Next-generation matrices for marine metapopulations: the case of sea lice on salmon farms[*]

Peter D. Harrington[†], Danielle L. Cantrell [‡] and Mark A. Lewis[†§]

April 19, 2022

Abstract

Classifying habitat patches as sources or sinks and determining metapopulation persistence requires coupling connectivity between habitat patches with local demographic rates. In this paper we show how next-generation matrices, originally popularized in epidemiology to calculate new infections after one generation, can be used in an ecological context to couple connectivity with local demography to calculate sources and sinks as well as metapopulation persistence in marine metapopulations. To demonstrate the utility of the method, we construct a next-generation matrix for a network of sea lice populations on salmon farms in the Broughton Archipelago, BC, an intensive salmon farming region on the west coast of Canada where certain salmon farms are currently being removed under an agreement between local First Nations and the provincial government. We identify the salmon farms which are acting as the largest sources of sea lice and show that in this region the most productive sea lice populations are also the most connected. We find that the farms which are the largest sources of sea lice have not yet been removed from the Broughton Archipelago, and that warming temperatures could lead to increased sea louse growth.

Keywords: next-generation matrix, metapopulation, source-sink dynamics, marine systems, sea lice, salmon farms

1 Introduction

Metapopulations consist of subpopulations located on isolated habitat patches that are connected via dispersal [1, 2, 3]. In most benthic marine species, this dispersal comes from the pelagic larval stage [4]. Larvae disperse between, and then settle on habitat patches, and once settled the remaining stages are sedentary and remain confined to a specific habitat patch. In marine systems the metapopulation concept, where subpopulations are connected but have their own demographic rates, has been used in the spatial planning of Marine Protected Areas and the siting of marine reserves [5, 6, 7, 8, 9, 10].

In a metapopulation framework, habitat patches are often classified into sources and sinks based on how the subpopulations on these patches contribute to the overall metapopulation. The source-sink classification of habitat patches was first described concretely in a terrestrial context by Pulliam

---

[11], where habitat patches were classified as sources if the local subpopulations could persist in isolation and sinks if they could not. However, this classification ignores the effect of dispersal, which is especially critical in marine metapopulations, and so Runge et al. [12] and Figueira and Crowder [13] updated the classification of source and sink patches to include both the local productivity of a patch, as well as the ability to disperse away from the habitat patch. Under this new classification a source patch is a patch on which an adult will more than self replace over the entire metapopulation and a sink is a patch on which an adult will not. Self-replacement need not occur on the same habitat patch as the adult originated, and thus under this classification a source patch may not be able to persist in isolation.

Due to the large scale larval dispersal that occurs in many marine species, it is common in marine metapopulations for sources patches to not be able to persist in isolation [14], and thus preserving a persistent metapopulation, especially in the context of MPAs, often requires more than simply preserving source patches. To maintain persistent marine metapopulations it is necessary to preserve sufficient larval exchange between closed loops of habitat patches so that an average adult can eventually self replace over multiple generations [15, 10]. If sources patches self-recruit enough larvae to persist in isolation then this can be accomplished by preserving only source patches. However if they do not it may require preserving both source and sink patches, as the sink patches may provide sufficient larval exchange back to the source patches to create a closed loop of habitat patches over which an adult can self replace. Evaluating the persistence of marine metapopulations is therefore difficult as it requires accurate measures of larval connectivity between habitat patches, as well as accurate local demographic rates of adult stages on each habitat patch [10]. Despite the difficulties, evaluating metapopulation persistence is critical in designing Marine Protected Areas in which the protected habitat patches can persist even when outside patches are exploited [16, 17, 18, 19, 20, 21].

Next-generation matrices are a useful tool that can both be used to evaluate the persistence of a metapopulation as well as identify the contribution of local habitat patches under a variety of modeling frameworks. Originally popularized in epidemiology as a simple method of calculating the basic reproduction number, R_0, in compartmental models [22], next-generation matrices convert population models into generational time, so that the entries are the number of new individuals produced in each compartment, or patch in the case of metapopulations, after one generation. The individual contribution of habitat patches or evaluation of metapopulation persistence can therefore be measured for different model structures (discrete time, continuous time, etc.) under the same framework of the next-generation matrix. The column sums can be used to measure the contribution of each habitat patch over a generation and the spectral radius can be used to evaluate metapopulation persistence [23]. Next-generation operators have previously been used in ecology to calculate source and sink regions in heterogeneous environments [24, 25, 26, 23], and to evaluate the level of control required to suppress invasive species [27].

In this paper we focus on using the next-generation matrix to evaluate local patch contribution and metapopulation persistence in marine metapopulations, but the framework used here is also applicable to many other birth-jump metapopulations [28], where there is a single juvenile stage which can disperse between habitat patches and the remaining stages remain on a single habitat patch. Examples of non-marine species that exhibit this structure include plant species where seeds are caried between habitat patches [29], or insect species with a single large dispersal event such as the spruce budworm [30] or mountain pine beetle [31] . In fact, the next-generation matrix approach can even be extended to metapopulations in which adults also disperse, though the calculations become more complicated and so here we focus on species with a single dispersing stage.

Specifically, to demonstrate the utility of the method, we use the next-generation matrix to calculate the contribution of a single salmon farm to the spread of sea lice in a salmon farming region on the west coast of British Columbia. Sea lice are a parasitic marine copepod that feed on the epidermal tissues, muscles, and blood of salmon [32]. With a free living larval stage they can disperse tens of kilometres, spreading between salmon farms in a region and between wild and farmed salmon [33, 34, 35]. Lesions and stress from high sea lice infestation make adult salmon more

2

susceptible to secondary infections, leading to large economic consequences for the salmon farming industry [36]. On wild juvenile salmon, infestation with sea lice can lead to mortality and elevated exposure to sea lice from salmon farms can contribute to population level declines in pink salmon [37]. In the context of sea lice on salmon farms we are not concerned with preserving a persistent metapopulation of sea lice parasites, but instead we use the next-generation matrix to calculate the relative contribution of each salmon farm and evaluate the effect of environmental variables on the overall growth of the sea louse metapopulation.

The specific salmon farming region that we focus on to calculate farm contribution is the Broughton Archipelago. The Broughton Archipelago is located on the west coast of Canada, between Vancouver Island and the mainland of British Columbia and has been central in evaluating the effect of sea lice from salmon farms on wild salmon [38, 35, 39, 40, 41, 37, 42, 43, 44, 45]. The area has historically had around 20 active salmon farms [46], though currently certain farms are being removed from this region in an agreement between the government of British Columbia and the Kwikwasut'inuxw Haxwa'mis, 'Namgis, and Mamalilikulla First Nations [47]. After 2023 many of the remaining farms must be approved by both the local First Nations and the government in order to continue to operate and thus determining the farms which are acting as the largest sources of sea lice is critical during this transition period.

The paper is structured as follows. First we demonstrate how to use the next-generation matrix to calculate the contribution of local habitat patches to the metapopulation and evaluate metapopulation persistence. Next, we highlight how to construct the next-generation matrix for different types of models. Then we calculate a next-generation matrix for sea louse populations in the Broughton Archipelago to identify which salmon farms are the largest sources of sea lice in this region, evaluate the effect of the current farm removals, and investigate the effect of environmental variables on metapopulation growth. Finally, we discuss how the calculations of patch contribution and metapopulation persistence from other studies compare to the calculations using the next-generation matrix.

2 Materials and Methods

In this section we present details on the construction of the next-generation matrix and show how it can be used to determine the contribution of the subpopulation on a single habitat patch to the metapopulation as well as determine the persistence of the metapopulation. We then present the explicit construction of the next-generation matrix for models with age dependent demography, and present the construction for ordinary differential equation models and discrete time models in the appendix. Finally, we detail the construction of the next-generation matrix for a system of sea lice populations on salmon farms in the Broughton Archipelago.

2.1 Next-generation matrices

Here we construct the next generation matrix for a single species marine metapopulation with a single larval stage that can disperse between patches and where the remaining stages are confined to the habitat patch on which the larvae settle. We assume that the last stage is the only stage that produces new larvae. This assumption can be relaxed, though the entries of the next-generation matrix become slightly more complicated. We construct the next-generation matrix for models where the metapopulation is divided into l patches and the m attached stages are modelled explicitly. The larval stage is modelled implicitly, so that the birth rate into the first attached stage includes both the birth rate of larvae and the probability of larvae successfully dispersing between patches and attaching on a new patch. The lifecycle diagram for such a metapopulation is shown in Figure 1.

We use the next-generation matrix to calculate local patch contribution to the metapopulation in the context of low population density. To calculate the next-generation matrix it is necessary to linearize a potentially density dependent model around the zero equilibrium so that the effect
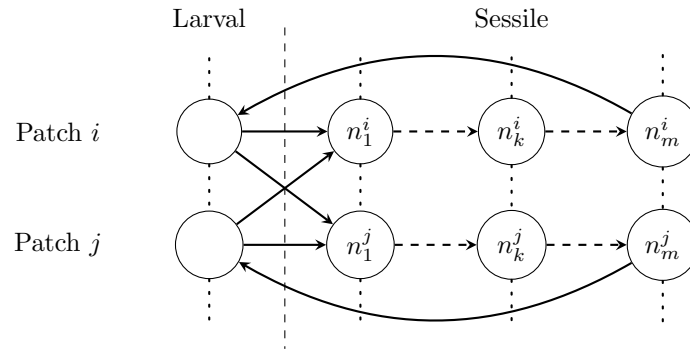
Figure 1: The lifecycle graph for two patches in a metapopulation of a species with a single larval stage that disperses between habitat patches and m sessile stages that remain on a habitat patch. The population on patch i in stage k is given by n_k^i.

of density dependence at higher population sizes is ignored when calculating patch contribution. This approach of ignoring density dependence is common when determining persistence or patch contribution of marine metapopulations, as the focus is either on determining if a metapopulation can persist at all, or determining which habitat patches are acting as population sinks and which are acting as population sources [15, 10, 24, 23]. Alternatively it is also useful for determining patch contribution in metapopulations of species that are being actively controlled to remain at low densities, such as sea lice on salmon farms, which is our focus in section 2.4. Another common assumption in the theory of persistence also made here is that there is no discernible Allee effect in any of the patches [15, 24, 10].

Under any model structure, the element in row i and column j of the next-generation matrix gives the number of new individuals produced on patch i by one new initial individual on patch j. However, how 'new' individuals are defined is subject to interpretation. Here, we construct the next-generation matrix under modeling frameworks that only consider the sessile stages explicitly, and so 'new' individuals will be newly attached stage 1 individuals. Then, when the number of new individuals on patch i produced from one new individual on patch j are tracked, the new individual must first survive and reproduce on patch j, before larvae disperse and arrive on patch i. In this way patch contribution, as will be calculated in section 2.2, is primarily a function of the local patch demography which is then coupled with dispersal to other patches. Under this framework, the entries of the next-generation matrix, K, for all model structures can be given by:

$$k_{ij} = \text{Pr}(\text{survival through sessile stages on patch } j) \times \# \text{ larvae produced on patch } j \times$$
$$\text{Pr}(\text{dispersal from patch } j \text{ to patch } i). \quad (1)$$

However if the next-generation matrix is constructed for models that explicitly model the larval stage, then the larval stage is often considered the first stage. The next-generation matrix will be slightly different in this case as well as the calculation of patch contribution, though the calculation of metapopulation persistence will be the same. We illustrate the differences between constructions of the next-generation matrix in the Discussion and show how the calculation of persistence remains the same under all constructions.

2.2 Determining patch contribution and metapopulation persistence

Here we show how to use the next-generation matrix, K, to determine the contribution of each patch to the metapopulation and evaluate metapopulation persistence. To determine the contribution of a specific patch to the metapopulation we track the total number of new individuals produced across the metapopulation after one generation from an initial individual starting on that patch. The entries of the next-generation matrix, k_{ij} give the number of new individuals produced on patch i from an initial individual on patch j. Therefore if we define

$$R_j = \sum_i k_{ij}, \tag{2}$$

so that R_j is the jth column sum of K, then R_j is the total number of new individuals produced across all patches from an initial individual starting on patch j and can be used to define the contribution of patch j to the entire metapopulation.

This definition of patch contribution easily lends itself to classifying local habitat patches as population sources or sinks. If $R_j < 1$, then an individual on patch j cannot replace itself over the entire metapopulation, and thus patch j is defined as a sink [23]. If $R_j > 1$ then one individual on patch j is producing more than one individual over the entire population and so patch j is defined as a source. To calculate persistence we can use the basic reproduction number, R_0, which can be calculated as

$$R_0 = \rho(K), \tag{3}$$

where $\rho()$ is the spectral radius. If $R_0 > 1$ then the metapopulation will persist and if $R_0 < 1$ then the metapopulation will go extinct, a relationship which holds under any of the model formulations considered here [22, 23, 48, 49]. The only condition required is that K be irreducible, which is biologically satisfied if there is some small positive probability that larvae leaving one patch can eventually arrive on any other patch.

There are several biologically reasonable properties that also exist mathematically under this framework. First, if the population on any single habitat patch can persist on its own, so that $k_{ii} > 1$, then the entire metapopulation will persist and $R_0 > 1$ [23]. Second, a metapopulation consisting only of sink patches cannot persist and a metapopulation consisting only of sources cannot go extinct. The mathematical underpinning of these relationships is that the spectral radius must be between the minimum and maximum column sums of a matrix, so in terms of our metapopulation quantities

$$\min_j R_j \le R_0 \le \max_j R_j. \tag{4}$$

Having defined patch contribution and persistence in terms of the next-generation matrix, we now demonstrate how to calculate the next-generation matrix for a model with age dependent demography, as this is the modeling structure commonly used for sea lice on salmon farms. We also present the construction of the next-generation matrix for discrete time models and ordinary differential equation models in the Appendices A and B, so that the details are present for the most commonly used population models

2.3 Calculating the next-generation matrix for models with age dependent demography

Here we calculate the next-generation matrix for models which allow for the maturation, survival, birth, and dispersal rates to depend not only on the stage and patch location of an individual, but also on the time that they have spent in a stage. This time is often referred to as stage-age and so in these models there are two time variables: the global time of the system, t, and the time that an individual has spent in a particular stage, their stage-age, a. These models can be specified either as McKendrick von-Foerster partial differential equations, integrodifferential equations, or

renewal equations [50]. In all cases, the dependence of the maturation rate on time spent in a stage allows for the addition of more realistic maturation functions where most individuals mature at some intermediate stage-age, or after some minimum time spent in the stage. In contrast, if models are formulated using ordinary differential equations the time in a stage is always exponentially distributed.

However, the specification of the model equations for these models can be rather complicated and so here we construct the next-generation matrix directly from the maturation, survival, birth, and dispersal functions, but the full model and the derivation of the next-generation matrix can be found in Appendix C as well as in [23]. Essentially the construction involves tracking the probability that an individual survives through the different stages on a specific patch, the number of larvae that they produce, and the probability that the larvae successfully disperse from one patch from another. If the probability of survival in stage k on patch j at stage-age a is $S_k^j(a)$, the maturation rate from stage k to $k+1$ is $m_k^j(a)$, the birth rate of larvae on patch j is $b^j(a)$, and the probability of a larva leaving patch j and successfully attaching on patch i is p^{ij}, then the entries of the next-generation matrix, K, are given by

$$k_{ij} = \underbrace{\left[\prod_{k=1}^{m-1} \overbrace{\left(\int_0^\infty S_k^j(a) M_k^j(a) m_k^j(a) da \right)}^{\text{Pr}(\text{stage } k \to k+1)} \right]}_{\text{survival through sessile stages}} \underbrace{\left(\int_0^\infty S_m^j(a) b^j(a) da \right)}_{\text{\# larvae produced}} \times \overbrace{p^{ij}}^{\text{Pr}(\text{patch } j \text{ to patch } i)} \quad (5)$$

where $M_k^j(a)$ is the probability that an individual has not yet matured from stage k to $k+1$ and can be calculated as $M_k^j(a) = \exp(-\int_0^a m_k^j(\tau)d\tau)$.

2.4 Application: sea lice on salmon farms in the Broughton Archipelago

In this section we construct a next-generation matrix to determine the contribution of a single salmon farm to the spread of sea lice in a salmon farming region on the west coast of British Columbia, the Broughton Archipelago. To determine the level of sea lice dispersal away from salmon farms, we use a hydrodynamic model, coupled with a particle tracking model, to track sea lice particles released from 20 historical farms in the Broughton Archipelago [51]. The coupled hydrodynamic and particle tracking models have been validated with sea louse counts on salmon farms in the Broughton [52]. We construct one next-generation matrix for the 20 historical farms in the region for which the hydrodynamic particle tracking model was run, as well as one for the 11 remaining farms in the area after 2023, subject to First Nations and governmental approval [47].

2.4.1 Modeling framework

Sea lice maturation through stages is often modelled with stage-age dependent maturation functions [53, 54, 55] and thus we construct the next-generation matrix for sea lice using age dependent demography, as shown in section 2.3.

2.4.2 Dispersal

The probability that a sea louse larvae leaves from one farm and successfully arrives on another depends on several factors including ocean current, temperature, and salinity. To accurately capture this probability it is necessary to use a computational hydrodynamic model that can track the spread of larvae originating from a given farm as well as the dependence of larval survival on temperature and salinity. In order to determine the probability of larvae dispersing between farms we use connectivity matrices from Cantrell et al. [51]. Details on these connectivity matrices can be found in [51]. Briefly, these connectivity matrices are calculated by applying Kernel Density Estimation (KDE) to

particle tracking simulations to calculate the infectious density of sea lice at each farm, originating from a given farm. The particle tracking simulations are run on output generated by a Finite Volume Community Ocean Model (FVCOM) which uses data on tides, wind surface-heating, and river discharge to simulate three-dimensional ocean velocity, temperature, and salinity [56]. In the particle tracking simulation the survival of sea louse particles is dependent on temperature and their maturation from non-infectious to infectious larvae is dependent on temperature.

The infectious densities of sea lice are calculated for each particle release day by taking daily snapshots of the particle locations of infectious lice for 11 days post release and then applying Kernel Density Estimation to the accumulated daily snapshots. A connectivity matrix is then calculated for each particle release day, where the entry in row i and column j of the connectivity matrix is the infectious density of larvae over farm i, produced by larvae initially leaving farm j. In [51] the infectious densities were calculated from 24 hours of particle releases, where 50 particles were released each hour and so to calculate the infectious density of one initial release particle we divide the entries in each connectivity matrix by 1200 (50×24). Then, to create a single connectivity matrix, C, for the 20 farms in the Broughton Archipelago we take the average over all the connectivity matrices created for particles released between March 14th and July 20th, 2009.

The necessary quantity to construct the next-generation matrix is p^{ij}, the probability that larvae leaving farm j will successfully attach on farm i. To estimate p^{ij} from the entries of the connectivity matrix, c_{ij}, there are several assumptions that need to be made. If we assume that the number of lice that arrive onto farms is small compared to the total number of lice in the water column, so that lice arriving onto farms do not significantly affect the density of lice in the water column, then

$$p^{ij} = \int_0^\infty \beta \int_{\text{farm } i} p_h^j(x,t)d\Omega dt,$$

where $p_h^j(x,t)$ is the two dimensional density of infectious lice produced from farm j that are still alive in the water column at position x and time t, and β is the arrival rate of lice moving over the farm arriving onto the farm. However the entries of the connectivity matrix, c_{ij}, are the infectious density of larvae over farm i, produced by larvae leaving farm j. The infectious densities are calculated by applying Kernel Density Estimation to daily snapshots of infectious particles over 11 days, starting at time $t = 0$, so roughly

$$c_{ij} = p_h^j(x_i,0) + p_h^j(x_i,1) + \cdots + p_h^j(x_1,11).$$

We can therefore roughly calculate p^{ij} from c_{ij} by assuming that the integrals over time and space can be approximated using their Riemann sums, and that the area of a farm is roughly 0.01 km^2, where

$$p^{ij} = \int_0^\infty \beta \int_{\text{farm } i} p_h^j(x,t)d\Omega dt \tag{6}$$

$$\approx \beta \times \text{Area of farm } i \times 1 \text{ day} \times \left(p_h^j(x_i,0) + p_h^j(x_i,1) + \cdots + p_h^j(x_i,11) \right) \tag{7}$$

$$= \beta \times 0.01 \times c_{ij}. \tag{8}$$

Therefore to calculate p^{ij} from c_{ij} we need to estimate the arrival rate β. However, very little is known about the arrival rate of lice dispersing from one farm to another, and so the estimate presented here is very uncertain and could be orders of magnitude off from the true arrival rate. We assume here that $\beta = 100/\text{day}$, and thus the average waiting time for infectious sea lice in the water column surrounding the farm to arrive on the farm is roughly 15 minutes ($1/\beta = (1 \text{ day}/100) \times (24 \text{ hrs}/1 \text{ day}) \times (60 \text{ minutes}/1 \text{ hr}) = 14.4$ minutes). However the waiting time could be as little as 1 minute or as long as 1 hour, depending on environmental conditions. Moreover, this is ignoring the fact that some infectious lice may never attach, and the estimates of the proportion of lice which successfully attach at all varies from 80% to 0.5% under different lab conditions [57].

It can be seen from equations (5) and (8) that each entry in the next-generation matrix, K, is a linear function of the arrival rate β, and therefore changing β will not affect the relative ordering of the contributions of different salmon farms, given by R_j (equation 2), it will only affect the absolute magnitude. The basic reproduction rate, $R_0 = \rho(K)$, is also a linear function of β and therefore any increase or decrease in β will result in the same proportional increase or decrease in R_0. However, the purpose of this application is not to accurately estimate the basic reproduction number, R_0, for the Broughton Archipelago, or to accurately estimate the contribution of an individual farm, R_j, but rather to compare the relative contributions of different salmon farms in the system, and to investigate the effect of environmental variables on the basic reproduction number. Therefore we present our estimate of the arrival rate for these purposes only, and our value of R_0 found in the results should not be taken as an accurate estimate.

2.4.3 Demography

Once infectious sea lice larvae attach to their salmonid hosts they must survive and mature through several attached life stages before they can produce offspring. These demographic rates are dependent on salinity and temperature, and thus some salmon farms may be more productive than others due to favourable environmental conditions. To capture the dependence of demography on salinity and temperature we simplify the attched sea lice life cycle down to three main stages: chalimus, pre-adult, and adult. Survival in each stage is salinity dependent, maturation is temperature dependent, and egg viability and production depends on both salinity and temperature. The demographic functions that we use are from models which have previously been fit to sea louse population data and are shown in Table 1.

We calculate the on-patch component of the elements of next-generation matrix, k_{ij}, by integrating the demographic functions over all time,

$$\underbrace{\left[\prod_{k=1}^{m-1}\overbrace{\left(\int_0^\infty S_k^j(a)M_k^j(a)m_k^j(a)da\right)}^{\Pr(k\to k+1)}\right]}_{\text{sessile stages}}\underbrace{\left(\int_0^\infty S_m^j(a)b^j(a)da\right)}_{\#\text{ larvae}}, \qquad (9)$$

and refer to this as the productivity of patch j. To determine the specific temperature and salinity dependent demographic rates at each farm we find the temperature and salinity that each sea louse particle experiences in the particle tracking simulation when initially released from a farm. We then use the average temperature and salinity of particles over all releases.

We also investigate the effect of varying temperature and salinity on the relative growth and persistence of the metapopulation. To do so, we must calculate survival and maturation rates for temperatures and salinities that farms may not experience in the period for which the FVCOM was run. To keep the variability of temperatures and salinities that exists between farms, we multiply the new temperature or salinity at which we want to evaluate persistence, by the ratio of the mean farm temperature or salinity divided by the mean total temperature or salinity experienced by all farms.

3   Results

In this section we present the next-generation matrix for sea lice populations on salmon farms in the Broughton Archipelago, the construction of which is detailed in section 2.4, and determine the relative patch contribution of each farm. We use this system to highlight to potential differences between the connectivity matrix, which only contains information surrounding the probability of dispersal from one farm to another, and the next-generation matrix, which combines dispersal between farms

Table 1: The maturation, survival, and birth functions used to create the next-generation matrix for sea louse populations on salmon farms in the Broughton Archipelago. The sea louse life cycle is simplified to three attached stages: chalimus, pre-adult and adult. For the maturation rate out of stage i, $i \in (c, p)$, where c refers to the chalimus stage and p refers to the pre-adult stage.

| Description	Function	Units	Reference
Survival probability of all stages	$S(a) = \exp(-(24a)^{(1/0.512)} \times \exp(-(1/0.512)(4.12 + 0.124S)))$		[58]
			[59]
Maturation rate out of stage i	$m_i(a) = \log(2)\delta_{si}((\delta_{mi}(10/T)^{\delta_{pi}} - \delta_{si})a^{\delta_{si}-1}$	$\frac{1}{\text{day}}$	[55]
	$\delta_{mc} = 18.869,\ \delta_{sc} = 7.945,\ \delta_{pc} = 1.305,\ \delta_{mp} = 10.742,$		
	$\delta_{sp} = 1.643,\ \delta_{pp} = 0.866$		
Egg batch size	$B = 2\exp(5.6 - 0.43\log(T/10) - 0.78(\log(T/10))^2)$	eggs	[60]
Egg development time	$D = (41.98/(T - 10 + 41.98 * 0.338))^2$	days	[60]
Egg viability (salinity)	$V_{sal} = \exp(-11.92 + 0.538S)/(1 + \exp(-11.92 + 0.538S))$		[60]
Egg viability (temperature)	$V_{temp} = \exp(-1.765 + 0.494T)/(1 + \exp(-1.765 + 0.494T))$		[60]
Birth rate	$b(a) = B \times V_{sal} \times V_{temp}/D$	$\frac{\text{larvae}}{\text{day}}$	[60]
Arrival rate	$\beta = 100$	$\frac{1}{\text{day}}$	
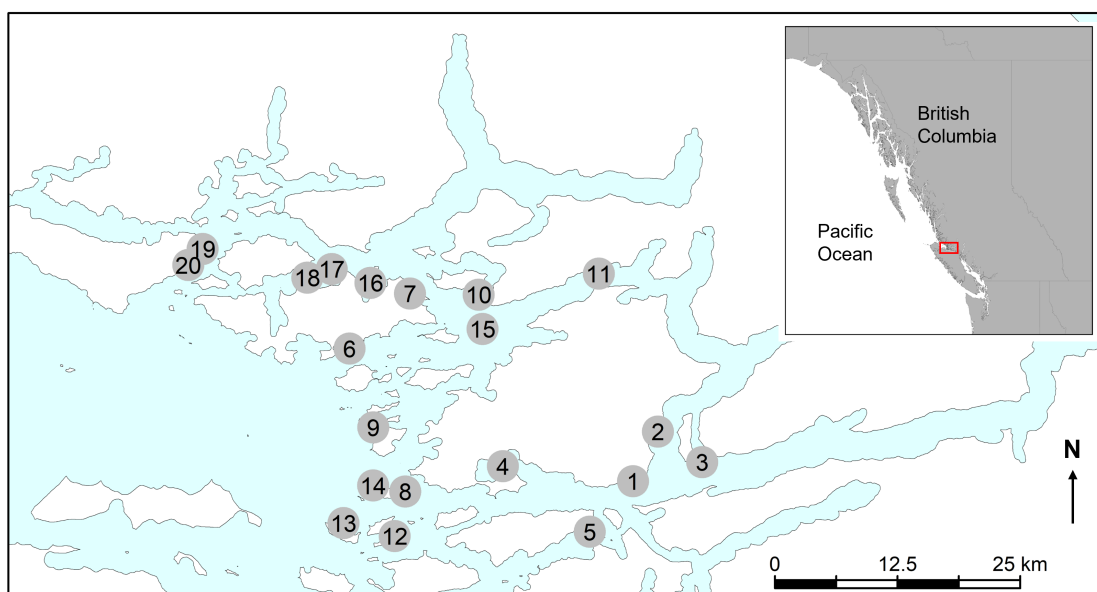
Figure 2: Map of the 20 historically active salmon farms in the Broughton Archipelago, for which the next-generation matrix is calculated.

and local productivity of sea lice on a salmon farm. We then demonstrate how the next-generation matrix can be used to investigate the effect of changing demographic rates on growth and persistence in this system. Finally, in the context of salmon farm removal from the Broughton Archipelago, we investigate how the removal of habitat patches affects patch contribution and persistence in this sea louse metapopulation.

The next-generation matrix for sea lice populations on salmon farms in the Broughton Archipelago is shown in Figure 3a). The patch contribution of each salmon farm is given by R_j, the jth column sum, which is presented at the bottom of each column. We also present the row sums to the left each row, to identify which farms are receiving the most sea lice from other farms in the region. Farm 2 is the largest contributor of sea lice to the metapopulation, followed by farms 17 and 18, whereas farm 12 has the lowest contribution. The farms receiving the most sea lice, in declining order, are farms 3, 16, and 7.

To better understand the details and construction of the next-generation matrix, we also present the connectivity matrix for this system in Figure 4 and the productivity (total number of new larvae produced from one attached chalimus louse) of each farm in Table 2. The (i, j)th entry of the connectivity matrix, c_{ij}, is the infectious density over farm i of lice leaving farm j and the (i, j)th entry of the next-generation matrix is constructed by multiplying the productivity of farm j (equation 2.4.2) $p^{ij} = \beta \times 0.01 \times c_{ij}$, the probability that a larvae leaving farm j attaches on farm i (8). The farms with the largest column sums of the connectivity matrix are, in declining order, farms 18, 2, and 17. However, farm 2 has a higher productivity than farms 17 or 18, and when the productivity of farm 2 is multiplied by the connectivity then it becomes the largest source of sea lice in the region, as identified by the next-generation matrix.

A further look into the connectivity matrix and productivity table provides more insight into the underlying drivers of the contribution of each farm to the sea louse population. Many of the farms with low connectivity also have low productivity, which may be due to the fact that temperature and salinity affect on farm demographic rates as well as survival and maturation in the particle tracking simulation which underlies the connectivity matrix. However, there are certain farms, such as 11

Table 2: The number of new larvae produced on each farm by a single louse starting in the chalimus stage. The first row is the farm number and the second row is the number of larvae produced.

Farm number	1	2	3	4	5	6	7	8	9	10
Number of larvae produced	1355	1582	1462	239	244	612	1064	111	56	1060
Farm number	11	12	13	14	15	16	17	18	19	20
Number of larvae produced	1529	33	331	175	1255	1143	1483	1348	584	464

and 15, which have a comparably high productivity compared to their connectivity. These farms are located in favourable environments with respect to temperature and salinity, but low connectivity due to either distance from other farms or unfavourable currents prevents these farms from acting as larger sources of sea lice.

Here we also examine how temperature and salinity affect the overall growth and relative persistence of the sea louse metapopulation, as shown in Figure 5. The persistence of the metapopulation is determined by the basic reproduction number, R_0, which is calculated as the spectral radius of the next-generation matrix. As we do not have an accurate estimate for β, we examine the effect of temperature and salinity on the relative change in persistence or growth of the metapopulation, but refrain from commenting on the absolute growth, as measured by R_0. We can see that as both temperature and salinity increases, the overall growth of the metapopulation increases, and that salinity has a larger effect on metapopulation growth than temperature. What is also interesting, but cannot be seen from the figure, is that as salinity increases, the farm that receives the most lice switches from farm 3 to farm 7.

In light of the removal of salmon farms in the Broughton Archipelago we also create a next-generation matrix consisting only of the farms which will remain after 2023 subject to First Nations and government approval, shown in Figure 3b), and examine the differences between this matrix and the next-generation matrix with all farms. Most of the farms which are acting as the largest sources of sea lice in this region remain, with the exception of farm 11, the fifth largest source in the original network, which has now been removed. Farm 11, while not the largest source, did have the largest betweenness score based solely on the connectivity matrix [51], and thus may have been acting as a connecting farm between the two large clusters of source farms. However, since none of the other large source farms have been removed, the overall growth of the metapopulation has only decreased from $R_0 = 2.33$ (original next-generation matrix) to $R_0 = 2.25$. Again these numbers are calculated using a very rough estimate of the arrival rate onto farms, detailed in section 2.4.2, and thus it is their relative similarity that is important, rather than the absolute magnitude.

4   Discussion

In this paper we demonstrated how to use the next-generation matrix to calculate the contribution of each habitat patch to the metapopulation and measure the overall persistence of a metapopulation. We detailed the construction of the next-generation matrix under different model structures to demonstrate the breadth of the approach to several systems. We then constructed the next-generation matrix under an age dependent modeling framework for sea lice populations on salmon farms in the Broughton Archipelago to illustrate how this approach can be applied to a real system. We determined which salmon farms may be acting as the largest sources of sea lice in this region, how the metapopulation will change once certain farms are removed, and examined the effect of temperature and salinity on the relative growth and persistence of this metapopulation.

Next-generation matrices have been used extensively in epidemiology to study the spread of infectious diseases but have recently been introduced in ecology [24, 61, 26, 27] and evolutionary analysis [62]. One of the key benefits of using next-generation matrices in epidemiology is that the basic reproduction number, R_0, for a disease can be calculated as the spectral radius of the next-
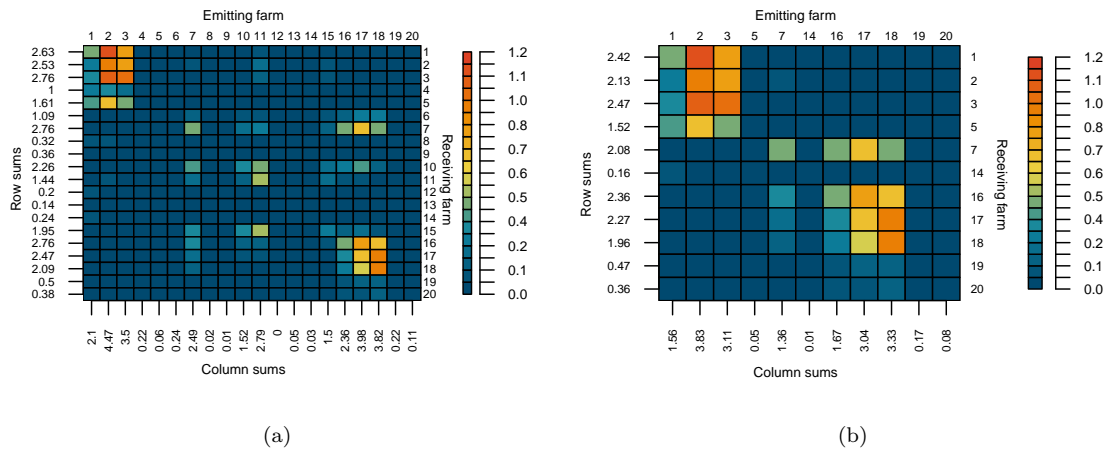
(a)   (b)

Figure 3: a) The next-generation matrix for the 20 historically active farms in the Broughton Archipelago, and b) the next-generation matrix containing only the farms remaining in the Broughton Archipelago after 2023, subject to First Nations and governmental approval [47]. The entries of the next-generations matrices, k_{ij} are the number of new chalimus stage lice produced on farm i from one initial chalimus on farm j. The column sums, R_j, are the total number of chalimus produced on all farms from an initial chalimus on farm j and are shown below each column. Likewise the row sums are the number of new chalimus received by each farm from all other farms and are shown on the left of each row. These numbers should be taken as relative, rather than absolute, as we do not have a very accurate estimate for the arrival rate of sea lice onto farms, β.
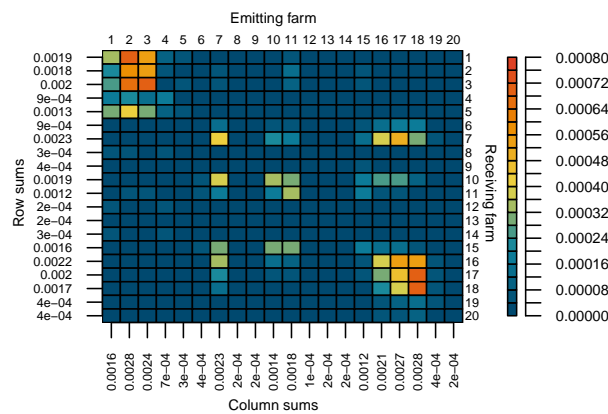

Figure 4: The connectivity matrix, C, for sea lice larvae dispersing between salmon farms, created by averaging all connectivity matrices from [51] over the simulation period. The (i, j)th entry is the infectious density of larvae $(1/\text{km}^2)$ over farm i that have left from farm j. Column and row sums are shown below and to the left of each column and row, respectively.
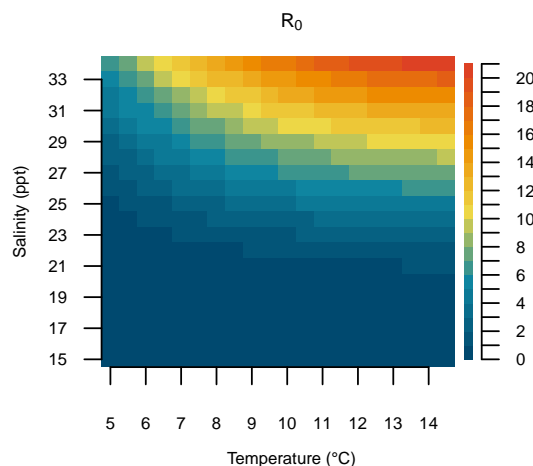
12

Figure 5: The effect of temperature and salinity on the overall growth or persistence of the original sea lice metapopulation of 20 farms, as described by the basic reproduction number, R_0. We do not have a good estimate for the arrival rate of sea lice onto farms, β, and so the R_0 values should only be interpreted relative to each other, rather than as absolute values.

generation matrix, which is often much simpler than calculating the eigenvalues of the full system to determine spread. In ecology, one main advantage of this approach is that the mathematical calculation of R_0 can be broken down into biologically relevant quantities, for example the contribution of different dispersal pathways to growth in a population [63] or the contribution of populations on different habitat patches [23]. While not novel, we hope next-generation matrices can be used more frequently as a simple and easily biologically interpretable method to measure the contribution of local habitat patches to a metapopulation and determine overall persistence.

We are also by no means the first to attempt to calculate the contribution of a local population, classify patches into sources and sinks, or attempt to measure the persistence of metapopulations. In the context of low densities Pulliam [11] defined a source as a habitat patch that would grow in the absence of immigration and emigration and a sink as a habitat patch that would decline in the absence of immigration and emigration. This is similar to only using the entries along the diagonal of the next-generation matrix to classify sources or sinks, except growth or decline was measured after one time step in a discrete time model, rather than one generation. However, as discussed at the end of section 2.2, it is possible to have a metapopulation composed only of sinks based on this definition ($k_{ii} < 1$ for all i), that persists.

Recognizing that dispersal between patches should also considered when classifying habitat patches as sources or sinks, both Runge [12] and Figueira and Crowder [13] defined new metrics to classify habitat patches that track the contribution of adults on a patch in one time step to the total population on all patches in the next timestep. These metrics are similar to our patch contribution metric R_j, except they measure the contribution over one time step rather than one generation, similar to using the dominant eigenvalue of the projection matrix A to determine the stability of the discrete system $n(t+1) = An(t)$, rather than the spectral radius of the next-generation matrix K. However, the calculations can become complicated if the population is stage structured and A is large (Appendix A, [12]) and the metrics do not easily generalize to systems described by ordinary differential equations.

There are other measures of persistence in metapopulations which do track the number of new individuals contributed to the metapopulation after one generation from an initial individual on one

patch, though they use different starting stages for the initial and new individuals. Krkošek and Lewis [24] define a next-generation operator for general heterogeneous populations which tracks the number of new adults produced in the population from one initial adult after one generation. If b_j is the reproductive output on patch j, p_{ij} is the probability of larvae dispersing between patch j successfully arrives on patch i, and a_i is the survival to adulthood on patch i, then the contribution of patch j to patch i according to Krkošek and Lewis [24] can be calculated as $b_j p_{ij} a_i$. This patch contribution metric cannot be calculated directly from the next-generation matrices shown in this paper unless there is only one stage. Burgess et al. [10], following Hastings and Botsford [15], track the number of new larvae on all patches produced by an initial larvae. The entries of their 'connectivity matrix' (similar to our next-generation matrix), c_{ij}, are given by $p_{ij} a_i b_i$. The entries of their connectivity matrix would be the entries of the next-generation matrix if it was calculated from a model with an explicit larval stage. However under this construction the contribution of patch j, as calculated by the jth column sum of their connectivity matrix, is primarily a function of the demography of the patches on which larvae dispersing from j settle, rather than the local patch demography of patch j itself, as it is when using the next-generation matrix as we have formulated it in this chapter.

The contribution of each habitat patch to the metapopulation will depend on the stage at which the generational output is measured, but the persistence of the metapopulation is equivalent under all of these frameworks. This is because there is only one component of the life cycle where movement can occur between patches (larval stage) and at all other stages individuals remain on a patch. Let P be a matrix with entries p_{ij}, B be a diagonal matrix with entries b_j and A be a diagonal matrix with entries a_j, where p_{ij}, b_j and a_j are the same as the preceding paragraph. If we measure generational output starting at the first attached stage, as we do in this chapter, the next-generation matrix can be written as PAB, if we measure generational output starting at the larval stage according to Burgess et al. [10], the matrix can be written as ABP, and if we measure generational output starting at the adult stage according to Krkošek and Lewis [24] then the matrix can be written as APB. This is because when P is multiplied by a diagonal matrix on the right, the entries of the diagonal matrix multiply each column of P and when P is multiplied on the left, the entries multiply each row. Now for any two matrices X and Y, XY and YX have the same eigenvalues, and because matrix multiplication is associative each of the matrices PAB, ABP, and APB all have the same eigenvalues as well, and therefore also the same spectral radius. Therefore in any of the formulations the metapopulation will only persist if $\rho(PAB) = \rho(ABP) = \rho(APB) > 1$.

While the metapopulation persistence criteria is equivalent to other formulations [24, 10, 15] we believe using the next-generation matrix provides several advantages. First, the framework is the same for discrete time, continuous time, and age structured systems of equations. Second, it is easy to convert between the full system of equations and the next-generation matrix, thus if a system of equations has already been parameterized for a given region, it is easy to calculate the contribution of each habitat patch using the next-generation matrix. Third, there is natural extension of the next-generation matrix to systems where multiple stages can reproduce, or when adults can migrate between patches. In discrete time and continuous time the composition of the matrix in terms of the fecundity and transition matrix still holds, only the formulas given in terms of the survival, maturation, and birth rates no longer hold. The next-generation matrix can also be extended as an operator to integro-difference equations [24, 26] if space is continuous, rather than divided into discrete habitat patches. Lastly, in continuous time systems it is possible to connect the patch contribution computed from the next-generation matrix, R_j, and the transient dynamics of the metapopulation (Harrington et al, in press).

In our application of the next-generation matrix to sea lice populations on salmon farms in the Broughton Archipelago, we believe there are two main sources of uncertainty when calculating the entries of the next-generation matrix: uncertainty in the arrival rate, β, and uncertainty in the number of larval sea lice that remain inside the salmon farm in the particle tracking model. In lab studies, the proportion of infective copepodid lice that attach to salmon varies enormously, from as

low as 0.5% to as high as $> 80\%$, with recent estimates between $12-56\%$ depending on temperature [57]. Moreover, the time during which copepodids are given to attach in these experiments also varies, from a few seconds, to 8 hours. If we assume an exponential waiting time model, $P = 1 - e^{-\beta t}$, where P is the proportion of copepodids that have attached to salmon, then these lab studies give estimates of β between $3/\text{day}$ and $47/\text{day}$, which are lower than the estimate of $100/\text{day}$ which we use in this paper. Any estimates less than $44/\text{day}$ would result in a basic reproduction number, $R_0 < 1$, which would mean that the overall sea louse population is not increasing.

However, there is abundant evidence that in the Broughton Archipelago sea lice populations on many salmon farms will grow exponentially in the absence of treatment by farm managers [64, 65, 66, 52, 67]. This discrepancy may be due to the fact that while the hydrodynamic and particle tracking models give good estimates of sea louse dispersal between farms [52], they likely underestimate the number of sea louse larvae that remain inside a salmon farm. In the hydrodynamic model the location of the farm does not alter the local hydrodynamics that occur within the farm, but in reality the the presence of salmon swimming within the net pens may alter the local hydrodynamics sufficiently that some proportion of larvae remain within the farm. To correct for this, it would be possible to add a diagonal matrix to the next-generation matrix presented in this paper, where each entry on the diagonal is the productivity of farm j, multiplied by the number of larvae which remain within and successfully attach to salmon on the farm. Because the most productive farms are also the most highly connected in the Broughton, this would likely not change the relative contribution of each farm, but would increase the absolute contribution of each farm as well as R_0.

In addition to the sources of uncertainty in the next-generation matrix, there are some other technical aspects which should be considered when applying the next-generation matrix to calculate patch contribution of metapopulation persistence in specific systems. In order to use the next-generation matrix to calculate patch contribution or persistence we are assuming that our system is autonomous, and that the demographic rates do not change with time. In reality for most systems, including sea lice on salmon farms, environmental variables will fluctuate over time, potentially changing the demographic rates of the population [68]. In our case temperature and salinity change over the course of the spring, but we calculate the next-generation matrix using the mean temperature and salinity that sea lice experience during the particle tracking simulation window. Therefore the entries in our next-generation matrix may be slightly different than the true number of newly attached lice produced on other farms from one initially attached louse, depending on the exact time that the louse began its lifecyle during the spring. For temporally oscillating systems it is possible to correct for this difference, though the entries of the next-generation matrix no longer have a simple form and must be computed computationally [59].

Over time periods where temperature and salinity are relatively constant, we can use the next-generation matrix to examine the overall effect of environmental variables on the growth of the metapopulation. The demographic rates at each stage depend explicitly on temperature and salinity but alone, or in a full system of equations, it can be difficult to examine the overall effect of changing environmental variables. However, the basic reproduction number R_0, calculated from the spectral radius of the next-generation matrix, provides a useful metric of the overall effect. We can infer how the growth of the metapopulation may change among seasons, or as the ocean warms. For sea lice on salmon farms in the Broughton Archipelago the effect that of temperature and salinity on R_0 is very similar to previous results found for a single farm [59]. With updated temperature and salinity at each farm, we could calculate the change in growth over years in the springtime, which may help explain the recent sea louse outbreaks during warm years in the Broughton Archipelago [66]. When examining the effect of temperature and salinity on R_0 we do not rerun the hydrodynamic and particle tracking models to recreate the connectivity matrices under new temperature and salinity scenarios, as this is very computationally intensive, but we expect connectivity to increase as temperature and salinity increases due to higher survival of sea lice and faster maturation. However, we believe it would be valuable to rerun hydrodynamic models under different projected ocean scenarios, to investigate the precise changes in connectivity that may occur.

15

Specific to management of sea lice populations on salmon farms in the Broughton Archipelago, there are several insights to be gained from our results. The first is that the farms that in the most productive environments are also the most highly connected, and thus become the largest contributors of sea lice to this sea louse metapopulation. They occur in two main clusters (shown in Figure 3) and both clusters of farms will remain in the Broughton Archipelago in the current removal plan, subject to First Nations and governmental approval [47]. It should be noted that these farms may not necessarily be producing the most number of lice compared to other farms in the region at a given time if their louse population is currently lower than other farms, rather they are the farms that have the largest potential to contribute to spread when sea louse numbers are even across farms. However, due to the highly connected nature of these clusters, coordinated treatment between farms in the clusters or all farms in the region could reduce the number of treatments required and number of sea lice produced on all farms [69]. An interesting avenue of future research would be to connect the productivity of the remaining farms in the Broughton Archipelago with the Kernel Density Estimates of sea louse dispersal from Cantrell et al. [51] to measure the exposure of migrating wild salmon to sea louse infection from these farms.

5   Acknowledgements

References

[1] Richard Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.*, 15(3):237–250, 1969.

[2] Ilkka Hanski. Metapolulation dynamics. *Nature*, 396:41–49.

[3] Jacob P. Kritzer and Peter F. Sale. Metapopulation ecology in the sea: From Levins' model to marine ecology and fisheries science. *Fish Fish.*, 5(2):131–140.

[4] Robert K. Cowen and Su Sponaugle. Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.*, 1:443–466, 2009.

[5] James R. Watson, David A. Siegel, Bruce E. Kendall, Satoshi Mitarai, Andrew Rassweiller, and Steven D. Gaines. Identifying critical regions in small-world marine metapopulations. *Proc. Natl. Acad. Sci. U. S. A.*, 108(43):907–913.

[6] Christopher Costello, Andrew Rassweiler, David Siegel, Giulio De Leo, Fiorenza Micheli, Andrew Rosenberg, and Steven D Gaines. The value of spatial information in MPA network design. *PNAS*, 107(43):18294–18299.

[7] L. B. Crowder, S. J. Lyman, W. F. Figueira, and J. Priddy. Source-sink population dynamics and the problem of siting marine reserves. *Bull. Mar. Sci.*, 66(3):799–820.

[8] M Bode, L Bode, and Paul R. Armsworth. Larval dispersal reveals regional sources and sinks in the Great Barrier Reef. *Mar. Ecol. Prog. Ser. Ser.*, 308:1–9.

[9] Will F. Figueira. Connectivity or demography: Defining sources and sinks in coral reef fish metapopulations. *Ecol. Modell.*, 220(8):1126–1137.

[10] Scott C. Burgess, Kerry J. Nickols, Chris D. Griesemer, Lewis A.K. Barnett, Allison G. Dedrick, Erin V. Satterthwaite, Lauren Yamane, Steven G. Morgan, J. Wilson White, and Louis W. Botsford. Beyond connectivity: How empirical methods can quantify population persistence to improve marine protected-area design. *Ecol. Appl.*, 24(2):257–270.

[11] H. R. Pulliam. Sources, sinks and population regulation. *Am. Nat.*, 132(5):652–661, 1988.

[12] Jonathan P. Runge, Michael C. Runge, and James D. Nichols. The role of local populations within a landscape context: Defining and classifying sources and sinks. *Am. Nat.*, 167(6):925–938.

[13] Will F. Figueira and Larry B. Crowder. Defining patch contribution in source-sink metapopulations: the importance of including dispersal and its relevance to marine systems. *Popul. Ecol.*, 48(3):215–224, 2006.

[14] J Wilson White, Mark H Carr, Jennifer E Caselle, Libe Washburn, C Brock Woodson, Stephen R Palumbi, Peter M Carlson, Robert R Warner, Bruce A Menge, John A Barth, Carol A Blanchette, Peter T Raimondi, and Kristen Milligan. Connectivity, Dispersal, and Recruitment. *Oceanography*, 32(3):50–59.

[15] Alan Hastings and Louis W. Botsford. Persistence of spatial populations depends on returning home. *Proc. Natl. Acad. Sci. U. S. A.*, 103(15):6067–6072.

[16] B. J. Puckett and D. B. Eggleston. Metapopulation dynamics guide marine reserve design: Importance of connectivity, demographics, and stock enhancement. *Ecosphere*, 7(6).

[17] Henry S Carson, Geoffrey S Cook, Paola C López-duarte, Lisa A Levin, Henry S Carson, Geoffrey S Cook, Paola C Lopez-duarte, and Lisa A Levin. Evaluating the importance of demographic connectivity in a marine metapopulation. *Ecology*, 92(10):1972–1984.

[18] Allison G. Dedrick, Katrina A. Catalano, Michelle R. Stuart, J. Wilson White, Humberto R. Montes, and Malin L. Pinsky. Persistence of a reef fish metapopulation via network connectivity: theory and data. *Ecol. Lett.*, 24(6):1121–1132.

[19] Seth J. Theuerkauf, Brandon J. Puckett, and David B. Eggleston. Metapopulation dynamics of oysters: sources, sinks, and implications for conservation and restoration. *Ecosphere*, 12(7).

[20] J. Wilson White, Louis W. Botsford, Alan Hastings, and John L. Largier. Population persistence in marine reserve networks: Incorporating spatial heterogeneities in larval dispersal. *Mar. Ecol. Prog. Ser.*, 398:49–67.

[21] Lysel Garavelli, J. Wilson White, Iliana Chollett, and Laurent Marcel Chérubin. Population models reveal unexpected patterns of local persistence despite widespread larval dispersal in a highly exploited species. *Conserv. Lett.*, 11(5).

[22] P. van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. 180(1-2):29–48.

[23] Peter D. Harrington and Mark A. Lewis. A next-generation approach to calculate source–sink dynamics in marine metapopulations. *Bull. Math. Biol.*, 82(1):1–44, 2020.

[24] Martin Krkošek and Mark A. Lewis. An R_0 theory for source-sink dynamics with application to Dreissena competition. *Theor. Ecol.*, 3(1):25–43, 2010.

[25] Qihua Huang, Yu Jin, and Mark A. Lewis. R_0 analysis of a benthic-drift model for a stream population. *SIAM J. Appl. Dyn. Syst.*, 15(1):287–321, 2016.

[26] H. W. Mckenzie, Y. Jin, J. Jacobsen, and M. A. Lewis. R_0 analysis of a spatiotemporal model for a stream population. *SIAM J. Appl. Dyn. Syst.*, 11(2):567–596, 2012.

[27] Mark A. Lewis, Zhisheng Shuai, and P. van den Driessche. A general theory for target reproduction numbers with applications to ecology and epidemiology. *J. Math. Biol.*, 78(7):2317–2339.

[28] T. Hillen, B. Greese, J. Martin, and G. de Vries. Birth-jump processes and application to forest fire spotting. *J. Biol. Dyn.*, 9:104–127, 2015.

[29] Brian C. Husband and Spencer C. H. Barrett. A metapopulation perspective in plant population biology. *J. Ecol.*, 84(3):461–469, 1996.

[30] David W. Williams and Andrew M. Liebhold. Spatial synchrony of spruce budworm outbreaks in eastern North America. *Ecology*, 81(10):2753–2766, 2000.

[31] L Safranyik and A L Carroll. *The biology and epidemiology of the mountain pine beetle in lodgepole pine forests.*, pages 3–66. Canadian Forest Service, Victoria, Canada, 2007.

[32] Mark J. Costello. Ecology of sea lice parasitic on farmed and wild fish. *Trends Parasitol.*, 22(10):475–483.

[33] D.J. Stucchi, M. Guo, M.G.G. Foreman, P. Czajko, M. Galbraith, D.L. Mackas, and P A Gillibrand. Modeling sea lice production and concentrations in the Broughton Archipelago, British Columbia. In R Jones, S., Beamish, editor, *Salmon lice: an integrated approach to understanding parasite abundance and distribution*, pages 117–150. Wiley, 1 edition.

[34] Stephanie J. Peacock, Martin Krkošek, Andrew W. Bateman, and Mark A. Lewis. Estimation of spatiotemporal transmission dynamics and analysis of management scenarios for sea lice of farmed and wild salmon. *Can. J. Fish. Aquat. Sci.*, 77(1):55–68.

[35] Martin Krkošek, Mark A. Lewis, Alexandra Morton, L. Neil Frazer, and John P. Volpe. Epizootics of wild fish induced by farm fish. *Proc. Natl. Acad. Sci. U. S. A.*, 103(42):15506–15510.

[36] Jay Abolofia, Frank Asche, and James E. Wilen. The cost of lice: Quantifying the impacts of parasitic sea lice on farmed salmon. *Mar Resour Econ*, 32(3):329–349.

[37] Martin Krkošek, Jennifer S. Ford, Alexandra Morton, Subhash Lele, Ransom A. Myers, and Mark A. Lewis. Declining Wild Salmon Populations in Relation to Parasites from Farm Salmon. *Science.*, 318(2):1772–1775.

[38] Martin Krkošek, Mark A Lewis, and John P Volpe. Transmission dynamics of parasitic sea lice from farm to wild salmon. *Proc. Biol. Sci.*, 272(1564):689–696.

[39] Kenneth M. Brooks. The effects of water temperature, salinity, and currents on the survival and distribution of the infective copepodid stage of sea lice (Lepeophtheirus salmonis) originating on atlantic salmon farms in the Broughton Archipelago of British Columbia, Canada. *Rev. Fish. Sci.*, 13(3):177–204.

[40] Martin Krkošek, Mark A. Lewis, John P. Volpe, and Alexandra Morton. Fish farms and sea lice infestations of wild juvenile salmon in the Broughton Archipelago - A Rebuttal to Brooks (2005). *Rev. Fish. Sci.*, 14(1-2):1–11.

[41] Gary D Marty, Sonja M Saksida, and Terrance J Quinn. Relationship of farm salmon, sea lice, and wild salmon populations. *Proc. Natl. Acad. Sci. U. S. A.*, 107(52):22599–22604.

[42] Martin Krkošek, Brendan M. Connors, Alexandra Morton, Mark A. Lewis, Lawrence M. Dill, and Ray Hilborn. Effects of parasites from salmon farms on productivity of wild salmon. *Proc. Natl. Acad. Sci. U. S. A.*, 108(35):14700–14704.

[43] Kenneth M. Brooks and Dario J. Stucchi. The effects of water temperature, salinity and currents on the survival and distribution of the infective copepodid stage of the salmon louse (Lepeophtheirus salmonis) originating on Atlantic salmon farms in the Broughton Archipelago of British Columbia, . *Rev. Fish. Sci.*, 14(1-2):13–23.

[44] Martin Krkošek, Jennifer S. Ford, Alexandra Morton, Subhash Lele, and Mark A. Lewis. Response to comment on "Declining wild salmon populations in relation to parasites from farm salmon". *Science*, 322(5909):1–4.

[45] Brian E Riddell, Richard J Beamish, Laura J Richards, and John R Candy. Comment on "Declining Wild Salmon Populations in Relation to Parasites from Farm Salmon". *Science*, 322(5909):1790 LP – 1790.

[46] M. G. G. Foreman, P. C. Chandler, D. J. Stucchi, K A Garver, M Guo, J Morrison, and D Tuele. The ability of hydrodynamic models to inform decisions on the siting and management of aquaculture facilities in British Columbia. Technical report, Fisheries and Oceans Canada.

[47] Lorne Brownsey and Robert Chamberlain. Collaborative solutions for finfish aquaculture farms in the Broughton Area, steering committee recommendations.

[48] J. M. Cushing and Zhou Yicang. The net reproductive value and stability in matrix population models. *Nat. Resour. Model.*, 8(4):297–333.

[49] Chi Kwong Li and Hans Schneider. Applications of Perron-Frobenius theory to population dynamics. *J. Math. Biol.*, 44(5):450–462.

[50] Zhilan Feng and Horst R. Thieme. Endemic Models with Arbitrarily Distributed Periods of Infection I: Fundamental Properties of the Model. *SIAM J. Appl. Math.*, 61(3):803–833.

[51] Danielle L. Cantrell, Erin E. Rees, Raphael Vanderstichel, Jon Grant, Ramón Filgueira, and Crawford W. Revie. The Use of Kernel Density Estimation With a Bio-Physical Model Provides a Method to Quantify Connectivity Among Salmon Farms: Spatial Planning and Management With Epidemiological Relevance. *Front. Vet. Sci.*, 5(269):1–14.

[52] Danielle Cantrell, Raphael Vanderstichel, Ramón Filgueira, Jon Grant, and Crawford W. Revie. Validation of a sea lice dispersal model: principles from ecological agent-based models applied to aquatic epidemiology. *Aquac. Environ. Interact.*, 13:65–79.

[53] Audun Stien, Pål Arne Bjørn, Peter Andreas Heuch, and David A. Elston. Population dynamics of salmon lice Lepeophtheirus salmonis on Atlantic salmon and sea trout. *Mar. Ecol. Prog. Ser.*, 290:263–275.

[54] Crawford W. Revie, C. Robbins, G. Gettinby, L. Kelly, and J W Treasurer. A mathematical model of the growth of sea lice, Lepeophtheirus salmonis, populations on farmed Atlantic salmon, Salmo salar L., in Scotland and its use in the assessment of treatment strategies. *J. Fish Dis.*, 28(10):603–614.

[55] M. Aldrin, R. B. Huseby, A. Stien, R. N. Grøntvedt, H. Viljugrein, and P. A. Jansen. A stage-structured Bayesian hierarchical model for salmon lice populations at individual salmon farms - Estimated from multiple farm data sets. *Ecol. Modell.*, 359:333–348.

[56] M. G. G. Foreman, P. Czajko, D. J. Stucchi, and M. Guo. A finite volume model simulation for the Broughton Archipelago, Canada. *Ocean Model.*, 30(1):29–47.

[57] Rasmus Skern-Mauritzen, Nini H. Sissener, Anne D. Sandvik, Sonnich Meier, Pål N. Sævik, Morten D. Skogen, Tone Vågseth, Sussie Dalvin, Mette Skern-Mauritzen, and Samantha Bui. Parasite development affect dispersal dynamics; infectivity, activity and energetic status in cohorts of salmon louse copepodids. *J. Exp. Mar. Bio. Ecol.*, 530-531:151429.

[58] B. M. Connors, E. Juarez-Colunga, and L. M. Dill. Effects of varying salinities on Lepeoph-theirus salmonis survival on juvenile pink and chum salmon. *J. Fish Biol.*, 72(7):1825–1830.

[59] Matthew A. Rittenhouse, Crawford W. Revie, and Amy Hurford. A model for sea lice (Lepeophtheirus salmonis) dynamics in a seasonally changing environment. *Epidemics*, 16:8–16.

[60] LC Stige, KO Helgesen, H Viljugrein, and L Qviller. A statistical mechanistic approach including temperature and salinity effects to improve salmon lice modelling of infestation pressure. *Aquac. Environ. Interact.*, 13:339–361.

[61] Qihua Huang and Mark A. Lewis. Homing fidelity and reproductive rate for migratory populations. *Theor. Ecol.*, 8(2):187–205, 2015.

[62] Amy Hurford, Daniel Cownden, and Troy Day. Next-generation tools for evolutionary invasion analyses. *J. R. Soc. Interface*, 7(45):561–571.

[63] T. De-Camino-Beck and M. A. Lewis. A new method for calculating net reproductive rate from graph reduction with applications to the control of invasive species. *Bull. Math. Biol.*, 69(4):1341–1354.

[64] Luke A. Rogers, Stephanie J. Peacock, Peter McKenzie, Sharon DeDominicis, Simon R.M. Jones, Peter Chandler, Michael G.G. Foreman, Crawford W. Revie, and Martin Krkošek. Modeling Parasite Dynamics on Farmed Salmon for Precautionary Conservation Management of Wild Salmon. *PLoS One*, 8(4).

[65] Sean C. Godwin, Martin Krkošek, John D. Reynolds, and Andrew W. Bateman. Bias in self-reported parasite data from the salmon farming industry. *Ecol. Appl.*, 31(1):1–10.

[66] Andrew W. Bateman, Stephanie J. Peacock, Brendan Connors, Zephyr Polk, Dana Berg, Martin Krkošek, and Alexandra Morton. Recent failure to control sea louse outbreaks on salmon in the Broughton Archipelago, British Columbia. *Can. J. Fish. Aquat. Sci.*, 73(8):1164–1172.

[67] Martin Krkošek, Andrew W. Bateman, Stan Proboszcz, and Craig Orr. Dynamics of outbreak and control of salmon lice on two salmon farms in the Broughton Archipelago, British Columbia. *Aquac. Environ. Interact.*, 1(2):137–146.

[68] Danielle Cantrell, Ramón Filgueira, Crawford W. Revie, Erin E. Rees, Raphael Vanderstichel, Ming Guo, Michael G.G. Foreman, Di Wan, and Jon Grant. The relevance of larval biology on spatiotemporal patterns of pathogen connectivity among open-marine salmon farms. *Can. J. Fish. Aquat. Sci.*, 77(3):505–519.

[69] Stephanie J Peacock, Andrew W Bateman, Martin Krkošek, and Mark A Lewis. The dynamics of coupled populations subject to control. *Theor. Ecol.*, pages 1–16.

[70] Abraham Berman and Robert J Plemmons. *Nonnegative matrices in the mathematical sciences.* Society for Industrial and Applied Mathematics, 1994.

[71] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface*, 7(47):873–885.

A Calculating the next-generation matrix for differential equation models

Here we briefly formalize the calculation of the next-generation matrix for differential equation models. Again, we are considering a model for a species with m sessile stages on l patches with a larval stage that can disperse between patches and all other stages are sedentary and confined a patch. Let n_k^j be the population in sessile stage k on patch j, then the population dynamics at low population densities can be described by

$$\frac{d\mathbf{n}}{dt} = A\mathbf{n} \tag{10}$$

where $\mathbf{n}$ is a population vector of size $m \times l$ describing the population size on all patches at all stages and A is a matrix describing the population dynamics at low population densities. We model the population size at low densities because we are interested in identifying which patches are acting as sources and supporting the population at low densities and which are acting as sinks. Therefore equation 10 should be thought of as the linearization about the zero equilibrium of a potentially more complex model that may include density dependence. The elements of A contain the maturation and death rates of sedentary stages confined to patches as well as the rates of larval dispersal and attachment between patches.

The next-generation matrix for this system can be calculated by first decomposing $A = F - V$ where F is a non-negative matrix with entries that describe the rates of larval birth and probability of dispersal between patches and attachment as the first sedentary phase, and V is a non-singular M matrix [70] with entries that describe the maturation rates between stages and death rates in a stage [22]. Because V is a non-singular M matrix, V^{-1} is non-negative. Following van den Driessche and Watmough [22] with notation from Diekmann et al. [71], the next-generation matrix with large domain, K_L, can then by calculated as

$$K_L = FV^{-1}. \tag{11}$$

The elements of K_L contain the number of new individuals in each stage and patch produced by one initial new individual in each stages and patch. However, since new individuals are only produced in stage 1, K_L will only have l non-zero rows. If equation 10 is arranged so that the populations of stage 1 individuals are in the first l rows, then K_L will have a $l \times l$ submatrix in the upper left-hand corner [71]. This submatrix is the next-generation matrix K, where the elements of K, k_{ij}, give the number of new (stage 1) individuals produced on patch i from one initial stage 1 individual on patch j. In equation 10, if the maturation rates from stage k to $k+1$ on patch j are m_k^j, the death rates in stage k are d_k^j, the birth rate in the last stage on patch j is b^j, and the probability of successful attachment on patch i as a larvae leaving patch j is p^{ij}, then the entries of K can be written as:

$$k_{ij} = \underbrace{\left[\prod_{k=1}^{m-1} \overbrace{\left(\frac{m_k^j}{m_k^j + d_k^j}\right)}^{\text{Pr(stage } k \to k+1)}\right]}_{\text{survival through sessile stages}} \times \underbrace{\left(\frac{b^j}{d_m^j}\right)}_{\text{\# larvae produced}} \times \overbrace{p^{ij}}^{\text{Pr(patch } j \text{ to patch } i)} . \tag{12}$$

Example of the next-generation matrix in for a two patch metapopulation

Here we illustrate the construction of the next-generation matrix using a two patch example for a species which has two sessile stages. The population of sessile stage k on path j is given by n_k^j and the maturation rates, death rates, birth rate, probability of dispersal are defined as in the preceding paragraph. The dynamics of the system can be described by

$$
\begin{bmatrix} n_1^1 \\ n_1^2 \\ n_2^1 \\ n_2^2 \end{bmatrix}' = \overbrace{\begin{bmatrix} -d_1^1 - m_1^1 & 0 & b^1 p^{11} & b^2 f^{12} \\ 0 & -d_1^2 - m_1^2 & b^1 p^{21} & b^2 f^{22} \\ m_1^1 & 0 & -d_2^1 & 0 \\ 0 & m_1^2 & 0 & -d_2^2 \end{bmatrix}}^{A} \begin{bmatrix} n_1^1 \\ n_1^2 \\ n_2^1 \\ n_2^2 \end{bmatrix},
$$

where we have arranged the system of equations so that individuals in the first stage appear in the top rows. We decompose $A = F - V$ where F contains all the births of new stage 1 individuals and V contains all the remaining transitions, so that F and V are given by

$$
F = \begin{bmatrix} 0 & 0 & b^1 p^{11} & b^2 p^{12} \\ 0 & 0 & b^1 p^{21} & b^2 p^{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, V = \begin{bmatrix} d_1^1 + m_1^1 & 0 & 0 & 0 \\ 0 & d_1^2 + m_1^2 & 0 & 0 \\ -m_1^1 & 0 & d_2^1 & 0 \\ 0 & -m_1^2 & 0 & d_2^2 \end{bmatrix}.
$$

The next-generation matrix with large domain, K_L, can then be calculated as $K_L = FV^{-1}$, where

$$
K_L = \begin{bmatrix} \frac{m_1^1}{m_1^1+d_1^1} \frac{b^1}{d_2^1} p^{11} & \frac{m_1^2}{m_1^2+d_1^2} \frac{b^2}{d_2^2} p^{12} & \frac{b^1}{d_2^1} p^{11} & \frac{b^2}{d_2^2} p^{12} \\ \frac{m_1^1}{m_1^1+d_1^1} \frac{b^1}{d_2^1} p^{21} & \frac{m_1^2}{m_1^2+d_1^2} \frac{b^2}{d_2^2} p^{22} & \frac{b^1}{d_2^1} p^{21} & \frac{b^2}{d_2^2} p^{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
$$

and then the next-generation matrix, K, will be given by the 2×2 submatrix in the upper left-hand corner of K_L, so

$$
K = \begin{bmatrix} \frac{m_1^1}{m_1^1+d_1^1} \frac{b^1}{d_2^1} p^{11} & \frac{m_1^2}{m_1^2+d_1^2} \frac{b^2}{d_2^2} p^{12} \\ \frac{m_1^1}{m_1^1+d_1^1} \frac{b^1}{d_2^1} p^{21} & \frac{m_1^2}{m_1^2+d_1^2} \frac{b^2}{d_2^2} p^{22} \end{bmatrix}.
$$

The contribution of patch 1 is therefore

$$
R_1 = \underbrace{\frac{m_1^1}{m_1^1 + d_1^1} \frac{b^1}{d_2^1}}_{\text{productivity}} \underbrace{\left(p^{11} + p^{12}\right)}_{\text{connectivity}}
$$

and the contribution of patch 2 is

$$
R_2 = \frac{m_1^2}{m_1^2 + d_1^2} \frac{b^2}{d_2^2} \left(p^{12} + p^{22}\right).
$$

In order for a patch to be a source $(R_j > 1)$, both the on-patch productivity and the connectivity to other patches need to multiply to be larger than 1. If either is too low, i.e. if a patch is highly productive but not well connected, or highly connected but not productive, then the patch will be a sink $(R_j < 1)$. There are two ways for the entire metapopulation to persist. Either a single patch can persist on its own $(k_{ii} > 1)$, or there must be sufficient production and connectivity within patches such that $R_0 = \rho(K) > 1$.

B Calculating the next-generation matrix for discrete time models

The construction of the next-generation matrix for discrete time models is very similar to that of continuous time, though with some minor differences. The population dynamics can now be

described by

$$\mathbf{n}_{t+1} = A\mathbf{n}_t \tag{13}$$

where again $\mathbf{n}$ is a population vector of size $m \times l$ (number of stages $\times$ number of patches) describing the population size on all patches at all sessile stages, but now A is a population projection matrix with transition and survival probabilities as well as births.

To calculate the next-generation matrix for the discrete time system, we now decompose $A = F + T$, where again F is a matrix that contains all the birth rates and probabilities of successful of larval dispersal between patches, and now T is a matrix that contains all of the survival probabilities and transition probabilities between stages. Under this decomposition, the next-generation matrix with large domain, K_L, can by calculated by

$$K_L = F(I - T)^{-1}. \tag{14}$$

If Equation 13 is arranged so that the populations of all stage 1 individuals are in the top l rows, then again the next-generation matrix, K, will be the $l \times l$ submatrix in the upper left hand corner of K_L. In the discrete time framework if the probability of transitioning from stage k to $k+1$ is m_k^j, the probability of survival in stage k is s_k^j, the fecundity in patch j is b^j, and the probability that a larvae leaving patch j successfully arrives on patch i is p^{ij}, then the entries of K can be written as

$$k_{ij} = \underbrace{\left[\prod_{k=1}^{m-1} \overbrace{\left(\frac{m_k^j}{1 - s_k^j} \right)}^{\text{Pr(stage } k \to k+1)} \right]}_{\text{survival through sessile stages}} \times \underbrace{\left(\frac{b^j}{1 - s_m^j} \right)}_{\text{\# larvae produced}} \times \overbrace{p^{ij}}^{\text{Pr(patch } j \text{ to patch } i)}. \tag{15}$$

C Details of the model with age dependent demography in section 2.3

The full set of age density equations for the density of individuals in sessile stage k on patch i at time t and age a, $n_k^i(t, a)$, is

$$n_k^i(t, a) = \begin{cases} \underbrace{B_k^i(t - a)S_k^i(a)M_k^i(a)}_{\text{entered at } t-a, \text{ survived to } a} & t > a \\ \underbrace{n_{k,0}^i(a - t)\dfrac{S_k^i(a)M_k^i(a)}{S_k^i(a - t)M_k^i(a - t)}}_{\text{present at } a-t, \text{ survived to } a} & 0 < t < a \\ \underbrace{n_{k,0}^i(a)}_{\text{initial density}} & t = 0 \end{cases} \qquad k = 1, \ldots, m-1, \tag{16}$$

$$n_m^i(t, a) = \begin{cases} B_m^i(t - a)S_m^i(a) & t > a \\ n_{m,0}^i(a - t)\frac{S_m^i(a)}{S_m^i(a-t)} & 0 < t < a \\ n_{m,0}^i(a) & t = 0 \end{cases} \qquad k = m,$$

$$B_1^i(t) = \sum_{j=1}^l \int_0^\infty n_m^i(t, a)b^i(a)p^{ij}da \qquad k = 1$$

$$B_k^i(t) = \int_0^\infty n_{k-1}^i(t, a)m_{k-1}^i(a)da \qquad k = 2, \ldots, m,$$

where $S_k^j(a)$ is the probability of survival in stage k on patch j at stage-age a, $m_k^j(a)$ is the maturation rate from stage k to $k+1$, $b^j(a)$ is the birth rate of larvae on patch j, and the probability of

a larva leaving patch j and successfully attaching on patch i is p^{ij}. $M_k^j(a)$ is the probability that an individual has not yet matured from stage k to $k+1$ and can be calculated as $M_k^j(a) = \exp(-\int_0^a m_k^j(\tau)d\tau)$. Additional details of the model equations can be found in [23], though there the first larval stage is modelled explicitly and so the set of equations is slightly different.

To construct the next-generation matrix for this system, we need to track the number of stage 1 individuals produced on each patch from one initial stage 1 individual on a given patch. We briefly present the idea of the construction here, and more details can be found in [23].

Let $\gamma_{ij}(t)$ be the rate of production of new stage 1 individuals on patch i from an initial stage 1 individual on patch j. For an initial individual to be producing offspring it must survive through each of the sessile stages and then produce larvae which settle on another patch. Let r_k be the time that an individual spends in stage k. Then in the first $m-1$ sessile stages $k = 1, \ldots, m-1$ the probability that the individual survives up to r_k is $S_k^j(r_k)M_k^j(r_k)$ and the rate that they are maturing to the next stage at r_k is $m_k^j(r_k)$. In the last stage the probability that they survive up to r_m is $S_m^j(r_m)$ and the rate that they are producing larvae from patch j which successfully arrive on patch i is $b^j(r_m)p^{ij}$.

To calculate the rate of production at time t, $\gamma_{ij}(t)$ we multiply all the survival probabilities, maturation rates, and birth rate in each stage and integrate over all possible r_k, where $0 \le \sum_{k=1}^m r_k \le t$. To ensure this bound we rewrite $r_m = t - \sum_{k=1}^{m-1} r_k$. Thus

$$\gamma_{ij}(t) = \int_0^t \int_0^{t-r_{m-1}} \int_0^{t-r_{m-1}-r_{m-2}} \cdots \int_0^{t-\sum_2^{m-1} r_k} S_1^j(r_1)M_1^j(r_1)m_1^j(r_1)\ldots$$

$$S_{m-1}^j(r_{m-1})M_{m-1}^j(r_{m-1})m_{m-1}^j(r_{m-1})S_m^j(t - \sum_1^{m-1} r_k)b^j(t - \sum_1^{m-1} r_k)p^{ij}dr_1\ldots dr_{m-1}. \quad (17)$$

Then we can integrate the rate of production, $\gamma_{ij}(t)$ over all time to find the total number of new individuals produced on i from an initial individual on j, this is the entry in the ith row and jth column of the next-generation matrix, k_{ij}. We can use the convolution theorem,

$$\int_0^\infty f(t) * g(t)dt = \int_0^\infty f(t)dt \int_0^\infty g(t)dt \quad (18)$$

to calculate $\int_0^\infty \gamma_{ij}(t)dt = k_{ij}$ as

$$k_{ij} = \left[\prod_{k=1}^{m-1} \left(S_k^j(t)M_k^j(t)m_k^j(t)dt \right) \right] \left(\int_0^\infty S_m^j(t)b^j(t)dt \right) p^{ij}. \quad (19)$$