Bulletin of Mathematical Biology (2020) 82:9
https://doi.org/10.1007/s11538-019-00674-1

Society for Mathematical Biology

ORIGINAL ARTICLE

A Next-Generation Approach to Calculate Source–Sink Dynamics in Marine Metapopulations

Peter D. Harrington[1] · **Mark A. Lewis**[1,2]

Received: 9 September 2019 / Accepted: 2 December 2019 / Published online: 14 January 2020

Abstract
In marine systems, adult populations confined to isolated habitat patches can be connected by larval dispersal. Source–sink theory provides effective tools to quantify the effect of specific habitat patches on the dynamics of connected populations. In this paper, we construct the next-generation matrix for a marine metapopulation and demonstrate how it can be used to calculate the source–sink dynamics of habitat patches. We investigate the effect of environmental variables on the source–sink dynamics and demonstrate how the next-generation matrix can provide useful biological insight into transient as well as asymptotic dynamics of the model.

Keywords Source–sink dynamics · Next-generation matrix · Metapopulation model · Marine systems

1 Introduction

Source–sink theory was developed to better understand population dynamics in connected populations. Originating from work by Levins (1969) using metapopulation models, source–sink theory attempts to explain how certain population patches in poor environments can be sustained by population patches in more favourable environments. Population patches in poor environments are labelled "sinks", because these populations could not be sustained in the absence of dispersal. "Sources" are then population patches that can sustain themselves in the absence of dispersal. Levins used metapopulation models to study source–sink dynamics, with patch occupancy as the state variable. These models as well as other types of related models have greatly

✉ Peter D. Harrington
harringt@ualberta.ca

[1] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada

[2] Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada

contributed to the rich body of work on source–sink theory (Amarasekare and Nisbet 2001; Figueira and Crowder 2006; Hanski 1998; Pulliam 1988).

Critical to the theory of source–sink dynamics is the concept of dispersal. Dispersal is the mechanism by which source populations can rescue sink populations from extinction. Some theoretical models have modelled dispersal implicitly and have investigated how dispersal rates can change the source–sink dynamics of a population (Gyllenberg and Hanski 1997). Others have modelled dispersal explicitly (Hastings 1982), which is important when the rates or mechanisms of dispersal are understood. The rates and mechanisms of dispersal also differ largely between terrestrial and marine systems. In terrestrial systems, it is often adults that are capable of dispersing between population patches. In many marine systems, adults are confined to habitat patches, and dispersal occurs through the release of pelagic larvae, which spread through the ocean to other patches (Cowen and Sponaugle 2009).

For marine species, modelling dispersal explicitly has led to advancements in understanding the degree of connectivity between different marine subpopulations (Figueira and Crowder 2006). While there is evidence that some larval populations are well mixed in an ocean environment (Cowen et al. 2000), both theoretical advection–diffusion models for larval movement between patches (Alexander and Roughgarden 1996; Botsford et al. 1994) and computational hydrodynamic models (Cowen and Sponaugle 2009; Watson et al. 2012), have been successful at reproducing patterns of connectivity observed in marine systems. While hydrodynamic models are very useful in understanding connectivity in the specific systems for which they are parameterized, advection–diffusion models can be applied more generally to give insights into the connectivity of subpopulations. In either framework, modelling dispersal explicitly can illuminate the level of connectivity between marine subpopulations of several species.

Corals and coral reef fish (Cowen et al. 2006; Jones et al. 2009), barnacles (Roughgarden et al. 1988), Dungeness crabs (Botsford et al. 1994), sea urchins (Botsford et al. 1994), and many benthic marine species (Cowen and Sponaugle 2009) have relatively sedentary adult stages that are confined to habitat patches, with larvae that disperse between patches. In fact, it is estimated that up to 70% of benthic invertebrates have a pelagic larval phase, capable of dispersal (Mileikovsky 1971). Adult subpopulations for these meroplanktonic species, species with a planktonic larval stage, then act as connected metapopulations which are connected through larval dispersal (Botsford et al. 1994). For these marine metapopulations, local environmental conditions determine survival and productivity of adult population patches, and regional environmental conditions determine the degree of pelagic larval dispersal, as environmental conditions between patches affect the growth and survival of larvae. Both local and regional environmental conditions will then affect the source–sink distribution of the different marine habitat patches. Accurately, modelling the source–sink distribution of marine metapopulations is especially important when this information is used to design conservation management actions, such as marine protected areas. When creating marine protected areas, determining the level of dispersal between population patches, as well as the productivity of local patches, is critical. Protecting productive source patches which are capable of dispersing to sink patches may be essential in sustaining the connected metapopulations (Planes et al. 2009).

However, for certain parasitic or invasive marine species, we may be interested in controlling the spread of the species, rather than conserving the existing population. One such species of importance on which we focus specifically in the Applications section is *Lepeophtheirus salmonis*, also known as sea lice. Sea lice are a marine ectoparasite that feed on the epidermal tissues of salmon (Costello 2006). When sea lice are present in high densities, their salmonid hosts can experience additional morbidity and mortality (Costello 2006; Krkošek et al. 2011), as well as reduced foraging ability (Godwin et al. 2017). Salmon farms in coastal ecosystems present stationary hosts for sea lice, on which sea lice can survive year round (Rogers et al. 2013). Sea lice are a large economic issue facing salmon farms worldwide and have previously cost the global aquaculture industry 6% of its product value a year (Costello 2009). These farms act as population patches on which sea lice can grow until maturity. Adult females exude egg strings and release larvae which can spread between salmon farms via coastal currents and ocean mixing. This larval dispersal between farms can connect sea lice populations on different farms within a specific region (Aldrin et al. 2017). The larval dispersal also transmits sea lice between farmed salmon and wild salmon migrating past farms (Krkošek et al. 2006) and has been shown to lead to population declines in Pacific pink salmon populations (Krkošek et al. 2007). To protect both farmed and wild salmon populations from the effects of sea lice, salmon farms now use a variety of treatment measures to reduce sea lice levels when populations outbreak (Aaen et al. 2015). However, in many regions sea lice have developed resistance to some of the most effective chemical treatments (Aaen et al. 2015), and even in regions without resistance sea lice continue to pose a threat to wild salmon (Bateman et al. 2016). Due to the economic and ecological importance of controlling sea lice on salmon farms, we use salmon farms as an example to study the source–sink distribution of habitat patches under different environmental conditions, as well as the effect of treatment.

To study source–sink distributions in sea lice and other meroplanktonic marine populations on habitat patches, we use a next-generation approach. Next-generation operators have a rich history in epidemiology (Diekmann et al. 1990; van den Driessche and Watmough 2002; Diekmann et al. 2010) and are often used to determine whether an infectious epidemic will occur in a population, by calculating the basic reproduction number, R_0. Recently, next-generation operators have been used to analyse heterogeneous aquatic populations (Huang and Lewis 2015; Huang et al. 2016b; Krkošek and Lewis 2010; McKenzie et al. 2012). The next-generation operator can be used to map the current number of individuals at each life stage in a heterogeneous environment to the new number of individuals at each life stage after one generation. If populations exist in discrete patches, then this next-generation operator can be formulated as a next-generation matrix. The next-generation approach distils the complex process of dispersal between patches and stage-structured survival and growth on a patch into a single operator. In ecological systems, the next-generation operator has been used to determine population persistence, but has also been used to determine the source–sink distribution of a population (Huang et al. 2016b; Krkošek and Lewis 2010; McKenzie et al. 2012).

Specifically, new measures of persistence, $R_\delta(x)$, $R_{\mathrm{loc}}(x)$, R_u and R_l have been defined using next-generation operators. $R_{\mathrm{loc}}(x)$ is the number of new individuals produced at location x in the absence of dispersal and thus can be used to measure

the fundamental niche of a population (Huang et al. 2016b). $R_\delta(x)$ is the number of new individuals produced over the entire population from one individual starting at spatial position x (Huang et al. 2016b). It takes into account both growth and survival at location x and dispersal from location x. If one individual at x produces less than one individual over the entire landscape, then $R_\delta(x) < 1$, and the location x is defined as a sink. If $R_\delta(x) > 1$, then x is defined as a source. Pulliam (1988) originally defined a source habitat as a habitat that can sustain itself in isolation and a sink habitat as one that cannot sustain itself in isolation, assuming low population density. However, this definition can give rise to connected patches of sink habitats that persist (Armsworth 2002). A benefit of the $R_\delta(x)$ measure is that it does not allow for connected patches of sinks to persist. Lastly, R_u and R_l are then defined as the maximum and minimum R_δ taken over all possible locations, respectively, and are shown to be the intergenerational growth rate under the best and worst possible initial conditions (Huang and Lewis 2015). They can therefore be useful in determining bounds for intergenerational growth, as well as R_0.

Another method of measuring the contribution that each patch provides to the total population is to look at the contribution of each patch to R_0 (Hurford et al. 2010). This approach uses the left and right eigenvectors associated with R_0 to determine the contribution of each patch if the population were distributed according the right eigenvector. In the Applications section, we build on and apply all of these persistence measures and next-generation theory to determine the source–sink distribution on the discrete population patches in our system under different environmental conditions.

The final concept in this paper borrowed from epidemiology is the type reproduction number (Heesterbeek and Roberts 2007; Lewis et al. 2019; Roberts and Heesterbeek 2003). The type reproduction number is a measure of the control effort needed when control is targeted at a certain type of individual in a heterogeneous population. For patch models, this is often the control effort required on a certain patch so that the total population cannot grow. Type reproduction numbers need not only be used to determine control, they can also determine the amount of enhancement effort required on a patch so that the entire population will grow. We use the concept of type reproduction number in the Applications section to determine the treatment level required for sea lice on a salmon farm so that the entire sea lice population is controlled.

In this paper, we use a next-generation approach to quantify the source–sink distribution of marine meroplanktonic populations where subpopulations are confined to local habitat patches and are connected via larval dispersal. First, we present a stage-structured model for a general marine population, composed of several sessile stages which survive and reproduce on a population patch. The final adult sessile stage gives birth to planktonic larvae, which disperse between patches. The dispersal of larvae is modelled explicitly by approximating hydrodynamic ocean movement using the Fokker–Planck equation. Next, we construct a next-generation matrix for this model and prove that the spectral radius of the next-generation matrix determines whether or not the species can persist. Lastly, we apply different persistence measures to sea lice populations on salmon farms, to answer several key questions around the source–sink dynamics of sea lice on salmon farms:

1. What is the source–sink distribution of salmon farms in a channel?

2. How does the source–sink distribution change with respect to environmental variables?
3. Are there certain parameter regions in which local outbreaks can occur, but not global outbreaks?
4. What is the effect of treating a single salmon farm for sea lice control on the transient and asymptotic dynamics?
5. What is the effect of an environmental gradient on patch contributions to R_0 and the source–sink distribution?

2 The Stage-Structured Patch Model

To study the source–sink distribution of a marine metapopulation, we consider a marine species with m life stages that is spreading between n spatial patches. We are interested in identifying which habitat patches are acting as sources and which are acting as sinks. Our focus is populations at low densities, typically near the extinction equilibrium. Determining the source–sink distribution can also uncover transient dynamics that may be substantially different than the long-term asymptotic dynamics of the population. To investigate the source–sink distribution, we construct a next-generation matrix for our model, and we show in a later section on applications (Sect. 4.2) how to use the column sums of the next-generation matrix to determine the source–sink distribution.

We are interested in modelling a marine meroplanktonic species where the larval (first) stage is the only stage capable of dispersing between population patches. The remaining juvenile and adult stages are confined to a single population patch. We call these stages confined to a patch sessile stages, though in reality they could be motile but restricted to the habitat patch; such is the case for sea lice or reef fish. We assume that the last adult stage is the only stage capable of reproduction. This is the case for sea lice, on which we focus specifically in the Applications section, but is also the case for other marine species mentioned in the introduction, depending on how stages are grouped. During the larval stage, we assume there is a latent period during which the newly released larva cannot attach to a new patch. Here, we define larvae not capable of attaching to a new patch as latent and larvae capable of attaching to a new patch as active. Larvae are therefore first released from a patch as latent larvae and then mature into active larvae, at which time they are capable of attaching to a new patch. Some marine species have larval stages that are active directly after release (Mileikovsky 1971), and so for these species, we can ignore the latent larval stage.

In both larval and sessile stages, an individual will either die or mature to the next stage, with the exception of the last sessile stage where individuals give birth instead of maturing. We choose to model the number of individuals in each stage using density equations, so that we are tracking the density of individuals in a given stage that have spent a time units in the stage at time t. Modelling the population using density equations allows for the probability of survival in a stage, as well as the probability of maturing to the next stage, to depend on a, the amount of time already spent in the stage. When modelling populations using ordinary differential equations (ODEs), it is

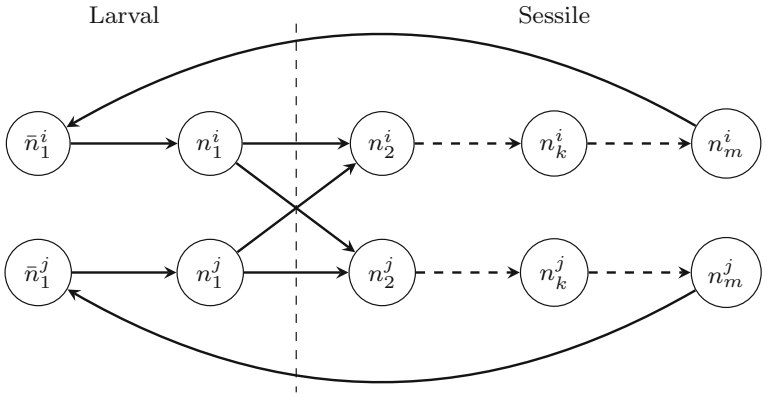


Fig. 1 The structure of the life cycle graph for a marine species with m stages on two patches. The top row shows the stages associated with patch i, and the bottom row shows the stages associated with patch j. The larval stages on the left have just left their respective patches, and the recruitment onto a patch occurs as the larval stage n_1 arrives on a patch as a sessile individual in stage n_2

assumed that stage durations are always exponentially distributed (Feng et al. 2007). However, this assumption is oversimplistic for many marine populations including sea lice.

While modelling using density equations allows for more generality in the survival and maturation of an individual, it is more difficult to include density dependence in this framework. However, source and sink populations are typically categorized in the context of low population densities, especially when calculated through a next-generation operator (Krkošek and Lewis 2010), and thus, we do not expect density dependence to play a crucial role in the context of source–sink dynamics for the problems considered. Our metapopulation model will therefore ignore the density dependent effects that could influence survival and maturation at both high and low densities. This model is then most useful for investigating the impact of connectivity among habitat patches at low population densities, in populations where habitat patches are not resource limited, or where populations are artificially managed to remain at low densities. It is in this context of negligible density dependence that we ask the five questions given at the end of the introduction.

2.1 Derivation of the Stage-Structured Patch Population Model

In this section, we derive a system of density equations that model the change in population density on each habitat patch in our connected metapopulation. The general structure of the model, consisting of sessile stages confined to a habitat patch connected by larval dispersal, can be seen from the life cycle graph for two subpopulations (Fig. 1). To mathematically describe our model, we first give the general structure for the density of sessile individuals in stage k on patch i.

Let $n_k^i(t, a)$ be the density of individuals with stage age a at time t, in stage k on patch i. Then,

$$n_k^i(t,a) = \begin{cases} B_k^i(t-a)S_k^i(a)M_k^i(a) & t > a \\ n_{k,0}^i(a-t)\frac{S_k^i(a)M_k^i(a)}{S_k^i(a-t)M_k^i(a-t)} & 0 < t < a \\ n_{k,0}^i(a) & t = 0 \end{cases} \tag{1}$$

for $k = 2, \ldots, m-1$. Here, $B_k^i(t)$ is the rate at which individuals are entering stage k on patch i at time t, $S_k^i(a)$ is the probability that an individual survives longer than a time units in a stage, given that they have not yet matured, $M_k^i(a)$ is the probability that an individual takes longer than a time units to mature to the next stage, given that they have survived, and $n_{k,0}^i(a)$ is the initial density of individuals with age a. We assume that both $S_k^i(a)$ and $M_k^i(a)$ are non-negative and non-increasing functions, with $S_k^i(0) = M_k^i(0) = 1$, and that $S_k^i(a)$ and $M_k^i(a)$ are L^1 functions, so $\int_0^\infty S_k^i(a)\mathrm{d}a < \infty$ and $\int_0^\infty M_k^i(a)\mathrm{d}a < \infty$. We also assume that survival and maturation in a given stage are independent. A selection of $M_k^i(a)$ functions that have been used for sea lice models is shown in Fig. 2.

Individuals with stage age $a > t$ must have entered the stage at time $t - a$, with rate $B_k^i(t-a)$, and then survived until stage age a with probability $S_k^i(a)M_k^i(a)$; the density of these individuals is given by the first line of Eq. 1. If individuals have stage age $a < t$, then they were already in the stage at $t = 0$, with density $n_{k,0}^i(a-t)$, and the probability that they survive to stage age a, given that they were present at stage age $a - t$, is $S_k^i(a)M_k^i(a)/S_k^i(a-t)M_k^i(a-t)$; the density of these individuals is given by the second line of Eq. 1. This formula for $n_k^i(t,a)$ is also the solution to the McKendrick–von Foerster partial differential equation (Keyfitz and Keyfitz 1997; McKendrick 1925):

$$\begin{aligned} \frac{\partial n_k^i(t,a)}{\partial t} + \frac{\partial n_k^i(t,a)}{\partial a} &= -\mu_k^i(a)n_k^i(t,a) \\ n_k^i(t,0) &= B_k^i(t) \\ n_k^i(0,a) &= n_{k,0}^i(a) \\ \mu_k^i(a) &= -\frac{\left(M_k^i(a)S_k^i(a)\right)'}{M_k^i(a)S_k^i(a)}, \end{aligned}$$

which can be found by integrating along the characteristic curves, as shown in Appendix A.

The larval (first) stage, which is capable of spreading between patches, includes both a latent and active stage. During the latent larval stage, individuals spread away from a patch, but are not capable of attaching to another patch. They then enter the active stage, where they are capable of attaching to another patch. To distinguish between the two stages, let $\bar{n}_1^i(t,a)$ be the density of latent larvae released from patch i with stage age a, and $n_1^i(t,a)$ be the density of active larvae. Let $\bar{S}_1^i(a)$ be the survival function for the latent stage, and let $M_1^i(a)$ the maturation function for the larval stage, so that $M_1^i(a)$ is the probability that a latent larva has not yet matured into an active larva, given it has survived.

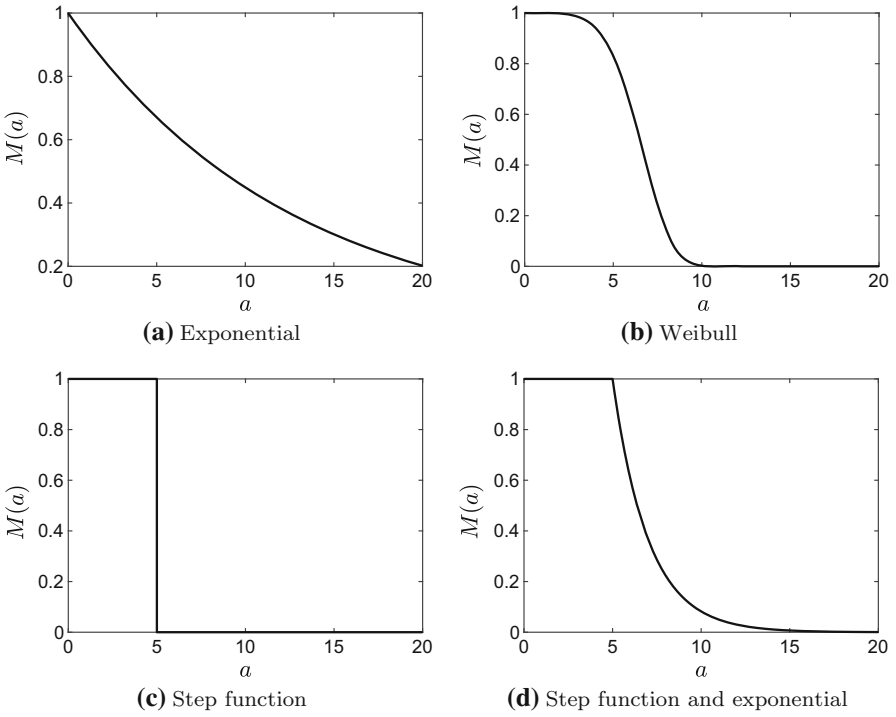


Fig. 2 A selection of $M_k^i(a)$ functions that have been used in sea lice population models. When a constant maturation rate is assumed, as in Krkošek et al. (2012a), $M_k^i(a)$ is represented as an exponential function, shown in (**a**). **b** The Weibull survival function was used by Aldrin et al. (2017) to avoid strict minimum development times and constant maturation rates. **c** The step function was used by Revie et al. (2005), where it is assumed that all sea lice in a stage mature at the same time. **d** A combination of a step function and exponential function was used by Stien et al. (2005), where the step function is used to enforce a minimum development time, after which the exponential is used to capture a constant maturation rate

During the active stage, instead of maturing, active larvae will be removed from this stage when they arrive on another patch. Let $f^{ij}(a)$ be the arrival time density function for an active larva spreading from patch j to patch i, where $\int_{a_1}^{a_2} f^{ij}(a)\mathrm{d}a$ is the probability that the active larva arrives on patch i between stage age a_1 and a_2. Let $F^{ij}(a) = \int_0^a f^{ij}(\tau)\mathrm{d}\tau$, then $1 - F^{ij}(a)$ is the probability that the larva has not yet arrived on a patch i by stage age a, given that it has not died. The physical process underlying the arrival time density function is shown in detail in Sect. 2.4, and a typical $f^{ij}(a)$ is shown in Fig. 3. In Sect. 2.4, we also show that if we define $f^j(a)$ as the arrival time density for an active larvae leaving patch j to arrive on any patch, then $f^j(a) = \sum_{i=1}^n f^{ij}(a)$, and $F^j(a) = \sum_{i=1}^n F^{ij}(a)$.

Let $S_1^i(a)$ be the survival function for the active larvae leaving patch i. The density of latent larvae in stage $k = 1$ leaving patch i is:

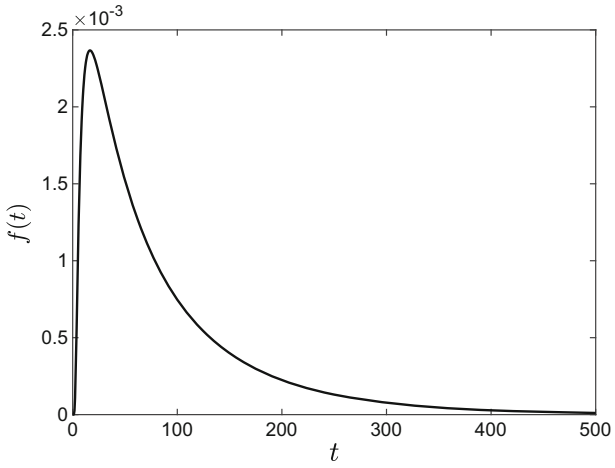


Fig. 3 The arrival time density, $f^{ij}(t)$, as a function of time, t, for larvae leaving patch j and arriving on patch i. Movement from patch j to patch i is described in Sect. 2.4. Here, patch j was located at $x = 0$, patch i at $x = 15$, with additional parameters $v = 1$, $D = 5, \alpha = 0.1$, $\Delta = 0.8$. A one-dimensional domain was used, with no other patches present

$$\bar{n}_1^i(t,a) = \begin{cases} \bar{B}_1^i(t-a)\bar{S}_1^i(a)M_1^i(a) & t > a \\ \bar{n}_{1,0}^i(a-t)\frac{\bar{S}_1^i(a)M_1^i(a)}{\bar{S}_1^i(a-t)M_1^i(a-t)} & 0 < t < a \\ \bar{n}_{1,0}^i(a) & t = 0 \end{cases},$$

and the density of active larvae is:

$$n_1^i(t,a) = \begin{cases} B_1^i(t-a)S_1^i(a)(1-F^i(a)) & t > a \\ n_{1,0}^i(a-t)\frac{S_1^i(a)(1-F^i(a))}{S_1^i(a-t)(1-F^i(a-t))} & 0 < t < a \\ n_{1,0}^i(a) & t = 0 \end{cases}.$$

The last sessile adult stage also requires a different density equation from the rest of the sessile stages; for during this stage, individuals cannot mature any longer and they can only survive. For $k = m$, the density of individuals on a patch i is:

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

To complete our model, we need to define $B_k^i(t)$, which is the rate that individuals enter each stage. For stages $3, \ldots, m$, $B_k^i(t)$ will be the rate at which individuals from stage $i-1$ are maturing to stage i:

$$B_k^i(t) = \int_0^\infty n_{k-1}^i(t,a) m_{k-1}^i(a) \mathrm{d}a,$$

where $m_k^i(a) = -M_k^{i'}(a)/M_k^i(a)$ is the instantaneous maturation rate for individuals in stage k with stage age a. Multiplying the current density by $m_k^i(a)$ and integrating across all stage ages gives the total density of individuals maturing at time t.

Individuals in the last stage give birth to latent larvae in the first stage. Let $b^i(a)$ be the stage age dependent birth rate in patch i, then $\bar{B}_1^i(t)$ is given by:

$$\bar{B}_1^i(t) = \int_0^\infty n_m^i(t,a) b^i(a) \mathrm{d}a.$$

The latent larvae then mature into active larvae in the first stage with rate:

$$B_1^i(t) = \int_0^\infty \bar{n}_1^i(t,a) m_1^i(a) \mathrm{d}a.$$

Individuals enter the second stage on patch i by arriving as active larvae, which are spreading from all patches. The instantaneous rate that an active larva, travelling from patch j to patch i, arrives on patch i as a larva in the second stage, is $f^{ij}(a)/(1-F^j(a))$. Therefore, we have

$$B_2^i(t) = \sum_{j=1}^n \int_0^\infty n_1^{ij}(t,a) f^{ij}(a)/(1-F^j(a)) \mathrm{d}a.$$

Combining all of these equations, the age density of individuals is

$$\bar{n}_1^i(t,a) = \begin{cases} \underbrace{\bar{B}_1^i(t-a)\bar{S}_1^i(a)M_1^i(a)}_{\text{entered at } t-a\text{, survived to } a} & t > a \\ \underbrace{\bar{n}_{1,0}^i(a-t)\dfrac{\bar{S}_1^i(a)M_1^i(a)}{\bar{S}_1^i(a-t)M_1^i(a-t)}}_{\text{present at a-t, survived to a}} & 0 < t < a \\ \underbrace{\bar{n}_{1,0}^i(a)}_{\text{initial density}} & t = 0 \end{cases} \qquad k = 1, \tag{2}$$

$$n_1^i(t,a) = \begin{cases} B_1^i(t-a)S_1^i(a)(1-F^i(a)) & t > a \\ n_{1,0}^i(a-t)\dfrac{S_1^{ij}(a)(1-F^i(a))}{S_1^i(a-t)(1-F^i(a-t))} & 0 < t < a \\ n_{1,0}^i(a) & t = 0 \end{cases} \qquad k = 1,$$

$$n_k^i(t,a) = \begin{cases} B_k^i(t-a)S_k^i(a)M_k^i(a) & t > a \\ n_{k,0}^i(a-t)\frac{S_k^i(a)M_k^i(a)}{S_k^i(a-t)M_k^i(a-t)} & 0 < t < a \quad k = 2,\ldots,m-1, \\ n_{k,0}^i(a) & t = 0 \end{cases}$$

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

$$\bar{B}_1^i(t) = \int_0^\infty n_m^i(t,a)b^i(a)\mathrm{d}a,$$

$$B_1^i(t) = \int_0^\infty \bar{n}_1^i(t,a)m_1^i(a)\mathrm{d}a,$$

$$B_2^i(t) = \sum_{j=1}^n \int_0^\infty n_1^j(t,a)f^{ij}(a)/(1-F^j(a))\mathrm{d}a,$$

$$B_k^i(t) = \textstyle\int_0^\infty n_{k-1}^i(t,a)m_{k-1}^i(a)\mathrm{d}a \qquad k = 3,\ldots,m,$$

These equations can also be expressed in an integral form, by substituting the equations for $n_k^i(t,a)$ into $B_k^i(t)$, and tracking the total number of parasites in each patch at each stage, $N_k^i(t) = \int_0^\infty n_k^i(t,a)\mathrm{d}a$. System 2 then becomes:

$$\bar{N}_1^i(t) = \underbrace{\int_0^t \bar{B}_1^i(t-a)\bar{S}_1^i(a)M_1^i(a)\mathrm{d}a}_{\text{entered at } t-a\text{, survived to } a} + \underbrace{\int_t^\infty \bar{n}_{1,0}^i(a-t)\frac{\bar{S}_1^i(a)M_1^i(a)}{\bar{S}_1^i(a-t)M_1^i(a-t)}\mathrm{d}a}_{\text{present at a-t, survived to a}}$$

$$N_1^i(t) = \int_0^t B_1^i(t-a)S_1^i(a)(1-F^i(a))\mathrm{d}a + \int_t^\infty n_{1,0}^i(a-t)\frac{S_1^i(a)(1-F^i(a))}{S_1^i(a-t)(1-F^i(a-t))}\mathrm{d}a$$

$$N_k^i(t) = \int_0^\infty B_k^i(t-a)S_k^i(a)M_k^i(a)\mathrm{d}a + \int_t^\infty n_{k,0}^i(a-t)\frac{S_k^i(a)M_k^i(a)}{S_k^i(a-t)M_k^i(a-t)}\mathrm{d}a, \quad k = 2,\ldots,m-1$$

$$N_m^i(t) = \int_0^t B_m^i(t-a)S_m^i(a)\mathrm{d}a + \int_t^\infty n_{m,0}^i(a-t)\frac{S_m^i(a)}{S_m^i(a-t)}\mathrm{d}a$$

$$\bar{B}_1^i(t) = \int_0^t B_m^i(t-a)S_m^i(a)b^i(a)\mathrm{d}a + \int_t^\infty n_{m,0}^i(a-t)\frac{S_m^i(a)b^i(a)}{S_m^i(a-t)}\mathrm{d}a$$

$$B_1^i(t) = \int_0^t \bar{B}_1^i(t-a)\bar{S}_1^i M_1^i(a)m_1^i(a)\mathrm{d}a$$

$$+\int_t^\infty \bar{n}^i_{1,0}(a-t)\frac{\bar{S}^i_1(a)M^i_1(a)m^i_1(a)}{\bar{S}^i_1(a-t)M^i_1(a-t)}\mathrm{d}a$$

$$B^i_2(t) = \sum_{j=1}^n \int_0^t B^j_1(t-a)S^j_1 f^{ij}(a)\mathrm{d}a$$

$$+\sum_{j=1}^n \int_t^\infty n^{ij}_{1,0}(a-t)\frac{S^j_1(a)f^{ij}(a)}{S^j_1(a-t)(1-F^j(a-t))}\mathrm{d}a$$

$$B^i_k(t) = \int_0^t B^i_{k-1}(t-a)S^i_{k-1}(a)M^i_{k-1}(a)m^i_{k-1}(a)\mathrm{d}a$$

$$+\int_t^\infty n^i_{k-1,0}(a-t)\frac{S^i_{k-1}(a)M^i_{k-1}(a)m^i_{k-1}(a)}{S^i_{k-1}(a-t)M^i_{k-1}(a-t)}\mathrm{d}a, \quad k=3,\ldots,m \qquad (3)$$

Here, we have constructed an stage-structured model for a marine population with several sessile stages on local habitat patches connected by larval dispersal. We have formulated our model as both a system of age density equations (system 2), which we find most intuitive, as well as a system of integral equations (system 3).

2.2 Reduction to a System of ODEs

It is also possible to reduce Eq. 2 to a system of ODEs under some strong assumptions. We do not believe that these assumptions are sufficiently realistic for benthic marine species with dispersing larvae. However, we include this reduction here as an illustrative example of how Eq. 2 can be connected to the more familiar ODE model structure. We start by assuming that time till maturation and time to death are both exponential waiting times. We also need to make this assumption for the arrival time. This arrival time distribution is no longer be directly solved through the more realistic advection–diffusion equation (7), given in Sect. 2.4 and shown in Fig. 3. However, the new exponential rate could be an approximation by using the average arrival time generated by the advection–diffusion equation.

Here, we will use the lower case letter as the exponential rate associated with the survival or maturation function. For example, $M^i_k(a) = e^{-m^i_k a}$, and $S^i_k(a) = e^{-s^i_k a}$. For the arrival time, we now assume that there is a constant rate of arrival of larvae, f^{ij}, from a source patch j to a receiving patch i. We can then formulate $f^j(a)$, the distribution of arrival times for a larva leaving patch j to arrive on any other patch, as the exponential function $f^j(a) = \sum_{i=1}^n f^{ij} e^{-\sum_{i=1}^n f^{ij} a}$. By formulating the arrival times using constant rates, we lose the spatial structure of the system, so that patches are now only distinguished by their arrival time rates, f^{ij}.

To reduce the system of density equations to a system of ODEs, it is easiest to use the McKendrick–von Foerster formulation of the model (Appendix A). We simply integrate the different versions of Eq. 14 of equations over all ages, so $N^i_k(t) = \int_0^\infty n^i_k(t,a)\mathrm{d}a$. The resulting model is as follows:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bar{N}_1^i(t) &= b^i N_m^i(t) - (\bar{s}_1^j + m_1^j)\bar{N}_1^j(t) && k = 1\\
\frac{\mathrm{d}}{\mathrm{d}t}N_1^j(t) &= m_1^j \bar{N}_1^j - (s_1^j + \sum_{i=1}^{n} f^{ij})N_1^j(t) && k = 1\\
\frac{\mathrm{d}}{\mathrm{d}t}N_2^i(t) &= \sum_{j=1}^{n} f^{ij} N_1^j(t) - (s_2^i + m_2^i)N_2^i(t) && k = 2\\
\frac{\mathrm{d}}{\mathrm{d}t}N_k^i(t) &= m_{k-1}^i N_{k-1}^i(t) - (s_k^i + m_k^i)N_k^i(t) && k = 3, \ldots, m-1\\
\frac{\mathrm{d}}{\mathrm{d}t}N_m^i(t) &= m_{m-1}^i N_{m-1}^i(t) - s_m^i N_m^i(t) && k = m \qquad (4)
\end{aligned}
$$

Here, we have also assumed the birth rate of latent larvae on each patch is constant. The reduction of our full system of equations (2) to a system of ODEs results in the loss of the age structure present in our original model as well as stronger model assumptions. However, formulating the model as a system of ODEs allows for a more familiar comparison between our model and other population models.

2.3 The Next-Generation Matrix for the Patch Model

In this section, we define the next-generation matrix for our model. In the Applications section, we show how this next-generation matrix can be used to identify source and sink habitat patches, using the column sums of this matrix. We also show how the source–sink distribution can be used to determine transient dynamics in our model. Next-generation matrices are often used to describe new infections in compartmental disease models (Diekmann et al. 1990, 2010; van den Driessche and Watmough 2002), though here we use the next-generation matrix to describe the growth and spread of marine organisms between patches. In the classic formulation of a next-generation matrix, the (i, j) entry describes the number of new infections in the ith compartment produced by one new infection in the jth compartment. In our model, since we are tracking individuals and not infections, we need to define "new" individuals. We choose to define "new" individuals as those first entering a patch at stage $k = 2$. We choose $k = 2$ as the first stage because this is the first sessile stage where individuals arrive on a patch and can be counted.

In our model, we have n patches and m stages, so in total we have $n \times m$ compartments. However, the only new individuals are produced in stage $k = 2$. We could create a next-generation matrix of size $nm \times nm$, but it would only have n nonzero rows, as there is only one stage on every patch where new individuals are produced. This next-generation matrix is referred to the next-generation matrix with large domain, K_L, by Diekmann et al. (2010). Instead, we can group all stages together for a single patch, so that the (i, j) entries of our next-generation matrix are the number of new individuals (stage $k = 2$) produced on patch i, from one new individual on patch j. This is referred to as the next-generation matrix K by Diekmann et al. (2010), as we have removed all compartments which cannot have "new" individuals, and are only

tracking the production of "new" individuals in compartments that start with "new" individuals. We will elaborate on the details of this process in Sect. 3.1.

For our system (2), the (i, j) entry of the next-generation matrix, K is

$$K(i,j) = \prod_{k=2}^{m-1} \left(\int_0^\infty S_2^j(t) M_2^j(t) m_2^j(t) \mathrm{d}t \right) \left(\int_0^\infty S_m^j(t) b^j(t) \mathrm{d}t \right) \left(\int_0^\infty \bar{S}_1^j(t) M_1^j(t) m_1^j(t) \mathrm{d}t \right) \left(\int_0^\infty S_1^j(t) f^{ij}(t) \mathrm{d}t \right). \tag{5}$$

In order to understand which patches may be acting as sources or sinks, we can look at the column sums of the different spatial patches. The sum of column j is the total number of new individuals (stage $k = 2$) produced on all patches from one new individual on patch j. If the sum of column j is larger than one, then each new individual on patch j is producing more than one new individual on all patches. Therefore, patch j is a source. Similarly, if the sum of column j is less than one, then patch j is a sink. We will expand on and formalize this quantification of source–sink dynamics in Sect. 4.2.

The general structure of our model and next-generation matrix allows it to be readily applied to a number of systems. However, in order to examine the effect of changing biological environments on the next-generation matrix, it is useful to look at specific survival and maturation functions, $S_k^j(t)$ and $M_k^j(t)$, for each stage k, and on each patch j. Suppose we let $S_k^j(t)$ be the survival function associated with the exponential distribution: $S_k^j(t) = e^{-\mu_k^j t}$. This means that in each stage and on each patch the instantaneous death rate, μ_k^j is constant and does not depend on the time spent in the stage. This is a common biological assumption, as mortality is often governed primarily the external environment and is often independent of age. For sea lice, this is assumed for most population models (Adams et al. 2015; Aldrin et al. 2017; Krkošek et al. 2006; Revie et al. 2005).

Next, we consider a simplifying case where the maturation probability density function (p.d.f.), $-M_k^j(t)'$, is the gamma distribution. There are several reasons for this choice. First, the gamma distribution can be unimodal, and therefore biologically represents a situation in which the highest probability of maturing to the next stage is at some intermediate age. This is the case for sea lice, which have a minimum required development time before they can mature through a stage (Johnson and Albright 1991). Second, the gamma distribution can reduce to the exponential distribution and, in a limiting case, to the delta distribution. Exponential and delta maturation p.d.f.s have both been used to model sea lice (Krkošek et al. 2012a; Revie et al. 2005), and their maturation functions $M_k^j(t)$ are shown in Fig. 2. When the gamma distribution is reduced to an exponential distribution, our system of density equations (2) can be reduced to an ODE system (Sect. 2.2). When the gamma distribution is reduced to a Dirac delta distribution, our system could be formulated as a system of discrete delay differential equations. The gamma distribution is also similar to the Weibull distribution, which has been used to model sea lice (Aldrin et al. 2017), as both distributions are continuous, unimodal, and can be reduced to exponential distributions. Lastly,

the integration of the gamma distribution multiplied by the exponential distribution is simple to evaluate analytically, and thus, our expression for the next-generation matrix does not become overly complicated. If we use β_k^j as the rate parameter and a_k^j as the shape parameter, then the maturation function becomes

$$-\frac{\mathrm{d}}{\mathrm{d}t}M_k^j(t) = M_k^j(t)m_k^j(t) = \frac{{\beta_k^j}^{a_k^j} x^{a_k^j-1} e^{-\beta_k^j x}}{\Gamma(a_k^j)}.$$

The last function to define explicitly is the age dependant birth rate, $b(t)$. Here, we assume that the birth rate is constant, $b(t) = b$, so larvae are produced at a constant rate as soon as an individual enters their final stage of maturation. This is a biologically reasonable assumption and again simplifies our calculations analytically. Our arrival time function, $f^{ij}(t)$, is derived in Sect. 2.4 and thus cannot be assumed to have any particular form.

Under all the stated assumptions, the next-generation matrix simplifies to:

$$K(i,j) = \prod_{k=2}^{m-1}\left(\frac{{\beta_k^j}^{a_k^j}}{(\beta_k^j+\mu_k^j)^{a_k^j}}\right)\left(\frac{b}{\mu_m^j}\right)\left(\frac{{\beta_1^j}^{a_1^j}}{(\beta_1^j+\bar{\mu}_1^j)^{a_1^j}}\right)\left(\int_0^\infty e^{-\mu_1^j t} f^{ij}(t)\mathrm{d}t\right). \tag{6}$$

Here, we have shown that the next-generation matrix distils the essential quantities of larval dispersal between population patches as well as growth and survival on local patches into a single operator. From the next-generation matrix, we can calculate the source–sink distribution of our connected metapopulation. By approximating the maturation functions as gamma distributions and survival functions as exponential functions the form of the matrix can be simplified, while maintaining sufficient generality to approximate several realistic biological systems.

2.4 The Arrival Time of Larvae Moving Between Patches

In this section, we formally define and demonstrate the calculation of the arrival time density, $f^{ij}(t)$. In previous sections, we have focused on stages that grow on distinct patches and now we turn our attention to the first, or larval stage, which is spreading between patches. We allow the larval stage to have a latent period, where the larvae cannot arrive at the second patch even if it passes by. The larvae then mature into an active larvae and during the active phase may arrive on a new patch. We include this latent period to allow our model to be applicable to various marine organisms where the first larval stage is not capable of attaching to a habitat patch, though the latency period can also easily be removed.

To formally define the arrival time distribution, let T be a random variable which describes the time from maturation that an active larva arrives on any patch, after it is released as a latent larva from patch j. If the larva does not arrive on a patch, then we say T is infinite. Then, $\int_{t_1}^{t_2} f^j(t)\mathrm{d}t = \Pr(t_1 < T < t_2)$, so that $f^j(t)$ is the distribution of arrival times for larvae leaving patch j. We will also show that

$f^j(t) = \sum_{i=1}^{n} f^{ij}(t)$, where $f^{ij}(t)$ is the distribution of arrival times for an active larvae from patch j arriving on patch i. In order to determine $f^j(t)$ and $f^{ij}(t)$, we will first solve an equation governing the movement of the larvae between patches, for both the latent and active larval stages.

We are interested in approximating the movement of larvae between habitat patches in a marine environment, so we approximate movement using the Fokker–Planck equation, or advection–diffusion equation in the case of constant diffusion. This equation has previously been used to model the dispersal of sea lice larvae away from salmon farms (Krkošek et al. 2006), as well as barnacle and crab larvae along the California coast (Alexander and Roughgarden 1996; Botsford et al. 1994). The diffusion term represents the effect of tides and ocean mixing, and the advection term represents any flow due to constant currents, potentially generated by river outflow, or other sources.

To allow for local hydrodynamic movement in a protected patch, we divide the total larvae produced at a patch into a fraction r that remain locally around the patch, and the remaining fraction q that enter the larger ocean environment and are then influenced by advection and diffusion. The fraction q that enter the channel are still able to rearrive on their natal patch.

To determine movement between patches and subsequently arrival on habitat patches in our model, we use the Fokker–Planck equation in a one-dimensional domain. We use a one-dimensional domain because coastlines can typically be approximated by a one-dimensional domain, and in the Applications section, we apply our model to salmon farms located in a channel. However, the advection–diffusion equation could easily be structured in its two-dimensional form, if the model was to be used to analyse marine species where patches did not simply lie in a channel or along a coastline.

Let $\bar{p}^j(x,t)$ be the probability density function for the location of the latent larvae leaving patch j into the channel environment as a function of time. The larvae are released from patch j, and then, we assume the movement of the larvae between patches is governed by the Fokker–Planck equation:

$$\begin{aligned} \frac{\partial}{\partial t}\bar{p}^j(x,t) &= -\frac{\partial}{\partial x}\left(v(x)\bar{p}^j(x,t)\right) + \frac{\partial^2}{\partial x^2}\left(D(x)\bar{p}^j(x,t)\right) \\ \bar{p}^j(x,0) &= qh(x-x_j)/\Delta, \end{aligned} \tag{7}$$

where Δ is the size of the patch, and $h(x) = 1$ when $x \in [-\Delta/2, \Delta/2]$ and $h(x) = 0$ otherwise. $D(x)$ is the diffusion coefficient and $v(x)$ is the advection coefficient of the environment.

Latent larvae then mature into active larvae. In Sect. 2, we defined $M_1^j(a)$ as the probability that a latent larvae released from patch j has not yet matured into an active larva, and $m_1^j(a)$ as the instantaneous rate of maturation. Using these two previously defined functions, $M_1^j(a)m_1^j(a)$ is therefore the probability density function associated with maturation. If $\bar{T}$ is the time it takes for the latent larva to mature into an active larva after it is released, then $\int_{t_1}^{t_2} M_1^j(t)m_1^j(t)\mathrm{d}t = \Pr(t_1 < \bar{T} < t_2)$. Once larvae mature into active larvae, they continue to move according to the Fokker–Planck equation, but now they arrive on patch i with rate α_i as they pass by. Let $p^j(x,t)$ be the probability density function of the active larvae travelling from patch j, then $p^j(x,t)$ is given by:

$$\frac{\partial}{\partial t}p^j(x,t) = -\frac{\partial}{\partial x}\left(v(x)p^j(x,t)\right) + \frac{\partial^2}{\partial x^2}\left(D(x)p^j(x,t)\right) - \sum_{i=1}^{n}\alpha_i h(x-x_i)p^j(x,t)$$
$$p^j(x,0) = \int_0^\infty M_1^j(\tau)m_1^j(\tau)\bar{p}^j(x,t)\mathrm{d}\tau. \tag{8}$$

Local larvae that remain at a patch also mature according to the same maturation function. As they do not spread between patches, the density of active larvae that are in local water column around patch j, P^j, can be described by the ordinary differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}P^j(t) = -\alpha_j^r P^j(t)$$
$$P(0) = r\int_0^\infty M_1^j(\tau)m_1^j(\tau)\mathrm{d}\tau, \tag{9}$$

where α_j^r, is the rate of arrival of the local larvae. We allow α_j^r to differ from α_j, the rate of arrival of larvae moving between patches.

We now consider the arrival time of an active larva. Let T be the time of arrival of the active larva on any patch, starting from the time it became active. Let the probability that the larva has not yet arrived on a patch at time t, $\Pr(T > t)$ be given by $A(t)$. Eq. 8 also describes the density of parasites that have not yet arrived on any patch, and so $A(t) = \int_{-\infty}^{\infty} p^j(x,t)\mathrm{d}x + P^j(t)$. To determine the relationship between $f^j(t)$ and $A(t)$, we can see that $\int_0^t f^j(\tau)\mathrm{d}\tau = \Pr(T < t)$, and $A(t) = \Pr(T > t)$, thus $\int_0^t f^j(\tau)\mathrm{d}\tau = 1 - A(t)$, or alternatively $f^j(t) = -\mathrm{d}A(t)/\mathrm{d}t$.

Integrating Eq. 8 over all space, adding Eq. 9, and substituting $f^j(t) = -\mathrm{d}A/\mathrm{d}t$ we find

$$f^j(t) = \sum_{i=1}^{n}\alpha_i\int_{x_i-\Delta/2}^{x_i+\Delta/2} p^j(x,t)\mathrm{d}x + \alpha_j^r P^j(t),$$

as we require $\lim_{x\to\pm\infty} p^j(x,t) = 0$ and $\lim_{x\to\pm\infty}\frac{\partial}{\partial x}p^j(x,t) = 0$ for Eq. 8 to be unique and well defined. We can then split the distribution for arrival time onto any patch, $f^j(t)$, into a sum of the arrival time distributions for each patch i. Let

$$f^{ij}(t) = \alpha_i\int_{x_i-\Delta/2}^{x_i+\Delta/2} p^j(x,t)\mathrm{d}x$$

be the distribution of arrival time for an active larvae produced from patch j arriving on patch i, for $i \neq j$. For $i = j$,

$$f^{jj}(t) = \alpha_j\int_{x_j-\Delta/2}^{x_j+\Delta/2} p^j(x,t)\mathrm{d}x + \alpha_r^j P^j(t).$$

Then, we can rewrite $f^j(t)$ as

$$f^j(t) = \sum_{i=1}^{n} f^{ij}(t).$$

Therefore, if we can solve for $p^j(x,t)$, we can solve for $f^j(t)$ and $f^{ij}(t)$. In the Applications section, we solve for the arrival time numerically. The approximate analytical solution of $f^{ij}(t)$ is the subject of a forthcoming paper.

Calculating the arrival time density function for larvae leaving one patch and arriving on another allows us to characterize the larval movement between a transmitting and receiving patch using a single function. This arrival time density function is incorporated into our full model to characterize larval dispersal between patches. In this section, we have presented the derivation for our full model (system 2) and demonstrated how larval dispersal between patches and growth and survival on patches determine metapopulation dynamics. We have also constructed a next-generation matrix for our model, which distils the essential information required to determine the source–sink distribution of our system.

3 Mathematical Analysis of the Model and Next-Generation Matrix

In this section, we present the details of the construction of the next-generation matrix, as well as proofs detailing the relationship between the stability of our model and the spectral radius of the next-generation matrix.

3.1 Constructing the Next-Generation Matrix

The next-generation matrix extracts the essential information of our model into a single operator. The elements of the matrix quantify the effect of an individual from one patch on other patches, and the column sums identify source and sink patches. Here, we present the details of the construction of the next-generation matrix, beginning with one patch and then abstracting to multiple patches.

3.1.1 The Next-Generation Matrix for a One-Patch System

To construct the next-generation matrix for a single patch, we need to calculate the number of new individuals produced on patch i from one initial individual on patch i in each stage. To calculate the number of new (stage $k = 2$) individuals produced on patch i by an initial individual at $t = 0$ starting on that same patch, we calculate the rate of production of new individuals at time t. We call this rate $\gamma(t)$. We then integrate $\gamma(t)$ over all t to calculate the total number of individuals produced.

For one initial individual in stage $k = 2$ to be producing new offspring, it must survive and mature through each stage and produce larvae which spread back to the patch. Let r_k be the time that the individual spends in stage k. For stages $k = 2, \ldots, m-1$ the probability that the individual survives to r_k in stage k is $S_k^i(r_k)M_k^i(r_k)$ and the

rate at which they mature to the next stage is $m_k^i(r_k)$. For stage $k = m$, the probability that they survive to r_m is $S_m^i(r_m)$ and the rate at which they are producing larvae is $b^i(r_m)$. The probability that latent larvae survive to $\bar{r}_1$ is $\bar{S}_1^i(\bar{r}_1)M_1^i(\bar{r}_1)$ and the rate at which they mature into active larvae is $m_1^i(\bar{r}_1)$. The probability that active larvae ($k = 1$) survive to r_1 is $S_1^i(r_1)(1 - F^i(r_1))$ and the rate at which they attach as $k = 2$ individuals is $f^{ii}(r_1)/(1 - F^i(r_1))$.

To calculate the rate of production at time t, $\gamma(t)$, we multiply the survival probabilities and maturation rates in each of the stages and integrate over all possible r_k. At time t, we must have $0 \le \bar{r}_1 + \sum_1^m r_k \le t$, so we can rewrite $r_m = t - \sum_1^{m-1} r_k - \bar{r}_1$ before integrating over all other possible r_k. We calculate

$$\gamma(t) = \int_0^t \int_0^{t-r_{m-1}} \int_0^{t-r_{m-1}-r_{m-2}} \cdots \int_0^{t-\sum_2^{m-1} r_k} S_2^i(r_2)M_2^i(r_2)m_2^i(r_2)\ldots$$
$$S_{m-1}^i(r_{m-1})M_{m-1}^i(r_{m-1})m_{m-1}^i(r_{m-1})S_m^i(t - \sum_1^{m-1} r_k - \bar{r}_1)b^i(t - \sum_1^{m-1} r_k - \bar{r}_1)$$
$$\times \bar{S}_1^i(\bar{r}_1)M_1^i(\bar{r}_1)m_1^i(\bar{r}_1)S_1^i(r_1)f^{ii}(r_1)d\bar{r}_1 dr_1 \ldots dr_{m-1}.$$

Then, integrating $\gamma(t)$ over all t and using the convolution theorem,

$$\int_0^\infty f(t) * g(t)\mathrm{d}t = \int_0^\infty f(t)\mathrm{d}t \int_0^\infty g(t)\mathrm{d}t$$

we calculate the number of new individuals on patch i produced from an initial individual in stage $k = 2$ on patch i to be:

$$\left(\prod_{k=2}^{m-1} \int_0^\infty S_k^i(t)M_k^i(t)m_k^i(t)\right)\left(\int_0^\infty S_m^i(t)b^i(t)\mathrm{d}t\right)$$
$$\left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right).$$

To calculate the next-generation matrix with large domain, K_L, (Diekmann et al. 2010), we can also calculate the number of new individuals ($k = 2$) produced on patch i, from initial individuals in the other stages. Repeating the process described above, we find that the number of new individuals produced on patch i from an initial individual in stage $k = l$, for $2 \le l \le m$, is

$$\left(\prod_{k=l}^{m-1} \int_0^\infty S_k^i(t)M_k^i(t)m_k^i(t)\right)\left(\int_0^\infty S_m^i(t)b^i(t)\mathrm{d}t\right)$$
$$\left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right).$$

We can repeat the same process for $k = 1$ (where now we group the latent and active larval stages for simplicity). The (i, j) entries of the next-generation matrix

with large domain, K_L, are the number of new (stage $k = 2$) individuals in stage i produced from an initial individual in stage j. Therefore,

$$\begin{aligned}
K_L(2,l) &= \left(\prod_{k=l}^{m-1}\int_0^\infty S_k^i(t)M_k^i(t)m_k^i(t)\right)\left(\int_0^\infty S_m^i(t)b^i(t)\mathrm{d}t\right)\\
&\quad\left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right) \quad 2 \le l \le m-1,\\
K_L(2,m) &= \left(\int_0^\infty S_m^i(t)b^i(t)\mathrm{d}t\right)\left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\\
&\quad\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right),\\
K_L(2,1) &= \left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right),\\
K_L(i,j) &= 0 \quad i \neq 2.
\end{aligned}$$

The next-generation matrix with large domain, K_L, can then be reduced to the next-generation matrix, K, through the process described by Diekmann et al. (2010). Essentially, we multiply K_L from the left and right by matrices which isolate the relevant entries where new individuals are produced from other new individuals. In this case, let E be an $n \times 1$ matrix with a 1 in row 2, and zeros elsewhere. Then, $K = E^T K_L E$. The entries of the next-generation matrix are the number of new individuals ($k = 2$) in stage i produced by an individual in stage j, though here we only include stages where new individuals can be produced. In our one patch example, because new individuals can only be produced in stage $k = 2$, our next-generation matrix, K, for one patch is simply the scalar:

$$\begin{aligned}
K &= \left(\prod_{k=2}^{m-1}\int_0^\infty S_k^i(t)M_k^i(t)m_k^i(t)\right)\left(\int_0^\infty S_m^i(t)b^i(t)\mathrm{d}t\right)\\
&\quad\left(\int_0^\infty \bar{S}_1^i(t)M_1^i(t)m_1^i(t)\right)\left(\int_0^\infty S_1^i(t)f^{ii}(t)\mathrm{d}t\right).
\end{aligned}$$

While the next-generation matrix, K, and next-generation matrix with large domain, K_L, have different sizes and different entries, their spectral radii are equal (Diekmann et al. 2010).

3.1.2 The Next-Generation Matrix for the Multiple-Patch System

Calculating the remaining entries of the next-generation matrix, K, for the multiple-patch model, using the same process as in the previous subsection, is relatively straightforward. In this case, we restrict ourselves to calculating K, and no longer K_L, though K_L can also easily be calculated in the same way as in the previous section. We want to calculate the number of new individuals produced on patch i from an

initial individual on patch j. In order for an individual to produce new individuals on patch i, it must first survive and mature on patch j and then produce larvae that successfully travel to patch i. The majority of $K(i, j)$ will be the same as $K(j, j)$, as the individual must mature on patch j before producing larvae. Only now, instead of the larvae travelling back to j, they must successfully spread to i. Therefore, the last multiplication factor which will now be $\int_0^\infty S_1^j(t) f^{ij}(t)\mathrm{d}t$, instead of $\int_0^\infty S_1^j(t) f^{jj}(t)\mathrm{d}t$. With this replacement, we can see that we have the same formula for $K(i, j)$ as given by (5). In the following sections we will also make the assumption that K is irreducible. Physically, this means that larvae have a positive probability of arriving on any patch when leaving from a given patch. Recall that $f^{ij}(t) = \alpha_i \int_{-\infty}^{\infty} p^j(x, t)\mathrm{d}x$, where $p^j(x, t)$ is the solution to the advection–diffusion equation from Sect. 2.4. The solution to the advection–diffusion equation is positive everywhere, so $p(x, t) > 0$, and thus $f^{ij}(t) > 0$ for all (i, j).

Here, we have presented the construction of the next-generation matrix for our stage-structured model on one patch as well as on multiple patches. We have explicitly shown how to reduce the next-generation matrix with large domain to the next-generation matrix on one patch, and we have demonstrated the process for multiple patches.

3.2 Model Stability Analysis

In this section, we will demonstrate that $R_0 = \rho(K)$ determines the stability of the zero equilibrium for the system (2).

3.2.1 Calculating the Model Equilibrium

First, we show that the zero equilibrium is the only equilibrium in our system. It should be noted here that system (2) is the solution to a linear system of age structured PDEs, and so we expect the zero equilibrium to be the only equilibrium. However, we include the details for completeness. Assume the system is at equilibrium, so that $n_k^i(t, a) = n_k^i(a)^*$, and $B_k^i(t) = B_k^{i*}$. Now we solve for B_1^{i*} using system (1):

$$
\begin{aligned}
B_1^{i*} &= \int_0^\infty \bar{n}_1(a)^* m_1^i(a)\mathrm{d}a \\
&= \int_0^\infty \bar{B}_1^{i*} \bar{S}_1^i(a) M_1^i(a) m_1^i(a)\mathrm{d}a \\
&= \int_0^\infty n_m^i(a)^* b^i(a)\mathrm{d}a \int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a)\mathrm{d}a \\
&= \int_0^\infty {B_m^i}^* S_m^i(a) b^i(a)\mathrm{d}a \int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a)\mathrm{d}a \\
&= \int_0^\infty n_{m-1}^i(a)^* m_{m-1}^i(a)\mathrm{d}a \int_0^\infty S_m^i(a) b^i(a)\mathrm{d}a \int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a)\mathrm{d}a \\
&= \int_0^\infty {B_{m-1}^i}^* S_{m-1}^i(a) M_{m-1}^i(a) m_{m-1}^i(a)\mathrm{d}a \int_0^\infty S_m^i(a) b^i(a)\mathrm{d}a
\end{aligned}
$$

$$
\begin{aligned}
&\int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= B_2^{i*} \left(\prod_{k=2}^{m-1} \int_0^\infty S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \left(\int_0^\infty S_m^i(a) b^i(a) \mathrm{d}a \right) \\
&\quad \left(\int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right) \\
&= \sum_{j=1}^n \int_0^\infty n_1^j(a)^* f^{ij}(a)/(1 - F^j(a)) \mathrm{d}a \left(\prod_{k=2}^{m-1} \int_0^\infty S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \\
&\quad \left(\int_0^\infty S_m^i(a) b^i(a) \mathrm{d}a \right) \left(\int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right) \\
&= \sum_{j=1}^n B_1^{j*} \left(\int_0^\infty S_1^j(a) f^{ij}(a) \mathrm{d}a \right) \left(\prod_{k=2}^{m-1} \int_0^\infty S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \\
&\quad \left(\int_0^\infty S_m^i(a) b^i(a) \mathrm{d}a \right) \left(\int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right).
\end{aligned}
$$

Let us define a matrix $\mathbf{S}$ entry-wise such that

$$
\begin{aligned}
S(i, j) = &\left(\int_0^\infty S_1^j(a) f^{ij}(a) \mathrm{d}a \right) \left(\prod_{k=2}^{m-1} \int_0^\infty S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \\
&\left(\int_0^\infty S_m^i(a) b^i(a) \mathrm{d}a \right) \left(\int_0^\infty \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right).
\end{aligned}
$$

Then, we can write the equations for B_1^{i*} for each farm i in matrix notation as

$$
B_1{}^* = \mathbf{S} B_1{}^*.
$$

This equation can only have a solution if $\det(\mathbf{S} - \mathbf{I}) = 0$. However, as we have general functions $f^{ij}(a)$, $S_k^i(a)$, and $M_k^i(a)$, we therefore require $B_1^* = 0$, from which we can recursively deduce that $B_k^* = 0$ for all k. Therefore, the zero equilibrium is the only equilibrium for this system.

3.2.2 Determining the Stability of the Equilibrium Using R_0

Next we prove that the spectral radius of the next-generation operator, $R_0 = \rho(K)$, determines the stability of the zero equilibrium of the full system (2). Again the result would be generally expected, based on the theory in Diekmann et al. (2010), but we include the details here for completeness.

Theorem 1 *Assume the next-generation matrix, K, is irreducible. Then,*

1. *if $R_0 < 1$, then the zero equilibrium of system 2 is globally stable.*

2. *if $R_0 > 1$, then the zero equilibrium of system 2 is unstable.*

Proof To analyse the stability of the zero equilibrium, we consider small perturbations to the equilibrium and examine their growth or decay. At the equilibrium, we have $B_k^i(t) = 0$ for all k, i. Consider a small perturbation of the form $B_k^i(t) = \bar{B}_k^i e^{\lambda t}$ to each of the $B_k^i(t)$. Similar to the calculation of the equilibrium, we will construct a recursive equation for $\bar{B}_1^i$ and then reformulate as a matrix equation for all patches. Using the equation for $B_k^i(t)$ in system (2), we find

$$
\begin{aligned}
\bar{B}_1^i &= e^{-\lambda t} B_1^i(t) \\
&= e^{-\lambda t} \int_0^\infty \bar{n}_1^i(t,a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \int_0^\infty \bar{B}_1^i(t-a) \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \int_0^\infty \bar{\bar{B}}_1^i e^{\lambda(t-a)} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= \bar{\bar{B}}_1^i \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \bar{B}_1^i(t) \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \int_0^\infty n_m^i(t,a) b^i(a) da \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \int_0^\infty B_m^i(t-a) S_m^i(a) b^i(a) \mathrm{d}a \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= e^{-\lambda t} \int_0^\infty \bar{B}_m^i e^{\lambda(t-a)} S_m^i(a) b^i(a) \mathrm{d}a \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= \bar{B}_m^i \int_0^\infty e^{-\lambda a} S_m^i(a) b^i(a) \mathrm{d}a \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \\
&= \bar{B}_2^i \left(\prod_{k=2}^{m-1} \int_0^\infty e^{-\lambda a} S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \\
&\quad \left(\int_0^\infty e^{-\lambda a} S_m^i(a) b^i(a) \mathrm{d}a \right) \left(\int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right) \\
&= \sum_{j=1}^n \bar{B}_1^j \left(\int_0^\infty e^{-\lambda a} S_1^j(a) f^{ij}(a) \mathrm{d}a \right) \left(\prod_{k=2}^{m-1} \int_0^\infty e^{-\lambda a} S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a \right) \\
&\quad \left(\int_0^\infty e^{-\lambda a} S_m^i(a) b^i(a) \mathrm{d}a \right) \left(\int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a \right).
\end{aligned}
$$

Define

$$\mathbf{L}(\lambda) = \begin{bmatrix} L_1^{11}(\lambda)\bar{L}_1^1(\lambda)\prod_{k=2}^m L_k^1(\lambda) & \dots & L_1^{1n}(\lambda)\bar{L}_1^1(\lambda)\prod_{k=2}^m L_k^1(\lambda) \\ \vdots & \ddots & \vdots \\ L_1^{n1}(\lambda)\bar{L}_1^n(\lambda)\prod_{k=2}^m L_k^n(\lambda) & \dots & L_1^{nn}(\lambda)\bar{L}_1^n(\lambda)\prod_{k=2}^m L_k^n(\lambda) \end{bmatrix}$$

with

$$L_k^i(\lambda) = \begin{cases} \int_0^\infty e^{-\lambda a} S_k^i(a) M_k^i(a) m_k^i(a) \mathrm{d}a & k = 2, \dots, m-1 \\ \int_0^\infty e^{-\lambda a} S_m^i(a) b^i(a) \mathrm{d}a & k = m \end{cases}$$

$$\bar{L}_1^i(\lambda) = \int_0^\infty e^{-\lambda a} \bar{S}_1^i(a) M_1^i(a) m_1^i(a) \mathrm{d}a$$

$$L_1^{ij}(\lambda) = \int_0^\infty e^{-\lambda a} S_1^{ij}(a) f^{ij}(a) \mathrm{d}a.$$

Then, again we can write the equations for each $\bar{B}_1^i$ for each patch i in matrix notation as

$$(\mathbf{L}(\lambda) - \mathbf{I})\bar{B}_1 = 0.$$

We are looking for non-trivial solutions where $\bar{B}_1 \neq 0$, and therefore require

$$\det(\mathbf{L}(\lambda) - \mathbf{I}) = 0. \tag{10}$$

This is the characteristic equation for our system. If the root λ satisfies $\Re(\lambda) < 0$, then the zero equilibrium is stable, and if $\Re(\lambda) > 0$, then the zero equilibrium is unstable. Furthermore, as our system 2 is linear, if the equilibrium is locally stable, it will be globally stable.

We know that because λ is a root of (10), then $1 \in \sigma(\mathbf{L}(\lambda))$, where $\sigma(\mathbf{L}(\lambda))$ is the spectrum of $\mathbf{L}(\lambda)$. We also know from the definition of system 2 that $\bar{B}_1$ must be non-negative, and from (10) that $\bar{B}_1$ is the eigenvector associated with an eigenvalue of 1. $\mathbf{L}(\lambda)$ is irreducible, because K is irreducible, and the eigenvalue associated with a non-negative eigenvector of an irreducible matrix is the spectral radius of the matrix (Theorem 2.1, Li and Schneider (2002)). Thus, $\rho(\mathbf{L}(\lambda)) = 1$.

(1) If $R_0 < 1$, then $\rho(\mathbf{L}(0)) = \rho(K) = R_0 < 1$. If $\Re(\lambda) > 0$, then $\mathbf{L}(0) \geq \mathbf{L}(\lambda)$ entry-wise, and so $\rho(\mathbf{L}(\lambda)) \leq \rho(\mathbf{L}(0)) < 1$ (Corollary 8.1.19, Horn and Johnson (2012)). Therefore, in order for $\rho(\mathbf{L}(\lambda)) = 1$ we require $\Re(\lambda) < 0$.
(2) If $R_0 > 1$, then $\rho(\mathbf{L}(0)) > 1$. If $\Re(\lambda) < 0$, then $\mathbf{L}(0) \leq \mathbf{L}(\lambda)$ entry-wise and so $\rho(\mathbf{L}(\lambda)) \geq \rho(\mathbf{L}(0)) > 1$. Therefore, in order for $\rho(\mathbf{L}(\lambda)) = 1$ we require $\Re(\lambda) > 0$. □

Corollary 1 *If the number of new individuals produced on a given patch k, from an initial individual starting on patch k, is greater than one, so $K(k,k) > 1$, then $R_0 > 1$.*

Proof If $K(k,k) > 1$, then we can decompose $K = A + B$, where

$$B(i,j) = \begin{cases} 1 & i = j = k \\ 0 & \text{otherwise,} \end{cases}$$

and A is still non-negative. Because A is non-negative, we can see that $K > B$, and so by Corollary 8.1.19 Horn and Johnson (2012), $\rho(K) > \rho(B) = 1$. □

Here, we have demonstrated that $R_0 = \rho(K)$ determines the stability of the zero equilibrium for system 2. Earlier in this section, we presented the details of the construction of the next-generation matrix, K, for the stage-structured population on both one and multiple habitat patches.

4 Applications

In this section, we discuss applications of the next-generation matrix and the effect of different environmental variables on the next-generation matrix. The next-generation matrix is a useful tool to quantify the effect of individuals from one patch on other patches. Here, we will show how to use the column sums in the matrix to determine the source–sink dynamics of the network and how this relates to R_0. We will also show how the column sums can be used to investigate the transient dynamics of the system and how these may differ from the asymptotic dynamics. We then investigate how the left and right eigenvectors can provide insight into the contributions of each patch to R_0. Using salmon farms as a motivating example, we structure the Applications section to answer the key questions posed at the end of the introduction.

4.1 Salmon Farms Distributed in a Channel

Here, we present an example of patches in a linear array and in the following sections will demonstrate how both the source–sink distribution and persistence measures change as a function of the distance between patches. This example is motivated by sea lice spreading between salmon farms in a channel. Salmon farms act as habitat patches for sea lice, as the non-larval sea lice stages require a salmonid host on which to feed, and the salmon are themselves confined to the net pens in the farms. The larval sea lice stages are released into the water column and are capable of spreading between salmon farms in a given region. In both Norway and Canada, many salmon farms are located in sheltered coastal channels or fjords (Aldrin et al. 2017; Krkošek et al. 2006). We therefore use a one-dimensional domain to calculate the arrival time of sea lice spreading between farms.

In the following examples, we consider 5 patches or farms of width Δ, each separated by some distance x_0 (Fig. 4). We consider systems where there is no maturation delay from the latent to active larval stage, and where all larvae enter the channel, so that none remain locally. The absorption rate for larvae when they pass by a patch is small, and in most examples, diffusion is larger than advection. This represents a

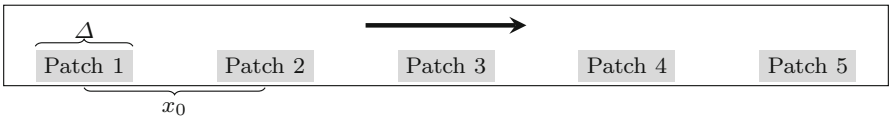


Fig. 4 Patches or salmon farms arranged in a channel. The width of each patch is Δ, and the distance between the centre of each patch is x_0. The arrow above the patches indicates the direction of advection

coastal channel environment in which ocean mixing is more prevalent than any constant currents generated by river outflow. It is in this environment that we answer the five questions posed at the end of the introduction.

4.2 What Is the Source–Sink Distribution of Salmon Farms in a Channel?

First, we show how to use the next-generation matrix to determine the source–sink distribution of farms. Recent work has used R_0 theory to define two measures of persistence of a species on a continuous landscape, $R_{\text{loc}}(x)$ and $R_\delta(x)$, using the next-generation operator (Huang et al. 2016b; Krkošek and Lewis 2010; McKenzie et al. 2012). $R_{\text{loc}}(x)$ is the number of new individuals produced at location x in the absence of dispersal and can be used as a measure of the fundamental niche in certain scenarios. In our model, dispersal is a key environmental feature for the larval stage of the marine organism, and while we allow some percentage of larvae to remain at a patch, it is not realistic that any large percentage would remain and avoid dispersal. Therefore, $R_{\text{loc}}(x)$ is not relevant for our model.

The second persistence measure, $R_\delta(x)$, is the number of new individuals produced over the entire network from one individual at location x. It takes into account both growth and survival at location x, and dispersal from location x. If one individual at x produces less than one individual over the entire landscape, then $R_\delta(x) < 1$, and the location x is defined as a sink. If $R_\delta(x) > 1$, then x is defined as a source. The spectral radius of the next-generation operator, R_0, determines species persistence over the entire landscape. When using this measure, it is not possible for connected patches of sinks to persist. In this section, we build on and apply this theory to determine the source–sink distribution on the discrete population patches in our system using the next-generation matrix.

In our system of n patches, let $R_\delta(j)$ be the number of new individuals on all patches produced from one individual on patch j. In terms of our next-generation matrix K (Eq. 5),

$$R_\delta(j) = \sum_{i=1}^{n} K(i, j).$$

If for patch j, $R_\delta(j) > 1$, then j is a source, and if $R_\delta(j) < 1$, then j is a sink. $R_0 = \rho(K)$ is still needed to determine if the populations on the connected patches will persist or perish, however there are some nice persistence results that follow directly from $R_\delta(j)$.

First, a connected network consisting only of sinks cannot persist. Using this new measure, for a connected network of sinks, we have $R_\delta(j) < 1$ for all j. There is a nice result concerning R_0 for non-negative irreducible matrices (Horn and Johnson 2012) which states that

$$\min_{1\le j\le n} \sum_{i=1}^{n} K(i, j) \le \rho(K) \le \max_{1\le j\le n} \sum_{i=1}^{n} K(i, j).$$

Substituting the definition of R_δ and R_0 this can be restated as

$$\min_{1\le j\le n} R_\delta(j) \le R_0 \le \max_{1\le j\le n} R_\delta(j). \tag{11}$$

Therefore, if $R_\delta(j) < 1$ for all j, then $R_0 < 1$ as well. Secondly, a connected network consisting only of sources cannot perish. Here, $R_\delta(j) > 1$ for all j, and so $R_0 > 1$ as well. Similarly, if any diagonal entry of the next-generation matrix is greater than one, $K(j, j) > 1$, then $R_0 > 1$. This can be seen from Corollary 1. Biologically, this result means a population on a network will persist if the network contains at least one patch which is self-sustaining in the absence of dispersal from other patches. However, there are also other situations in which the network can persist. In summary, $R_\delta(j)$, the jth column sum of the next-generation matrix, is necessary to determine whether a patch j is a source or a sink, and R_0, the spectral radius of the next-generation matrix, is necessary to determine the persistence of the total population.

4.3 How Does the Source–Sink Distribution Change with Respect to Environmental Variables?

In this section, we examine the effect of advection and diffusion on R_0 and R_δ, using the example of 5 patches in a channel, shown in Fig. 4. First, we study the effect of varying diffusion on R_0 across different interfarm separation distances, x_0. In Fig. 5, the diffusion coefficient, D, is decreased from 5 to 1 and the change in R_0 is shown as a function of x_0. For small values of x_0, R_0 is larger when there is less diffusion. When there is less diffusion, each patch has a greater probability of self-infecting, because there is less immediate dispersal away from the patch. When the patches are overlapping, they act as one patch, which is why R_0 is larger at small x_0 when there is less diffusion. This is also why the horizontal asymptote for R_0 is larger for smaller diffusion. For intermediate values of x_0, an interesting phenomenon occurs. As the separation distance, x_0 increases, R_0 for low diffusion drops below R_0 for high diffusion. When diffusion is low, it is more difficult for individuals to disperse against the direction of advection, and so as x_0 increases, individuals from Patch 5 begin to only contribute to other individuals on Patch 5. As will be shown when examining the effect of advection, Patch 5 becomes a sink for small values of x_0, and so the other patches are contributing individuals to Patch 5, which cannot sustain them.

Next, we study the effect that advection has on R_0 and R_δ across different interfarm distances, x_0. For different values of advection, both R_0 and R_δ, as functions of x_0, have the same shape as R_0 shown in Fig. 5. Therefore, we find it most illuminating

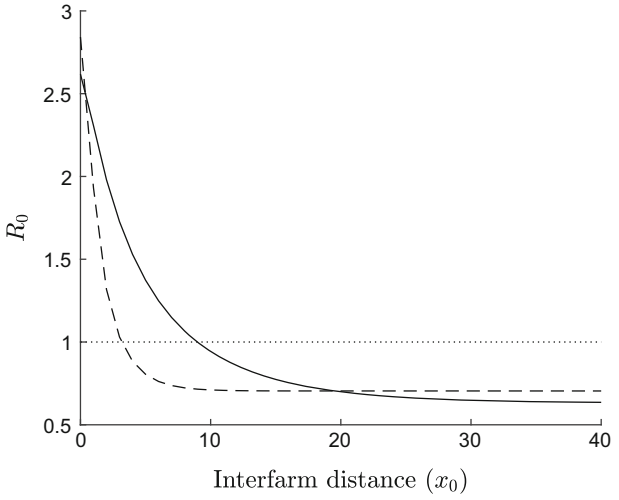
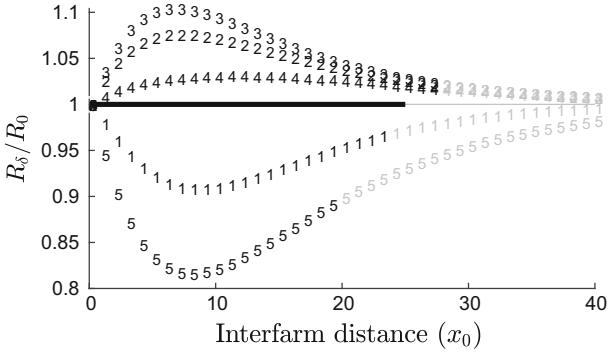


Fig. 5 R_0 as a function of x_0, for $D = 5, v = 1$, solid line; $D = 1, v = 1$, dashed line. The remaining parameter values are $\alpha = 0.1, g^{ij}(t) = 0, \Delta = 0.8, S^{ij}(t) = e^{-0.05t}$. The survival, maturation, and birth functions for the sessile stages were combined so that $\prod_{k=2}^{m-1}\left(\int_0^\infty S_2^j(t)M_2^j(t)m_2^j(t)\mathrm{d}t\right)\left(\int_0^\infty S_m^j(t)b^j(t)\mathrm{d}t\right) = 10$

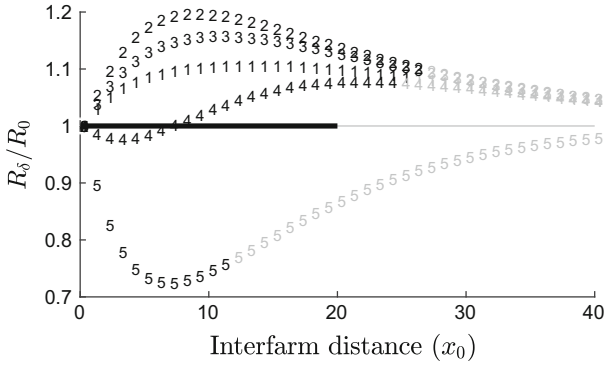
to examine the effect of advection on the ratio of R_δ/R_0. Biologically, R_0 is the number of new individuals produced in the population, from one typical individual, and $R_\delta(j)$ is the number of new individuals produced in the population, from one individual starting on patch j. Therefore, the ratio $R_\delta(j)/R_0$ can be seen as the relative multiplication factor of the number of new individuals in the population produced by one individual starting on patch j, compared to one typical individual. If $R_\delta(j)/R_0 > 1$, then an individual on patch j is contributing more than the typical individual, and if $R_\delta(j)/R_0 < 1$, then it is contributing less. Of course, it is also important to also track if each R_δ and R_0 are greater or less than one, so that it is known which patches are sources, which are sinks, and whether or not the total population is growing.

In Fig. 6, R_δ/R_0 is plotted as a function of the interfarm separation distance, x_0, for different values of v. The switch in each curve from black to grey marks where $R_\delta \leq 1$. Using the R_δ measure, we can see that for low x_0 all patches are sources. As x_0 increases, Patch 5 becomes a sink, with $R_0 > 1$, and then as x_0 continues to increase, R_0 falls below 1. Therefore, in this linear array, there is some critical separation distance, beyond which the population patches are not sufficiently connected for the total population to persist. Even after this critical distance, some patches are still sources with $R_\delta > 1$, and it takes a larger separation distance x_0 for all patches to become sinks.

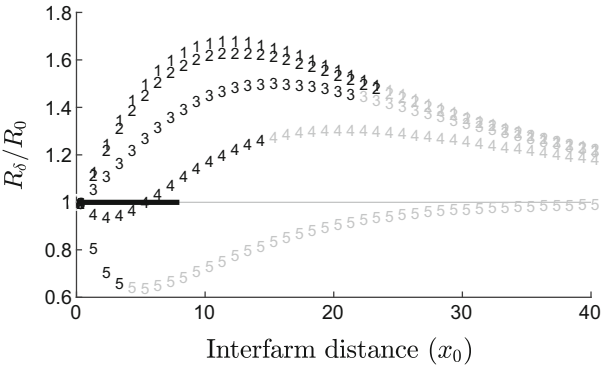
Both R_0 and R_δ change when v increases from 0.1 to 1. As v increases, it takes a much smaller separation distance, x_0, for R_0 to fall below 1. Increasing v not only reduces retention of individuals on each patch, but also inhibits individuals from better dispersing against the direction of advection. Therefore, when v is lowered, there is greater dispersal among neighbouring patches, but less long distance dispersal in the direction of advection, from Patch 1 to Patch 5. There are two other interesting



(a) Low advection $(v = 0.1)$



(b) Medium advection $(v = 0.4)$



(c) High advection $(v = 1)$

Fig. 6 R_δ/R_0 for each patch when $D = 5$. When each curve is black $R_\delta > 1$, and when the curve is grey $R_\delta \leq 1$. The switch from black to grey on the solid line indicates when $R_0 = 1$. The remaining parameters are the same as in Fig. 5

behaviours that should also be highlighted. First is that $\max_{1\leq i\leq n} R_\delta(i)$ is achieved at Patch 3 when advection is low, compared to Patch 1 when advection is high. We can see the transition as Patch 1 becomes a larger source as v increases. Second, the critical separation distances, for which each R_δ falls below one and becomes a sink, come closer together, as the decrease in advection makes the behaviour on all patches more similar.

In both examples where diffusion and advection were changed, for large separation distances, $R_0 < 1$. However, this is due to the parameters controlling the birth, survival, and maturation functions for these examples. For other parameter values, certain values of v and D could result in $R_0 > 1$ for large x_0 and some could result in $R_0 < 1$. Here, we have shown how R_0 and the source–sink distribution, quantified by R_δ, change as a function of the diffusion and advection in the system, as well as the interfarm separation distance, x_0.

4.4 Are There Certain Parameter Regions in Which Local Outbreaks Can Occur, But Not Global Outbreaks?

Interesting transient dynamics can occur if we consider networks of patches with $R_0 < 1$, but where some patches are sources, and networks with $R_0 > 1$, but where some patches are sinks. In these network arrangements, the value of R_0 still determines the global persistence of the total population, but the initial conditions determine whether the population will begin by increasing or decreasing. To consider these dynamics, we let $N(g)$ be the vector of sea lice populations on each patch, in generation g. Then, in terms of generational time, the population will update according to

$$N(g+1) = KN(g),$$

with the initial condition N_0, which is a vector of the initial sea lice populations on each patch. For example, consider the next-generation matrix K with parameter values that are the same as for Fig. 6c. At $x_0 = 8$, $R_0 > 1$, but $R_\delta(5) < 1$. Therefore, the total population ($\sum_i N_i(g)$) will increase eventually, but will start by decreasing if our initial population is all in Patch 5 ($N_0 = [0\ 0\ 0\ 0\ 1]^T$). In fact, for $x_0 = 8$, if we begin with 1 individual in Patch 5, it takes 23 generations before the total population increases above 1, as shown in Fig. 7. Similarly, for $x_0 = 10$, $R_0 < 1$ but $R_\delta(1) > 1$. In this case, the population will eventually decrease, but will begin by increasing if the initial population is in Patch 1 ($N_0 = [1\ 0\ 0\ 0\ 0]^T$). If we start with 1 individual in Patch 1, it takes 41 generations before the total population falls below 1. In this configuration, this means that there would be a local outbreak, but not a global sea lice outbreak.

To attempt to quantify the effect of the source and sink patches on transient dynamics more formally, we use notation from Huang and Lewis (2015). They define

$$R_l = \min_{1\leq j\leq n} \sum_{i=1}^{n} K(i,\, j) = \min_{1\leq j\leq n} R_\delta(j),$$

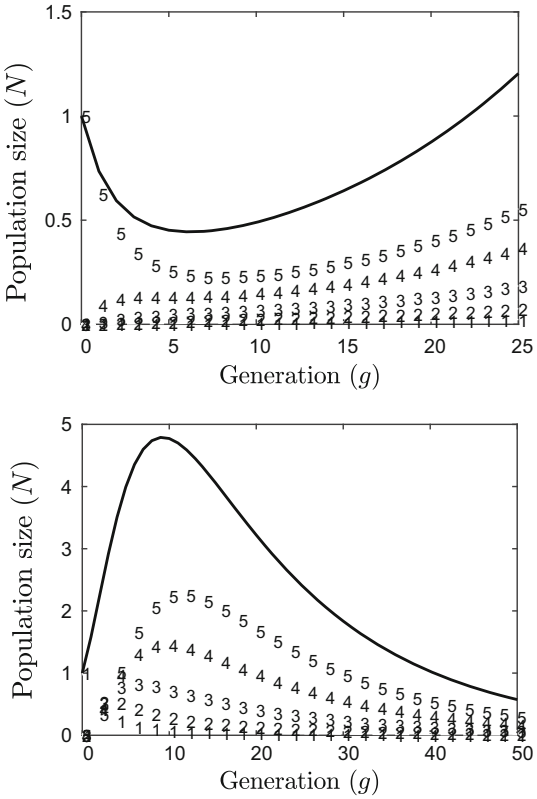


Fig. 7 The population of sea lice ($N_i(g)$) on each patch i in each generation (g) after one initial individual is released on a patch. Let $N(g)$ be a vector of patch populations, then in generational time the population updates according to $N(g+1) = KN(g)$, with the initial condition N_0, which is a vector of the initial sea lice populations on each patch. The black line shows the total population size ($\sum_i N_i(g)$). Parameter values are the same as Fig. 6. In a), we fix $x_0 = 8$, so that $R_0 > 1$, and release the initial individual on patch 5 ($N_0 = [0\ 0\ 0\ 0\ 1]^T$). In b), we fix $x_0 = 10$, so that $R_0 < 1$, and release the initial individual on patch 1 ($N_0 = [1\ 0\ 0\ 0\ 0]^T$)

which is shown to be the intergenerational growth rate under the worst possible initial conditions, and

$$R_u = \max_{1\le j\le n} \sum_{i=1}^{n} K(i,\, j) = \max_{1\le j\le n} R_\delta(j),$$

which is shown to be the intergenerational growth rate under the best possible initial conditions (Huang and Lewis 2015). Eq. 11 can then be restated using R_l and R_u as

$$R_l \le R_0 \le R_u.$$

In essence, R_l is the growth rate in the first generation if our population is initially at the worst sink patch, and R_u is the growth rate if we are at the best source patch. Therefore, R_l and R_u are useful measures to quantify potential transient dynamics and also retain key information about the source–sink distribution. If $R_l < 1$, then there is at least one patch acting as a sink, and if $R_u > 1$, then there is at least one patch acting as a source. In the following section, we examine how R_l, R_u, and R_0 change with different variables, instead of considering R_δ for every patch.

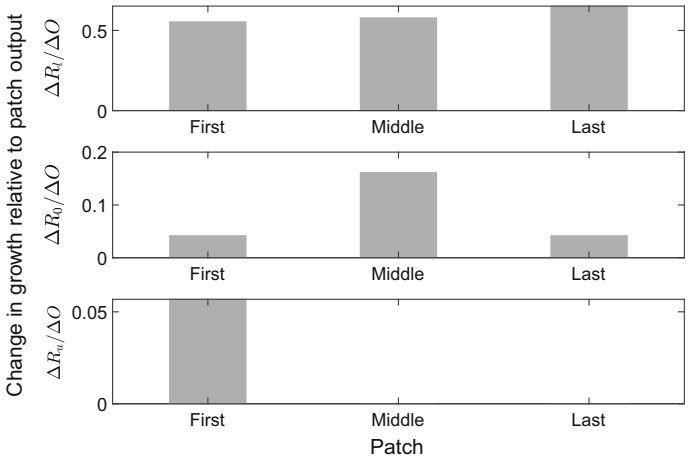


Fig. 8 The change in R_l, R_0, and R_u when the output from the First Patch (Patch 1), Middle Patch (Patch 3) or Last Patch (Patch 5) is reduced from 10 to 1. Parameter values for this figure are $\alpha = 0.1$, $D = 5$, $v = 1$, $g^{ij}(t) = 0$, $\Delta = 0.8$, $S^{ij}(t) = e^{-0.05t}$, $x_0 = 8.4$. The survival, maturation, and birth functions for the sessile stages were combined so that $O = \prod_{k=2}^{m-1}\left(\int_0^\infty S_2^j(t)M_2^j(t)m_2^j(t)\mathrm{d}t\right)\left(\int_0^\infty S_m^j(t)b^j(t)\mathrm{d}t\right) = 10$, and for the reduced patch $O = \prod_{k=2}^{m-1}\left(\int_0^\infty S_2^j(t)M_2^j(t)m_2^j(t)\mathrm{d}t\right)\left(\int_0^\infty S_m^j(t)b^j(t)\mathrm{d}t\right) = 1$

4.5 What Is the Effect of Treating a Single Farm on the Transient and Asymptotic Dynamics?

In this section, we examine the effect that treating specific farms have on R_l, R_u, and R_0. We define treatment as reduced survival and maturation on a patch or farm, affecting stages k through m, but not affecting the larval stage. On salmon farms, treatment is used to reduce sea lice levels and is typically administered orally to farmed salmon (Rogers et al. 2013). Reduced survival and maturation could also be the result of poor environmental conditions at a patch, such as low salinity and temperature in the case of sea lice. We also view treatment through the lens of type reproduction numbers, and the effort required for control on a patch to reduce $R_0 = 1$.

To first examine the effect of treatment on the transient and asymptotic dynamics, we treat Patch 1, Patch 3, and Patch 5 separately. In the direction of advection, these patches are the first, middle, and last patches, respectively (Fig. 4). Figure 8 shows the change in R_l, R_u, and R_0 for the system when either the first, middle, or last patch has a reduced output (ΔO), from treatment. Perhaps the most interesting result is that if either the first or last patch has reduced output, the change in the R_0 value remains same. The first patch is the patch that produces the most individuals on other patches, and the last patch is the patch that receives the most individuals from other patches. If we consider the reduced output as treatment, then if we treat either the patch that produces the most individuals, or the patch that receives the most individuals, the effect on R_0 will be the same. However, if we treat the middle patch, we have a larger change in R_0. Therefore, treating the middle patch is more effective if we want to reduce long-term population growth.

If treating the first or last patch has the same effect on R_0, then how might it change the transient dynamics of the system? We can see that if we treat the first patch, then the change in R_u is larger than if we treat either the middle or last patch, where there is no change. Therefore, if we treat the first patch, the maximum possible growth rate is reduced, and thus, we can reduce the severity of a local outbreak. However, if we treat the first patch, the change in R_l is less than if we treat either the middle or last patch. Therefore, treating the first patch results in a larger minimum possible growth rate.

We can also examine the effect of treatment using the type reproduction number (Heesterbeek and Roberts 2007; Lewis et al. 2019; Roberts and Heesterbeek 2003). The type reproduction number measures the control effort required to control a certain patch to reduce $R_0 = 1$. The type reproduction number can also been generalized to the target reproduction number, if control is not applied to an entire patch, but to specific inter patch infections. To define the type reproduction number, we divide the next-generation matrix $K = [k_{ij}]$ into two matrices, $K = B + C$, where C is the target matrix associated with control and B is the residual matrix. If we are interested in controlling patch i, then $C_i = [c_{ij}]$, with $c_{ij} = k_{ij}$ for $1 \leq i \leq n$ and $j = i$, and $c_{ij} = 0$ otherwise. Then, $B = K - C$. To control patch i such that we can reduce $R_0 = 1$, we require $\rho(B) < 1$. The type reproduction number T_{C_i} is then defined by $\rho(C_i(I - B)^{-1})$, and $1 - 1/T_{C_i}$ is the fraction that output from patch i must be reduced in order for $R_0 = 1$. The controlled next-generation matrix would then be $\frac{1}{T_{C_i}} C_i + (K - C_i)$ and would have $\rho(K) = 1$.

Consider the patch arrangement and parameter values as shown in Fig. 8. In this case $R_0 > 1$, but when we are considering control of either the first, middle, or last patch, and creating target matrices C, we still have $\rho(B) < 1$. To create the target matrix C_1 associated with control on the first patch, we take the first column of K as the first column of C_1, and put zeros in all other entries. Likewise to treat the middle patch, we take the third column of K as the third column of C_3, with zeros elsewhere, and to treat the last patch we take the last column of K as the last column of C_5 with zeros elsewhere. Using the formula for the type reproduction number given above, we calculate $T_{C_1} = T_{C_5} = 3.6$, and $T_{C_3} = 1.2$. This demonstrates that it requires a greater control effort to treat the first or last patch, to reduce $R_0 = 1$, than is required if treating the middle patch.

Here, we have shown that either using the type reproduction number or directly examining the effect of treatment, that more treatment is required to control the first or last patch than the middle patch. Moreover, treating the first patch produces the largest change in R_u, the largest intergeneration growth rate, and treating the last patch produces the largest change in R_l, the minimum intergenerational growth rate.

4.6 What Is the Effect of an Environmental Gradient on Patch Contributions to R_0 and the Source–Sink Distribution?

A different approach to measuring the contribution that each can patch have on the population is to use the right and left eigenvectors associated with R_0 (Hurford et al. 2010). This approach measures the contributions that each patch has on R_0 if the

population is proportioned relative to the right eigenvector of the next-generation matrix.

First, from the Perron–Frobenius theorem, we know that the eigenvalue for which the spectral radius of K is achieved is real. Then, we can write this eigenvalue, R_0, using the left eigenvector, w, and the right eigenvector, v, as

$$R_0 = \frac{w^T K v}{w^T u}.$$

If we rescale w and v so that $w^T v = 1$, then we can rewrite R_0 as

$$R_0 = v_1 \sum_{i=1}^{n} K(i,1)w_i + v_2 \sum_{i=1}^{n} K(i,2)w_i + \cdots + v_n \sum_{i=1}^{n} K(i,n)w_i. \tag{12}$$

Equation 12 can then be interpreted as the sum of the contributions of each patch to R_0, when the population is proportioned relative to the right eigenvector (Hurford et al. 2010). If we look at the first term in this sum, which is the contribution of patch 1, v_1 is the relative proportion of the population that is in patch 1. This proportion, v_1, is then multiplied by $\sum_{i=1}^{n} K(i,1)w_i$, where each $K(i,1)$ is the number of new individuals produced on patch i from one individual on patch 1, and w_i is the reproductive value of patch i. Therefore, the first term in Eq. 12 measures the proportion of the population that is in patch 1 multiplied by the effect that individuals from patch 1 will have on future growth after they give birth to larvae which disperse to other patches. For the ease of future reference, we define

$$R_c(j) = v_j \sum_{i=1}^{n} K(i,j)w_i$$

as the contribution of patch j to R_0. We can therefore rewrite Eq. 12 as

$$R_0 = \sum_{j=1}^{n} R_c(j) \tag{13}$$

Interestingly, the relative source–sink measure of a patch, $R_\delta(j)$, can be very different from the relative contribution measure, $R_c(j)$. In certain instances, for a given patch j, we can have $R_\delta(j) = \min_{1\le i\le n} R_\delta(i) < 1$, so that patch j is the largest sink in the population. However, that same patch j, may have $R_c(j) = \max_{1\le i\le n} R_c(i)$, with $R_0 > 1$, so that if the population is distributed according to v, patch j has the largest contribution to R_0. R_0 is greater than 1, so the population is growing. This means that in one generation, one individual from patch j is contributing the least to the total population, but over several generations, patch j is having the largest contribution to the total growth of the population.

As an example, we examine how the intergenerational growth measures change if we put the patches in an environmental gradient. Salmon farms are often located

along ocean channels, where rivers feed into the source of these channels. This creates a salinity gradient along the channel, where farms located closest to the river have the lowest salinity and farms furthest from the river have the highest salinity. Lower salinity results in a reduction in sea lice survival at each stage (Johnson and Albright 1991). Often the river output at the source of the channel is also the source of advection in the channel, though there may be systems in which the average advection is in the opposite direction due to strong ocean currents. First, we consider the less likely case, where the first patch (or farm) has the largest output, and the last patch has the least. In this case, output decreases in the direction of advection. Next we will consider the more realistic case where output increases in the direction of advection. This would be the common case for salmon farms in a channel where the first farm is located closest to the river output, which is where salinity is lowest and so the patch output is also the lowest, and the river is the source of advection in the system. For reference, the average output from all patches is the same as the output for a single patch when there is no gradient.

We use R_δ/R_0 and R_c/R_0 as the intergenerational growth measures, for which we examine the effect of an environmental gradient. While R_δ/R_0 can be thought of as the relative multiplication factor of the number of new individuals in the population produced by one individual on patch j, compared to a typical individual, R_c/R_0 is simply the relative contribution of the patch to R_0, within the framework of left and right eigenvectors. In Fig. 9, both R_δ/R_0 and R_c/R_0 are plotted as a function of x_0, for an environmental gradient in the direction of advection and for a gradient in the opposite direction.

First, what cannot be seen easily from Fig. 9 is that R_0 is the same when patch output increases or decreases in the direction of advection. Interestingly, both of these R_0 values are larger than for patches without any gradient. When comparing R_δ/R_0 values, the relative spread of R_δ/R_0 is larger when patch output decreases in the direction of advection. Here, Patch 1 has the largest R_δ/R_0 value. Interestingly, when we look at R_c/R_0, the relative contribution to R_0, we can see that there is an intermediate distance x_0 where if the population is distributed relative to the right eigenvector, Patch 2 would be contributing more to R_0 than Patch 1, even though Patch 1 is the larger source.

We observe even more interesting behaviour when we look at R_δ/R_0 and R_c/R_0 when patch output increases in the direction of advection. Here, the relative ordering of $R_\delta(j)/R_0$ values changes for different values of x_0. Patches 3 and 4 are the largest sources for most x_0, until they become sinks (when numbers switch from black to grey). Patch 5 starts as the largest source when $x_0 = 0$, but is the first to become a sink (along with Patch 1), around $x_0 = 9$. However, if we look at the plot of R_c/R_0, then for some x_0, Patch 5 is the largest contributor to R_0. In fact, at $x_0 = 9$, Patch 5 is a sink with the smallest R_δ, $R_\delta(5) = \min_{1\le i\le 5} R_\delta(i) < 1$. Therefore, in the absence of dispersal from other patches, the population on Patch 5 would perish. However, for $x_0 = 9$, $R_0 > 1$ and $R_c(5) = \max_{1\le i\le 5} R_c(i)$. Therefore, the population is growing, and moreover, Patch 5 would be the largest contributor to growth if the population was distributed according the right eigenvector.

What we can also see from Fig. 9a, b is that when the patch output decreases in the direction of advection, the relative ordering of the R_δ/R_0 values of the patches is the same as when there is no gradient (Fig. 6c). However, when patch output increases

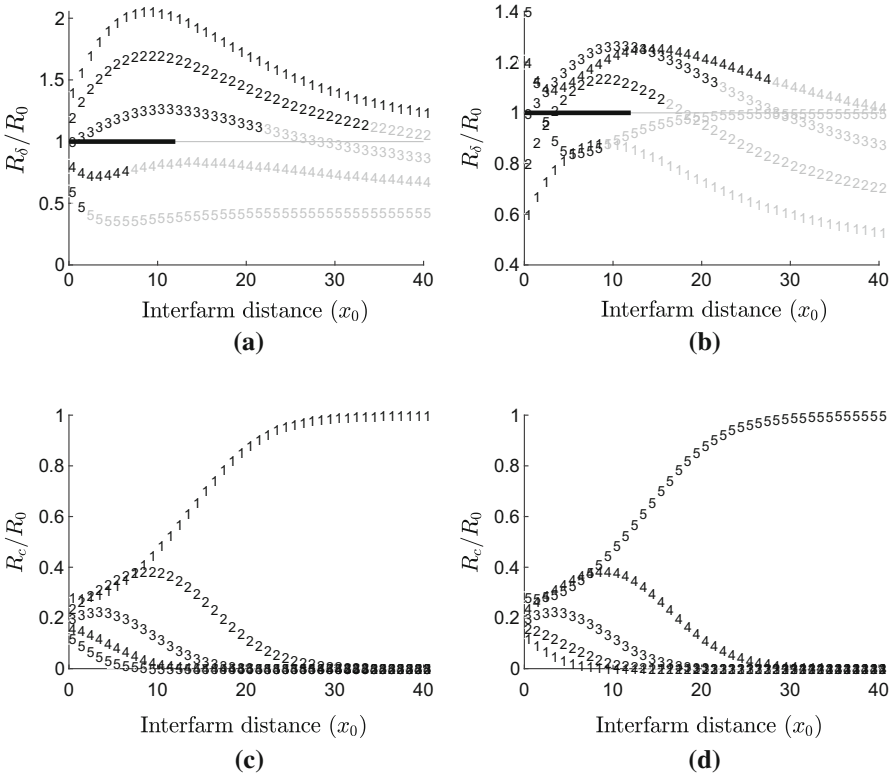


Fig. 9 R_δ/R_0 and R_c/R_0 are shown as a function of x_0. **a**, **c** are when the output decreases in the direction of advection, and **b**, **d** are when output increases. For **a**, **b**, when each curve is black $R_\delta > 1$, and when the curve is grey $R_\delta \leq 1$. The switch from black to grey on the solid line indicates when $R_0 = 1$. Parameter values for this figure are $\alpha = 0.1$, $D = 5$, $v = 1$, $g^{ij}(t) = 0$, $\Delta = 0.8$, $S^{ij}(t) = e^{-0.05t}$. The survival, maturation, and birth functions for the sessile stages were combined so that in the case of constant patch output, the output is $O = \prod_{k=2}^{m-1}\left(\int_0^\infty S_2^j(t)M_2^j(t)m_2^j(t)\mathrm{d}t\right)\left(\int_0^\infty S_m^j(t)b^j(t)\mathrm{d}t\right) = 10$. To construct the environmental gradient, the largest output was $1.4 \times O$, then $1.2 \times O$, then O, then $0.8 \times O$, and then $0.6 \times O$

in the direction of advection, the relative ordering of R_δ depends on the interfarm separation distance x_0. In this case, the farm with the largest output is not the largest source, nor is the farm with the lowest output the largest sink. Here, knowing the local environmental conditions that determine sessile output does not directly inform the source–sink distribution of the patch network.

In the Applications section, we have demonstrated how the next-generation matrix can be used to determine the source–sink distribution of a metapopulation, as well as other persistence measures, each of which quantifies some useful information about our population. We applied these different persistence measures to examine populations of sea lice on salmon farms and to answer the five questions posed at the end of the introduction.

5 Discussion

In this paper, we constructed a model for a meroplanktonic marine species, in which the larval stage is capable of dispersing between habitat patches, and the later sessile stages remain confined to a single habitat patch. This type of model is applicable to corals and coral reef fish (Cowen et al. 2006; Jones et al. 2009), barnacles (Roughgarden et al. 1988), Dungeness crabs (Botsford et al. 1994), sea urchins (Botsford et al. 1994), and many benthic marine species (Cowen and Sponaugle 2009). We modelled the growth and survival of sessile stages on a habitat patch using arbitrary survival and maturation functions so that our model is applicable to a breadth of different systems. To model the dispersal between patches in the larval stage, we approximated hydrodynamic movement, so that rates of larval movement between patches have an underlying mechanistic model. We then constructed the next-generation matrix, K, for this model. The next-generation matrix distils the key elements of the model into a matrix from which we can determine the source–sink distribution among patches using the column sums. We denote the jth column sum by $R_\delta(j)$ and showed that if $R_\delta(j) < 1$, then patch j is a sink and if $R_\delta(j) > 1$ then patch j is a source. We also proved that the basic reproduction number $R_0 = \rho(K)$ determines the stability of the zero equilibrium of our model, so that if $R_0 > 1$, then the population grows, and if $R_0 < 1$, then the population goes extinct.

Using salmon farms as an example, we investigated how the source–sink distribution can change as a function of patch separation distance and how often there is a critical separation distance for each patch, at which point a patch changes from a source to a sink. We demonstrated that increasing the ratio of advection to diffusion between patches increases the difference in critical separation distance between patches. We also demonstrate how $R_\delta(j)$ can be used to determine the transient dynamics of the salmon farm system and how these transient dynamics can persist over several generations, and differ from the asymptotic dynamics determined by R_0 (Sect. 4.4, Fig. 7). We investigated the effect of treatment of a single patch on the patch dynamics using the concept of type reproduction numbers and found that treating the middle patch in a channel results in the greatest reduction in R_0, but that treating the first patch results in the greatest reduction in the maximum R_δ. Lastly, we looked at differing local productivity on patch dynamics, determined the contribution that each patch has to R_0, and demonstrated how this can also differ from $R_\delta(j)$.

Next-generation matrices have a long history in epidemiology, where they have been used to calculate the number of new infections produced in one compartment when a newly infectious individual is introduced in another compartment (Diekmann et al. 1990, 2010; van den Driessche and Watmough 2002). Our approach in constructing a next-generation matrix for a stage-structured model with arbitrary stage durations and larval flow between patches extends recent use of next-generation operators in ecology (Huang et al. 2016a; Huang and Lewis 2015; Krkošek and Lewis 2010; McKenzie et al. 2012). Much of the previous work has used continuous space next-generation operators to determine the source–sink distribution of populations in streams or lakes. There, the movement of individuals through the water is also described by partial differential equations, though individuals can be produced at any point in space. In our work, we describe the movement of larvae between patches using advection–diffusion

equations, though as larvae can only be produced on certain population patches, our linear operator can be formulated as a matrix. Our work also extends work of Huang and Lewis (2015), where the next-generation matrix was used to determine the transient dynamics in a system and how they differ from the asymptotic dynamics determined by R_0 in a model of salmonids. The minimum R_δ was shown to determine whether it is possible for the population to initially decline, even if it eventually grows, and the maximum R_δ was shown to determine whether it is possible for the population to initially grow, even if it declines. We re-emphasize the finding that while R_0 determines the long-term dynamics of a system, it cannot also characterize the transient dynamics. We also advocate for further use of the next-generation approach in ecology, as the next-generation operators are often able to distil relevant ecological information into a simple operator.

Our model formulation as a set of age density equations, rather than a set of distributed delay equations, or partial differential equations, follows the work of Feng and Thieme (2000). There, arbitrary survival and maturation functions were used to model the progression of an infection with a finite number of infections stages, all of which have general length distributions. The generality of the survival and maturation functions used in our model allow us to calculate the next-generation matrix for a wide breadth of model formulations. For example, in models of sea lice populations on salmon, studies have used a variety of different maturation functions. When discrete differential equations are used (Adams et al. 2015; Revie et al. 2005), all lice of a given stage mature at the same age. The maturation function, $M_k^i(a)$, can then be formulated using step functions, as $M_k^i(a) = 1 - H(a - \tau)$. Here, $H(a)$ is the Heaviside function and τ is the development time. When linear delay differential equations are used to model sea lice development (Stien et al. 2005), there is some minimum development time, after which sea lice mature at a constant rate. The maturation function for our model could then be written as $M_k^i(a) = 1 - H(a - \tau)(1 - e^{-m(a-\tau)})$, where τ is now the minimum development time, and m is the constant rate of maturation. Weibull functions have the nice property that the probability of maturing is largest at some intermediate age and have thus also been used to describe maturation functions, without requiring a fixed minimum development time or fixed maturation time (Aldrin et al. 2017). Here, the maturation function can be written as $M_k^i(a) = e^{(-\lambda a)^p}$, where λ is the scale parameter, and p is the shape parameter for the Weibull distribution. Even though all of these models are formulated using different equations, by identifying the maturation and survival functions used, we can reformulate these models as age density equations, given by system 2 and therefore calculate the next-generation matrix for all these different types of models using Eq. 5. However, our model explicitly calculates rates of larval movement using the Fokker–Planck equation, and therefore, the arrival time component of our model and next-generation matrix will remain different from the above-mentioned models.

While the general structure of our model allows us to calculate the next-generation matrix for a variety of survival and maturation functions, one of the limitations of our model is that we do not include density dependence. To include density dependence in our model, the partial differential equation formulation of the model would no longer be set a of McKendrick–von Foerster equations (Keyfitz and Key-

fitz 1997; McKendrick 1925), as shown in Appendix A. They could, however, be reformulated as a series of Gurtin–McCamy equations (Gurtin and MacCamy 1974), where now $\mu_k^i(a) = \mu_k^i(a, N_1^i, \ldots, N_m^i)$, and $B_k^i(t) = B_k^i(t, N_1^i, \ldots, N_m^i)$ where $N_k^i = \int_0^\infty n_k^i(t, a)\mathrm{d}a$. Most of the sea lice population models previously mentioned (Adams et al. 2015; Aldrin et al. 2017; Revie et al. 2005; Stien et al. 2005) do not include density dependence in their model formulations, as the assumption is that sea lice are regulated before they reach high enough densities to exhibit negative density dependence and that there is no Allee affect at low densities. However, using Anderson–May host parasite equations, Krkošek et al. (2012b), demonstrated evidence of an Allee effect of sea lice on wild salmon, so that at low densities there is mate limitation, and thus a reduced birth rate. Modifying the birth rate of larvae to include mate limitation at low densities would be interesting future work. In modelling the populations of other marine species on habitat patches, it has been shown that if external recruitment to populations is much larger than self-recruitment, then it is not necessary to include negative density dependence at high population densities to control for unbounded growth (Armsworth 2002). However, the inclusion of negative density dependence is necessary to prevent unbounded growth of populations when self-recruitment to a population is large. The addition of density dependence to our model would therefore allow it to be applicable to a broader set of species at equilibrium densities.

With respect to sea lice on salmon farms, we use the advection–diffusion equation to approximate hydrodynamic ocean flow between farms due to the success of modelling the transmission of nauplii onto wild salmon with the same advection–diffusion equation. In the Broughton Archipelago, a region at the centre of the debate of the effect of salmon farms on wild salmon, advection–diffusion equations were used to model nauplii and copepodid movement, where nauplii were released as point sources from salmon farms (Krkošek et al. 2006). Copepodids could then attach to wild salmon migrating past these salmon farms, and Krkošek et al. (2006) were able to correlate the spatial distribution of sea lice on wild salmon with the spatial position of salmon farms. The accuracy of the approximation of ocean current using an advection–diffusion equation in the Broughton Archipelago has been debated (Brooks 2005), as well as the use of a constant maturation rate from nauplii to copepodids. We believe that the advection–diffusion equation is a useful approximation to ocean circulation in channels, especially when hydrodynamic models are not available, though we include a general maturation delay in our model, so that the maturation of larvae can parameterized accurately to different species.

In the context of sea lice on salmon farms, we also estimated the effect of sea lice treatment on a salmon farm network using measures of both transient and asymptotic dynamics. In the next-generation framework, we defined treatment as a reduction in the survival and/or maturation in the sessile stages on a farm. Salmon farms typically apply a parasiticide, emamectin benzoate, into the salmon feed to treat sea lice infestations (Rogers et al. 2013). We assumed that treating a salmon farm is therefore equivalent to reducing the survival through the different sessile stages on farm. We investigated the effect of treating the first, middle, and last farm in a channel. We found that treating the middle farm resulted in the greatest reduction in R_0 and that if either the first or last

farm where treated, then the reduction in R_0 was the same. However, treating the first farm resulted in the greatest reduction in R_u, the maximum intergenerational growth rate, and treating the last farm resulted in the greatest reduction in R_l, the minimum intergenerational growth rate.

If we are interested in preventing long-term outbreaks, reducing R_0 is important. However, in the case of the salmon aquaculture industry, frequent sea lice treatments prevent long-term growth of sea lice. In this case, it may be more important to prevent local outbreaks, as even local outbreaks of sea lice on salmon farms can have negative effects on migrating wild salmon (Bateman et al. 2016). Treating the first farm would thus most reduce the magnitude of a local outbreak. This result contradicts simulation studies of salmon farm dynamics in Scotland (Adams et al. 2015). This study found the farm influx (number of lice coming into a farm) was a better predictor of management impact than farm outflux (number of lice coming from a farm), even though when unmanaged lice density was accounted for, influx and management impact were only weakly correlated. In our work, relative influx can be calculated using the difference in row sums, and relative outflux could be calculated using relative column sums. In our model, the first farm has the highest outflux, and the last farm in the channel has the highest influx. Perhaps this difference is due to the fact that the farm with the highest influx is most likely to outbreak, and thus, treating that farm will be the most effective at reducing the total sea lice population. If the farm with the highest outflux is treated, then the worst possible initial outbreak decreases, even if this initial outbreak is less likely to happen. The difference between the results found in this paper and from Adams et al. (2015) highlights the complexity of designing effective management actions to control sea lice.

The largest limitation of our model, if applied to specific biological systems, is the use of the advection–diffusion equation to approximate ocean movement, rather than a hydrodynamic model. While the advection–diffusion equation may be a good approximation in a channel environment (Krkošek et al. 2006), the use of a hydrodynamic model to approximate larval movement between patches rather than an advection–diffusion equation would greatly improve the accuracy and relevance of the model to a specific region. Recently, there have been several studies which have used hydrodynamic models to accurately model the transmission of sea lice between salmon farms (Adams et al. 2015; Cantrell et al. 2018; Foreman et al. 2009), as well as the transmission of other marine larvae between population patches. These studies often quantify the amount of larval connectivity between patches by pairing particle tracking models with ocean circulation models, such as FVCOM (Chen et al. 2006). Connectivity matrices can then be constructed by tracking the number of particles released from one patch that pass by another patches. Most recently, Cantrell et al. (2018) have used kernel density estimation on the output of the particle tracking model to quantify infection pressure of sea lice from a particular salmon farm. Depending on the method that these models use to estimate larval connectivity between patches, the arrival time that we calculate in our paper from the advection–diffusion equation, could easily be calculated from these detailed hydrodynamic models. The connectivity matrices often calculated in these papers could then be reformulated as next-generation matrices using survival and maturation functions specific to the species studied, so that the source–sink distribution of populations can be directly calculated, and so that the

entries have a more relevant biological meaning. We believe that the combination of connectivity matrices from hydrodynamic models with next-generation matrices is an exciting area of future work to understand the population dynamics of specific systems.

Acknowledgements The authors would like to thank the members of the Lewis Lab for many helpful discussions and suggestions. PDH gratefully acknowledges an NSERC CGS-M scholarship and Queen Elizabeth II scholarship, and MAL gratefully acknowledges an NSERC Discovery Grant and a Canada Research Chair. We thank an anonymous reviewer for their helpful comments.

A Derivation from McKendrick–von Foerster PDE

Here, we derive Eq. 1 by solving the McKendrick–von Foerster PDE:

$$\begin{aligned} \frac{\partial n_k^i(t,a)}{\partial t} + \frac{\partial n_k^i(t,a)}{\partial a} &= -\mu_k^i(a) n_k^i(t,a) \\ n_k^i(t,0) &= B_k^i(t) \\ n_k^i(0,a) &= \tilde{n}_k^i(a) \\ \mu_k^i(a) &= -\frac{\left(M_k^i(a) S_k^i(a)\right)'}{M_k^i(a) S_k^i(a)}. \end{aligned} \tag{14}$$

First, for simplicity we drop the indexes k and i so that $n_k^i(t,a) = n(t,a)$. Then, we solve this linear partial differential equation using the method of characteristics. The goal is to reduce the partial differential equation into an ordinary differential equation of one variable along certain characteristic curves in a and t. To do this, we parameterize $a = a(s)$ and $t = t(s)$, so that $n(t(s), a(s))$ is now a function of the single variable s. Differentiating $n(t(s), a(s))$ with respect to s:

$$\frac{\mathrm{d}n}{\mathrm{d}s} = \frac{\partial n}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}s} + \frac{\partial n}{\partial a}\frac{\mathrm{d}a}{\mathrm{d}s}. \tag{15}$$

Now we choose the characteristic curves $a(s)$ and $t(s)$ such that

$$\frac{\mathrm{d}a}{\mathrm{d}s} = 1 \quad \text{and} \quad \frac{\mathrm{d}t}{\mathrm{d}s} = 1.$$

Then, substituting Eq. 14 into Eq. 15 we arrive at the ordinary differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}s} n(t(s), a(s)) = -\mu(a(s)) n(t(s), a(s)). \tag{16}$$

Solving for the characteristic curves, $t(s)$ and $a(s)$, we find

$$t(s) = s + t_0 \quad \text{and} \quad a(s) = s + a_0.$$

Then, solving for $n(t(s), a(s))$ in Eq. 16 we find:

$$\begin{aligned} n(t(s), a(s)) &= n(t(0), a(0))e^{-\int_0^s \mu(x+a_0)dx}\mathrm{d}x \\ &= n(t(0), a(0))e^{-\int_{a_0}^{a_0+s} \mu(y)dy} \\ &= n(t(0), a(0))e^{\int_{a_0}^{a_0+s} \frac{(M(y)S(y))'}{M(y)S(y)}dy} \\ &= n(t(0), a(0))e^{\int_{a_0}^{a_0+s} \frac{d}{dy}\log(M(y)S(y))dy} \\ &= n(t(0), a(0))\frac{M(a_0+s)S(a_0+s)}{M(a_0)S(a_0)}. \end{aligned} \tag{17}$$

Now we have two boundary conditions to impose, one at $t = 0$ and one at $a = 0$. Together, the two boundaries intersect all characteristic curves, and so Eq. 17 is the unique solution to Eq. 14 for all $a \geq, t \geq 0$. From the form of our characteristic equations for $a(s)$ and $t(s)$, it is clear that all characteristics are lines $t = a + b$ in the a-t plane. The line $t = a$ divides the a-t plane into two regions: $t \leq a$ and $t > a$. Characteristic curves for which $t \leq a$ intersect the boundary $t = 0$ at some point $(t, a) = (0, a_0)$. Substituting $t = s$ and $a = s + a_0$ into Eq. 17, we find

$$\begin{aligned} n(t, a) &= n(0, a - t)\frac{M(a)S(a)}{M(a-t)S(a-t)} \\ &= n_0(a - t)\frac{M(a)S(a)}{M(a-t)S(a-t)}. \end{aligned}$$

Similarly, characteristic curves for which $t > a$ intersect the $a = 0$ boundary at some point $(t, a) = (t_0, 0)$. Substituting $t = s + t_0$ and $a = s$ into Eq. 17, we find

$$\begin{aligned} n(t, a) &= n(t - a, 0)M(a)S(a) \\ &= B(t - a)M(a)S(a). \end{aligned}$$

Therefore, together we have

$$n(t, a) = \begin{cases} B(t-a)M(a)S(a) & t > a \\ n_0(a-t)\frac{M(a)S(a)}{M(at)S(a-t)} & 0 < t < a. \\ n_0(a) & t = 0 \end{cases}$$

References

Aaen SM, Helgesen KO, Bakke MJ, Kaur K, Horsberg TE (2015) Drug resistance in sea lice: a threat to salmonid aquaculture. Trends Parasitol 31(2):72–81

Adams TP, Proud R, Black KD (2015) Connected networks of sea lice populations: dynamics and implications for control. Aquac Environ Interact 6(3):273–284

Aldrin M, Huseby RB, Stien A, Grøntvedt RN, Viljugrein H, Jansen PA (2017) 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

Alexander SE, Roughgarden J (1996) Larval transport and population dynamics of intertidal barnacles: a coupled benthic/oceanic model. Ecol Monogr 66(3):259–275

Amarasekare P, Nisbet RM (2001) Spatial heterogeneity, source–sink dynamics, and the local coexistence of competing species. Am Nat 158(6):572–584
Armsworth PR (2002) Recruitment limitation, population regulation, and larval connectivity in reef fish metapopulations. Ecology 83(4):1092–1104
Bateman AW, Peacock SJ, Connors B, Polk Z, Berg D, Krkošek M, Morton A (2016) Recent failure to control sea louse outbreaks on salmon in the Broughton Archipelago, British Columbia. Can J Fish Aquat Sci 73(8):1164–1172
Botsford LW, Moloney CL, Hastings A, Largier JL, Powell TM, Higgins K, Quinn JF (1994) The influence of spatially and temporally varying oceanographic conditions on meroplanktonic metapopulations. Deep Res II 41(1):107–145
Brooks KM (2005) 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
Cantrell DL, Rees EE, Vanderstichel R, Grant J, Filgueira R, Revie CW (2018) 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
Chen C, Beardsley R, Cowles G (2006) An unstructured grid, finite-volume coastal ocean model (FVCOM) system. Oceanography 19(1):78–89
Costello MJ (2006) Ecology of sea lice parasitic on farmed and wild fish. Trends Parasitol 22(10):475–483
Costello MJ (2009) The global economic cost of sea lice to the salmonid farming industry. J Fish Dis 32(1):115–118
Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Ann Rev Mar Sci 1:443–466
Cowen RK, Lwiza KM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or closed? Science 287(5454):857–859
Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations. Science 311(5760):522–527
Diekmann O, Heesterbeek JAP, Metz JAJ (1990) On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. J Math Biol 28(4):365–382
Diekmann O, Heesterbeek JAP, Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models. J R Soc Interface 7(47):873–885
Feng Z, Thieme HR (2000) Endemic models with arbitrarily distributed periods of infection I: fundamental properties of the model. SIAM J Appl Math 61(3):803–833
Feng Z, Xu D, Zhao H (2007) Epidemiological models with non-exponentially distributed disease stages and applications to disease control. Bull Math Biol 69(5):1511–1536
Figueira WF, Crowder LB (2006) Defining patch contribution in source-sink metapopulations: the importance of including dispersal and its relevance to marine systems. Popul Ecol 48(3):215–224
Foreman MGG, Czajko P, Stucchi DJ, Guo M (2009) A finite volume model simulation for the Broughton Archipelago, Canada. Ocean Model 30(1):29–47
Godwin SC, Krkošek M, Reynolds JD, Rogers LA, Dill LM (2017) Heavy sea louse infection is associated with decreased stomach fullness in wild juvenile sockeye salmon. Can J Fish Aquat Sci 75(10):1587–1595
Gurtin ME, MacCamy RC (1974) Non-linear age-dependent population dynamics. Arch Ration Mech Anal 54(3):66–76
Gyllenberg M, Hanski I (1997) Habitat deterioration, habitat destruction, and metapopulation persistence in a heterogenous landscape. Theor Popul Biol 52(3):198–215
Hanski I (1998) Metapopulation dynamics. Nature 396:41–49
Hastings A (1982) Dynamics of a single species in a spatially varying environment: the stabilizing role of high dispersal rates. J Math Biol 16(1):49–55
Heesterbeek JA, Roberts MG (2007) The type-reproduction number T in models for infectious disease control. Math Biosci 206(1):3–10
Horn RA, Johnson CR (2012) Matrix analysis, 2nd edn. Cambridge University Press, New York
Huang C, Cao J, Wen F, Yang X (2016a) Stability analysis of SIR model with distributed delay on complex networks. PLoS ONE 11(8):1–23
Huang Q, Lewis MA (2015) Homing fidelity and reproductive rate for migratory populations. Theor Ecol 8(2):187–205

Huang Q, Jin Y, Lewis MA (2016b) R_0 analysis of a Benthic-Drift model for a stream population. SIAM J Appl Dyn Syst 15(1):287–321
Hurford A, Cownden D, Day T (2010) Next-generation tools for evolutionary invasion analyses. J R Soc Interface 7(45):561–571
Johnson SC, Albright LJ (1991) Development, growth, and survival of lepeophtheirus salmonis (Copepoda: Caligidae) under laboratory conditions. J Mar Biol Assoc United Kingdom 71(2):425–436
Jones G, Almany G, Russ G, Sale P, Steneck R, van Oppen M, Willis B (2009) Larval retention and connectivity among populations of corals and reef fishes : history, advances and challenges. Coral Reefs 28:307–325
Keyfitz BL, Keyfitz N (1997) The McKendrick partial differential equation and its uses in epidemiology and population study. Math Comput Model 26(6):1–9
Krkošek M, Lewis MA (2010) An R_0 theory for source-sink dynamics with application to Dreissena competition. Theor Ecol 3(1):25–43
Krkošek M, Lewis MA, Morton A, Frazer LN, Volpe JP (2006) Epizootics of wild fish induced by farm fish. Proc Natl Acad Sci USA 103(42):15506–15510
Krkošek M, Ford JS, Morton A, Lele S, Myers RA, Lewis MA (2007) Declining wild salmon populations in relation to parasites from farm salmon. Science 318(2):1772–1775
Krkošek M, Connors BM, Morton A, Lewis MA, Dill LM, Hilborn R (2011) Effects of parasites from salmon farms on productivity of wild salmon. Proc Natl Acad Sci U S A 108(35):14700–14704
Krkošek M, Bateman AW, Proboszcz S, Orr C (2012a) 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
Krkošek M, Connors BM, Lewis MA, Poulin R (2012b) Allee effects may slow the spread of parasites in a coastal marine ecosystem. Am Nat 179(3):401–412
Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull Entomol Soc Am 15(3):237–250
Lewis MA, Shuai Z, van den Driessche P (2019) A general theory for target reproduction numbers with applications to ecology and epidemiology. J Math Biol 78(7):2317–2339
Li CK, Schneider H (2002) Applications of Perron-Frobenius theory to population dynamics. J Math Biol 44(5):450–462
McKendrick AG (1925) Applications of mathematics to medical problems. Proc Edinb Math Soc 44:98–130
McKenzie HW, Jin Y, Jacobsen J, Lewis MA (2012) R_0 analysis of a spatiotemporal model for a stream population. SIAM J Appl Dyn Syst 11(2):567–596
Mileikovsky SA (1971) Types of larval development in marine bottom invertebrates, their distribution and ecological significance: a re-evaluation. Mar Biol Int J Life Ocean Coast Waters 10(3):193–213
Planes S, Jones GP, Thorrold SR (2009) Larval dispersal connects fish populations in a network of marine protected areas. Proc Natl Acad Sci 106(14):5693–5697
Pulliam HR (1988) Sources, sinks and population regulation. Am Nat 132(5):652–661
Revie CW, Robbins C, Gettinby G, Kelly L, Treasurer JW (2005) 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
Roberts MG, Heesterbeek JA (2003) A new method for estimating the effort required to control an infectious disease. Proc R Soc B Biol Sci 270(1522):1359–1364
Rogers LA, Peacock SJ, McKenzie P, DeDominicis S, Jones SR, Chandler P, Foreman MG, Revie CW, Krkošek M (2013) Modeling parasite dynamics on farmed salmon for precautionary conservation management of wild salmon. PLoS ONE 8(4)
Roughgarden J, Gaines S, Possingham H (1988) Recruitment dynamics in complex life cycles. Science 241(4872):1460–1466
Stien A, Bjørn PA, Heuch PA, Elston DA (2005) Population dynamics of salmon lice Lepeophtheirus salmonis on Atlantic salmon and sea trout. Mar Ecol Prog Ser 290:263–275
van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180(1–2):29–48
Watson JR, Kendall BE, Siegel DA, Mitarai S (2012) Changing seascapes, stochastic connectivity, and marine metapopulation dynamics. Am Nat 180(1):99–112