Title: A partly stage-structured model for the abundance of salmon lice in salmonid farms

Author: Aldrin M. Jansen P.A. Stryhn H.

PII: S1755-4365(18)30070-7
DOI: https://doi.org/doi:10.1016/j.epidem.2018.08.001
Reference: EPIDEM 311

To appear in:

Received date: 17-4-2018
Revised date: 12-7-2018
Accepted date: 16-8-2018

Please cite this article as: Aldrin M., Jansen P.A., Stryhn H., A partly stage-structured model for the abundance of salmon lice in salmonid farms, <*![CDATA[Epidemics]]*> (2018), https://doi.org/10.1016/j.epidem.2018.08.001

A partly stage-structured model for the abundance of salmon lice in salmonid farms

Aldrin, M.[a*], Jansen,P.A.[b], Stryhn, H.[c]

[a] Norwegian Computing Center, P.O.Box 114 Blindern, N-0314, Oslo, Norway

[b]Norwegian Veterinary Institute, P.O. Box 750 Sentrum N-0106, Oslo, Norway

[c] University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island, Canada C1A 4P3

* Corresponding author: Tel: +47 22 85 26 58, fax: +47 22 69 76 60,

Email address: magne.aldrin@nr.no

July 12, 2018

Abstract The parasitic salmon louse constrains growth in the Norwegian salmon farming industry through density dependent host-parasite interaction. Hence, there is a need for insight into how increases in salmon production, varying spatial organisation of the production and pest control strategies affect salmon louse population dynamics. Here we present a new salmon louse model for exploring effects of varying salmon farming conditions on spatio-temporal abundances of the parasite. The salmon louse model is partly stage-structured, comprising of i) adult female lice and ii) other mobile stages of lice. The abundance of adult

females depend on survival of females from previous weeks and recruitment from the other mobile group of lice. The other mobiles also depend on survival of other mobiles from previous weeks, as well as recruitment from the previous generation of reproducing adult females from the same farm or from farms in the neighbourhood. In addition, expected abundances of the two stage-groups are modelled as functions of biological and physical covariates. The model is fitted to weekly salmon farm data covering all marine farms producing salmonids along the Norwegian cost over the years 2012-2016. Among novel results from fitting the model are estimates of the time-lag structure representing recruitment of other mobile lice from the previous generation adult females for different temperatures. Furthermore, the model estimates how various factors affect the susceptibility of fish on farms to louse infection and effects of treatments to control infection. Finally, the model estimates density dependent effects of increasing the number of fish in farms and of increasing the numbers of farms, on the rate of recruitment of other mobile lice. Analytically, the parameters representing density dependencies suggest that few farms with many fish and large between farm distances is effective in terms of minimising the overall output of salmon lice infection.

KEY WORDS: Salmon lice; stage-structured model; aquaculture

ï»¿

1 Introduction

Farming of salmonids has grown to become one of the major export industries in Norway. The parasitic salmon louse, however, threatens the sustainability of the salmon farming industry (Taranger et al., 2015). In particular, the salmon louse,

Lepeophtheirus salmonis, constrains growth in the Norwegian salmon farming industry due to large scale host-parasite density dependence (Jansen et al., 2012; Kristoffersen et al., 2017). As an instrument to promote sustainable growth in the industry, the Norwegian coast was recently divided into 13 production-zones in which environmental criteria will determine future growth in the production of farmed salmonids. For the time being, the only criterion to decide future growth in salmon production within the production-zones is the local burden of salmon lice (Kristoffersen et al., 2017). The reasoning behind this is that spillover of salmon lice from farmed salmon represents a major threat to the viability of wild salmonid stocks, and that there is a density dependent association between farmed salmon hosts and the abundance of the salmon louse parasite (Forseth et al., 2017; Jansen et al., 2012). Accordingly, there is a strong incentive for the salmon farming industry to reduce salmon lice infections in salmon farms. Numerous new control methods and control strategies are adopted, with costs of louse control accelerating (Abolofia et al., 2017; Liu and Bjelland, 2014), but with limited insight into the cost-efficiency of various preventive actions. The latter is especially true when considering scales of multiple farms interacting through between farm spread of infectious lice (Aldrin et al., 2013). Key questions in this regard are how the population dynamics of the salmon louse will be affected by increasing production of farmed salmonids, the spatial organisation of farms or pest control strategies.

Exploring these questions is not simply a question of comparing different areas or different periods of production since each sample of the production history is the product of its own set of conditions. However, if you are able to implement the major driving forces for the host-parasite population dynamics in a predictive model, you can explore such questions in simulation experiments (Tildesley

et al., 2012; Brooks-Pollock et al., 2014; Pettersen et al., 2016). The Norwegian farmed salmon host and louse parasite association is well suited for this approach. Due to strict regulations of this production system, a wealth of data are reported to document farm populations of salmonids, abundances of parasites on these fishes, environmental data and anti-parasitic treatments, in space and time. Furthermore, farm management determines salmonid host population dynamics in this system. This means that we only need to model the parasite population, whereas the host population can be pre-determined. Finally, due to the massive dominance of farmed salmonids over wild salmonids in coastal areas of Norway (Jansen et al., 2012; Torrissen et al., 2013), it is reasonable to neglect any contribution from hosts other than farmed salmonids to the population dynamics of the salmonid specific salmon louse.

Thus motivated, the main goal in the present study is to develop a simulation model for exploring effects of varying farming conditions on spatio-temporal abundances of salmon lice on farmed salmon along the Norwegian coast. Variable farming conditions could be production volumes, the location of farms or pest management strategies. The salmon-louse model is partly stage structured, comprising of i) adult female lice and ii) other mobile stages of lice. Other mobiles consist of parasitic pre-adult stages and adult males, all capable of moving about on fish hosts. The expected abundances of the two stage-groups on farmed salmonids are modelled as functions of biologically motivated co-variates. Essentially, the adult females for a given week and farm depend on the survival of females from previous weeks and recruitment of females from the other mobile group of lice.

We do not anticipate that parasitic lice spread naturally between hosts (Hamre

et al., 2013). Although it has been suggested that parasitic lice may be spread in conjunction with anti-parasitic treatments, we restrict recruitment into adult females to be from other mobiles on the same fish and farm. The other mobiles depend on survival of other mobiles from previous weeks at the same farm. Recruitment into this group of stages, however, depends on reproduction from the prior generation adult females, along with the planktonic spread of infectious copepodites (Hamre et al., 2013). These recruits may therefore come from the same farm or from farms in the neighbourhood. For neighbourhood farms we scale the relative contribution according to seaway distance to the given neighbourhood farm using a function with parameters estimated from the data. All biological processes in the two sub-models are subject to temperature dependency. Furthermore, farms are subject to pest control in the form of anti-parasitic treatments with instant effects. The susceptibility of the fish in a farm to lice infection depends on fish size, the numbers of fish on the farms and other covariates. The sub-models for adult females and other mobiles are fitted to farm data on a weekly resolution covering all marine farms producing salmonids along the Norwegian coast over the years 2012 - 2016.

In this paper, we first present the data on reported louse abundances and farmed fish populations. We go on to present the sub-models for the two stage-groups of lice. We continue by defining how we account for correlations in the time-series data of the two stage-groups, in order to calculate correct uncertainty in forecasts from the model. In the results and discussion sections, we present parameter-estimates for the two sub-models and discuss important aspects with regard to the population dynamics of salmon lice on farmed salmonids. Among such aspects, we explore i) the contribution of various time-lags of adult female lice to the recruitment into the stage-group of other mobiles; ii) how various covariates affect

the susceptibility of the fish on a farm; iii) and how host density dependence affects the expected abundances of the two stage-groups one week ahead under standardised conditions. Finally, we illustrate how the model performs when used for predicting louse abundances one to eight weeks ahead.

2 Materials and methods

The present study-system consists of marine farms in coastal waters of Norway nurturing salmonids in open water net-pens. The fish are typically stocked as smolts produced in freshwater farms and on-grown for roughly 18 months, until slaughtering for food consumption. During the marine on-growing phase, the fish are exposed to infectious salmon louse copepodites that are spread about as plankton in the water current. To approximate this spread we use a distance based relative risk function estimated from the data. This does not account for water currents, but can be regarded to represent the normalised spread of the planktonic salmon lice between farms in the present data set, regardless of the environmental conditions that force water currents. Once attached to a fish, the copepodites develop through two larval stages attached to the fish, two pre-adult mobile stages and finally adult males and females who reproduce the next generation planktonic stages (Hamre et al., 2013). More detailed descriptions of salmon farming and the salmon louse parasite appear in previous papers (Jansen et al., 2012; Aldrin et al., 2013).

2.1 Data

The data used for the present study consists of i) farm locations, collated from an open source Aquaculture register hosted by the Norwegian Directorate for Fisheries (Anonymous, 2017a). From these data we have calculated pairwise seaway distances between all farms, but truncated at 100 km (Jansen et al., 2012). Farm populations of fish ii), i.e. total numbers of fish, fish size and species (Atlantic salmon, *Salmo salar*, or rainbow trout, *Onchorhyncus mykiss*). These data are reported on a monthly frequency, but are not publicly available due to stock marked sensitivity. The farm population data are available to The Norwegian Veterinary Institute for surveillance and research purposes, through the responsible Norwegian Food Safety Authority (Jansen et al., 2012). Salmon louse related data iii), i.e. abundances of adult females and other mobiles of salmon lice reported from lice counts, anti-parasitic treatments and water temperatures. Counting of lice and reporting is done on a weekly frequency and these data are now publicly available (Anonymous, 2017b).

The present data include all 991 Norwegian marine fish farms with standing stocks of either Atlantic salmon or rainbow trout (salmonids) in any week from week 1 in 2012 to week 44 in 2016. The data from the first sixteen weeks were only used to construct lagged explanatory variables, whereas the model was fitted to data from the 17th week in 2012 and onward. In the following description, we therefore only summarise data from the latter period. The number of active farms was then reduced to 979. Each salmonid farm normally had several consecutive periods of fish production, interrupted by periods of fallowing (i.e. no fish on the farm). Each farm could produce either Atlantic salmon or rainbow trout, or both species at the same time. The fish population within a production period

is termed a cohort and the present data consists of 2 598 cohorts with a total of 137 595 farm-weeks of fish production. Farms produced between one and seven consecutive cohorts. For each salmonid farm, it is obligatory to report the salmon lice abundance on a sample of fish to the responsible Norwegian Food Safety Authority. We assumed that the number of sampled fish was always 40, corresponding to ten fish sampled from each of four cages, which is the minimum required sample size for a farm of normal size with eight cages (Jansen et al., 2012). The reported salmon lice abundances were therefore multiplied by 40 and implemented as the response variable in the present model. In reality, the number of sampled fish was probably often larger than 40, but this did not necessarily affect the mean number of lice per fish. However, lice abundance data were missing for about 16% of the total number of weeks with farmed fish, giving 115 132 farm-weeks with reported salmon lice abundance. The number of missing lice counts is particularly high during the first weeks after stocking, since the fish should be disturbed as little as possible in this period and because the lice abundance is expected to be low. Ignoring the first four weeks of each cohort, the proportion of missing lice data was reduced to 13%.

Lice abundances were recorded for both for adult females (A) and other mobile stages of lice (O). The A category include adult females with and without egg strings attached. The O category include adult males and pre-adult stages of both males and females. The abundance of adult females was between 0 and 1.2 in 95% of the farm-weeks, with a mean of 0.18 lice per fish (Table 1). The distribution of adult female abundance was profoundly right-skewed, being exactly 0 in 41% of the farm-weeks, whereas the highest reported abundance was 29 lice per fish. The abundance of other mobiles was higher than for adult females, with a mean of 0.82 lice per fish, and being 0 in 17% of the farm-weeks. Summary statistics

for lice abundance and several other quantities are found in Table 1.

Table 1: Summary statistics for various quantities over all weeks (over all farms for No. neighbouring farms) with observed values of the quantity in question.

Quantity	Unit	Mean	2.5% percentile	97.5% percentile	Minimum	Maximum
A abundance in sample	lice per fish	0.18	0.00	1.23	0.00	29.5
O abundance in sample	lice per fish	0.82	0.00	5.15	0.00	86.6
No. fish	millions	0.700	0.028	1.81	0.00021	5.10
Average weight of fish	kg	2.4	0.117	6.6	0.029	15.0
Proportion of salmon*	%	93	0	100	0	100
Seawater temperature	°C	9.3	3.4	16.2	0.1	22.6
No. neighbouring farms within 10 km		4.3	0	6	0	18

A: Adult females
O: Other mobiles
*: Farmed salmonids being Atlantic Salmon

The black curves in Figure 1 show how the observed lice abundances vary over the data period averaged over all farms and for two selected farms. The lice counts have a clear seasonal pattern driven by the seawater temperature.

Figure 2 shows the locations of all Norwegian salmonid farms that actively produced salmonids any week during the study period, with a closer look at farms that were active in autumn 2013 (week 44) in an area in the South-West of Norway.

The seawater temperature was missing for 7.7% of the farm-weeks. Each missing temperature was imputed by a weighted mean of all observed temperatures the same week, with weights proportional to the inverse of the seaway distance to the current farm with the missing temperature. For 1.4% of the farm-weeks, there were no fish at the same farm in the previous week, and for 3/4 of these weeks, the mean fish weight was less than 250 g. This indicates that the fish cohort was stocked for the first time on the given farm, whereas a start weight

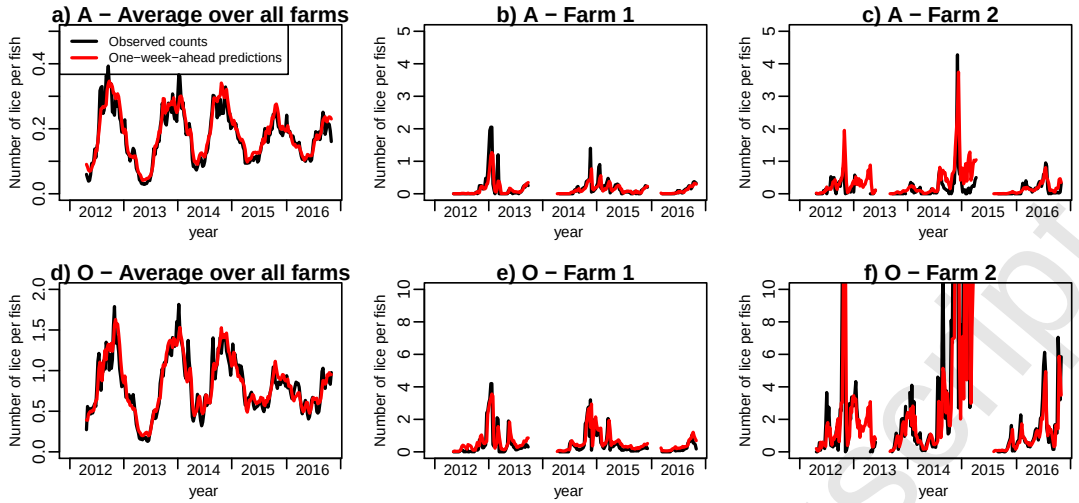

Figure 1: Time plots of observed counts (black) and one-week-ahead predictions (red) of the abundance of adult females (upper panels) and other mobiles (lower panels) in average over all farms (panels a and d), and for two separate farms (panels b, c, e and f).

of 250 g or larger indicates that the fish cohort had been relocated, i.e. moved from another marine farm. Medical salmon lice treatments were applied in 4.3% of the farm-weeks.

In addition, the use of non-medical treatments and novel methods to shield farmed fish from infection has increased towards the end of the data-period. Shielding methods include the use of plankton nets surrounding the net pens, whereas non-medical treatments include mechanical removement of lice, or the use of warm- or fresh-water. The present model does not account for this since we lack sufficient data. For the same reason, we do not account for the use of cleaner fish to control lice, although this has been shown to have an effect in a more detailed data set (Aldrin et al., 2017). Finally, we have also ignored effects of salinity since we lack sufficient data, although salinity is known to affect the infection process of salmon lice (Bricknell et al., 2006).

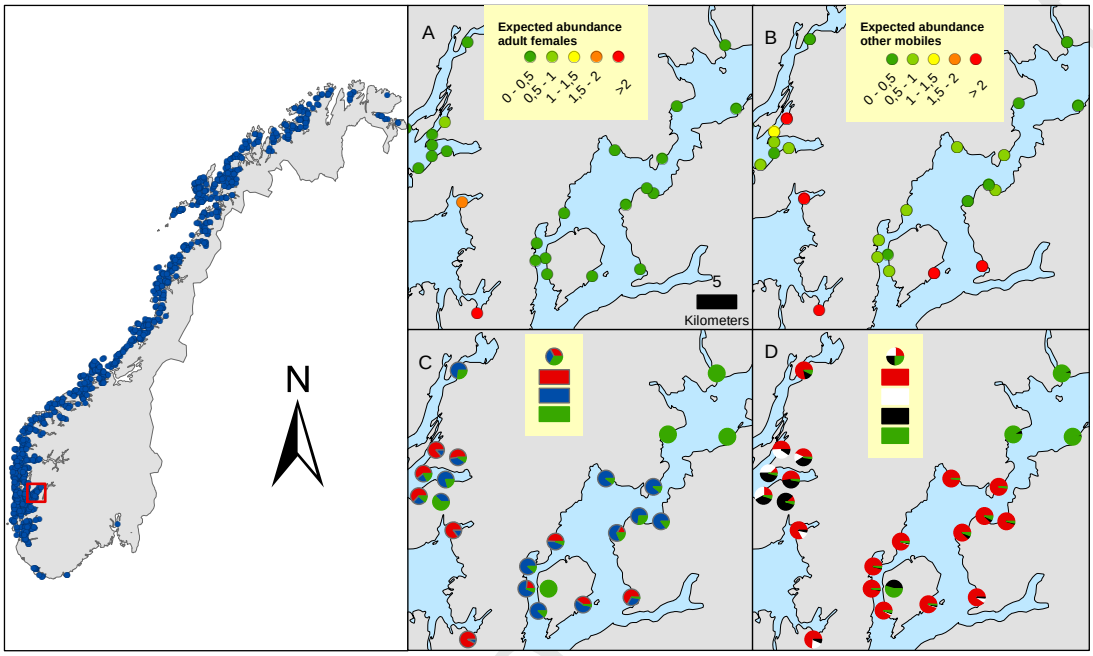

Figure 2: The left panel shows Norway with all farms having reported lice at some time during the data-period 2012-2016 (blue circles). The red square shows the area enlarged in panels A-D. The top panels show the model's expected abundances of adult female lice (A) and other mobiles (B) in farms for week 44 in 2013. The lower panels show pies representing the proportional contribution of various additive sources of infection comprising the model expected abundances in the top panels. For panel C, red colour represents surviving adult female lice; blue represents adult females developed from other mobiles; and green represents non-observed adult females. For panel D, red colour represents surviving other mobiles; white represents recruitment of other mobiles from the prior generation adult females within the same farm; black represents recruitment of other mobiles from the prior generation adult females in neighbouring farms; and green represents non-observed other mobiles (note that some farms have been slightly moved in the lower panels to avoid overlapping symbols).

2.2 Model

2.2.1 Model structure common for both adult females and other mobiles

The present salmon lice model consists of two sub-models, one for adult females and one for other mobiles. They share some of the same model-structure, but have separate parameter estimates. Let y_{it} denote the counted abundance of lice in a given stage at farm i in week t, that is, either $y_{it} = y_{it}^A$ for adult females or $y_{it} = y_{it}^O$ for other mobiles. This is the average number of lice in the respective stage counted on n fish, where $n = 40$ (as discussed in the previous section). We first present the model structure that is common for adult females and other mobiles, and then present the stage-specific parts of the models. Let further μ_{it} denote the expected abundance, i.e. $\mu_{it} = E(y_{it})$. This expectation is modelled as a function of i) observed lice abundances in previous weeks at the current farm, ii) observed lice abundances in previous weeks at neighbouring farms, and iii) other factors such as seaway distance.

We define $n \cdot y_{it}$, the total number of lice (in either stage A or O) counted on n fish, as our response variable, and model it as a zero-inflated, negative binomially distributed variable (Zuur et al., 2009) with expectation $n \cdot \mu_{it}$. Let p_{it}^z denote the probability of excess zero observations in this compound distribution. Thus, y_{it} comes from a negative binomial distribution with probability $1 - p_{it}^z$, and has an excess zero value with probability p_{it}^z. Let further μ_{it}^{NB} denote the expectation in the negative binomial part of the distribution. Then, the expectation of y_{it} can

be expressed as

$$\mathrm{E}(n \cdot y_{it}) = n \cdot \mu_{it} = (1 - p^z_{it}) \cdot \mu^{NB}_{it}. \quad (1)$$

Our main focus is on modelling the expected salmon lice abundance μ_{it}, and we give a detailed description of the two sub-models for μ_{it}, one for adult females and another for other mobiles, in the next subsections. The probability of excess zero observations, p^z_{it}, is also modelled as a function of μ_{it}. From Eq. (1), the expectation in the negative binomial part of the compound distribution is given as $\mu^{NB}_{it} = n \cdot \mu_{it}/(1 - p^z_{it})$. The variance of the negative binomial distribution can be written as $\mu^{NB}_{it} + (\mu^{NB}_{it})^2/\nu_{it}$, where $\nu_{it} > 0$ controls the over-dispersion and is here modelled as a function of μ_{it}. The parameter ν_{it} is sometimes called a heterogeneity or aggregation parameter and has often been denoted k in parasitological literature (Grenfell et al., 1995; Irvine et al., 2000). We present the models for p^z_{it} and ν_{it} in the Supplementary material, since these only affect the shape of the distribution of y_{it}, which is not our focus here.

The motivations for using a zero-inflated, negative binomial distribution are that i) it allows for over-dispersion due to large variability from fish to fish, and ii) that previous analyses of similar data showed that there was an excess frequency of zeroes (Jansen et al., 2012; Aldrin et al., 2013) compared to a negative binomial distribution, and including zero-inflation therefore gave a better fit to the data in those studies as well as in the present one.

2.2.2 The expected lice abundance for adult females

The model for the expected abundance of adult females at farm i in week t has the following additive-multiplicative form:

$$\mu_{it}^{A} = S_{it} \cdot \kappa_{it}^{susc} \cdot \kappa_{it}^{treat} \cdot (\lambda_{it}^{As} + \lambda_{it}^{Od} + \lambda_{it}^{int}), \tag{2}$$

where all terms are non-negative.

The three multiplicative terms in Eq. (2) are:

- S_{it} is an "at-risk" indicator that is 1 when farm i is active (has a positive number of farmed fish) at week t and 0 otherwise.
- κ_{it}^{susc} is a factor proportional to the susceptibility of farm i. It depends on explanatory variables that characterise the conditions for the fish at farm i at week t, for instance seawater temperature. It has the form

$$\kappa_{it}^{susc} = \kappa^{susc\bullet} \cdot \exp\left(\sum_{k} \beta_{k}^{susc} x_{ikt}^{susc}\right), \tag{3}$$

where x_{ikt}^{susc} denotes the k-th explanatory variable for farm i at week t and β_{k}^{susc} denotes the corresponding regression coefficient. The explanatory variables are centred at given reference values close to the average of each variable, such that the parameter $\kappa^{susc\bullet}$ is the value of κ_{it}^{susc} when all the explanatory variables are equal to their reference values. Specifically, this means when the time is week 30 in 2014, the seawater temperature is 9°C, the farm is located at latitude 64° north, the fish weight is 2 kg and the fish are neither stocked or relocated in week t, see also Table 2. We will also later use the symbol $^\bullet$ in the superscript for parameters that represent

a "standard" value of a quantity when modifying explanatory variables are at given reference values.

- κ_{it}^{treat} is a factor that accounts for mortality due to medical treatments, which is modelled as

$$\kappa_{it}^{treat} = 1 - \theta^{tm} \cdot I_{i(t-1)}^{t}, \tag{4}$$

where θ^{tm} is a parameter representing the treatment mortality and $I_{i(t-1)}^{t}$ is an indicator variable that is 1 if a treatment has been applied in week $t-1$ at farm i and 0 otherwise. With this formulation, the treatment effect is assumed to occur the week after the treatment was carried out. This is supported by the data, but see the Supplementary material for a more flexible formulation. Since the mortality parameter θ^{tm} must be between 0 and 1, another unrestricted parameter θ^{tm*} is modelled on the logit scale and θ^{tm} is given by $\theta^{tm} = \exp(\theta^{tm*})/(1 + \exp(\theta^{tm*})$.

The three additive terms represent three possible sources of lice infection:

1. λ_{it}^{As} is based on counts of adult females in the two previous weeks within the farm of interest, and represents *adult females* that have *survived* from previous weeks, see details below.

2. λ_{it}^{Od} is based on counts of other mobiles in the two previous weeks within the farm of interest, and represents *other mobiles* that have *developed* into adult females, see details below.

3. λ_{it}^{int} is an *intercept* that, when multiplied by $S_{it} \cdot \kappa_{it}^{susc} \cdot \kappa_{it}^{treat}$, represents the expected abundance of adult females when the counted numbers of both

adult females and other mobiles previous weeks are zero. This term must be included because lice may be present in previous weeks even if the counts of adult females and other mobiles are zero.

Details on how each of these terms are modelled are given in the following:

Surviving adult females are modelled as

$$\lambda_{it}^{As} = \varphi_1^{As} \cdot y_{i(t-1)}^{A*} + \varphi_{2it}^{As} \cdot y_{i(t-2)}^{A*} \cdot S_{i(t-1)}, \tag{5}$$

where:

- $y_{i(t-l)}^{A*} = \left(y_{i(t-l)}^{A}\right)^{\alpha}$, where α is a positive parameter that allows for a non-linear dependency of the previous weeks' lice counts.
- $\varphi_1^{As} = 1$ is an autoregressive coefficient set to 1 for identifiability, since this coefficient is statistically confounded with the factor $\kappa^{susc\bullet}$ in (3).
- φ_{2it}^{As} is another positive autoregressive coefficient that depends on the seawater temperature, and therefore varies by time, through the following equation

$$\varphi_{2it}^{As} = \varphi_2^{As\bullet} \cdot \exp(\beta_2^{As} \cdot (T_{it} - 9)), \tag{6}$$

where $\varphi_2^{As\bullet}$ and β_2^{As} are parameters and $T_{it} - 9$ is the seawater temperature centred around 9°C.

- $S_{i(t-1)}$ has the same definition as before, and is included to ensure that lice from time $t-2$ cannot survive until time t if the farm is inactive the week in between.

The term representing *other mobiles developed* into adult females is modelled as

$$\lambda_{it}^{Od} = \varphi_{1it}^{Od} \cdot y_{i(t-1)}^{O*} + \varphi_{2it}^{Od} \cdot y_{i(t-2)}^{O*} \cdot S_{i(t-1)}, \tag{7}$$

where:

- $y_{i(t-l)}^{O*} = \left(y_{i(t-l)}^{O}\right)^{\alpha}$.
- φ_{1it}^{Od} is an autoregressive coefficient, depending on the seawater temperature in the same way as φ_{2it}^{As} in Eq. (6), but with parameters $\varphi_1^{Od\bullet}$ and β_1^{Od}, where $\varphi_1^{Od\bullet}$ is the autoregressive coefficient when the seawater temperature is 9°C.
- φ_{2it}^{Od} is another autoregressive coefficient, also depending on the seawater temperature in the same way as φ_{2it}^{As} in Eq. (6), but with parameters $\varphi_2^{Od\bullet}$ and β_2^{Od}.

The *intercept*, λ_{it}^{int}, representing the scaled expected abundance when the counted numbers of lice previous weeks are zero, is modelled as a function of the seawater temperature in the same way as φ_{1it}^{Od} in Eq. (6), but with parameters $\lambda^{int\bullet}$ and β^{int}.

The coefficients φ_{2it}^{As}, φ_{1it}^{Od}, φ_{2it}^{Od} and λ_{it}^{int}, which depend on seawater temperature, can be interpreted relative to the fixed value of $\varphi_1^{As} = 1$. However, because all of these are multiplied by κ_{it}^{susc}, which also depends on seawater temperature, the net temperature dependency is more complex.

2.2.3 The expected lice abundance for other mobiles

The model for the expected abundance of other mobiles at farm i in week t has the following additive-multiplicative form:

$$\mu_{it}^{O} = S_{it} \cdot \kappa_{it}^{susc} \cdot \kappa_{it}^{treat} \cdot (\lambda_{it}^{Os} + \lambda_{it}^{Adw} + \lambda_{it}^{Adn} + \lambda_{it}^{int}), \tag{8}$$

where again all terms are non-negative. Below, we only explain terms that are different from the model for adult females. Other terms, with the same notation as in the model for adult females, have the same structure and interpretation as in that model, but different parameter values. The terms that differ from the model for adult females are:

1. λ_{it}^{Os} based on counts of other mobiles in the two previous weeks within the farm of interest, and represents *other mobiles* that have *survived* from previous weeks, see details below.

2. λ_{it}^{Adw} is based on counts of adult females in the previous sixteen weeks within the farm of interest, and represents offspring of previous *adult females within* the farm that have *developed* into other mobiles, typically several weeks later, see details below.

3. λ_{it}^{Adn} is based on counts of adult females in the previous sixteen weeks at neighbouring farms, and represents offspring of previous *adult females* at *neighbouring* farms, see details below.

Details on how each of these terms are modelled are given in the following:

Surviving other mobiles are modelled as

$$\lambda_{it}^{Os} = \varphi_1^{Os} \cdot y_{i(t-1)}^{O*} + \varphi_{2it}^{Os} \cdot y_{i(t-2)}^{O*} \cdot S_{i(t-1)}, \tag{9}$$

where:

- $\varphi_1^{Os} = 1$ is an autoregressive coefficient set to 1 for identifiability.
- φ_{2it}^{Os} is another autoregressive coefficient that depends on the seawater temperature in the same way as φ_{2it}^{As} in Eq. (6), but with parameters $\varphi_2^{Os\bullet}$ and β_2^{Os}.

The term representing *adult females within* the farm, whose offspring have *developed* into other mobiles, is modelled as

$$\lambda_{it}^{Adw} = \sum_{l=1}^{l=16} \varphi_{lit}^{Ad} \cdot y_{i(t-l)}^{A*} \cdot n_{i(t-l)}^{\theta^{rep}} / n_{it}^{\theta^{inf}}, \tag{10}$$

where

- $\varphi_{lit}^{Ad}, l = 1, \ldots 16$ are coefficients that vary smoothly by increasing lags, through the dependency of only three underlying coefficients in the equation

$$\varphi_{lit}^{Ad} = \gamma_{it}^{a} \cdot [1 - \gamma_{it}^{s} \cdot (l - \gamma_{it}^{c})^2]_+ / c_{it}, l = 1, \ldots, 16. \tag{11}$$

  Here, $[\cdot]_+$ means the non-negative part, γ_{it}^{c} is the centre in a quadratic curve, $\gamma_{it}^{s} > 0$ controls the spread of the curve, γ_{it}^{a} is its amplitude and c_{it} is a normalising constant such that the coefficients sum to γ_{it}^{a}, i.e. $\sum_{l=1}^{l=16} \varphi_{lit}^{Ad} = \gamma_{it}^{a}$. See Figure 4 in Section 3.1 for examples on how the lag structure may look. The coefficients γ_{it}^{s} and γ_{it}^{a} depend on the seawater temperature in

the same way as φ_{1t}^{Od} in Eq. (6), but with parameters $\gamma^{s\bullet}$, β^s and $\gamma^{a\bullet}$,2 β^a, respectively. Furthermore, γ_{it}^c depends on the seawater temperature as

$$\gamma_{it}^c = 1 - 1/\sqrt{\gamma_{it}^s} + \gamma^{c\bullet} \cdot \exp(\beta^c \cdot (T_{it} - 9)), \tag{12}$$

where $\gamma^{c\bullet}$ and β^c are parameters. This functional form ensures that the minimum development time from eggs to other mobiles increases by decreasing seawater temperatures, see details in the Supplementary Material.

- n_{it} is the number of fish at farm i at time t, measured in million fish, and $n_{i(t-l)}$ is the number of fish l weeks before.

- θ^{rep} and θ^{inf} are non-negative parameters that allow for non-linear dependencies of the numbers of fish at previous lags, i.e. the hosts for reproducing lice at previous lags, and the number of fish that becomes infected by the new lice. Furthermore, θ^{rep} is re-parameterised as

$$\theta^{rep} = \theta^{inf} + \theta^0. \tag{13}$$

If the parameter θ^0 is positive, the infection pressure per fish will increase if the overall number fish increases.

The term representing *adult females* at *neighbouring* farms, whose offspring have *developed* into other mobiles, is modelled as

$$\lambda_{it}^{Adn} = \left(\sum_{j \neq i} \lambda_{ijt}^{Adn} \right)^{\delta}, \tag{14}$$

where λ_{ijt}^{Adn} is the contribution from the j-th farm given by

$$\lambda_{ijt}^{Adn} = \left(\sum_{l=1}^{l=16} \varphi_{ljt}^{Ad} \cdot y_{j(t-l)}^{A*} \cdot n_{j(t-l)}^{\theta^{rep}} / n_{it}^{\theta^{inf}} \right) \cdot \pi_0 \cdot \exp\left(\pi_1 \cdot ((d_{ij}^{\pi_2} - 1)/\pi_2)\right). \tag{15}$$

Here,

- δ is a positive parameter that allows for non-linear dependency of the sum of the contributions from the neighbouring farms. The values of δ and θ^{rep} quantifies the consequences of introducing new neighbouring farms versus increasing the number of fish at the existing neighbouring farms, see the discussion in Section 3.1.

- $n_{j(t-l)}$ is the number of fish at a neighbouring farm j at time $t-l$, i.e. the hosts for reproducing lice.

- π_0 quantifies the importance of neighbouring infection compared to the other sources of infection.

- d_{ij} is the seaway distance between farms i and j.

- π_1 and π_2 are parameters that reflect the effect of the seaway distances to neighbouring farms, and the transformation $(d_{ij}^{\pi_2} - 1)/\pi_2$ is the Box-Cox transformation, which allows for different shapes of the distance function.

2.2.4 Correlations

So far, we have treated the lice counts of adult females (y_{it}^{A}) and other mobiles (y_{it}^{O}) at each farm as separate time series. The expected values of one series is modelled as a function of previous values of the other series at the same farm,

and for other mobiles also on the series at neighbouring farms. However, we have not considered potential mutually stochastic dependency within the same week between the various time series of lice counts. It turns out that some of the series are positively correlated (Section 3). It is essential to include such correlations to calculate the correct uncertainty for forecasts more than one step ahead for single time series and even for one step ahead for the average of several time series (Section 2.2.5).

Thus, we need to define a "multivariate zero-inflated negative binomial distribution" (several choices are possible, but none of them are the single obvious one), and we need to select a suitable model structure. Our approach is to start with a multivariate standard normal distribution with a given correlation structure, consisting of q correlated univariate standard normal variables $z_j, j = 1, \dots q$ (one for each farm). Then each univariate standard normal variable z (dropping the farm index for a while) is transformed to a zero-inflated negative binomial variable with a given set of parameters μ, p^z and ν (who depend on the farm and the time point). This procedure induces a correlation structure at the zero-inflated negative binomial scale. The transformation can be divided into two steps: First, z is transformed to a uniform variable u by $u = \Phi(z)$, where Φ is the cumulative distribution function for a standard normal variable. In the next step, u is transformed to a zero-inflated negative binomial variable y by $y = F^{-1}_{\mu,p^z,\nu}(u)$, where $F^{-1}_{\mu,p^z,\nu}$ is the inverse cumulative distribution function for a zero-inflated negative binomial variable with the given set of parameters. These two steps can be condensed to $y = F^{-1}_{\mu,p^z,\nu}(u) = F^{-1}_{\mu,p^z,\nu}(\Phi(z))$.

We consider four types of pairwise correlations; i) between adult females and other mobiles within farms, and ii) between adult females, iii) between other

mobiles and iv) between adult females and other mobiles at different farms. Let $\rho_{ij}^{s,s'}$ denote the correlation at the standard normal scale between transformed lice counts at stages s and s' at farms i and j. We expect that correlations between farms i and j depend on the their corresponding seaway distances d_{ij}, but assume that between farm correlations are exact 0 for seaway distances larger than 100 km. Our correlation model at the standard normal scale are then

$$\rho_{ii}^{wAO} = \rho^{wAO}, \tag{16}$$

$$\rho_{ij}^{bAA} = \rho^{bAA\bullet} \cdot \exp(\beta_1^{bAA} d_{ij}) \text{ for } i \neq j \text{ and } d_{ij} < 100 \text{ km}, \tag{17}$$

$$\rho_{ij}^{bOO} = \rho^{bOO\bullet} \cdot \exp(\beta_1^{bOO} d_{ij}) \text{ for } i \neq j \text{ and } d_{ij} < 100 \text{ km}, \tag{18}$$

$$\rho_{ij}^{bAO} = \rho^{bAO\bullet} \cdot \exp(\beta_1^{bAO} d_{ij}) \text{ for } i \neq j \text{ and } d_{ij} < 100 \text{ km}, \tag{19}$$

where the ρ's and β's are parameters to be estimated. The parameters $\rho^{bAA\bullet}$, $\rho^{bOO\bullet}$ and $\rho^{bAO\bullet}$ are interpreted as correlations between farms with seaway distances exactly 0, but in practice there are usually at least 1-2 km between farms.

For estimating these parameters, we follow the opposite procedure, i.e. we transform from the observed zero-inflated negative binomial variables to standard normal variables, and then estimate the correlation parameters. If an observation y was sampled from a continuous variable with cumulative distribution function $F(y)$, the transformation to the standard normal scale would be unique by $z = \Phi^{-1}(F(y))$. However, since y is discrete, the transformation of y to u on the 0-1 scale is only determined up to an interval, and all values of u within this interval would give the same y. We handle this by simulating a value u^{sim} uniformly within this interval, i.e. $u^{sim} \sim Uniform(p_1, p_2)$, where $p_1 = P(Y \leq y-1) = F_{\mu,p^z,\nu}(y-1)$ and $p_2 = P(Y \leq y) = F_{\mu,p^z,\nu}(y)$, and then $z = \Phi^{-1}(u^{sim})$. Here, $P(\cdot)$ means probability, Y is a stochastic variable and y

a possible observed value. Each observed lice count is transformed separately in this way.

After transforming all observations to the standard normal scale, we calculate pairwise empirical correlations $r_{ij}^{s,s'}$ between all variables (time series) that overlap in time. Then, the parameters are estimated by the least squares fits of the three variants of the non-linear regression equations $r_{ij}^{ss'} = \exp(\beta_0^{bss'} + \beta_1^{bss'} d_{ij}) + \varepsilon_{ij}$ where ε_{ij} is random noise, and of the simpler variant $r_{ii}^{wAO} = \exp(\beta_0^{wAO}) + \varepsilon_{ii}$ for the within farm correlations. This is a Gaussian model with a logarithmic link function and standard software for generalised linear models may be therefore be used. The estimates of ρ parameters of the right sides of Eqs. (16)-(19) are then given by back-transformations of the form $\hat{\rho}^{wAO} = \exp(\hat{\beta}_0^{wAO})$ and $\hat{\rho}^{bss'\bullet} = \exp(\hat{\beta}_0^{bss'})$.

Since each variable is transformed independently of the others, the simulation step eliminates some of the correlation structure, resulting in biased parameter estimates giving too low correlations. To adjust for this we apply bias correction by the following parametric bootstrap procedure:

1. Estimate the β parameters by the regression procedure above and denote the estimates by $\tilde{\beta}$.

2. Simulate a data set from the estimated model on the zero-inflated negative binomial scale with the same size as the observed data.

3. Transform the simulated observations to the standard normal scale and estimate the parameters, denoting them by $\tilde{\beta}^b$.

4. Repeat 1-3 for $b = 1, \ldots B = 20$, and compute the average of each parameter estimate as $\tilde{\beta}^* = 1/B \sum_b \tilde{\beta}^b$.

5. The final estimate is then given by $\hat{\beta} = 2\tilde{\beta} - \tilde{\beta}^*$. This standard bootstrap bias correction formula (Efron and Tibshirani, 1993, see e.g.) is based on the assumption that $(\tilde{\beta} - \beta)$ has approximately the same distribution as $(\tilde{\beta}^b - \tilde{\beta})$, where β is the true parameter value.

For investigating the correlations between the original lice counts y_{it} $(i = 1, 2, \ldots)$ in the same week t, with corresponding parameters μ_{it}, p^z_{it} and ν_{it}, we consider the Pearson residuals $(y_{it} - \mu_{it})/s_{it}$, where $s_{it} = \sqrt{n\mu_{it} + [(n^2\mu^2_{it})/(1 - p^z_{it})]\,(1/\nu_{it} + p^z_{it})}$ is the standard deviation of y_{it}, see the Supplementary Material. After substituting the parameter values by their estimates, we calculate empirical correlations between all pairs of residual series. We summarise the correlations by regressing them on the corresponding seaway distances with the same structure as Eqs. (16)-(19). We compare this empirical correlation structure with the corresponding structure in data simulated from the model, to ensure that our model is able to reconstruct the correlation structure in the observed data.

2.2.5 Forecasts

For forecasting ahead in time, we use the following procedure:

1. With observations up to and including time t, calculate μ_{t+1}, p^z_{t+1} and ν_{t+1} for all time series.

2. Simulate multivariate normal variables with correlations given from Eqs. (16)-(19).

3. Transform the standard normal variables to zero-inflated negative binomial variables, using the parameters calculated in 1.

4. Go one step ahead to time $t+1$ and let the simulated zero-inflated negative binomial variables from 3 take the role as real observations.

5. Repeat 1-4 until the maximum forecast horizon.

6. Repeat 1-5 $nsim = 200$ times

7. Take the average of all simulations as the point forecast and calculate the 95% prediction interval as the 2.5 and the 97.5 quantiles of the simulated values. This is done for adult females and other mobiles at each farm and for the sum over farms.

2.2.6 Estimation

The various models were estimated by fitting them to observed lice counts between week 17 in 2012 until week 44 in 2016, whereas data from the first sixteen weeks in 2012 were used to construct lagged variables. The models for adult females and other mobiles were estimated by maximizing the log likelihood (see Supplementary material for the likelihood expression) using the function *optim* in the statistical software R, using the method of Byrd et al. (1995). Parameter uncertainties were based on the observed Fisher information matrix (Pawitan, 2001) and are reported as 95% Wald confidence intervals. Some parameters were estimated on the logit or the logarithmic scale and transformed back to the original scale used in the model descriptions, giving non-symmetrical confidence intervals on the original scale. The correlation models were estimated using the *glm* function in R.

3 Results and discussion

3.1 Estimated models for expected lice abundance

The estimated parameters with 95% confidence intervals for the expected abundances are given in Tables 2 and 3. The estimates for the corresponding submodels for the excess zero probabilities p^z and the aggregation parameters ν are found in the Supplementary materials. Below, we first illustrate the results by figures and then comment on the parameters that are not covered by these figures.

Figure 1 shows time plots of the expected abundance of lice and the corresponding observed lice abundance averaged over all farms and for two selected farms. The expected abundance of adult females (panel A) and other mobiles (panel B), along with the proportional contribution of the various additive sources of infections (panels C,D), are shown for a small area in the South West of Norway in week 44 in 2013 in Figure 2. In panel C, the proportions are based on Eq. (2), and for instance the contributions from surviving adult females are calculated as $\lambda_{it}^{As}/(\lambda_{it}^{As}+\lambda_{it}^{Od}+\lambda_{it}^{int})$. In panel D, the proportions are based on Eq. (8), and are calculated similarly.

Figure 3 shows the estimated lag structure in the model for adult females at various seawater temperatures. When the temperature increases, the first order lag becomes increasingly more important than the second order lag, and the importance of the development from other mobiles to adult females increases compared to the surviving adult females. Both these effects are explained by the well-known fact that the rate of development from one stage to the next increases by increasing seawater temperature. When the seawater temperature is 4 °C, φ_{2it}^{Od} is larger than φ_{1it}^{Od} (upper right panel), which suggests that the model could be

improved by including a third lag. However, it turns out that the improvement is not significant, so we keep only two lags in the final model.

Figure 4 corresponds to Figure 3, but shows the lag structure for the other mobiles model. Again, the increasing development rates by increasing temperature explain the different patterns in the panels. The lag structure for the development from adult females to the next generation other mobiles is centred around 10 days at 4°C, 6 days at 9°C and 4 days at 14°C (Figure 4, right panels). These time delays are comparable with some previous studies. Stien et al. (2005) presented a population model for salmon lice with equations for *minimum* development times for each stage to the next, where the estimated parameters were based on a synthesis of laboratory experiments. Their model did not include the time it takes for an infective copepodid to infect a fish host, but if we assume that this takes 4 days, corresponding to the expected lifetime of the copepodids (1/daily mortality, Stien et al. (2005)), the sum of the minimum development times over all relevant stages from eggs to pre-adults, amount to 11.3, 5.3 and 3.2 weeks for seawater temperatures 4, 9 and 14°C, respectively. A more detailed population model estimated on full scale production data from 32 Norwegian salmon farms, was developed by Aldrin et al. (2017). This study presented estimates for the *median* development times for each stage, and the sum of these amount to 12.0, 5.7 and 4.3 weeks for seawater temperatures 4, 9 and 14°C, respectively. A third comparison can be made by combining results from two recent laboratory studies (Eichner et al., 2015; Samsing et al., 2016). If we still assume that it takes 4 days for a copepodid to infect a fish host and one day more before it develops to the chalimus stage, their results indicate that it takes 3.5-4.1 weeks to develop from eggs to pre-adults at 10°C.

Now, we continue by commenting on the effect of explanatory variables included in the factors κ_{it}^{susc} in Eqs. (2) and (8), i.e. time, seawater temperature, latitude and fish weight. The various effects are presented as relative effects compared to a reference value, i.e. as $\exp(\hat{\beta}x)/\exp(\hat{\beta}x^{ref})$ if we consider an explanatory variable x with reference value x^{ref} and estimated regression coefficient $\hat{\beta}$.

Everything else unchanged, the expected lice abundances decrease over time for both adult females and other mobiles (Figure 5). One plausible explanation for this is that lice control practices have changed during the data period. In 2012, all lice treatments were medical treatments. However, as medical treatments have shown reduced effects due to louse resistance (Aaen et al., 2015; Jansen et al., 2016), new methods to shield farms from infection and non-medical treatments are increasingly being applied. These are not accounted for in our model due to the lack of sufficient data.

Figure 6 shows the relative effects of seawater temperatures at two different latitudes. For both adult females and other mobiles, the relative effect increases with increasing temperatures for low and medium temperatures. For temperatures above 11-12°C, however, the relative effect tends to decrease again for other mobiles. Furthermore, the effect of temperature difference between weeks t and $t-1$ is similar in the two models, with slightly lower relative effect in periods with increasing seawater temperatures (spring and summer). In contrast, the latitude has opposite effects in the two models. For adult females the relative effect increases with latitude, whereas for other mobiles it decreases (remember that this is conditional on temperatures and everything else being equal).

The mean fish weight at susceptible farms is also an important explanatory variable. The relative effect increases with increasing weight of the fish (Figure 7).

For other mobiles, one plausible explanation may be that the increasing surface of the fish facilitates host finding and the infection process for the parasite, as has repeatedly been suggested previously (Jansen et al., 2012; Aldrin et al., 2013). However, this does not explain the even more pronounced effect of fish weight for adult females, since they recruit from other mobiles and not through an infection process. Demographic processes that could contribute to the observed effect of fish weight on adult female lice could either be fish size dependent survival or recruitment of the adult females. Although speculative, we suggest that adult female mortality may be high on small fish due to the extensive use of cleaner fish while the farmed salmon are relatively small. With regard to recruitment from other mobiles, we suspect that this may be under-estimated for large fish. Other mobile lice are comparably small of size and easy to oversee under standard counting procedures and under-reporting related to fish size has also been suggested earlier (Aldrin et al., 2013).

The relative effect of infection pressure from neighbouring farms decreases (only relevant for other mobiles) rapidly by increasing seaway distances to the neighbours (Figure 8). Furthermore, the infection pressure from a neighbouring farm infinitesimally close is lower than the internal infection pressure within a farm. Interestingly, the form of the function in Figure 8 resembles the relative lice abundances in Figures 3 and 4 in Salama et al. (2016), obtained from simulating the spread of salmon lice in Loch Linnhe using a hydrodynamic model and particle tracking.

We end this subsection by commenting on the remaining parameter estimates in Tables 2 and 3. The estimated treatment mortalities are almost equal in the two models, being 44% for adult females and 47% for other mobiles. This seems to be

very low. However, the estimated treatment effects include all the various medical treatments and it is not clear in the data whether a given treatment covers all or only parts of the farm cages. This implies that lice counts post treatment may in some cases be from non-treated cages leading to underestimated treatment effects. The estimated values of the α parameters, which allow for non-linear dependency of previous weeks lice counts, are also very similar (0.874 and 0.889). The estimated value of the intercept λ_{it}^{int} in Eq.(2) for adult females is small ($\lambda^{int\bullet}$ is small compared to 1) and decreases by increasing seawater temperatures ($\beta^{int} < 0$). The intercept is small also for other mobiles, but the estimate of β^{int} for this model is close to 0 and non-significant. The parameter θ^0 is estimated to be positive, which means that the infection pressure per fish will increase if the number fish increases in general, i.e. if the number of fish increases by the same factor for both a susceptible farm and its infective neighbours. The estimate of the exponent $\theta^{rep} = \theta^{inf} + \theta^0$ in Eq.(13) is $0.391 + 0.171 = 0.562$, and the estimate of the exponent δ in Eq.(15) is 0.709. This means that if the numbers of fish are doubled at all neighbouring farms, the infection pressure for the susceptible farm will increase by 32% ($(2^{0.562})^{0.709} = 1.32$), as long as the lice abundance at the neighbouring farms as well as the number of fish at the susceptible farm are held constant. On the other hand, if the number of neighbouring farms are doubled, but with unchanged seaway distances to the susceptible farm, the infection pressure will increase by 63% ($2^{0.709} = 1.63$). We have no explanation for the asymmetric result of increasing the number of fish at the neighbouring farms compared to increasing the number of neighbouring farms. However, symmetry may be achieved by setting $\theta^{rep} = 1$, i.e. $\theta^0 = 1 - \theta^{inf}$. For a model with this restriction, the infection pressure will increase by 62% both by doubling the number of fish at the neighbours and by doubling the number of

neighbours. However, this model gave a significantly (p value ≈ 0) worse fit to the data, decreasing the log likelihood by 107. Therefore, by the configuration of farms in the present data set, an analytic result suggests that comparably few farms with many fish and large between farms distances is effective in terms of minimising the output of salmon lice. One should, however, be careful to interpret this as a causual effect, since there are many potential confounding factors related to farm size, for instance how the farms are managed or in which regions they are located.

If a farm is active in week t, but not in the previous week, the fish cohort at the farm is considered as stocked if the mean fish weight is less than 250 g and relocated otherwise. This is modelled by including a corresponding indicator variable in the expression for κ_{it}^{susc} in Eq.(3). Remember that previous lice counts in this case by our definition is 0. If a fish cohort is stocked, its expected abundance of adult females is reduced by 86% (since $1 - \exp(-2.002) = 0.86$)) compared to a farm that was active in the previous week, but with no observed lice. The corresponding reduction for other mobiles is 76%. On the other hand, if a fish cohort is relocated, the comparable expected lice abundance the first week after relocation is a factor 5.2 ($\exp(1.653) = 5.2$) for adult females and 3.5 for other mobiles higher than in an active farm without observed lice the previous week. This is probably because relocated fish bring lice with them from their previous farm.

Concerning the zero-inflation part of the model, the probabilities for excess zero lice counts, compared to a pure negative exponential distribution, are neglible in situations with positive lice counts the previous week. However, if there either was zero lice counts the previous week or if the farm were not active last week,

659 the probabilities for excess zeroes are high and in the range 0.25-0.6 (Figure 1 in
660 the Supplementary material).

Table 2: Estimated parameters in the model for expected abundance μ_{it} of adult females with 95% confidence intervals for the selected model.

Parameter group	Variable name or parameter description	Parameter symbol	Est.	Lower	Upper	Trans.
λ_{it}^{As},λ_{it}^{Od}	Non-linear dependency of counts	α	0.874	0.866	0.882	
λ_{it}^{As}	AR-coefficient, lag 1	φ_1^{As}	1	Fixed		
λ_{it}^{As}	Value of φ_{2it}^{As} when $T_{it}=9$	$\varphi_2^{As\bullet}$	0.314	0.294	0.336	log
λ_{it}^{As}	Temperature dependency of φ_{2it}^{As}	β_2^{As}	-0.063	-0.081	-0.045	
λ_{it}^{Od}	Value of φ_{1it}^{Od} when $T_{it}=9$	$\varphi_1^{Od\bullet}$	0.107	0.100	0.114	log
λ_{it}^{Od}	Temperature dependency of φ_{1it}^{Od}	β_1^{Od}	0.064	0.049	0.079	
λ_{it}^{Od}	Intercept in φ_{2it}^{Od}	$\varphi_2^{Od\bullet}$	0.055	0.051	0.060	log
λ_{it}^{Od}	Temperature dependency of φ_{2it}^{Od}	β_2^{Od}	-0.080	-0.103	-0.058	
λ^{int}	Value of λ_{it}^{int} when $T_{it}=9$	$\lambda^{int\bullet}$	0.014	0.013	0.015	log
λ^{int}	Temperature dependency of λ_{it}^{int}	β^{int}	-0.057	-0.068	-0.046	
susceptibility κ_{it}^{susc}	Reference value of κ_{it}^{susc}	$\kappa^{susc\bullet}$	0.480	0.468	0.492	log
-"-	(time-week 30 2014)*	β_k^{susc}	-0.026	-0.031	-0.021	
-"-	(temp-9)*	-"	0.040	0.034	0.045	
-"-	$(\text{temp-9})^2$*	-"-	-0.0024	-0.0029	-0.0018	
-"-	$\text{temp}_t - \text{temp}_{t-1}$*	-"-	-0.028	-0.035	-0.022	
-"-	(latitude-64)*	-"-	0.023	0.021	0.025	
-"-	log(weight)-log(2)*	-"-	0.246	0.238	0.255	
-"-	1 if stocked in week t, 0 otherwise	-"-	-2.002	-6.258	2.253	
-"-	1 if relocated in week t, 0 otherwise	-"-	1.653	0.961	2.346	
treatment κ_{it}^{treat}	treatment mortality	θ^{tm}	0.439	0.424	0.455	logit

* Seawater temperature is measured in °C, latitude in °N, weight in kg and time in years
Est.: Estimate
Lower: Lower bound of 95% confidence interval
Upper: Upper bound of 95% confidence interval
Trans: Transformation used in optimisation

Table 3: Estimated parameters in the the model for expected abundance μ_{it} of other mobiles with 95% confidence intervals for the selected model.

Parameter group	Variable name or parameter description	Parameter symbol	Est.	Lower	Upper	Trans.
λ_{it}^{Adn}	Non-linear dependency of infection from neighbours	δ	0.709	0.676	0.741	
λ_{it}^{Os},λ_{it}^{Adw},λ_{it}^{Adn}	Non-linear dependency of counts	α	0.889	0.883	0.896	
λ_{it}^{Os}	AR-coefficient, lag 1	φ_1^{Os}	1	Fixed		
λ_{it}^{Os}	Value of φ_{2it}^{Os} when $T_{it}=9$	$\varphi_2^{Os\bullet}$	0.157	0.148	0.168	
λ_{it}^{Os}	Temperature dependency of φ_{2it}^{Os}	β_2^{Os}	-0.130	-0.151	-0.109	
λ_{it}^{Adw},λ_{it}^{Adn}	Amplitude of lag curve when $T_{it}=9$	$\gamma^{a\bullet}$	0.262	0.243	0.283	log
λ_{it}^{Adw},λ_{it}^{Adn}	Temperature dependency of amplitude	β^a	0.165	0.153	0.177	
λ_{it}^{Adw},λ_{it}^{Adn}	Constant in centre of lag curve when $T_{it}=9$	$\gamma^{c\bullet}$	7.445	7.322	7.570	log
λ_{it}^{Adw},λ_{it}^{Adn}	Temperature dependency of centre	β^c	-0.118	-0.122	-0.114	
λ_{it}^{Adw},λ_{it}^{Adn}	Spread of lag curve when $T_{it}=9$	$\gamma^{s\bullet}$	0.165	**		log
λ_{it}^{Adw},λ_{it}^{Adn}	Temperature dependency of spread	β^s	0.217	0.216	0.219	
λ_{it}^{Adw},λ_{it}^{Adn}	Non-linear dependency of no. fish	θ^{inf}	0.391	0.356	0.427	
λ_{it}^{Adw},λ_{it}^{Adn}	Non-linear dependency of no. fish	θ^0	0.171	0.129	0.213	
λ_{it}^{Adn}	Relative importance of neighbouring infection	π_0	0.503	0.426	0.595	log
λ_{it}^{Adn}	Parameter in Box-Cox transformation	π_1	-0.374	-0.313	-0.436	
λ_{it}^{Adn}	Parameter in Box-Cox transformation	π_2	0.692	0.620	0.764	
λ^{int}	Value of $\lambda^{int}it$ when $T_{it}=9$	$\lambda^{int\bullet}$	0.018	0.016	0.020	log
λ^{int}	Temperature dependency of λ_{it}^{int}	β^{int}	0.0045	-0.0083	0.0173	
susceptibility κ_{it}^{susc}	Reference value of κ_{it}^{susc}	$\kappa^{susc\bullet}$	0.908	0.896	0.920	
-"-	(time-week 30 2014)*	β_k^{susc}	-0.039	-0.043	-0.035	
-"-	(temp-9)*	-"-	0.033	0.029	0.037	
-"-	$(\text{temp-9})^2$*	-"-	-0.0078	-0.0083	-0.0073	
-"-	$\text{temp}_t - \text{temp}_{t-1}$*	-"-	-0.034	-0.040	-0.028	
-"-	(latitude-64)*	-"-	-0.014	-0.016	-0.013	
-"-	log(weight)-log(2)*	-"-	0.103	0.096	0.110	
-"-	1 if stocked in week t, 0 otherwise	-"-	-1.429	-2.129	-0.729	
-"-	1 if relocated in week t, 0 otherwise	-"-	1.259	0.815	1.703	
treatment κ_{it}^{treat}	treatment mortality	θ^{tm}	0.469	0.456	0.482	logit

* Seawater temperature is measured in °C, latitude in °N, weight in kg and time i year

** Confidence limits not available due to unknown standard error. The optimisation routine were not able to compute the standard errors for these coefficients, probably due to to high correlations which give numerical instabilities while computing the Hessian matrix.

Est.: Estimate

Lower: Lower bound of 95% confidence interval

Upper: Upper bound of 95% confidence interval

Trans: Transformation used in optimisation

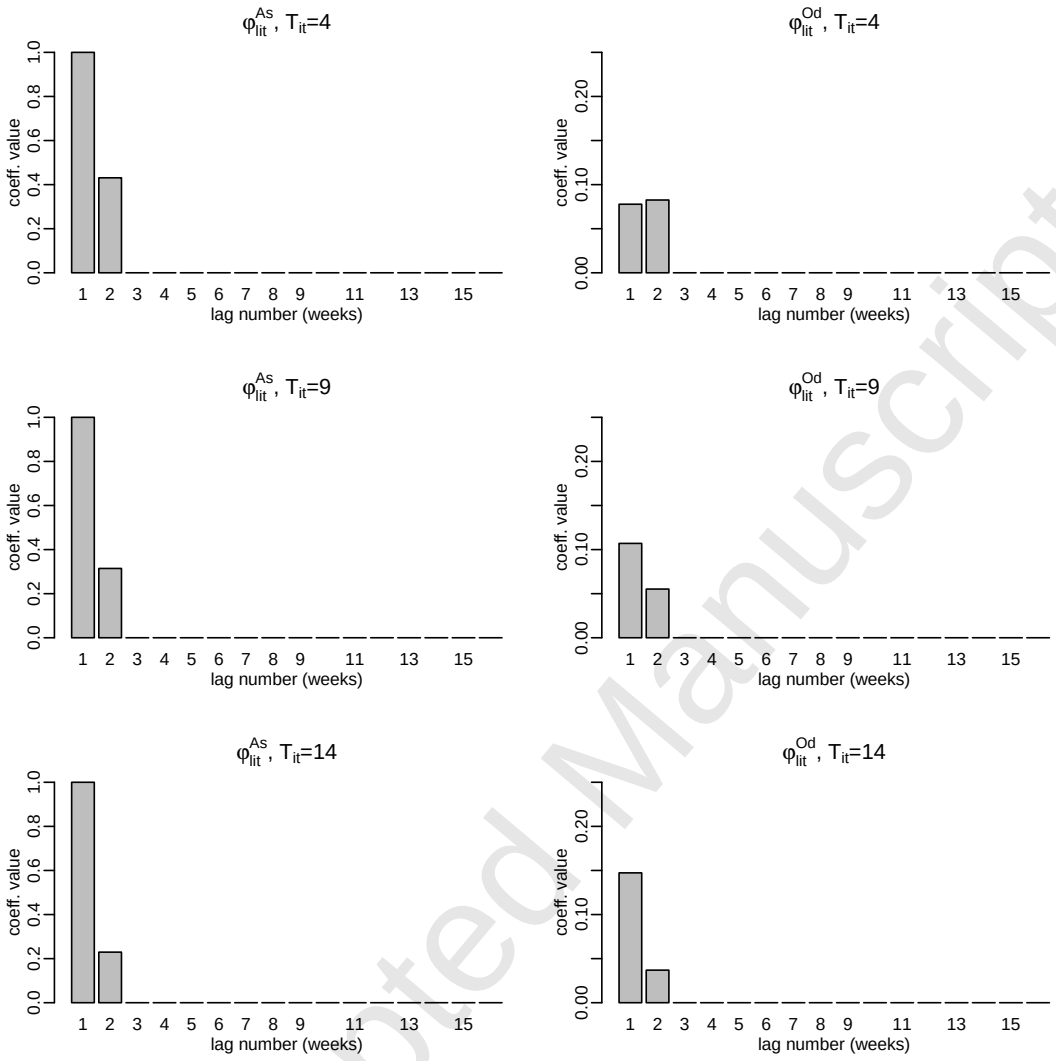

Figure 3: Autocorrelation coefficients in the A model, for λ_{it}^{As} (left panels) and λ_{it}^{Od} (right panels) and for seawater temperatures 4, 9 and 14°C (upper, mid and lower panels, respectively).

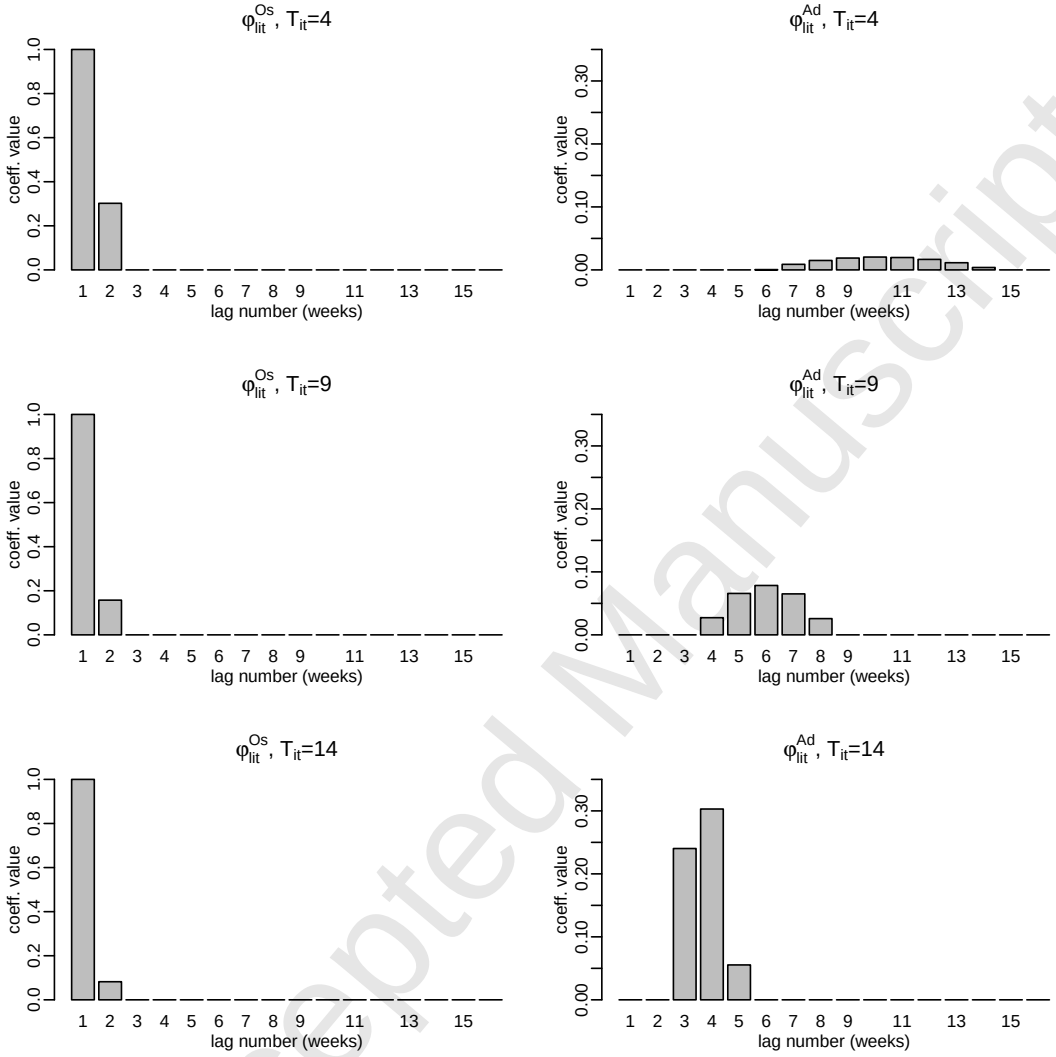

Figure 4: Autocorrelation coefficients in the O model, for λ_{it}^{Os} (left panels) and λ_{it}^{Adw} (right panels) and for seawater temperatures 4, 9 and 14°C (upper, mid and lower panels, respectively). The lag structure for infection from the previous generation adult females at a neighbouring farm (λ_{it}^{Adn}, not shown) is similar to the lag structure for infection from the previous generation adult females within the farm (λ_{it}^{Adw}), but it is first multiplied by a factor that depends on the seaway distance to the neighbour through $\pi_0 \cdot \exp\left(\pi_1 \cdot ((d_{ij}^{\pi_2} - 1)/\pi_2)\right)$, and then raised to the power of δ, see Eqs. (15) and (14).

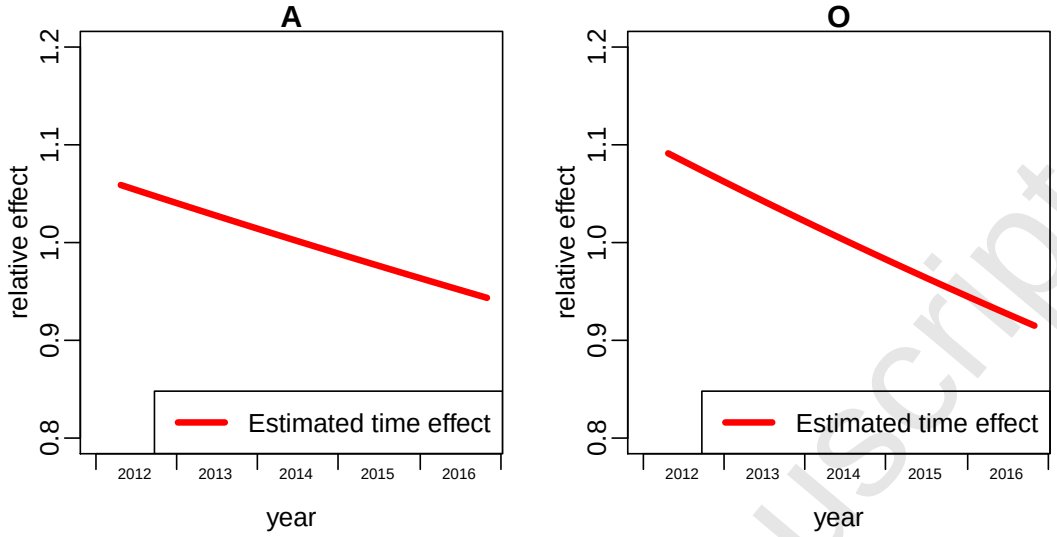

Figure 5: Relative effects of time the A and O models, with the effects set to 1 at week 30 2014.

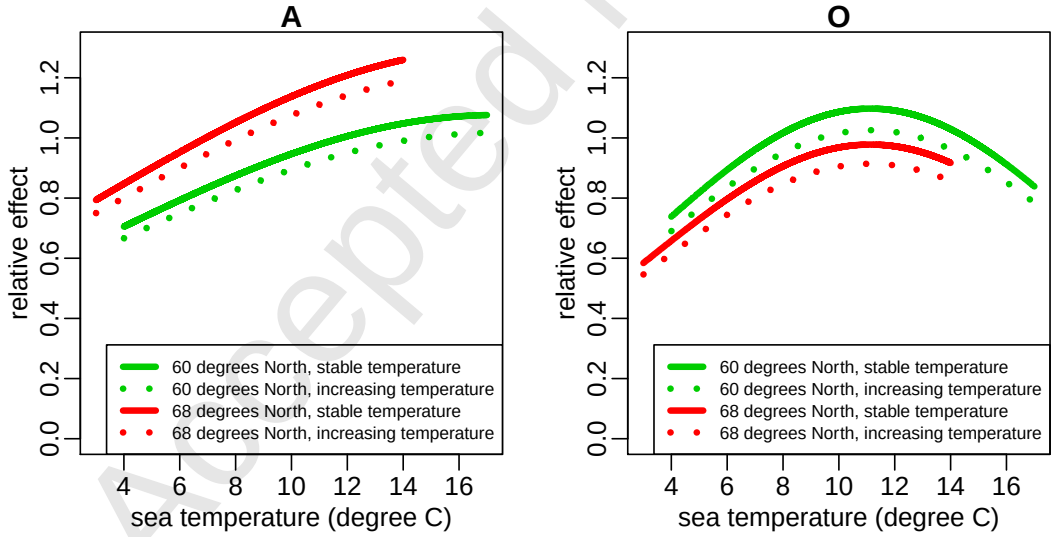

Figure 6: Relative effects of seawater temperatures and latitude in the A and O models, with the effects set to 1 for a constant seawater temperature 9°C at 64°N. The solid curves show the effects when the temperatures in the current and the previous week are equal, whereas the dotted curves show the effects when the temperature is increased by 2°C from the previous to the current week.

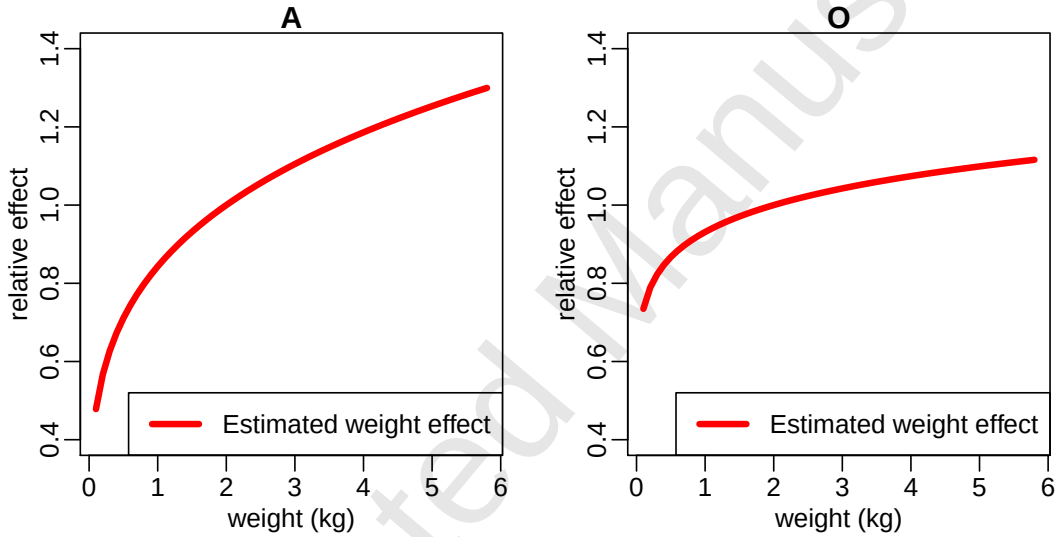

Figure 7: Relative effects of fish weight in the A and O models, with the effects set to 1 for fish at 2 kg.

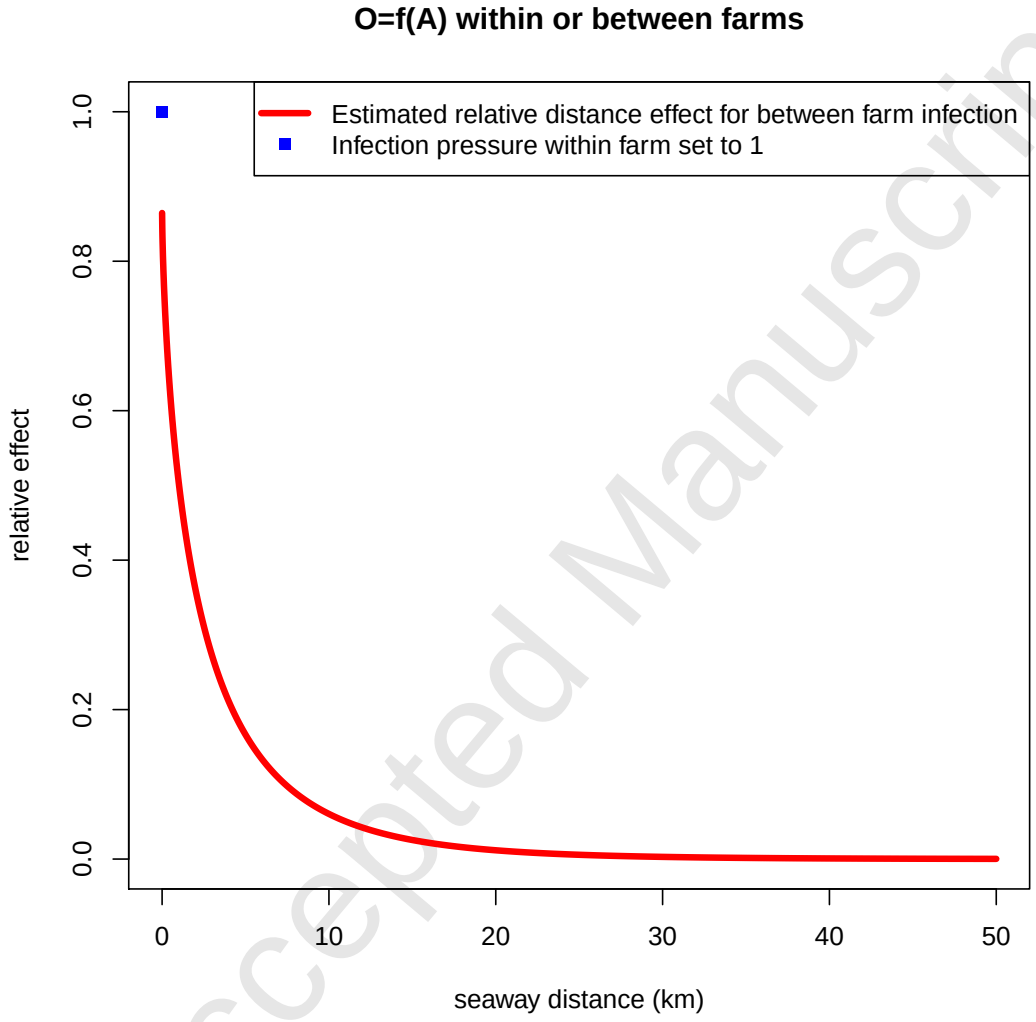

Figure 8: Relative effect of seaway distance for infection from neighbouring farms in the O model, where the within farm infection is the reference point set to 1. The relative effect is defined as the expression $\pi_0 \cdot \exp\left(\pi_1 \cdot ((d_{ij}^{\pi_2} - 1)/\pi_2)\right)$ from Eq. (15).

3.2 The correlation model

The estimated correlation parameters are given in Table 4. The estimated model parameters defined in Eqs. (17)-(19) are given in the leftmost columns. The β_1 parameters are negative, which means that the correlations decrease by increasing seaway distances between farms. The bias corrected parameters, which have been used in the final model, imply higher correlations than the ordinary ones, with higher values of ρ (correlations at zero distances) and lower or equal values of β_1. The induced correlations on the original scale are given in the rightmost columns, and we see that the induced correlations from the bias corrected model parameters match the observed correlations well.

Table 4: Estimated correlation parameters.

	Standard normal scale				Lice count scale					
	Ordinary		Bias corrected		Empirical		Ordinary		Bias corrected	
Correlation type	ρ	β_1	ρ	β_1	ρ	β_1	ρ	β_1	ρ	β_1
wAO	0.347		0.484		0.375		0.268		0.339	
bAA	0.069	-0.036	0.107	-0.034	0.107	-0.057	0.045	-0.036	0.072	-0.031
bOO	0.138	-0.024	0.158	-0.024	0.159	-0.028	0.118	-0.025	0.131	-0.027
bAO	0.045	-0.027	0.060	-0.027	0.044	-0.026	0.040	-0.034	0.040	-0.029

3.3 Predictions

To illustrate how the model can be used for prediction, we re-estimated the model without the data for the last eight weeks (weeks 37-44 in 2016), and made predictions 1-8 weeks ahead using lice counts up to and including week 36 in 2016. Figure 9 shows the predictions with 95% prediction limits in average over all farms and for the same two farms as was shown in Figure 1. Here, these predictions are conditioned on the observed seawater temperatures as well as the lice

treatments in the prediction period. In practice, one has to predict the seawater temperatures as well, but this should be quite easy, since the temperatures are rather predictable by season. Regarding lice treatments, one could compute predictions both with and without treatments, and use these predictions as a guide to whether one should treat or not during the next weeks to avoid lice abundance becoming too high.

The predictions for the average lice abundances (left panels in Figure 9) tend to be too high at the end of the prediction period, for both adult females and other mobiles. One reason for this is probably that the predictions are conditioned on the true lice treatments ahead in time. In real life, high abundances of lice induce treatments. In the simulations, however, the true lice treatments may occur when lice abundances by chance are rather low. Treatment on comparatively lower lice abundance will reduce the total effect of treatments in the simulations compared to the real data.

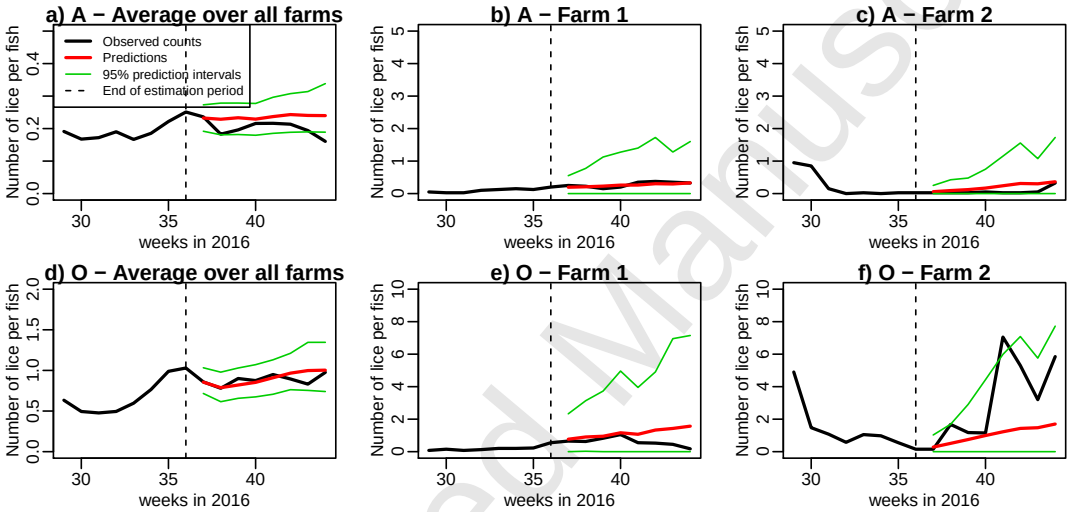

Figure 9: Observed (black) and predicted (red) lice abundance of adult females (upper panels) and other mobiles (lower panels) in average over all farms (panels a and d), and for two separate farms (panels b, c, e and f). The predictions for weeks 37-44 2016, with corresponding 95% prediction intervals (green), are based on lice counts up to week 36 2016.

3.4 Conclusions

The present paper presents a model for salmon lice abundance covering all marine salmon farms along the Norwegian coast. Fitting the model to weekly data over a five year period reveals biologically relevant characteristics of the salmon host-salmon louse parasite association in salmon farming. The model will be used as a simulation tool to explore large-scale effects of various salmon louse control strategies.

Several refinements of the present model can be suggested. Salinity is known to affect salmon lice abundance but is not a part of the model at present, since salinity measurements on farms are not available. However, salinity calculated by hydrodynamic models has now become available for the whole coast of Norway, and may be included as an extra variable in the κ^{susc} term.

In our model, we have used the seaway distance d_{ij} as the distance measure between two farms i and j. However, without any change in the model, this may be replaced by a "sea current distance" c_{ij}, i.e. an index of average contact between farms calculated by simulating the spread of sea lice larvae from a hydrodynamic model. This allows for non-symmetrical distances, i.e. the contact c_{ij} from farm j to i may differ from the contact c_{ji} in the opposite direction. However, the use of such distances require that they are calculated for all pairs of farms covered by the model and over the time covered by the model since the hydrodynamic forcing will change over time. At present, such data are not available.

Acknowledgements

This work was funded by the Research Council of Norway through the project: 254830/E40 "Host density and pathogen transmission in salmon aquaculture: effects of increased production and pathogen control policies". In addition, Henrik Stryhn was supported by The Research Council of Norway with a personal grant (Project: 268051/E40 "Stochastic models on pathogen transmission in aquaculture").

References

Aaen SM, Helgesen KO, Bakke MJ, Kaur K, Horsberg TE. 2015. Drug resistance in sea lice: a threat to salmonid aquaculture. *Trends in Parasitology* **31**: 72–81, doi:10.1016/j.pt.2014.12.006.

Abolofia J, Asche F, Wilen JE. 2017. The Cost of Lice: Quantifying the Impacts of Parasitic Sea Lice on Farmed Salmon. *Marine Resource Economics* **32**: 329–349, doi:10.1086/691981.

Aldrin M, Huseby RB, Stien A, Grøntvedt RN, Viljugrein H, Jansen PA. 2017. A stage-structured Bayesian hierarchical model for salmon lice populations at individual salmon farms - Estimated from multiple farm data sets. *Ecological Modelling* **359**: 333–348, doi:10.1016/j.ecolmodel.2017.05.019.

Aldrin M, Storvik B, Kristoffersen AB, Jansen PA. 2013. Space-time modelling of the spread of salmon lice between and within Norwegian marine salmon farms. *PLoS ONE* **8**: e64039, doi:10.1371/journal.pone.0064039.

Anonymous. 2017a. *Akvakulturregisteret (in Norwegian)*,

The Norwegian Ministry of Trade, Industry and Fisheries, Oslo, http://www.fiskeridir.no/Akvakultur/Registre-og-skjema/Akvakulturregisteret, downloaded 20th January 2017.

—. 2017b. *Fiskehelse (in Norwegian)*, Barentswatch, https://www.barentswatch.no/fiskehelse/.

Bricknell IR, Dalesman SJ, O'Shea B, Pert CC, Luntz AJ. 2006. Effect of environmental salinity on sea lice *Lepeophtheirus salmonis* settlement success. *Dis Aquat Organ* **71**: 201:2012, doi:10.3354/dao071201.

Brooks-Pollock E, Roberts GO, Keeling MJ. 2014. A dynamic model of bovine tuberculosis spread and control in Great Britain. *Nature* **511**, doi:10.1038/nature13529.

Byrd RH, Lu P, Nocedal J, Zhu C. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **16**: 1190–1208.

Efron B, Tibshirani RJ. 1993. *An introduction to the bootstrap*, Chapman & Hall/CRC, Boca Raton.

Eichner C, Hamre LA, Nilsen F. 2015. Instar growth and molt increments in *Lepeophtheirus salmonis* (Copepoda: Caligidae) chalimus larvae. *Parasitology International* **64**: 86–96, doi:10.1016/j.parint.2014.10.006.

Forseth T, Barlaup BT, Finstad B, Fiske P, Gjøsæter H, Falkegård M, Hindar A, Mo TA, Rikardsen AH, Thorstad EB, Vøllestad V L. A. Wennevik. 2017. The major threats to Atlantic salmon in Norway. *ICES Journal of Marine Science* **74**: 1496–1513, doi:10.1093/icesjms/fsx020.

Grenfell BT, Wilson K, Isham VS, Boyd HEG, Dietz K. 1995. Modelling patterns of parasite aggregation in natural populations: trichostrongylid namatode-ruminant interactions as a case study. *Parasitology* **111 (Suppl.)**: S135–S51.

Hamre LA, Eichner C, Caipang CMA, Dalvin ST, Bron JE, Nilsen F, Boxshall G, Skern-Mauritzen R. 2013. The salmon louse *Lepeophtheirus salmonis* (Copepoda:Caligidae) life cycle has only two chalimus stages. *PLoS ONE* **8**: e73539.

Irvine RJ, Stien A, Halvorsen O, Langvatn R, Albon SD. 2000. Life-history strategies and population dynamics of abomasal nematodes in Svalbard reindeer (*Rangifer tarandus plathyrhynchus*). *Parasitology* **120**: 297–311.

Jansen P, Kristoffersen AB, Viljugrein H, Jimenez D, Aldrin M, Stien A. 2012. Sea lice as a density dependent constraint to salmonid farming. *Proc R Soc B* Doi:10.1098/rspb.2012.0084.

Jansen PA, Grøntvedt RN, Tarpai A, Helgesen KO, Horsberg TE. 2016. Surveillance of the Sensitivity towards Antiparasitic Bath-Treatments in the Salmon Louse (*Lepeophtheirus salmonis*). *PLoS ONE* **11**: e0149006, doi.org/10.1371/journal.pone.0149006.

Kristoffersen AB, Qviller L, Helgesen KO, Vollset KW, Viljugrein H, Jansen PA. 2017. Quantitative risk assessment of salmon louse-induced mortality of seaward-migrating post-smolt Atlantic salmon. *Epidemics* Available online 2 December 2017, doi:10.1016/j.epidem.2017.11.001.

Liu YJ, Bjelland HV. 2014. Estimating costs of sea lice control strategy in Norway. *Preventive Veterinary Medicine* **117**: 469–477, doi:10.1016/j.prevetmed.2014.08.018.

781 Pawitan Y. 2001. *In all likelihood - statisctical modelling and inference using*
782 *likelihood*, Oxford, UK: Crarendon Press.

783 Pettersen JM, Brynildsrud OB, Huseby RB, Rich K, Aunsmo A, Jensen BB,
784 Aldrin M. 2016. The epidemiological and economic effects from systematic
785 depopulation of Norwegian marine salmon farms infected with pancreas disease
786 virus. *Preventive Veterinary Medicine* **132**: 113–124.

787 Salama N, Murray A, Rabe B. 2016. Simulated environmental transport dis-
788 tances of *Lepeophtheirus salmonis* in Loch Linnhe, Scotland, for informing
789 aquaculture area management structures. *Journal of Fish Diseases* **39**: 419–
790 428, doi:10.1111/jfd.12375.

791 Samsing F, Oppedal F, Dalvin S, Johnsen I, Vågseth T, Dempster T.
792 2016. Salmon lice (*Lepeophtheirus salmonis*) development times, body size,
793 and reproductive outputs follow universal models of temperature depen-
794 dence. *Canadian Journal of Fisheries and Aquatic Science* **73**: 1841–1851,
795 doi:10.1139/cjfas-2016-0050.

796 Stien A, Bjørn PA, Heuch PA, Elston DA. 2005. Population dynamics of salmon
797 lice *Lepeophtheirus salmonis* on Atlantic salmon and sea trout. *Mar Ecol Prog*
798 *Ser* **290**: 263–275.

799 Taranger G, Karlsen Ø, Bannister R, Glover K, Husa V, Karlsbakk E, Kvamme
800 B, Boxaspen K, BjÃ¸rn P, Finstad B, Madhun A, Morton H, Svåsand T. 2015.
801 Risk assessment of the environmental impact of Norwegian Atlantic salmon
802 farming. *ICES Journal of Marine Science* **72**: 997–1021.

803 Tildesley MJ, Smith G, Keeling MJ. 2012. Modeling the spread and control of

foot-and-mouth disease in Pennsylvania following its discovery and options for control. *Prev. Vet. Med.* **104**: 224–239.

Torrissen O, Jones S, Asche F, Guttormsen A, Skilbrei OT, Nilsen F, Horsberg TE, Jackson D. 2013. Salmon lice – impact on wild salmonids and salmon aquaculture. *Journal of Fish Diseases* **36**: 171–194, doi:10.1111/jfd.12061.

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM. 2009. *Mixed Effects Models and Extensions in Ecology with R*, Springer, 1st ed., 574 p., doi:10.1007/978-0-387-87458-6.

- A partly stage-structured model for salmon lice in a network of fish farms is introduced
- Adult females lice and other mobile stage are treated separately
- The time for recruitment of other mobile lice from the previous generation adult females increases by increasing temperatures
- The infection pressure from neighbouring farms decreases by increasing seaway distances to the neighbours
- The model can be used for prediction and for scenario simulations