

# Next-generation matrices for marine metapopulations: The case of sea lice on salmon farms

Peter D. Harrington, Danielle L. Cantrell, and Mark A. Lewis

2023

Faculty of Science

Faculty Publications

© 2023 Harrington et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC BY:  
<http://creativecommons.org/licenses/by/4.0/>

Original citation:

Harrington, P. D., Cantrell, D. L., & Lewis, M. A. (2023). Next-generation matrices for marine metapopulations: The case of sea lice on salmon farms. *Ecology and Evolution*, 13(4). <https://doi.org/10.1002/ece3.10027>

---

Downloaded from UVicSpace Research & Learning Repository

[dspace.library.uvic.ca](https://dspace.library.uvic.ca)



University  
of Victoria

Libraries

## RESEARCH ARTICLE

# Next-generation matrices for marine metapopulations: The case of sea lice on salmon farms

Peter D. Harrington<sup>1,2</sup>  | Danielle L. Cantrell<sup>3</sup> | Mark A. Lewis<sup>1,4,5,6</sup>

<sup>1</sup>Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada

<sup>2</sup>Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada

<sup>3</sup>California Department of Fish and Wildlife, Marine Region's Fisheries Analytics Project, Monterey, California, USA

<sup>4</sup>Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

<sup>5</sup>Department of Mathematics and Statistics, University of Victoria, Victoria, British Columbia, Canada

<sup>6</sup>Department of Biology, University of Victoria, Victoria, British Columbia, Canada

## Correspondence

Peter D. Harrington, Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver, BC V6T 1Z2, Canada.  
Email: [harrington@math.ubc.ca](mailto:harrington@math.ubc.ca)

## Funding information

Alberta Graduate Excellence Scholarship; Natural Sciences and Engineering Research Council of Canada

## Abstract

Classifying habitat patches as sources or sinks and determining metapopulation persistence requires coupling connectivity between habitat patches with local demographic rates. While methods to calculate sources, sinks, and metapopulation persistence exist for discrete-time models, there is no method that is consistent across modeling frameworks. In this paper, we show how next-generation matrices, originally popularized in epidemiology to calculate new infections after one generation, can be used in an ecological context to calculate sources and sinks as well as metapopulation persistence in marine metapopulations. To demonstrate the utility of the method, we construct a next-generation matrix for a network of sea lice populations on salmon farms in the Broughton Archipelago, BC, an intensive salmon farming region on the west coast of Canada where certain salmon farms are currently being removed under an agreement between local First Nations and the provincial government. The column sums of the next-generation matrix can determine if a habitat patch is a source or a sink and the spectral radius of the next-generation matrix can determine the persistence of the metapopulation. With respect to salmon farms in the Broughton Archipelago, we identify the salmon farms which are acting as the largest sources of sea lice and show that in this region the most productive sea lice populations are also the most connected. The farms which are the largest sources of sea lice have not yet been removed from the Broughton Archipelago, and warming temperatures could lead to increased sea louse growth. Calculating sources, sinks, and persistence in marine metapopulations using the next-generation matrix is biologically intuitive, mathematically equivalent to previous methods, and consistent across different modeling frameworks.

## KEYWORDS

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

## TAXONOMY CLASSIFICATION

Conservation Ecology, Population Ecology, Theoretical Ecology

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Metapopulations consist of subpopulations located on isolated habitat patches that are connected via dispersal (Hanski, 1998; Kritzer & Sale, 2004; Levins, 1969). In most benthic marine species, this dispersal comes from the pelagic larval stage (Cowen & Sponaugle, 2009). Larvae disperse between and then settle on habitat patches, and once settled the remaining stages are sedentary and remain confined to a specific habitat patch. In marine systems, the metapopulation concept, where subpopulations are connected but have their own demographic rates, has been used in the spatial planning of marine protected areas (MPA) and the siting of marine reserves (White et al., 2013).

In a metapopulation framework, habitat patches are often classified into sources and sinks based on how the subpopulations on these patches contribute to the overall metapopulation. The source-sink classification of habitat patches was first described concretely in a terrestrial context by Pulliam (1988), where habitat patches were classified as sources if the local subpopulations could persist in isolation and sink if they could not. However, this classification ignores the effect of dispersal, which is especially critical in marine metapopulations, and so Runge et al. (2006) and Figueira and Crowder (2006) updated the classification of source and sink patches to include both the local productivity of a patch, as well as the ability to disperse away from the habitat patch. Under this new classification, a source patch is a patch on which an adult will more than self-replace over the entire metapopulation, and a sink is a patch on which an adult will not. Self-replacement need not occur on the same habitat patch as the adult originated, and thus under this classification, a source patch may not be able to persist in isolation.

Due to the large-scale larval dispersal that occurs in many marine species, it is common in marine metapopulations for source patches to not be able to persist in isolation (White et al., 2019), thus preserving a persistent metapopulation, especially in the context of MPAs, often requires more than simply preserving source patches. To maintain persistent marine metapopulations it is necessary to preserve sufficient larval exchange between closed loops of habitat patches so that an average adult can eventually self-replace over multiple generations (Burgess et al., 2014; Hastings & Botsford, 2006). If sources patches self-recruit enough larvae to persist in isolation then this can be accomplished by preserving only source patches. However, if they do not it may require preserving both source and sink patches, as the sink patches may provide sufficient larval exchange back to the source patches to create a closed loop of habitat patches over which an adult can self-replace. Evaluating the persistence of marine metapopulations is therefore difficult as it requires accurate measures of larval connectivity between habitat patches, as well as accurate local demographic rates of adult stages on each habitat patch (Burgess et al., 2014). Despite the difficulties, evaluating metapopulation persistence is critical in designing marine reserves in which the protected habitat patches can persist even when outside patches

are exploited (Costello et al., 2010; Garavelli et al., 2018; Puckett & Eggleston, 2016; White et al., 2010, 2013).

Next-generation matrices are a useful tool that can both be used to evaluate the persistence of a metapopulation as well as identify the contribution of local habitat patches under a variety of modeling frameworks. Originally popularized in epidemiology (Diekmann et al., 1990), they have since been used in ecology to calculate the basic reproduction number,  $R_0$ , of structured population models (Cushing & Yicang, 1994; Li & Schneider, 2002). Next-generation matrices convert structured population models into generational time, so that the entries of the next-generation matrix are the number of new individuals produced in each age, stage, or patch, depending on how the model is structured. The individual contribution of habitat patches or evaluation of metapopulation persistence can therefore be measured for different model structures (discrete time, continuous time, etc.) under the same framework of the next-generation matrix. The column sums can be used to measure the contribution of each habitat patch over a generation and the magnitude of the dominant eigenvalue can be used to evaluate metapopulation persistence if the next-generation matrix is primitive (otherwise the spectral radius can be used to determine persistence if the matrix is non-negative and irreducible; Harrington & Lewis, 2020). Next-generation operators have also been used in ecology to calculate source and sink regions in heterogeneous environments (Harrington & Lewis, 2020; Huang et al., 2016; Krkošek & Lewis, 2010; McKenzie et al., 2012), and to evaluate the level of control required to suppress invasive species (Lewis et al., 2019).

In this paper, we focus on using the next-generation matrix to evaluate local patch contribution and metapopulation persistence in marine metapopulations, but the framework used here is also applicable to many other birth-jump metapopulations (Hillen et al., 2015), where there is a single juvenile stage which can disperse between habitat patches and the remaining stages remain on a single habitat patch. Examples of non-marine species that exhibit this structure include plant species where seeds are carried between habitat patches (Husband & Barrett, 1996), or insect species with a single large dispersal event such as the spruce budworm (Williams & Liebhold, 2000) or mountain pine beetle (Safranyik & Carroll, 2007). In fact, the next-generation matrix approach can even be extended to metapopulations in which adults also disperse, though the calculations become more complicated and so here we focus on species with a single dispersing stage.

Specifically, to demonstrate the utility of the method, we use the next-generation matrix to calculate the contribution of a single salmon farm to the spread of sea lice in a salmon farming region on the west coast of British Columbia. Sea lice (*Lepeophtheirus salmonis*) are a parasitic marine copepod that feed on the epidermal tissues, muscles, and blood of salmon (Costello, 2006). With a free-living larval stage, they can disperse tens of kilometers, spreading between salmon farms in a region and between wild and farmed salmon (Krkošek et al., 2006; Peacock et al., 2020; Stucchi et al., 2011). Lesions and stress from high sea lice infestation make adult salmon

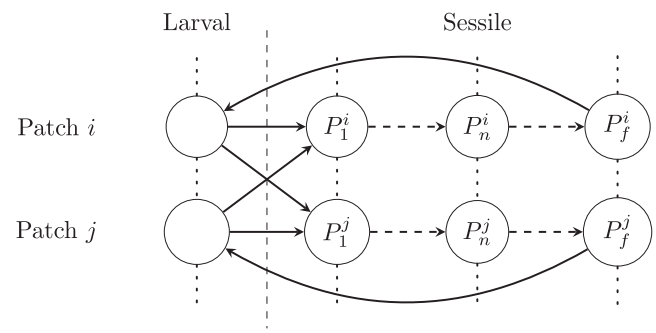
more susceptible to secondary infections, leading to large economic consequences for the salmon farming industry (Abolofia et al., 2017). On wild juvenile salmon, infestation with sea lice can lead to mortality and elevated exposure to sea lice from salmon farms can contribute to population-level declines in pink salmon (Krkošek et al., 2007). In the context of sea lice on salmon farms, we are not concerned with preserving a persistent metapopulation of sea lice parasites, but instead, we use the next-generation matrix to calculate the relative contribution of each salmon farm and evaluate the effect of environmental variables on the overall growth of the sea louse metapopulation.

The specific salmon farming region that we focus on to calculate farm contribution is the Broughton Archipelago. The Broughton Archipelago is located on the west coast of Canada, between Vancouver Island and the mainland of British Columbia, and has been central in evaluating the effect of sea lice from salmon farms on wild salmon (Brooks, 2005; Krkošek et al., 2005, 2007, 2011; Marty et al., 2010). The area has historically had around 20 active salmon farms (Foreman et al., 2015), though currently certain farms are being removed from this region in an agreement between the government of British Columbia and the Kwikwasut'inuxw Haxwa'mis, Namgis, and Mamalilikulla First Nations (Brownsey & Chamberlain, 2018). After 2023 many of the remaining farms must be approved by both the local First Nations and the government in order to continue to operate and thus determining the farms which are acting as the largest sources of sea lice is critical during this transition period.

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

## 2 | MATERIALS AND METHODS

In this section, we present details on the construction of the next-generation matrix and show how it can be used to determine the contribution of the subpopulation on a single habitat patch to the metapopulation as well as determine the persistence of the metapopulation. We then present the explicit construction of the next-generation matrix for models with age-dependent demography and present the construction for ordinary differential equation models and discrete-time models in Appendices 1 and 2, respectively. Finally, we detail the construction of the next-generation matrix for



**FIGURE 1** The lifecycle graph for two patches in a metapopulation of a species with a single larval stage that disperses between habitat patches and  $f$  sessile stages that remain on a habitat patch. The population on patch  $i$  in stage  $n$  is given by  $P_n^i$ .

a system of sea lice populations on salmon farms in the Broughton Archipelago.

### 2.1 | Next-generation matrices

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

We use the next-generation matrix to calculate the local patch contribution to the metapopulation in the context of low population density. To calculate the next-generation matrix it is necessary to linearize a potential density-dependent model around the zero equilibrium so that the effect of density dependence at higher population sizes is ignored when calculating patch contribution. This approach of ignoring density dependence is common when determining persistence or patch contribution of marine metapopulations, as the focus is either on determining if a metapopulation can persist at all, or determining which habitat patches are acting as population sinks and which are acting as population sources (Burgess et al., 2014; Harrington & Lewis, 2020; Hastings & Botsford, 2006; Krkošek & Lewis, 2010). Alternatively, it is also useful for determining patch contribution in metapopulations of species that are being actively controlled to remain at low densities, such as sea lice on salmon farms, which is our focus in Section 2.4. Another common assumption in the theory of persistence also made here is that there is no

discernible Allee effect in any of the patches (Burgess et al., 2014; Hastings & Botsford, 2006; Krkošek & Lewis, 2010).

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

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

where the product of the first two terms is the lifetime reproduction of an individual on patch  $j$  and the third term is the probability that larvae leaving patch  $j$  successfully arrive on patch  $i$ .

However, if the next-generation matrix is constructed for models that explicitly model the larval stage, then the larval stage is often considered the first stage. The next-generation matrix will be slightly different in this case as well as the calculation of patch contribution, though the calculation of metapopulation persistence will be the same. We illustrate the differences between constructions of the next-generation matrix in the Section 4 and show how the calculation of persistence remains the same under all constructions.

## 2.2 | Determining patch contribution and metapopulation persistence

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

$$R_j = \sum_i k_{ij}, \quad (2)$$

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

This definition of patch contribution easily lends itself to classifying local habitat patches as population sources or sinks. If  $R_j < 1$ , then an individual on patch  $j$  cannot replace itself over the entire metapopulation, and thus patch  $j$  is defined as a sink (Harrington & Lewis, 2020). If  $R_j > 1$ , then one individual on patch  $j$  is producing more than one individual over the entire population and so patch  $j$  is defined as a source. To calculate persistence we can use the basic reproduction number,  $R_0$ , which can be calculated as

$$R_0 = \rho(K), \quad (3)$$

where  $\rho()$  is the spectral radius. If  $R_0 > 1$ , then the metapopulation will persist and if  $R_0 < 1$  then the metapopulation will go extinct, a relationship that holds under any of the model formulations considered here (Cushing & Yicang, 1994; Harrington & Lewis, 2020; Li & Schneider, 2002; van den Driessche & Watmough, 2002). The only conditions required are that  $K$  be irreducible, which is biologically satisfied if there is some small positive probability that larvae leaving one patch can eventually arrive on any other patch, and that  $K$  be non-negative, which is always satisfied as the entries are numbers of individuals and cannot be negative.

There are two biologically reasonable properties that also exist mathematically under this framework. First, if the population on any single habitat patch can persist on its own, so that  $k_{jj} > 1$ , then the entire metapopulation will persist and  $R_0 > 1$  (Harrington & Lewis, 2020). Second, a metapopulation consisting only of sink patches cannot persist and a metapopulation consisting only of sources cannot go extinct. The mathematical underpinning of these relationships is that the spectral radius must be between the minimum and maximum column sums of a matrix, so in terms of our metapopulation quantities

$$\min_j R_j \leq R_0 \leq \max_j R_j. \quad (4)$$

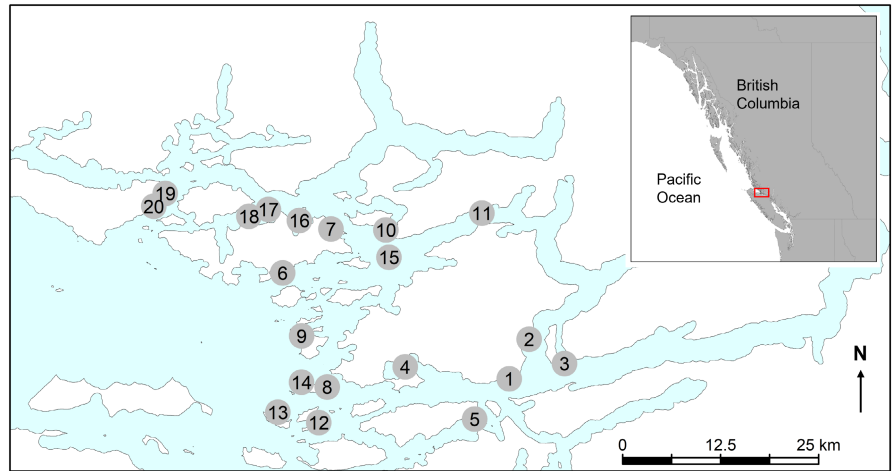
It should be noted that both of these properties exist under a deterministic modeling framework; when stochasticity is added neither the persistence nor extinction results are guaranteed.

Having defined patch contribution and persistence in terms of the next-generation matrix, we now demonstrate how to calculate the next-generation matrix for a model with age-dependent demography, as this is the modeling structure commonly used for sea lice on salmon farms. We also present the construction of the next-generation matrix for ordinary differential equation models and discrete time models in Appendices 1 and 2, so that the details are present for the most commonly used population models.

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

Here, we calculate the next-generation matrix for models which allow for the maturation, survival, birth, and dispersal rates to

**FIGURE 2** Map of the 20 historically active salmon farms in the Broughton Archipelago, for which the next-generation matrix is calculated.



depend not only on the stage and patch location of an individual, but also on the time that they have spent in a stage. This time is often referred to as stage age and so in these models there are two time variables: the global time of the system,  $t$ , and the time that an individual has spent in a particular stage, their stage-age,  $a$ . These models can be specified either as McKendrick von-Foerster partial differential equations, integrodifferential equations, or renewal equations (Feng & Thieme, 2000). In all cases, the dependence of the maturation rate on time spent in a stage allows for the addition of more realistic maturation functions where most individuals mature at some intermediate stage-age, or after some minimum time spent in the stage. In contrast, if models are formulated using ordinary differential equations the time in a stage is always exponentially distributed.

However, the specification of the model equations for these models can be rather complicated and so here we construct the next-generation matrix directly from the maturation, survival, birth, and dispersal functions, but the full model and the derivation of the next-generation matrix can be found in Appendix 3 as well as in Harrington and Lewis (2020). Essentially the construction involves tracking the probability that an individual survives through the different stages on a specific patch, the number of larvae that they produce, and the probability that the larvae successfully disperse from one patch from another.

If the maturation rate from stage  $n$  to  $n + 1$  is  $m_n^j(a)$ , then the probability that an individual has not yet matured from stage  $n$  to  $n + 1$  at stage-age  $a$  can be given by  $M_n^j(a) = \exp\left(-\int_0^a m_n^j(\tau) d\tau\right)$ . Then, if the probability that an individual survives up to stage-age  $a$  in stage  $n$  is given by  $S_n^j(a)$ , the total probability that an individual is still alive and has not yet matured at age  $a$  is  $S_n^j(a)M_n^j(a)$ . To calculate the probability that an individual starting in stage  $n$  eventually reaches stage  $n + 1$ , we multiply the rate at which they are maturing at stage-age  $a$ ,  $m_n^j(a)$ , by the probability that they are still alive and have not yet matured at  $a$ ,  $S_n^j(a)M_n^j(a)$ , and integrate over all stage-ages at which they could have left the stage:

$$\Pr(\text{stagen} \rightarrow n + 1) = \int_0^\infty S_n^j(a)M_n^j(a)m_n^j(a)da. \quad (5)$$

If the rate of maturation is constant so that  $m_n^j(a) = m_n^j$  and the mortality rate is constant, so that  $S_n^j(a) = e^{-\mu_n^j a}$ , this reduces to

$$\Pr(\text{stagen} \rightarrow n + 1) = \int_0^\infty e^{-\mu_n^j a} e^{-m_n^j a} m_n^j da = \frac{m_n^j}{m_n^j + \mu_n^j}$$

which is the case if the model is formulated as a system of ordinary differential equations.

In the terminal stage, if individuals at stage-age  $a$  produce larvae at rate  $b^j(a)$  and the probability that they survive up to  $a$  is  $S_f^j(a)$ , then to calculate the overall number of larvae that an individual produces over their lifetime we multiply  $b^j(a)$  by  $S_f^j(a)$  and integrate over all  $a$ :

$$\# \text{ larvae produced} = \int_0^\infty S_f^j(a)b^j(a)da.$$

Therefore, the entries of the next-generation matrix,  $k_{ij}$ , which are the total number of new individuals produced on patch  $i$  over the lifetime of an initial individual on patch  $j$ , will be given by the probability that an initial individual survives to the terminal stage on patch  $j$ , multiplied by the total number of larvae it produces, multiplied by the probability that the larvae successfully disperse from patch  $j$  to patch  $i$ :

$$k_{ij} = \underbrace{\prod_{n=1}^{f-1} \left( \int_0^\infty S_n^j(a)M_n^j(a)m_n^j(a)da \right)}_{\text{survival to terminal stage}} \underbrace{\left( \int_0^\infty S_f^j(a)b^j(a)da \right)}_{\# \text{ larvae produced}} \times \underbrace{p^{ji}}_{\text{Pr(patch } j \text{ to patch } i)}. \quad (6)$$

## 2.4 | Application: Sea lice on salmon farms in the Broughton archipelago

In this section, we construct a next-generation matrix to determine the contribution of a single salmon farm to the spread of sea lice in a salmon farming region on the west coast of British Columbia, the

Broughton Archipelago (Figure 2). To determine the level of sea lice dispersal away from salmon farms, we use a hydrodynamic model, coupled with a particle tracking model, to track sea lice particles released from 20 historical farms in the Broughton Archipelago (Cantrell et al., 2018). The coupled hydrodynamic and particle tracking models have been validated with sea louse counts on salmon farms in the Broughton (Cantrell et al., 2021). We construct one next-generation matrix for the 20 historical farms in the region for which the hydrodynamic particle tracking model was run (Figure 3a), as well as one for the 11 remaining farms in the area after 2023 (Figure 3b), subject to First Nations and governmental approval (Brownsey & Chamberlain, 2018).

### 2.4.1 | Modeling framework

Sea lice maturation through stages is often modeled with stage-age-dependent maturation functions (Aldrin et al., 2017; Revie et al., 2005; Stien et al., 2005) and thus we construct the next-generation matrix for sea lice using age-dependent demography, as shown in Section 2.3.

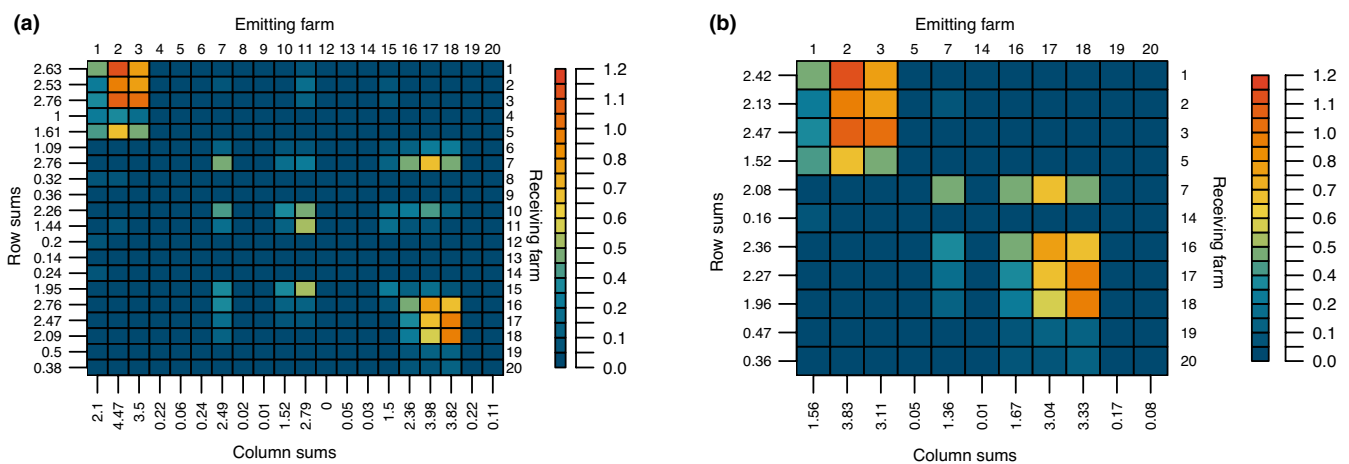
### 2.4.2 | Dispersal

The probability that a sea louse larvae leaves from one farm and successfully arrives on another depends on several factors including ocean current, temperature, and salinity. To accurately capture this probability, it is necessary to use a computational hydrodynamic model that can track the spread of larvae originating from a given farm as well as the dependence of larval survival on temperature and salinity. In order to determine the probability of

larvae dispersing between farms we use connectivity matrices from Cantrell et al. (2018). Details on these connectivity matrices can be found in Cantrell et al. (2018). Briefly, these connectivity matrices are calculated by applying Kernel Density Estimation (KDE) to particle tracking simulations to calculate the infectious density of sea lice at each farm, originating from a given farm. The particle tracking simulations are run on the output generated by a Finite Volume Community Ocean Model (FVCOM) which uses data on tides, wind surface heating, and river discharge to simulate three-dimensional ocean velocity, temperature, and salinity (Foreman et al., 2009). In the particle tracking simulation, the survival of sea louse particles is dependent on temperature and their maturation from non-infectious to infectious larvae is dependent on temperature.

The infectious densities of sea lice are calculated for each particle release day by taking daily snapshots of the particle locations of infectious lice for 11 days post-release and then applying Kernel Density Estimation to the accumulated daily snapshots. A connectivity matrix is then calculated for each particle release day, where the entry in row  $i$  and column  $j$  of the connectivity matrix is the infectious density of larvae over farm  $i$ , produced by larvae initially leaving farm  $j$ . In Cantrell et al. (2018) the infectious densities were calculated from 24 h of particle releases, where 50 particles were released each hour and so to calculate the infectious density of one initial release particle we divide the entries in each connectivity matrix by 1200 ( $50 \times 24$ ). Then, to create a single connectivity matrix,  $C$ , for the 20 farms in the Broughton Archipelago, we take the average over all the connectivity matrices created for particles released between March 14 and July 20, 2009.

The necessary quantity to construct the next-generation matrix is  $p^{ij}$ , the probability that larvae leaving farm  $j$  will successfully attach on farm  $i$ . To estimate  $p^{ij}$  from the entries of the connectivity matrix,  $c_{ij}$ , there are several assumptions that need to be made. If we assume



**FIGURE 3** (a) The next-generation matrix for the 20 historically active farms in the Broughton Archipelago, and (b) the next-generation matrix containing only the farms remaining in the Broughton Archipelago after 2023, subject to First Nations and governmental approval (Brownsey & Chamberlain, 2018). The entries of the next-generations matrices,  $k_{ij}$  are the number of new chalimus stage lice produced on farm  $i$  from one initial chalimus on farm  $j$ . The column sums,  $R_j$ , are the total number of chalimus produced on all farms from an initial chalimus on farm  $j$  and are shown below each column. Likewise, the row sums are the number of new chalimus received by each farm from all other farms and are shown on the left of each row. These numbers should be taken as relative, rather than absolute, as we do not have a very accurate estimate for the arrival rate of sea lice onto farms,  $\beta$ .

that the number of lice that arrive onto farms is small compared to the total number of lice in the water column, so that lice arriving onto farms do not significantly affect the density of lice in the water column, then

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

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

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

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

$$p^{ji} = \int_0^\infty \beta \int_{\text{farm } i} p_h^j(x, t) d\Omega dt \quad (7)$$

$$\approx \beta \times \text{Area of farm } i \times 1 \text{ day} \times (p_h^j(x_i, 0) + p_h^j(x_i, 1) + \dots + p_h^j(x_i, 11)) \quad (8)$$

$$= \beta \times 0.01 \times c_{ij}. \quad (9)$$

Therefore, to calculate  $p^{ji}$  from  $c_{ij}$  we need to estimate the arrival rate  $\beta$ . However, very little is known about the arrival rate of lice dispersing from one farm to another, and so the estimate presented here is very uncertain and could be orders of magnitude off from the true arrival rate. We assume here that  $\beta=100/\text{day}$ , and thus the average waiting time for infectious sea lice in the water column surrounding the farm to arrive on the farm is roughly 15 min ( $1/\beta=(1 \text{ day}/100) \times (24 \text{ h}/1 \text{ day}) \times (60 \text{ min}/1 \text{ h})=14.4 \text{ min}$ ). However, the waiting time could be as little as 1 min or as long as 1 h, depending on environmental conditions. Moreover, this is ignoring the fact that some infectious lice may never attach, and the estimates of the proportion of lice which successfully attach at all varies from 80% to 0.5% under different lab conditions (Skern-Mauritzen et al., 2020).

It can be seen from Equations (6) and (9) that each entry in the next-generation matrix,  $K$ , is a linear function of the arrival rate  $\beta$ , and therefore changing  $\beta$  will not affect the relative ordering of the contributions of different salmon farms, given by  $R_j$  (Equation 2), it will only affect the absolute magnitude. The basic reproduction rate,  $R_0=\rho(K)$ , is also a linear function of  $\beta$  and therefore any increase or decrease in  $\beta$  will result in the same proportional increase or decrease in  $R_0$ . However, the purpose of this application is not

to accurately estimate the basic reproduction number,  $R_0$ , for the Broughton Archipelago, or to accurately estimate the contribution of an individual farm,  $R_j$ , but rather to compare the relative contributions of different salmon farms in the system, and to investigate the effect of environmental variables on the basic reproduction number. Therefore, we present our estimate of the arrival rate for these purposes only, and our value of  $R_0$  found in the results should not be taken as an accurate estimate.

### 2.4.3 | Demography

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

We calculate the on-patch component of the elements of next-generation matrix,  $k_{ij}$ , by integrating the demographic functions over all time, and refer to this as the productivity of patch  $j$ :

$$\text{Productivity of patch } j = \underbrace{\prod_{n=1}^{f-1} \left( \int_0^\infty S_n^j(a) M_n^j(a) m_n^j(a) da \right)}_{\text{survival to terminal stage}} \underbrace{\left( \int_0^\infty S_f^j(a) b^j(a) da \right)}_{\text{\# larvae produced}}, \quad (10)$$

so that  $k_{ij}=(\text{Productivity of patch } j) \times p^{ji}$ , where  $p^{ji}$  is the probability that larvae successfully disperse from patch  $j$  to patch  $i$ . To determine the specific temperature and salinity-dependent demographic rates at each farm we find the temperature and salinity that each sea louse particle experiences in the particle tracking simulation when initially released from a farm. We then use the average temperature and salinity of particles over all releases.

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

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

Description	Function	Units	References
Survival probability of all stages	$S(a) = \exp\left(- (24a)^{(1/0.512)} \times \exp(- (1/0.512)(4.12 + 0.1245S))\right)$		Connors et al. (2008) Rittenhouse et al. (2016)
Maturation rate out of stage $i$	$m_i(a) = \log(2)\delta_{si} \left( (\delta_{mi}(10/T)^{\delta_{pi}})^{-\delta_{si}} \right) a^{\delta_{si}-1}$ $\delta_{mc} = 18.869, \delta_{sc} = 7.945, \delta_{pc} = 1.305, \delta_{mp} = 10.742, \delta_{sp} = 1.643,$ $\delta_{pp} = 0.866$	$\frac{1}{\text{day}}$	Aldrin et al. (2017)
Egg batch size	$B = 2\exp\left(5.6 - 0.43\log(T/10) - 0.78(\log(T/10))^2\right)$	eggs	Stige et al. (2021)
Egg development time	$D = (41.98/(T - 10 + 41.98 * 0.338))^2$	days	Stige et al. (2021)
Egg viability (salinity)	$V_{sal} = \exp(- 11.92 + 0.538S) / (1 + \exp(- 11.92 + 0.538S))$		Stige et al. (2021)
Egg viability (temperature)	$V_{temp} = \exp(- 1.765 + 0.494T) / (1 + \exp(- 1.765 + 0.494T))$		Stige et al. (2021)
Birth rate	$b(a) = B \times V_{sal} \times V_{temp} / D$	$\frac{\text{larvae}}{\text{day}}$	Stige et al. (2021)
Arrival rate	$\beta = 100$	$\frac{1}{\text{day}}$	

### 3 | RESULTS

In this section, we present the next-generation matrix for sea lice populations on salmon farms in the Broughton Archipelago, the construction of which is detailed in Section 2.4, and determine the relative patch contribution of each farm. We use this system to highlight to potential differences between the connectivity matrix, which only contains information surrounding the probability of dispersal from one farm to another, and the next-generation matrix, which combines dispersal between farms and local productivity of sea lice on a salmon farm. We then demonstrate how the next-generation matrix can be used to investigate the effect of changing demographic rates on growth and persistence in this system. Finally, in the context of salmon farm removal from the Broughton Archipelago, we investigate how the removal of habitat patches affects patch contribution and persistence in this sea louse metapopulation.

The next-generation matrix for sea lice populations on salmon farms in the Broughton Archipelago is shown in Figure 3a. The patch contribution of each salmon farm is given by  $R_j$ , the  $j$ th column sum, which is presented at the bottom of each column. We also present the row sums to the left each row, to identify which farms are receiving the most sea lice from other farms in the region. The farms producing the most lice per generation are 2, 17, 18, and 3, and the farms receiving the most lice are 3, 16, and 7.

To better understand the details and construction of the next-generation matrix, we also present the connectivity matrix for this system in Figure 4 and the productivity (total number of new larvae produced from one attached chalimus louse) of each farm in Table 2. The  $(i, j)$ th entry of the connectivity matrix,  $c_{ij}$ , is the infectious density over farm  $i$  of lice leaving farm  $j$ . The  $(i, j)$ th entry of the next-generation matrix,  $k_{ij}$  (Equation 6), is constructed by multiplying the productivity of farm  $j$  (Equation 10) by  $p^{ij} = \beta \times 0.01 \times c_{ij}$ , the probability that a larvae leaving farm  $j$  attaches on farm  $i$  (Equation 9). The farms with the largest column sums of the connectivity matrix are, in declining order, farms 18, 2, and 17. However, farm 2 has a higher

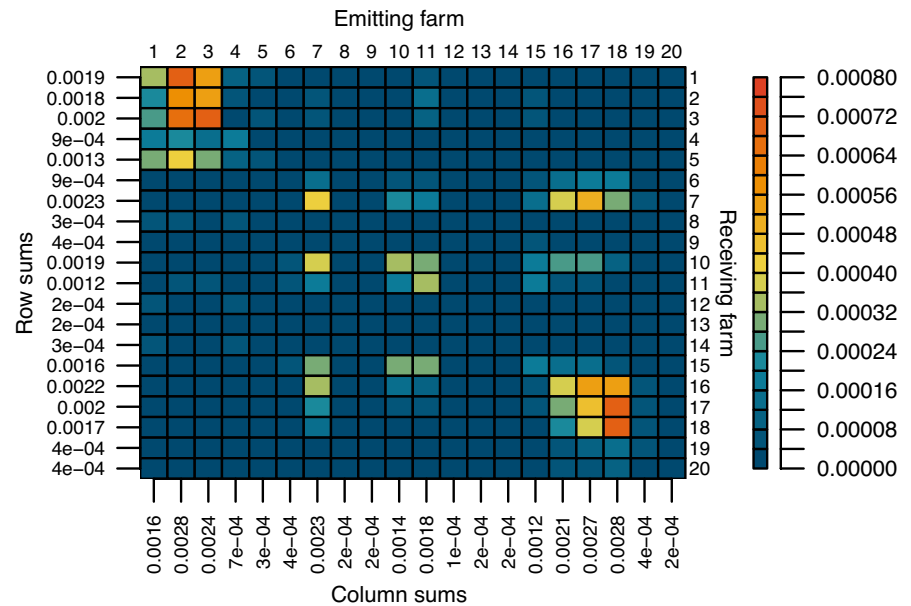
productivity than farms 17 or 18, and when the productivity of farm 2 is multiplied by the connectivity then it becomes the largest source of sea lice in the region, as identified by the next-generation matrix.

A further look into the connectivity matrix and productivity table provides more insight into the underlying drivers of the contribution of each farm to the sea louse population. Many of the farms with low connectivity also have low productivity, which may be due to the fact that temperature and salinity affect on farm demographic rates as well as survival and maturation in the particle tracking simulation which underlies the connectivity matrix. However, there are certain farms, such as 11 and 15, which have a comparably high productivity compared to their connectivity. These farms are located in favorable environments with respect to temperature and salinity, but low connectivity due to either distance from other farms or unfavorable currents prevents these farms from acting as larger sources of sea lice.

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

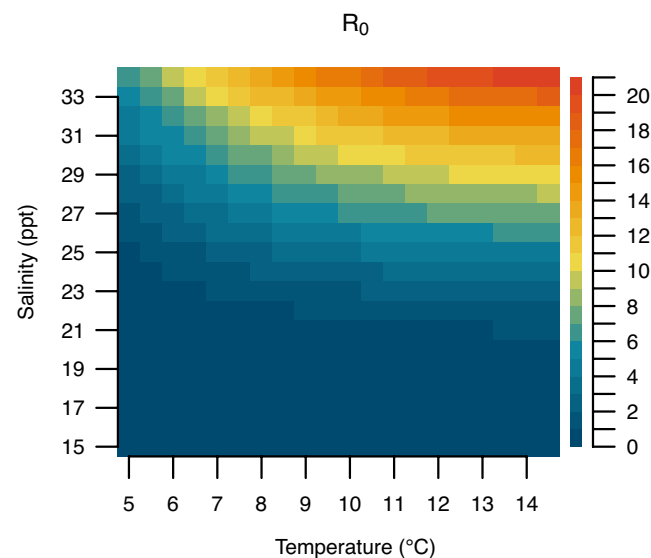
In light of the removal of salmon farms in the Broughton Archipelago we also create a next-generation matrix consisting only of the farms which will remain after 2023 subject to First Nations and government approval, shown in Figure 3b, and examine the differences between this matrix and the next-generation matrix with all farms. Most of the farms which are acting as the largest sources

**FIGURE 4** The connectivity matrix,  $C$ , for sea lice larvae dispersing between salmon farms, created by averaging all connectivity matrices from Cantrell et al. (2018) over the simulation period. The  $(i, j)$ th entry is the infectious density of larvae ( $1/\text{km}^2$ ) over farm  $i$  that have left from farm  $j$ . Column and row sums are shown below and to the left of each column and row, respectively.



**TABLE 2** The number of new larvae produced on each farm by a single louse starting in the chalimus stage. The first and third rows contain the farm numbers and the second row and fourth rows contain the number of larvae produced.

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



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

of sea lice in this region remain, with the exception of farm 11, the fifth largest source in the original network, which has now been removed. Farm 11, while not the largest source, did have the largest betweenness score based solely on the connectivity matrix (Cantrell et al., 2018), and thus may have been acting as a connecting farm between the two large clusters of source farms. However, since none of the other large source farms have been removed, the overall growth of the metapopulation has only decreased from  $R_0 = 2.33$  (original next-generation matrix) to  $R_0 = 2.25$ . Again these numbers are calculated using a very rough estimate of the arrival rate onto farms, detailed in Section 2.4.2, and thus it is their relative similarity that is important, rather than the absolute magnitude.

## 4 | DISCUSSION

In this paper, we demonstrated how to use the next-generation matrix to calculate the contribution of each habitat patch to the metapopulation and measure the overall persistence of a metapopulation. We detailed the construction of the next-generation matrix under different model structures to demonstrate the breadth of the approach to several systems. We then constructed the next-generation matrix under an age-dependent modeling framework for sea lice populations on

salmon farms in the Broughton Archipelago to illustrate how this approach can be applied to a real system. We determined which salmon farms may be acting as the largest sources of sea lice in this region, how the metapopulation will change once certain farms are removed, and examined the effect of temperature and salinity on the relative growth and persistence of this metapopulation.

Next-generation matrices have been used extensively in epidemiology to study the spread of infectious diseases but have recently been introduced in spatial ecology (Huang & Lewis, 2015; Krkošek & Lewis, 2010; McKenzie et al., 2012) and evolutionary analysis (Hurford et al., 2010). One of the key benefits of using next-generation matrices in epidemiology is that the basic reproduction number,  $R_0$ , for a disease can be calculated as the spectral radius of the next-generation matrix, which is often more analytically tractable than calculating the dominant eigenvalue of the full system to determine the spread. In ecology, one main advantage of this approach is that the mathematical calculation of  $R_0$  can be broken down into biologically relevant quantities, for example, the contribution of different dispersal pathways to growth in a population (De-Camino-Beck & Lewis, 2007) or the contribution of populations on different habitat patches (Harrington & Lewis, 2020). While not novel, we hope next-generation matrices can be used more frequently as a simple and easily biologically interpretable method to measure the contribution of local habitat patches to a metapopulation and determine overall persistence.

We are also by no means the first to attempt to calculate the contribution of a local population, classify patches into sources and sinks, or attempt to measure the persistence of metapopulations. In the context of low densities, Pulliam (1988) defined a source as a habitat patch that would grow in the absence of immigration and emigration and a sink as a habitat patch that would decline in the absence of immigration and emigration. This is similar to only using the entries along the diagonal of the next-generation matrix to classify sources or sinks, except growth or decline was measured after one time step in a discrete time model, rather than one generation. However, as discussed at the end of Section 2.2, it is possible to have a metapopulation composed only of sinks based on this definition ( $k_{ij} < 1$  for all  $i$ ), that persists.

Recognizing that dispersal between patches should also be considered when classifying habitat patches as sources or sinks, both Runge et al. (2006) and Figueira and Crowder (2006) defined new metrics to classify habitat patches that track the contribution of adults on a patch in one time step to the total population on all patches in the next timestep. These metrics are similar to our patch contribution metric  $R_j$ , except they measure the contribution over one time step rather than one generation, similar to using the dominant eigenvalue of the projection matrix  $A$  to determine the stability of the discrete system  $n(t+1) = An(t)$ , rather than the spectral radius of the next-generation matrix  $K$ . However, the calculations can become complicated if the population is stage-structured and the number of stages and/or patches is large (Appendix A, Runge et al. (2006)).

There are other measures of persistence in metapopulations which also track the number of new individuals contributed to the

metapopulation after one generation from an initial individual on one patch. Krkošek and Lewis (2010) define a next-generation operator for general heterogeneous populations which tracks the number of new adults produced in the population from one initial adult after one generation. If  $b_j$  is the reproductive output on patch  $j$ ,  $p_{ij}$  is the probability of larvae dispersing between patch  $j$  successfully arrives on patch  $i$ , and  $a_i$  is the survival to maturity on patch  $i$ , then the contribution of patch  $j$  to patch  $i$  according to Krkošek and Lewis (2010) can be calculated as  $b_j p_{ij} a_i$ . This patch contribution metric cannot be calculated directly from the next-generation matrices shown in this paper unless there is only one stage. Burgess et al. (2014), following Hastings and Botsford (2006), track the number of new recruits on all patches produced over the lifetime of an initial recruit to determine metapopulation persistence in discrete time models. If recruits are defined to be individuals that have survived to maturity, then the entries of their "realized connectivity matrix" are the same as from Krkošek and Lewis (2010):  $b_j p_{ij} a_i$ . If recruits are defined to be newly arrived individuals, then the entries are the same as the entries of the next-generation matrix in this paper:  $a_i b_j p_{ij}$ . One could also track the number of larvae produced on patch  $i$  from one initially released larva from patch  $j$ , in which case the generational output of patch  $j$  on patch  $i$  would be  $p_{ij} a_i b_j$ .

The contribution of each habitat patch to the metapopulation will depend on the stage at which the generational output is measured, but the persistence of the metapopulation is equivalent under all of these frameworks. This is because there is only one component of the life cycle where movement can occur between patches (larval stage) and at all other stages individuals remain on a patch. Let  $P$  be a matrix with entries  $p_{ij}$ ,  $B$  be a diagonal matrix with entries  $b_j$  and  $A$  be a diagonal matrix with entries  $a_j$ , where  $p_{ij}$ ,  $b_j$ , and  $a_j$  are the same as the preceding paragraph. If we measure generational output starting at the first attached stage, as we do in this paper, the next-generation matrix can be written as  $PAB$ , if we measure generational output starting at the larval stage, the matrix can be written as  $ABP$ , and if we measure generational output starting at the adult stage according to Krkošek and Lewis (2010) or Burgess et al. (2014) then the matrix can be written as  $APB$ . This is because when  $P$  is multiplied by a diagonal matrix on the right, the entries of the diagonal matrix multiply each column of  $P$  and when  $P$  is multiplied on the left, the entries multiply each row. Now for any two matrices  $X$  and  $Y$ ,  $XY$  and  $YX$  have the same eigenvalues, and because matrix multiplication is associative each of the matrices  $PAB$ ,  $ABP$ , and  $APB$  all have the same eigenvalues as well, and, therefore, also the same spectral radius. Therefore, in any of the formulations, the metapopulation will only persist if  $\rho(PAB) = \rho(ABP) = \rho(APB) > 1$ .

While the spectral radius of the next-generation matrix determines the overall persistence of the metapopulation and the column sums determine whether a habitat patch is a source or a sink, it is important to note that these quantities ( $R_j$  and  $R_0$ ) are measured in generational time, rather than chronological time. Therefore, if there are large differences in generation time between patches, there could be situations in which one patch is producing more new individuals per generation than another patch, even if it is producing fewer

individuals in chronological time. This may occur if one patch has a higher productivity, but longer generation time, than another patch. One method to estimate whether patch contribution is the same on generation and chronological time is to estimate the growth rate of the patch,  $r_j$ , by assuming that at low population densities, the population will grow exponentially and that  $R_j$  individuals will be produced after one generation, so  $e^{r_j T_j} = R_j$ , where  $T_j$  is the generation time of the patch. If the relative ordering of the growth rates,  $r_j$ , is the same as the relative ordering of the column sums of the next-generation matrix,  $R_j$ , then the relative contributions of the habitat patches will be the same in generational and chronological time. If the relative ordering is different, then the classification of habitat patches into sources and sinks using the column sums of the next-generation matrix still applies, only in this case one source patch could be producing more larvae per generation than another source patch, but less in chronological time.

In the case where there are large differences in the number of larvae produced from different patches when measured in generational and chronological time, and the ordering of source or sink patches is important, it may be more suitable to simulate the metapopulation model directly to understand which patches are contributing the most to the growth or decline of the metapopulation. If the metapopulation model is formulated in discrete time, or in continuous time with no age structure, then this is straightforward assuming the number of stages and patches is not too large. In discrete time, patch contribution could also be calculated directly following Runge et al. (2006) and Figueira and Crowder (2006). However, if it is necessary to capture age-dependent maturation and survival, as in the case of sea lice, then simulating the metapopulation model requires either simulating many coupled partial differential equations or converting the model into discrete time. If converting to discrete time, then the number of equations is given by # stages  $\times$  # patches  $\times$  age in the longest stage. In the case of sea lice in the Broughton where some stages can last 100 days, this would mean simulating a system of  $100 \times 4 \times 20 = 8000$  equations. Additionally, to simulate the model accurately, the time of dispersal between patches must be estimated.

In the case where either the classification of habitat patches into sources or sinks is of primary importance, or when the relative ordering of patch contribution is the same in generational or chronological time, then using the next-generation matrix to determine sources and sinks has many advantages over simulating the metapopulation model. The entries of the next-generation matrix have a direct, simple, biological interpretation in terms of number of new individuals produced after one generation and the framework is the same for discrete time, continuous time, and age-structured systems of equations. The next-generation matrix is easy to calculate, even when maturation and survival in each stage are dependent on stage age. Perhaps most importantly, there may be cases where quantities required to calculate the next-generation matrix are known but the individual rates required to calculate the full-time-dependent system are not. For example, if the overall lifetime productivity of a patch is known, but not the stage-specific vital

rates, or if the overall probability that larvae successfully disperse between patches has been estimated, but not the timing of dispersal, which is the case for sea lice populations on salmon farms in the Broughton Archipelago.

For sea lice on salmon farms in the Broughton, the relative ordering of the population growth rate of each patch,  $r_j$ , and the column sums of the next-generation matrix,  $R_j$ , is the same for all but three pairs of farms: farms 2 and 17, 11 and 16, and 4 and 19. The difference in the calculated growth rate  $r_j$  was  $<0.0008$  between each pair and most farms treat their salmon yearly to reduce sea lice (Bateman et al., 2016; Peacock et al., 2013), so we expect the differences in population size between pairs to be small. Here we used a generation time of  $T_j = \left( \sum_{n=1}^2 \int_0^\infty S_j(a) M_n^j(a) da \right) + \int_0^\infty S_j(a) da$  to calculate  $r_j$ . If we were to split the farms into categories, where large sources are those with  $R_j > 3$ , small sources are those with  $1 < R_j < 3$ , and 'sinks' are those with  $R_j < 1$ , then no farms would switch categories if the rankings were instead based on growth rate. Therefore while a few slight differences in ordering may exist if the output is measured on a chronological scale, the overall findings of which farms are the largest sources of sea lice in the Broughton remain the same irrespective of the timescale used for measurement.

When calculating the next-generation matrix for sea lice populations on salmon farms in the Broughton Archipelago, we believe there are two main sources of uncertainty: uncertainty in the arrival rate,  $\beta$ , and uncertainty in the number of larval sea lice that remain inside the salmon farm in the particle tracking model. In lab studies, the proportion of infective copepodid lice that attach to salmon varies enormously, from as low as 0.5% to as high as  $>80\%$ , with recent estimates between 12% and 56% depending on temperature (Skern-Mauritzen et al., 2020). Moreover, the time during which copepodids are given to attach in these experiments also varies, from a few seconds, to 8 h. If we assume an exponential waiting time model,  $P = 1 - e^{-\beta t}$ , where  $P$  is the proportion of copepodids that have attached to salmon, then these lab studies give estimates of  $\beta$  between 3/day and 47/day, which are lower than the estimate of 100/day which we use in this paper. Any estimates less than 44/day would result in a basic reproduction number,  $R_0 < 1$ , which would mean that the overall sea louse population is not increasing.

However, there is abundant evidence that in the Broughton Archipelago sea lice populations on many salmon farms will grow exponentially in the absence of treatment by farm managers (Bateman et al., 2016; Cantrell et al., 2021; Godwin et al., 2021; Krkošek et al., 2012; Rogers et al., 2013). This discrepancy may be due to the fact that while the hydrodynamic and particle tracking models give good estimates of sea louse dispersal between farms (Cantrell et al., 2021), they likely underestimate the number of sea louse larvae that remain inside a salmon farm. In the hydrodynamic model, the location of the farm does not alter the local hydrodynamics that occurs within the farm, but in reality, the presence of salmon swimming within the net pens may alter the local hydrodynamics sufficiently that some proportion of larvae remains within the farm. To correct this, it would be possible to add a diagonal matrix to the next-generation matrix presented in this paper, where each entry

on the diagonal is the productivity of farm  $j$ , multiplied by the number of larvae which remain within and successfully attach to salmon on the farm. Because the most productive farms are also the most highly connected in the Broughton, this would likely not change the relative contribution of each farm, but would increase the absolute contribution of each farm as well as  $R_0$ .

In addition to the sources of uncertainty in the next-generation matrix, there are some other technical aspects which should be considered when applying the next-generation matrix to calculate patch contribution of metapopulation persistence in specific systems. In order to use the next-generation matrix to calculate patch contribution or persistence, we are assuming that our system is autonomous and that the demographic rates do not change with time. In reality for most systems, including sea lice on salmon farms, environmental variables will fluctuate over time, potentially changing the demographic rates of the population (Cantrell et al., 2020). In our case temperature and salinity change over the course of the spring, but we calculate the next-generation matrix using the mean temperature and salinity that sea lice experience during the particle tracking simulation window. Therefore the entries in our next-generation matrix may be slightly different than the true number of newly attached lice produced on other farms from one initially attached louse, depending on the exact time that the louse began its lifecycle during the spring. For temporally oscillating systems it is possible to correct for this difference, though the entries of the next-generation matrix no longer have a simple form and must be computed computationally (Rittenhouse et al., 2016).

Over time periods where temperature and salinity are relatively constant, we can use the next-generation matrix to examine the overall effect of environmental variables on the growth of the metapopulation. The demographic rates at each stage depend explicitly on temperature and salinity but alone, or in a full system of equations, it can be difficult to examine the overall effect of changing environmental variables. However, the basic reproduction number  $R_0$ , calculated from the spectral radius of the next-generation matrix, provides a useful metric of the overall effect. We can infer how the growth of the metapopulation may change among seasons, or as the ocean warms. For sea lice on salmon farms in the Broughton Archipelago, the effect of temperature and salinity on  $R_0$  is very similar to previous results found for a single farm (Rittenhouse et al., 2016). With updated temperature and salinity at each farm, we could calculate the change in growth over years in the springtime, which may help explain the recent sea louse outbreaks during warm years in the Broughton Archipelago (Bateman et al., 2016). When examining the effect of temperature and salinity on  $R_0$  we do not rerun the hydrodynamic and particle tracking models to recreate the connectivity matrices under new temperature and salinity scenarios, as this is very computationally intensive, but we expect connectivity to increase as temperature and salinity increase due to higher survival of sea lice and faster maturation. However, we believe it would be valuable to rerun hydrodynamic models under different projected ocean scenarios, to investigate the precise changes in connectivity that may occur.

Specific to the management of sea lice populations on salmon farms in the Broughton Archipelago, there are several insights to be gained from our results. The first is that the farms in the most productive environments are also the most highly connected, and thus become the largest contributors of sea lice to this sea louse metapopulation. They occur in two main clusters (shown in Figure 3) and both clusters of farms will remain in the Broughton Archipelago in the current removal plan, subject to First Nations and governmental approval (Brownsey & Chamberlain, 2018). It should be noted that these farms may not necessarily be producing the most number of lice compared to other farms in the region at a given time if their louse population is currently lower than other farms, rather they are the farms that have the largest potential to contribute to spread when sea louse numbers are even across farms. However, due to the highly connected nature of these clusters, coordinated treatment between farms in the clusters or all farms in the region could reduce the number of treatments required and number of sea lice produced on all farms (Peacock et al., 2016). An interesting avenue of future research would be to connect the productivity of the remaining farms in the Broughton Archipelago with the Kernel Density Estimates of sea louse dispersal from Cantrell et al. (2018) to measure the exposure of migrating wild salmon to sea louse infection from these farms.

## AUTHOR CONTRIBUTIONS

**Peter D. Harrington:** Conceptualization (lead); data curation (supporting); formal analysis (lead); investigation (lead); software (equal); writing – original draft (lead); writing – review and editing (equal). **Danielle L. Cantrell:** Data curation (lead); software (equal); writing – original draft (supporting); writing – review and editing (equal). **Mark A. Lewis:** Conceptualization (supporting); formal analysis (supporting); supervision (lead); writing – original draft (supporting); writing – review and editing (equal).

## ACKNOWLEDGMENTS

The authors would like to thank the members of the Lewis lab for many helpful discussions and suggestions. PDH gratefully acknowledges an NSERC-CGSM scholarship, Queen Elizabeth II scholarship, and an Alberta Graduate Excellence Scholarship; and MAL gratefully acknowledges the Canada Research Chair program and an NSERC Discovery Grant.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The data and code used to conduct the analyses reported in this paper are available through the Dryad Digital Repository: <https://doi.org/10.5061/dryad.jm63xsjfv> (Harrington et al., 2023).

## ORCID

Peter D. Harrington  <https://orcid.org/0000-0002-5968-8984>

## REFERENCES

- Abolofia, J., Asche, F., & Wilen, J. E. (2017). The cost of lice: Quantifying the impacts of parasitic sea lice on farmed salmon. *Marine Resource Economics*, 32(3), 329–349. <https://doi.org/10.1086/691981>
- Aldrin, M., Huseby, R. B., Stien, A., Grøntvedt, R. N., Viljugrein, H., & Jansen, P. A. (2017). A stage-structured Bayesian hierarchical model for salmon lice populations at individual salmon farms – Estimated from multiple farm data sets. *Ecological Modelling*, 359, 333–348.
- Bateman, A. W., Peacock, S. J., 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. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(8), 1164–1172.
- Berman, A., & Plemmons, R. J. (1994). *Nonnegative matrices in the mathematical sciences*. Society for Industrial and Applied Mathematics.
- Brooks, K. M. (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. *Reviews in Fisheries Science & Aquaculture*, 13(3), 177–204.
- Brownsey, L., & Chamberlain, R. (2018). *Collaborative solutions for fin-fish aquaculture farms in the Broughton area, Steering Committee Recommendations*. Government of British Columbia. [https://news.gov.bc.ca/files/Steering\\_Committee\\_Broughton\\_Recommendations](https://news.gov.bc.ca/files/Steering_Committee_Broughton_Recommendations)
- Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W., & Botsford, L. W. (2014). Beyond connectivity: How empirical methods can quantify population persistence to improve marine protected-area design. *Ecological Applications*, 24(2), 257–270. <https://doi.org/10.1890/13-0710.1>
- Cantrell, D., Filgueira, R., Revie, C. W., Rees, E. E., Vanderstichel, R., Guo, M., Foreman, M. G., Wan, D., & Grant, J. (2020). The relevance of larval biology on spatiotemporal patterns of pathogen connectivity among open-marine salmon farms. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(3), 505–519. <https://doi.org/10.1139/cjfas-2019-0040>
- Cantrell, D., Vanderstichel, R., Filgueira, R., Grant, J., & Revie, C. W. (2021). Validation of a sea lice dispersal model: Principles from ecological agent-based models applied to aquatic epidemiology. *Aquaculture Environment Interactions*, 13, 65–79. <https://doi.org/10.3354/AEI00390>
- Cantrell, D. L., Rees, E. E., Vanderstichel, R., Grant, J., Filgueira, R., & Revie, C. W. (2018). The use of kernel density estimation with a biophysical model provides a method to quantify connectivity among Salmon farms: Spatial planning and management with epidemiological relevance. *Front Vet Sci*, 5(269), 1–14.
- Connors, B. M., Juarez-Colunga, E., & Dill, L. M. (2008). Effects of varying salinities on *Lepeophtheirus salmonis* survival on juvenile pink and chum salmon. *Journal of Fish Biology*, 72(7), 1825–1830. <https://doi.org/10.1111/j.1095-8649.2008.01839.x>
- Costello, C., Rassweiler, A., Siegel, D., De Leo, G., Micheli, F., Rosenberg, A., & Gaines, S. D. (2010). The value of spatial information in MPA network design. *PNAS*, 107(43), 18294–18299. <https://doi.org/10.1073/pnas.0908057107>
- Costello, M. J. (2006). Ecology of sea lice parasitic on farmed and wild fish. *Trends in Parasitology*, 22(10), 475–483.
- Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, 1, 443–466. <https://doi.org/10.1146/annurev.marine.010908.163757>
- Cushing, J. M., & Yicang, Z. (1994). The net reproductive value and stability in matrix population models. *Natural Resource Modeling*, 8(4), 297–333. <https://doi.org/10.1111/j.1939-7445.1994.tb00188.x>
- De-Camino-Beck, T., & Lewis, M. A. (2007). A new method for calculating net reproductive rate from graph reduction with applications to the control of invasive species. *Bulletin of Mathematical Biology*, 69(4), 1341–1354. <https://doi.org/10.1007/s11538-006-9162-0>
- Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4), 365–382.
- Diekmann, O., Heesterbeek, J. A. P., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface*, 7(47), 873–885.
- Feng, Z., & Thieme, H. R. (2000). Endemic models with arbitrarily distributed periods of infection I: Fundamental properties of the model. *SIAM Journal on Applied Mathematics*, 61(3), 803–833.
- Figueira, W. F., & Crowder, L. B. (2006). Defining patch contribution in source-sink metapopulations: The importance of including dispersal and its relevance to marine systems. *Population Ecology*, 48(3), 215–224. <https://doi.org/10.1007/s10144-006-0265-0>
- Foreman, M. G. G., Chandler, P. C., Stucchi, D. J., Garver, K. A., Guo, M., Morrison, J., & Tuele, D. (2015). *The ability of hydrodynamic models to inform decisions on the siting and management of aquaculture facilities in British Columbia*. Tech. Rep., Fisheries and Oceans Canada. [https://publications.gc.ca/collections/collection\\_6912015/mpo-dfo/Fs70-5-2015-005-eng.pdf](https://publications.gc.ca/collections/collection_6912015/mpo-dfo/Fs70-5-2015-005-eng.pdf)
- Foreman, M. G. G., Czajko, P., Stucchi, D. J., & Guo, M. (2009). A finite volume model simulation for the Broughton archipelago, Canada. *Ocean Modelling*, 30(1), 29–47.
- Garavelli, L., White, J. W., Chollett, I., & Chérubin, L. M. (2018). Population models reveal unexpected patterns of local persistence despite widespread larval dispersal in a highly exploited species. *Conservation Letters*, 11(5), e12567. <https://doi.org/10.1111/conl.12567>
- Godwin, S. C., Krkošek, M., Reynolds, J. D., & Bateman, A. W. (2021). Bias in self-reported parasite data from the salmon farming industry. *Ecological Applications*, 31(1), 1–10. <https://doi.org/10.1002/eap.2226>
- Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396, 41–49.
- Harrington, P. D., Cantrell, D. L., & Lewis, M. A. (2023). Data from: Next-generation matrices for marine metapopulations: The case of sea lice and salmon farms. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.jm63xsjfv>
- Harrington, P. D., & Lewis, M. A. (2020). A next-generation approach to calculate source-sink dynamics in marine metapopulations. *Bulletin of Mathematical Biology*, 82(1), 1–44. <https://doi.org/10.1007/s11538-019-00674-1>
- Hastings, A., & Botsford, L. W. (2006). Persistence of spatial populations depends on returning home. *Proceedings of the National Academy of Sciences of the United States of America*, 103(15), 6067–6072. <https://doi.org/10.1073/pnas.0506651103>
- Hillen, T., Greese, B., Martin, J., & de Vries, G. (2015). Birth-jump processes and application to forest fire spotting. *Journal of Biological Dynamics*, 9, 104–127. <https://doi.org/10.1080/17513758.2014.950184>
- Huang, Q., Jin, Y., & Lewis, M. A. (2016).  $R_0$  analysis of a benthic-drift model for a stream population. *SIAM Journal on Applied Dynamical Systems*, 15(1), 287–321. <https://doi.org/10.1137/15M1014486>
- Huang, Q., & Lewis, M. A. (2015). Homing fidelity and reproductive rate for migratory populations. *Theoretical Ecology*, 8(2), 187–205. <https://doi.org/10.1007/s12080-014-0243-7>
- Hurford, A., Cownden, D., & Day, T. (2010). Next-generation tools for evolutionary invasion analyses. *Journal of the Royal Society Interface*, 7(45), 561–571.
- Husband, B. C., & Barrett, S. C. H. (1996). A metapopulation perspective in plant population biology. *Journal of Ecology*, 84(3), 461–469. <https://doi.org/10.2307/2261207>

- Kritzer, J. P., & Sale, P. F. (2004). Metapopulation ecology in the sea: From Levins' model to marine ecology and fisheries science. *Fish and Fisheries*, 5(2), 131–140. <https://doi.org/10.1111/j.1467-2979.2004.00131.x>
- Krkošek, M., Bateman, A. W., Probošycz, S., & Orr, C. (2012). Dynamics of outbreak and control of salmon lice on two salmon farms in the Broughton archipelago, British Columbia. *Aquaculture Environment Interactions*, 1(2), 137–146. <https://doi.org/10.3354/aei00014>
- Krkošek, M., Connors, B. M., Morton, A., Lewis, M. A., Dill, L. M., & Hilborn, R. (2011). Effects of parasites from salmon farms on productivity of wild salmon. *Proceedings of the National Academy of Sciences of the United States of America*, 108(35), 14700–14704.
- Krkošek, M., Ford, J. S., Morton, A., Lele, S., Myers, R. A., & Lewis, M. A. (2007). Declining wild salmon populations in relation to parasites from farm Salmon. *Science*, 318(2), 1772–1775.
- Krkošek, M., & Lewis, M. A. (2010). An  $R_0$  theory for source-sink dynamics with application to Dreissena competition. *Theoretical Ecology*, 3(1), 25–43. <https://doi.org/10.1007/s12080-009-0051-7>
- Krkošek, M., Lewis, M. A., Morton, A., Frazer, L. N., & Volpe, J. P. (2006). Epizootics of wild fish induced by farm fish. *Proceedings of the National Academy of Sciences of the United States of America*, 103(42), 15506–15510.
- Krkošek, M., Lewis, M. A., & Volpe, J. P. (2005). Transmission dynamics of parasitic sea lice from farm to wild salmon. *Proceedings of the Biological Sciences*, 272(1564), 689–696. <https://doi.org/10.1098/rspb.2004.3027>
- Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15(3), 237–250. <https://doi.org/10.1088/2058-9565/aaba1a>
- Lewis, M. A., Shuai, Z., & van den Driessche, P. (2019). A general theory for target reproduction numbers with applications to ecology and epidemiology. *Journal of Mathematical Biology*, 78(7), 2317–2339.
- Li, C. K., & Schneider, H. (2002). Applications of Perron-Frobenius theory to population dynamics. *Journal of Mathematical Biology*, 44(5), 450–462.
- Marty, G. D., Saksida, S. M., & Quinn, T. J. (2010). Relationship of farm salmon, sea lice, and wild salmon populations. *Proceedings of the National Academy of Sciences of the United States of America*, 107(52), 22599–22604. <https://doi.org/10.1073/pnas.1009573108>
- Mckenzie, H. W., Jin, Y., Jacobsen, J., & Lewis, M. A. (2012).  $R_0$  analysis of a spatiotemporal model for a stream population. *SIAM Journal on Applied Dynamical Systems*, 11(2), 567–596. <https://doi.org/10.1137/100802189>
- Peacock, S. J., Bateman, A. W., Krkošek, M., & Lewis, M. A. (2016). The dynamics of coupled populations subject to control. *Theoretical Ecology*, 9, 1–16. <https://doi.org/10.1007/s12080-016-0295-y>
- Peacock, S. J., Krkošek, M., Bateman, A. W., & Lewis, M. A. (2020). Estimation of spatiotemporal transmission dynamics and analysis of management scenarios for sea lice of farmed and wild salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(1), 55–68. <https://doi.org/10.1139/cjfas-2019-0036>
- Peacock, S. J., Krkošek, M., Probošycz, S., Orr, C., & Lewis, M. A. (2013). Cessation of a salmon decline with control of parasites. *Ecological Applications*, 23(3), 606–620. <https://doi.org/10.1890/12-0519.1>
- Puckett, B. J., & Eggleston, D. B. (2016). Metapopulation dynamics guide marine reserve design: Importance of connectivity, demographics, and stock enhancement. *Ecosphere*, 7(6), e01322. <https://doi.org/10.1002/ecs2.1322>
- Pulliam, H. R. (1988). Sources, sinks and population regulation. *The American Naturalist*, 132(5), 652–661. <https://doi.org/10.1086/284880>
- Revie, C. W., Robbins, C., Gettinby, G., Kelly, L., & Treasurer, J. W. (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. *Journal of Fish Diseases*, 28(10), 603–614.
- Rittenhouse, M. A., Revie, C. W., & Hurford, A. (2016). A model for sea lice (*Lepeophtheirus salmonis*) dynamics in a seasonally changing environment. *Epidemics*, 16, 8–16. <https://doi.org/10.1016/j.epidem.2016.03.003>
- Rogers, L. A., Peacock, S. J., McKenzie, P., DeDominicis, S., Jones, S. R., Chandler, P., Foreman, M. G., Revie, C. W., & Krkošek, M. (2013). Modeling parasite dynamics on farmed Salmon for precautionary conservation Management of Wild Salmon. *PLoS One*, 8(4), e60096.
- Runge, J. P., Runge, M. C., & Nichols, J. D. (2006). The role of local populations within a landscape context: Defining and classifying sources and sinks. *The American Naturalist*, 167(6), 925–938. <https://doi.org/10.1086/503531>
- Safranyik, L., & Carroll, A. L. (2007). *The biology and epidemiology of the mountain pine beetle in lodgepole pine forests* (pp. 3–66). Canadian Forest Service.
- Skern-Mauritzen, R., Sissener, N. H., Sandvik, A. D., Meier, S., Sævik, P. N., Skogen, M. D., Vågseth, T., Dalvin, S., Skern-Mauritzen, M., & Bui, S. (2020). Parasite development affect dispersal dynamics; infectivity, activity and energetic status in cohorts of salmon louse copepodids. *Journal of Experimental Marine Biology and Ecology*, 530–531, 151429. <https://doi.org/10.1016/j.jembe.2020.151429>
- Stien, A., Bjørn, P. A., Heuch, P. A., & Elston, D. A. (2005). Population dynamics of salmon lice *Lepeophtheirus salmonis* on Atlantic salmon and sea trout. *Marine Ecology Progress Series*, 290, 263–275.
- Stige, L., Helgesen, K., Viljugrein, H., & Qviller, L. (2021). A statistical mechanistic approach including temperature and salinity effects to improve salmon lice modelling of infestation pressure. *Aquaculture Environment Interactions*, 13, 339–361. <https://doi.org/10.3354/aei00410>
- Stucchi, D., Guo, M., Foreman, M., Czajko, P., Galbraith, M., Mackas, D., & Gillibrand, P. A. (2011). Modeling Sea lice production and concentrations in the Broughton archipelago, British Columbia. In S. Jones & R. Beamish (Eds.), *Salmon lice: An integrated approach to understanding parasite abundance and distribution* (1st ed., pp. 117–150). Wiley.
- van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1–2), 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)
- White, J. W., Botsford, L. W., Hastings, A., & Largier, J. L. (2010). Population persistence in marine reserve networks: Incorporating spatial heterogeneities in larval dispersal. *Marine Ecology Progress Series*, 398, 49–67. <https://doi.org/10.3354/meps08327>
- White, J. W., Carr, M. H., Caselle, J. E., Washburn, L., Woodson, C. B., Palumbi, S. R., Carlson, P. M., Warner, R. R., Menge, B. A., Barth, J. A., Blanchette, C. A., Raimondi, P. T., & Milligan, K. (2019). Connectivity, dispersal, and recruitment. *Oceanography*, 32(3), 50–59.
- White, J. W., Scholz, A. J., Rassweiler, A., Steinback, C., Botsford, L. W., Kruse, S., Costello, C., Mitarai, S., Siegel, D. A., Drake, P. T., & Edwards, C. A. (2013). A comparison of approaches used for economic analysis in marine protected area network planning in California. *Ocean and Coastal Management*, 74, 77–89. <https://doi.org/10.1016/j.ocecoaman.2012.06.006>
- Williams, D. W., & Liebhold, A. M. (2000). Spatial synchrony of spruce budworm outbreaks in eastern North America. *Ecology*, 81(10), 2753–2766.

**How to cite this article:** Harrington, P. D., Cantrell, D. L., & Lewis, M. A. (2023). Next-generation matrices for marine metapopulations: The case of sea lice on salmon farms. *Ecology and Evolution*, 13, e10027. <https://doi.org/10.1002/ece3.10027>

## APPENDIX 1

## Calculating the next-generation matrix for differential equation models

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

$$\frac{dn}{dt} = An \quad (11)$$

where  $\mathbf{n}$  is a population vector of size  $m \times l$  describing the population size on all patches at all stages and  $A$  is a matrix describing the population dynamics at low population densities. We model the population size at low densities because we are interested in identifying which patches are acting as sources and supporting the population at low densities and which are acting as sinks. Therefore, Equation (11) should be thought of as the linearization about the zero equilibrium of a potentially more complex model that may include density dependence. The elements of  $A$  contain the maturation and death rates of sedentary stages confined to patches as well as the rates of larval dispersal and attachment between patches.

The next-generation matrix for this system can be calculated by first decomposing  $A = F - V$  where  $F$  is a non-negative matrix with entries that describe the rates of larval birth and probability of dispersal between patches and attachment as the first sedentary phase, and  $V$  is a non-singular  $M$  matrix (Berman & Plemmons, 1994) with entries that describe the maturation rates between stages and death rates in a stage van den Driessche and Watmough (2002). Because  $V$  is a non-singular  $M$  matrix,  $V^{-1}$  is non-negative. Following van den Driessche and Watmough (2002) with notation from Diekmann et al. (2010), the next-generation matrix with large domain,  $K_L$ , can then be calculated as

$$K_L = FV^{-1}. \quad (12)$$

The elements of  $K_L$  contain the number of new individuals in each stage and patch produced by one initial new individual in each stages and patch. However, since new individuals are only produced in stage 1,  $K_L$  will only have  $l$  non-zero rows. If Equation (11) is arranged so that the populations of stage 1 individuals are in the first  $l$  rows, then  $K_L$  will have a  $l \times l$  submatrix in the upper left-hand corner (Diekmann et al., 2010). This submatrix is the next-generation matrix  $K$ , where the elements of  $K$ ,  $k_{ij}$ , give the number of new (stage 1) individuals produced on patch  $i$  from one initial stage 1 individual on patch  $j$ . In Equation (11), if the maturation rates from stage  $k$  to  $k+1$  on patch  $j$  are  $m_k^j$ , the death rates in stage  $k$  are  $d_k^j$ , the birth rate in the last stage on patch  $j$  is  $b^j$ , and the probability of successful attachment on patch  $i$  as a larvae leaving patch  $j$  is  $p^{ij}$ , then the entries of  $K$  can be written as:

$$k_{ij} = \underbrace{\prod_{k=1}^{m-1} \left( \frac{m_k^j}{m_k^j + d_k^j} \right)}_{\text{survival through sessile stages}} \times \underbrace{\left( \frac{b^j}{d_m^j} \right)}_{\text{\# larvae produced}} \times \underbrace{p^{ij}}_{\text{Pr(patch } j \text{ to patch } i)}. \quad (13)$$

## EXAMPLE OF THE NEXT-GENERATION MATRIX IN FOR A TWO PATCH METAPOPULATION

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

$$\begin{bmatrix} n_1^1 \\ n_2^1 \\ n_1^2 \\ n_2^2 \end{bmatrix}' = \begin{bmatrix} -d_1^1 - m_1^1 & 0 & b^1 p^{11} & b^2 f^{12} \\ 0 & -d_1^2 - m_1^2 & b^1 p^{21} & b^2 f^{22} \\ m_1^1 & 0 & -d_2^1 & 0 \\ 0 & m_2^1 & 0 & -d_2^2 \end{bmatrix} \begin{bmatrix} n_1^1 \\ n_2^1 \\ n_1^2 \\ n_2^2 \end{bmatrix},$$

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

$$F = \begin{bmatrix} 0 & 0 & b^1 p^{11} & b^2 p^{12} \\ 0 & 0 & b^1 p^{21} & b^2 p^{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} d_1^1 + m_1^1 & 0 & 0 & 0 \\ 0 & d_1^2 + m_1^2 & 0 & 0 \\ -m_1^1 & 0 & d_2^1 & 0 \\ 0 & -m_2^1 & 0 & d_2^2 \end{bmatrix}.$$

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

$$K_L = \begin{bmatrix} \frac{m_1^1}{m_1^1 + d_1^1} \frac{b^1}{d_2^1} p^{11} & \frac{m_2^1}{m_1^1 + d_1^1} \frac{b^2}{d_2^2} p^{12} & \frac{b^1}{d_2^1} p^{11} & \frac{b^2}{d_2^2} p^{12} \\ \frac{m_1^1}{m_1^1 + d_1^1} \frac{b^1}{d_2^1} p^{21} & \frac{m_2^1}{m_1^1 + d_1^1} \frac{b^2}{d_2^2} p^{22} & \frac{b^1}{d_2^1} p^{21} & \frac{b^2}{d_2^2} p^{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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

$$K = \begin{bmatrix} \frac{m_1^1}{m_1^1 + d_1^1} \frac{b^1}{d_2^1} p^{11} & \frac{m_2^1}{m_1^1 + d_1^1} \frac{b^2}{d_2^2} p^{12} \\ \frac{m_1^1}{m_1^1 + d_1^1} \frac{b^1}{d_2^1} p^{21} & \frac{m_2^1}{m_1^1 + d_1^1} \frac{b^2}{d_2^2} p^{22} \end{bmatrix}.$$

The contribution of patch 1 is, therefore,

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

and the contribution of patch 2 is

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

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

## APPENDIX 2

### Calculating the next-generation matrix for discrete time models

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

$$\mathbf{n}_{t+1} = \mathbf{A}\mathbf{n}_t \quad (14)$$

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

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

$$K_L = \mathbf{F}(\mathbf{I} - \mathbf{T})^{-1}. \quad (15)$$

If Equation (14) is arranged so that the populations of all stage 1 individuals are in the top  $l$  rows, then again the next-generation matrix,  $K$ , will be the  $l \times l$  submatrix in the upper left hand corner of  $K_L$ . In the discrete time framework if the probability of transitioning from stage  $k$  to  $k+1$  is  $m_k^j$ , the probability of remaining in stage  $k$  is  $s_k^j$  (diagonal entries of  $\mathbf{A}$ ), the fecundity in patch  $j$  is  $b^j$ , and the probability that a larvae leaving patch  $j$  successfully arrives on patch  $i$  is  $p^{ji}$ , then the entries of  $K$  can be written as

$$k_{ij} = \underbrace{\prod_{k=1}^{m-1} \left( \frac{m_k^j}{1 - s_k^j} \right)}_{\text{survival through sessile stages}} \times \underbrace{\left( \frac{b^j}{1 - s_m^j} \right)}_{\text{\# larvae produced}} \times \underbrace{p^{ji}}_{\text{Pr(patch } j \text{ to patch } i)}. \quad (16)$$

## APPENDIX 3

### Details of the model with age-dependent demography in

#### Section 2.3

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

$$n_k^i(t, a) = \begin{cases} \underbrace{B_k^i(t-a)S_k^i(a)M_k^i(a)}_{\text{entered at } t-a, \text{ survived to } a} & t > a \\ \underbrace{n_{k,0}^i(a-t) \frac{S_k^i(a)M_k^i(a)}{S_k^i(a-t)M_k^i(a-t)}}_{\text{present at } a-t, \text{ survived to } a} & 0 < t < a \\ \underbrace{n_{k,0}^i(a)}_{\text{initial density}} & t = 0 \end{cases} \quad k = 1, \dots, m-1, \quad (17)$$

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

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

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

where  $S_k^j(a)$  is the probability of survival in stage  $k$  on patch  $j$  at stage-age  $a$ ,  $m_k^j(a)$  is the maturation rate from stage  $k$  to  $k+1$ ,  $b^j(a)$  is the birth rate of larvae on patch  $j$ , and the probability of a larva leaving patch  $j$  and successfully attaching on patch  $i$  is  $p^{ji}$ .  $M_k^j(a)$  is the probability that an individual has not yet matured from stage  $k$  to  $k+1$  and can be calculated as  $M_k^j(a) = \exp\left(-\int_0^a m_k^j(\tau) d\tau\right)$ . Additional details of the model equations can be found in Harrington and Lewis (2020), though there the first larval stage is modeled explicitly and so the set of equations is slightly different.

To construct the next-generation matrix for this system, we need to track the number of stage 1 individuals produced on each patch from one initial stage 1 individual on a given patch. We briefly present the idea of the construction here, and more details can be found in Harrington and Lewis (2020).

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

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

$$\gamma_{ij}(t) = \int_0^t \int_0^{t-r_{m-1}} \int_0^{t-r_{m-1}-r_{m-2}} \dots \int_0^{t-\sum_{k=1}^{m-1} r_k} S_1^j(r_1)M_1^j(r_1)m_1^j(r_1) \dots S_{m-1}^j(r_{m-1})M_{m-1}^j(r_{m-1})m_{m-1}^j(r_{m-1})S_m^j\left(t - \sum_{k=1}^{m-1} r_k\right)b^j\left(t - \sum_{k=1}^{m-1} r_k\right)p^{ij} dr_1 \dots dr_{m-1}. \quad (18)$$

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

$$\int_0^\infty f(t) * g(t) dt = \int_0^\infty f(t) dt \int_0^\infty g(t) dt \quad (19)$$

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

$$k_{ij} = \left[ \prod_{k=1}^{m-1} \left( \int_0^\infty S_k^j(t)M_k^j(t)m_k^j(t) dt \right) \right] \left( \int_0^\infty S_m^j(t)b^j(t) dt \right) p^{ij}. \quad (20)$$