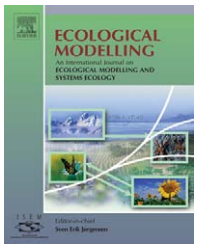
available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevier.com/locate/ecolmodel



The estimation of dispersal rates using the covariance of local populations

Anne Schneeberger[a,b], *Vincent A.A. Jansen*[c,*]

[a] *Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom*
[b] *Department of Infectious Disease Epidemiology, Imperial College London, Norfolk Place, London W2 1PG, United Kingdom*
[c] *School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom*

ARTICLE INFO

Article history:
Received 7 June 2005
Received in revised form 17 January 2006
Accepted 9 February 2006
Published on line 17 April 2006

Keywords:
Time series analysis
Autoregressive process
Migration
Dispersal rate
Metapopulation

ABSTRACT

Dispersal has important implications on an individual, population and metapopulation level. Dispersal rates, however, are difficult to measure. In this paper we introduce a method that uses census data, i.e. repeated counts of number of individuals per location and point in time, to estimate dispersal rates. The rationale underlying this method is that local stochastic disturbances which dissipate in subdivided populations create a covariance structure, the details of which depend on how strongly the local populations are coupled. This covariance structure can be used to estimate dispersal rates. We describe this process using a stochastic model for growth and dispersal which explicitly accounts for the geometry of the patchy population. A regression of the covariance structure of this model is then used to infer the growth rate near equilibrium and the dispersal rate. We study the distribution of the estimated parameters and obtain confidence intervals using a bootstrap analysis and a Monte Carlo technique. We study how the confidence intervals depend on the model's parameters, the robustness of the estimating scheme, and discuss the applicability of our method.

1. Introduction

Many natural populations are subdivided and occur in habitat patches (Husband and Barrett, 1996; Hanski, 1999; Thrall et al., 2000; Barton, 2001; Chesson, 2001). Ecologists and geneticists have formulated models describing metapopulations, i.e. populations consisting of "local populations", each of which have a substantial probability of extinction, but which can persist at a regional level (Wright, 1940; Levins, 1970; Hanski, 1999, 2001). Dispersal between these local populations is an essential feature of a spatially structured population.

The importance of dispersal has been recognised in empirical and theoretical studies. However, quantification of dispersal is difficult as measuring dispersal rates in natural populations is a time and labour consuming enterprise. (For recent reviews on this issue see Clobert et al., 2001; Nathan et al., 2003; Cain et al., 2003 and other papers in same issue of *Ecology*.) So far two different methodologies have been used to measure dispersal: (1) one can observe marked *individuals* and track movement and redistribution (Stensteth and Lidicker, 1992; Kaiser, 1995), (2) one can infer dispersal by the redistribution of a *population* of markers (Slatkin, 1985; Barton, 2001).

Individual movement-redistribution methods use data on observations of individuals. In most cases, animals will be captured, a mark applied, and the animal released. Subsequent reobservation generates (a) recovery data, in which animals are recovered dead, (b) recapture/resighting data, or (c) known-status data, in which marked animals are re-observed alive or dead at specified times. Evaluation of the data using an appropriate model yields dispersal-related parameters (Bennets et al., 2001). Direct tracking of individuals involves the use of

* *Corresponding author*. Tel.: +44 1784 443179; fax: +44 1784 470756.
E-mail address: vincent.jansen@rhul.ac.uk (V.A.A. Jansen).
0304-3800/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2006.02.031

radio tags, satellite tags or electronic data storage tags (see references in Nathan (2001)). Although this approach potentially gives good results, its main disadvantages are associated technical difficulties, the labour-intensive way of collecting the data, and above all the high costs of the tags. Moreover, the data gathered using these direct methods might not represent the full distribution of individual movements and are temporarily and spatially restricted (Koenig et al., 1996).

In case of the redistribution of a population of markers, the markers used are mostly genetic e.g. microsatellite loci, mitochondrial DNA, ribosomal DNA or allozymes. The distribution of neutral genetic markers can be predicted from a model and the observed distribution can be compared to the distribution predicted by the model. The data necessary for these indirect techniques are easy to gather: a single sampling event in time generates a snapshot of the current state of genetic variation which is used to calculate the number of migrants under a certain model. Coalescence methods and *F*-statistics are mostly based on the island model (Wright, 1940) and are widely used to measure gene flow within and among populations and infer dispersal rates. This model is the most widely used, and is based on a large number of assumptions, which are all too often violated in real populations. This is the main reason why studies using *F*-statistics to infer dispersal patterns have been widely criticized (Whitlock and McCauley, 1999; Cain et al., 2000; Barton, 2001; Rousset, 2001 and references therein). While a stepping stone model is often more realistic, and is also used as a basis to calculate statistics of interest in theoretical papers (Slatkin and Barton, 1989; Rousset, 1996), it is rarely applied to determine dispersal rates in studies of natural populations.

A more fundamental shortcoming of such methods is that for population ecological purposes one usually needs estimates of dispersal rates at the timescale of generations, such as the "instantaneous" dispersal rates determined by direct measurements. The rates obtained by genetic approaches reflect an average over a period of time whose length depends on mutation rates and genetic drift. This means that these indirect (genetic) methods of estimating dispersal rates are of limited use for population ecologists (Hanski, 2001).

Here, we will illustrate how spatial census data, i.e. repeated counts of number of individuals per location and point in time, can be used to estimate dispersal rates. The rationale underlying this method is that natural populations are subject to local stochastic disturbance due to, for instance, demographic stochasticity or the effect of local weather. The creation and dissipation of such perturbations will create a typical covariance structure. These perturbations will dissipate quickly through the population if the population is highly connected and there is much dispersal, if there is limited dispersal the dissipation will be slow.

Various formalisms to describe dispersal have been used in the formulation of ecological models (see Czaran, 1998 for a comprehensive review). Among the different approaches are models that describe a continuum of space in the form of partial differential equations, individuals based models, and multi-patch or metapopulation models. The last category forms a convenient middle ground between realism and tractability and this explains the popularity of this approach (for recent examples in the ecological modelling literature see Etienne (2004), Hein et al. (2004), Hovestadt and Poethke (2006), Matter (2001), Metzger (2005), Reed and Levine (2005), Pfenning et al. (2004) and Singh et al. (2004).) Here, we describe a generic model for dispersal in coupled populations using a model for population growth and dispersal and from this description derive a method to infer dispersal rates from census data. The model describes an ecological system of coupled local populations which are reasonably close to their equilibrium. As the equilibrium value, its stability properties and the topology of the spatial system can be chosen freely, this approach provides a generic description of diffusive systems. The model thus generalises the description of dispersal used previously (see e.g. Hassell et al., 1991; Rohani et al., 1996; Czaran, 1998; Bascompte and Sole, 1998). Our model extends earlier work in that it explicitly accounts for a large class of topologies and thus avoids simplifying assumptions on the geometry of the habitat. We develop a statistical method to analyze these data and estimate parameters from the covariance structure, this provides a novel and generic way to estimate dispersal rates and is an advance compared to other methods, in particular those based on genetics, in that this provides an estimate at an ecological relevant timescale. We study the distribution of the estimator, and we obtain confidence intervals of the dispersal rate. Finally, we discuss the robustness of this estimation scheme.

2. A generic model for dispersal

We will start this section with a brief explanation of the model structure. For ease of explanation we will start with a strictly deterministic model. We will next use this model as a basis for a stochastic model. For clarity we have formulated the model as a non-structured, single species model without density-dependent dispersal. These model can be easily generalised, in Appendix A we outline how this can be done.

2.1. *The deterministic model*

The model we will use is a simple linear model describing the changes in population size. The model describes the population dynamics of a local population in the vicinity of its equilibrium.

Let us assume that the population consists of n subpopulations. In the absence of dispersal the linearised dynamics take the form

$$x_{j,t+1} = a x_{j,t}, \quad j = 1, \ldots, n \tag{1}$$

where $x_{j,t}$ is the deviation of the density of population j from its equilibrium density at time t. Note that the first index refers to location, and the second to time. The parameter a can be interpreted as the *per capita* growth rate of the population near equilibrium. This is a generic description that can cover a large class of growth models. To see this, suppose that the local population grows according to some growth function F: so that $N_{j,t+1} = F(N_{j,t})$. Suppose there exists an equilibrium population size $\tilde{N}$ defined as $F(\tilde{N}) = \tilde{N}$. The constant a is the derivative of F evaluated at the equilibrium point: $a = (\mathrm{d}F/\mathrm{d}N)|_{N=\tilde{N}}$. If $x_{j,t} = N_{j,t} - \tilde{N}$ is the deviation from the equilibrium density $\tilde{N}$,

the deviation from equilibrium approximately changes over time as described by Eq. (1) which describes the exponential decay of the deviation from the equilibrium. Small deviations from the equilibrium die away if the absolute value of a is smaller than 1, so in that case the equilibrium is stable. If $0 < a < 1$ the deviations converge monotonously to zero, if $-1 < a < 0$ damped oscillations to the equilibrium are found. The characteristic return time for the deviation is given by $-1/\ln|a|$. Note that we have assumed that all populations have the same equilibrium density and the same stability properties, such an assumption is easily justified if the environmental conditions are equal in all patches. This does not imply that all patches are of the same size; all we require is that the equilibrium density is the same in all patches.

Next we include dispersal in our model. We will assume, as is frequently done for this class of models (e.g. Hassell et al., 1991), that during each time step reproduction and dispersal occur sequentially. The reproduction process is as described above: the population linearly approaches its equilibrium. Next a fraction of the population disperses from a patch and is distributed over the other patches in the system. Let m be the fraction of the local population that disperses. Every time step a fraction m of the local population is removed and distributed over all patches. We assume that this distribution is known and reflects the way the different locations are connected. The strength of the connection between location i and j is given by c_{ij}. Thus, with m being the total fraction of the population that disperses, $c_{ij}m$ denotes the fraction in location i that disperses to location j and $1 + c_{ii}m$ the fraction in location i that stays at location i. Dispersers can return to their patch of origin, so we do not require $c_{ii} = -1$. Since dispersal will not increase the numbers in the patch of origin, we require $c_{ii} \leq 0$ for all i, and dispersers arriving at other patches increase the number there, so count positively, $c_{ij} \geq 0$ for all $i \neq j$. Because it is reasonable to assume that the dispersal cannot redistribute more individuals than are locally available we require $1 + mc_{ii} \geq 0$. For further details of the connectivity matrix see below. The combined process of reproduction and dispersal are described by

$$x_{j,t+1} = ax_{j,t} + m\sum_{i=1}^{n} c_{ij}\, ax_{i,t} \tag{2}$$

To analyze this model it is convenient to express this model in vector notation, following Jansen and Lloyd (2000). Let $\mathbf{x}_t$ be a row vector that has $x_{j,t}$ as elements, i.e., $\mathbf{x}_t = (x_{1,t}, \ldots, x_{n,t})$. Using this we can rewrite (1) as

$$\mathbf{x}_{t+1} = a\mathbf{x}_t. \tag{3}$$

To express dispersal in terms of vectors we first gather the c_{ij} in the $n \times n$ connectivity matrix $C = \{c_{ij}\}$ (the first subscript of c_{ij} refers to the rows of C, the second to the columns). The dispersal term in (2) can now be written as $am\mathbf{x}_tC$. Note that because we defined $\mathbf{x}$ as a row vector, dispersal is described by a post-multiplication. The complete spatial model with dispersal (2) can be rewritten as

$$\mathbf{x}_{t+1} = a\mathbf{x}_t(I + mC) \tag{4}$$

The advantage of this notation is that it allows us to bring to bear tools of linear algebra. In what follows, we will transform the vector $\mathbf{x}$, using a similarity transform, such that the n coupled processes described by Eq. (4) decouple (Jansen and Lloyd, 2000). Before we explain this in detail we will first describe the stochastic extension of this generic model for dispersal.

2.2. The stochastic model

Consider ecological or epidemiological population census data $N_{j,t}$ with $1 \leq j \leq n$ and $0 \leq t \leq T$ of a single species, in n locations and regularly spaced in time. These data can be species densities or number of individuals. Such data, recorded at different times, are an example of a time series. Here, we will use techniques for time series analysis based on description of the data by an autoregressive process. This is a stochastic process in which the next value is generated by multiplying the current value of the process with a constant and adding to it a random number with zero mean (Box and Jenkins, 1970).

Because our model describes deviations from equilibrium, the first step in the analysis is centering the data around zero by subtracting the mean value of the data ($\mu = (1/nT)\sum_{j=1}^{n}\sum_{t=1}^{T} N_{j,t}$), from all data $N_{j,t}$. If there is a clear transient in the data it is advisable to discard the data in the transient. In this way we get centered measurements $x_{j,t} = N_{j,t} - \mu$. We can now describe these centered measurements by a first-order autoregressive model (AR1), i.e. We will follow the deterministic model in the previous section and assume that, in the absence of dispersal, the density at location j at any point in time depends only on the value of the previous time point (times a proportionality constant a) and an additive noise term:

$$x_{j,t+1} = ax_{j,t} + \xi_{j,t} \quad j = 1, \ldots, n \tag{5}$$

The fluctuations $\xi_{j,t}$ are assumed to have time-independent variance σ^2 and mean zero and to be uncorrelated over time, although they may spatially be correlated. We assume that the noise at time t, $\xi_{j,t}$, is independent of the local population sizes at the same time point, $x_{j,t}$. This assumption is likely to be justified even if the noise depends on the population size, as is the case for demographic stochasticity, as long as the dependence is weak and the population remains close to its equilibrium value. It is easy to show that for $x_{j,t}$ generated by this process the mean is 0. The process has a stationary distribution if and only if $|a| < 1$ (see e.g. Box and Jenkins, 1970; Karlin and Taylor, 1975). The stochastic term ξ includes all non-deterministic processes such as responses to changes in the environment. As above we assume that the environment is homogeneous, i.e., the growth rate near equilibrium, a, is the same at all locations. In our simulations we assumed that the noise is normally distributed, yet our general conclusions do not depend on this assumption.

Next we include dispersal in this process in a similar way as in the previous section: we assume that a fraction m is removed from a location and redistributed over all patches according to the parameters in the connectivity matrix. By combining the autoregressive process with dispersal we obtain the following equation:

$$x_{j,t+1} = ax_{j,t} + \xi_{j,t} + m\sum_{i=1}^{n} c_{ij}\,(ax_{i,t} + \xi_{i,t}) \tag{6}$$

Using vector notation we can rewrite (6) as

$$\mathbf{x}_{t+1} = (a\mathbf{x}_t + \xi_t)(I + mC) \tag{7}$$

where ξ_t is a row vector that has $\xi_{j,t}$ as elements, i.e., $\xi_t = (\xi_{1,t}, \ldots, \xi_{n,t})$. This notation is an economic way to describe all n autoregressive processes at once.

In our model dispersal is realised in a deterministic fashion, i.e. of the fraction m of the population that disperses, mc_{ij} migrate to any adjacent patch. Particularly if the number of individuals in a patch is low, there could be considerable stochasticity associated with the dispersal process. Simulations studies (not shown here) illustrated that even for relatively small number of individuals (5–10), the process as described provides an adequate description of this additional stochasticity and the method described here to estimate dispersal rates gives good results even for small number of individuals.

2.3. *The connectivity matrix*

The connectivity matrix holds the information on the geometry of the habitat patches. Because the geometry is a central feature of spatial systems we will discuss possible forms of the connectivity matrix in some detail (see Fig. 1). In the analysis presented in this paper we assume that this matrix is known. This requires that one needs to have knowledge about the relative amount of dispersal taking place between local populations. This could be inferred from the distance between local populations or other information about the topology, or it could be known through marking and recapturing individuals (for instance, through ringing birds). We do not assume that the total fraction of the population involved in dispersal is known. This is exactly the parameter we try to infer.

The method presented here works for a very large class of geometries. However, there are some cases in which it does not work and we need one mild additional assumption on the form of the matrix C. For the method presented here to work we require that the connections between the patches to be such that, if the densities or numbers in all patches are equal, redistribution through dispersal will not change this. In many cases this will follow naturally from the geometry of the spatial systems. Below we will present a simple condition on the matrix C by which this can be tested.

A row of C defines where the dispersers from a particular location end up. Similarly, from the entries in the columns one can determine from what location a particular patch receives dispersers. If all densities or numbers are equal, and dispersal does not change the sum of dispersers received per patch, the columns of C should sum up to zero (independent of patch size). In mathematical terms this means that $(1, \ldots, 1)$ is a left eigenvector of C. If we also require that dispersal conserves numbers the eigenvalue associated with this eigenvector is 0. Thus the rows of C sum up to 0, i.e. $\sum_j c_{ij} = 0$ if the patches are all of the same size. When locations are not of the same size the rows of C need not sum up to zero and instead we require that $\sum_{j=1}^{n} w_{1j}c_{ij} = 0$, or that $\mathbf{w}_1 = (w_{11}, \ldots, w_{1n})^{\mathrm{T}}$ is a right eigenvector $\mathbf{w}_1$ of C with eigenvalue 0. The analysis of systems with different location sizes is otherwise completely analogous to the case in which locations are of equal size (Jansen and Lloyd, 2000). Using the additional properties of C mentioned above, it follows (see, for instance, Marcus and Minc, 1964; Jansen and Lloyd, 2000; Lloyd and Jansen, 2004) that all

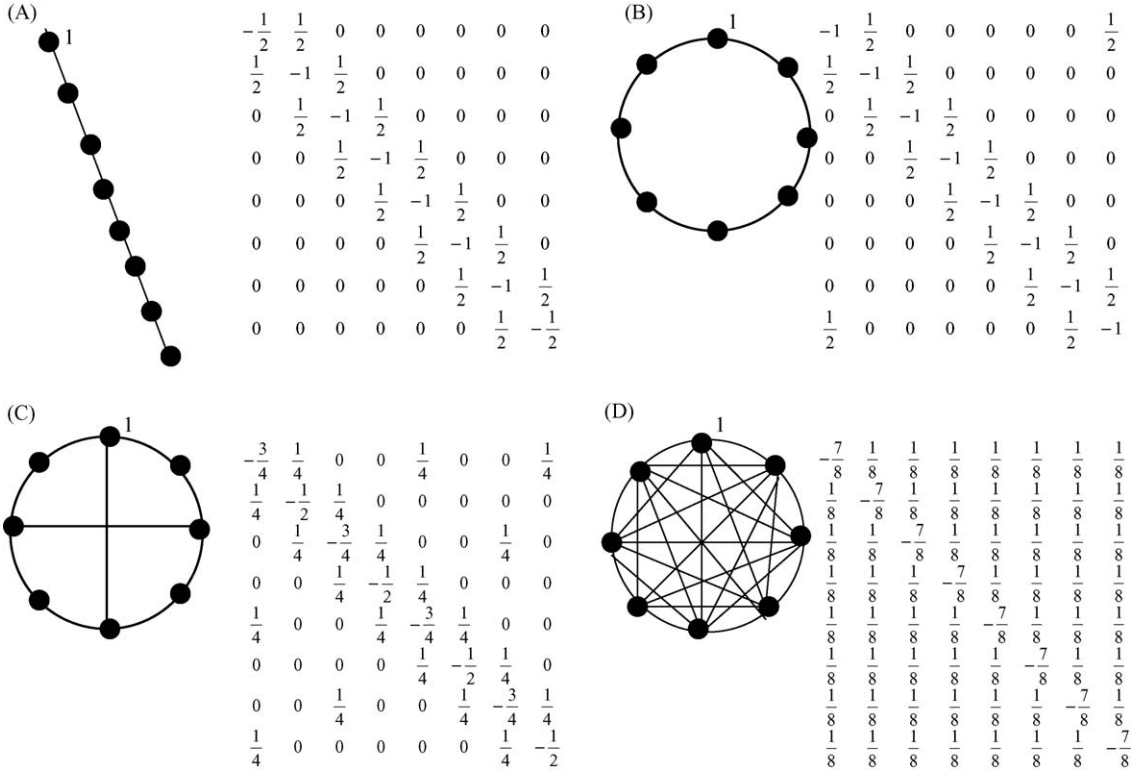


Fig. 1 – Different geometric organizations of metapopulations and their respective connectivity matrices.

non-zero eigenvalues of C have negative real part. For a detailed description of the eigenvalues of the connectivity matrix see Lloyd et al. (2006).

To illustrate the construction of the connectivity matrix C, Fig. 1 gives several possible geometric organizations and the respective connectivity matrices, under the assumption that patches are equally spaced and of identical size, that dispersers move only to neighbouring patches, and that they distribute equally and without any preferences between all the adjacent patches. This means in case A of a chain of patches, that from e.g. patch three on average 50% ($c_{34} = 1/2$) of the dispersers move to patch four and 50% ($c_{32} = 1/2$) to patch two and so forth. Since all the dispersers are lost to their patch, $c_{33} = -1$. Patches one and eight at the edges lose only 50% of their original dispersers ($c_{11} = c_{88} = -1/2$), since the remaining 50% are reflected on the boundaries. A more complicated case is depicted in C, where the inner four patches two, four, six and eight have connections to two other patches, so lose 2/4 of their dispersers. The outer patches one, three, five and seven lost 1/4 of their dispersers in either of the four directions, while the 25% which tried to migrate into the fourth direction are reflected back into their original patch. We assume that no losses are incurred during dispersal, as all dispersers lost for one patch reappear in other patches, thus the rows of C sum up to zero (for identical patch size). The simpler cases A and B can be interpreted as patches on a shoreline or edge of a forest, cases C and D as more complicated spatially structured environments. All these cases correspond to very simple dispersal kernels with a single dispersal distance. If it is possible for dispersers to reach further than the neighbouring patch within one time step, then the connectivity matrix C will contain more non-zero elements and the corresponding dispersal kernel will look accordingly. Note that the diagonal entries heights are equal to $1 + mc_{ii}$ and the matrix has non-zero entries mc_{ij} for $j = i \pm 1, i \pm 2$. In the following we assume that the connectivity matrix can be inferred directly from the geometrical organization and that dispersal occurs only to the neighbouring patches. If that is not the case, it is possible to estimate the connectivity matrix by constructing various matrices C and assess through maximum-likelihood techniques which matrix, respectively, is most suitable to describe the data. This, however, is beyond the scope of the current paper. We chose the circle as an example of low to intermediate complexity for further study, so if not stated otherwise, all following results were obtained using the circle.

3. Analysis

Although the model with dispersal (7) looks deceptively simple, it is still of considerable complexity as it describes n different variables and the model, as it stands, is difficult to use for statistical analysis. In this section we will show how to calculate statistics of interest from this model which are easier to deal with.

To reduce the complexity of the spatial data we will transform the spatial system, in a way that is analogous to the Fourier transformation used for the analysis of diffusive processes. We do this by applying a similarity transformation (Jansen and Lloyd, 2000). This transformation decouples the n-dimensional coupled autoregressive processes into a system of n decoupled autoregressive processes. For this transformation to be applied we require C to have n linearly independent eigenvectors. For most spatial systems this will be the case, and this condition is mostly unrestrictive. If this condition is fulfilled, C can be diagonalised by a similarity transform, i.e., there exists an invertible matrix B such that $B^{-1} C B = \Lambda$, where Λ is a diagonal matrix. This is a standard technique from linear algebra (see, for example, Hirsch and Smale, 1974). The diagonal elements, λ_i of the matrix Λ are the eigenvalues of C, and the matrix B can be constructed using the eigenvectors of C. In particular, we set $B = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$, where $\mathbf{w}_i$ is a right eigenvector of C, i.e. $C\mathbf{w}_i = \lambda_i \mathbf{w}_i$, and $B^{-1} = (\mathbf{v}_1, \ldots \mathbf{v}_n)^T$, where $\mathbf{v}_i$ is a left eigenvector of C, i.e., $\mathbf{v}_i C = \lambda_i \mathbf{v}_i$. We shall choose $\mathbf{v}_1 = (1, \ldots, 1)$ as the left eigenvector associated with $\lambda_1 = 0$.

In a Fourier transform the spatial system is described as a summation of sines and cosines which describes modes which are conserved under diffusion. Similarly, the eigenvectors of C describe modes which are conserved under dispersal. We will use the matrix B to transform the vector $\mathbf{x}_t$ into the vector $\psi_t = \mathbf{x}_t B$. The back transform is given by $\mathbf{x}_t = \psi_t B^{-1}$. The transformed vector at the next time point is given by

$$\psi_{t+1} = \mathbf{x}_{t+1}B = (a\mathbf{x}_t + \xi_t)(I + mC)B = a\psi_t + ma\psi_t B^{-1}CB + \xi_t(I + mC)B = a\psi_t(I + m\Lambda) + \omega_t$$

where $\omega_t = \xi_t(I + mC)B$. Note that the elements of ω_t, given by $\omega_{i,t}$, have zero mean. Because the matrix Λ is diagonal, this transformation decouples the coupled autoregressive processes into n decoupled AR1's. The individual autoregressive processes are given by

$$\psi_{j,t+1} = l_j \psi_{j,t} + \omega_{j,t}$$

where $l_j = a(1 + m\lambda_j)$ and λ_j is the jth diagonal element of Λ.

We can now bring to bear the tools for the analysis of autoregressive processes. To estimate the parameters l_j from the temporal covariance we use the Yule–Walker formalism (Karlin and Taylor, 1975; Box and Jenkins, 1970). Multiplication with ψ_j, summing up and division by T yields

$$\frac{1}{T}\sum_{1}^{T-1} \psi_{j,t+1}\psi_{j,t} = \frac{l_j}{T}\sum_{1}^{T-1} \psi_{j,t}\psi_{j,t} + \frac{1}{T}\sum_{1}^{T-1} \omega_{j,t}\psi_{j,t}$$

Since the noise is uncorrelated with x, it follows that also $\psi_{i,t}$ and $\omega_{i,t}$ are uncorrelated (Karlin and Taylor, 1975), thus, the last term in the equation above vanishes for $T \to \infty$. In the limit of $T \to \infty$ we find

$$E(\psi_{j,t+1}\psi_{j,t}) = l_j E(\psi_{j,t}^2)$$

This allows us to estimate the parameter $l_j = \hat{a}(1 + \hat{m}\lambda_j)$

$$\hat{l}_j = \frac{E(\psi_{j,t+1}\psi_{j,t})}{E(\psi_{j,t}^2)} \tag{8}$$

The expression above can be interpreted as covar$(\psi_{j,t+1}\psi_{j,t})$/var(ψ), i.e. the temporal correlation coefficient.

Note that the residual noise is given by

$$\hat{\omega}_{j,t} = \psi_{j,t+1} - \hat{l}_j \psi_{j,t}$$

The variance of the residual noise is minimised by this estimate of the multiplication factor $\hat{l}_j$

Estimates for the parameters $\hat{m}$ and $\hat{a}$ can be found by a linear regression of $\hat{l}_j$ on λ_j. The parameter $\hat{a}$ is the intercept of the regression line, while $\hat{a}\hat{m}$ corresponds to the slope. Thus $\hat{m}$ is not determined separately but through division of the slope by $\hat{a}$. This means that the inaccuracy in the determination of $\hat{a}$ is transferred to $\hat{m}$. This is especially important if $\hat{a}$ is small.

For a useful estimate we need an indication of the quality of the estimate. This can be done by assigning confidence intervals to the estimate or significance levels. For instance, in order to establish that space plays an important role in a system, one needs to show that $\hat{m}$ differs significantly from 0. Because our estimation technique is non-linear we were not able to derive analytical results and therefore constructed confidence intervals by numerical means. We did this in two different ways: firstly, by designing a bootstrap method and, secondly, by reconstructing the distribution of the estimated dispersal rates according to the autoregressive model (7) using a Monte Carlo method.

3.1. *Bootstrap method*

Confidence intervals and significance levels can be calculated from sample data using a computational method known as the bootstrap (Efron and Tibshirani, 1993). The main idea behind the bootstrap is that without any further information on the distribution of the noise ω_t, we will take the residual noise $\hat{\omega}_t$ as a sample of the true noise. We can reconstruct a large number of replicate time series, $\psi_t^* = (\psi_{1,t}^*, \ldots, \psi_{n,t}^*)$, called bootstrap time series, according to the scheme:

$$\psi_{t+1}^* = \hat{L}\psi_t^* + \omega_t^*$$

where $\hat{L}$ is a diagonal matrix with $\hat{l}_i$ as elements. The bootstrap noise ω_t^* is at each time drawn as a random sample from the set $\hat{\Omega} = \{\hat{\omega}_2, \ldots, \hat{\omega}_T\}$ and $\psi_1^* = \psi_1$. It is important to realise that we draw the vector ω_t^* rather than its elements $\omega_{i,t}^*$ from the estimated distribution in order to preserve the spatial correlation of the noise. From the bootstrap time series we can now estimate the n parameters $\hat{l}_i^*$. From these estimates we can find $\hat{m}^*$ and $\hat{a}^*$ through regression of $\hat{l}_i^*$ on λ_i. By running the bootstrap process a large number of times we can generate a population of $\hat{m}^*$ and $\hat{a}^*$. The bootstrapped estimates can be used to find the bootstrap variation of $\hat{m}$ given by $E((\hat{m}^* - \hat{m})^2)$ and to find a confidence interval. The boundaries of a single sided g% upper (lower) confidence intervals is the value for which only g% of the realised $\hat{m}^*$ is larger (smaller).

The bootstrap method can also be used to assign significance levels. For instance, if one would like to know whether or not the dispersal rate differs significantly from 0 one only needs to establish the proportion of $\hat{m}^*$ that is negative. To illustrate this, Fig. 2 shows how the 90, 95 and 99% confidence intervals decrease with increasing number of points in time. For an intermediate dispersal rate of 10% the width of the confidence interval is equal to 200 at around 300 time points. Since we calculated the confidence interval in percent of the true m, a deviation of 100% means that dispersal is not significantly different from zero.

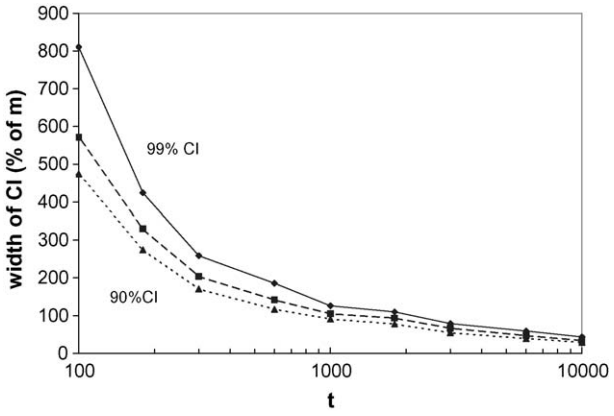


Fig. 2 – The width of the 90, 95 and 99% confidence interval decreases with increasing number of points in time using a bootstrap analysis for a circle of eight patches, $m = 0.1$, $a = 0.5$.

3.2. *Monte Carlo method*

To assign confidence intervals and to establish whether the estimation method described in the previous section is biased it is useful to know the distribution of the estimator. We did not obtain analytical results on this distribution and therefore generated the distribution by numerical means. Although computationally intensive, the rationale behind this is straightforward. Since the vector autoregressive model (7) includes the stochastic element ξ (the noise), the time series generated by the autoregressive model will generally differ. As a result, the corresponding estimates $\hat{a}$ and $\hat{m}$ will differ between realised time series even if the underlying parameters, and therefore the true values of the estimated parameters, are the same. One can use this to generate distributions of $\hat{a}$ and $\hat{m}$. These distributions can be used to construct confidence intervals and assign significance levels. This method differs from the bootstrap method in that the bootstrap method resamples the residuals to generate noise, whereas the Monte Carlo method requires an a priori assumption about the form of the noise.

4. Results

Dispersal can create a pattern in the population that is not obvious to the casual observer but that can be revealed using the method presented in this paper. Fig. 3a shows a time series generated using Eq. (7) and a connectivity matrix corresponding to eight patches connected in circular fashion (see Fig. 1b). The time series does not show an obvious pattern and gives no hint of the process of redistribution that is hidden within it. Fig. 3b shows the regression of $\hat{l}_j$ on λ_j according to Eq. (8). Whereas the time series does not reveal the presence of dis-

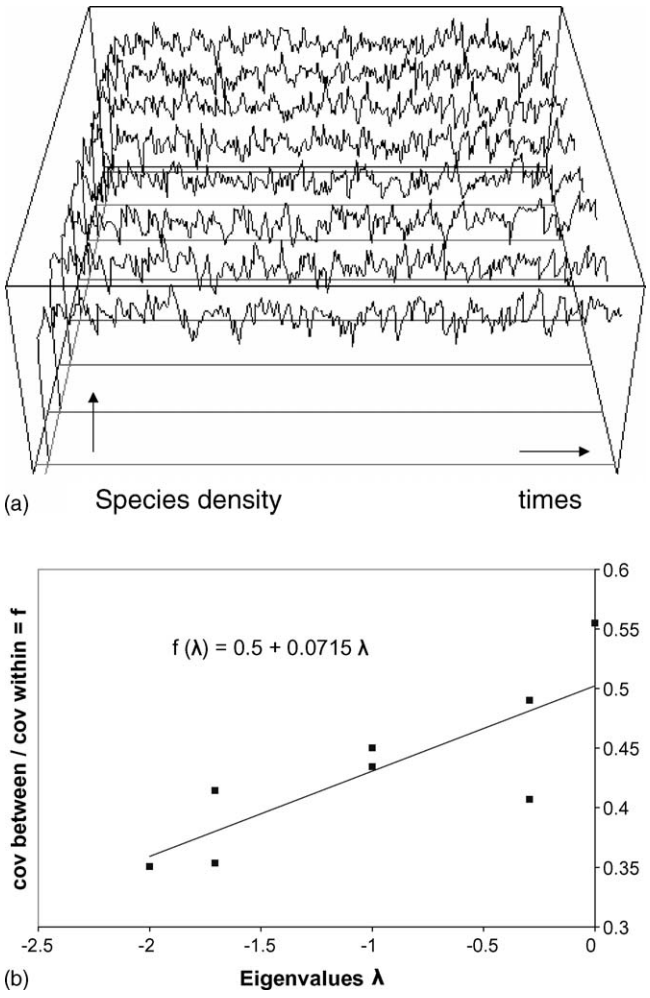


Fig. 3 – (a) A time series generated according to Eq. (7) for a circle of eight patches with local dispersal, $m = 0.1$, $a = 0.5$, $x_{0,i} = -10$, $T = 300$. (b) Analysis of the same data after transformation as explained in the text. The figure shows the regression of $\hat{l}$ on λ according to Eq. (8). In the time series data (a) the effect of dispersal is near impossible to see, whereas the effect is obvious in the regression (b).

persal, the regression clearly shows that dispersal has left its mark on the population dynamics. Although the plot of $\hat{l}_j$ versus λ_j shows considerable scatter, the linear component gives a good estimate of the dispersal rate. This illustrates how to infer the dispersal rates from correlation of noise in census data.

In order to get an impression of the quality of this estimation procedure we generated realisations of the autoregressive process (6) and analyzed how well the parameters a and m can be recovered with our estimation technique. We used the Monte Carlo method, in which we repeated this many times for the same parameters, to reconstruct the distribution of the estimators. In Fig. 4a the distribution for $\hat{m}$ and $\hat{a}$ are shown. Note that the maximum coincides well with the true value, but that in the case of very short time series the distributions are slightly skewed to the left. We calculated confidence intervals as the intervals which contain q% of the probability mass. Next we will investigate how these intervals depend on the system's parameters.

Fig. 4b shows how the confidence intervals of the estimated dispersal rate $\hat{m}$ depends on the true value of m. The width of the confidence intervals remains roughly constant and is independent of the amount of dispersal. This means that dispersal rates can be estimated with approximately equal precision-independent of the amount of dispersal, thus small values of the dispersal rate can be estimated with the same absolute precision as large dispersal rates. However, for small values of m it is not easy to establish that dispersal rates differ significantly from zero as the relative accuracy is larger for small values of m. For $a = 0.5$ and 1000 points in time, the significance level of m lies in the order of 5% (95% confidence interval); with a per capita growth rate of $a = 0.9$ dispersal rates as low as 1.8% (95% confidence interval) can be significantly distinguished from 0.

The parameter a has also a considerable impact on the accuracy of both its own estimation and especially on the estimation of m. Especially if a is close to zero the confidence intervals have considerable width. The reason for this is that m is calculated from the slope of the regression line, which is am, divided by a. Small values of a present a problem as they will make the accurate determination of m difficult. If a is small this will lead to considerable error. The closer the absolute of a is to one, the smaller the part of the deterministic decay $|a\mathbf{x_t} - \mathbf{x_{t+1}}|$ as compared to the stochastic part ξ_t and the better the quality of the estimate. A strong stochastic component is advantageous to the accuracy of the determination of both a and m since the method makes use of the propagation of noise between patches and depends on how quickly the local relaxes back to equilibrium. The sign of a appears to have little influence, only its absolute value is important.

Fig. 5a shows the dependence on the number of time points of the parameter estimation. Shown is the width of the confidence intervals of m against the reciprocal of the square root of the number of time points ($1/\sqrt{T}$). It can be seen that this yields an approximately linear relation. A four-fold increase in the number of observations will therefore lead to a halving of the confidence intervals.

The accuracy of the parameter estimation also depends on the number of patches involved. The higher the number of patches, the smaller the confidence intervals, as seen in Fig. 5b. The confidence intervals appear to depend on the number of patches through a power law with exponent −0.4. However, this is much less clear than in the case of the number of time points. In this example, a doubling of sampling effort would result in a greater effect if sampling is carried out for twice the number of points in time than if twice the number of patches are included. The difference, however, is marginal and we expect this to depend on the details of the study.

Different geometrical organizations lead to different matrices C and their corresponding eigenvalues which then directly influence the regression of the variances and covariances of the transformed densities of the time series against the eigenvalues (see Eq. (8)). Table 1 shows how the variances s of the distributions of the estimates of a and m increase with decreasing variance of the eigenvalues themselves. So the confidence intervals of dispersal rates in a metapopulation, i.e. a geometry in which all patches are connected to all other

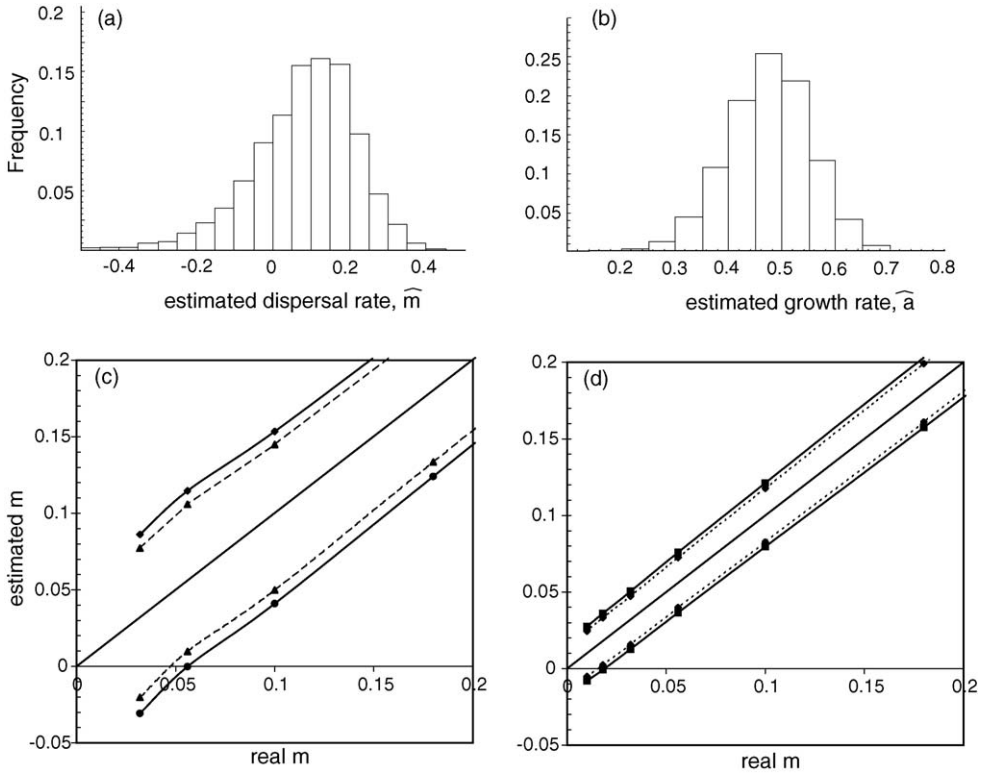


Fig. 4 – The distributions of (a) the estimated dispersal rates, $\hat{m}$, and (b) the estimated growth rates, $\hat{a}$, for a circle of eight patches, $t = 60$, $a = 0.5$, $m = 0.1$. (c) The width of the estimated confidence intervals against the real value of m for $a = 0.5$. Dashed lines: 90% confidence intervals, solid lines: 95% confidence intervals, $T = 1000$. (c) as (d) with $a = 0.9$.

patches with equal strength ($\text{var}(\lambda_j) = 7/64$) are roughly twice as wide as of the circle ($\text{var}(\lambda_j) = 32/64$). This suggests that the smaller the variance in the eigenvalues of C, the larger the confidence intervals will be, presumably the effect is through the statistics of the linear regression used. We chose the circle as an example of low to intermediate complexity for further study, so unless stated otherwise, all following results were obtained using the circle.

The method applied in this paper is based on the system having reached a stationary state (to be precise, for the distribution of the local densities to have converged to a stationary distribution). We assessed how transients interfere with the parameter estimation and found that a pronounced transient in time will lead to an overestimation of both m and a, the level of overestimation increases with the proportion of the transient, relative to the total number of time points. Note that because the underlying model is linear, we did not study in detail the effect of long transients, or transient chaos.

4.1. *Robustness*

Next we will demonstrate how this estimation method performs if the model assumptions are not justified. Our model (6) has a linear population growth rate, a. This does not imply that the method works only in that special case. To demonstrate that also non-linear models can easily be analyzed using the method described above, we chose the familiar logistic equation $F(N_t) = rN_t(1 - N_t/K)$. The results, i.e. the estimated parameters, are virtually identical with those of the linear model in the case of the carrying capacity K being large as compared to the noise. Even if the variation in the noise is in the order of the carrying capacity, whilst preventing the densities to become negative, this led only to a slight overestimation of the parameters. The main reason underlying this robustness is that the densities are most of the time in the vicinity of the equilibrium, so that the linearisation provides a good approximation. Even though large excursions occur from time to

Table 1 – Dependence of the variance of the distribution of $\hat{m}$ on the variance of the eigenvalues of C for systems of eight patches

							
$\sigma^2(\lambda)$	32/64	32/64	13.75/64	12/64	12/64	11/64	7/64
$\text{var}(m) \times 10^4$	7.2	8.5	14.65	20.8	23.2	27.9	40.1

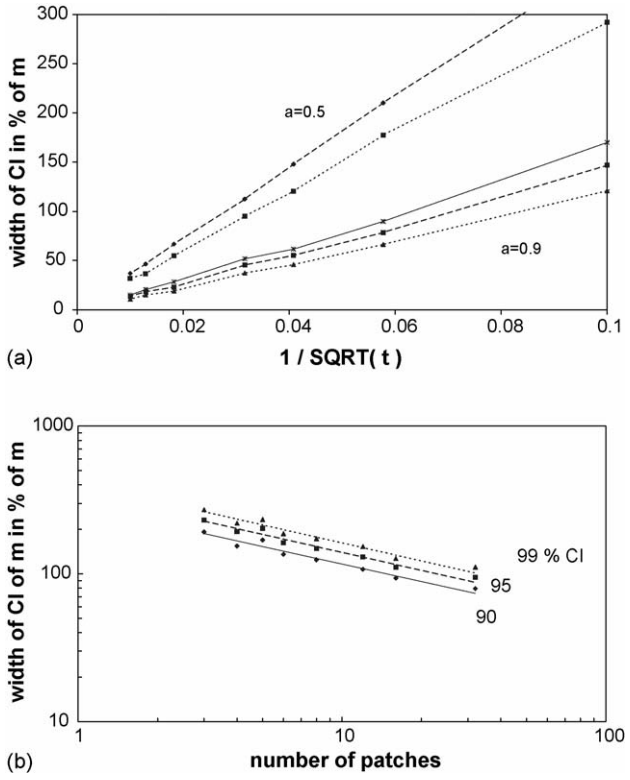


Fig. 5 – (a) The width of the confidence interval against the reciprocal of the square root of the number of time points ($1/\sqrt{T}$). Shown are the 90% (dotted lines), 95% (dashed lines) and 99% (solid lines) confidence intervals. Results for a circle of eight patches, $m = 0.1$ (b) The dependence of the width of the confidence intervals on the number of patches, $a = 0.9$, $T = 100$.

time, these will have a relatively short duration and tend not to distort the analysis too much. This shows that this method is robust against a change in the population growth model.

In natural populations, growth and dispersal rates might not be constant but vary over time. Thus we investigated whether such variations would influence the estimation procedure. As a simple example we chose a linear increase in the value of m or a. An increase (50–150%) of the respective value in a led only to a negligible overestimation of both (averaged) $\hat{a}$ and $\hat{m}$ values, an increase of m left the parameter estimation virtually unaffected.

As described above, we assumed that the connectivity between patches and thus the matrix C is known. In reality, one might overlook existent connections or assume dispersal between patches where for some reasons it cannot occur. Thus we explored how analyzing simulated data with a wrong matrix C will influence the outcome of the parameter estimation. As an example we used a circular geometry to generate data and then analyzed them using the matrix of a circle where in addition patch 1 is connected with patch 5 and patch 3 with patch 7. This led to an overestimation of m by 50–65% ($a = 0.5$; $m = 0.1$; $t = 100, 300, 1000$) and an overestimation of the width of the confidence interval of m by more than 50%. The estimation of a remained entirely unchanged. The opposite case, where we used this circle with two extra connections to generate the data and the matrix of the circle to analyze them, lead to opposite results: The value of m was underestimated by 50% and the width of the confidence interval was underestimated by more than 50%. Again, the estimation of a remained untouched. Despite the shifted estimates for m, the confidence intervals still contained the true value of m.

Since patches of suitable habitat might well be subjected to similar environmental fluctuations, i.e. the noise that patches experience might be correlated between them. We explored how the presence of correlated noise influences the estimation of the parameters. As an example we used blocks of two patches with the same noise, so in the circle made of eight patches, at any one time there would be only four different noise terms. Fig. 6a shows that this spatial correlation between the patches leads only to a small underestimation of the width of the confidence intervals of m, whereas the estimation of a is unaltered. We repeated this for several points in time. We further investigated the influence of temporal correlation in the noise. Fig. 6b shows how temporal correlation of the noise, i.e. the same noise term for several points in time, leads to an underestimation of m as well as a pronounced underestimation of the width of its confidence interval which increases with increasing correlation. In extreme cases, the confidence interval does not include the true value any more. The opposite applies to the estimation of a, which is markedly overestimated.

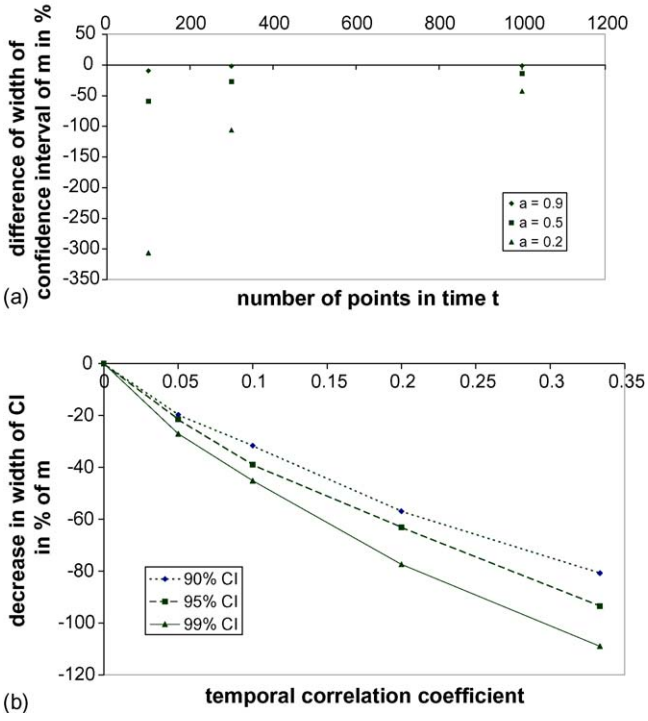


Fig. 6 – (a) The increase in width of confidence intervals with spatially correlated (spatial correlation coefficient 0.25) noise compared to the confidence intervals when the noise was spatially uncorrelated, for a circle of eight patches at different number of points in time, $m = 0.1$. (b) The decrease in width of the confidence intervals of m compared to the simulations in which the noise was uncorrelated in time. Circle of patches, $a = 0.5$, $T = 300$.

5. Discussion

Dispersal has important implications on an individual, population and meta-population level. Dispersal rates however are difficult to measure. We presented a method to estimate dispersal and per capita growth rates from time series data using a first-order autoregressive model to describe and generate these data. The model explicitly takes the geometrical organization of the patches into account and is as such an improvement compared to frequently used models such as e.g. the island model or stepping stone model.

We show that the parameter estimation is robust against a variety of possible violations of assumptions of the model. Non-linear growth rates had little influence on the accuracy of the estimation. Also transients of the parameters had little effect on the estimation of the parameters; spatially correlated noise has only a limited effect. Using incorrect connectivities, i.e. the wrong matrix C, led to mis-estimation of m, but the respective confidence intervals still included the true values. Only pronounced temporally correlated noise as well as a marked transient led to mis-estimates of the involved parameters and present a problem. This can be dealt with by cutting off any transients, and in the case of temporally highly correlated noise another model than the AR1 model may be needed. This is a standard method in time series analysis (Box and Jenkins, 1970).

We have assumed that there are no losses incurred during dispersal, hence, that there is no dispersal mortality. This is a simplification, since the crossing of unsuitable areas in search of a new habitable patch is often dangerous and therefore could incur considerable losses. It is easy to include a certain fixed mortality (equal to all dispersers) into the dispersal process by simply discounting the growth rate by that mortality. We have also assumed that censuses as well as dispersal would take place once per generation and that generations do not overlap. This is a reasonable assumption for a number of species, particularly insect species. In the case of overlapping generations one would need an continuous model, which can be constructed along the lines outlined in this paper. Also if the population has a pronounced structure the model will need to be adapted accordingly. For instance, if only certain life stages are able to disperse, as is the case for aphids where only the adults are winged and can fly, this is particularly important. The method described, however, is also applicable to the case of more than one species. Multiple species scenarios could include complete amensalism, (cyclic) predator–prey dynamics and competition for a resource. One could also include a density-dependent dispersal rate which would account for dispersal as a reaction of e.g. overcrowding. As shown in the results part, a non-constant dispersal rate is not a problem per se. Analysis of density-dependent dispersal can be performed in a similar way to density-independent dispersal (Huang and Diekmann, 2003). The method presented here can with little effort be generalised to include all these scenarios.

Our method makes no a priori assumptions about the way in which patches are connected, but provides a method that can be applied to any population geometry. In that respect it is a substantial advance over the commonly used island models. Interestingly, we found that island models generally perform badly in that they tend to result in large confidence intervals. Stepping stone models perform better in this respect, but will introduce error if the geometry is different. The fact that our model can describe any geometry, comes with a price: the connectivity of patches needs to be known. Ideally, this information should be gained from independent sources, such as, from ringing data or other forms of marking and recapturing at other locations. If such information is not available, one can make informed choices about possible topologies using travel distance between patches. An interesting extension of the work described here would be to develop methods to choose the most likely of a number of plausible connectivity matrices, using the data. This is possible, if one would consider the matrix that would lead to the smallest residuals, and hence that explains most of the variance, as the most likely description of the connectivity matrix. In a similar fashion one could investigate the details of a dispersal kernel. This could be done by constructing various matrices C with dispersal to the next, second, third-distant or other patches, and assess through maximum-likelihood techniques which dispersal kernel and connectivity matrix, respectively, are the most suitable to describe the data.

The model described in this paper relates directly to the issues of synchrony between populations (Lloyd and Jansen, 2004; Matter, 2001; Li et al., 2005). The techniques discussed here could form the basis for a method to quantify synchrony in ensembles of linked populations. Although such methods have been developed they are often based on simplifying assumptions such as a linear or exponential decay of the synchrony with distance (Myers et al., 1995; Liebhold et al., 2004). The ideas put forward in this paper could be used to generically quantify the effect of synchrony in a network of coupled patches. Similarly, one could envision applying such ideas to the quantification of the Moran effect; synchronisation due to the correlation of environmental noise between patches (Moran, 1953; Royama, 1992; Ripa, 2000). There is, however, a practical problem: the Moran effect requires the population dynamics to be correlated beyond the correlation that one would expect from the dispersal process. If one has access to only one data set one cannot estimate both the dispersal rates and the environmental correlation from the same data. If the amount of dispersal is known, for instance if there is no dispersal between subpopulations, one can, in principle, disentangle the relative contributions of dispersal and environmental correlation.

The main disadvantage of the method discussed here is the considerable amount of points in time needed to get accurate estimates of the dispersal rate m. In order to get small confidence intervals one would need a large number of observations, and this might be problematic in ecological research. Having said this, we have applied this method in an experimental setting with as little as 20 data points and obtained useful results. Also note that this disadvantage is likely to hold for most methods which are based on the redistribution of a population of markers and that rely on repeated observations. This method is therefore, in terms of confidence intervals, of similar quality as other methods but has the advantage over other methods that it allows the use of information about the topology. One way in which a large number of observations

can be achieved is by automated observation. The method described here allows for the analysis of such data.

An interesting finding with respect to the limited length of time series in ecology is that in estimating the dispersal rate it is possible to replace space for time, i.e., by halving the time points sampled and doubling the number of locations sampled, by studying a larger area, one can determine the dispersal rate with slightly greater precision. This does assume that the patches are connected in such a way that changing the size of the overall area would not cause major changes in the local dispersal structure and that edge effects play a minor role. This will be the case if dispersal is mainly to nearby localities. This observation is useful for the design of studies which are limited in time. It also is somewhat reminiscent of ideas on time series embedding which suggest that observations on a single location can be used to infer information on the population dynamics on the full spatial system (Takens, 1980; Rand and Wilson, 1995; Keeling et al., 2000).

There is currently much interest in invasions and species shifting their ranges in response to climate change. Many of the details of these processes depend on the quantification of dispersal and the approach outlined here offers a way to estimate this parameters. An expanding population would increase its range, and will normally do so relatively slowly compared to its characteristic return time. This implies that in most locations the population dynamics will still be around the equilibrium density. Our method therefore allows the estimation of the dispersal rate in an expanding population in the part of the population where it is reasonably close to equilibrium.

Acknowledgement

AS thanks the DAAD (German Academic Exchange Service) for support.

Appendix A

In case the population within one patch is structured, or contains more than one species, a similar formalism an be applied (Jansen and Lloyd, 2000). Let us assume there are m different classes or species to be kept track off. The local densities in patch j at time t in such a case can be described by the $m \times 1$ column vector $\mathbf{x}_{j,t} = (x_{1j,t}, x_{2j,t}, x_{3j,t}, \ldots)^{\mathrm{T}}$. The stochastic model takes the form

$$\mathbf{x}_{j,t+1} = A\mathbf{x}_{j,t} + \xi_{j,t} + M\sum_{i=1}^{n} c_{ij}\,(\mathbf{x}_{i,t} + \xi_{i,t})$$

The $m \times m$ matrix A is the Jacobian at equilibrium, and the $m \times m$ matrix M contains the dispersal rates. Note that the off-diagonal terms m_{ij} represent the impact of the density of class j on the dispersal of class i. The column vector $\xi_{j,t}$ describes the noise in patch j at time t.

This same formalism can be applied to describe density-dependent dispersal. If this is the case the total amount of dispersal not only depends the product of the density and the dispersal rate, but may depend on higher order terms. To allow for this we will describe the local population in patch j at time t by a $m \times 1$ column vector $\mathbf{x}_{j,t} = (x_{j,t}, x_{j,t}^2, x_{j,t}^3, \ldots)^{\mathrm{T}}$. With this vector we can now describe the dispersal as $x_{j,t}m_{11} + x_{j,t}^2 m_{12} + x_{j,t}^3 m_{13} + x_{j,t}^4 m_{14} + \cdots$ and, hence, the density dependence dispersal is captured by the coefficients in the first row of M. Note that this is a generic but not necessarily the most efficient way to describe density-dependent dispersal. If the functional form of the dependence is known other description could follow naturally.

We will now proceed by using a tensor notation to capture the descriptions of the local densities. In this tensor we glue together all the vectors that describe the local densities so the $m \times n$ tensor

$$X_t = (\mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \ldots, \mathbf{x}_{n,t}) = \begin{bmatrix} x_{1,1,t} & \cdots & x_{1,n,t} \\ \vdots & \ddots & \vdots \\ x_{m,1,t} & \cdots & x_{m,n,t} \end{bmatrix}$$

has the density of class i in patch j at position i, j. In a similar fashion are the noise terms now arranged in tensor

$$\Xi_t = (\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{n,t})$$

Using tensor notation we can efficiently rewrite the dynamics as

$$X_{t+1} = AX_t + \Xi_t + M(AX_t + \Xi_t)C$$

where Ξ_t is a row vector that has $\xi_{j,t}$ as elements, i.e., $\Xi_t = (\xi_{1,t}, \ldots, \xi_{n,t})$ (see Jansen and De Roos, 2000; Jansen and Lloyd, 2000; Hunter and Caswell, 2006). This notation is an economic way to describe all n autoregressive processes at once. Note that all operation concerning the interactions within a patch are given by pre-multiplication of a tensor with a matrix, whereas operations concerning interactions between patches are given by post-multiplication of a tensor with a matrix.

The analysis of such system is completely analogue to the analysis given in this paper. To do so we will use the same matrix B as described in the text to transform the tensor X_t into the tensor $\Psi_t = X_t B$. The back transform is given by $X_t = \Psi_t B^{-1}$. The transformed vector at the next time point is given by

$$\Psi_{t+1} = X_{t+1}B = (AX_t + \Xi_t + M(AX_t + \Xi_t)C)B = A\Psi_t + MA\Psi_t\Lambda + \Omega_t$$

where $\Omega_t = (\Xi_t + M\Xi_t C)B$. Note that the elements of Ω_t again have zero mean. This operation decouples the coupled autoregressive processes into n decoupled AR1's as it did before. The individual autoregressive processes are given by

$$\psi_{j,t+1} = L_j\psi_{j,t} + \omega_{j,t}$$

where $L_j = (I + \lambda_j M)A$ and λ_j is the jth diagonal element of Λ.

To estimate the matrices L_j we apply the Kronecker product with ψ_j on both sides. The Kronecker product is an operator which takes two arbitrarily dimensioned matrices A with dimension $v \times w$ and B with dimension $p \times q$ and results into a

$vp \times wq$ matrix with the structure

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \dots & a_{1,w}B \\ \vdots & \ddots & \vdots \\ a_{v,1}B & \dots & a_{v,w}B \end{bmatrix}$$

This yields

$$\psi_{j,t+1} \otimes \psi_{j,t} = L_j \psi_{j,t} \otimes \psi_{j,t} + \omega_{j,t} \otimes \psi_{j,t}$$

After summing up over t and dividing by T we find

$$\frac{1}{T}\psi_{j,t+1} \otimes \psi_{j,t} = \frac{1}{T}L_j \sum_{1}^{T-1} \psi_{j,t} \otimes \psi_{j,t} + \frac{1}{T}\sum_{1}^{T-1} \omega_{j,t} \otimes \psi_{j,t}$$

Since the noise is uncorrelated with x, it follows that also $\psi_{i,t}$ and $\omega_{i,t}$ are uncorrelated (Karlin and Taylor, 1975), thus, the last term in the equation above vanishes for $T \to \infty$. In the limit we find

$$E(\psi_{j,t+1} \otimes \psi_{j,t}) = L_j E(\psi_{j,t} \otimes \psi_{j,t})$$

This allows us to estimate the matrix $\hat{L}_j = (I + \lambda_j \hat{M})\hat{A}$ as

$$\hat{L}_j = E(\psi_{j,t+1} \otimes \psi_{j,t})E(\psi_{j,t} \otimes \psi_{j,t})^{-1}$$

The matrices in the expression above are covariance matrices.

This provides a recipe for calculation dispersal rates in case the rates depend on the presence of other species, if the local population is structured, and if dispersal is density-dependent. In all these cases, but for density dependence, extra information is needed in terms of the number in each class at each time point. For density-dependent dispersal this information can be derived from a single measurement. Although this formalism possibly has some pitfalls in that there is dependence of the elements in the A and M matrices, which could complicate matters, it is not necessarily the case that such procedures require much more data points than the case presented in the main text. Especially if the extended description offers a much better description of the process, there could be an advantage of this description over the one-dimensional one. In principle, this can be tested by constructing a series of nested models (for instance, in case of density dependence by including second powers, third powers, etc.) in the model description and test if such description is sufficiently better to justify the inclusion of an extra parameter.

REFERENCES

Barton, N.H., 2001. The evolutionary consequences of gene flow and local adaptation: future approaches. In: Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D. (Eds.), Dispersal. Oxford University Press, Oxford, pp. 329–340.

Bascompte, J., Sole, R.V., 1998. Modeling Spatiotemporal Dynamics in Ecology. Springer Verlag, New York.

Bennets, R.E., Nichols, J.D., Lebreton, J.-D., Pradel, R., Hines, J.E., Kitchens, W.M., 2001. Methods for estimating dispersal probabilities and related parameters using marked animals. In: Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D. (Eds.), Dispersal, pp. 2–17.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. McGraw-Hill, San Francisco.

Cain, M.L., Milligan, B.G., Strand, A.E., 2000. Long-distance seed dispersal in plant populations. Am. J. Bot. 87, 1217–1227.

Cain, M.L., Nathan, R., Levin, S.A., 2003. Long-distance dispersal. Ecology 84, 1943–2020.

Chesson, P., 2001. Metapopulations. In: Levin, S.A. (Ed.), Encyclopedia of Biodiversity, vol. 4. Academic Press, pp. 161–176.

Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D., 2001. Dispersal. Oxford University Press, Oxford.

Czaran, T., 1998. Spatiotemporal Models of Population and Community Dynamics. Chapman & Hall, London.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Etienne, R.S., 2004. On optimal choices in increase of patch area and reduction of interpatch distance for metapopulation persistence. Ecol. Model. 179, 77–90.

Hanski, I., 1999. Metapopulation Ecology. Oxford Series in Ecology and Evolution. Oxford University Press.

Hanski, I., 2001. Population dynamic consequences of dispersal in local populations and in metapopulations. In: Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D. (Eds.), Dispersal, pp. 283–298.

Hassell, M.P., Comins, H.N., May, R.M., 1991. Spatial structure and chaos in insect populations-dynamics. Nature 353, 255–258.

Hein, S., Pfenning, B., Hovestadt, T., Poethke, H.J., 2004. Patch density, movement pattern, and realised dispersal distances in a patch-matrix landscape—a simulation study. Ecol. Model. 174, 411–420.

Hirsch, M., Smale, S., 1974. Differential Equations, Dynamical Systems and Linear Algebra. Academic Press, New York.

Hovestadt, T., Poethke, H.J., 2006. The control of emigration and its consequences for the survival of populations. Ecol. Model. 190, 443–453.

Huang, Y., Diekmann, O., 2003. Interspecific influence on mobility and turing instability. Bull. Math. Biol. 65, 143–156.

Hunter, C.M., Caswell, H., 2006. The use of the vec-permutation matrix in spatial matrix population models. Ecol. Model. 188, 15–21.

Husband, B.C., Barrett, S.C.H., 1996. A metapopulation perspective in plant population biology. J. Ecol. 84, 461–469.

Jansen, V.A.A., Lloyd, A.L., 2000. Local stability analysis of spatially homogeneous solutions of multi-patch systems. J. Math. Biol. 41, 232–252.

Jansen, V.A.A., De Roos, A.M., 2000. The role of space in reducing predator–prey cycles. In: Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge.

Kaiser, A., 1995. Estimating turnover, movements and capture parameters of resting passerines in standardized capture-recapture studies. J. Appl. Statist. 22, 1039–1047.

Koenig, W.D., Van Vuren, D., Hooge, P.N., 1996. Detectability, philopatry, and the distribution of dispersal distances in vertebrates. Trends Ecol. Evol. 11, 514–517.

Levins, R., 1970. Extinction. Lect. Notes Math. 2, 75–107.

Li, Z.Z., Gao, M., Hui, C., Han, X.Z., Shi, H.H., 2005. Impact of predator pursuit and prey evasion on synchrony and spatial patterns in metapopulation. Ecol. Model. 185, 245–254.

Liebhold, A., Koenig, W.D., Bjørnstad, O.N., 2004. Spatial synchrony in population dynamics. Annu. Rev. Ecol. Syst. 35, 467–490.

Lloyd, A.L., Jansen, V.A.A., 2004. Spatiotemporal dynamics of epidemics: synchrony in metapopulation models. Math. Biosci 188, 1–16.

Lloyd, A.L, Funk, G.A, Jansen, V.A.A., 2006. Coupling matrices in metapopulation models. Theor. Pop. Biol., in preparation.
Matter, S.F., 2001. Synchrony, extinction, and dynamics of spatially segregated, heterogeneous populations. Ecol. Model. 141, 217–226.
Keeling, M.J., Wilson, H.B., Pacala, S.W., 2000. Reinterpreting space, time lags, and functional responses in ecological models. Science 290, 1758–1761.
Karlin, S., Taylor, H.M., 1975. A First Course in Stochastic Processes. Academic Press, San Diego.
Marcus, M., Minc, H., 1964. A Survey of Matrix Theory and Matrix Inequalities. Allyn and Bacon, Inc., Boston.
Metzger, R., 2005. Animal dispersal modelling: handling landscape features and related animal choices. Ecol. Model. 190, 159–170.
Moran, P.A.P., 1953. The statistical analysis of the Canadian lynx cycle II. Aust. J. Zool. 1, 291–298.
Myers, R.A., Mertz, G., Barrowman, N.J., 1995. Spatial scales of variability in cod recruitment in the North Atlantic. Can. J. Fish. Aquat. Sci. 52, 1849–1862.
Nathan, R., 2001. The challenges of studying dispersal. Trends Ecol. Evol. 16, 481–483.
Nathan, R., Perry, G., Cronin, J.T., Strand, A.E., Cain, M.L., 2003. Methods for estimating long-distance dispersal. Oikos 103, 261–273.
Pfenning, B., Hovestadt, T., Poethke, H.J., 2004. The effect of patch constellation on the exchange of individuals between habitat-islands. Ecol. Model. 180, 515–522.
Rand, D.A., Wilson, H.B., 1995. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. Proc. R. Soc. Lond. B 259, 111–117.
Reed, J.M., Levine, S.H., 2005. A model for behavioral regulation of metapopulation dynamics. Ecol. Model. 183, 411–423.
Ripa, J., 2000. Analysing the Moran effect and dispersal: their significance and interaction in synchronous population dynamics. Oikos 89, 175–187.
Rohani, P., May, R.M., Hassell, M.P., 1996. Metapopulations and equilibrium stability: the effects of spatial structure. J. Theor. Biol. 181, 97–109.
Rousset, F., 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142, 1357–1362.
Rousset, F., 2001. Genetic approaches to the estimation of dispersal rates. In: Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D. (Eds.), Dispersal 18–28.
Royama, T., 1992. Analytical population dynamics. Chapman & Hall, London.
Singh, B.K., Rao, J.S., Ramaswamy, R., Sinha, S., 2004. The role of heterogeneity on the spatiotemporal dynamics of host-parasite metapopulation. Ecol. Model. 180 (2/3), 435–443.
Slatkin, M., 1985. Gene flow in natural populations. Ann. Rev. Ecol. Syst. 16, 393–430.
Slatkin, M., Barton, N.H., 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43, 1349–1368.
Stensteth, N.C., Lidicker Jr., W.J., 1992. Where do we stand methodologically about experimental design and methods of analysis in the study of dispersal. In: Stensteth, N.C., Lidicker, W.J. (Eds.), Animal Dispersal: Small Mammals as a Model. Chapman & Hall 295–318.
Takens, F., 1980. Detecting strange attractors in turbulence. In: Rand, D.A., Young, L.-S. (Eds.), Dynamical Systems and Turbulence. Lecture Notes in Mathematics No. 898. Springer, Warwick.
Thrall, P.H., Burdon, J.J., Murray, B.R., 2000. The metapopulation paradigm: a fragmented view of conservation biology. In: Young, A.G., Clarke, G.M. (Eds.), Conservation Biology 4: Genetics, Demography and Vitality of Fragmented Populations. Cambridge University Press.
Whitlock, M.C., McCauley, D.E., 1999. Indirect measures of gene flow and migration: Fst $\pm 1/(4Nm + 1)$. Heredity 82, 117–125.
Wright, S., 1940. Breeding structure of populations in relation to speciation. Am. Nat. 74, 232–248.