BAYESIAN ANALYSIS OF METAPOPULATION DATA

R. B. O'Hara,[1,2,3] E. Arjas,[2] H. Toivonen,[2,4] and I. Hanski[1]

[1]*Metapopulation Research Group, Division of Population Biology, Department of Ecology and Systematics,*
P.O. Box 65 (Viikinkaari 1), FIN-00014, University of Helsinki, Finland
[2]*Rolf Nevanlinna Institute, P.O. Box 4 (Yliopistonkatu 5), FIN-00014, University of Helsinki, Finland*

Abstract. A Bayesian approach is used to develop a method for fitting a metapopulation model (the incidence function model) to data on habitat patch occupancy, providing estimates of the five model parameters. Parameter estimation is carried out using a Markov chain Monte Carlo sampler, and data augmentation is used to include the effect of missing data in the analysis. The Bayesian approach allows us to take into account uncertainty about the parameter estimates when making predictions with the model.

We demonstrate the methods of parameter estimation and prediction with simulated data. We first simulated metapopulation dynamics in real habitat patch networks with given parameter values and sampled the simulated data. Parameters were estimated both from full data sets, and from data sets with data for many years treated as missing. These estimates were then used to predict the distribution of time to extinction in modified networks, where patch areas had been reduced so that the real parameter values led to metapopulation extinction within ~30 yr. We were successfully able to fit the model and found that, in some cases, the predictions can be sensitive to one of the parameters.

Key words: Bayesian inference; data augmentation; habitat patch occupancy; Markov chain Monte Carlo (MCMC); metapopulation model; population viability analyses (PVA).

INTRODUCTION

Much of the research into metapopulation dynamics is aimed at elucidating the factors influencing long-term metapopulation persistence (Hanski 1999). This research has involved the study of empirical systems, and relating their dynamics to mathematical models. Most models have assumed that the patches are either occupied or empty, with colonizations and extinctions being stochastic and determined by a variety of factors, such as the area of the habitat patch (which is used as a surrogate for the population size). Fitting the models to empirical data is an important part of determining which factors are important in the real world.

Logistic regression (e.g., Sjögren-Gulve and Ray 1996, Ericson et al. 1999) has been used to analyze metapopulation data by estimating the probability of a patch becoming occupied or extinct at time *t* as a function of the patch occupancy pattern at a previous time, and other covariates such as patch area and isolation. With complete data, and regression being performed on the previous year's data, this approach is similar to that taken below (although the shapes of the functions are different).

Hanski (1994) constructed the incidence function

model (IFM, see *Methods*: *Incidence function model*) and derived expressions for the probability that a particular patch will be occupied, assuming quasi-stationarity. This approach has been criticized for not taking into account correlations between patch occupancies; it maximizes a pseudo-likelihood, not the true likelihood (ter Braak et al. 1998).

More recently, Moilanen (1999) developed an estimation method based on a Monte Carlo approach. Moilanen took missing years into account by simulating them, and using the transition from the last simulated year to the next (observed) year to estimate the model parameters. This approach is somewhat ad hoc, particularly as it introduces turnover limitation, which constrains the estimation by fixing the maximum number of changes in patch occupancy of the metapopulation. This constraint is not estimated from the data, but has to be chosen from other information. The method has also yet to be extended to provide confidence limits, or to the situation when some, but not all, patches are observed in a year.

In this paper we develop an approach to metapopulation model fitting that overcomes the shortcomings of the previous approaches. We adopt a Bayesian approach, which allows us to take into account uncertainty about the processes being modeled. We illustrate our approach with simulated data, using several real habitat patch networks, and demonstrate how the approach can be used to make predictions about metapopulation response to changes in the structure of the fragmented landscape.

[3] E-mail: bob.ohara@helsinki.fi

[4] Present address: Department of Computer Science, P.O. Box 26 (Teollisuuskatu 23), FIN-00014, University of Helsinki, Finland.

METHODS

Incidence function model

We will recast the incidence function model in a slightly different way from Hanski (1994). A metapopulation consists of a set, I, of patches. The state of the metapopulation at time t is denoted by the vector $\mathbf{X}(t) = (X_i(t), i \in I)$, where $X_j(t) = 0$ if patch j is unoccupied and $X_j(t) = 1$ if j is occupied. The sequence $\{\mathbf{X}(t), t = 0, 1, 2, \ldots\}$ is a Markov chain on a finite (if large) state space with $2^{|I|}$ states. The probabilities of transition from $\mathbf{X}(t - 1)$ to $\mathbf{X}(t)$ are summarized in the transition matrix, $\mathbf{P}$. This is a square matrix of size $2^{|I|} \times 2^{|I|}$, and becomes too large to handle directly if there are more than about 30 patches.

We assume that the state of each patch at time t is independent of the states of the other patches at time t, given the state of the entire metapopulation at time $t - 1$. The elements of $\mathbf{P}$ can then be calculated as the product of the probabilities for the state of each individual patch.

The probability of an unoccupied patch being colonized at time t is dependent on the rate of migration into the patch, which is measured as the connectivity of the patch, $S_i(t)$:

$$S_i(t) = \sum_{\forall j \neq i} X_j(t - 1)e^{-\alpha d_{ij}}A_j^b \qquad (1)$$

where d_{ij} is the distance between patches i and j, α is a dispersal scale parameter, A_j is the area of patch j, and b is a parameter scaling emigration to area. To convert $S_i(t)$ to the probability of colonization, $C_i(t)$, a link function is used:

$$C_i(t) = \frac{S_i(t)^2}{S_i(t)^2 + y^2} \qquad (2)$$

where y is a parameter.

If patch i is occupied at time t, the population will go extinct with probability $E_i(t)$:

$$E_i(t) = \min\left(1, \left(\frac{A_0}{A_i}\right)^x\right)(1 - C_i(t)) \qquad (3)$$

where A_0 is the minimum patch area for viable populations (i.e., $E_i(t) = (1 - C_i(t))$ for $A_i < A_0$), x is a parameter and the $1 - C_i(t)$ term takes into account the rescue effect (Hanski 1999). The model has five parameters, which can all be estimated from patch occupancy data.

Bayesian estimation

Here we take a Bayesian approach to estimation (e.g., Gelman et al. 1995). A Bayesian analysis of data can be seen as a process of using data to change our knowledge about the values of parameters of a model. This is done by applying Bayes' formula:

$$\text{pr}(\boldsymbol{\theta}|\text{data}) \propto \text{pr}(\boldsymbol{\theta})\text{pr}(\text{data}|\boldsymbol{\theta}) \qquad (4)$$

where $\text{pr}(\boldsymbol{\theta}|\text{data})$ is the posterior density of the param-

eters ($\boldsymbol{\theta}$), given the data. The expression is made up of the prior density ($\text{pr}(\boldsymbol{\theta})$), a description of where the parameters are thought to lie before the data are analyzed, and the likelihood ($\text{pr}(\text{data}|\boldsymbol{\theta})$), which is the contribution to our knowledge given by the data.

One strength of the Bayesian approach is that it deals with uncertainty explicitly. Of particular relevance in this paper is the way missing data and prediction are handled. Nuisance variables (whether they are nuisance parameters, or missing data) can be accounted for by integrating them out:

$$\text{pr}(\theta_1 | \text{data}) = \int \text{pr}(\theta_1, \theta_2 | \text{data}) \, d\theta_2$$

$$= \int \text{pr}(\theta_1 | \theta_2, \text{data})\text{pr}(\theta_2 | \text{data}) \, d\theta_2 \qquad (5)$$

where θ_2 is a nuisance parameter. This can be interpreted as taking a weighted average of the conditional density for θ_1 ($\text{pr}(\theta_1|\theta_2, \text{data})$), weighted by the probability density for the value of θ_2 ($\text{pr}(\theta_2|\text{data})$). A similar trick is used for making predictions based on a fitted model. The parameters of the model can be viewed in a similar way to nuisance parameters, and they can also be integrated out:

$$\text{Pr}(\text{prediction} | \text{data})$$

$$= \int \text{Pr}(\text{prediction} | \boldsymbol{\theta}, \text{data})\text{Pr}(\boldsymbol{\theta} | \text{data}) \, d\boldsymbol{\theta}. \qquad (6)$$

This has a similar interpretation to Eq. 5.

Model fitting

Typical metapopulation data consist of a set of observations of patch occupancies, $\mathbf{X}(0), \mathbf{X}(t_1), \mathbf{X}(t_2), \ldots, \mathbf{X}(t_m)$, made at m times $0 < t_1 < t_2 < \ldots < t_m$. The corresponding likelihood of these observations is the product of the likelihoods for the transitions between the states, and the likelihood of reaching the first state, i.e.,

$$\mu(\mathbf{X}(0))\mathbf{P}^{(t_1)}(\mathbf{X}(0), \mathbf{X}(t_1)) \ldots \mathbf{P}^{(t_m - t_{m-1})}(\mathbf{X}(m) - 1), (\mathbf{X}(m)) \qquad (7)$$

where μ is the (initial) distribution of $\mathbf{X}(0)$ and $\mathbf{P}^{(t)}(x, x')$ is the (x, x')-element of the tth power of $\mathbf{P}$. In this paper statistical inference will be done conditionally on the initial state $\mathbf{X}(0)$, which amounts to ignoring the corresponding contribution $\mu(\mathbf{X}(0))$. Fitting this model to data requires a numerical approach. Here we use Markov chain Monte Carlo (MCMC) (Gilks et al. 1996b), which draws a large number of points from the posterior distribution.

Of particular importance here is how missing data are handled. If the data are not from consecutive years, then we need to compute elements of the matrices $\mathbf{P}^t$ with $t > 1$. This is unrealistic in practice (because of the size of $\mathbf{P}$, as noted above), so instead we do the

computations approximately, by applying data augmentation methods. In essence, this means repeatedly filling in the missing data, and using this augmented data set to estimate the parameters. The trick is to fill in the missing data from the correct distribution. For Bayesian inference, we are aiming to obtain a distribution that is a numerical approximation of Eq. 5.

For an illustration, consider the case in which $\mathbf{X}(0)$ and $\mathbf{X}(2)$ are observed, but the value of $\mathbf{X}(1)$ is missing. The conditional distribution of the unknown $\mathbf{X}(1)$ given the entire data will depend only on the neighboring values in time, that is, $\mathbf{X}(0)$ and $\mathbf{X}(2)$:

$$\Pr(\mathbf{X}(1)\,|\,\mathbf{X}(0),\,\mathbf{X}(2))$$
$$\propto \Pr(\mathbf{X}(0),\,\mathbf{X}(1))\Pr(\mathbf{X}(1),\,\mathbf{X}(2)). \qquad (8)$$

Straightforward simulation of the vector $\mathbf{X}(1)$ is impossible in practice because the occupancy of each patch $\mathbf{X}_j(1)$, $j \in I$, cannot be drawn independently of the other patch states. Instead, a version of the Metropolis-Hastings (M-H) algorithm can be used to give a sequence $\mathbf{X}^{(1)}(1), \mathbf{X}^{(2)}(1), \ldots, \mathbf{X}^{(n)}(1)$ of n simulated occupancy patterns, whose limiting distribution will coincide with that given by Eq. 8. If more than one year's data are missing, then the procedure is easily extended. For example, if years 1 and 2 are missing, then we augment the data by simulating initial values $\mathbf{X}^{(0)}(1)$ and $\mathbf{X}^{(0)}(2)$, then use the M-H algorithm to produce a new value $\mathbf{X}^{(1)}(1)$ given $\mathbf{X}(0)$ and $\mathbf{X}^{(0)}(2)$, and then a new value $\mathbf{X}^{(1)}(2)$ given $\mathbf{X}^{(1)}(1)$ and $\mathbf{X}(3)$. Again, this can be iterated for n steps, and the limiting distributions will be correct. The extension of this procedure to the situation in which data are missing for more than two consecutive years is obvious. If only some of the data from a year are incomplete, then we only use data augmentation for the data that are missing.

From the estimates of the missing data, we can compute numerical approximations of the likelihood function, and use these likelihoods for the estimation of the parameters of the model. This scheme can be used to obtain maximum likelihood (ML) estimates via a stochastic EM algorithm (Diebolt and Ip 1996). As we are taking a Bayesian approach, we combine the estimation of the missing data with that of the parameters directly, to produce a set of n vectors of the parameters, whose limiting distribution is their posterior distribution (Eq. 4).

Implementation

Since the IFM specifies the likelihood for the observations, all we need to do to set up a Bayesian model is to specify the joint prior distribution of the model parameters. We defined the priors for all five parameters to be independent exponential distributions, with means equal to 100. Although these are not uninformative, they are vague, i.e., they are almost flat around the parameter space where the (known) parameter values lie.

The model was fitted to the data with a MCMC algorithm, using the Bassist, version 0.8.4 (Toivonen et al. 1998). Parameters were updated using single component M-H, proposing from a normal distribution centered at the current value, with a variance that had been tailored to give reasonable mixing (from 10^{-5} for A_0 to 10^{-2} for α and y). Augmented data were updated using a Gibbs sampler (Gilks et al. 1996a).

For the MCMC computation, we ran five chains simultaneously. For each chain, a burn-in of 5000 iterations was used, after which the next 8000 iterations were thinned by taking every fourth iteration, to give a total of 2000 samples. These runs gave a five-dimensional cloud of 10 000 points representing the posterior distribution. Convergence was assessed by eye. We also took 10 000 samples from the prior distribution, to be used as a reference.

Simulated data

The fitting of the model was investigated by fitting it to a metapopulation simulated in the network of habitat patches of the silver-spotted skipper butterfly (*Hesperia comma*) described in Hanski and Thomas (1994). We used two sets of parameters (Table 1). Set 1 is close to values that had been estimated previously (Hanski 1994), and set 2 was chosen to allow the simulated metapopulation to persist, but with dynamics different to those determined by set 1. Both sets of parameters gave long-term persistence in the patch network. The metapopulation was simulated from an initial random configuration for 100 years, after which the following 50 years of data were sampled. Some statistics about the data sets are shown in Table 2. In both cases there was a reasonable amount of population turnover between years, though the turnover was greater with the first set of parameters. The parameters of the model were then estimated using the first 10 years, the first 20 years, and the full 50 years of the data, as well as just the first and 10th years present, and all of the intervening years being treated as missing.

In order to illustrate the application of a Bayesian analysis to prediction, we used the posterior distributions to make predictions about the dynamics of the metapopulation in an altered patch network. We did this by simulating the dynamics in the same network from which the parameters were estimated but with the areas of all of the patches having been reduced by the same factor (Loss in Table 2, perhaps representing a deterioration in habitat quality) so that with the real parameters the average time to extinction was ~30 generations. The simulations were started at the observed final year's occupancy pattern, and run for up to 150 years, with the time to extinction being recorded. We made 10 000 simulations using the real parameter values, and one simulation for each parameter set from the posterior distribution.

We compared our method with the one proposed by Moilanen (1999) by fitting the model to several 10-yr-long simulated time series. As well as the data sets generated above, we used data sets generated by sim-

TABLE 1. Means and standard deviations (σ) of posterior distributions estimated from data sets simulated in a *Hesperia comma* patch network.

Data used	Statistic	x	α	y	b	A_0
Set 1						
Real parameters		0.98	1	2	0.5	0.22
Years 2–9 removed	mean	3.95	1.00	3.02	0.92	11.9
	σ	3.38	0.37	3.89	0.75	10.0
All 10 years	mean	1.04	1.11	2.03	0.44	0.17
	σ	0.24	0.17	0.55	0.18	0.047
20 years	mean	1.08	0.95	2.20	0.42	0.21
	σ	0.20	0.11	0.42	0.13	0.05
50 years	mean	0.92	0.93	2.23	0.44	0.21
	σ	0.10	0.069	0.25	0.077	0.027
Set 2						
Real parameters		0.98	1	8.61	0.5	0.011
Years 2–9 removed	mean	2.48	0.67	14.95	0.96	13.66
	σ	3.14	0.41	12.90	0.75	19.22
All 10 years	mean	1.11	0.98	14.05	0.92	0.017
	σ	0.20	0.38	9.08	0.47	0.007
All 20 years	mean	0.95	0.88	9.12	0.48	0.014
	σ	0.090	0.21	2.89	0.25	0.003
All 50 years	mean	0.98	0.88	10.0	0.51	0.014
	σ	0.064	0.13	1.84	0.17	0.002

ulating metapopulations in four networks (Table 3), three networks of butterflies (networks B, C, and D), and one of the American pika (network E). Parameters were chosen in the same fashion as for network A and are also given in Table 3.

We estimated x, y, and A_0 (keeping b and α at their true values) using our method and the TMC method of Moilanen (1999). The TMC method was used as the assumptions of the model are identical to ours. The true values for x, y, and A_0 were used as initial values in the TMC estimation. Estimation was carried out at least twice, and continued until convergence appeared to have been achieved, and the value with the largest likelihood taken. The estimated and real parameters were then used to simulate the metapopulation in a reduced network as above, with the point estimate from the TMC method being used 10 000 times.

RESULTS

The means and variances of the estimated posterior distributions for the estimations for the data sets created in network A are given in Table 1. With only the first and last years' data, the posterior means of several parameters (particularly A_0) are located some way from the real values. However, the standard deviations are large, so this is a result of the lack of precision, rather than a bias in the estimation. As more data are added, the variances decrease and the posterior means are located closer to the real values.

The predictions for the reduced networks are shown in Fig. 1, where the proportion of simulations that had not gone extinct by time t is plotted against t. The posterior predictions for set 2 are close to the predictions obtained with the real parameter values, whereas the predictions for set 1 are very different. As more

information is added, the curves only slowly move from the prior prediction toward the prediction based on the real parameter values.

We investigated whether the data set itself was somehow abnormal by repeating the process of simulation of 10 years of data, estimation, and prediction in a reduced network a further 20 times (a single chain of 5000 values was used for the MCMC estimation, with burn-in and thinning as above). The survival curves are plotted in Fig. 1, and it is clear that the poor estimates for set 1 are not atypical. Larger values of b generally lead to shorter times to extinction (Fig. 2), because larger values of b lead to smaller connectivities, and hence the overall amount of migration is less. The parameter set chosen is close enough to an area in the parameter space

TABLE 2. Summary statistics for two simulated 10-yr-long metapopulations data sets; averages are per year.

Statistic	Set 1	Set 2
Loss parameter	3.5	30
Average proportion of occupied		
Patches	0.87	0.79
Area	0.93	0.92
Average number of		
Colonizations	4.7	2.0
Extinctions	5.2	2.3
Changes between first and last year		
Colonizations	6	4
Extinctions	11	7

Note: The loss parameter is the amount by which the patch areas were divided when the areas were reduced to demonstrate the predicted dynamics from posterior estimates of the parameters.

TABLE 3. Parameter values used to simulate 10 data sets in patch networks of five different species.

Species	Source	Data set	x	α	y	b	A_0
Hesperia comma	Hanski and Thomas (1994)	A1	0.98	1	2	0.5	0.22
		A2	0.98	1	8.61	0.5	0.011
Scoliantides orion	Hanski (1994)	B1	0.96	1	4.5	0.5	0.01
		B2	0.96	1.3	4.5	0.5	0.01
Melitaea cinxia	Hanski (1994)	C1	0.95	1	3.0	0.5	0.2
		C2	0.95	1	0.86	0.5	0.2
Melitaea diamina	Wahlberg et al. (1996)	D1	0.9	1	3.3	0.5	0.01
		D2	0.9	0.5	3.3	0.5	0.01
Ochotona princeps	Smith and Gilpin (1997)	E1	1.7	2	2.5	0.5	0.01
		E2	0.5	2	2.5	0.5	0.0025

where the dynamics of the reduced network are sensitive to changes in b, so different parts of the posterior distributions cover the regions giving different outcomes of the simulations.

The predictive survival curves from our method and the TMC method are shown in Fig. 3. As above, as more data are added into the Bayesian predictions, the estimations tend to converge to the real value, and the slope of the survival curve becomes steeper, even when

there is little change in the average time to extinction, reflecting a greater certainty in the predictions. The TMC method also tends to do better with all the data, but because the predictions are based on point estimates, the slope is determined solely by the mean time to extinction, and there is no indication of uncertainty. With missing data the TMC method often provides parameter estimates and predictions that are inaccurate (e.g., C1, C2, and D1). This may be because the dif-
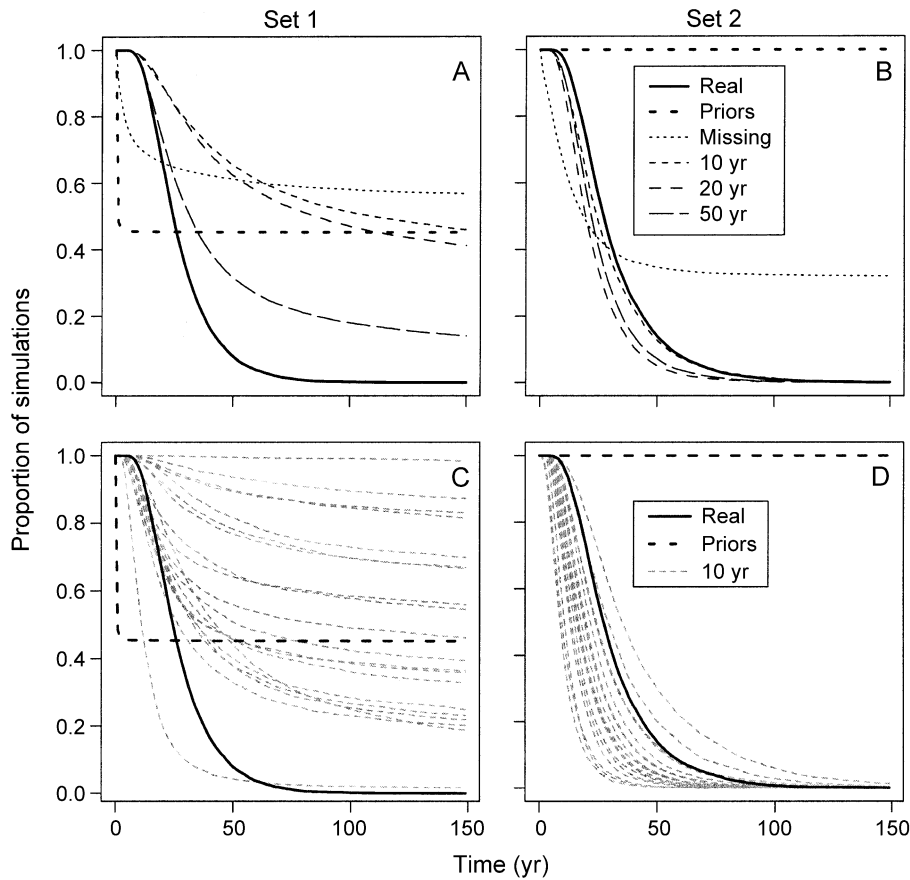

FIG. 1. Predictive survival curves (i.e., the proportion of simulations in which the metapopulation had not gone extinct) simulated from posterior distributions estimated from data simulated with two sets of parameters in a *Hesperia comma* patch network. (A, B) predictions from estimations made using different amounts of data; (C, D) predictions from replicate 10-yr simulations.
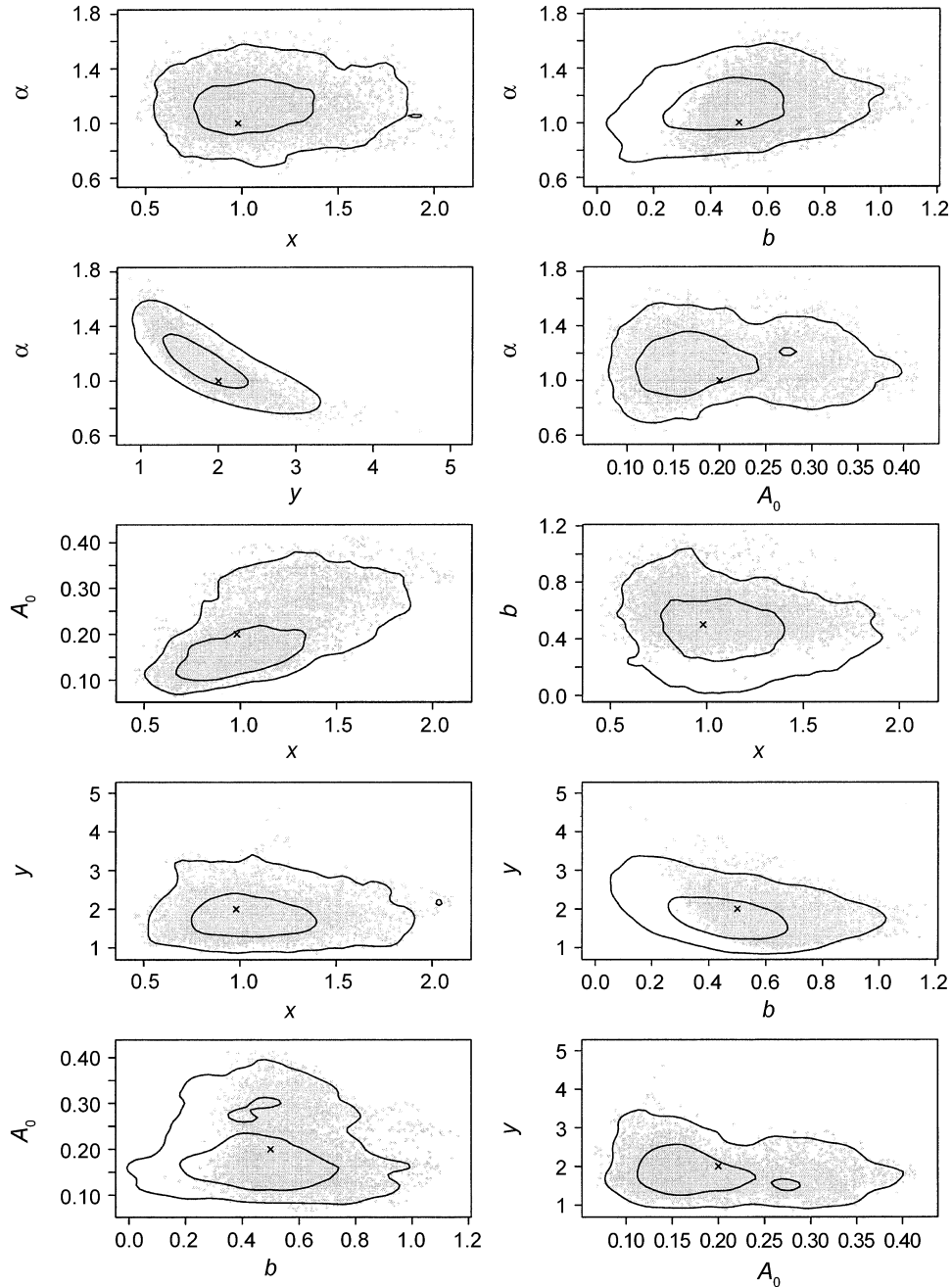
FIG. 2. Scatter plots of posterior samples for the set-1 data set, all 10 years of data used in the estimation. Contours give the 50% and 95% posterior density regions. Grey points are the posterior samples for which the corresponding simulation went extinct within 150 generations. The black cross shows the location of the true parameter values.

ference in the likelihood between the true ML estimate and the reported estimate is so small that it is obscured by the stochastic noise in the simulations of the missing data that are being used to estimate the parameters.

DISCUSSION

By taking a Bayesian approach, we have been able to directly quantify our uncertainty in the parameters of a metapopulation model and the predictions arising from the parameter estimates. One property of the Bayesian approach is that no additional confidence statement about the predictive survival curves is necessary; they represent the predicted probabilities of survival, given the observed data and the model. A greater knowledge about the parameters leads to a curve with a steeper slope, whereas a curve based on a point estimate would
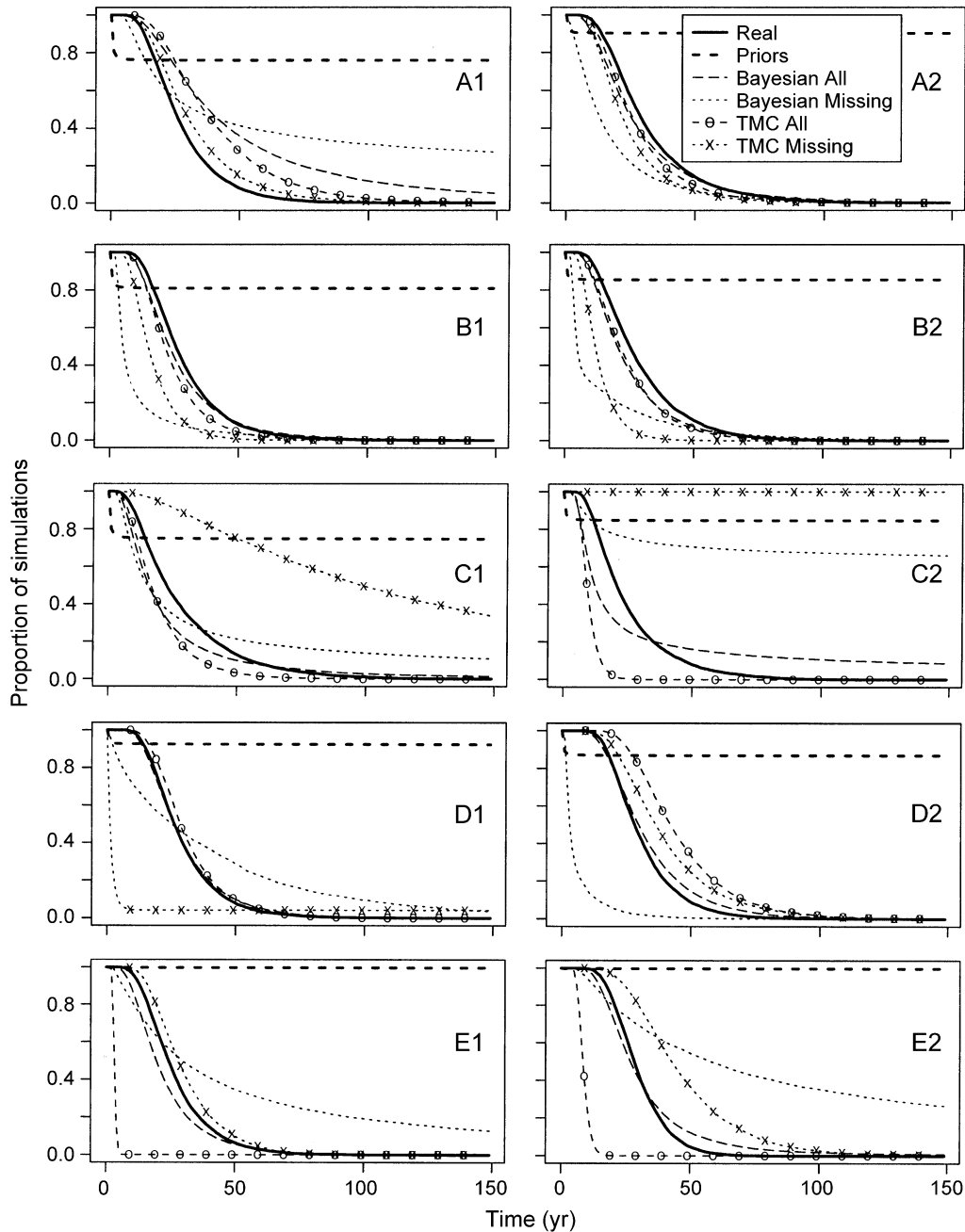
FIG. 3. Predictive survival curves for ten 10-yr long data sets simulated using patch networks and parameter sets listed in Table 3, comparing the estimation methods detailed here and in Moilanen (1999).

have the same slope regardless of the amount of information, but the uncertainty in the predictions would be seen in the width of the confidence limits.

The reason for having uncertainty in the parameter estimates is that different realizations of the real metapopulation process will give slightly different estimates of the posterior distribution. As suggested by Ludwig (1996), and shown here, the uncertainty in the estimated parameters can be magnified to give very different predictions. Although the *b* parameter can have a large influence on the predictions in our examples, it is possible that this is an artifact of the combination of model, parameters, and mechanism we used to reduce the network size rather than a biological phenomenon. Unfortunately, with real data we do not have the benefit of knowing the real underlying process, so model uncertainty should also be assessed.

With our emphasis on prediction, it is clear that our

approach can be used for population viability analyses (PVA) of metapopulations. Although a Bayesian approach has been advocated for PVAs (e.g., Ludwig 1996, Virtala et al. 1998), we are not aware of any real applications in a conservation context. We would hope that this will change as ecologists become aware of the advantages of a Bayesian approach in allowing us to tackle uncertainty in a rigorous fashion, and of the use of numerically intensive tools for model fitting (e.g., MCMC).

LITERATURE CITED

Diebolt, J., and E. H. S. Ip. 1996. Stochastic EM: method and application. Pages 259–273 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall, London, UK.

Ericson, L., J. J. Burdon, and W. J. Müller. 1999. Spatial and temporal dynamics of the rust fungus *Uromyces valerianae* on populations of its host *Valeriana salina*. Journal of Ecology **87**:649–658.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, London, UK.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996*a*. Introducing Markov chain Monte Carlo. Pages 1–19 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall, London, UK.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996*b*.

Markov chain Monte Carlo in practice. Chapman and Hall, London, UK.

Hanski, I. 1994. A practical model of metapopulation dynamics. Journal of Animal Ecology **63**:151–162.

Hanski, I. 1999. Metapopulation ecology. Oxford University Press, Oxford, UK.

Hanski, I., and C. D. Thomas. 1994. Metapopulation dynamics and conservation: a spatially explicit model applied to butterflies. Biological Conservation **68**:167–180.

Ludwig, D. 1996. Uncertainty and the assessment of extinction probabilities. Ecological Applications **6**:1067–1076.

Moilanen, A. 1999. Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. Ecology **80**:1031–1043.

Sjögren-Gulve, P., and C. Ray. 1996. Large scale forestry extirpates the pool frog: using logistic regression to model metapopulation dynamics. Pages 111–137 *in* D. L. McCullough, editor. Metapopulations and wildlife conservation. Island, Washington, D.C., USA.

Smith, A. T., and M. Gilpin. 1997. Spatially correlated dynamics in a pika metapopulation. Pages 407–428 *in* I. Hanski and M. E. Gilpin, editors. Metapopulation biology. Academic Press, London, UK.

ter Braak, C. J. F., I. Hanski, and J. Verboom. 1998. The incidence function approach to the modeling of metapopulation dynamics. Pages 167–188 *in* J. Bascompte and R. V. Solé, editors. Modeling spatiotemporal dynamics in ecology. Springer-Verlag, Berlin, Germany.

Toivonen, H., H. Mannila, M. Salmenkivi, and K.-P. Laakso. 1998. Bassist—a tool for MCMC simulation of statistical model. Pages 590–595 *in* Proceedings of the Eurosim '98 Simulation Congress, (Helsinki, Finland, April 1998). Federation of European Simulation Societies, University of Vienna, Vienna, Austria.

Virtala, M., S. Kuikka, and E. Arjas. 1998. Stochastic virtual population analysis. ICES Journal of Marine Science **55**: 892–904.

Wahlberg, N., A. Moilanen, and I. Hanski. 1996. Predicting the occurrence of endangered species in fragmented landscapes. Science **273**:1536–1538.