

Modelling quasi-stationary behaviour in metapopulations

P.K. Pollett¹

Department of Mathematics, The University of Queensland, Queensland 4072, Australia

Abstract

We consider a Markovian model proposed by Gyllenberg and Silvestrov for studying the behaviour of a *metapopulation*: a population that occupies several geographically separated habitat patches. Although the individual patches may become empty through extinction of local populations, they can be recolonized through migration from other patches. There is considerable empirical evidence (in the work of Gilpin and Hanski, for example) which suggests that a balance between migration and extinction is reached which enables these populations to persist for long periods. The Markovian model predicts extinction in a finite time. Thus, there has been considerable interest in developing methods which account for the persistence of these populations and which provide an effective means of studying their long-term behaviour before extinction occurs. We shall compare and contrast the methods of Gyllenberg and Silvestrov (pseudo-stationary distributions) and those of Day and Possingham, which are based on the classical notion of a quasi-stationary distribution. We present here a convincing rationale for the latter, using limits of conditional probabilities.

1 INTRODUCTION

Suppose that we are using a stochastic process $(X(t), t \geq 0)$ to model a population which might eventually become extinct. The state $X(t)$ at time t might be something as simple as the number in the population, but it could be more complicated: $X(t)$ might be a vector indicating the numbers of various species, or the numbers occupying various geographical regions. If we are to use our model to explain observed phenomena, to make predictions, or if, in the first instance, we wish simply to refine it in order that it might faithfully capture

¹ Supported by the Australian Research Council (Grant No. A69130032)

the behaviour of the population, then *the very best we can hope to extract from our model* is the complete set of state probabilities: $p_x(t) = \Pr(X(t) = x)$, $x \in \mathcal{S}$, $t > 0$, where \mathcal{S} is the set of states. This can usually be done, at least in principle, by solving a set of difference equations or differential equations. More often than not, an exact solution cannot be obtained analytically, and so either analytical approximations, computational methods or asymptotic methods are used. But, for the moment, let us imagine that we have complete information: to be emphatic, we know $p_x(t)$ for every x and t .

Now suppose that we observe the population at an arbitrary time u and we see that extinction has not yet occurred. We know nothing more. How can we incorporate this new information? We should evaluate a *conditional state distribution*, that is, the state probabilities at time u conditioned on non-extinction:

$$m_x(u) = \Pr(X(u) = x | X(u) \neq 0) = \frac{p_x(u)}{1 - p_0(u)}, \quad x \in \mathcal{C}, \quad (1)$$

where 0 is the state corresponding to extinction, and \mathcal{C} comprises the remaining states ($\mathcal{S} = \{0\} \cup \mathcal{C}$).

Our purpose here is to use this conditional state distribution to better understand the metapopulation models introduced by Gyllenberg and Silvestrov [4]. If the numbers of states is not too large, as is the case in all the examples studied in [4], then $m_x(t)$ can be calculated numerically. We shall compare the conditional state distribution, $m(t) = (m_x(t), x \in \mathcal{S})$, with quasi-stationary distributions (later called pseudo-stationary distributions in [5]) introduced by those authors. We shall see that, as t becomes large, $m(t)$ approaches a limit m , called a *limiting conditional distribution* (traditionally called a quasi-stationary distribution), and that m is the eigenvector corresponding to the largest eigenvalue of the transition matrix restricted to \mathcal{C} . Thus, when the convergence of $m(t)$ to m is rapid, m provides a simple means of assessing the long-term behaviour of the metapopulation before extinction occurs. Day and Possingham [2] have used this approach in analyzing a similar, more refined model. They comment that the quasi-stationary distribution and the pseudo-stationary distribution appear not to coincide. We shall amplify this observation, by showing that the two ways of analyzing the model can give rise to opposing conclusions.

2 A STOCHASTIC MODEL FOR METAPOPOPULATIONS

Suppose that there are n distinct geographical regions, or *patches*, and let $X_i(t)$ be 1 or 0 according as patch i is occupied or not at time t , where $t = 0, 1, 2, \dots$. Let $X = (X_1, X_2, \dots, X_n)$ and suppose that $(X(t), t \geq 0)$ is

a (discrete-time) Markov chain taking values in $\mathcal{S} = \{0, 1\}^n$. Its transition structure is defined as follows.

Define an *interaction matrix* $Q = (q_{ij}, i, j \in \mathcal{N})$, where $\mathcal{N} = \{1, 2, \dots, n\}$, which governs the behaviour of the patches over a single time step: q_{ij} , for $j \neq i$, is the probability that patch j will *not* be colonized by migration from patch i , and q_{ii} is the probability that, in the absence of immigration, patch i will become extinct. We shall assume that the interaction probabilities depend on the distance d_{ij} between patches i and j (note that $d_{ii} = 0$ and that $d_{ij} = d_{ji}$) and the area A_i of patch i in the following way (see [4]):

$$q_{ij} = \exp(-e^{-ad_{ij}} A_i), \quad i, j \in \mathcal{N},$$

where $a(\geq 0)$ measures how badly individuals are at migrating. Thus, the larger the area of a given patch, the more likely that patch is to survive and to successfully colonize other patches, while the larger the distance between patches the smaller the chance of colonization between them. In the extreme case $a = 0$, colonization does not depend on the distance between colonies; otherwise, the larger the value of a , the more pronounced is the effect of distance between patches.

The various colonization processes and local extinction processes are assumed to be independent. Hence, we can define $q_i(x)$, where $x = (x_1, x_2, \dots, x_n)$, by

$$q_j(x) = \prod_{i=1}^n q_{ij}^{x_i}, \quad j \in \mathcal{N}, x \in \mathcal{S},$$

to be the probability that patch j will become extinct at the next epoch given a present configuration x . Thus it is clear not only that the model accounts for spatial structure in the population but also that the local extinction probability may depend on the effect of migration.

Finally, the transition matrix $P = (p(x, y), x, y \in \mathcal{S})$ can be written as

$$p(x, y) = \prod_{i=1}^n q_i(x)^{1-y_i} (1 - q_i(x))^{y_i}, \quad x, y \in \mathcal{S}.$$

Notice that, since $q_i(0) = 1$ for all $i \in \mathcal{N}$, state $0 = (0, 0, \dots, 0)$ (corresponding to the extinction of all patches) is an absorbing state for the chain:

$$p(0, y) = \begin{cases} 1, & \text{if } y = 0, \\ 0, & \text{otherwise.} \end{cases}$$

We shall assume that the remaining states $\mathcal{C} = \{x \in \mathcal{S} : x \neq 0\}$ form an irreducible, aperiodic class (Assumptions A1-A3 of [4] guarantee this). We shall also assume that $q_{ii} > 0$ for all $i \in \mathcal{N}$, so that, locally, every patch has

a positive probability of extinction. This ensures that $p(x, 0) > 0$ for some $x \in \mathcal{C}$, so that the absorbing state is accessible from all states in \mathcal{C} , and hence, since \mathcal{C} is finite, eventual extinction will occur with probability 1.

A salient feature of this model is that it helps account for the persistence of metapopulations. Figure 1 shows a simulation of a 5-patch metapopulation. The number of occupied patches is plotted against time up to extinction

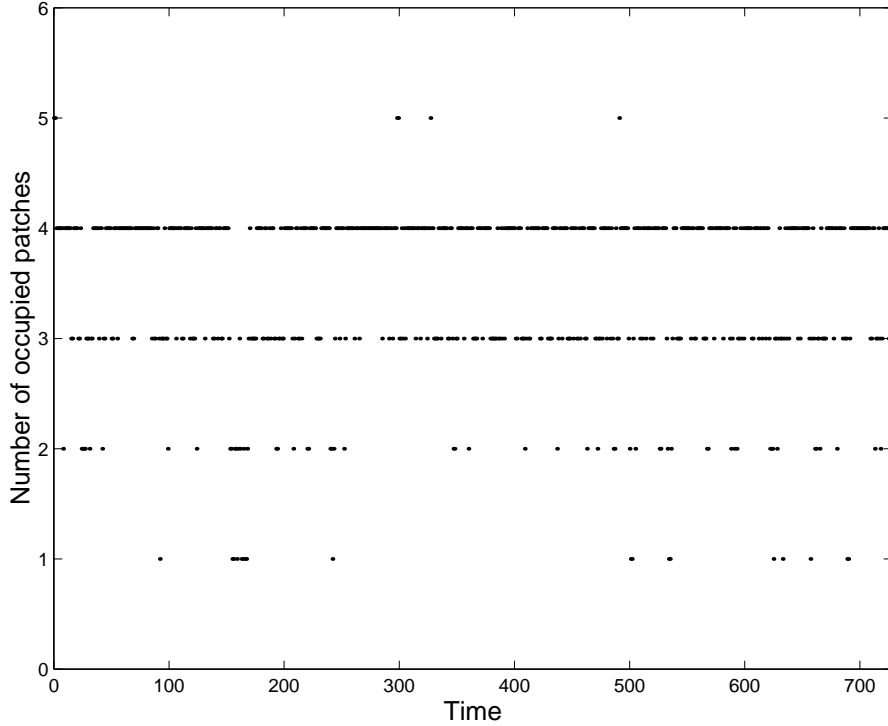


Fig. 1. Simulation of a 5-patch metapopulation with $a=7$ and initially all patches occupied.

at $t = 728$. All patches have the same area ($A_i = 1$). The distance between Patches 2,3,4 and 5 is the same ($d_{ij} = 0.1$), while Patch 1 is 10 times that distance away from each of the others ($d_{1j} = 1$).

Do the conditional state probabilities account for the observed behaviour? Figure 2 compares the observed frequencies for the sample path illustrated in Figure 1, with the conditional state distribution $m_x(t)$ at $t = 1, 2, 5$ and 10. In each case, the black bar is the distribution of the number of occupied patches evaluated using $m_x(t)$, while the white bar is the proportion of time for which i patches were occupied ($i = 1, 2, \dots, 5$) during the period of the simulation. The conditional state distribution was evaluated using (1) and the iteration

$$p(t+1) = p(t)P, \quad t = 0, 1, \dots, \quad (2)$$

with $p(0)$ assigning all its mass to the initial state. Notice that when t is large the conditional state distribution is in reasonable agreement with empirical

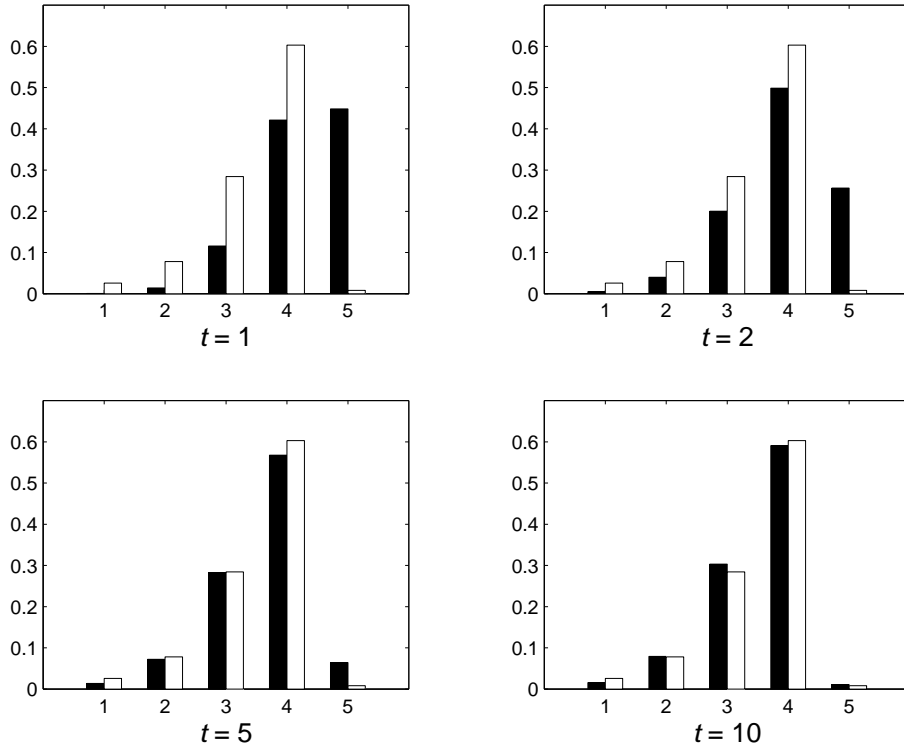


Fig. 2. Comparison between the conditional state distribution (black) and the simulated proportions of occupied patches (white).

frequencies. This agreement is not surprising from a theoretical point of view, for in a sense which can be made very precise, the set of frequencies is a “good” estimate of the state distribution for t large. My purpose here is to endorse the reader’s common sense: that using conditional state probabilities is a sensible approach to modelling the behaviour of the population before extinction occurs.

3 LIMITING CONDITIONAL DISTRIBUTIONS

The trend illustrated in Figure 2 has a simple theoretical explanation. Since \mathcal{C} is a finite set, the limit

$$\lim_{t \rightarrow \infty} m_x(t) = m_x \quad (3)$$

exists and defines a proper distribution $m = (m_x, x \in \mathcal{C})$, called a *limiting conditional distribution*, and m is the left eigenvector of $P_{\mathcal{C}}$ (P restricted to \mathcal{C}) corresponding to the eigenvalue, ρ_1 , with maximal modulus. This is true for any aperiodic Markov chain with a finite transient class \mathcal{C} ; see Darroch and Seneta [1]. Indeed, classical (Perron-Frobenius) matrix theory guarantees that, under our assumptions, ρ_1 has multiplicity 1, it is real and strictly less than 1, and, the corresponding left and right eigenvectors have strictly positive entries.

Figure 3 illustrates the rapid convergence of $m_x(t)$ to m_x for the 5-patch metapopulation model. The corresponding distributions of the number of occupied patches are plotted and compared at times $t = 2, 4, 9$ and 15 .

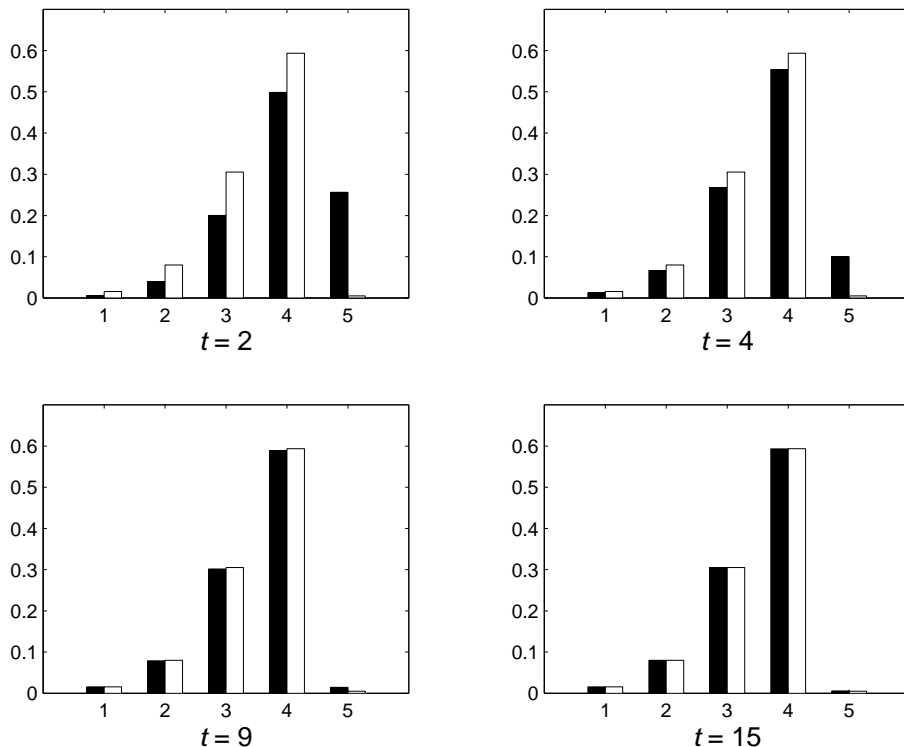


Fig. 3. Comparison between the conditional state distribution (black) and the limiting conditional distribution (white) for the number of occupied patches.

Again since \mathcal{C} is finite, we can be precise about the *rate* of convergence in (3) by examining the eigenvalue, ρ_2 , of $P_{\mathcal{C}}$ with *second-largest* modulus. This eigenvalue might not be real, and it might have a multiplicity, κ , which is greater than 1, but, for simplicity, let us suppose that $\kappa = 1$. It can be shown (see [1]) that

$$m_x(t) = m_x + O(\beta^t) \quad \text{as } t \rightarrow \infty,$$

where $\beta = |\rho_2|/\rho_1 (< 1)$. Thus, the smaller $|\rho_2|$ is compared with ρ_1 , the faster the convergence of $m_x(t)$ to m_x . Further, the expected time till absorption, τ , is approximately $\rho_1/(1 - \rho_1)$ and so if, in addition, ρ_1 is close to 1, we should expect m to faithfully describe the behaviour of the population before extinction. For the 5-patch metapopulation model with $a = 7$, we find that $\rho_1 \simeq 0.9979$, $\rho_2 \simeq 0.6312$ (real with multiplicity 1), $\beta (= \rho_2/\rho_1) \simeq 0.6325$ and $\tau \simeq 488$.

Before we proceed to examine the notion of a pseudo-stationary distribution, we remark that usage of the term *quasi-stationary distribution* for limiting conditional distribution is commonplace. However, in recent times, *this* term

has been reserved for any initial distribution $(\alpha_x, x \in \mathcal{C})$ such that the conditional probability $m_x(t)$ *does not depend on t* : $m_x(t) = \alpha_x$ for all $t > 0, x \in \mathcal{C}$. Every limiting conditional distribution is a quasi-stationary distribution in this sense, but the converse is not true; see for example Nair and Pollett [7], van Doorn [9], and van Doorn and Schrijner [10]. The interchangeability of the terms arose perhaps because, for Markov chains with \mathcal{S} irreducible, the *stationary distribution* and *limiting distribution* coincide. The problem of identifying conditions under which a quasi-stationary distribution is also a limiting conditional distribution has been the subject of much recent research; see [6].

4 PSEUDO-STATIONARY DISTRIBUTIONS

If we had assumed that a given patch, say Patch 1, had a zero local extinction probability ($q_{11} = 0$), that patch would behave, in an obvious sense, as a *mainland*. State 0 would no longer be accessible from all states and indeed \mathcal{C} would decompose into two irreducible classes, \mathcal{C}_0 and \mathcal{C}_1 , consisting of those states in \mathcal{C} which have, respectively, $x_1 = 0$ and $x_1 = 1$: either the process would start in \mathcal{C}_1 (the mainland is inhabited) and remain there, or, start in \mathcal{C}_0 (the mainland is uninhabited) and eventually enter either \mathcal{C}_1 or the absorbing state.

The method of Gyllenberg and Silvestrov depends on being able to identify a “quasi-mainland”, namely a single patch i with q_{ii} small; take this to be Patch 1. By considering a sequence of processes indexed by $\epsilon = q_{ii}$, Gyllenberg and Silvestrov were able to invoke the powerful perturbation theory of Markov chains by treating ϵ as a perturbation. It is worth reviewing their results in some detail.

Let $\epsilon \in (0, 1]$ (now arbitrary) and suppose that our interaction matrix depends on ϵ in the following way:

$$q_{ij}^{(\epsilon)} = q_{ij} + \epsilon \hat{q}_{ij} + o(\epsilon) \quad \text{as } \epsilon \rightarrow 0,$$

where

$$q_{ij} = \lim_{\epsilon \rightarrow 0} q_{ij}^{(\epsilon)} \quad \text{and} \quad \hat{q}_{ij} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (q_{ij}^{(\epsilon)} - q_{ij}),$$

the latter assumed to be non-negative and finite, and, that $Q = (q_{ij}, i, j \in \mathcal{N})$ satisfies $q_{11} = 0$. (This notation might cause some confusion; it is important to realize that $Q^{(\epsilon)}$ is the original interaction matrix, with ϵ chosen appropriately (say $\epsilon = q_{11}$), and that here Q is the interaction matrix obtained in the limit as $\epsilon \rightarrow 0$.) Then, in an obvious notation,

$$p^{(\epsilon)}(x, y) = p(x, y) + \epsilon \hat{p}(x, y) + o(\epsilon), \quad x, y \in \mathcal{S},$$

where $P^{(\epsilon)} = (p^{(\epsilon)}(x, y), x, y \in \mathcal{S})$ is the transition matrix corresponding to $Q^{(\epsilon)}$ and $P = (p(x, y), x, y \in \mathcal{S})$ is the transition matrix corresponding to Q . It is worth emphasizing at this stage that, of necessity, $q_{11}^{(\epsilon)} \rightarrow 0$, and so the choice $\epsilon = q_{11}$ is a natural one, and one which Gyllenberg and Silvestrov adopted in all their examples (see Section 7 of [4]).

Next, Gyllenberg and Silvestrov examined the asymptotic behaviour of the state probabilities, obtained from $P^{(\epsilon)}$, by letting $\epsilon \rightarrow 0$ and $t(= t_\epsilon) \rightarrow \infty$ in such a way that $\epsilon t_\epsilon \rightarrow s$, where $0 \leq s \leq \infty$. This is an intriguing idea. By arguing that the expected lifetime of the quasi-mainland is of order $1/\epsilon$, they were able to study the process on different time scales: smaller than, larger than, and of the same order as, the expected lifetime (these corresponding, respectively, to $s = 0$, $s = \infty$, and $0 < s < \infty$). They proved that, for $x, y \in \mathcal{C}$, the limit

$$\lim_{\epsilon \rightarrow 0} \Pr(X(t_\epsilon) = y | X(0) = x)$$

exists and is given by a mixture of the limiting probabilities $\pi(x, y)$ for the (ergodic) chain generated by Q and the degenerate distribution $\delta(y, 0)$ which assigns all its mass to state 0, the mixing probability being $e^{-\lambda s}$, where λ is a positive constant which is specified in terms of $\hat{p}(x, y)$. Gyllenberg and Silvestrov called this mixture a *quasi-stationary* distribution; later, in [5], they coined the term *pseudo-stationary* distribution to distinguish it from the distributions described in Section 3. Note that $\pi(y) = \pi(x, y)$ is the same for all $x \in \mathcal{C}_1$, that $\pi(y) = 0$ for $y \in \{0\} \cup \mathcal{C}_0$ and strictly positive otherwise, and, that for $x \in \mathcal{C}_0$ and $y \in \mathcal{C}_1$, $\pi(x, y) = h(x, \mathcal{C}_1)\pi(y)$, $h(x, \mathcal{C}_1)$ is the probability that the chain reaches \mathcal{C}_1 starting in x .

The most interesting case for the practitioner is $s = 0$, where the population is observed before extinction occurs; when t_ϵ grows *more slowly* than the expected lifetime, the limiting state probabilities converge to the distribution obtained by setting $q_{11} = 0$. In all the examples given in [4], only this case was examined.

The case $s = \infty$ gives rise to an obvious conclusion: that if the population is observed on a time scale with t_ϵ growing *more rapidly* than the expected lifetime, then the limiting distribution would be degenerate.

The most interesting case from a mathematical point of view is $0 < s < \infty$, where the process is observed on the same time scale as expected lifetime—though the author contends that the observer could not possibly discern this. Here, the limiting distribution is a genuine mixture of π and δ .

5 A COMPARISON

In this section we shall compare the two approaches to modelling quasi-stationary behaviour in metapopulations.

First let us return to the 5-patch model studied in Section 2. Recall that all patches had the same area ($A_i = 1$), Patches 2,3,4 and 5 (“islands”) were a distance 0.1 from one another, while Patch 1 (the quasi-mainland) was a distance 1 from each of the others. Figure 4 compares the limiting conditional distribution, the simulated proportions of occupied patches and the pseudo-stationary distribution. The disparity is marked: for this example, the two

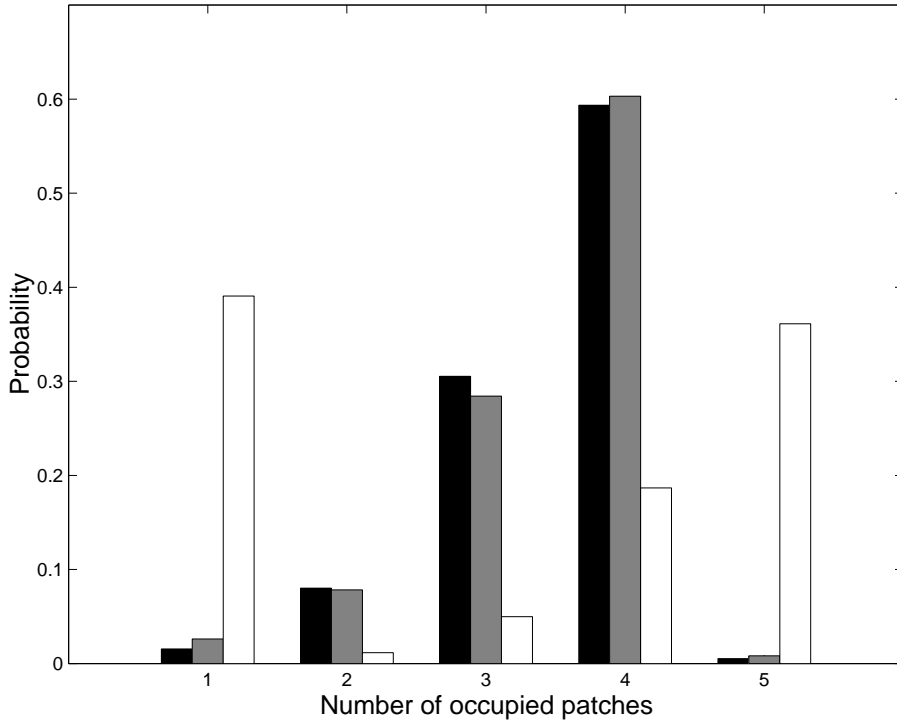


Fig. 4. Comparison between the limiting conditional distribution (black), the simulated proportions of occupied patches (grey) and the pseudo-stationary distribution (white).

ways of analysing the model lead to quite different predictions. And, as illustrated in Figure 5, this disparity becomes worse as the time-scale parameter s increases. The source of the disparity is easy to identify. The limiting conditional distribution assigns mass to *all* states, whereas the pseudo-stationary distribution assigns mass to only those states x with $x_1 = 1$, and, for these states the distributions are markedly different. For example, under the limiting conditional distribution, state $(1, 0, 0, 0, 0)$ has negligible mass, whereas the pseudo-stationary distribution assigns nearly half its mass to this state.

To illustrate this graphically, it is simpler to consider the corresponding 3-

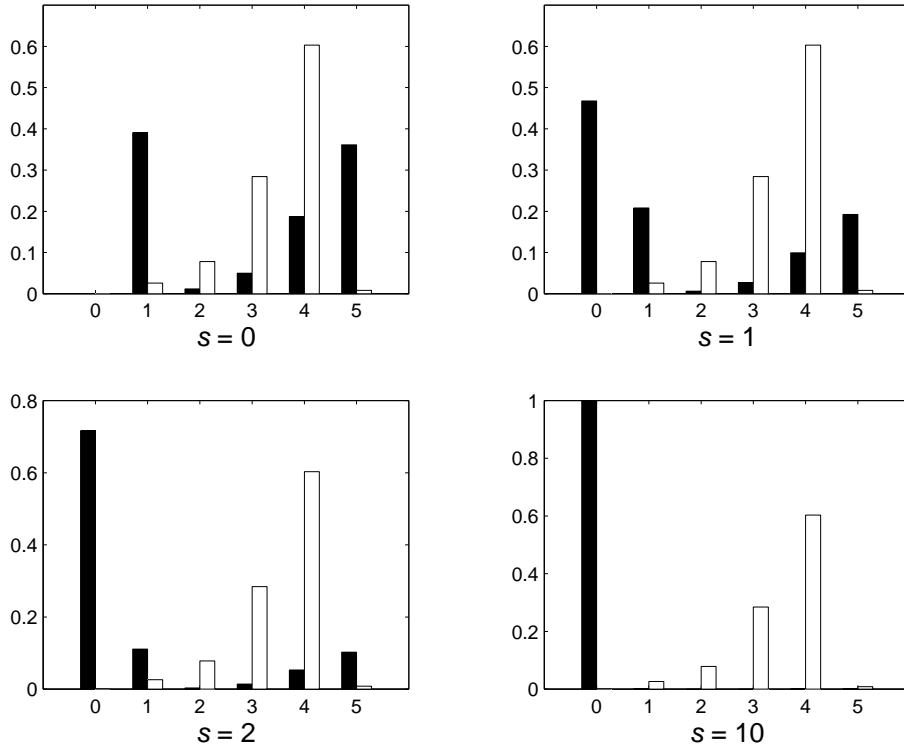


Fig. 5. The effect of varying s on the pseudo-stationary distribution (black). The white bar represents the simulated proportions of occupied patches.

patch model. All patches have the same area, the distance between the islands (Patches 2 and 3) is 0.1, and the quasi-mainland (Patch 1) is 10 times that distance away from each of 2 and 3. This is precisely the first example of [4]. Figures 6 and 7 compare the limiting conditional distribution, the simulated proportions of occupied patches, and the pseudo-stationary distribution for the 3-patch metapopulation model with $a = 3$. Figure 6 shows the distribution of the number of patches, while Figure 7 shows that actual state distribution. The white bar in Figure 6 can be compared with Figure 1(d) of [4]. The disparity is particularly noticeable in the state distribution. Observe, in particular, the unexpectedly high probability assigned to state $(1, 0, 0)$ under the pseudo-stationary distribution.

We remark that for the 3-patch model, repeated runs are needed to accurately estimate the proportions of occupied patches, for extinction occurs quickly (the expected time till extinction being $\tau \simeq 12.94$ starting, as we did, in state $(0, 1, 1)$). Our results are based on 1000 runs. For the $a = 3$ model we have $\rho_1 \simeq 0.9155$, $\rho_2 \simeq 0.6005$ and $\beta \simeq 0.6560$. The anticipated rapid convergence of the conditional state distribution is observed.

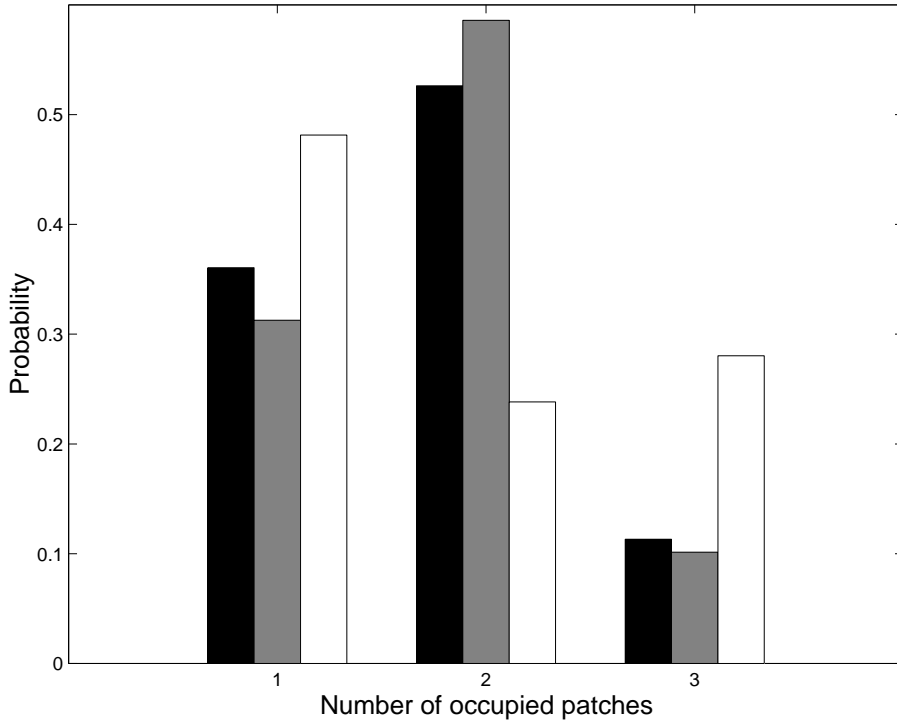


Fig. 6. Comparison between the limiting conditional distribution (black), the simulated proportions of occupied patches (grey) and the pseudo-stationary distribution (white) for a 3-patch metapopulation model with $a = 3$.

6 DISCUSSION

In a recent paper, Gosselin [3] has attempted to reconcile the two approaches by establishing an intuitively obvious result. Denote the state probabilities corresponding to $P^{(\epsilon)}$ by $p^{(\epsilon)}(t) = (p_x^{(\epsilon)}(t), x \in \mathcal{S})$, so that

$$p^{(\epsilon)}(t+1) = p^{(\epsilon)}(t)P^{(\epsilon)}, \quad t = 0, 1, \dots,$$

and denote the corresponding conditional probabilities by $m_x^{(\epsilon)}(t)$ (cf (1)). Gosselin proved that

$$\lim_{\epsilon \rightarrow 0} \lim_{t \rightarrow \infty} m_x^{(\epsilon)}(t) = \begin{cases} \pi(x), & \text{if } x \in \mathcal{C}_1, \\ 0, & \text{if } x \in \mathcal{C}_0, \end{cases}$$

which he compared with Theorem 6.2 of [4]:

$$\lim_{\epsilon \rightarrow 0} m_x^{(\epsilon)}(t_\epsilon) = \begin{cases} \pi(x), & \text{if } x \in \mathcal{C}_1, \\ 0, & \text{if } x \in \mathcal{C}_0. \end{cases}$$

(These results have been stated here in slightly more generality than in [3] and [4]; they assume, or condition on, a particular initial state, but since \mathcal{C}

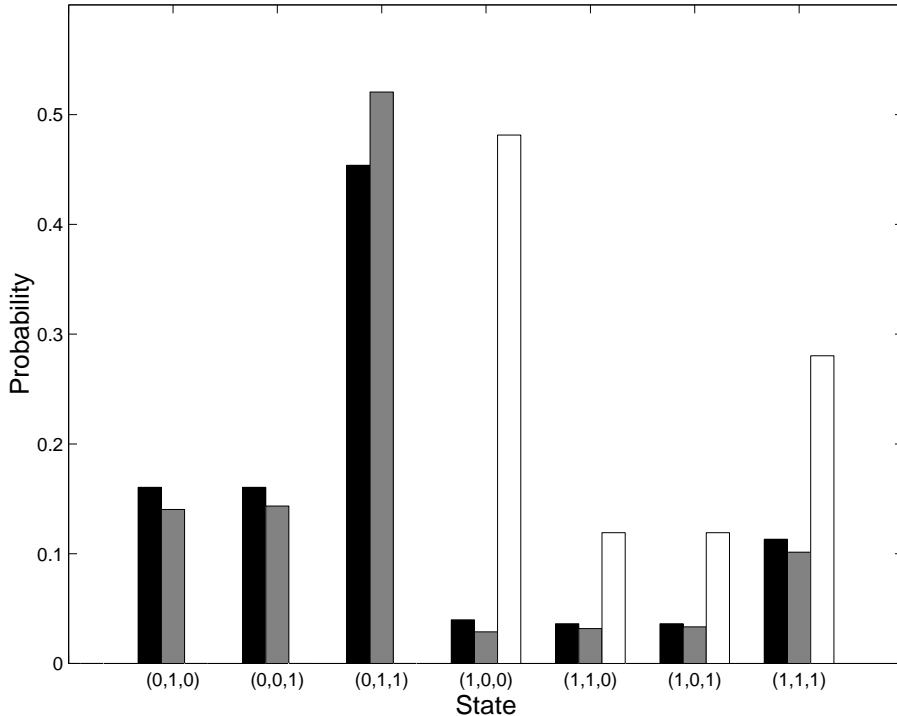


Fig. 7. Comparison between the limiting conditional distribution (black), the simulated proportions of occupied patches (grey) and the pseudo-stationary distribution (white) for a 3-patch metapopulation model with $a = 3$.

is a finite set, we may use an arbitrary initial *distribution* $p(0)$ over states, in which case the limits do not depend on $p(0)$.) Thus, in the important case $s = 0$, the limiting conditional distribution and the pseudo-stationary agree when ϵ is small. The problem with the models examined earlier is that $\epsilon (= q_{11}) \simeq 0.3679$, independently of the number of patches.

However, quasi-stationary behaviour is a *property of the model* and *not* the means of analysing it. The 5-patch model exhibits quasi-stationarity, demonstrated emphatically in Figure 1, yet q_{11} is not small. The pseudo-stationary distribution does not capture this behaviour. On the other hand, the conditional state distribution $m(t)$ does: after all, it is the *most information our model can provide* at any time t given that we know extinction has not occurred by time t . In cases when the convergence of $m(t)$ to the limiting conditional distribution m is rapid, *this* distribution can be used instead.

A case for which q_{11} is small, is the second example of [4]. It differs from the 3-patch example described above only in that $A = (10, 5, 0.001)$ and $d_{23} = d_{32} = 0.001$, so that the quasi-mainland is well away from the two islands, a distance 1000 times that between the islands, and, of the two islands, one is half the size of the quasi-mainland and 5000 times that of the other island. When $a = 7.5$ we find that $\epsilon \simeq 0.000045$, and that $\rho_1 \simeq 0.9999$, $\rho_2 \simeq 0.9905$ and $\beta \simeq 0.9906$. Thus, the expected time till extinction is large, while convergence

of the conditional state distribution is slow. Figure 8 shows the distribution of the number of occupied patches. The white bar in Figure 8 can be compared with Figure 3(c) of [4].

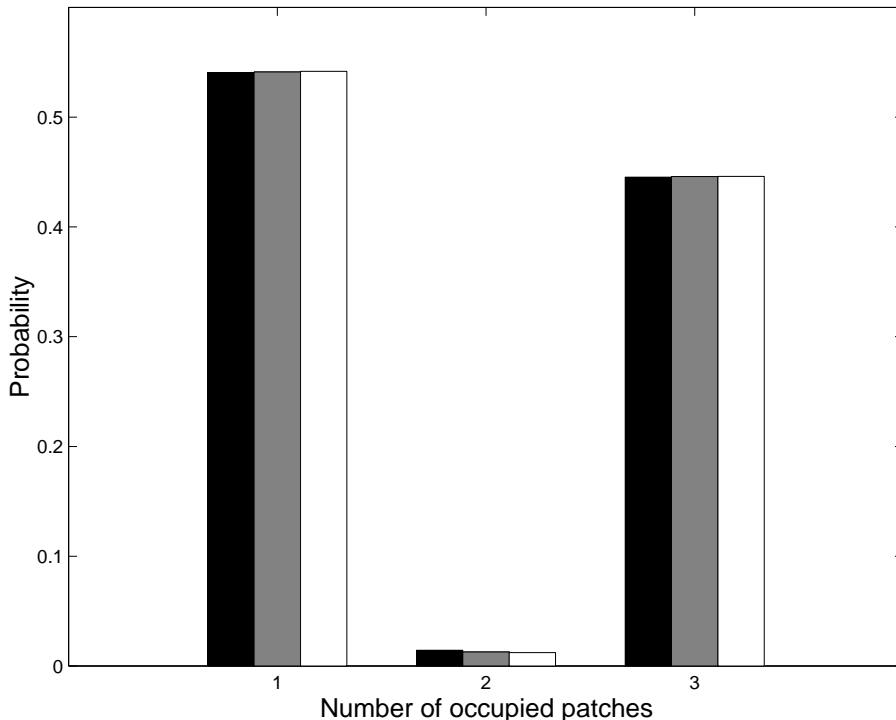


Fig. 8. Comparison between the limiting conditional distribution (black), the simulated proportions of occupied patches (grey) and the pseudo-stationary distribution (white) for the second 3-patch metapopulation model with $a = 7.5$.

We shall conclude with a brief discussion of computational issues. There are many numerical software libraries available which include routines for performing matrix computations and, in particular, for evaluating eigenvalues and eigenvectors. Perhaps the most widely used is MATLAB. In using this package, we have three obvious methods available to us. We could use the function `eig`, which evaluates all eigenvalues and/or eigenvectors using the QR algorithm, or the function `eigs` (available in Version 5 or above), which evaluates particular eigenvalues and/or eigenvectors using the Arnoldi algorithm, or, we could simply use iteration based on (1) and (2). We note that the Arnoldi algorithm is normally used for large-sparse systems (see for example [8]), but can be used in the present context (P is always dense), though its advantages will not be fully realized. Table 6 compares these methods. The size of the system, measured by the number of states (equivalently, the size of P), is given in Column 1. Columns 2,3 and 4 list the numbers of *flops* (floating point operations) used by each method; the *iteration* was stopped (at t given in the last column) when the maximum elementwise difference between the state vector and the corresponding result of the Arnoldi method was smaller than 10^{-6} (this being also the default tolerance of the Arnoldi method). The

2^n	Arnoldi	QR	Iteration	t
4	3543	1182	932	29
8	41966	12610	5768	45
16	260232	64996	15888	31
32	740295	527710	59424	29
64	1116134	3554160	245824	30
128	2264912	30466852	983168	30
256	5425752	223682282	3932416	30
512	17161871	–	15729152	30
1024	65372004	–	62915584	30

Table 1. Comparison between various computational methods.

QR algorithm works well for small systems, while the Arnoldi algorithm appears to be better for large systems. The performance of the simplest method, namely iteration based on (1) and (2), is always the best.

Acknowledgement

The author is grateful to Nigel Bean, Jemery Day, Andrew Hart and Pauline Schrijner for valuable comments on this work, and to Frédéric Gosselin and Matts Gyllenberg for making their work available prior to publication.

References

- [1] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time Markov chains. *J. Appl. Probab.*, 2:88–100, 1965.
- [2] J.R. Day and H.P. Possingham. A stochastic metapopulation model with variable patch size and position. *Theoret. Pop. Biol.*, 48:333–360, 1995.
- [3] F. Gosselin. Reconciling theoretical approaches to stochastic patch-occupancy metapopulation models. *Bull. Math. Biol.*, 1998. To appear.

- [4] M. Gyllenberg and D.S. Silvestrov. Quasi-stationary distributions of a stochastic metapopulation model. *J. Math. Biol.*, 33:35–70, 1994.
- [5] M. Gyllenberg and D.S. Silvestrov. Exponential asymptotics for limits of renewal equations and pseudo-stationary phenomena for stochastic systems. *Umeå University, Department of Mathematical Statistics, Research Report 3*, 1997.
- [6] H. Kesten. A ratio limit theorem for (sub) Markov chains on $\{1, 2, 3, \dots\}$ with bounded jumps. *Adv. Appl. Probab.*, 27:652–691, 1995.
- [7] M.G. Nair and P.K. Pollett. On the relationship between μ -invariant measures and quasistationary distributions for continuous-time Markov chains. *Adv. Appl. Probab.*, 25:82–102, 1993.
- [8] P.K. Pollett and D.E. Stewart. An efficient procedure for computing quasistationary distributions of Markov chains with sparse transition structure. *Adv. Appl. Probab.*, 26:68–79, 1994.
- [9] E.A. van Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. Appl. Probab.*, 23:683–700, 1991.
- [10] E.A. van Doorn and P. Schrijner. Geometric ergodicity and quasi-stationary distributions in discrete-time birth-death processes. *J. Austral. Math. Soc.*, 37:121–144, 1995.