  which is a function of  $a$  and  $b$ ). These parameters may vary over space. If they do, we would like to know how they vary over space. We may also want to know the relationship be-

tween these spatially varying parameters and various habitat characteristics.

2) Dispersal parameter  $\delta$ . This will suggest how fast the organisms are spreading over the space.

3) Environmental noise parameters  $\rho$  and  $\sigma^2$ .

In passing we note that if the dispersal parameter  $\delta = 0$ , we recover the "joint density dependence" model proposed by Dennis et al. (1997). Notice that if the environmental noise correlation parameter  $\rho = 0$ , but the dispersal parameter  $\delta \neq 0$ , we will observe spatial dependence in the process. This kind of spatial dependence is not captured by the joint density dependence model. Spatial dependence can be due to shared environment, or dispersal, or both.

#### *Model selection*

We have described various ways to model population dynamics in space and time. For a given data set, one can decide to use a particular model based on various possible criteria such as cross validation based root mean squared prediction error (RMSPE). In this paper, we have used RMSPE for model selection. We prefer to conduct the calculations of the prediction error on the log scale, because, according to the model, on the log scale the variances are homoscedastic. Hence all the prediction differences can be weighted equally.

#### ESTIMATION PROCEDURES

The most commonly used estimation procedure for a stochastic model is the maximum likelihood estimation. For the models considered here, writing the explicit likelihood and maximizing it to obtain the maximum likelihood estimates is a difficult task. Moreover, likelihood estimators tend to be quite sensitive to departures from the model. To avoid the numerical difficulties as well as to gain robustness against model misspecification, there is a growing trend in the statistical literature to use a generalization of the maximum likelihood procedure called "estimating functions" (Godambe 1991). Because this literature tends to be highly technical, we provide a very brief review of some basic ideas in this approach. For more details and mathematical discussions, readers may refer to Godambe (1991).

#### *Estimating functions and their advantages*

An estimating function is any function of the data and the parameters whose mean value is zero. To obtain the estimates of the parameters using an estimating function, one equates it to zero and solves the resulting equation for the parameters.

We will illustrate this idea using simple linear regression. Consider a linear regression with a single covariate and no intercept. In this case the model is  $Y_i = X_i\beta + \epsilon_i$ , where the  $\epsilon_i$  are zero mean random variables. The least squares method of obtaining the estimator of the parameter  $\beta$  finds a value of  $\beta$  that minimizes  $\sum_{i=1}^n (Y_i - X_i\beta)^2$ . Using the method of calculus,

it is known that this is equivalent to finding the solution of the equation  $\sum_{i=1}^n X_i(Y_i - X_i\beta) = 0$  for  $\beta$ . This, in fact, is a particular case of estimating function approach.

Notice that  $\sum_{i=1}^n X_i(Y_i - X_i\beta)$  is a function of the data  $(X_i, Y_i)$  and parameter  $\beta$ . Moreover under the model's assumptions,  $Y_i$  has mean  $X_i\beta$ , and hence the mean of  $(Y_i - X_i\beta)$  is zero. We have  $n$  such functions, namely,  $(Y_1 - X_1\beta)$ ,  $(Y_2 - X_2\beta)$ ,  $\dots$ ,  $(Y_n - X_n\beta)$ , but we have only one parameter. Thus it is intuitively advantageous to combine these  $n$  functions. The most informative such combination is given by  $\sum_{i=1}^n X_i(Y_i - X_i\beta)$ . Notice that this is also a function of the data and the parameter. Because each component has mean zero, the sum also has mean zero. Hence we have a function of the data  $(X_i, Y_i)$  and parameter  $\beta$  whose mean is zero. Thus, by definition, this is an estimating function. An estimator is obtained by equating this function to zero and solving the resultant equation for the parameter. Similar arguments show that finding the maximum likelihood estimator also corresponds to finding a solution to an estimating function. Estimating function methods are a generalization of the least squares and maximum likelihood methods.

There are several advantages to considering estimating functions over a maximum likelihood approach. Perhaps the most important advantage is computational. In complicated stochastic models, such as the one considered in this paper, the exact likelihood function is analytically difficult. It involves inversion of large covariance matrices of dimension equal to the number of spatial locations. The number of parameters is also large. In such a situation, numerical maximization of the likelihood tends to be unstable and computationally intensive. However, constructing estimating functions that are computationally simple to solve, is feasible. It will be demonstrated here that the estimating functions we construct mostly involve solving a simple system of linear equations. This is computationally stable and simple even for very large data sets. See Lindsay (1988) for many such examples where likelihood is extremely difficult, but one can construct estimating functions that are analytically and computationally simple.

Not only can one construct estimating functions that are simple analytically and computationally, but the estimators obtained by solving the estimating functions, under simple regularity conditions, are consistent; that is they tend to the true value of the parameter as the sample size increases (Crowder 1986). Moreover, they are asymptotically normal; that is, the distribution of the parameters can be approximated by a normal distribution.

The second important feature of estimating functions is model robustness. Recall that the least squares estimators are valid even if the error structure is not normal. The quasi-likelihood methods, which are a particular case of the estimating function methodology,

used in Generalized Linear Models (McCulloch and Nelder 1989) are robust against model misspecification. Only part of the model, such as the mean and the variance has to be correctly specified for the estimating functions to provide correct estimators. For example, it will be shown that our estimation procedure is valid whether or not environmental variation is normal.

ESTIMATING FUNCTIONS FOR THE SPATIAL STOCHASTIC RICKER MODEL

We will now derive the estimating functions for the Spatial Stochastic Ricker Model which is a particular case of the dispersive joint density dependence models proposed in this paper. The local density dependence model used is the Ricker model; hence the name Spatial Stochastic Ricker Model. To begin with, we will consider the model for which the parameters  $a$ ,  $b$ , and  $\delta$  are spatially homogeneous. Spatially varying parameters are considered in the *Data analysis* section.

Consider the following function of the data and the parameters  $a$ ,  $b$  and  $\delta$ :

$$g(N_{i,t+1}, N_{i,t}, a, b, \delta) = \ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^* \quad (1)$$

where  $N_{i,t}^* = N_{i,t} - \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{j,t} + \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{j,t}$ . It is easy to check that under the assumptions on the environmental noise, this has mean zero. Hence, according to the definition, these are estimating functions. Also notice that there are  $n(T - 1)$  of these estimating functions. We only have three parameters  $a$ ,  $b$ , and  $\delta$  to estimate. We need to combine these estimating functions so that there are three equations in three unknowns. A nearly optimal set of linear combinations of these estimating functions (details in the Appendix) is:

$$\sum_{t=1}^{T-1} \sum_{i=1}^n (\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*) = 0 \quad (2)$$

$$\sum_{t=1}^{T-1} \sum_{i=1}^n N_{i,t}^* (\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*) = 0 \quad (3)$$

$$\sum_{t=1}^{T-1} \sum_{i=1}^n w_{i,t} (\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*) = 0 \quad (4)$$

where the weights,  $w_{i,t}$ , are given by

$$w_{i,t} = \frac{\sum_{j \neq i} \|s_i - s_j\| \delta^{\|s_i - s_j\|} N_{j,t} - \sum_{j \neq i} \|s_i - s_j\| \delta^{\|s_i - s_j\|} N_{i,t}}{N_{i,t} - \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{i,t} + \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{j,t}}$$

The above equations are designed (see the Appendix) to maximize the information extracted from the data about  $a$ ,  $b$ , and  $\delta$  respectively.

To obtain the estimates, we must find that value of the vector  $(a, b, \delta)$  which satisfies the above three equations simultaneously. We use the following procedure.

*Step 1.*—Choose an initial value of  $\delta_0$  in the interval  $[0, 1)$ . For this value of  $\delta$  solve Eqs. 2 and 3 to obtain

$(a_0, b_0)$ . Notice that this is achieved very easily. It involves only two linear equations in two unknowns and has a closed form solution. The estimators  $\hat{a}$  and  $\hat{b}$ , for a fixed value of  $\delta$ , are intercept and slope of a least squares regression of  $\ln(N_{i,t+1}) - \ln(N_{i,t}^*)$  on  $N_{i,t}^*$ .

*Step 2.*—Consider the vector  $(a_0, b_0, \delta_0)$ . Substitute these values in Eq. 4 and check if it equals zero. If it does, then our estimates are  $(a_0, b_0, \delta_0)$ . If not, change the value of  $\delta$  and try again. Since the range of  $\delta$  is bounded, such a search is feasible. We utilize the root bisection method (Press et al. 1986) to organize the search.

The above three equations lead us to an estimate of  $(a, b, \delta)$ . Let us denote it by  $(\hat{a}, \hat{b}, \hat{\delta})$ . Notice that a Gaussian error structure is not essential in the derivation of these estimators.

Estimation of the stochastic noise parameter  $\sigma^2$  is also fairly straightforward. We use the model assumption that environmental noise

$$\epsilon_{i,t} = \ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*$$

has mean zero and variance  $\sigma^2$ . Thus an estimator of  $\sigma^2$  is given by

$$\widehat{\sigma^2} = \frac{1}{n(T-1)} \sum_{t=1}^{T-1} \sum_{i=1}^n [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - \hat{a} - \hat{b}N_{i,t}^*]^2$$

where  $N_{i,t}^*$  is evaluated using  $\hat{\delta}$ .

Estimating the spatial correlation  $\rho$  is more difficult. Consider the stochastic noise vector  $\epsilon$ . Suppose we now assume that this vector comes from a multivariate elliptically symmetric distribution. This class of distributions includes multivariate Gaussian distribution and scale mixtures of the multivariate Gaussian distribution, such as multivariate  $t$ , multivariate Cauchy, etc. Under certain conditions (Kelker 1970), for such a vector valued random variable, the following result holds:

$$E(\epsilon_i | \epsilon_j) = \rho^{\|s_i - s_j\|} \epsilon_j$$

where  $E(\cdot)$  denotes the expectation operator. Thus an estimating function for  $\rho$  can be easily formulated as

$$\sum_{i=2}^T \sum_{i < j} \|s_i - s_j\| \rho^{\|s_i - s_j\|} \epsilon_{j,t} (\epsilon_{i,t} - \rho^{\|s_i - s_j\|} \epsilon_{j,t}) = 0.$$

Of course, in practice,  $\epsilon_{i,t}$  are not observable. We replace them by their estimates,

$$\epsilon_{i,t} = \ln(N_{i,t+1}) - \ln(N_{i,t}^*) - \hat{a} - \hat{b}N_{i,t}^*.$$

Also it should be noted that since  $\rho$  lies in the interval  $[0, 1)$ , weights corresponding to distant locations are very small. To reduce the computational burden substantially without losing much efficiency, one can consider only locations that fall within a specified distance of each focal location. Solving the above estimating function provides the estimate of the environmental correlation parameter  $\rho$ .

The values of  $\delta$  and  $\rho$  depend on the spatial units chosen. For estimation purposes we scale distances so the minimum distance is 1 (by dividing by the smallest

of all possible pairwise distances between the county centroids). For numerical stability, we also scale the population abundances by dividing by the largest abundance value in the data set. However we report the dispersal parameter  $\delta$  and the spatial correlation parameter  $\rho$  in the units of kilometers per year. The average dispersal distance per year for the double exponential dispersion model is given by  $-1/\ln \delta$ .

*Accommodating spatial variability*

The above estimating functions were derived under the model where the parameters ( $a$ ,  $b$ ) were spatially homogeneous. In the real world, these parameters may vary over space. One can presume a model where  $a$  varies spatially, but other parameters do not, and estimate  $a_i$  corresponding to each location. If some habitat-related covariates are available, one can easily estimate the effect of these covariates on  $a$  by modeling  $a$  as a function of the habitat covariates. A similar tactic can be followed for the spatially-varying  $b$ . In this paper, we will not use particular covariates but only study how the estimates of site-specific  $a$  vary spatially.

Studying this variability pictorially helps understand the process. Plotting of the spatially-varying carrying capacity  $K_i$  is also quite informative. Geographical plotting of the spatially-varying  $K_i$  can help identify possible covariates that determine variation in the carrying capacity. For example, if the map of  $K_i$  and a map of a food plant species are similar, then that may indicate an important covariate. Inclusion of some explicit covariates and the effect of habitat characteristics on the carrying capacity,  $K_i$ , and the growth parameters  $a$  and  $b$  is studied elsewhere. It is interesting to note that negative carrying capacity estimates are quite possible. In the language of source-sink theory, locations with positive carrying capacities are sources while those with negative carrying capacities are sinks (Pulliam 1988).

*Obtaining confidence intervals for the estimates*

There are various approaches to obtaining confidence intervals and standard errors. One can use the asymptotic normality of the estimators to obtain the confidence intervals. But when the time series length is short, other methods such as the jackknife and the parametric bootstrap (Efron and Tibshirani 1991) might be better. For the present paper, we use the parametric (Monte Carlo) bootstrap. This approach needs an assumption for the distributional form of the stochastic noise. We assume Gaussian noise. The steps involved in obtaining the Monte Carlo confidence intervals are as follows.

Step 1: Generate space-time series of the same size and structure as the original data under the Gaussian model, using the estimated parameters  $\hat{a}$ ,  $\hat{b}$ ,  $\hat{\delta}$ ,  $\hat{\rho}$ , and  $\hat{\sigma}^2$ .

Step 2: Reestimate the parameters using the generated data.

Step 3: Repeat Steps 1 and 2 large number of times, say 1000.

Step 4: Find the histogram of the estimated values from Step 2. This is the Monte Carlo estimate of the distribution of the estimators. Percentiles of this distribution are used to obtain the confidence intervals for the parameters.

Note that if the distribution of the bootstrapped parameter estimates appears to be non-symmetric or skewed, biased corrected or calibrated bootstrap intervals (Efron and Tibshirani 1991) may be more desirable.

*Prediction error calculation*

We calculate the prediction error using jackknifing. Because of the spatial correlation in environmental noise, instead of jackknifing observations, we jackknife the estimating functions (Lele 1991) in groups of size  $n$  as follows:

Step 1: Delete all  $n$  estimating functions corresponding to growth in the  $k$ th year. Solve the following set of equations to obtain jackknifed estimate of the parameters:

$$\sum_{t \neq k}^{T-1} \sum_{i=1}^n [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] = 0$$

$$\sum_{t \neq k}^{T-1} \sum_{i=1}^n N_{i,t}^* [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] = 0$$

$$\sum_{t \neq k}^{T-1} \sum_{i=1}^n w_{i,t} [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] = 0.$$

Step 2: Use these estimated parameters, denoted by  $(\hat{a}_{(-k)}, \hat{b}_{(-k)}, \hat{\delta}_{(-k)})$ , to predict the log population densities for the  $k$ th year from the population densities of year  $(k - 1)$  as follows. First calculate

$$N_{i,k-1}^* = N_{i,k-1} - \sum_{j \neq i} \hat{\delta}_{(-k)}^{\|s_i - s_j\|} N_{j,k-1} + \sum_{j \neq i} \hat{\delta}_{(-k)}^{\|s_i - s_j\|} N_{j,k-1}$$

and then calculate the predicted population sizes as

$$N_{i,k} = \exp(\hat{a}_{(-k)} + \hat{b}_{(-k)} N_{i,k-1}^*) N_{i,k-1}.$$

Step 3: Calculate the average squared difference between the log-predicted values and the logarithm of actual values. Averages can be taken either over all sites in a given year or over all sites and years. The square root of the average is reported as the root mean squared prediction errors (RMSPE) on the log scale. The log scale is used because this is the scale in which errors can be expected to be homoscedastic under the model assumptions.

ANALYSIS OF GYPSY MOTHS IN MICHIGAN'S LOWER PENINSULA

The purpose of this section is to illustrate the use of the described techniques in a real-life situation and not to provide an in-depth analysis of this particular data set.

The gypsy moth, *Lymantria dispar* L., is considered one of the most destructive forest defoliating pests in North America. Since its introduction in Massachusetts in 1869, it has expanded its range over 10<sup>6</sup> km<sup>2</sup> (Liebhold et al. 1992), and it is likely to invade most of North America. In many areas, epidemics are common, and resultant defoliation can cause substantial ecological and economic effects.

As with other pests, an ability to predict the occurrence of outbreaks is crucial for pest management. Unfortunately, gypsy moth populations tend to be erratic through space and time, and outbreaks are consequently difficult to forecast. Recently there have been attempts at incorporating the spatial component in models for studying the gypsy moth population dynamics (Campbell 1973, 1976, Liebhold and Elkinton 1989, Zhou and Liebhold 1995). For a general review and a detailed bibliography, see *Spatial analysis and forest pest management* (Liebhold and Barrett 1993).

Gypsy moths seem to have density-dependent growth that could possibly be modeled by the Stochastic Ricker Model. Females in North America are unable to fly, and thus the primary mechanism for gypsy moth dispersal is thought to be wind-borne passive movement of first instars (Mason and McManus 1981). The expansion of the generally infested area and the founding of isolated populations may also occur when egg masses or other life stages are accidentally transported by humans (Talerico 1981). The model proposed in this paper thus seems to be appropriate for the gypsy moth population dynamics. We illustrate the use of the suggested spatio-temporal model to analyze the spread of gypsy moths in the state of Michigan.

The gypsy moth was first discovered in Michigan in 1954 (O'Dell 1955), and first observed defoliation occurred in the state in 1984 in Midland county, located in the center of the state. In 1992, 76 of Michigan's 83 counties reported the presence of the gypsy moth (Witter et al. 1992). As much as 3000 m<sup>2</sup> of defoliation was recorded in 1992. This is a 10-fold increase in defoliation over the amount of defoliation in 1986 (Simmons 1987).

In order to follow the population distribution of these pests and to assess the risk to forests in Michigan, a statewide population monitoring program was established in 1985 using pheromone-baited traps (Gage et al. 1990). Pheromone traps attract only the male gypsy moth. Over 3300 traps are monitored annually at permanent stations, and the number of moths in the trap are counted and reported by the Michigan Department of Agriculture each year. Unfortunately, not all trap locations have data recorded for the complete 10-yr period under consideration. Thus for this analysis we have aggregated the trap counts within a county. The county abundance for a year is taken to be the average number of moths caught per trap in reporting traps multiplied by the area of the county. Because an effort was made to distribute traps uniformly in space, this

number should approximately be proportional to the total number of moths occurring in that county in that year. We consider here only the 68 contiguous counties of the lower peninsula. A space-time series of these counts is shown in Fig. 1. The spatial aspect of this spread is quite obvious.

We present the models used and the results of the analysis. This analysis provides parameter estimates, Monte Carlo confidence intervals, and a measure of the performance of the model based on prediction error. Comparison of models with different numbers of parameters is attained by comparing root mean squared prediction errors.

*Model I: Spatially homogeneous model*

This is the simpler of the models. We assume Stochastic Ricker Growth and double exponential dispersal. The model components are given by

$$f(N_{i,t}^*; a, b) = a + bN_{i,t}^*$$

$$k(s_i, s_j; \delta) = \delta^{\|s_i - s_j\|}$$

$$\text{cov}(\epsilon_i, \epsilon_j; \rho, \sigma^2) = \sigma^2 \rho^{\|s_i - s_j\|}$$

Thus the parameters of interest are  $(a, b, \delta, \rho, \sigma^2)$ . The parameter estimates and the corresponding confidence intervals are presented in Table 1.

The values of  $\delta$  and  $\rho$  depend on the spatial units chosen. For estimation purposes we scale distances so the minimum distance is 1 (by dividing by the smallest of all possible pairwise distances between the county centroids). For numerical stability, we also scale the population abundances by dividing by the largest abundance value in the data set.

For reporting (Table 1), values of the scale-dependent parameter  $b$  are converted back to the original scale by again dividing by the largest observation. We report the dispersal parameter  $\delta$  and the spatial correlation parameter  $\rho$  at a distance of 1 km. The average dispersal distance per year for the double exponential dispersion model is given by  $-1/\ln(\delta)$ . The environmental noise correlation  $\rho$  at distance 1 km is high, but this should be expected because a distance of 1 km is very small on the meteorological scale. Locations that are only 1 km apart tend to have very similar weather patterns and other climatic conditions. The confidence intervals reported in Table 1 are based on the simple percentile method. The bootstrap distributions are reasonably symmetric for this model.

As indicated in Table 2, the overall RMSPE of log-population size for Model I is 1.02. This means that 90% of our predictions are within a factor of 5.54 ( $=\exp(1.68 \times 1.02)$ , where 1.68 is the  $z$ -value for 90% cutoff point of a standard normal distribution) of the true population sizes. While this figure may seem high, it should be remembered that this is an eruptive species whose population densities fluctuated dramatically over the study period. Over the course of the decade, maximum population densities at a site were, on av-



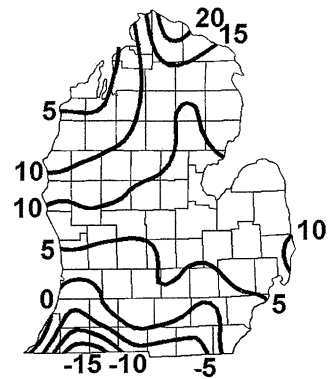
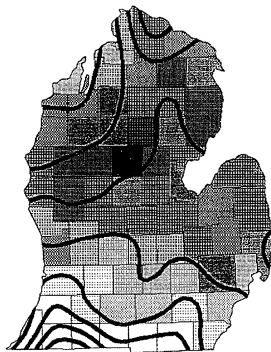
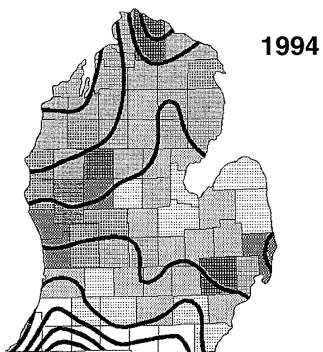
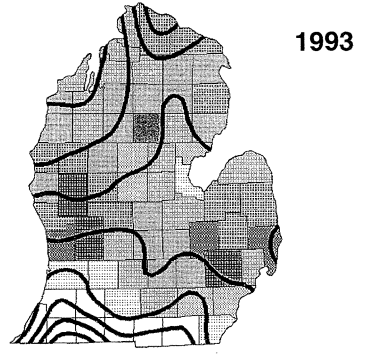
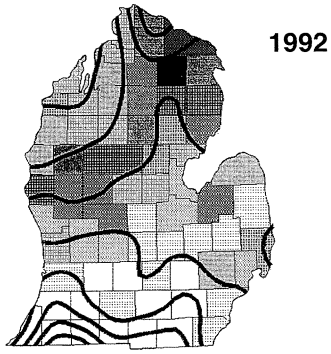
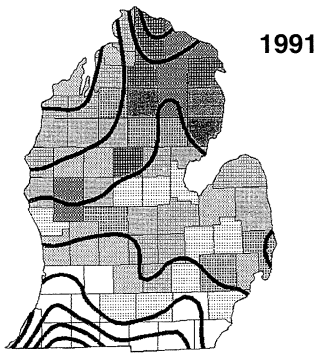
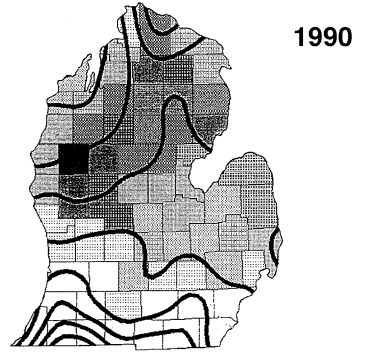
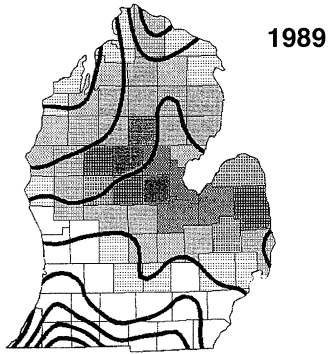
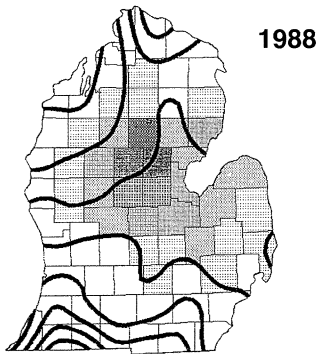
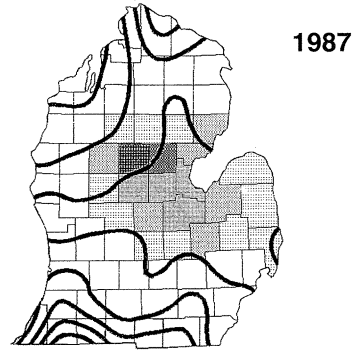
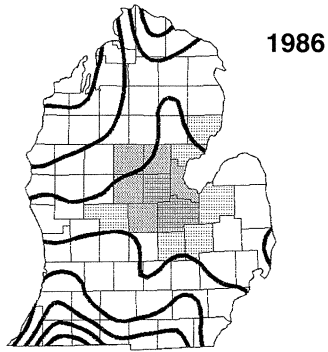
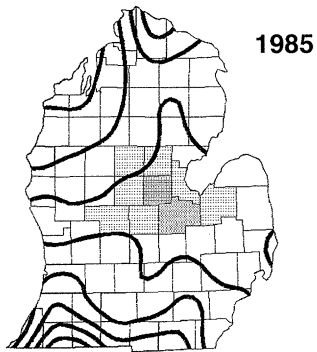


TABLE 1. The parameter estimates (and 90% bootstrap CI) for spatially homogeneous (Model I) and spatially varying  $a$  (Model II) models per location for the log population densities using dispersive joint density dependence (DJDD) and joint density dependence (JDD). Use of dispersion reduces the estimate of the variance parameter substantially, suggesting a better fit for DJDD models than for JDD models.

Parameter	DJDD		JDD	
	Model I	Model II	Model I	Model II
$a$ (yr <sup>-1</sup> )	0.555 (0.304, 0.800)	...	0.654 (0.430, 0.862)	...
$b$ (yr <sup>-1</sup> ·individual <sup>-1</sup> )	-3.448 (-4.864, -2.548)	-4.849 (-6.711, -3.215)	-3.810 (-4.137, -3.590)	-5.776 (-6.235, -5.732)
$\delta$ (km <sup>-1</sup> )	0.928 (0.919, 0.989)	0.941 (0.935, 0.952)	...	...
$\rho$ (km <sup>-1</sup> )	0.986 (0.981, 0.989)	0.987 (0.981, 0.990)	0.971 (0.964, 0.975)	0.970 (0.963, 0.975)
$\sigma^2$ (yr <sup>-1</sup> )	0.904 (0.725, 1.074)	0.705 (0.489, 0.756)	2.620 (2.327, 2.899)	2.368 (1.852, 2.354)

erage, a thousand times greater than that site's nonzero minimum density.

*Model II: Spatially nonhomogeneous intrinsic growth rate  $a$*

Next we fit a model where the parameter  $a$  is allowed to vary spatially. This more general model should provide with a better prediction error rate. Also, geographical plotting of the spatially varying  $a$  can hint at the possible covariates that could be of importance for the intrinsic growth rate. For example, if the map of  $a$  and map of potential food plant are similar, then that may indicate an important covariate.

TABLE 2. Yearly root mean squared prediction errors for spatially homogeneous (Model I) and spatially varying  $a$  (Model II) models per location for the log population densities using dispersive joint density dependence (DJDD) and joint density dependence (JDD). Use of dispersion improves the predictions especially in the expansion phase of population growth.

Year	DJDD		JDD	
	Model I	Model II	Model I	Model II
1986	1.69	1.64	4.03	4.14
1987	1.48	1.44	1.47	1.69
1988	0.84	0.70	1.44	1.46
1989	1.03	0.91	1.23	1.34
1990	0.85	0.70	0.66	0.64
1991	0.56	0.52	0.61	0.88
1992	0.83	0.76	0.79	0.92
1993	0.62	0.84	0.66	0.84
1994	0.70	0.62	0.63	0.68
Overall	1.02	0.97	1.64	1.74

The exact description of Model II is

$$f(N_{i,t}^*, a, b) = a_i + b \times N_{i,t}^*$$

$$k(s_i, s_j; \delta) = \delta^{\|s_i - s_j\|}$$

$$\text{cov}(\epsilon_i, \epsilon_j; \rho, \sigma^2) = \sigma^2 \times \rho^{\|s_i - s_j\|}$$

Thus, in this model,  $a$  depends on the site location. It has no temporal trend. Everything else remains the same as in the previous model.

The parameter estimates for ( $b, \delta, \rho, \sigma^2$ ) are provided in Table 1. Estimated  $a_i$  values range from -0.78 to 1.65. Instead of tabulating the estimates of all 68 carrying capacity values, we present them in a pictorial fashion by superimposing contours of  $K$  values over shaded plots of population densities by county (see Fig. 1). It is evident that there is spatial heterogeneity in the carrying capacity. The estimated carrying capacities match well with the population densities observed, averaged over the entire time period. See the lower right sub-figure of Fig. 1. The estimate of the average dispersal distance for the DJDD Model II of about 17 km/yr ( $16.44 = -1/\ln(0.941)$ ) matches reasonably well with the reported empirical rates of dispersal (21 km/yr) for North America outside Michigan (Liebhold et al. 1992).

We compare the mean squared prediction errors (RMSPE) under the two models in Table 2. The spatially heterogeneous DJDD model has a RMSPE = 0.97 (90% predictions within a factor of 5.1). This is slightly smaller than the spatially homogeneous DJDD model, with RMSPE = 1.02 (factor of 5.54). Despite the greatly increased number of parameters to be estimated, we have decreased the RMSPE. This indicates that spatial variation in  $a$  is important. By regressing on habitat characteristics, one could reduce the number of param-

←

FIG. 1. Contours of carrying capacity estimated from dispersive joint density dependence Model II superimposed on shaded plots of gypsy moth population abundances by county. Deeper shading indicates a higher abundance. Population abundances averaged over all 10 yr are plotted in the subfigure in the lower right corner.

eters to be estimated, which might reduce the prediction error further.

We consider Model II important for data exploration and hypothesis generation. There is a separate parameter  $a$  corresponding to each location. Clearly this is an undesirable situation from the statistical perspective (Neyman and Scott 1948) because of the low ratio of number of parameters to the available number of observations. Although one can estimate all these parameters, the efficiency of these estimates can be very low. One cannot expect too much from a model with so many parameters. We suggest that these parameter estimates should not be used at their face values, but the researcher should plot these estimates against possible covariates such as temperature gradients or dominant tree species, etc. Such a procedure should suggest a useful covariate that may explain the spatial variability of parameter  $a$ . Such covariates may then be incorporated into the model structure through regression, effectively reducing the number of parameters. Statistical analysis can then be carried out efficiently. This idea will be explored and illustrated elsewhere.

We also analyzed the gypsy moth space–time series using the joint density dependence model (Dennis et al. 1997). The parameter estimates are different (Table 1). Notice particularly that the variance estimate is substantially larger under the JDD model than DJDD. This indicates a possible bad fit for the joint density dependence model and supports the importance of dispersal in this system. This, in general, also will have an impact on the length of the prediction interval, or equivalently the uncertainty of the predictions, that may have relevance to pest management decisions. The RMSPE is also larger for the JDD model as is evident from Table 2. However the reader may notice that the predictions based on JDD do substantially worse than those based on DJDD in the first four years, whereas they are comparable in the last few years. A possible explanation for this is that during the expansion phase DJDD is better than JDD; whereas during the equilibrium phase JDD and DJDD behave similarly at this county-level resolution.

DISPERSAL AND ITS EFFECTS

There are two concerns that may be raised against the dispersive joint density dependence model. The first one relates to the possible confounding of the parameters, and the other relates to the (possible) unimportance of dispersal. We address both of these concerns.

*Confounding of the parameters*

Using standard arguments (see for example, Yasui and Lele 1997) in the theory of estimating functions and conditions on the rate of decay of the environmental correlation, it is straightforward to prove that the estimation procedure described yields consistent estimators; that is, as the sample size increases the estimates converge to the true parameter values. Here

TABLE 3. Nonconfounding and estimability of the parameters in the dispersive joint density dependence Model I.

	True parameters					Mean of the estimates
Intrinsic growth rate $a = 0.55$						0.57
Density dependence $b = -3.45 \times 10^{-6}$						$-3.7 \times 10^{-6}$
Dispersal parameter $\delta = 0.93$						0.93
Environmental correlation $\rho = 0.986$						0.985
Environmental noise $\sigma^2 = 0.90$						0.88
Standard deviations (along diagonal) and correlations (off diagonal)						
	$a$	$b$	$\delta$	$\rho$	$\sigma^2$	
$a$	0.157	-0.249	-0.385	-0.025	-0.043	
$b$		$6.83 \times 10^{-7}$	0.188	-0.016	-0.014	
$\delta$			$4.64 \times 10^{-3}$	-0.010	0.047	
$\rho$				$2.36 \times 10^{-3}$	0.785	
$\sigma^2$					0.104	

*Note:* The results are based on the average of 250 simulations. They clearly show that the parameters are estimable unbiasedly. True values, means, standard deviations, and correlations between the estimates are given. The estimators are clearly unbiased, indicating estimability, and the correlations are not too high, indicating nonconfounding.

we report simulation evidence for the unbiasedness of these estimators. We generate a data set using DJDD Model I under given parameter values and estimate the parameters using the estimating function procedure. We repeated this procedure 250 times. Table 3 reports the average values of these 250 parameter estimates and their standard errors and correlations. It is self-evident that the estimators are almost unbiased. There seems to be neither a theoretical nor a practical reason to believe that the parameters are confounded. The parameters related to dispersion,  $\delta$ , and spatial correlation,  $\rho$ , in the environment are estimable. The parameter estimates, however, are correlated, and hence there is some variance inflation. The estimates most affected are  $\hat{a}$  and  $\hat{\sigma}^2$ . But even for these, the error is manageable.

*Importance of dispersal*

Suppose that the space–time dynamics of the populations do include dispersal, but one chooses to ignore it; that is one fits the joint density dependence model proposed by Dennis et al. (1997). In simulations with true parameters set to our Model I estimates, we find that the intrinsic growth rate parameter  $a$  is substantially biased. With a mean estimate of 0.99,  $\hat{a}$  has a bias of 180% (250 trials). Given that the bias in  $b$  is small, this translates to estimating the carrying capacity to be  $1.8K$  instead of the true carrying capacity of  $K$ . On the other hand, suppose there is no dispersal (that is, the underlying model is of joint density dependence), and we fit the dispersive joint density dependence model. It is clear that joint density dependence model is a particular case of the dispersive joint density dependence model, and hence the estimates should be identical. This was also found to be true in simulations

(not reported here). We can see in our example study (Table 2) of gypsy moths in Michigan that excluding dispersal from our analysis decreases the accuracy of our predictions.

In conclusion, one can say that there is no loss in fitting the dispersive joint density dependence model to the space-time series of populations. In fact, it can be advantageous.

DISCUSSION

Stochastic models for studying a single set of time series data of an ecological population have proved quite successful in applied ecology. Our goal was to extend the use of stochastic models to studying spatio-temporal phenomena. Dennis et al. (1997) incorporate spatial aspects through the inclusion of spatially correlated environmental noise. We have extended their model to explicitly incorporate dispersal. Stochastic models incorporating the spatial dimension may prove useful for pest management and other applied ecological problems.

These stochastic models and statistical inference procedures can be refined to consider several other practical issues.

*Sampling variability.*—Most of the stochastic models in ecology are formulated in terms of population densities or population abundances. However in practice, all that is available is an estimate of these quantities. When fitting the models to the data, we tend to behave as if these estimates are the true values. This, however, makes the confidence statements about the parameter estimates incorrect. This can be corrected by considering hierarchical modeling approach (Clayton and Kaldor 1987, Carlin and Louis 1996). Although conceptually simple to model, the statistical inference becomes very involved for such models.

*Spatial aggregation.*—Many theoretical models work at the level of spatial continuum, but the data are available only at the spatially aggregated level such as a township or a county. There is a need for both theoretical and empirical analyses of the effect of such aggregation on the parameter estimates and their relationship with the model parameters. A preliminary simulation study was conducted to get a feel for the effects of aggregation. For this study, we generated a single space-time series with 1024 spatial locations and 10 time steps. Different aggregations were achieved by summing the values of the neighboring sites. For example, average of four neighboring sites led to 256 resultant sites. These data were then used to reestimate the parameters. The results for four levels of aggregation are reported in Table 4. As one would expect, large amount of aggregation leads to larger biases in the estimation of underlying parameters. These biases are true to the intuition. For example, for the aggregated data, dispersal of individuals from the boundary of one area to the boundary of the adjacent area is taken to be dispersal from the center of one area to the center

TABLE 4. The effects of aggregation are seen for a space-time series generated for 10 time steps at 1024 sites. Parameters were estimated at the original scale and at increasing levels of aggregation. Aggregation was achieved by summing the observations from contiguous sites. Systematic and interpretable biases are generated by aggregation.

True values	1024 sites	256 sites	64 sites	16 sites
$a = 0.555$	0.486	0.521	0.604	0.71
$b = -3.448 \times 10^{-6}$	-3.355	-3.100	-3.208	-3.719
$\delta = 0.066$	0.069	0.202	0.356	0.692
$\rho = 0.606$	0.606	0.677	0.750	0.758
$\sigma^2 = 0.904$	0.947	0.675	0.433	0.217

of the other. Hence, the dispersal parameter is larger than the finer resolution data. Thus, estimated parameter values are most reasonably interpreted on the scale at which the analysis is undertaken. We also think that although aggregation introduces such biases, it may reduce the effect of sampling variability, thus reducing the variance in the estimators. It would be worthwhile studying the optimal aggregation scale for this bias-variance trade-off.

*Spatially and temporally varying parameters.*—Only recently (Zeng et al. 1997) have researchers considered temporally varying parameters in population dynamics. Our estimation procedure can be easily generalized to the case where the parameters vary both spatially and temporally, when they are modeled explicitly as functions of spatially and temporally varying known covariates such as habitat characteristics or environmental factors such as temperature, amount of rainfall, etc.

In summary this paper on “dispersive joint density dependence” achieves the following:

- 1) It generalizes the joint density dependence model to include dispersal.
- 2) It proposes a statistical inference procedure so that the model can be connected to real data.
- 3) It allows for spatially and temporally varying parameters.
- 4) Spatially and temporally varying parameters can be explicitly modeled as functions of the spatially and temporally varying known covariates such as habitat characteristics or environmental factors such as temperature, amount of rainfall, etc.
- 5) The empirical conclusions about the spread of gypsy moths (about 21 km/yr) match with the DJDD Model II-based estimate (17 km/yr) reasonably well.
- 6) Dispersal is known to be an essential part of population dynamics. Diffusion is generally thought of as a stabilizing influence, one that homogenizes populations and moderates temporal fluctuations. It is easy to see that many of the location-specific  $a_i$  suggest nearly chaotic behavior or rapid local extinction, but the actual data does not exhibit such a behavior. It is possible that the dispersal is moderating the temporal fluctua-

tions as predicted by various reaction–diffusion theories.

Although there is still much work to be done, we view this paper as a first step towards a useful and exciting methodology for analyzing ecological space–time data.

#### ACKNOWLEDGMENTS

Subhash Lele was partially supported by a grant from the US-DOE (95-841) and US-EPA (CR-820086) to Daniel Goodman. M. L. Taper was partially supported by grants from US-EPA (CR-820086), US-NSF (DEB9411770) and Mountain Research Center of the Montana State University. We are grateful to two anonymous referees, the associate editor, Brian Dennis, Sandy Liebhold, Charles McCulloch and Bill Link for their comments on the manuscript.

#### LITERATURE CITED

- Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, Illinois, USA.
- Bulmer, M. G. 1975. The statistical analysis of density dependence. *Biometrics* **31**:901–911.
- Campbell, R. W. 1973. Numerical behavior of a gypsy moth population system. *Forest Science* **19**:162–167.
- . 1976. Comparative analysis of numerically stable and violently fluctuating gypsy moth populations. *Environmental Entomology* **5**:1218–1224.
- Carlin, B. P., and T. A. Louis. 1996. Bayes and empirical Bayes methods for data analysis. Chapman and Hall, London, UK.
- Chepko-Sade, B. D., and Z. T. Halpin, editors. 1987. Mammalian dispersal patterns. University of Chicago Press, Chicago, USA.
- Cigliano, M. M., W. P. Kemp, and T. M. Kalaris. 1995. Spatiotemporal characteristics of rangeland grasshopper (Orthoptera: Acrididae) regional outbreaks in Montana. *Journal of Orthoptera Research* **4**:111–126.
- Clayton, D. G., and J. M. Kaldor. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* **43**:671–681.
- Cressie, N. A. C. 1991. Statistics for spatial data. John Wiley and Sons, New York, New York, USA.
- Crowder, M. 1986. On consistency and inconsistency of estimating equations. *Econometric Theory* **2**:305–330.
- Dennis, B. 1989. Allee effects: population growth, critical density, and the chance of extinction. *Natural Resource Modeling* **3**:481–538.
- Dennis, B., W. P. Kemp, and M. L. Taper. 1998. Joint density dependence. *Ecology* **79**(2):426–441.
- Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. *Ecological Monographs* **64**:205–224.
- Diggle, P. J. 1991. Time series. Oxford University Press, Oxford, UK.
- Efron, B., and R. J. Tibshirani. 1991. An introduction to the bootstrap. Chapman and Hall, New York, New York, USA.
- Fisher, R. A. 1937. The wave of advance of advantageous genes. *Annals of Eugenics* **7**:355–369.
- Gage, S. H., Wirth, T. and G. Simmons. 1990. Predicting regional gypsy moth (Lymantriidae) population trends in an expanding population using pheromone trap catch and spatial analysis. *Environmental Entomology* **19**(2):370–377.
- Godambe, V. P. 1985. The foundations of finite sample estimation in stochastic processes. *Biometrika* **72**:419–428.
- , editor. 1991. Estimating functions. Oxford University Press, Oxford, UK.
- Goldsworthy, G. J., and C. H. Wheeler, editors. 1989. Insect flight. CRC Press, Boca Raton, Florida, USA.
- Goldwasser, L., J. Cook, and E. D. Silverman. 1994. The effects of variability on metapopulation dynamics and rates of invasion. *Ecology* **75**:40–47.
- Gompertz, B. 1825. On the nature of the function expressive of the law of human mortality, and a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* **115**:513–585.
- Hamilton, W. D., and R. M. May. 1977. Dispersal in stable habitats. *Nature* **269**:578–581.
- Harper, J. L. 1977. Population biology of plants. Academic Press, New York, New York, USA.
- Holmes, E. E., and M. A. Lewis. 1994. Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology* **75**:17–29.
- Hooten, M. 1995. Distinguishing forms of statistical density dependence and independence in animal time series data using information criteria. Dissertation. Montana State University, Bozeman, Montana, USA.
- Kareiva, P. 1994. Space: the final frontier for ecological theory. *Ecology* **75**:1.
- Kelker, D. 1970. Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya A* **32**:419–430.
- Kemp, W. P., and B. Dennis. 1993. Density dependence in rangeland grasshoppers (Orthoptera: Acrididae). *Oecologia* **96**:1–8.
- Kettle, D. S. 1951. Spatial distribution of *Culoides Impunctatus* Goet under woodland and moorland conditions and its flight range through woodland. *Bulletin of Entomological Research* **42**:239–291.
- Kot, M., and W. M. Schaffer. 1986. Discrete-time growth-dispersal models. *Mathematical Biosciences* **80**:109–136.
- Lele, S. 1991. Jackknifing linear estimating equations: asymptotic theory and applications in stochastic processes. *Journal of the Royal Statistical Society Series B* **53**:253–267.
- Liang K. Y., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* **73**:13–22.
- Liebhold, A. M., and H. R. Barrett, editors. 1993. Proceedings: spatial analysis and forest pest management. ARS, U.S. Forest Service Technical Report **NE-175**.
- Liebhold, A. M., and J. S. Elkinton. 1989. Characterizing spatial patterns of gypsy moth regional defoliation. *Forest Science* **35**:557–568.
- Liebhold, A. M., J. A. Halverson, and G. A. Elmes. 1992. Gypsy moth invasion in North America: a quantitative analysis. *Journal of Biogeography* **19**:513–520.
- Liebhold, A., E. Luzader, J. Halverson, and G. Elmes. 1993. The spatial dynamics of invasions by exotic pests. Pages 125–132 in A. M. Liebhold and H. R. Barrett, editors. Proceedings: spatial analysis and forest pest management. ARS, U.S. Forest Service Technical Report **NE-175**.
- Lindsay, B. G. 1988. Composite likelihood methods. *Contemporary Mathematics*, **80**:221–236.
- Mason, C. J., and M. L. McManus. 1981. Larval dispersal of the gypsy moth. The gypsy moth: research toward integrated pest management. Pages 161–202 in C. C. Doane and M. L. McManus, editors. U.S. Forest Service Technical Bulletin 1584.
- McCulloch P. and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, London, UK.
- Molofsky, J. 1994. Population dynamics and pattern formation in theoretical populations. *Ecology* **75**:30–39.
- Neyman, J., and E. L. Scott. 1948. Consistent estimates based on partially consistent observations. *Econometrica* **16**:1–32.

O'Dell, W. V. 1955. The gypsy moth outbreak in Michigan. *Journal of Economic Entomology* **48**:170-172.  
 Okubo, A. 1980. Diffusion and ecological problems: mathematical models. *Biomathematics* 10. Springer-Verlag, Berlin, Germany.  
 Okubo, A., P. K. Minani, M. H. Williamson, and J. D. Murray. 1989. On the spatial spread of grey squirrel in Britain. *Proceedings of the Royal Society of London*, **B 238**:113-125.  
 Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1986. *Numerical recipes in C: the art of scientific computing*. Second edition. Cambridge University Press, New York, New York, USA.  
 Pulliam, R. H. 1988. Sources, sinks and population regulation. *American Naturalist* **132**(5):652-661.  
 Ricker, W. E. 1954. Stock and recruitment. *Journal of Fisheries Research Board of Canada* **11**:559-623.  
 Royama, T. 1977. Population persistence and density dependence. *Ecological Monographs* **47**:1-35.  
 Simmons, G. A. 1987. History of the gypsy moth in Michigan. Pages 8-9 in B. A. Montgomery, editor. *Gypsy moth in Michigan: the first annual report of the gypsy moth technical committee*. Michigan Cooperative Forest Management Annual Report 87-2, East Lansing, Michigan, USA.

Skellam, J. G. 1951. Random dispersal in theoretical populations. *Biometrika* **38**:196-218.  
 Talerico, R. L. 1981. Methods for gypsy moth detection and evaluation. Pages 31-63 in C. C. Doane and M. L. McManus, editors. *The gypsy moth: research toward integrated pest management*. U.S. Forest Service Technical Bulletin 1584.  
 Tilman, D. 1994. Competition and biodiversity in spatially structured habitats. *Ecology* **75**:2-16.  
 Witter, J. A., J. L. Stoyenoff, and F. Sapio. 1992. Impacts of the gypsy moth in Michigan. *Michigan Academician* **XXV**:67-90.  
 Yasui, Y., and S. Lele. 1997. A regression method for spatial disease rates: an estimating function approach. *Journal of American Statistical Association* **92**(437):21-32.  
 Zeng, Z., R. M. Nowierski, M. L. Taper, B. Dennis, and W. P. Kemp. 1998. Complex population dynamics in the real world: modeling the influence of time-varying parameters and time lags. *Ecology, in press*.  
 Zhou, G., and A. M. Liebhold. 1995. Forecasting the spatial dynamics of gypsy moth outbreaks using cellular transition models. *Landscape Ecology* **10**:177-189.

APPENDIX

OPTIMAL LINEAR COMBINATIONS OF ESTIMATING FUNCTIONS

Let  $Y_1, Y_2, \dots, Y_n$  be independent random variables. Let  $g(Y_i; \theta)$  be an estimating function corresponding to the observation  $Y_i$  for estimating the parameter  $\theta$ . Consider a linear combination of these estimating functions  $\sum_{i=1}^n w_i g(Y_i; \theta)$ . We want to find the coefficients  $w_i$  such that the asymptotic variance of  $\theta$  is minimized. It can be shown (for example, Lindsay 1988) that the optimal weights are given by

$$w_i = \frac{E\left[\frac{d}{d\theta}g(Y_i; \theta)\right]}{E[g^2(Y_i; \theta)]}$$

where  $E(\cdot)$  denotes the expectation operator.

Let us consider the linear regression case discussed in the text. Let  $g(Y_i, X_i; \beta) = Y_i - X_i\beta$  denote the estimating functions corresponding to the  $n$  data points  $(X_i, Y_i)$ . Consider a linear combination of these estimating functions  $\sum_{i=1}^n w_i g(Y_i, X_i; \beta)$ . We want to find the coefficients,  $w_i$ , such that the asymptotic variance of  $\beta$  is minimized. Then the optimal weights are given by

$$w_i = \frac{E\left[\frac{d}{d\beta}g(Y_i, X_i; \beta)\right]}{E[g^2(Y_i, X_i; \beta)]}$$

In the linear regression situation these turn out to be  $w_i = X_i/\text{var}(Y_i)$ . The optimal linear combination is given by  $\sum_{i=1}^n [X_i/\text{var}(Y_i)](Y_i - X_i\beta) = 0$ . Notice that we give more weight to those observations that have small variance and less weight to observations with large variances.

In the context of time series data Godambe (1985) shows that the optimal weights are given by

$$w_i = \frac{E\left[\frac{d}{d\theta}g(Y_i, Y_{i-1}; \theta) \mid Y_{i-1}\right]}{E[g^2(Y_i, Y_{i-1}; \theta) \mid Y_{i-1}]}$$

where the expectations are taken conditionally, that is, behaving as if the previous time period observation,  $Y_{i-1}$ , is fixed and not random.

A simple example can be constructed for autoregressive time series for the estimation of the autocorrelation. The mod-

el is  $Y_i = \rho Y_{i-1} + \epsilon_i$  where  $\rho$  is the autocorrelation parameter. The errors,  $\epsilon_i$ , are assumed to be independent random variables with mean zero (Diggle 1991). To estimate the parameter  $\rho$ , one can use the estimating functions  $g(Y_i, Y_{i-1}; \theta) = (Y_i - \rho Y_{i-1})$ . Notice that these are functions of the data and the parameter, with expected value zero. It is easy to show that  $E[(d/d\rho)g(Y_i, Y_{i-1}; \rho) \mid Y_{i-1}] = Y_{i-1}$  and  $E[g^2(Y_i, Y_{i-1}; \rho) \mid Y_{i-1}] = \text{var}(Y_i \mid Y_{i-1})$ . Hence the optimal linear combination is given by

$$\sum_{i=1}^n \frac{Y_{i-1}}{\text{var}(Y_i \mid Y_{i-1})} (Y_i - \rho Y_{i-1}) = 0.$$

If we assume that the error variances are constant, that is  $\text{var}(Y_i \mid Y_{i-1}) = \text{var}(\epsilon_i) = \sigma^2$ , this reduces to the usual estimating function  $\sum_{i=1}^n Y_{i-1}(Y_i - \rho Y_{i-1}) = 0$ ; if the variances depend on the previous observation, each estimating function gets weighted inversely with its variance. For example, if  $\text{var}(Y_i \mid Y_{i-1}) = \text{var}(\epsilon_i) = \sigma^2 Y_{i-1}^2$ , then the optimal linear combination is given by  $\sum_{i=1}^n (Y_{i-1}/\sigma^2 Y_{i-1}^2)(Y_i - \rho Y_{i-1}) = 0$ .

An important fact to note is that suboptimal weights only reduce efficiency and that consistency and asymptotic normality of the estimators is retained (Liang and Zeger 1986). The calculation of the optimal weights can be extended to vector-valued parameters in a straightforward fashion.

ESTIMATING FUNCTIONS FOR SPATIAL STOCHASTIC RICKER MODEL

Consider the basic estimating function given in Eq. 1:

$$g(N_{i,t+1}, N_{i,t}, a, b, \delta) = \ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*$$

where  $N_{i,t}^* = N_{i,t} - \sum_{j \neq i} \delta^{|i-j|} N_{j,t} + \sum_{j \neq i} \delta^{|i-j|} N_{j,t}$ .

It is easy to check that under the assumptions on the environmental noise, this has expectation zero. Also notice that there are  $n(T - 1)$  of these estimating functions. We only have three parameters  $a, b$  and  $\delta$  to estimate. We need to combine these estimating functions so that there are three equations in three unknowns.

Consider weighted linear combinations of the basic estimating functions:

$$\sum_{t=1}^{T-1} \sum_{i=1}^n w_{i,t} (\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*).$$

Assuming that the environmental noise has constant vari-

ance, that is,  $\text{var}(\epsilon_{i,t}) = \sigma^2$ , the weights for different linear combinations for different parameters are given by

$$w_{i,t}(a) = E \left\{ \frac{d}{da} [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] \mid \mathbf{N}_t \right\}$$

$$w_{i,t}(b) = E \left\{ \frac{d}{db} [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] \mid \mathbf{N}_t \right\}$$

$$w_{i,t}(\delta) = E \left\{ \frac{d}{d\delta} [\ln(N_{i,t+1}) - \ln(N_{i,t}^*) - a - bN_{i,t}^*] \mid \mathbf{N}_t \right\}$$

where  $\mathbf{N}_t$  denotes the population vector at time  $t$ .

These weights correspond approximately to the information each basic estimating function has for the corresponding parameter (Godambe 1985). The straightforward calculations show that

$$w_{i,t}(a) = -1$$

$$w_{i,t}(b) = -N_{i,t}^*$$

and

$$w_{i,t}(\delta) \cong \frac{\sum_{j \neq i} \|s_i - s_j\| \delta^{\|s_i - s_j\|} N_{j,t} - \sum_{j \neq i} \|s_i - s_j\| \delta^{\|s_i - s_j\|} N_{i,t}}{N_{i,t} - \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{i,t} + \sum_{j \neq i} \delta^{\|s_i - s_j\|} N_{j,t}}$$

We prefer to use the approximation for  $w_{i,t}(\delta)$  because of its computational simplicity. Recall that approximate weights only affect the efficiency of the estimators but not the consistency, that is, as sample size increases, under mild regularity conditions, the estimators converge to the true population parameters.

This leads to the three estimating functions in Eqs. 2, 3, and 4.