Ecological Modelling 359 (2017) 333–348

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A stage-structured Bayesian hierarchical model for salmon lice populations at individual salmon farms – Estimated from multiple farm data sets

M. Aldrin[a,*], R.B. Huseby[a], A. Stien[b], R.N. Grøntvedt[c], H. Viljugrein[d], P.A. Jansen[d]

[a] Norwegian Computing Center, P.O. Box 114 Blindern, NO-0314 Oslo, Norway
[b] Norwegian Institute for Nature Research, P.O. Box 6606 Langnes, 9296 Tromsø, Norway
[c] INAQ AS, P.O. Box 1223 Sluppen, NO-7462 Trondheim, Norway
[d] Norwegian Veterinary Institute, P.O. Box 750 Sentrum, NO-0106 Oslo, Norway

ARTICLE INFO

Article history:
Received 11 January 2017
Received in revised form 18 May 2017
Accepted 23 May 2017

Keywords:
Population model
Aquaculture
Stochastic model
Sea lice counts

ABSTRACT

Salmon farming has become a prosperous international industry over the last decades. Along with growth in the production farmed salmon, however, an increasing threat by pathogens has emerged. Of special concern is the propagation and spread of the salmon louse, *Lepeophtheirus salmonis*. To gain insight into this parasite's population dynamics in large scale salmon farming system, we present a fully mechanistic stage-structured population model for the salmon louse, also allowing for complexities involved in the hierarchical structure of full scale salmon farming. The model estimates parameters controlling a wide range of processes, including temperature dependent demographic rates, fish size and abundance effects on louse transmission rates, effect sizes of various salmon louse control measures, and distance based between farm transmission rates. Model parameters were estimated from data including 32 salmon farms, except the last production months for five farms, which were used to evaluate model predictions. We used a Bayesian estimation approach, combining the prior distributions and the data likelihood into a joint posterior distribution for all model parameters. The model generated expected values that fitted the observed infection levels of the chalimus, adult female and other mobile stages of salmon lice, reasonably well. Predictions for the periods not used for fitting the model were also consistent with the observational data. We argue that the present model for the population dynamics of the salmon louse in aquaculture farm systems may contribute to resolve the complexity of processes that drive this host-parasite relationship, and hence may improve strategies to control the parasite in this production system.

1. Introduction

Salmon farming has become a large and economically prosperous international industry over the last decades. Norway holds a leading position as a producer of farmed salmonids with an annual production of about 1.2 million tonnes, which is roughly half of the worldwide production (Anonymous, 2015). Further growth in the production of salmonids is in demand (Anonymous, 2015), but this will come at the cost of increasing risks of pathogen propagation and transmission. Large-scale host density dependence acting on pathogen transmission has been demonstrated in salmon farming production systems, both for macro parasites (Aldrin et al., 2013; Jansen et al., 2012; Kristoffersen et al., 2014) and viruses (Aldrin et al., 2011, 2010; Kristoffersen et al., 2009). Of special concern, is the propagation and spread of the salmon louse, *Lepeophtheirus salmonis*, which is implicitly responsible for regulating the salmon farming industry through density dependent host parasite interactions (Frazer et al., 2012; Jansen et al., 2012; Groner et al., 2016b) Consequently, this parasite plays a dominant role in the formulation of management policies (Anonymous, 2015), dominates among salmon pathogens in the scientific literature (Murray et al., 2016) and is perceived as a major threat to wild salmon populations (Taranger et al., 2015; Vollset et al., 2015; Forseth et al., 2017), all testifying to the gravity of detrimental effects of the salmon louse on salmon farming.

Mathematical and statistical models are increasingly being used to evaluate infection pathways and risk factors for pathogen propagation and disease development, both in aquatic and terrestrial

* Corresponding author.
E-mail address: magne.aldrin@nr.no (M. Aldrin).

http://dx.doi.org/10.1016/j.ecolmodel.2017.05.019

animal farming (Aldrin et al., 2013, 2011, 2010; Salama and Murray, 2013; Murray and Salama, 2016; Jonkers et al., 2010; Diggle, 2006; Höhle, 2009; Keeling et al., 2001; Scheel et al., 2007). When such models are capable to reproduce the main patterns in the host-pathogen population dynamics, including the spread within and between farms, they can be used to predict future infection levels as well as simulate the outcomes of disease mitigation scenarios, examples being interventions to mitigate bovine tuberculosis in Great Britain (Brooks-Pollock et al., 2014) and long term effects of infection control measurements to mitigate salmonid alphavirus (SAV) incidences causing pancreas disease (PD) outbreaks (Aldrin et al., 2015). The Norwegian salmonid production system is exceptionally well suited for developing models for salmon lice infection dynamics because of the wealth of surveillance time-series that document both the spatial locations and population sizes of host populations at risk of infection, as well as salmon lice abundances in these host populations. Coupling these host and parasite population data have provided insights into e.g. how salmon lice spread between farms depending on between-farm distances, and how transmission and parasite abundances depend on local host biomasses (Jansen et al., 2012; Aldrin et al., 2013; Kristoffersen et al., 2014). However, previous models that describe both between and within farm parasite population dynamics have for simplicity typically been autoregressive statistical models focusing on single aggregated measures of parasite infection levels (Jansen et al., 2012; Aldrin et al., 2013; Kristoffersen et al., 2013, 2014). Alternatively, models have been developed for simulation purposes only (Groner et al., 2014) or focused on the population dynamics on single farms over a limited period (Krkošek et al., 2010). Most of these approaches have relied heavily on estimates of demographic rates obtained in the laboratory (Stien et al., 2005; Revie et al., 2005; Gettinby et al., 2011; Groner et al., 2013, 2016a; Rittenhouse et al., 2016).

The aim of the present paper is to formulate a fully mechanistic stage-structured population model for the salmon louse, that also allows for the complexities inherent in full scale salmon farming. Furthermore, the model accounts for the hierarchical structure of the data obtained from the production system where salmon lice are counted on subsamples of fish, the fish being aggregated into separate cages and the cages being aggregated to farm. The model estimates parameters controlling a wide range of processes, including effects of temperature on demographic rates, fish size and abundance effects on transmission rates, the different effect sizes, temporal and stage specific effects of a wide range of salmon lice control measures, and distance-based transmission rates between farms. The objectives for developing such a complex population model for the salmon louse was to cover several needs in modern salmon farming: (1) To develop a tool that keeps account of the salmon louse populations at the production unit level in salmon farms, based on the successive counting of salmon louse infections, and produce more reliable estimates of salmon louse population sizes at any point in time than the individual salmon louse counts. (2) To produce short term predictions of future salmon louse infection levels to enable proactive use of salmon louse management actions. (3) To evaluate the efficiency of different control measures. (4) To evaluate whether estimates of demographic rates obtained in the laboratory seems applicable in full scale production settings. (5) To explore importance of different sources of infection (e.g. internal versus external sources). (6) Finally, to develop a sufficiently realistic model that can be used for scenario-simulations exploring the effects of various parasite control strategies. In this paper, we describe the model in detail and discuss the results in relation to the first five objectives. Scenario simulations from the model will be the focus in later work.

2. Materials and methods

2.1. Data

Farm production of salmon comprises a freshwater juvenile phase, being followed by a marine grow out phase, the latter which is the focus of this study. The production of salmon on a Norwegian marine farm is initiated by stocking juvenile smolts to cages (or net-pens) either in spring or in autumn. Salmon are kept in the marine farms for about 1.5 years after which they are slaughtered for food consumption. In Norway, only fish of the same year class of age are kept on a given farm and we term this a cohort throughout the present paper. After slaughtering, it is mandatory to fallow the farm for a period of at least two months before stocking a new cohort of salmon.

The main body of data in the present study consist of cage-level data from 32 marine salmon farms in Norway, of which 12 farms are aggregated just north of the island Frøya in Mid-Norway (Fig. 1). For each farm, the data covers a full production cycle for farmed salmon, from stocking as smolts to slaughtering as adult Atlantic salmon (*Salmo salar* L.), including fish production data, lice counts, temperatures and louse control efforts. Salmon were stocked between 2011 and 2013 and slaughtered about 1 1/2 year after stocking, between 2012 and 2014. The number of production units (cages) per farm varied from 3 to 12, but were usually around 8 (mean 7.7). For 9 farms, the fish were moved between cages within the farm during the production period.

Seawater temperatures were measured at 3 m depth at the farms. The average temperature was 9.1 °C, and 95% of the temperatures were between 3.6 and 15.0 °C. Data on salinity were unavailable in sufficient detail and have therefore not been used.

The production data consist of daily numbers and mean weights of salmon per cage during the production period, information on movement of salmon between cages within farms and information on antiparasitic lice treatment using chemotherapeutic medicals (day of application and type of medical). Furthermore, the data contain information on stocking of cleaner fish (day and number of cleaner fish stocked), but with limited information on their mortality, and hence on the number of cleaner fish present at a given day. The cleaner fish are usually either lumpsuckers (*Cyclopterus lumpus*), ballan wrasses (*Labrus bergylta*) or goldsinny wrasses (*Ctenolabrus rupestris*) or a mix of these, but we do not distinguish between various species of cleaner fish in the model.

The production cycles lasted on average 16.5 months per cage, with on average 140 000 fish per cage, typically more in the beginning of a production cycle and less towards the end. The average minimum and maximum fish weights during a production cycle was 140 g and 5.7 kg, respectively. 89% of the cages contained cleaner fish in parts of the production period.

As a main rule, lice counts were performed on a sample of at least 10 fish every second week for each cage. The salmon lice were divided into three categories according to developmental stages, *i.e.* (i) chalimus (CH), (ii) other mobiles (OM), which consist of preadults of both sexes and adult males, and (iii) adult females (AF). There were on average 41 dates with lice counts per cage, with averages (abundance) of 0.23 CH, 0.76 OM and 0.18 AF per fish. Six different types of antiparasitic medicals were used (Table 1), and there were on average 4.6 events of medical treatments per cage. The medicals emamectin benzoate and diflubenzuron are given through the feed, typically over a period of around two weeks. These treatments have a relatively low daily effect, but effects last over a prolonged period. The other medicals are applied as bath treatments over a duration of a few hours, with a larger daily effect, but lasting over a shorter period (see Section 2.2.3).

In addition to the detailed cage-level data on the 32 farms, we have more aggregated data on all other Norwegian marine salmon

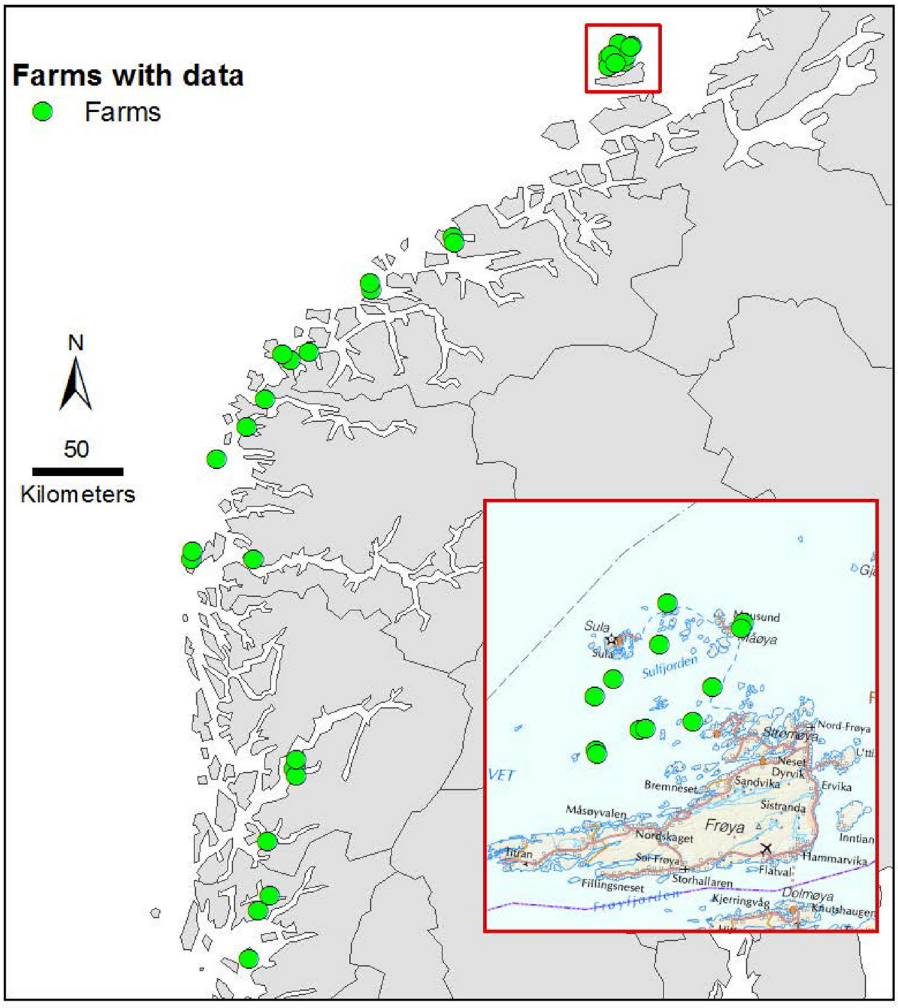

Fig. 1. Geographical positions of the 32 salmon farms on the West coast of Norway (green circles). The highlighted area contains the 12 farms north of the island Frøya. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Overview of types of medical treatments used. Codes: Y = yes, N = no, NA = missing information. The table content is based on information in Nygaard (2010) and Ottesen et al. (2012). The delay Δ^{del} and the duration constant Δ^{dur} are defined in Section 2.2.3. For hydrogen peroxide, the unit for δ^{dur} is days.

Medical	Product name	Delay Δ^{del} (days)	Duration constant δ^{dur} (°C·days or days)	Temperature dependency	Effect on CH	Effect on PA	Effect on A	Effect on egg
Hydrogen peroxide (H_2O_2)		0	7	N	N	Y	Y	NA
Deltamethrin	Alphamax	2	84	Y	Y	Y	Y	N
Cypermethrin	Betamax	2	84	Y	Y	Y	Y	N
Azamethiphos	Salmosan	1	42	Y	N	Y	Y	N
Emamectin benzoate	Slice	5	210	Y	Y	Y	Y	NA
Diflubenzuron	Releeze	10	126	Y	Y	Y	N	N

farms. For a farm f at day t, we know the number of salmon, denoted by N^{SAL}_{tf}. We also have an estimate $\hat{A}^{AF}_{tf}$ of the abundance of AF lice at the farm, based on weekly lice counts on a sample of fish, and therefore also an estimate of the total number of AF lice, given by $\hat{N}^{AF}_{tf} = \hat{A}^{AF}_{tf} N^{SAL}_{tf}$. Finally, we have the seaway distances between all farms, and we let $d_{ff'}$ denote the seaway distance between a farm f and another farm f'. These data were used to calculate an external infection pressure index(see Section 2.2.9).

Fig. 2 shows the most relevant data for one cage at one farm. The upper panel shows time plots of the seawater temperature (on the left y-axis) and the external infection pressure index. In addition, the first stocking of salmon in this cage is indicated by the vertical pink line and the various medical treatments are shown as blue vertical lines. Finally, the stocking of cleaner fish is also shown as vertical lines. The vertical extension of these lines is proportional to the stocked cleaner fish ratio (on the right y-axis), *i.e.* the number of stocked cleaner fish divided by the number of salmon. The lower three panels show the counted abundance of lice in the CH, OM and AF categories.

Fig. 2. Counted lice abundance and other information for one cage at one farm. Upper panel: Seawater temperature (green line), external infection pressure index (red dotted curve), time of stocking (pink vertical line), treatments (blue vertical lines) and stocked cleaner fish ratio (black vertical lines). Three lower panels: Counts of chalimi, other mobiles or adult females shown as green circles connected by straight lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Model framework

2.2.1. Modelling background and overview

Many authors have previously presented models for salmon louse population dynamics. Most of these models use parameter values obtained in a laboratory (Stien et al., 2005; Groner et al., 2013, 2016a) or from small scale experimental units in the marine environment (Krkošek et al., 2009). When full scale farm-production data have been used for estimation, these have been aggregated and only a few of the model parameters were estimated from these data (Revie et al., 2005; Gettinby et al., 2011). The present estimating approach is fundamentally different. All the parameters are estimated by fitting our model to the individual lice counts collected through the production period in commercial fish farms. However, to ensure that the final parameter values are within biological plausible ranges, we use laboratory data, mostly those summarised in Stien et al. (2005), to specify informative prior distributions for many of the parameters. The priors are then updated to posterior distributions by the full scale farm data using Bayesian methods.

Since the model is estimated on real data from many different farms under various conditions, it must simultaneously incorporate many features to handle activities or events that affect lice abundance, including various types of treatments, external infection from neighbouring farms and the movement of fish (and then also lice) between cages at the same farm. Our model is therefore more complex than the aforementioned models.

Biologically, the life cycle of the salmon louse consists of eight developmental stages (Hamre et al., 2013). These are aggregated into the following five stages in our model: (i) recruits (R, eggs and nauplii larvae), (ii) copepodids (CO, infective planktonic larvae), (iii) chalimi (CH, sessile lice on fish), (iv) pre-adults (PA, mobile lice on fish) and (v) adults (A, also mobile lice on fish). The adults are further divided into adult females (AF) and adult males (AM).

The main stage structure and the focus on temperature dependencies in development and recruitment are aspects of the model that are inspired by the model in Stien et al. (2005). Stien et al. (2005) suggest a delay integro-differential equation model, a model structure commonly used for parasitic gastrointestinal nematodes of farm animals (*e.g.* Grenfell et al., 1987) and free living insects populations (*e.g.* Nisbet and Gurney, 1983; Nelson et al., 2013). We have chosen a discrete time version that uses a one day time step. This could indicate that matrix models and associated methods and theory could be used to analyse the model (Caswell, 2001). However,

Table 2
Overview of the model notation. When relevant, quantities may be used with subscripts *f*, *t*, *a* and *c*, and with superscripts *R*, *CO*, *CH*, *PA*, *AF* or *AM*.

f	index for farm
t	index for time (day)
a	index for stage-age
c	index for cage
N^R	total number of lice recruits
N^{CO}	total number of copepodids
N^{CH}	total number of chalimus larvae on fish in a given cage
N^{PA}	total number of pre-adult lice on fish in a given cage
N^{AF}	total number of adult female lice on fish in a given cage
N^{AM}	total number of adult male lice on fish in a given cage
s	survival rate (proportion per day)
m	mortality rate ($m = 1 - s$)
d	development rate (proportion per day)
r	reproduction rate (numbers per day and per AF lice)
e^{Ext}	modifying factor for external recruitment
N^{SAL}	number of salmon
W	average weight of salmon
$M^{SAL}_{c'c}$	number of salmon moved from cage c' to cage c
$w_{c'c}$	proportion of salmon moved, $M^{SAL}_{c'c}/N^{SAL}_{c'}$
$A = N^{AF}/N^{SAL}$	abundance of adult female lice on fish in a given cage
N^{CLF}	number of cleaner fish
S^{CLF}	number of cleaner fish stocked
Y	Number of lice counted on a sample of n fish
λ	parameters related to mortality
δ	parameters related to development
μ	expected values
$\beta, \gamma, \kappa, \rho$	various parameters
σ^2	variances
z	autoregressive processes
ϕ	autoregressive coefficients

the strong temperature dependence in development and recruitment rates makes this difficult. As the temperature varies through the year, the temperature lice experience while at a given stage needs to be kept account of. Accordingly, we will use the expression "stage-age" throughout the manuscript to denote that the age of a cohort of salmon lice within a stage is kept track of throughout, and used to determine development rates to the next stage.

The general idea is that for stage-age $a = 0$, the lice have developed into the given stage from the previous stage, and that for $a > 0$, the lice can develop into the subsequent stage. We further assume that within a day, in the following order;

(i) lice may be counted on a sample of fish,
(ii) lice may die due to natural mortality or treatment,
(iii) the surviving lice might develop to the next stage, and finally,
(iv) fish, including their sessile and mobile lice, can be moved to another cage within the farm or be removed from the farm.

Removement of fish from the farm may be due to slaughter, fish mortality of other reasons or, occasionally, movement of fish to another farm. Note that we have data on stocking of fish, on movement of fish between cages within the farm and on removal of fish from the farm, and there is therefore no need to model the population dynamics of the farmed salmonids themselves. When fish are stocked to marine farms as smolt, they are free of lice, and hence the initial lice transmission is caused by external infections. Not until some lice at the farm have developed into adults, can the internal infection process start. The population model for a farm with two cages is illustrated in Fig. 3. In the R and CO stages, the lice are associated with the farm, but not with any specific cage. From the CH stage and beyond, however, the lice have infected the fish and are therefore associated with specific cages. In the following subsections we describe the various aspects of the population model and how the model is related to lice count data. Table 2 gives an overview of the main notation we use.

2.2.2. Population model

Below, we present the population model for lice at a given farm (but for simplicity without the farm index *f*). The model consists of two equations per stage. The first equation handles lice entering a given stage at stage-age 0, typically from the preceding model stage. The second per-stage equation handles lice ageing into higher stage-ages without developing into another stage.

The submodels for survival, development and reproduction rates are presented later in Sections 2.2.3–2.2.6.

Model for the recruitment stage

At the R stage, lice are not associated with a specific cage, but rather seen as a reservoir of recruits with the potential to infect a fish host at the farm in question in the future. The model is:

$$N^R_{t(a=0)} = e^{Ext}_t N^{AFExt}_{t-1} r^{Ext}_{t-1} + \sum_c \sum_{a'} [N^{AF}_{(t-1)a'c} s^{AF}_{(t-1)a'c} r_{(t-1)a'c}], \quad (1)$$

$$N^R_{t(a>0)} = N^R_{(t-1)(a-1)} s^R_{(t-1)(a-1)} [1 - d^R_{(t-1)(a-1)}]. \quad (2)$$

The first term in Eq. (1) represents recruitment (into stage-age 0) from neighbouring farms, also called external recruitment. Here, N^{AFExt}_{t-1} is a weighted sum of adult females at neighbouring farms at time $t-1$ (defined in Section 2.2.9). Furthermore, we assume that these reproduce with a rate r^{Ext}_{t-1} (see Section 2.2.7). Then, $N^{AFExt}_{t-1} r^{Ext}_{(t-1)}$ can be interpreted as a preliminary estimate of the number of external recruits reaching the farm. However, this accounts for seaway distances to neighbouring farms, but not for the sea currents in the area that may be more or less favourable for a given farm, and which also may vary over time. Therefore, we have introduced the modifying factor e^{Ext}_t, which is a farm-dependent and time varying modifying factor, see Section 2.2.8 for an exact definition. Note that Eq. (1) only takes into account infections from salmon lice on farmed fish. One could include an extra term to account for infections from wild salmon and trout, but these are so few compared to the farmed fish that they can be neglected (Johansen et al., 2011).

The second term in (1) represents recruitment (into stage-age 0) from adult female lice at the same farm, also called internal recruitment. The product $N^{AF}_{(t-1)a'c} s^{AF}_{(t-1)a'c}$ is the number of adult females at stage-age a' in cage c that survives at time $t-1$ and they reproduce with a rate $r_{(t-1)a'c}$ (see Section 2.2.6). The new recruits are summed over all possible stage-ages of the adult females and over all cages.

Eq. (2) keeps track of the number of recruits of stage-age $a > 0$ that (i) survives from the previous time point with survival rate $s^R_{(t-1)(a-1)}$ (see Section 2.2.3) and (ii) do not develop into the infective CO stage, where $d^R_{(t-1)(a-1)}$ is the development rate, i.e. the proportion of recruits that develop into the CO stage (see Section 2.2.4).

The equations for the next stages use similar notation for survival rates, development rates and numbers of lice.

Model for the copepodid stage

$$N^{CO}_{t(a=0)} = \sum_{a'} N^R_{(t-1)a'} s^R_{(t-1)a'} d^R_{(t-1)a'}, \quad (3)$$

$$N^{CO}_{t(a>0)} = N^{CO}_{(t-1)(a-1)} s^{CO}_{(t-1)(a-1)} [1 - \sum_c d^{CO}_{(t-1)(a-1)c}]. \quad (4)$$

Here, the development rate $d^{CO}_{(t-1)(a-1)c}$ (see Section 2.2.5) represents the infection rate, i.e. the proportion of the available copepodids that during a day infect fish in cage c and thus enter the CH stage. The sum over cages, $d^{CO}_{(t-1)(a-1)} = \sum_c d^{CO}_{(t-1)(a-1)c}$ is then the total infection rate at the farm. This is modelled as independent of stage-age. We note that when an infective copepodid (stage CO) attaches to a fish host, it takes approximately 24 hours before it

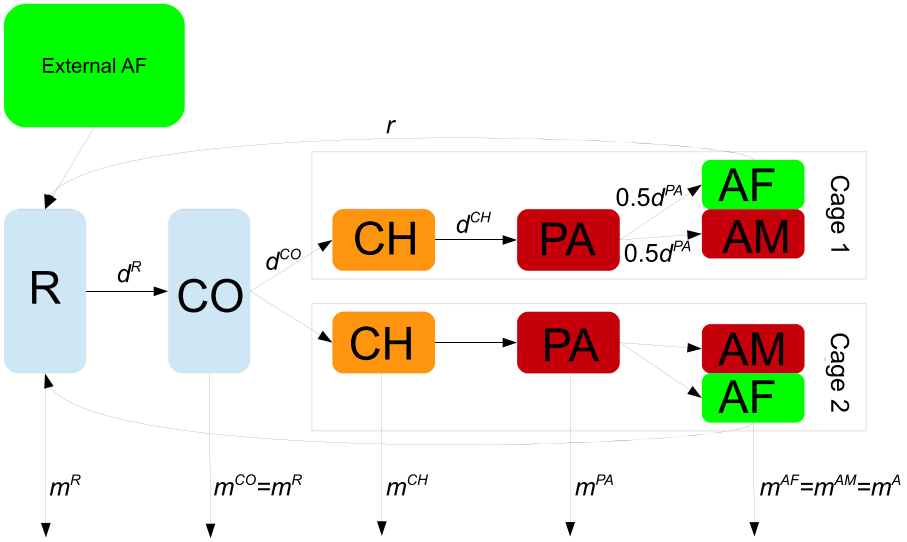

Fig. 3. Overview of the population model for the salmon louse. Lice in the orange, red and green stages are counted, whereas lice in the blue stages are not counted. Lice are associated with a cage from the chalimus stage, here illustrated by a farm with two cages. The d-s, m-s and r-s symbolise development, mortality and recruitment, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

moults into the CH stage (Krkošek et al., 2009). However, in the model we ignore this short period and assume that a copepodid enters the CH stage immediately upon attachment to a fish host.

Model for the chalimus stage

$$N^{CH}_{t(a=0)c} = \sum_{a'} N^{CO}_{(t-1)a'} s^{CO}_{(t-1)a'} d^{CO}_{(t-1)a'c}, \tag{5}$$

$$N^{CH}_{t(a>0)c} = \sum_{c'} N^{CH}_{(t-1)(a-1)c'} s^{CH}_{(t-1)(a-1)c'} [1 - d^{CH}_{(t-1)(a-1)}]. \tag{6}$$

From the CH stage on, the lice are attached to a fish, and therefore associated with a specific cage. We assume that the attached lice follow the fish if the fish are moved to another cage or if the fish are removed from the farm (including slaughtering and other fish mortality). To handle this, the equations given here for the CH, PA and AF stages are extended slightly (see Section 1.2 in the Supplementary material).

Model for the pre-adult stage

$$N^{PA}_{t(a=0)c} = \sum_{a'} \sum_{c'} N^{CH}_{(t-1)a'c'} s^{CH}_{(t-1)a'c'} d^{CH}_{(t-1)a'}, \tag{7}$$

$$N^{PA}_{t(a>0)c} = \sum_{c'} N^{PA}_{(t-1)(a-1)c'} s^{PA}_{(t-1)(a-1)c'} [1 - d^{PA}_{(t-1)(a-1)}]. \tag{8}$$

Model for the adult stages

For the adult stage, we distinguish in principle between males and females. However, we assume that males and females have the same survival and development rates and therefore each constitute 50 % of the adults. The main reason for this is that we do not have data on the number of adult males on the fish, and therefore do not have the information necessary for separate estimation of adult male demographic rates. The equations for adult females are then

$$N^{AF}_{t(a=0)c} = 0.5 \sum_{a'} N^{PA}_{(t-1)a'c'} s^{PA}_{(t-1)a'c'} d^{PA}_{(t-1)a'}, \tag{9}$$

$$N^{AF}_{t(a>0)c} = \sum_{c'} N^{AF}_{(t-1)(a-1)c'} s^{AF}_{(t-1)(a-1)c'}, \tag{10}$$

while the number of adult males is equal to the number of adult females:

$$N^{AM}_{tac} = N^{AF}_{tac}. \tag{11}$$

2.2.3. Survival rates

We assume that the survival rates may be farm-specific for some stages, and therefore use the index f when convenient, but we sometimes drop the superscript that indicates the stage name. We assume that the total survival rate is the product of three terms;

$$s_{tfac} = s^{nat}_{tfac} \cdot s^{clf}_{tfac} \cdot s^{cht}_{tfac} = (1 - m^{nat}_{tfac}) \cdot (1 - m^{clf}_{tfac}) \cdot (1 - m^{cht}_{tfac}), \tag{12}$$

where s^{nat}_{tfac} is survival after natural mortality, s^{clf}_{tfc} is survival after additional mortality due to cleaner fish predation (independent of stage-age) and s^{cht}_{tfc} is survival after additional mortality due to chemotherapeutic treatment (independent of stage-age). The m-s denote the corresponding mortalities. The two latter terms are relevant only for the CH, PA and A stages. All the three mortality (and survival) terms must lie between 0 and 1, but they have different structures.

Natural mortality

For the R and CO stages, we simply assume that the natural mortality is a constant that is common for both stages, *i.e.*

$$m^{Rnat}_{tfac} = m^{Rnat} = \lambda^{RCOnat}, \tag{13}$$

$$m^{COnat}_{tfac} = m^{COnat} = \lambda^{RCOnat} \tag{14}$$

For each of the CH, PA and A stages, we assume that the natural mortalities are stochastic processes that can vary over time and between farms, but are common for all cages within a farm and independent of stage-age. This may account for factors that differ between farms and change over time, for instance salinity, which is not included in the model. Furthermore, including these mortalities as farm-specific and time-varying terms improves the fit of the model to data. For each farm and louse stage, the mortality is assumed to follow an autoregressive model of order 1 (AR(1)) on the logit-scale as

$$m^{nat}_{tfac} = m^{nat}_{tf} = \exp(z^{nat}_{tf}/(1 + \exp(z^{nat}_{tf})), \tag{15}$$

$$(z^{nat}_{tf} - \lambda^{nat}_0) = \phi^{nat} \cdot (z^{nat}_{(t-1)f} - \lambda^{nat}_0) + \varepsilon^{nat}_{tf}, \tag{16}$$

$$\mathrm{Var}(\varepsilon^{nat}_{tf}) = (\sigma^{nat})^2. \tag{17}$$

Here $z^{nat}_{tf} = \mathrm{logit}(m^{nat}_{tf}) = \log(m^{nat}_{tf}/(1 - m^{nat}_{tf}))$. Furthermore, λ^{nat}_0 is the expected value on the logit-scale, ϕ^{nat} an autoregressive coefficient and ε^{nat}_{tf} a white noise process with variance $(\sigma^{nat})^2$. These

parameters have separate values for each stage. In addition, the time-varying mortalities m^{nat}_{tfac} are restricted to lie within specified intervals, which are (0.0006–0.02) for CH, (0.002–0.21) for PA and (0.0003–0.70) for A. These limits are motivated from the various studies summarised in Stien et al. (2005), and are simply the most extreme limits of the intervals given in their Table 4.

In addition, we assume $m = 1$ from stage-age 80 days for adults and from stage-age 60 days for the other stages. This is an approximation made to save computer time.

Mortality due to cleaner fish

We assume that cleaner fish feed on lice at the PA and A stages only (Leclercq et al., 2014), and that the corresponding mortality for these two stages are equal. Let $x^{clf}_{tfc} = N^{CLF}_{tfc}/N^{SAL}_{tfc}$ be the ratio of the number of cleaner fish to the number of salmon in cage c, at farm f and time t. This ratio depends among others on the mortality of cleaner fish, which has to be estimated, and the model for this is described later in Section 2.2.10. Lice mortality in the PA and A stages due to cleaner fish is then given by

$$m^{clf}_{tfac} = m^{clf}_{tfc} = 1 - \exp(-\lambda^{clf} x^{clf}_{tfc}), \tag{18}$$

where the parameter λ^{clf} is non-negative, such that the mortality always is between 0 and 1. One reason for assuming such a simple model for the effect of cleaner fish (as opposed to the effect of medical treatments discussed below) is that we also must estimate the cleaner fish ratio, which is multiplied by the cleaner fish effect.

Mortality due to chemotherapeutic treatment

We assume that the chemotherapeutic treatments introduce extra mortality of lice in some or all the stages CH, PA and A, depending on the type of treatment. However, for simplicity we assume that lice are only affected as long as they stay in the stage they were at the time of treatment. If they manage to develop to the next stage, they are clear of the treatment effect. Let the set of subscripts *fcbi* denote the ith application of a chemotherapeutic of type b in cage c at farm f. Assuming that this treatment was given at time t^0_{fcbi}, we define an indicator variable x^{cht}_{tfacbi} that is 1 when the treatment is active (a period after the treatment is given) for lice at stage-age a, i.e. when

$$t \in [t^0_{fcbi} + \Delta^{del}_b, t^0_{fcbi} + \Delta^{del}_b + \Delta^{dur}_{fcbi}) - 1] \text{and} a \geq t - t^0_{fcbi}. \tag{19}$$

Here Δ^{del} is a time delay (in days) from application until a treatment gives a visible effect, which varies between treatments (Table 1). Furthermore, Δ^{dur} is the duration of the effect, which depends on the seawater temperature according to

$$\Delta^{dur} = \delta^{dur}/T_{t^0}, \tag{20}$$

where δ^{dur} is a constant (with unit degree days, i.e. °C·days) given in Table 1 and T_{t^0} is the seawater temperature when the medical is applied. One exception is when hydrogen peroxide was applied, for which Δ^{dur} is temperature independent and given by $\Delta^{dur} = \delta^{dur}$ (in days).

The mortality due to chemotherapeutic treatment is given by

$$m^{cht}_{tfac} = 1 - \exp(\sum_b - u^{cht}_{fcbi} x^{cht}_{tfacbi}), \tag{21}$$

where u^{cht}_{fcbi} is a regression coefficient expressing the effect of the specific application of the treatment. These regression coefficients vary systematically between treatment types, accounting for varying efficiency of different types of treatments. In addition, they vary randomly between different applications of the same treatment type, which for instance may be due to a varying degree of resistance in the lice populations (Jansen et al., 2016). This is handled by the following formulation

$$u^{cht}_{fcbi} = \log(1 + \exp(u^{cht*}_{fcbi})), \tag{22}$$

$$u^{cht*}_{fcbi} \sim \mathrm{N}(\lambda^{cht}, (\sigma^{cht})^2). \tag{23}$$

In general, the parameters λ^{cht} and σ^{cht} differ between various types of treatment, but are set equal for some treatment types. These parameters are equal for the stages for which a treatment has effect, but the effect of a specific treatment *fcbi*, represented by the random coefficient u^{cht}_{fcbi}, may vary between stages (this is for simplicity omitted from the notation above).

We assume that the effects of deltamethrin and cypermethrin are equal, since these are similar compounds. Furthermore, when azamethiphos is used in combination with deltamethrin or cypermethrin, we assume it has the same effect as using deltamethrin or cypermethrin alone, since this combination is used when reduced treatment effect is expected due to resistance towards the medicals.

2.2.4. Development rates

We consider here the development rate from one stage to the next, for the stages R, CH and PA. The development rate for CO can be interpreted as an infection rate, and is treated in the next subsection. In all the population models for lice that we mentioned in the introduction, the development rate is 0 until some, minimum stage-age and afterwards positive and constant. In our opinion, the concept of a strict and absolute minimum development time can be questioned in a population with millions of individuals, and the assumption of a constant development thereafter may be unrealistic. We have chosen to consider the development to the next stage as a time-to-event process, and model it as a discretised version of a Weibull distribution. The Weibull distribution is widely used in statistical models for time-to-event or survival analysis (Aalen et al., 2008). It also gave a better fit to our data than the minimum development time model mentioned above (data not shown).

When the time to an event is continuous and Weibull distributed, the event rate (often called hazard in survival analysis) is $(\delta^{sc})^{-\delta^s}\delta^s a^{\delta^s - 1}$, where a is the stage-age or time, δ^s is a shape parameter and δ^{sc} is a scale parameter (sometimes $(\delta^{sc})^{-\delta^s}$ is termed the scale parameter). In our case, it is convenient to re-parameterise this as a function of the median time to event, δ^m, and the shape parameter. The event rate then becomes $\log(2)(\delta^m)^{-\delta^s}\delta^s a^{\delta^s - 1}$, since the median in the Weibull distribution is $\delta^{sc}(\log(2))^{(1/\delta^s)}$.

We use a discretised version of this, *i.e.* our development rate is the probability to develop to the next stage within a day, and it must therefore also be restricted to be at most one. We assume that the median time to develop may vary over time and between farms, and introduce therefore the subscripts *tfa* on it. The model for the development rate is then

$$d_{tfa} = \min(\log(2)(\delta^m_{tfa})^{-\delta^s}\delta^s a^{\delta^s - 1}, 1) \text{for} a = 0, 1, \ldots. \tag{24}$$

We further assume that the median development time depends on the temperature history as $\delta^m_{tfa} = c/(\bar{T}_{tfa})^{\delta^p}$, where $\bar{T}_{tfa}$ is the average temperature that lice at stage-age a at farm f have experienced, *i.e.* the average temperature from time $t - a$ to time t, c is a constant and δ^p is another constant that performs a power transformation of $\bar{T}_{tfa}$. To get a more clear interpretation of the constant c, we parameterise it as a function of the median development time at 10 ° C, denoted by δ^{m10}. The final model for the median development time then becomes

$$\delta^m_{tfa} = (10^{\delta^p}\delta^{m10})/(\bar{T}_{tfa})^{\delta^p} = \delta^{m10}(10/\bar{T}_{tfa})^{\delta^p}. \tag{25}$$

The development rate defined by Eqs. (24) and (25) also depends on the stage in the way that the parameters δ^{m10}, δ^s and δ^p are stage-specific. One motivation for introducing δ^{m10} as a basic parameter is that we use results on development times around 10 °C from other studies as prior information, to ensure that our estimates are within biological plausible ranges. For the R stage, which consists of eggs

and nauplii, this prior information is given separate for eggs and nauplii. Therefore, for the R stage, δ^{Rm10} is the sum of one parameter δ^{Em10} for eggs and another quantity δ^{Nm10} for nauplii, *i.e.*

$$\delta^{Rm10} = \delta^{Em10} + \delta^{Nm10}, \tag{26}$$

and δ^{Em10} is also contained in the reproduction rate introduced later in Section 2.2.6.

In the estimation, we restrict δ^s to be larger than 1, and the development rate will then be 0 at stage-age 0 and then increase by increasing stage-age. Furthermore, the larger δ^s is, the more steep will the development rate increase from 0 to 1 around the median. When $\delta^s > 2$, the difference between the mean and the median will be less than 7 %. It should further be noted that the parameter δ^{m10} is only approximately the median development time, since we consider a time-discrete version of the Weibull distribution.

Assuming a constant temperature T, Stien et al. (2005) modelled the minimum development time as $c_1/(T+c_2)^{c_3}$, where c_1, c_2 and c_3 are constants. They further assumed that $c_3 = 2$ and estimated c_1 and c_2. We use a similar formulation for the median development time, but assume $c_2 = 0$ and estimate c_1 and c_3. In practice, these two formulations are quite similar for the relevant temperatures and for the estimated values of $c_3 = \delta^p$ (between 0.4 °C and 1.3 °C, see Table 4).

2.2.5. Infection rate

The infection rate is the proportion of the copepodids that infect fish during a day and thus develop into the CH stage. It is farm- and cage-dependent, but does not depend on stage-age a, except that we assume that development may only happen for $a \geq 1$. This is modelled as

$$d^{CO}_{tfc} = \exp(\eta^{CO}_{tfc})/(1 + \sum_c \exp(\eta^{CO}_{tfc})), \tag{27}$$

where

$$\eta^{CO}_{tfc} = \delta^{CO}_{0fc} + \log(N^{SAL}_{tfc}) + \delta^{CO}_1(\log(W_{tfc}) - 0.55). \tag{28}$$

Here N^{SAL}_{tfc} and W_{tfc} are the number (in millions) and the average weight (in kg), respectively, of fish in cage c at farm f and time t, and 0.55 is roughly the mean of the natural logarithm of the weight of fish. With this formulation, $d^{CO}_{tf} = \sum_c d^{CO}_{tfc}$ will be the proportion of copepodids that infect fish in any cage during day t, and this will always be between 0 and 1. Furthermore, when the proportions or rates are small, the rate d^{CO}_{tfac} for each cage will approximately be proportional to the number of fish N^{SAL}_{tfc} in the cage and to $W^{\delta^{CO}_1}_{tfc}$.

The parameter δ^{CO}_{0fc} controls the magnitude of the infection rate conditioned on the number and weight of fish within a given cage. In our model δ^{CO}_{0fc} depends on cage and farm, reflecting that some farms or cages may be more exposed to infection than others due to for instance sea current conditions. This is handled by the following hierarchical structure:

$$\delta^{CO}_{0fc} \sim \mathrm{N}(\delta^{CO}_{0f}, (\sigma^{COdf})^2), \tag{29}$$

$$\delta^{CO}_{0f} \sim \mathrm{N}(\delta^{CO}_{0}, (\sigma^{COd})^2), \tag{30}$$

where δ^{CO}_{0f} is a farm-specific mean and δ^{CO}_0 an overall mean. Furthermore, σ^{COdf} reflects the variability between cages at the same farm, whereas σ^{COd} reflects the variability between farms.

2.2.6. Reproduction rate

The recruitment model, Eq. (1), includes the internal reproduction rate r_{tac}, which is modelled taking into account the following factors: Female adults extrude pairs of egg strings. They can extrude a new set of egg strings within 24 hour after the previous set was hatched, but hatching can take several days (Stien et al., 2005). The number of eggs per string may increase for each consecutive extrusion, which we approximate with stage-age. Finally, not all eggs are viable. In addition, we allow for density dependence in recruitment, due to potentially reduced probability of mate finding at low lice abundance, as suggested by Stormoen et al. (2013), Krkošek et al. (2012) and Groner et al. (2014).

The reproduction rate r_{tac} for internal recruitment at time t, stage-age a and cage c is thus modelled as

$$r_{tac} = \beta^r_0 \cdot (a+1)^{\beta^r_1} \cdot 1/(\delta^{Em}_t + 1) \cdot (1 - \exp(-\gamma^r \cdot A_{tc})). \tag{31}$$

The first term in Eq. (31), β^r_0, represents the number of viable eggs for the first extrusion. The next term, $(a+1)^{\beta^r_1}$ models how the number of viable eggs per extrusion increases by stage-age. The third term, $1/(\delta^{Em}_t + 1)$, represents the rate of pairs of egg strings produced per day, which is the inverse of average time between each egg extrusion, which further is approximately the median hatching time plus one day for developing new egg strings. The median hatching time is given by

$$\delta^{Em}_t = \delta^{Em10}(10/T_t)^{\delta^{Rp}}, \tag{32}$$

where T_t is the seawater temperature and δ^{Em10} and δ^{Rp} are parameters defined in Section 2.2.4. Finally, the term $(1 - \exp(-\gamma^r \cdot A_{tc}))$ allows for density dependent recruitment. Here, $A_{tc} = N^{AF}_{tc}/N^{SAL}_{tc}$ is the abundance of adult females in cage c at time t. A very large value of γ^r corresponds to a model without density dependent recruitment. Of the parameters involved in Eq. (31), we estimate δ^{m10E}, δ^{pR} and γ^r and fix β^r_0 and β^r_1 to 172.5 and 0.2, respectively (see Section 2.5 in the Supplementary material for a motivation of these values).

2.2.7. Reproduction rate for external recruitment

The reproduction rate r^{Ext}_t for external recruitment in Eq. (1) is similar to the internal one, but the female lice abundance A_{tc} in Eq. (31) is replaced by a weighted average of the counted abundance at neighbouring farms, A^{AFExt}_t (Section 2.2.9). Furthermore, we assume that all these female lice at neighbouring farms are at stage-age $a = 10$. The assumed stage-age of 10 is rather arbitrary, but the results are insensitive to this choice.

2.2.8. Modifying factor in the external recruitment

The modifying factor e^{Ext}_t for external recruitment is farm-specific, so we include the farm index f as well. At the log-scale, it varies over time around a farm-specific level according to the following AR(1) model:

$$e^{Ext}_{tf} = \exp(z^{Ext}_{tf}), \tag{33}$$

$$(z^{Ext}_{tf} - \mu^{Ext}_f) = \phi^{Ext} \cdot (z^{Ext}_{(t-1)f} - \mu^{Ext}_f) + \varepsilon^{Ext}_{tf}, \tag{34}$$

$$(\varepsilon^{Ext}_{tf}) \sim \mathrm{N}(0, (\sigma^{Extar})^2), \tag{35}$$

$$\mu^{Ext}_f \sim \mathrm{N}(\mu^{Ext}, (\sigma^{Ext})^2). \tag{36}$$

Here, μ^{Ext}_f is the farm-specific expected value on the log-scale, ϕ^{Ext} the autoregressive coefficient and $(\sigma^{Extar})^2$ the residual variance. Furthermore, μ^{Ext} is the overall expected value and $(\sigma^{Ext})^2$ the between-farm variance of μ^{Ext}_f.

2.2.9. Definitions of N^{AFExt}_{tf} and A^{AFExt}_{tf}

The weighted sum of adult females at neighbouring farms used in Eq. (2), denoted by N^{AFExt}_{tf} for farm f at time t, is given by

$$N^{AFExt}_{tf} = \sum_{f' \neq f} g(d_{ff'}) \hat{N}^{AF}_{tf'}, \tag{37}$$

where $g(\cdot)$ is a function decreasing by increasing distance given by

$$g(d) = \exp(-0.618 d^{0.568}). \quad (38)$$

This distance function is taken from Aldrin et al. (2013), and is based on a data-driven model for lice abundance estimated from more than eight years of data on all 1400 Norwegian salmon farms that were active in the data period.

We have also calculated a corresponding weighted average of the counted abundance of adult females at neighbouring farms, A_{tf}^{AFExt}, which replaces A_{tc} in Eq. (31) when the reproduction rate r_t^{Ext} for external recruitment is computed (Section 2.2.7). It is given by

$$A_{tf}^{AFExt} = \sum_{f' \neq f} g(d_{ff'}) \hat{A}_{tf'}^{AF} / \sum_{f' \neq f} g(d_{ff'}). \quad (39)$$

2.2.10. Cleaner fish model

Let S_{tc}^{clf} denote the number of cleaner fish stocked and N_{tc}^{clf} the total number of cleaner fish in cage c at time t. S_{tc}^{clf} is observed, whereas N_{tc}^{clf} is unknown and modelled as

$$N_{tc}^{clf} = N_{(t-1)c}^{clf}(1 - \kappa^{clf}) + S_{tc}^{clf}, \quad (40)$$

where κ^{clf} is the daily constant mortality rate of cleaner fish, common for all farms.

2.2.11. Data model and model fitting

In this subsection, we describe how the population model is related to the lice count data. Let Y_{tc}^{CG} be the number of lice in count group CG found on n_{tc} counted fish at time t and cage c, where the count groups are either chalimus (CH), adult females (AF) or other mobiles (OM, *i.e*, pre-adults and adult males). We assume that these follow a negative binomial distribution with mean $\mu_{tc}^{CG} = \mathrm{E}(Y_{tc}^{CG})$ and a heterogeneity or aggregation parameter $n_{tc}\rho^{CG}$, such that the variance of Y_{tc}^{CG} is $\mu_{tc}^{CG} + (\mu_{tc}^{CG})^2/(n_{tc}\rho^{CG})$. Deleting the superscript CG and subscript tc for a moment, the probability distribution of Y is

$$P(Y = y) = \frac{\Gamma(y + n\rho)}{y!\Gamma(n\rho)} \left(\frac{n\rho}{n\rho + \mu}\right)^{n\rho} \left(\frac{\mu}{n\rho + \mu}\right)^{y}. \quad (41)$$

We get the total likelihood for each count group by multiplying over all counts, cages and farms. We further assume independence between count groups and get the total likelihood by multiplying the contribution from each count group.

The expected numbers of the various Y_{tc}^{CG}'s are given from the population model as

$$\mathrm{E}(Y_{tc}^{CH}) = n_{tc} \cdot p_{tc}^{CHcount} \cdot N_{tc}^{CH}/N_{tc}^{SAL}, \quad (42)$$

$$\mathrm{E}(Y_{tc}^{AF}) = n_{tc} \cdot N_{tc}^{AF}/N_{tc}^{SAL}, \quad (43)$$

$$\mathrm{E}(Y_{tc}^{OM}) = n_{tc} \cdot (N_{tc}^{PA} + N_{tc}^{AM})/N_{tc}^{SAL}, \quad (44)$$

where the role of the factor $p_{tc}^{CHcount}$ is to adjust for under-reporting of CH lice, since they are very small and difficult to count, especially on large fish. We assume that this factor is farm-specific (for instance, the staff at some farms may be more trained or motivated than staff at other farms), and we introduce from now on the index f for farm. Then, the model for $p_{tfc}^{CHcount}$ is

$$p_{tfc}^{CHcount} = \exp(\eta_{tfc}^{CHcount})/(1 + \exp(\eta_{tfc}^{CHcount})), \quad (45)$$

where

$$\eta_{tfc}^{CHcount} = \beta_{0f}^{CHcount} + \beta_1^{CHcount}(W_{ftc} - 0.1), \quad (46)$$

where W_{ftc} as before is the mean weight of fish in cage c at farm f at time t. The constant 0.1 is chosen to make it easier to specify prior distributions for $\beta_{0f}^{CHcount}$ and $\beta_1^{CHcount}$. Here, $\beta_1^{CHcount}$ is common for all farms, but $\beta_{0f}^{CHcount}$ varies between farms according to the following hierarchical model:

$$\beta_{0f}^{CHcount} \sim \mathrm{N}(\beta_0^{CHcount}, (\sigma^{CHcount})^2). \quad (47)$$

The model was estimated from the data including 32 farms, except the last months (3–11) of data for five of the farms that were used for evaluating conditional predictions. We used a Bayesian estimation approach, combining the prior distributions and the data likelihood into a joint posterior distribution for all model parameters. Many of the prior distributions used were informative, based on results from laboratory experiments, *e.g.* those reported in Stien et al. (2005). For some priors, however, we use more vague settings (see Section 2 in the Supplementary material for details). The model was fitted to data using Markov Chain Monte Carlo (MCMC) simulations (Gilks et al., 1996). First, several initial chains were run to identify a rough range for plausible parameter values. Then four independent chains were started from slightly different starting values within this range. The first 25000 iterations were used as burn-in to establish convergence, and the posterior distributions were calculated by combining 100 thinned samples from the last 14 000 iterations from each of the chains. See Section 4 in the Supplementary Material for more details on the MCMC algorithm.

3. Results and discussion

3.1. Fitted and predicted values

The model generated expected values that fitted the observed infection levels of chalimus (CH), adult female (AF) and other mobile stages (OM) well (Table 3, Fig. 4, and Section 3 in the Supplementary material with results for seven other farms). Predictions for the periods not used for fitting the model were also consistent with the data with respect to the timing of population growth of adult female (AF) and other mobile stages (OM) (Fig. 4, and Figure 1–5 in Section 3 in the Supplementary material), even though the prediction errors naturally tend to be larger in the prediction periods than in the estimation periods (Table 3). These results support the notion that there is a substantial deterministic component in the transmission pathways and population dynamics of salmon lice in fish farms. This emphasises a potential for utilising the massive body of data gathered by the salmon farming industry to gain control of salmon louse infections in farms, which is a prerequisite for sustainable growth in Norwegian salmon farming (Anonymous, 2015).

For periods with elevated predicted population sizes, however, abundances of salmon lice were sometimes over-estimated (e.g. AF abundance in August 2013, Figure 3, Section 3 in Supplementary material) and sometimes under-estimated (e.g. AF and OM in first part of September 2013, Fig. 4). These large deviations in some predictions are likely to reflect (1) that there are predictor variables that have not been included in the present model (e.g. salinity), (2) substantial uncertainty in some predictor variables like the abundance of cleaner fish in the cages, and (3) that stochasticity, in particular with respect to the infection process, limits our ability to make precise predictions. Accordingly, also the credible intervals for the predictions were wide when elevated abundances of infection were predicted (e.g. Fig. 4). The credible intervals were well calibrated in the estimation periods for all three counting categories of lice. This was also the case for AF and OM in the first month of each prediction period, in that the actual coverage was close to the nominal 95% (Table 3). However, for predictions more than one month ahead, the actual coverage was slightly too low, indicating that the credible intervals for long term forecasts were slightly too narrow.

Table 3
Mean absolute percentage error (MAPE) and percentage coverage of the 95% credible intervals for lice counts in the CH, OM and AF categories, over (i) the estimation periods, (ii) the first 30 days of the five prediction periods and (iii) the remaining parts of the prediction periods. MAPE is here defined as $\sum_i |Y_i - \hat{Y}_i| / \sum_i (\hat{Y})_i$ where Y_i is an observed lice abundance and $\hat{Y}_i$ is the corresponding fitted or predicted abundance, and the sum is taken over all farms, cages and times with observations. The percentage coverage is the percentage of observations at or within the outer intervals (the outer limits of the grey areas) in Fig. 4 and similar figures.

Stage	MAPE			Coverage (%)		
	Estimation periods	First 30 days of each prediction periods	Remaining parts of prediction periods	Estimation periods	First 30 days of each prediction periods	Remaining parts of prediction periods
CH	83	262	199	97	84	87
OM	47	76	96	96	94	86
AF	76	74	104	98	97	87

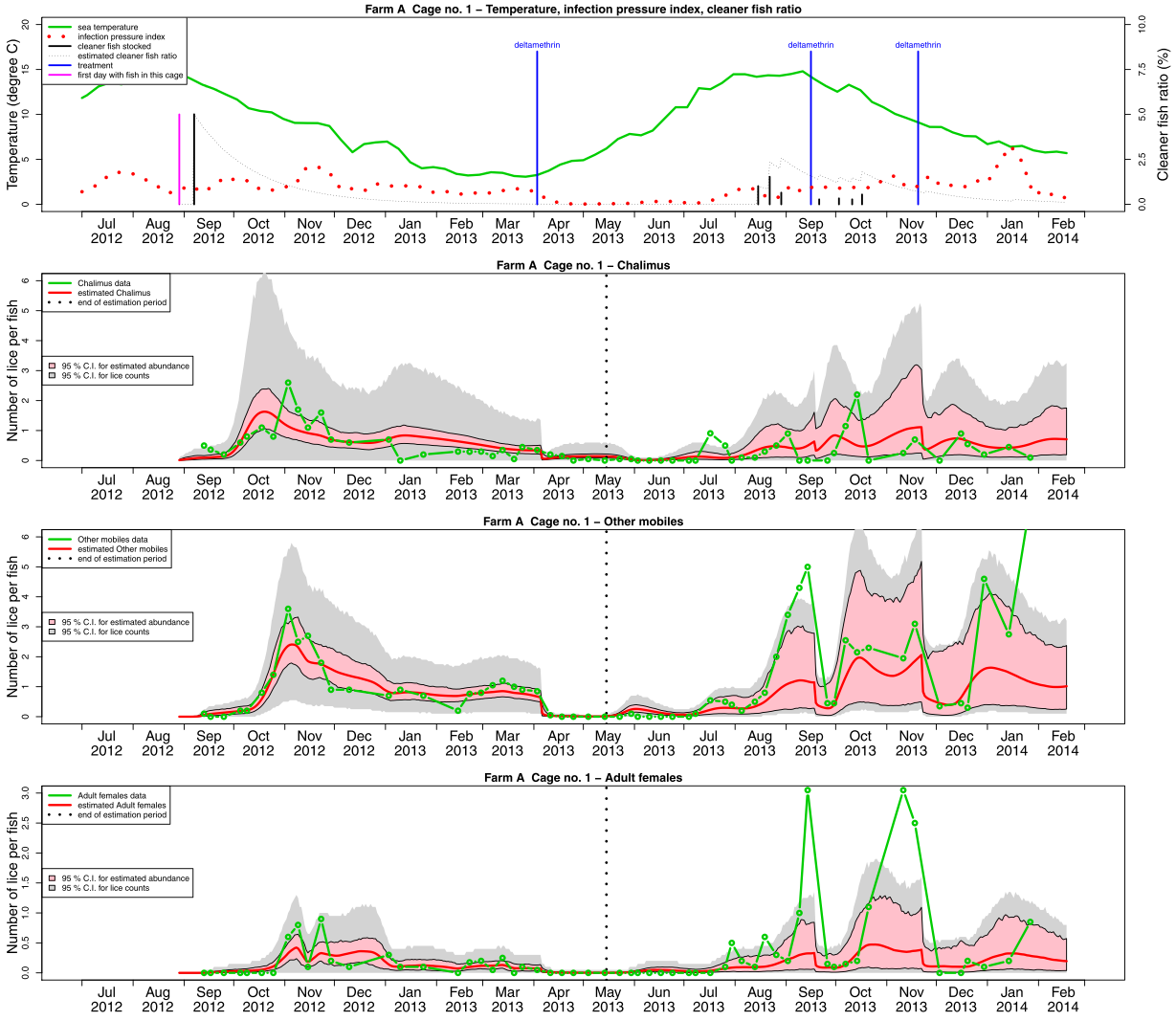

Fig. 4. Fitted (until 15 May 2013) and predicted (from 16 May 2013) values for the lice and the cleaner fish populations. Symbols are as given in Fig. 2 with the following additions: Upper panel: Fitted (posterior mean) cleaner fish ratio (dotted black curve). Three lower panels: (i) fitted values (red curves to the left of the vertical black dotted line), (ii) predictions conditioned on known temperature, external infection pressure index and number and weight of salmon (red curves to the right of the vertical black dotted line), (iii) corresponding 95% credible interval for the lice population (pink area) and iv) additional 95% credible interval for lice counts (grey area), *i.e.* including the randomness in the negative binomial distribution for lice counts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

There was substantial underreporting of the number of lice at the CH stage. Depending on the size of the fish, the model estimates suggested that on average only 9% to 19% of the CH lice were counted (Fig. 5). In addition, there was substantial between farm variability in this counting error (Fig. 5). The relationship between observed abundances of lice at the CH stage and predicted values was poorer than for the other stages (OM and AF), even when the underreporting was accounting for (the grey "counting error" area for CH in Fig. 4 is wide and includes zero). This can be quantified by the elevated prediction error, especially in the prediction periods (Table 3) and by the aggregation parameter ρ, which was 50–75% lower than for OM and AF (Table 4). This indicates that the information content in the counts of CH stage lice is limited.

Table 4
Posterior means with 95% credible intervals of parameters in the static parts of the model.

Part of model	Stage	Parameter interpretation	Parameter symbol	Section	Posterior mean	95% C.I. lower	95% C.I. upper
Natural mortality	R,CO	Mortality rate	λ^{RCOnat}	2.2.3	0.302	0.287	0.315
Mortality cl.fish	PA, A	Regression coeff.	λ^{clf}	2.2.3	0.823	0.620	1.046
Development	Egg	Median at 10 °C	δ^{Em10}	2.2.4	4.866	4.432	5.728
Development	Nauplii	Median at 10 °C	δ^{Nm10}	2.2.4	3.948	3.088	4.422
Development	R	Shape parameter	δ^{Rs}	2.2.4	18.869	15.915	19.975
Development	R	Power parameter	δ^{Rp}	2.2.4	0.401	0.400	0.405
Development	CH	Median at 10 °C	δ^{CHm10}	2.2.4	18.934	18.314	19.487
Development	CH	Shape parameter	δ^{CHs}	2.2.4	7.945	6.999	9.022
Development	CH	Power parameter	δ^{CHp}	2.2.4	1.305	1.257	1.354
Development	PA	Median at 10 °C	δ^{PAm10}	2.2.4	10.742	10.174	11.322
Development	PA	Shape parameter	δ^{PAs}	2.2.4	1.643	1.430	1.898
Development	PA	Power parameter	δ^{PAp}	2.2.4	0.866	0.784	0.950
Development	CO	Expectation	δ_0^{CO}	2.2.5	−2.576	−2.958	−2.207
Development	CO	Regression coeff.	δ_1^{CO}	2.2.5	0.082	0.040	0.120
Development	CO	Variance within farm	$(\sigma^{COdf})^2$	2.2.5	0.035	0.026	0.044
Development	CO	Variance between farms	$(\sigma^{COd})^2$	2.2.5	0.357	0.206	0.606
Reproduction	AF to R	Basic number of eggs	β_0^r	2.2.6	172.500	Fixed	
Reproduction	AF to R	Age dependence	β_1^r	2.2.6	0.200	Fixed	
Reproduction	AF to R	Density dependence	γ^r	2.2.6	493	482	498
Cleaner fish model		Mortality rate	κ^{clf}	2.2.10	0.028	0.022	0.034
Data model	CH	Aggregation parameter	ρ^{CH}	2.2.11	0.051	0.048	0.054
Data model	OM = PA + AM	Aggregation parameter	ρ^{OM}	2.2.11	0.194	0.181	0.207
Data model	AF	Aggregation parameter	ρ^{AF}	2.2.11	0.120	0.109	0.132
Data model	CH	Expectation	$\beta_0^{CHcount}$	2.2.11	−1.566	−1.819	−1.339
Data model	CH	Variance	$(\sigma^{CHcount})^2$	2.2.11	0.411	0.238	0.682
Data model	CH	Regression coeff.	$\beta_1^{CHcount}$	2.2.11	−0.164	−0.188	−0.136

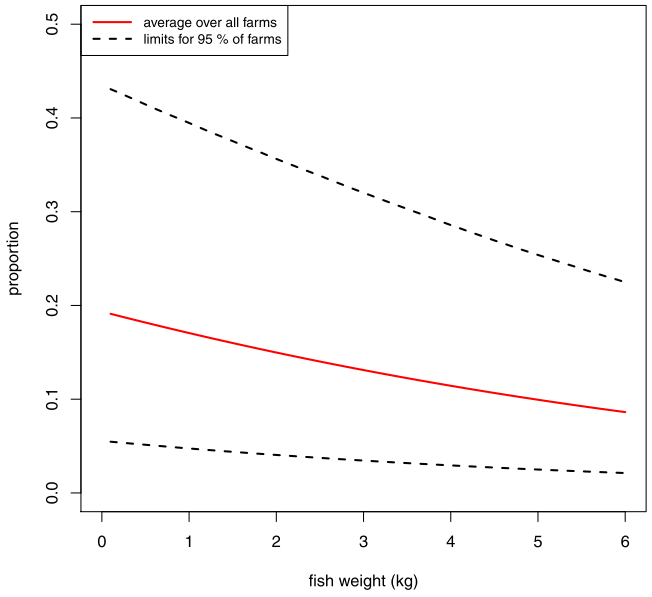

Fig. 5. Posterior mean of the proportion of true CH lice counted as a function of fish weight, on average over all farms (solid line) and 95% limits for the between-farm variability (dashed lines).

3.2. Parameter estimates

Posterior mean estimates and credible intervals for parameters in the model are given in Tables 4 and 5. Note that for those parts of the model where we have no data, covariation between parameters in the model may lead to potential bias, i.e. high estimates of one parameter may be compensated for by an associated change in the value of another parameter. An example of this is that mortality and development rates for the R and CO stages and the reproduction rate are related, but without relevant observations to tease them apart. For instance, if the reproduction rate is overestimated, this can be compensated either by increasing the mortality in the R and CO stages (λ^{RCOnat}) or by reducing the infection rate (δ_0^{CO}, which controls the development from CO to CH). Also, for a given reproduction rate, an increase in λ^{RCOnat} can be compensated by an increase in δ_0^{CO}, which is confirmed by a correlation of 0.70 between the estimates of these parameters. Another obvious example is the development time parameters δ^{Em10} and δ^{Nm10} for the egg and nauplii stages, which were negatively correlated (-0.98) and can not separated by the data, only by the priors.

When compared to previously published estimates on stage-specific mortality, there are some notable differences (Table 6). The estimate of the mean mortality rate common to the R and CO stages (λ^{RCOnat}) was higher than the previous estimate (Table 6). However, note that in our model, this quantity also accounts for nauplii and copepodids that drift away from the farm, in addition to the pure natural mortality. For the CH, PA and A stages, the mortality rates vary over time, but we calculated their overall expectations (averages in the long run) by simulation. The overall expectation for the mortality rates of the CH and PA stages tended to be lower than previous estimates, while the estimate for the adult stage was within the range of previous studies (Table 6). These estimates must be interpreted with caution since the model assumes equal development rates between genders and, furthermore, since adult males are pooled with the PA stages in the observational data. More detailed figures on mortality, including parameter uncertainties, are presented in Figures 9–13 in the Supplementary material.

We also used simulations to find the estimated median development times, since the δ^{m10} parameters given in Table 4 have exact interpretations only in the continuous Weibull distribution. Estimated median development times at 10 °C were similar to mean and minimum estimates from previous studies (Table 7), although with a slightly higher estimate at the CH stage and fairly low estimate for the PA stage. Panel (a) in Fig. 6 show how the estimated development rates increase by stage-age at a temperature of 10 °C. These curves differ in principle from the development rates used in all population models mentioned in Section 2.2.1, since all these models use step functions with a development rate of 0 until a minimum development time and then a constant rate afterwards. For the R and CH stages, however, the estimated *cumulative* proportion of lice developed to the next stage (panel b) in Fig. 6) resemble these step functions. For the PA stage, however, the estimated devel-

Table 5
Posterior means with 95% credible intervals of parameters in the time-varying parts of the model.

Part of model	Stage	Parameter interpretation	Parameter symbol	Section	Posterior mean	95% C.I. lower	95% C.I. upper
Natural mortality	CH	Expectation in AR(1)	λ_0^{CHnat}	2.2.3	−6.953	−7.076	−6.842
Natural mortality	CH	Coefficient in AR(1)	ϕ^{CHnat}	2.2.3	0.012	0.001	0.028
Natural mortality	CH	Variance in AR(1)	$(\sigma^{CHnat})^2$	2.2.3	0.018	0.010	0.026
Natural mortality	PA	Expectation in AR(1)	λ_0^{PAnat}	2.2.3	−5.011	−5.234	−4.725
Natural mortality	PA	Coefficient in AR(1)	ϕ^{PAnat}	2.2.3	0.036	0.001	0.092
Natural mortality	PA	Variance in AR(1)	$(\sigma^{PAnat})^2$	2.2.3	0.114	0.073	0.171
Natural mortality	A	Expectation in AR(1)	λ_0^{Anat}	2.2.3	−2.408	−2.487	−2.342
Natural mortality	A	Coefficient in AR(1)	ϕ^{Anat}	2.2.3	0.693	0.676	0.707
Natural mortality	A	Variance in AR(1)	$(\sigma^{Anat})^2$	2.2.3	0.728	0.686	0.772
Mortality ch.tr.	CH, PA, A	Expectation, deltamethrin	λ^{DMcht}	2.2.3	2.434	1.574	3.354
Mortality ch.tr.	CH, PA, A	Variance, deltamethrin	$(\sigma^{DMcht})^2$	2.2.3	9.084	5.905	12.565
Mortality ch.tr.	PA, A	Expectation, azamethiphos	λ^{AZcht}	2.2.3	0.158	−0.755	1.106
Mortality ch.tr.	PA, A	Variance, azamethiphos	$(\sigma^{AZcht})^2$	2.2.3	$(\sigma^{DMcht})^2$		
Mortality ch.tr.	PA, A	Expectation, H_2O_2	λ^{HPcht}	2.2.3	4.034	3.119	5.071
Mortality ch.tr.	PA, A	Variance, H_2O_2	$(\sigma^{HPcht})^2$	2.2.3	$(\sigma^{DMcht})^2$		
Mortality ch.tr.	CH, PA, A	Expectation, emamectin	λ^{EMcht}	2.2.3	−4.786	−5.589	−4.251
Mortality ch.tr.	CH, PA, A	Variance, emamectin	$(\sigma^{EMcht})^2$	2.2.3	1.912	0.908	4.004
Mortality ch.tr.	CH, PA	Expectation, diflubenzuron	λ^{Dlcht}	2.2.3	−8.399	−12.702	−5.099
Mortality ch.tr.	CH, PA	Variance, diflubenzuron	$(\sigma^{Dlcht})^2$	2.2.3	$(\sigma^{EMcht})^2$		
External recr.	AF to R	Expectation in AR(1)	μ^{Ext}	2.2.8	0.301	0.177	0.410
External recr.	AF to R	Coefficient in AR(1)	ϕ^{Ext}	2.2.8	0.934	0.925	0.945
External recr.	AF to R	Variance in AR(1)	$(\sigma^{Extar})^2$	2.2.8	0.163	0.152	0.174
External recr.	AF to R	Variance	$(\sigma^{Ext})^2$	2.2.8	0.007	0.001	0.028

Table 6
Posterior means with 95% credible intervals of daily mortality from this study together with point estimates or ranges from previous studies.

Stage	Point estimate or range	95% C.I.	Sex	Comment	Reference
Nauplii	0.30	0.29–0.32		For R = eggs + nauplii, includes drifting away	This paper
	0.17			"Plausible value"	Stien et al. (2005)
CO	0.30	0.29–0.32		The same as for R	This paper
	0.22			"Plausible values"	Stien et al. (2005)
CH	0.0010	0.0008–0.0011			This paper
	0.002–0.01			"Plausible values"	Stien et al. (2005)
	0.0006–0.020			Outer interval limits from 4 reported studies	Stien et al. (2005)
	0.0002–0.026			Range over 7 trials on juvenile Pacific salmon	Krkošek et al. (2009)
PA	0.0071	0.0055–0.0095			This paper
	0.02–0.18		Males	"Plausible values"	Stien et al. (2005)
	0.002–0.21		Males	Outer interval limits from 4 reported studies	Stien et al. (2005)
	0.03–0.07		Females	"Plausible values"	Stien et al. (2005)
	0.011–0.102		Females	Outer interval limits from 4 reported studies	Stien et al. (2005)
	0.14–0.34			PA + A combined,range over 7 trials on juvenile Pacific salmon	Krkošek et al. (2009)
A	0.12	0.11–0.13			This paper
	0.03–0.06		Males	"Plausible values"	Stien et al. (2005)
	0.008–0.26		Males	Outer interval limits from 4 reported studies	Stien et al. (2005)
	0.02–0.04		Females	"Plausible values"	Stien et al. (2005)
	0.003–0.70		Females	Outer interval limits from 3 reported studies	Stien et al. (2005)
	0.14–0.34			PA + A combined,range over 7 trials on juvenile Pacific salmon	Krkošek et al. (2009)

Table 7
Posterior means with 95% credible intervals of development times (in days) at 10 °C from this study together with point estimates or ranges from previous studies.

Stage	Point estimate or range	95% C.I.	Estimate of what	Sex	Comment	Reference
Eggs	5.0	4.5–5.8	Median			This paper
	8.8		Minimum		From their Eq. (8) and Table 3	Stien et al. (2005)
	4.6		Mean			Samsing et al. (2016)
Nauplii	4.0	3.2–4.5	Median			This paper
	3.6		Minimum		From their Eq. (8) and Table 3	Stien et al. (2005)
	3.8		Mean			Samsing et al. (2016)
CH	18.8	18.0–19.0	Median			This paper
	15.4		Minimum	Males	From their Eq. (8) and Table 3	Stien et al. (2005)
	16.5		Minimum	Females	From their Eq. (8) and Table 3	Stien et al. (2005)
	11–13		Range	Males		Eichner et al. (2015)
	13–15		Range	Females		Eichner et al. (2015)
	11–14		Minimum		5 trials on juvenile Pacific salmon at 9–11 °C	Krkošek et al. (2009)
PA	10.5	10.0–11.0	Median			This paper
	10.4		Minimum	Males	Calculated as difference of their time from CH to A and from CH to PA	Stien et al. (2005)
	15.4		Minimum	Females	Calculated as difference of their time from CH to A and from CH to PA	Stien et al. (2005)

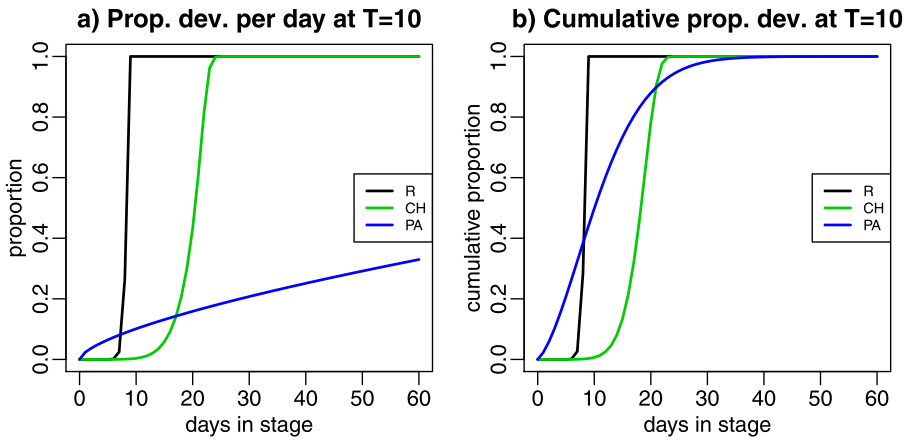

Fig. 6. Posterior means of daily development (a) and cumulative proportion developed to next stage (b) at temperature 10 °C for stages R, CH and PA. The curves are the posterior means of the development rates, conditioned on each stage-age, and may therefore be slightly different from curves based the posterior means of the parameter values plugged into Eq. (24).

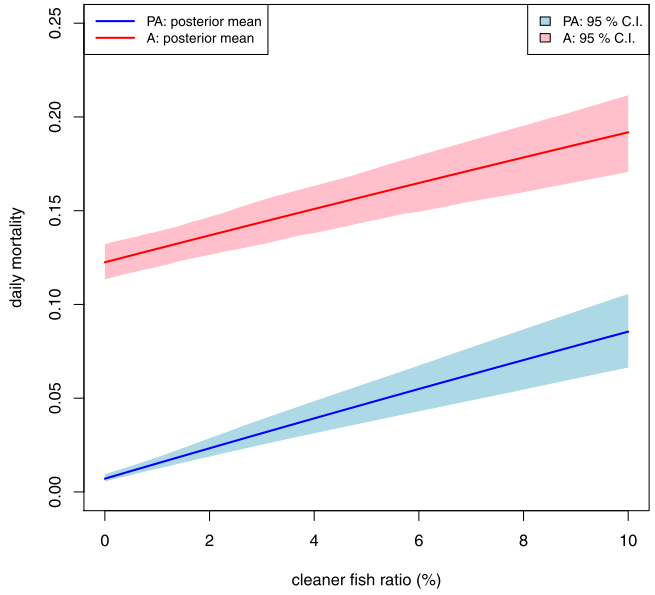

Fig. 7. Posterior mean and 95% credible intervals for the total mortality for the PA and A stages when the cleaner fish ratio is increased from 0 (only natural mortality) to 10%.

opment rate is fundamentally different from those used in other models, since the estimated development rate for PA is non-zero already after one day. This implies that some lice in the PA stage may develop to the A stage very quickly in the present model.

Note that these curves ignore mortality, and the cumulative mortalities may be high for stage-ages where the development rates still are quite low, especially at low temperatures. More detailed figures on development times, including parameter uncertainties and for different temperatures (5, 10 and 15 °C), are presented in Figures 14–16 in the Supplementary material.

We obtain an estimate of the daily cleaner fish mortality of 0.028 (C.I. 0.022–0.034, Table 4). This suggests that the cleaner fish population is reduced to its half about 1 month after release. The model confirms that there is increased mortality of lice associated with the use of cleaner fish (Fig. 7 and Table 4). With a 10% cleaner fish to salmon ratio, the estimated daily lice mortality due to the use of cleaner fish, is 0.079 (C.I. 0.060–0.099). This implies a reduction in the life expectancy for adult lice from 8.2 to 5.2 days with an increase in cleaner fish ratio from 0 to 10%, and a decrease in the life expectancy for pre-adult lice (PA) going from 141 days to 12 days for the same change in cleaner fish ratio. Therefore, in particular for PA lice, the use of cleaner fish is estimated to have a substantial effect on lice survival. However, in the present data, the estimated cleaner fish ratio seldom amounted to more than 5%. This seems to be too low to avoid additional treatments, since medical treatments were applied in almost all cages in the data set. Hence, given the estimated rates of decay of the populations, cleaner fish louse control demands large numbers of fish. Clearly, this raises many sustainability issues, for example with regard to the welfare of cleaner fish in salmon farms and the resilience of wild cleaner fish populations to fishery (Halvorsen et al., 2017; Skiftesvik et al., 2014).

The effects of the various chemotherapeutic treatments are difficult to compare since the assumed duration of the effects varies between treatments and by temperature, and because they affect different stages of lice. Furthermore, since lice develop resistance towards such treatments (Aaen et al., 2015; Jansen et al., 2016), we expect that the effect will decrease over time. Nevertheless, the estimated expected cumulative mortality of lice (found by simulation) in the PA or A stages due to bath treatments (*i.e.* non-feed) ten days post treatment at 10 °C, were high for hydrogen peroxide (0.99, C.I. 0.97–1.00), and deltamethrin/cypermethrin (0.94, C.I. 0.89–0.98) and somewhat lower for azamethiphos (0.75, C.I. 0.64–0.86). The first two of these are similar to what others have reported for non-resistant lice populations, being 99% for hydrogen peroxide (Groner et al., 2013) and 95% for deltamethrin and cypermethrin (Revie et al., 2005). The estimated effects of all treatment types are further illustrated in Figures 9–13 in the Supplementary material.

Both external and internal recruitment varied substantially over time, and one of the two may dominate the other in certain periods. The estimated proportion of internal recruitment for each farm averaged over the whole production cycle varies from about 4 to 73%. On average over all farms, this proportion is 25% (C.I. 23–26%), with a median of 19. When the abundance of adult female lice was low in farms, naturally internal recruitment tended to be low. Internal recruitment was, however, estimated to be very important in farms with a high abundance of adult females, especially for farms without any neighbouring farms (Fig. 8). The present proportion of internal recruitment is substantially below the estimates presented by Aldrin et al. (2013). This model, however, did not tease apart actual new recruitment and the contribution to lice numbers from lice surviving over monthly time increments. Furthermore, it was fitted to farm level data originating from different reporting requirements. Therefore, we have more confidence in the present

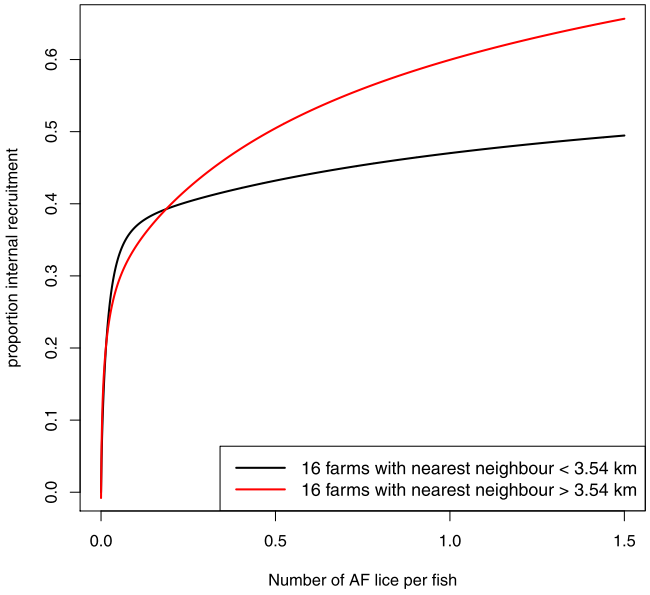

Fig. 8. Posterior mean of the average proportion of recruitment that has an internal source as a function of abundance of adult females, averaged over all time points and for (i) the 16 farms with at least one neighbour within 3.54 km (black curve) and (ii) the 16 farms with more than 3.54 km to its nearest neighbour (red curve). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

results, which are also in accordance with results in Ådlandsvik (2015). This report presented a median of 18% for the proportion of internal recruitment, based on a simulation of the spread of lice larvae between 591 farms using a hydrodynamic model accounting for sea currents (Johnsen et al., 2014). The importance of internal versus external recruitment of lice has been under considerable debate in the salmon farming industry, probably since it has direct bearings on the motivation for controlling abundances of adult female lice at the farm level.

Concerning the reproduction rate, we notice that the estimate of γ_0^r is 493 (C.I 482–498). In practice, this means that there is no evidence of density dependent reproductive rates (Allee effect) in the estimated model, contrary to the hypothesis suggested by Stormoen et al. (2013), Krkošek et al. (2012) and Groner et al. (2014). One reason for this discrepancy is that recruitment from external sources is not considered in the aforementioned models. In our model, internal recruitment is on average of less importance than external recruitment at low abundances of AF lice. Reduced reproductive rates at low abundances of infection are therefore likely to be masked by external recruitment.

The role of the modifying factors e_t^{Ext} in Eq. (1) is to adjust the preliminary estimates of external recruits, which are based on distances to neighbouring farms, and then indirectly taking into account local hydrodynamics. Averaging the estimates of these factors over time within each farm gives a farm-wise index for how exposed each farm is to external infection. The largest farm-wise factor was 61% higher than the smallest one, indicating that there are local variations that cannot be explained by distances to the neighbouring farms alone. Furthermore, the cage-specific parameters δ_{0fc}^{CO} from Eq. (28) are related both to local hydrodynamics and how cages are located relative to each other, and $\exp(\delta_{0fc}^{CO})$ can serve as a cage-wise exposure index. The most exposed cage within a farm was between 17% and 203% more exposed than the least exposed cage in the same farm, and 48% more exposed on average over all farms.

As pointed out previously, count data are missing completely for the R stage of the model, aggregated for the PA and AM stages, and biased towards low numbers for the CH stage. Therefore, some of the model parameters cannot be estimated from the data, and to overcome this we have chosen to set some of the parameters to fixed values. The most important is probably that we have set the proportion of females among the adults to 50%, which may influence estimates of other parameters. In addition, we have simplified the model and assumed that the development times from PA to A and adult mortality are the same for both sexes. These simplifications contradict observations of somewhat shorter development times for AM than AF lice (Stien et al., 2005; Eichner et al., 2015). We could have chosen other ways to handle this, but it would anyways have been necessary with strong prior assumptions to be able to estimate gender differences in the model.

In total, the model contains 44 parameters that are common for all farms and cages and that are estimated from the data. A few of these are confounded, as pointed out in the beginning of this section, but most of them are estimated with credible intervals much tighter than the corresponding intervals given by the prior distributions. The large number of parameters on different hierarchical levels in the model gives a possibility for overfitting. However, the size of prediction errors in the prediction intervals were reasonable compared to those in the estimation periods, at least for the AF and OM categories, and did not indicate dramatic overfitting (Table 3). A more informal indication of a balanced fit was given by the farmers at four of the farms: The cage-wise exposure indexes were shown to these farmers, and they concluded that the estimates were consistent with their own experience on exposure differences between cages.

4. Conclusion

The presented process model for the population dynamics of salmon lice in aquaculture farm systems combine (1) a model for the main stage structure of salmon lice; (2) model terms that describe internal and external recruitment of salmon lice to the farms; (3) models for the impact of management strategies adopted to control salmon lice infection levels; and (4) allows for stochasticity in aspects of these processes. In addition, we describe a model for the link between the process model and data collected on a routine basis in modern aquaculture. This allows the process model to be fitted to data and updated when new production and salmon lice data from fish farms become available. The model produces estimates of lice abundances in fish farms that correspond well with observed infection levels. Furthermore, it allows reasonable predictions of pre-adult and adult sea lice infection levels to be made for several weeks into the future (Fig. 4). We note, however, that the observational data on the number of lice at the chalimus stage vary substantially, and is underestimated to varying degrees between fish farms. This suggests that the data collected on the abundance of lice at the chalimus stages contain limited information in terms of real infection levels in the farms. Nevertheless, observation of chalimus stage larvae may give salmon farmers early warnings of pre-adult and adult lice abundances to come, even though their quantification is uncertain.

The model generates stage specific estimates of development and mortality rates based on real production data. For the planktonic stages, the estimates of mortality rates were high relative to earlier field and laboratory studies, but these are not directly comparable, since our definition of mortality include lice that drift away from the farm. For the chalimus and pre-adult stages, the mortality rates were low compared to previous studies, while it was rather high for the adult stage. One reason could be the poor data support for the early life stages, since we have no data on the planktonic stages and poor data on the chalimus stages. The poor support by data makes it difficult to separate survival-estimates at these

stages from other parameters in the model, and this problem may also propagate to the pre-adult stage. The median development times for all stages seem reasonable. However, we note that the distribution of development times from the pre-adult to the adult stages include the possibility of very short development times. It is unclear why this happens, but one possible explanation could be movement of adult lice between cages within a fish farm. Another explanation could be that some of the non-gravid female adults have been mis-classified as pre-adults in the lice count data.

The model is well suited to quantify the effect of different treatments on lice infection levels. This aspect may be useful in monitoring development of resistance to treatments in sea lice populations. Furthermore, to our knowledge this is a first data-based quantification of the effect of using cleaner fish to manage lice levels in full scale salmon farms.

The model allows both internal and external infection processes to be quantified. When there are no adult female lice in a fish farm, all new infections will necessarily come from external sources. When reproducing adult female lice are present, however, they may contribute substantially to infection pressure at the farm. Their impact on local recruitment will depend both on their numbers and their reproductive rate, but also on the external infection pressure.

The emphasis on salmon louse control in salmon farming has been intensified in later years, resulting in the implementation of a variety of innovative control methods. Some of these methods aim to reduce salmon louse recruitment from external sources, whereas other methods aim at increasing the mortality of parasitic stages of the lice, such as the use of cleaner fish. To tease out actual effects of such control efforts at farm levels, or more so for farms in a production area, is not a simple task, given the interconnected production-system of fish farms and planktonic spread of this parasite. The present model for the population dynamics of salmon lice in aquaculture farm systems may resolve the complexity of processes to a degree that will improve evaluations of effects of various control efforts at farm levels. We believe that using the model as “mathematical laboratory” to explore effects of different salmon louse control strategies in interconnected farms in production areas has a large potential for rationalising area-wise control strategies in an informed manner. Models for scenario simulations have previously been used to investigate effects of various measures to control foot-and-mouth disease (Tildesley et al., 2012) or bovine tuberculosis Brooks-Pollock et al. (2014) in cattle, and lately also for controlling the spread of pancreas disease in salmon farming (Pettersen et al., 2016).

Acknowledgements

This work was funded by the Norwegian Seafood Research Fund through the project FHF 900970 “Populasjonsmodell for lakselus på merd og lokalitetsnivå”. We thank the fish farming companies Marine Harvest, Salmar and Måsøval for supplying us with their detailed production and lice count data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2017.05.019.

References

Aaen, S., Helgesen, K., Bakke, M., Kaur, K., Horsberg, T., 2015. Drug resistance in sea lice: a threat to salmonid aquaculture. Trends Parasitol. 31, 72–81.
Aalen, O., Borgan Ø, Gjessing, H., 2008. Survival and Event History Analysis – A Process Point of View. Springer, New York.
Ådlandsvik, B., 2015. Forslag til produksjonsområder (in Norwegian). Tech. rep. Institute of Marine Research, Bergen, Norway, report from IMR no. 20-2015.
Aldrin, M., Huseby, R., Jansen, P., 2015. Space-time modelling of the spread of pancreas disease (PD) within and between Norwegian marine salmonid farms. Prev. Vet. Med. 121, 132–141.
Aldrin, M., Lyngstad, T., Kristoffersen, A., Storvik, B., Borgan Ø, Jansen, P., 2011. Modelling the spread of infectious salmon anaemia (ISA) among salmon farms based on seaway distances between farms and genetic relationships between ISA virus isolates. J. R. Soc. Interface 8, 1346–1356.
Aldrin, M., Storvik, B., Frigessi, A., Viljugrein, H., Jansen, P., 2010. A stochastic model for the assessment of the transmission pathways of heart and skeleton muscle inflammation, pancreas disease and infectious salmon anaemia in marine fish farms in Norway. Prev. Vet. Med. 93, 51–61.
Aldrin, M., Storvik, B., Kristoffersen, A., Jansen, P., 2013. Space-time modelling of the spread of salmon lice between and within Norwegian marine salmon farms. PLoS ONE 8, e64039.
Anonymous, 2015. Forutsigbar og miljømessig bærekraftig vekst i norsk lakse- og ørretoppdrett (in Norwegian). Tech. rep. Ministry of Trade, Industry and Fisheries, Oslo, Norway, meld. St 16 (2014-2015).
Brooks-Pollock, E., Roberts, G., Keeling, M., 2014. A dynamic model of bovine tuberculosis spread and control in Great Britain. Nature 511, 228–231.
Caswell, H., 2001. Matrix Population Models: Construction, Analysis and Interpretation, 2nd ed. Sinauer Associates Inc., Sunderland.
Diggle, P., 2006. Spatio-temporal point processes, partial likelihood, foot and mouth disease. Stat. Methods Med. Res. 15, 325–336.
Eichner, C., Hamre, L., Nilsen, F., 2015. Instar growth and molt increments in *Lepeophtheirus salmonis* (Copepoda: Caligidae) chalimus larvae. Parasitol. Int. 64, 86–96.
Forseth, T., Barlaup, B., Finstad, B., Fiske, P., Gjøsæter, H., Falkegård, M., Hindar, A., Mo, T., Rikardsen, A., Thorstad, E., Vøllestad, L., Wennevik, V., 2017. The major threats to Atlantic salmon in Norway. ICES J. Mar. Sci. fsx020.
Frazer, L., Morton, A., Krkošek, M., 2012. Critical thresholds in sea lice epidemics: evidence, sensitivity and subcritical estimation. Proc. Biol. Sci. 279, 1950–1958.
Gettinby, G., Robbins, C., Lees, F., Heuch, P., Finstad, B., Malkenes, R., Revie, C., 2011. Use of a mathematical model to describe the epidemiology of *Lepeophtheirus salmonis* on farmed Atlantic salmon *Salmo Salar* in the Hardangerford, Norway. Aquaculture 320, 164–170.
Gilks, W., Richardson, S., Spiegelhalter De, 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall/CRC, Boca Raton.
Grenfell, B.T., Smith, G., Anderson, R.M., 1987. A mathematical model of the population biology of *Ostertagia ostertagi* in calves and yearlings. Parasitology 95, 389–406.
Groner, M., Cox, R., Gettinby, G., Revie, C., 2013. Use of agent-based modelling to predict benefits of cleaner fish in controlling sea lice, *Lepeophtheirus salmonis*, infestations on farmed Atlantic salmon, *Salmo Salar* L. J. Fish Dis. 36, 195–208.
Groner, M., Gettinby, G., Stormoen, M., Revie, C., Cox, R., 2014. Modelling the impact of temperature-induced life history plasticity and mate limitation on the epidemic potential of a marine ectoparasite. PLoS ONE 9, e88465.
Groner, M., McEwan, G., Rees, E., Gettinby, G., Revie, C., 2016a. Quantifying the influence of salinity and temperature on the population dynamics of a marine ectoparasite. Can. J. Fish. Aquatic Sci. 73, 1–11.
Groner, M., Rogers, L., Bateman, A., Connors, B., Frazer, L., Godwin, S., Krkosek, M., Lewis, M., Peacock, S., Rees, E., Revie, C., Schlagel, U., 2016b. Lessons from sea louse and salmon epidemiology. Philos. Trans. R. Soc. B: Biol. Sci. 371, 20150203.
Halvorsen, K., Larsen, T., Sørdalen, T., Vøllestad, L., Knutsen, H., Olsen, E., 2017. Impact of harvesting cleaner fish for salmonid aquaculture assessed from replicated coastal marine protected areas. Mar. Biol. Res. 0, 1–11.
Hamre, L., Eichner, C., Caipang, C., Dalvin, S., Bron, J., 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.
Höhle, M., 2009. Additive-multiplicative regression models for spatio-temporal epidemics. Biom. J. 51, 961–978.
Jansen, P., Grøntvedt, R., Tarpai, A., Helgesen, K., Horsberg, T., 2016. Surveillance of the sensitivity towards antiparasitic bath-treatments in the salmon louse (*Lepeophtheirus salmonis*). PLOS ONE 11, e0149006.
Jansen, P., Kristoffersen, A., Viljugrein, H., Jimenez, D., Aldrin, M., Stien, A., 2012. Sea lice as a density dependent constraint to salmonid farming. Proc. R. Soc. B 279, 2330–2338.
Johansen, L., Jensen, I., Mikkelsen, H., Bjørn, P., Jansen, P., Bergh, Ø., 2011. Disease interaction and pathogen exchange between wild and farmed fish populations with special reference to Norway. Aquaculture 315, 167–186.
Johnsen, I., Fiksen, O., Sandvik, A., Asplin, L., 2014. Vertical salmon lice behaviour as a response to environmental conditions and its influence on regional dispersion in a fjord system. Aquacult. Environ. Interact. 5, 127–141.
Jonkers, A., Sharkey, K., Thrush, M., Turnbull, J., Morgan, K., 2010. Epidemics and control strategies for diseases of farmed salmonids: a parameter study. Epidemics 2, 195–206.
Keeling, M., Woolhouse, M., Shaw, D., Matthews, L., Chase-Topping, M., Haydon, D., Cornell, S., Kappey, J., Wilesmith, J., Grenfell, B., 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science 294, 813–817.
Kristoffersen, A.D.J., Viljugrein, H., Grøntvedt, R., Stien, A., Jansen, P., 2014. Large scale modelling of salmon lice (*Lepeophtheirus salmonis*) infection pressure based on lice monitoring data from Norwegian salmonid farms. Epidemics 9, 31–39.

Kristoffersen, A., Rees, E., Stryhn, H.R.I., Campisto, J., Revie, C., St-Hilaire, S., 2013. Understanding sources of sea lice for salmon farms in Chile. Prev. Vet. Med. 111, 165–175.
Kristoffersen, A., Viljugrein, H., Kongtorp, R., Brun, E., Jansen, P., 2009. Risk factors for pancreas disease (PD) outbreaks in farmed Atlantic salmon and rainbow trout in Norway during 2003–2007. Prev. Vet. Med. 90, 127–136.
Krkošek, M., Bateman, A., Proboszcz, S., Orr, C., 2010. Dynamics of outbreak and control of salmon lice on two salmon farms in the Broughton Archipelago. Aquacult. Environ. Interact. 1, 137–146.
Krkošek, M., Connors, B., Lewis, M., Poulin, R., 2012. Allee effects may slow the spread of parasites in a coastal marine ecosystem. Am. Nat. 179, 401–412.
Krkošek, M., Morton, A., Volpe, J., Lewis, M., 2009. Sea lice and salmon population dynamics: effects of exposure time for migratory fish. Proc. Biol. Sci. 276, 2819–2828.
Leclercq, E., Davie, A., Migaud, H., 2014. Delousing efficiency of farmed ballan wrasse (*Labrus bergylta*) against *Lepeophtheirus salmonis* infecting Atlantic salmon (*Salmo salar*) post-smolts. Pest Manag. Sci. 70, 1274–1282.
Murray, A., Salama, N., 2016. A simple model of the role of area management in the control of sea lice. Ecol. Model. 337, 39–47.
Murray, A., Wardeh, M., McIntyre, K., 2016. Using the H-index to assess disease priorities for salmon aquaculture. Prev. Vet. Med. 126, 199–207.
Nelson, W.A., Bjørnstad, O.N., Yamanaka, T., 2013. Recurrent insect outbreaks caused by temperature-driven changes in system stability. Science 341, 796–799.
Nisbet, R.M., Gurney, W.S.C., 1983. The systematic formulation of population models for insects with dynamically varying instar duration. Theor. Popul. Biol. 23, 114–135.
Nygaard, S., 2010. Slides - Lusekurs for røktere - Bokn 26. mars 2010. Tech. rep. Fiskehelse og Miljø AS.
Ottesen, K., Rykhus, K., et al., 2012. Terapiveileder - Medikamentell behandling mot lakselus, revidert utgave våren 2012, Tech. rep. Arbeidsgruppe.
Pettersen, J., Brynildsrud, O., Huseby, R., Rich, K., Aunsmo, A., Jensen, B., Aldrin, M., 2016. The epidemiological and economic effects from systematic depopulation of Norwegian marine salmon farms infected with pancreas disease virus. Prev. Vet. Med. 132, 113–124.
Revie, C., Robbins, C., Gettinby, G., Kelly, L., Treasurer, J., 2005. A mathematical model of the growth of sea lice, *Lepeophtheirus salmonis*, populations on farmed Atlantic salmon, *Salmo Salar* L., in Scotland and its use in the assessment of treatment strategies. J. Fish Dis. 28, 603–613.
Rittenhouse, M., Revie, C., Hurford, A., 2016. A model for sea lice (*Lepeophtheirus salmonis*) dynamics in a seasonally changing environment. Epidemics 16, 8–16.
Salama, N., Murray, A., 2013. A comparison of modelling approaches to assess the transmission of pathogens between Scottish fish farms: the role of hydrodynamics and site biomass. Prev. Vet. Med. 108, 285–293.
Samsing, F., Oppedal, F., Dalvin, S., Johnsen, I., Vågseth, T., Dempster, T., 2016. Salmon lice (*Lepeophtheirus salmonis*) development times, body size, and reproductive outputs follow universal models of temperature dependence. Can. J. Fish. Aquatic Sci. 73, 1841–1851.
Scheel, I., Aldrin, M., Frigessi, A., Jansen, P., 2007. A stochastic model for infectious salmon anemia (ISA) in Atlantic salmon farming. J. R. Soc. Interface 4, 699–706.
Skiftesvik, A., Blom, G., Agnalt, A.B., Durif, C., HI, B., Bjelland, R., Harkestad, L., Farestveit, E., Paulsen, O., Fauske, M., Havelin, T., Johnsen, K., Mortensen, S., 2014. Wrasse (Labridae) as cleaner fish in salmonid aquaculture – the Hardangerfjord as a case study. Mar. Biol. Res. 10 (3), 289–300.
Stien, A., Bjørn, P., Heuch, P., Elston, D., 2005. Population dynamics of salmon lice *Lepeophtheirus salmonis* on Atlantic salmon and sea trout. Mar. Ecol. Prog. Ser. 290, 263–275.
Stormoen, M., Skjerve, E., Aunsmo, A., 2013. Modelling salmon lice, *Lepeophtheirus salmonis*, reproduction on farmed Atlantic salmon, *Salmo salar* L. J. Fish Dis. 36, 25–33.
Taranger, G., Karlsen, O., Bannister, R., Glover, K., Husa, V., Karlsbakk, E., Kvamme, B., Boxaspen, K., Bjørn, P., Finstad, B., Madhun, A., HC, M., Svasand, T., 2015. Risk assessment of the environmental impact of Norwegian Atlantic salmon farming. ICES J. Mar. Sci. 72, 997–1021.
Tildesley, M., Smith, G., Keeling, M., 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.
Vollset, K., Krontveit, R., Jansen, P., Finstad, B., Barlaup, B., Skilbrei, O., Krkošek, M., Romunstad, P., Aunsmo, A., Jensen, A., Dohoo, I., 2015. Impacts of parasites on marine survival of Atlantic salmon: a meta-analysis. Fish Fish. 17, 714–730.

A population model for salmon lice - Additional results and modelling details

Supplementary Material to
A stage-structured Bayesian hierarchical model for salmon lice populations at individual salmon farms - Estimated from data from a set of farms

Aldrin, M., Huseby, R.B., Stien, A., Grøntvedt, R.,Viljugrein, H., Jansen, P.A.

May 3, 2017

Contents

1 Modelling details

1.1 Population model with movement of salmon between cages

From the CH stage on, the lice are attached to a fish, and therefore associated with a specific cage. We assume that the attached lice follow the fish if the fish are moved to another cage or if the fish are removed from the farm (including slaughter and other fish mortality). To handle this, we have introduced the factor $w_{(t-1)c'c}$, which is the proportion of salmon moved from cage c' to cage c at the end of time $t-1$, and if $c' = c$ it is the proportion of fish that stay in the cage. The equations below for the CH, PA and AF stages then replace the corresponding equations in the main article.

Model for the chalimus stage

$$N^{CH}_{t(a=0)c} = \sum_{a'} N^{CO}_{(t-1)a'} s^{CO}_{(t-1)a'} d^{CO}_{(t-1)a'c}, \quad (1)$$

$$N^{CH}_{t(a>0)c} = \sum_{c'} w_{(t-1)c'c} N^{CH}_{(t-1)(a-1)c'} s^{CH}_{(t-1)(a-1)c'} [1 - d^{CH}_{(t-1)(a-1)}]. \quad (2)$$

Model for the pre-adult stage

$$N^{PA}_{t(a=0)c} = \sum_{a'} \sum_{c'} w_{(t-1)c'c} N^{CH}_{(t-1)a'c'} s^{CH}_{(t-1)a'c'} d^{CH}_{(t-1)a'}, \quad (3)$$

$$N^{PA}_{t(a>0)c} = \sum_{c'} w_{(t-1)c'c} N^{PA}_{(t-1)(a-1)c'} s^{PA}_{(t-1)(a-1)c'} [1 - d^{PA}_{(t-1)(a-1)}]. \quad (4)$$

Model for the adult female stage

$$N^{AF}_{t(a=0)c} = 0.5 \sum_{a'} \sum_{c'} w_{(t-1)c'c} N^{PA}_{(t-1)a'c'} s^{PA}_{(t-1)a'c'} d^{PA}_{(t-1)a'}, \quad (5)$$

$$N^{AF}_{t(a>0)c} = \sum_{c'} w_{(t-1)c'c} N^{AF}_{(t-1)(a-1)c'} s^{AF}_{(t-1)(a-1)c'}. \quad (6)$$

2 Priors

The priors for the parameters defined in the main article are specified in Table 1, including those set to fixed values, whereas the priors for the parameters defined in this Supplementary Material are specified in Table 2. We give motivations for some of these priors in the following subsections.

Table 1: Prior distributions for parameters in the data model and in the deterministic parts of the population model. G(s,r): gamma distribution with shape s and rate r, U: uniform distribution, U*: approximate uniform distribution by the algorithm described in Section 2.6, N: normal distribution, N*: normal distribution truncated at mean ± 3 standard deviations, invlogit(x): $\exp(x)/(1+\exp(x))$.

Part of model	Stage	Parameter interpretation	Parameter symbol	Prior
Natural mortality	R,CO	Mortality rate	λ^{RCOnat}	$\lambda^{RCOnat} \sim$ U*(0.02,0.99)
Mortality cl.f.	PA, A	Regression coeff.	λ^{clf}	$\lambda^{clf} \sim$ U(0,5)
Development	Egg	Median at 10°C	δ^{Em10}	$\delta^{Em10} \sim$ N*(10, 2^2)
Development	Nauplii	Median at 10°C	δ^{Nm10}	$\delta^{Nm10} \sim$ N*(5, 1^2)
Development	R	Shape parameter	δ^{Rs}	$\delta^{Rs} \sim$ U(1,20)
Development	R	Power parameter	δ^{Rp}	$\delta^{Rp} \sim$ N*(1, 0.2^2)
Development	CH	Median at 10°C	δ^{CHm10}	$\delta^{CHm10} \sim$ N*(16, 3^2)
Development	CH	Shape parameter	δ^{CHs}	$\delta^{CHs} \sim$ U(1,20)
Development	CH	Power parameter	δ^{CHp}	$\delta^{CHp} \sim$ N*(1, 0.2^2)
Development	PA	Median at 10°C	δ^{PAm10}	$\delta^{PAm10} \sim$ N*(16, 3^2)
Development	PA	Shape parameter	δ^{PAs}	$\delta^{PAs} \sim$ U(1,20)
Development	PA	Power parameter	δ^{PAp}	$\delta^{PAp} \sim$ N*(1, 0.2^2)
Development	CO	Expectation	δ_0^{CO}	invlogit(δ_0^{CO})$\sim$ U*(10^{-6},0.25)
Development	CO	Regression coeff.	δ_1^{CO}	$\delta_1^{CO} \sim$ U(0,1)
Development	CO	Variance within farm	$(\sigma^{COdf})^2$	$1/(\sigma^{COdf})^2 \sim$ G(0.001,0.001) and $\sigma^{COdf} < 5$
Development	CO	Variance between farms	$(\sigma^{COd})^2$	$1/(\sigma^{COd})^2 \sim$ G(0.001,0.001) and $\sigma^{COd}) < 5$
Reproduction	AF to R	Basic number of eggs	β_0^r	172.500, fixed
Reproduction	AF to R	Age dependence	β_1^r	0.200,fixed
Reproduction	AF to R	Density dependence	γ^r	log(γ^r)$\sim$ U(0,log(500))
Cleaner fish model		Mortality rate	κ^{clf}	$\kappa^{clf} \sim$ U(0,0.5)
Data model	CH	Aggregation parameter	ρ^{CH}	log(ρ^{CH})$\sim$ U(log(0.01),log(100))
Data model	OM=PA+AM	Aggregation parameter	ρ^{OM}	log(ρ^{OM})$\sim$ U(log(0.01),log(100))
Data model	AF	Aggregation parameter	ρ^{AF}	log(ρ^{AF})$\sim$ U(log(0.01),log(100))
Data model	CH	Expectation	$\beta_0^{CHcount}$	invlogit($\beta_0^{CHcount}$) $\sim$ U*(0.03,0.98)
Data model	CH	Variance	$(\sigma^{CHcount})^2$	$1/(\sigma^{CHcount})^2 \sim$ G(0.001,0.001) and $\sigma^{CHcount} < 5$
Data model	CH	Regression coeff.	$\beta_1^{CHcount}$	$\beta_1^{CHcount} \sim$ U(-1,0)

Table 2: Prior distributions for parameters in the stochastic parts of the population model. G(s,r): gamma distribution with shape s and rate r, U: uniform distribution, U*: approximate uniform distribution by the algorithm described in Section 2.6, N: normal distribution, N*: normal distribution truncated at mean $\pm$ 3 standard deviations, invlogit(x): $\exp(x)/(1+\exp(x))$.

Part of model	Stage	Parameter interpretation	Parameter symbol	Prior
Natural mortality	CH	Expectation in AR(1)	λ_0^{CHnat}	invlogit(λ_0^{CHnat})$\sim$ U*(0.0006,0.02)
Natural mortality	CH	Coefficient in AR(1)	ϕ^{CHnat}	$\phi^{CHnat} \sim$ U(0,0.99)
Natural mortality	CH	Variance in AR(1)	$(\sigma^{CHnat})^2$	$1/(\sigma^{CHnat})^2 \sim$ G(0.001,0.001) and $\sigma^{CHnat} < 5$
Natural mortality	PA	Expectation in AR(1)	λ_0^{PAnat}	invlogit(λ_0^{PAnat})$\sim$ U*(0.002,0.21)
Natural mortality	PA	Coefficient in AR(1)	ϕ^{PAnat}	$\phi^{PAnat} \sim$ U(0,0.99)
Natural mortality	PA	Variance in AR(1)	$(\sigma^{PAnat})^2$	$1/(\sigma^{PAnat})^2 \sim$ G(0.001,0.001) and $\sigma^{PAnat} < 5$
Natural mortality	A	Expectation in AR(1)	λ_0^{Anat}	invlogit(λ_0^{Anat})$\sim$ U*(0.003,0.70)
Natural mortality	A	Coefficient in AR(1)	ϕ^{Anat}	$\phi^{Anat} \sim$ U(0,0.99)
Natural mortality	A	Variance in AR(1)	$(\sigma^{Anat})^2$	$1/(\sigma^{Anat})^2 \sim$ G(0.001,0.001) and $\sigma^{Anat} < 5$
Mortality ch.tr.	CH, PA, A	Expectation, deltamethrin	λ^{DMcht}	$\lambda^{DMcht} \sim$ N(0,100^2)
Mortality ch.tr.	CH, PA, A	Variance, deltamethrin	$(\sigma^{DMcht})^2$	$1/(\sigma^{DMcht})^2 \sim$ G(0.001,0.001) and $\sigma^{DMcht} < 5$
Mortality ch.tr.	PA, A	Expectation, azamethiphos	λ^{AZcht}	$\lambda^{AZcht} \sim$ N(0,100^2)
Mortality ch.tr.	PA, A	Variance, azamethiphos	$(\sigma^{AZcht})^2$	$(\sigma^{AZcht})^2 = (\sigma^{DMcht})^2$
Mortality ch.tr.	PA, A	Expectation, H_2O_2	λ^{HPcht}	$\lambda^{HPcht} \sim$ N(0,100^2)
Mortality ch.tr.	PA, A	Variance, H_2O_2	$(\sigma^{HPcht})^2$	$\sigma^{HPcht})^2 = (\sigma^{DMcht})^2$
Mortality ch.tr.	CH, PA, A	Expectation, emamectin	λ^{EMcht}	$\lambda^{EMcht} \sim$ N(0,100^2)
Mortality ch.tr.	CH, PA, A	Variance, emamectin	$(\sigma^{EMcht})^2$	$1/(\sigma^{EMcht})^2 \sim$ G(0.001,0.001) and $\sigma^{EMcht} < 5$
Mortality ch.tr.	CH, PA	Expectation, diflubenzuron	λ^{DIcht}	$\lambda^{DIcht} \sim$ N(0,100^2)
Mortality ch.tr.	CH, PA	Variance, diflubenzuron	$(\sigma^{DIcht})^2$	$(\sigma^{DIcht})^2 = (\sigma^{EMcht})^2$
External recr.	AF to R	Expectation in AR(1)	μ^{Ext}	$\log(\mu^{Ext}) \sim$ N(0,1) and $0.01 < \mu^{Ext} < 100$
External recr.	AF to R	Coefficient in AR(1)	ϕ^{Ext}	$\phi^{PAnat} \sim$ U(0,0.99)
External recr.	AF to R	Variance in AR(1)	$(\sigma^{Extar})^2$	$1/(\sigma^{Extar})^2 \sim$ G(0.001,0.001) and $\sigma^{Extar} < 5$
External recr.	AF to R	Variance	$(\sigma^{Ext})^2$	$1/(\sigma^{Ext})^2 \sim$ G(0.001,0.001) and $\sigma^{Ext} < 5$

2.1 Natural mortality

The natural mortalities for the R and CO stages are set equal (to λ^{RCOnat}), since it is impossible to separate the mortalities in these stages as long as we have no data for them. Furthermore, λ^{RCOnat} is restricted to lie between 0.02 and 0.99. The high upper bound is chosen to take into account that larvae may drift away from a farm, which in our model has the same effect as mortality. The lower bound is chosen to account for a possible underestimation of eggs produced.

The natural mortalities for the CH, PA and A stages are AR(1)-processes on the logit scale. The priors for the expected values on the logit scale (λ_0^{nat}) are set such that the prior for these, transformed to the mortality scale, are approximately uniform within limits given as the most extreme limits of the intervals in Table 4 in Stien et al. (2005). Furthermore, as mentioned above, the processes themselves are restricted to lie within the same limits. The AR coefficients are restricted to be non-negative since negative values are implausible, and at most 0.99 to ensure stationarity. The priors for the variances are vague (inverse gamma distributed), but with a upper limit at 25 to ensure numerical stability. These priors are used for all variances.

2.2 Mortality due to treatment

By construction, the mortalities due to treatment are always between 0 and 1. The priors for the expectations (λ^{cht}) at the transformed scale are vague (normal distributed with a large variance), which means that the expected mortalities can take any value between 0 and 1.

2.3 Development rates

The priors for median development times at 10°C for eggs, nauplii, and the CH, PA and A stages (δ^{m10}) are truncated normal distributed with means taken from the minimum development times in Stien et al. (2005), and with a subjective chosen standard deviation of around 20%. In addition, each δ^{m10} must lie within 3 standard deviations from its prior mean. Note that the minimum development times referred to in Stien et al. (2005) are close to the corresponding median (and mean) development times, since the corresponding development rates after the minimum development times are quite high.

The priors for the shape parameters (δ^s) are uniform between 1 and 20. The lower bound ensures that the development rates start at 0 for stage-age=0, and the upper bound is set for numerical reasons, since higher values are practically equivalent.

The priors for the power parameters (δ^p) are truncated normal with mean 0 and standard deviation 0.2, which ensures that these parameters are not too far from 1, but which are flexible enough to allow some non-linearity.

2.4 Infection rate

Since larger fish have larger surfaces, we assume that δ_1^{CO} is non-negative. Furthermore, since the surface of a fish usually increases less than proportional with its weight (Niimi (1975)), we assume that δ_1^{CO} is at most 1.

2.5 Reproduction

In the reproduction equation, β_0^r, represents the number of viable eggs for the first extrusion. According to Stien et al. (2005), a female adult louse can produce a set of two egg strings with 70-152 eggs per string for the first set. Furthermore, the proportion of viable eggs can be in the range 0.5-0.9. So the number of viable eggs may be in the range 70-275 for the first set of egg strings, and we have simply fixed β_0^r to the mid point in this interval, *i.e.* 172.5. Note that if this is too low or too high, it can be compensated by changing the mortalities in the R and CO stages or the infection rate.

The term $(a+1)^{\beta_1^r}$ models how the number of viable eggs per extrusion increases by stage-age. Stien et al. (2005) stated that the number of eggs can be increased by a factor 2.1 (150/70) from the first to the fifth set of egg strings. and that hatching take about 9 days at 10°C. Further assuming that it takes one more day to extrude a new set of egg strings, the fifth set of egg strings are extruded after $(4 \cdot (9+1) = 40$ days. A reasonable value for β_1^r is therefore $\log(2.1)/\log(40) = 0.20$, and we have fixed β_1^r to that value.

The parameter γ^r controls the density dependency due to mate limitation Allee effects, and is restricted to lie between 0 (extreme density dependency) and 500 (ignorable density dependency).

2.6 Approximate uniform prior

In some parts of the MCMC algorithm (Section 4) it is computationally simpler to approximate uniform priors by truncated normal priors. Especially, if the uniform prior is given on the probability scale for a parameter p, the prior for the corresponding parameter $\beta_0 = \text{logit}(p)$ on the logit scale will sometimes be specified as a truncated normal distribution. This will not give exact uniform priors for p, but the following procedure approximates this reasonably well (denoted by $U*$ in Tables 1 and 2):

Specify the lower and upper limits p_l and p_u on the probability scale, and let

$$\begin{aligned} p_m &= (p_l + p_u)/2), && (7)\\ \beta_{0l} &= \log(p_l/(1-p_l)), && (8)\\ \beta_{0u} &= \log(p_u/(1-p_u)), && (9)\\ \beta_{0m} &= \log(p_m/(1-p_m)), && (10)\\ \sigma &= (\beta_{0u} - \beta_{0l})/2. && (11) \end{aligned}$$

Then, let prior for β_0 be

$$\begin{aligned} \beta_0 &\sim \mathrm{N}(\beta_{0m}, \sigma^2), && (12)\\ & && (13) \end{aligned}$$

truncated at β_{0l} and β_{0u}.

3 Additional results

3.1 Fitted values and conditional predictions

Figures 1- 7 give examples of data and fitted and conditional predicted values. Figure 1 corresponds to Figure 4 in the main article, but is for the average over all eight cages at that farm instead of only a single cage. Figures 2- 5 show similar results for single cages at the four other farms where the last months of lice data were not used for model fitting. Finally, Figures 6 and 7 show data fitted values for single cages for two farms where all data were used for estimation, *i.e.* no data were held back for predictions.

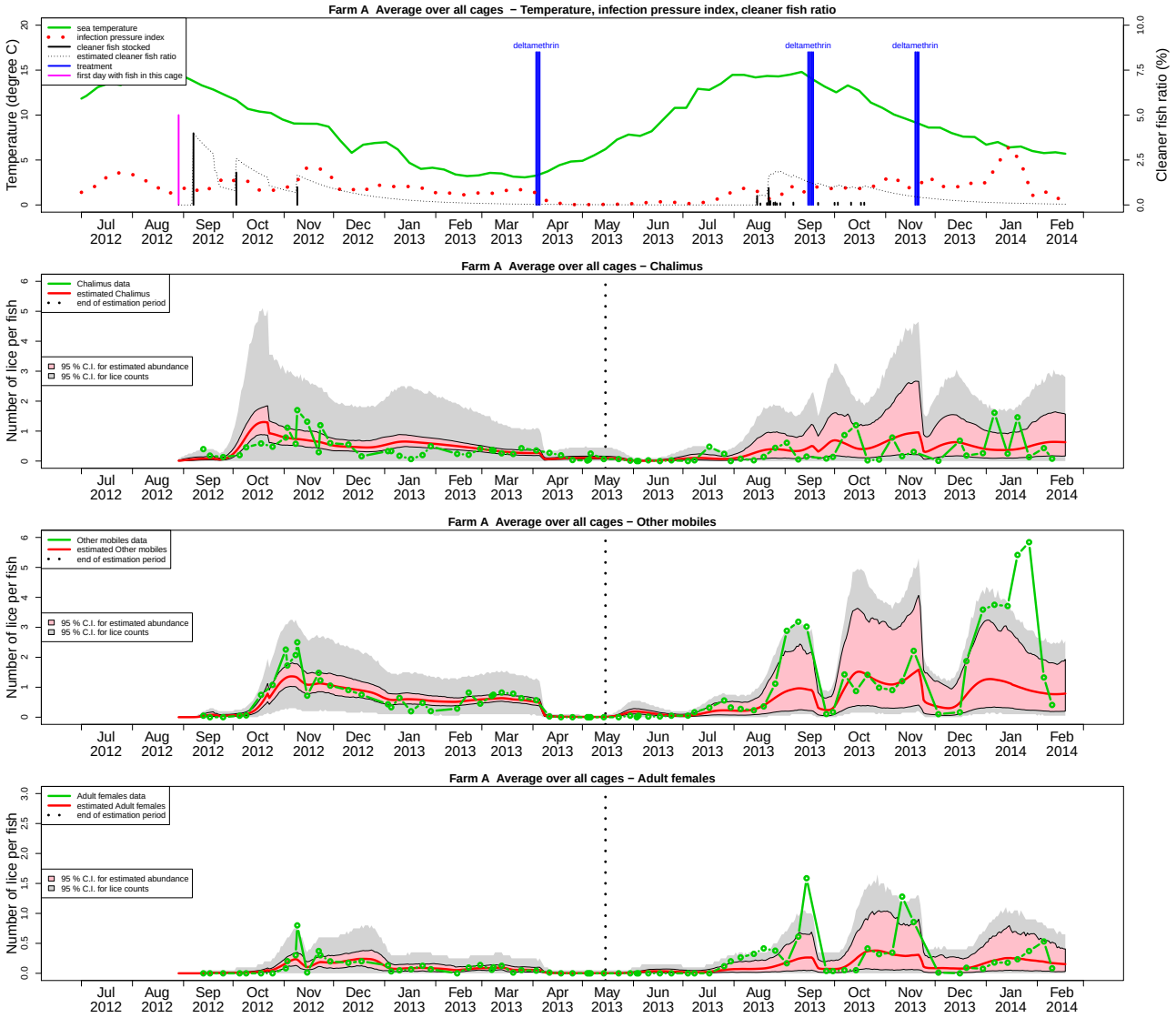

Figure 1: Fitted (until 15. May 2013) and predicted (from 16. May 2013) values for the lice and the cleaner fish populations in averages over all eight cages for farm A, *i.e.* the same farm as in Figures 2 and 4 in the main article. The symbols are described in Figure 4 in the main article.

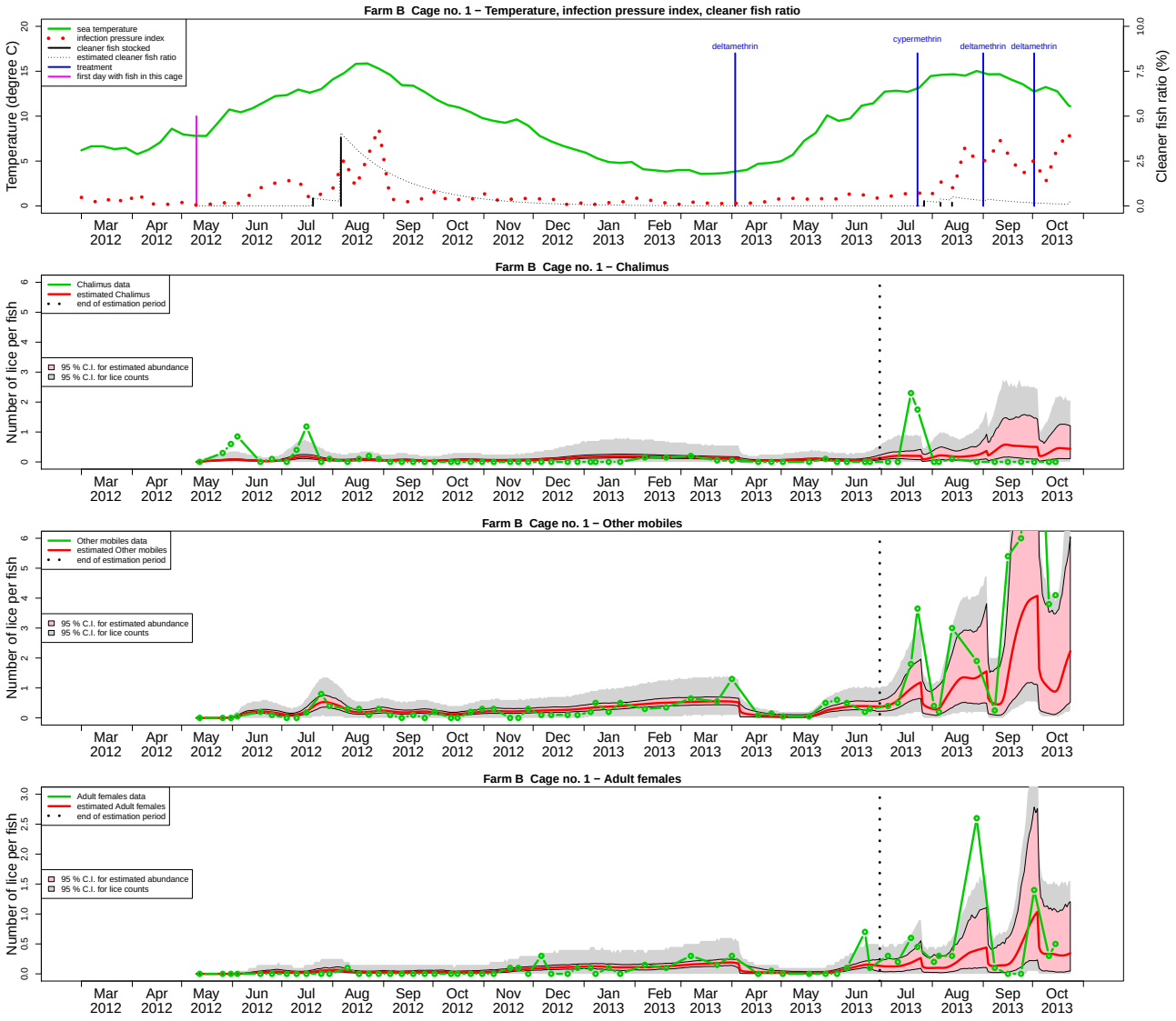

Figure 2: Fitted (until 30. June 2013) and predicted (from 1. June 2013) values for the lice and the cleaner fish populations in cage 1 in farm B. The symbols are described in Figure 4 in the main article.

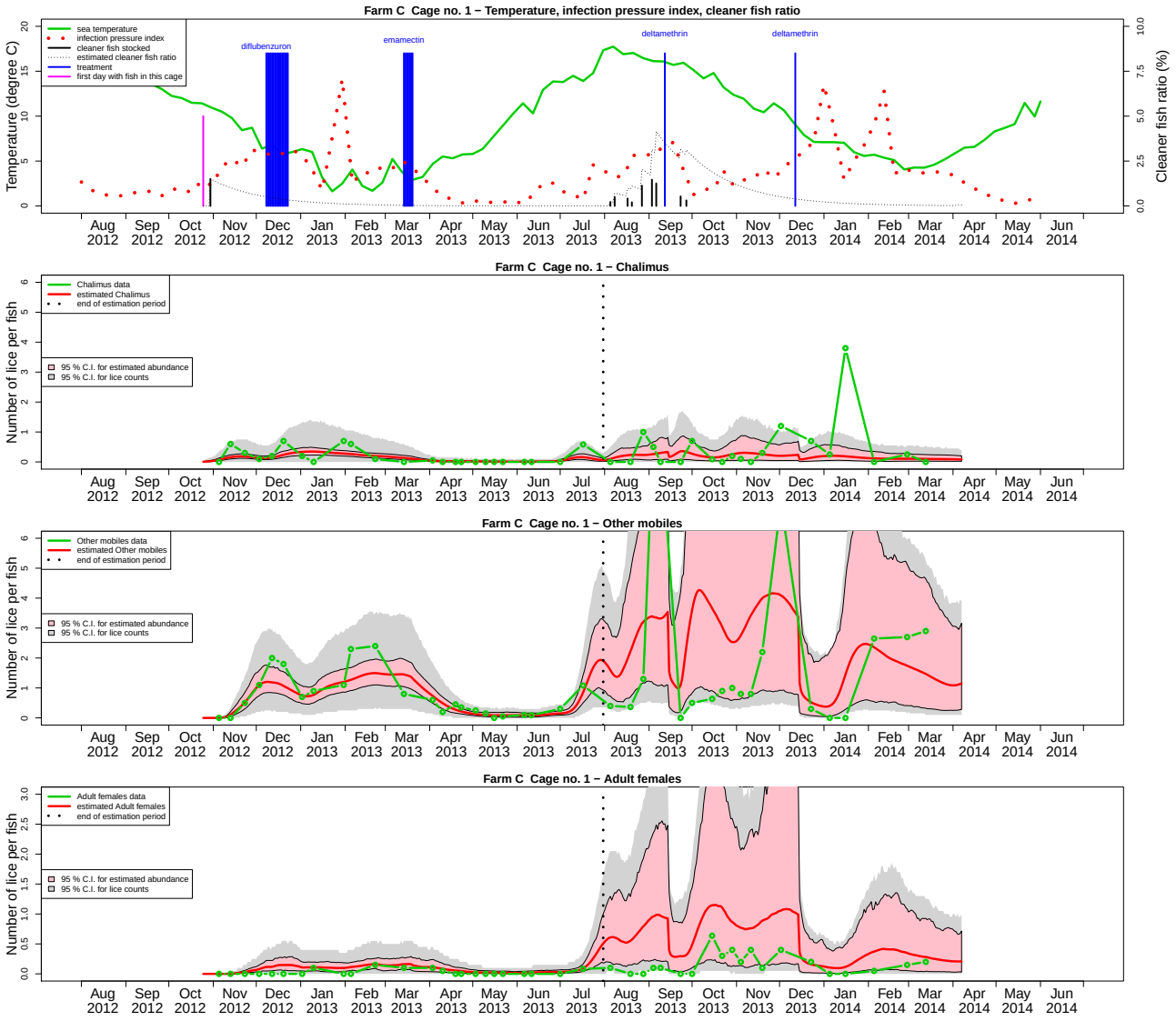

Figure 3: Fitted (until 31. July 2013) and predicted values (from 1. August 2013) for the lice and the cleaner fish populations in cage 1 in farm C. The symbols are described in Figure 4 in the main article.

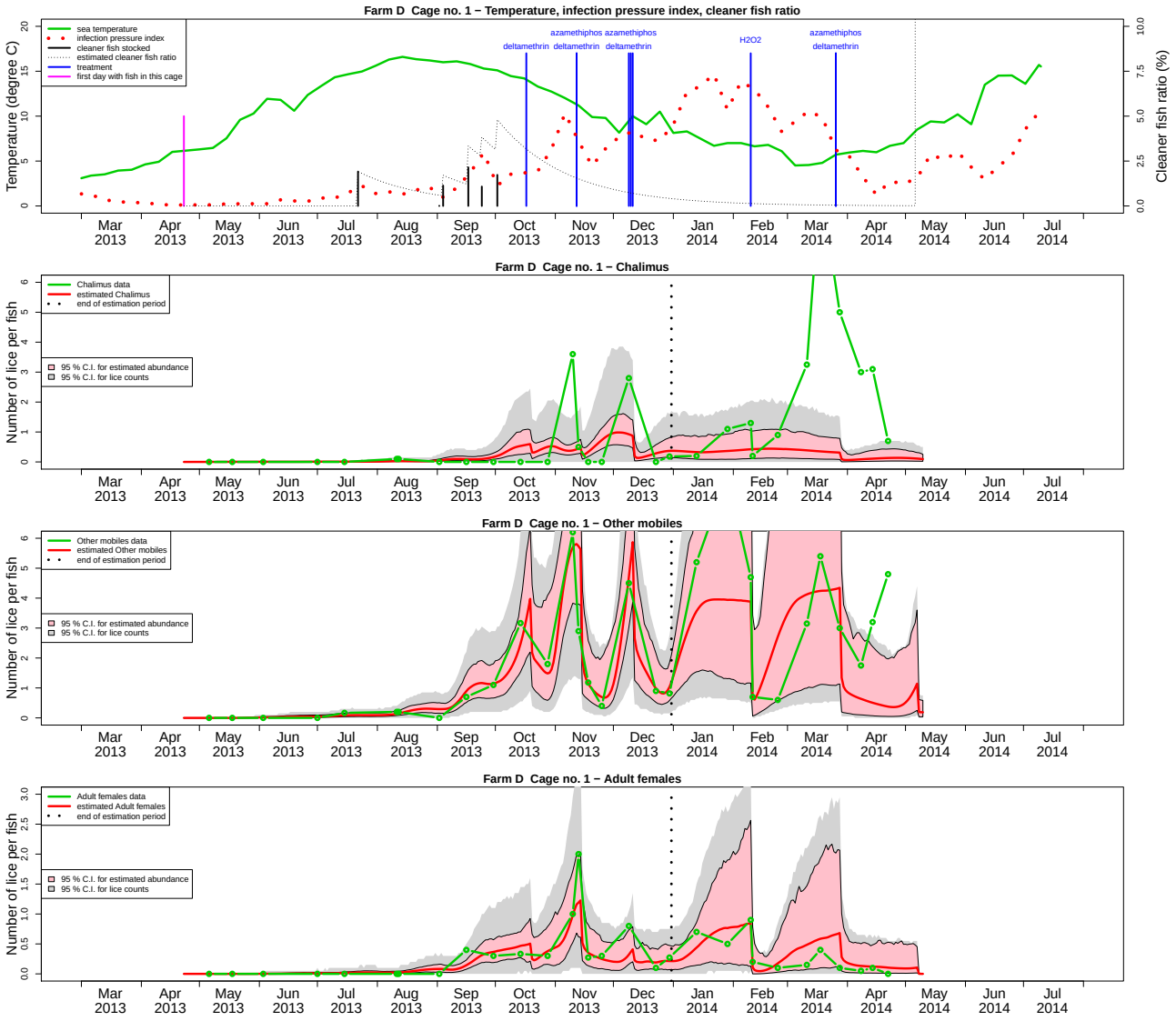

Figure 4: Fitted (until 31. December 2013 and predicted values (from 1. January 2014) for the lice and the cleaner fish populations in cage 1 in farm D. The symbols are described in Figure 4 in the main article.

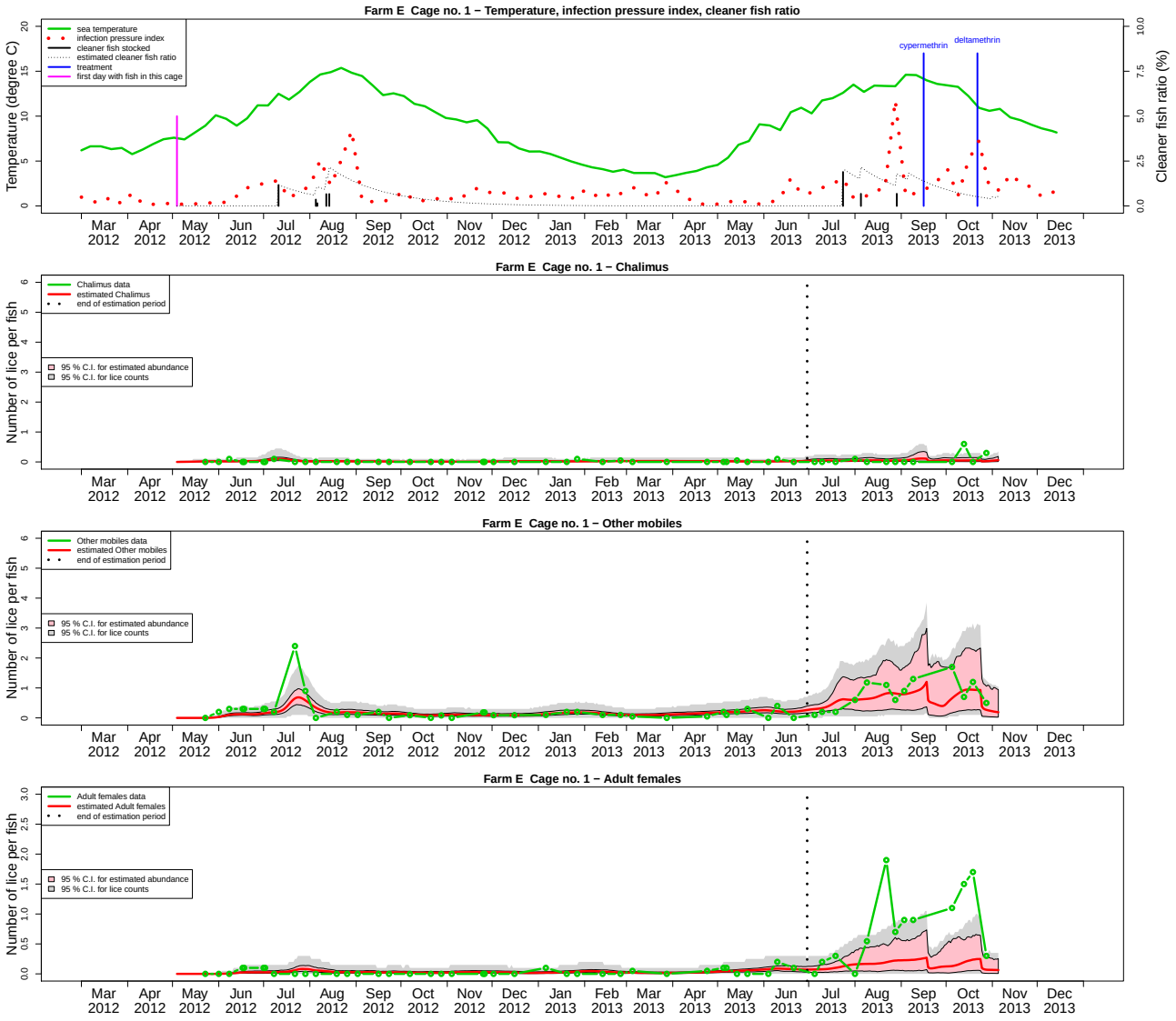

Figure 5: Fitted (until 30. June 2013) and predicted values (from 1. July 2013) for the lice and the cleaner fish populations in cage 1 in farm E. The symbols are described in Figure 4 in the main article.

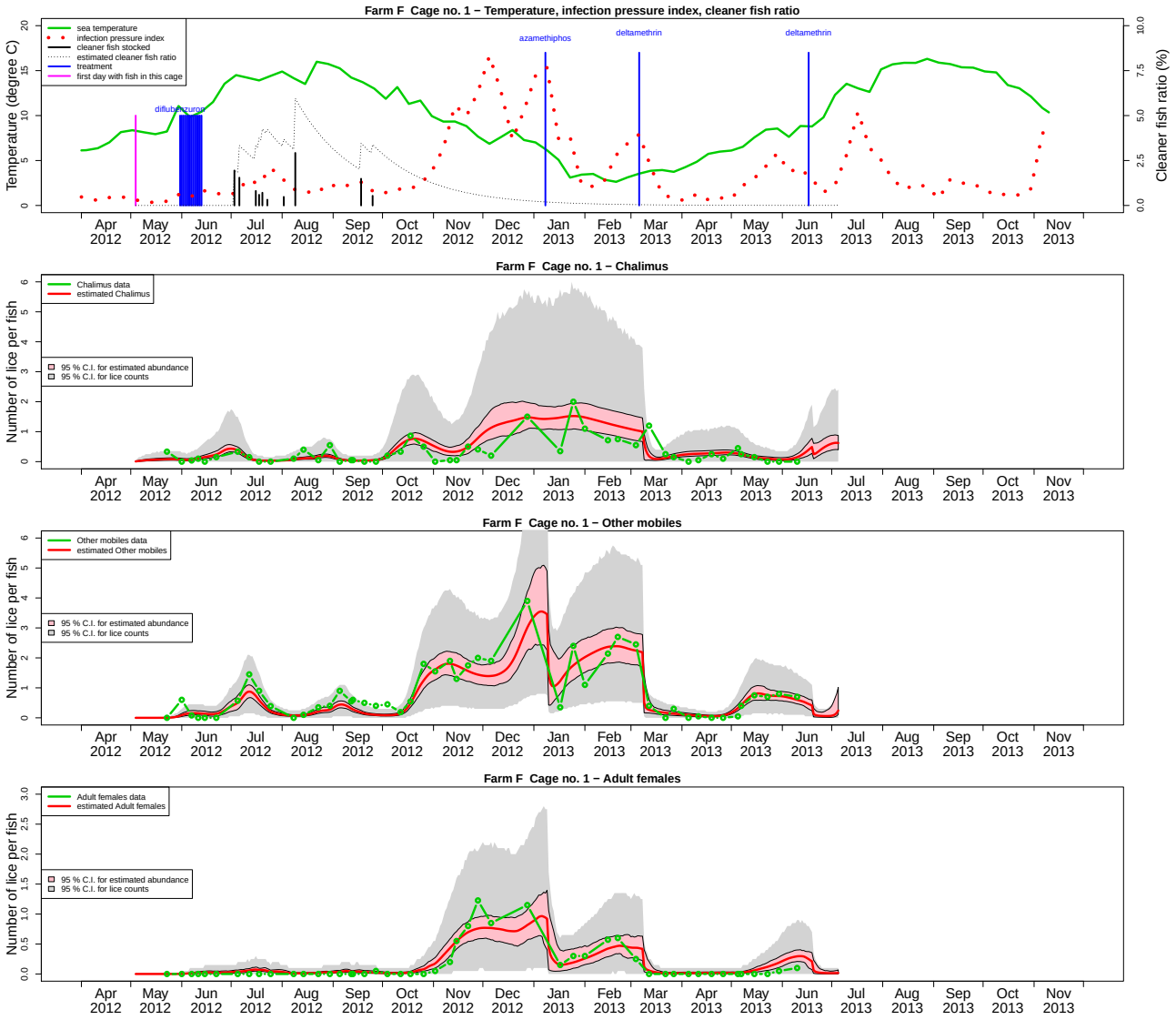

Figure 6: Fitted values for the lice and the cleaner fish populations in cage 1 in farm F. The symbols are described in Figure 4 in the main article.

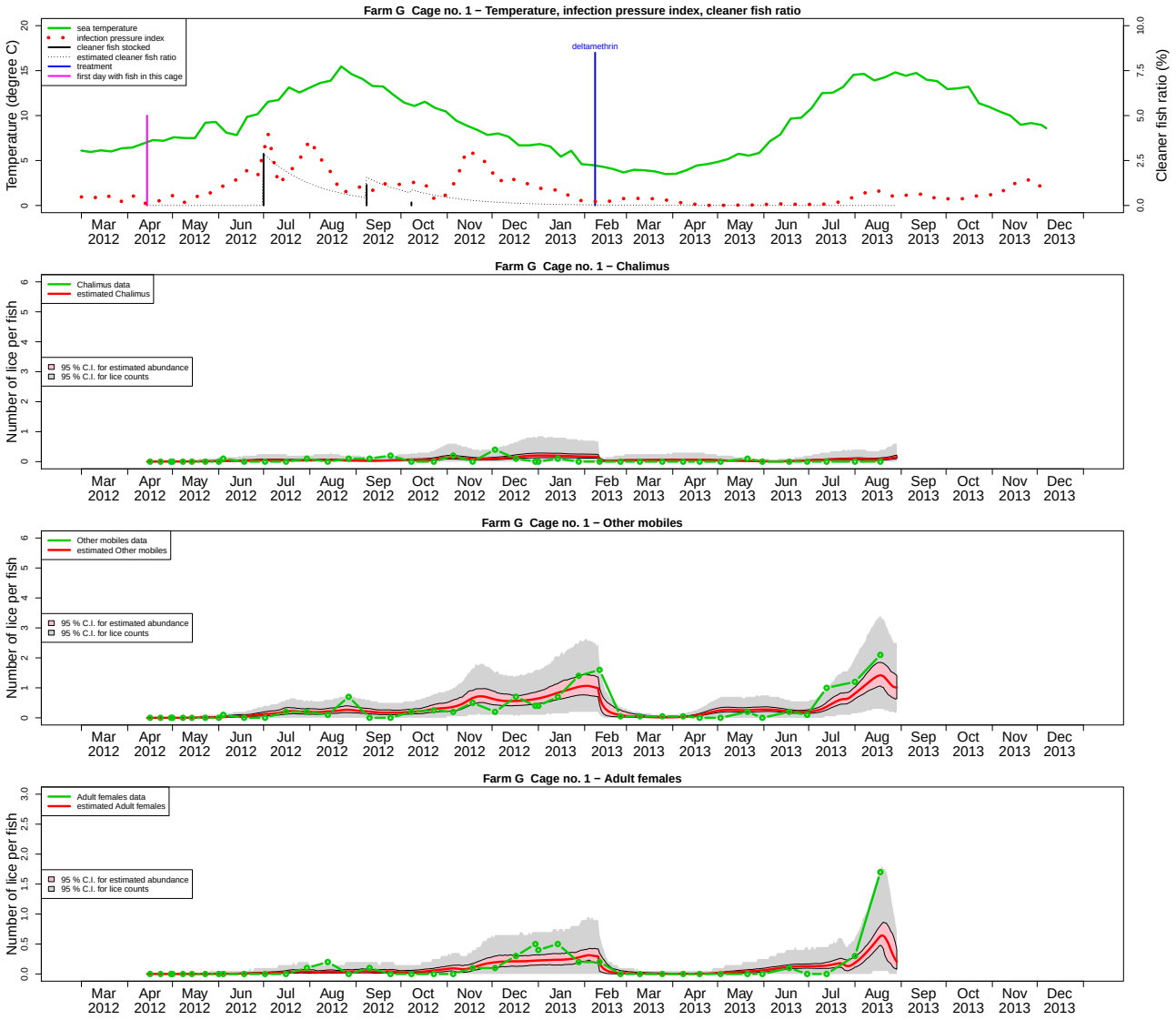

Figure 7: Fitted values for the lice and the cleaner fish populations in cage 1 in farm G. The symbols are described in Figure 4 in the main article.

3.2 Mortality due to cleaner fish

Figure 8 shows the estimated lice mortality due to the use of cleaner fish with a constant 10% cleaner fish ratio.

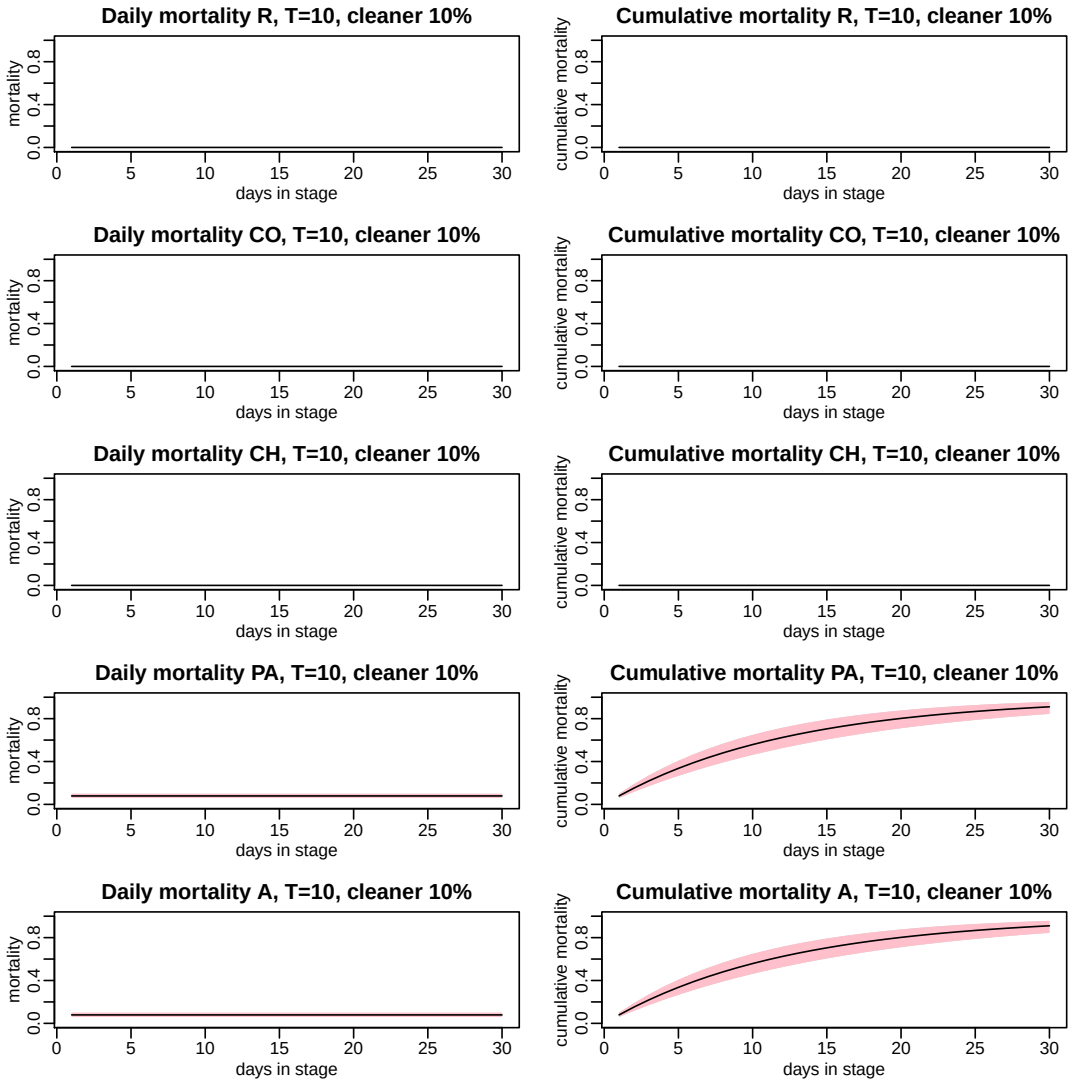

Figure 8: Daily mortality and cumulative mortality due to a cleaner fish ratio of 10%. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval.

3.3 Mortality due to chemotherapeutic treatment

Figures 9-13 shows the estimated lice mortality due to the use of various chemotherapeutic treatments. Remember that treatment effects for individual treatments are modelled as random effects, that may vary between stages within the same treatment and within different treatments with the same treatment type. The grey area in the figures reflect this random variability, whereas the solid lines are the estimated expected effect with a corresponding 95% credible interval given by the pink area.

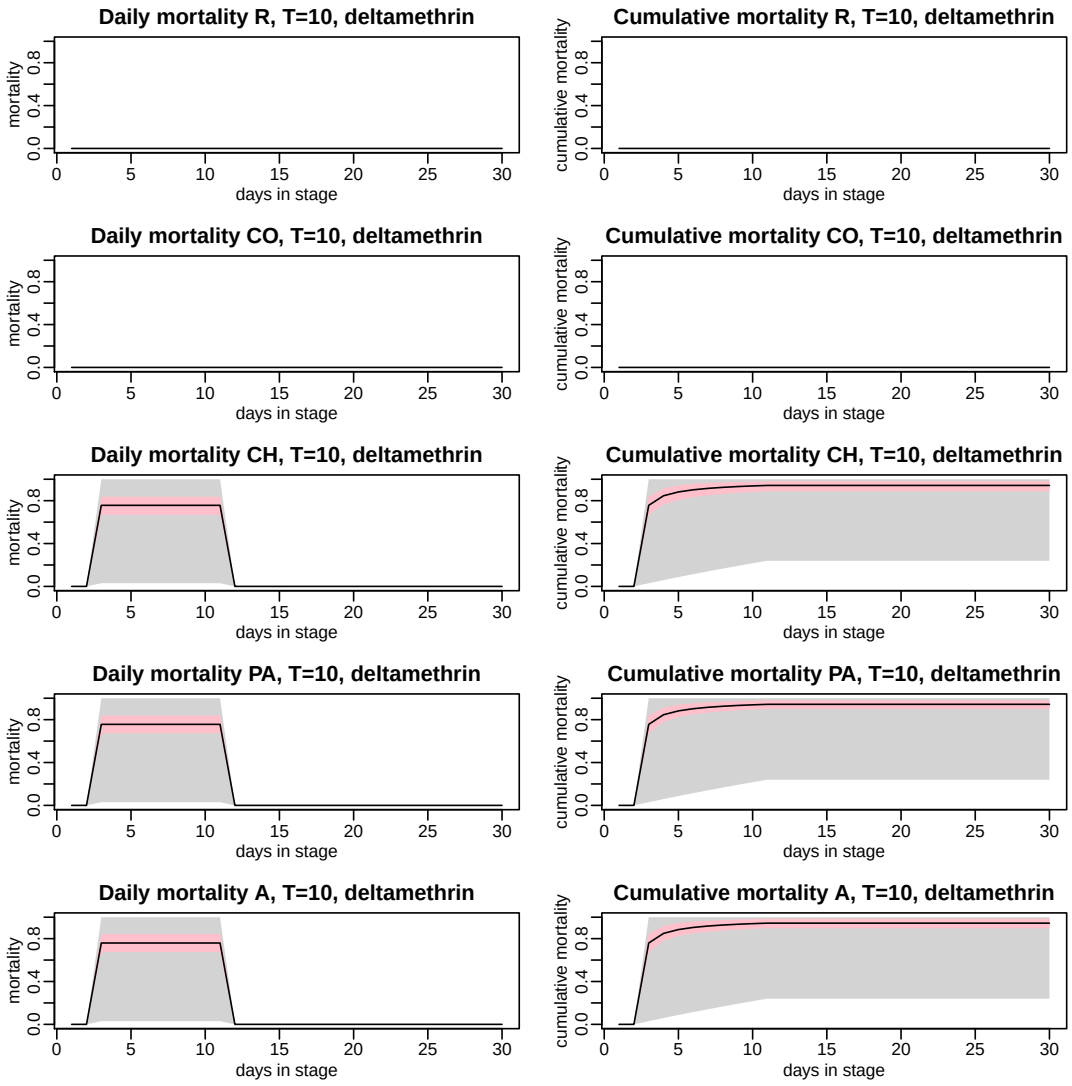

Figure 9: Daily mortality and cumulative mortality due to treatment with deltamethrin or cypermethrin. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

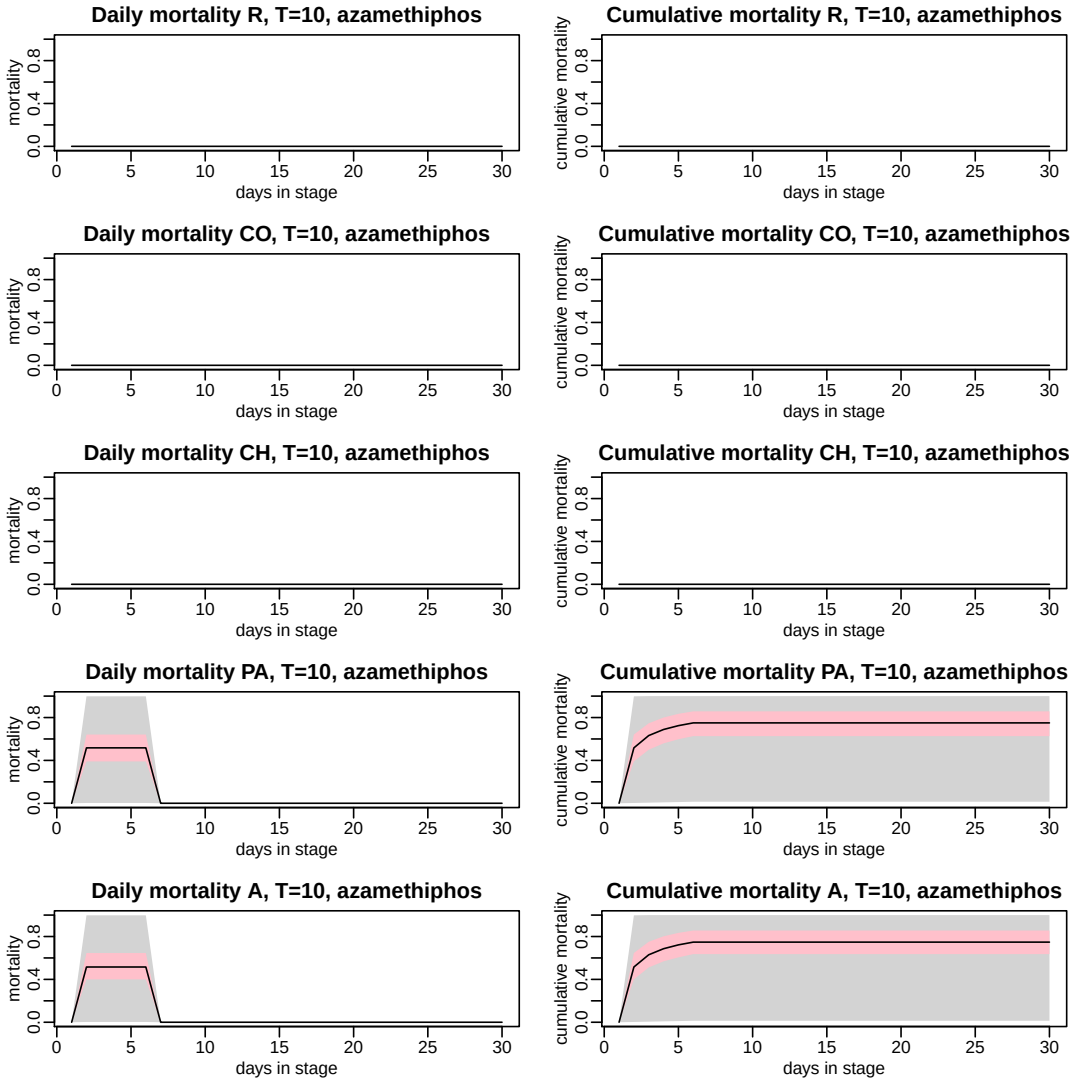

Figure 10: Daily mortality and cumulative mortality due to treatment with azamethiphos. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

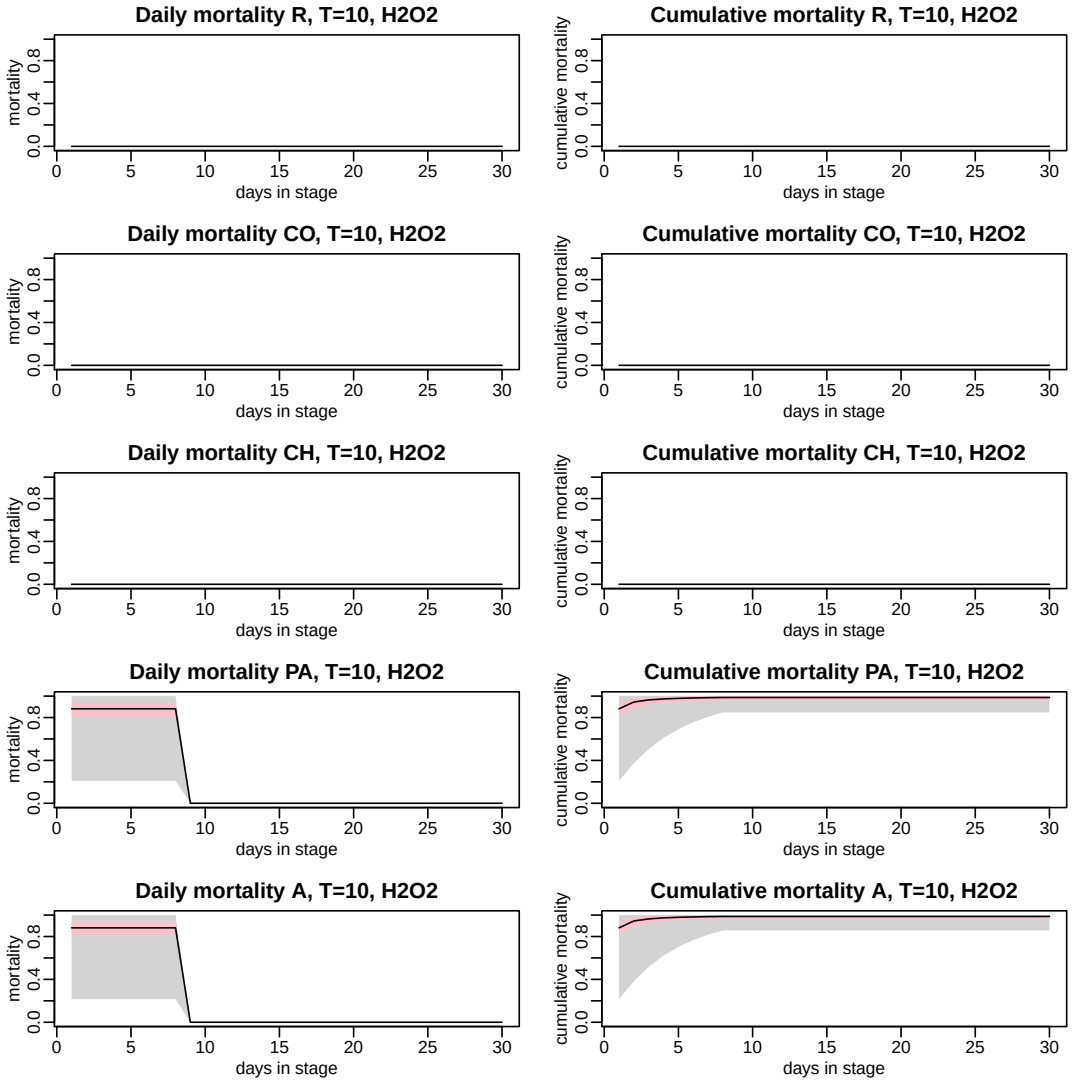

Figure 11: Daily mortality and cumulative mortality due to treatment with hydrogen peroxide. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

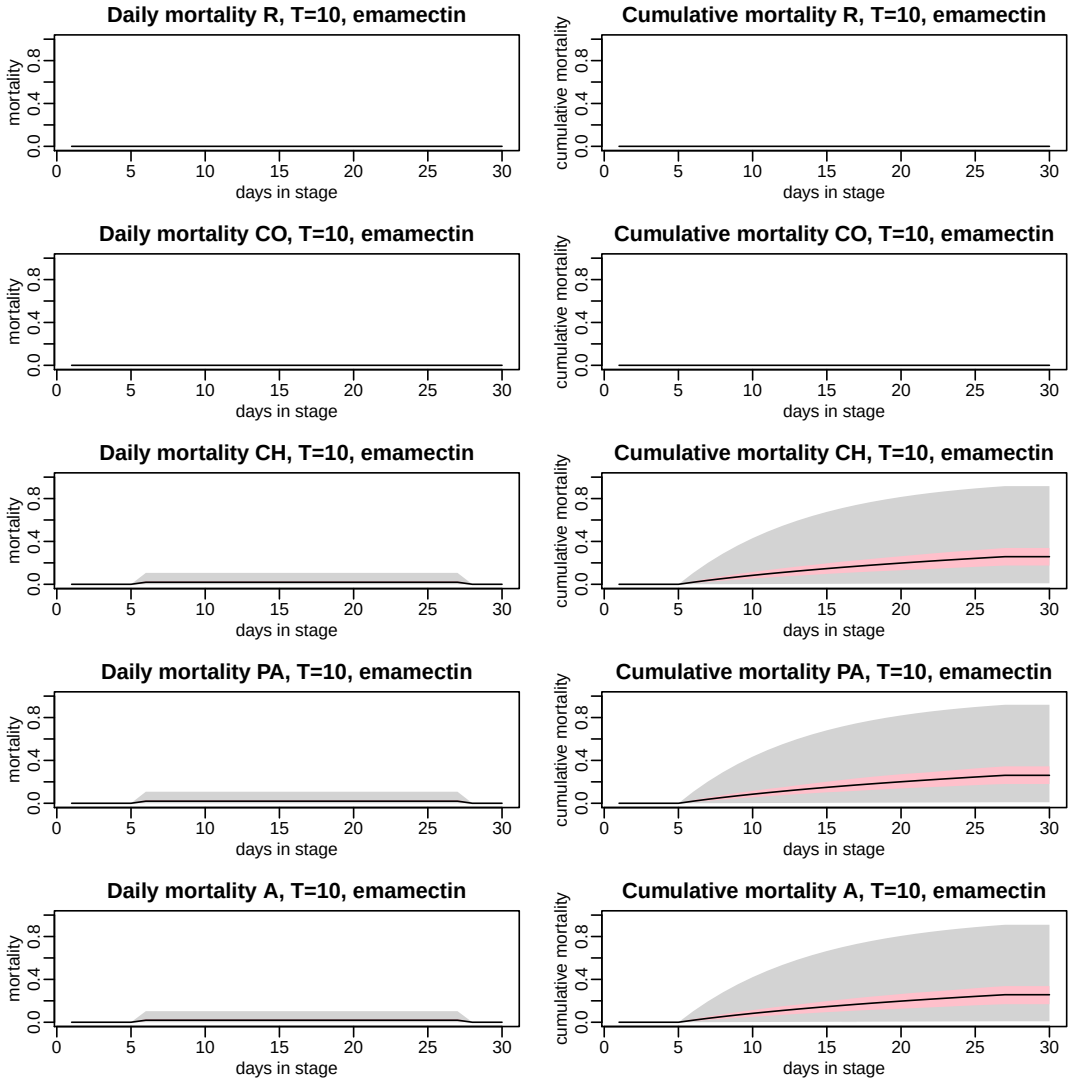

Figure 12: Daily mortality and cumulative mortality due to treatment with emamectin benzoate. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

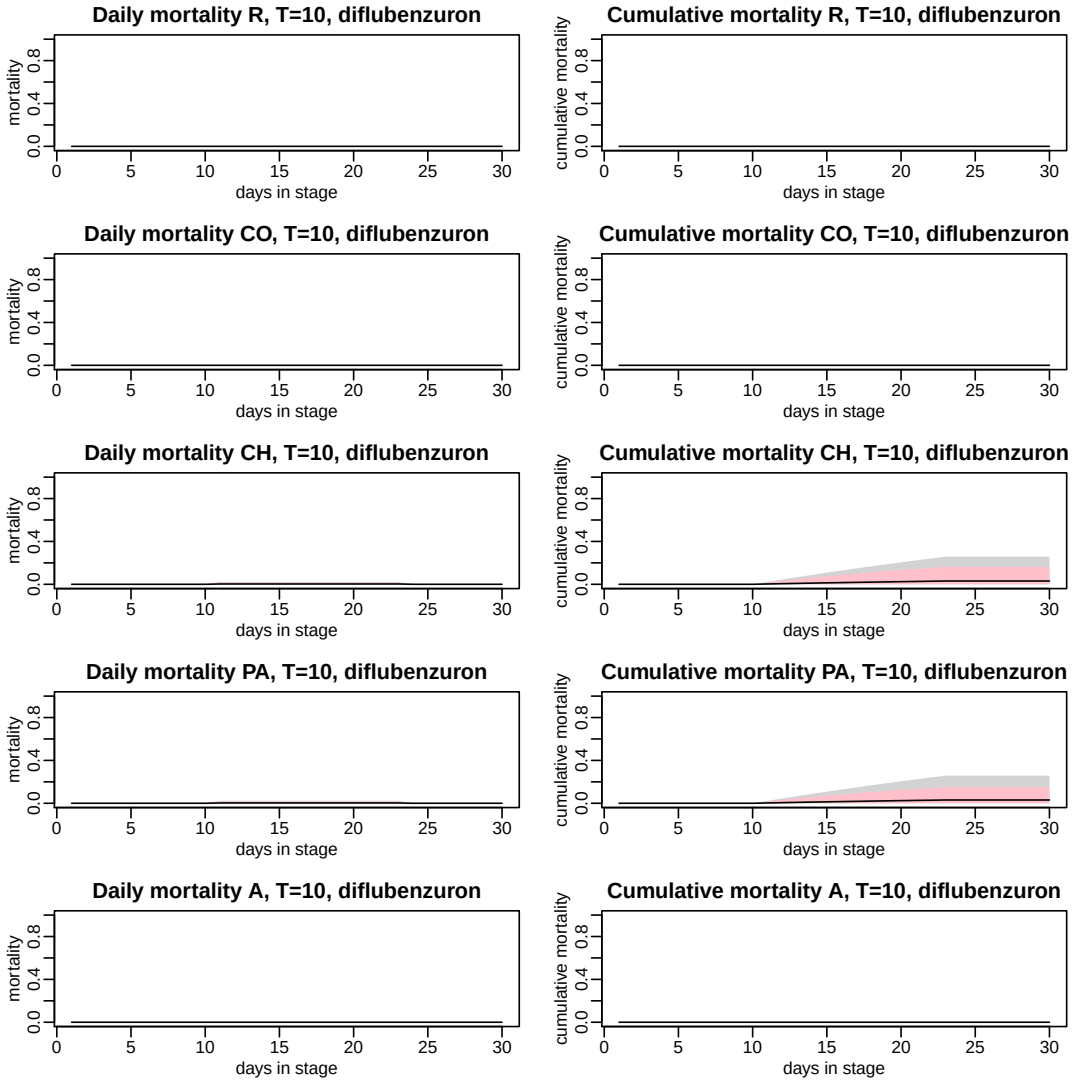

Figure 13: Daily mortality and cumulative mortality due to treatment with diflubenzuron. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

3.4 Development and reproduction

Figures 14-16 shows estimated development rates and egg production, including parameter uncertainties, for seawater temperatures 5, 10 and 15°C. The the infection rates (development from CO to CH) are calculated for a farm with ten cages with 100 000 fish in each cage.

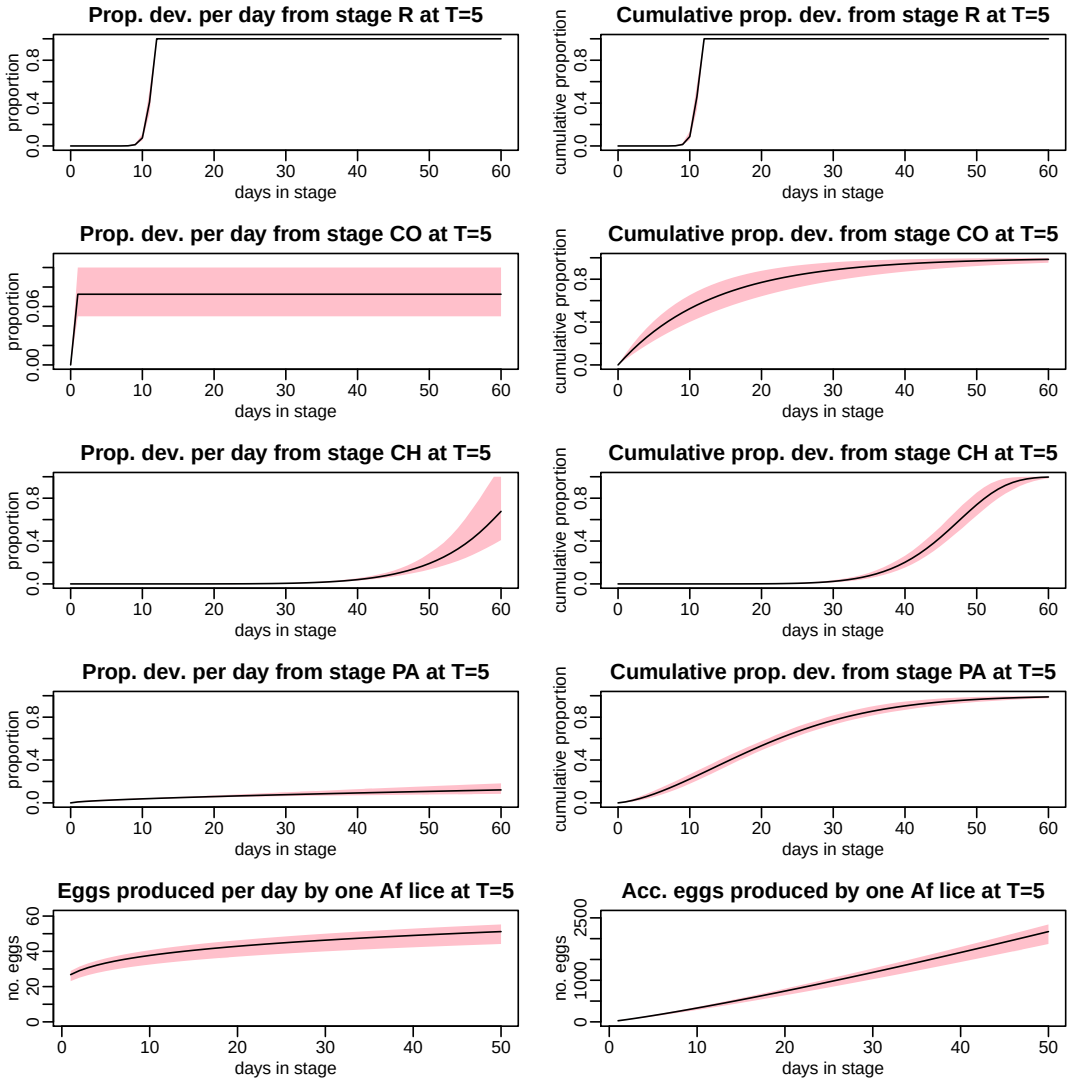

Figure 14: Daily and cumulative development rates and egg production at seawater temperature 5°C. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

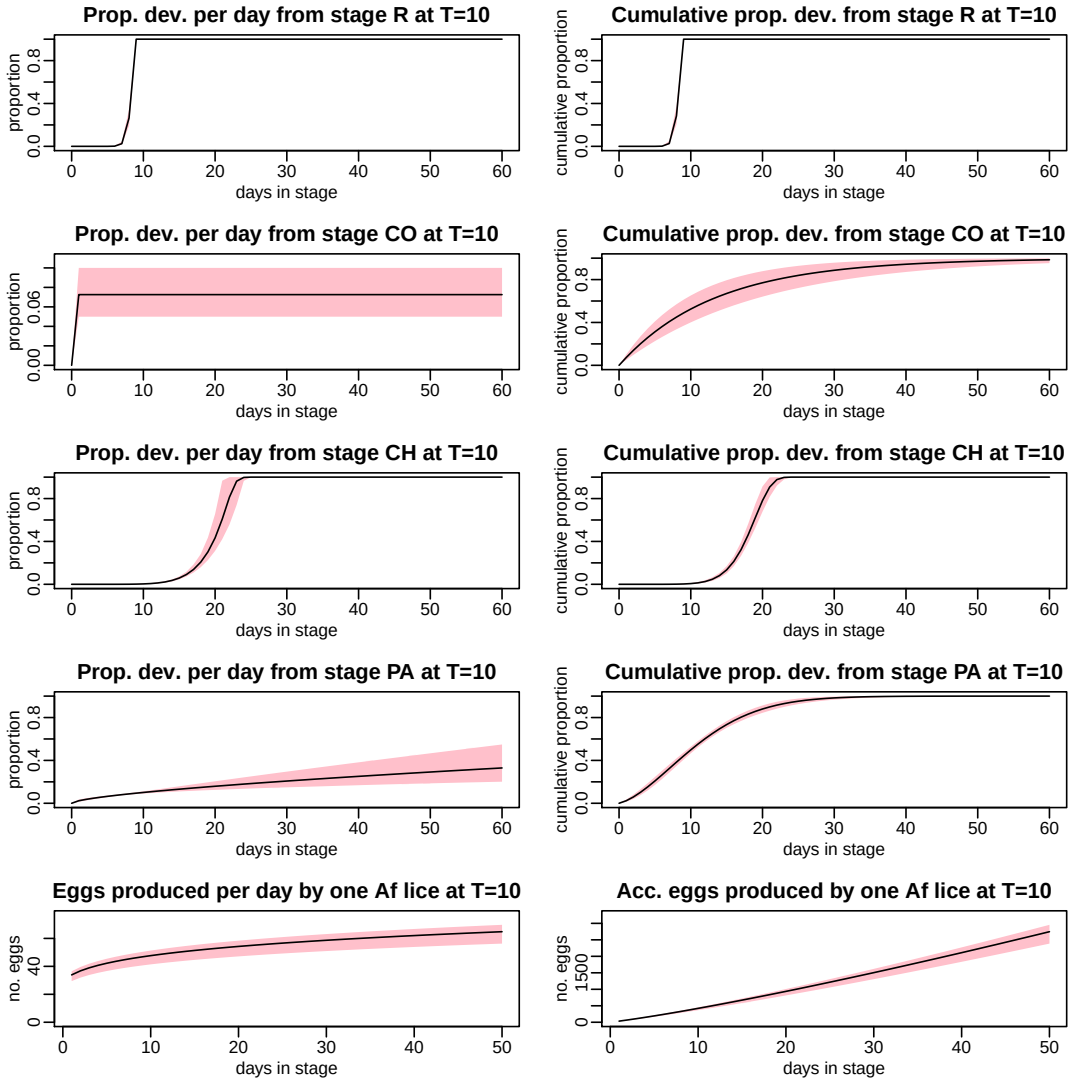

Figure 15: Daily and cumulative development rates and egg production at seawater temperature 10°C. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

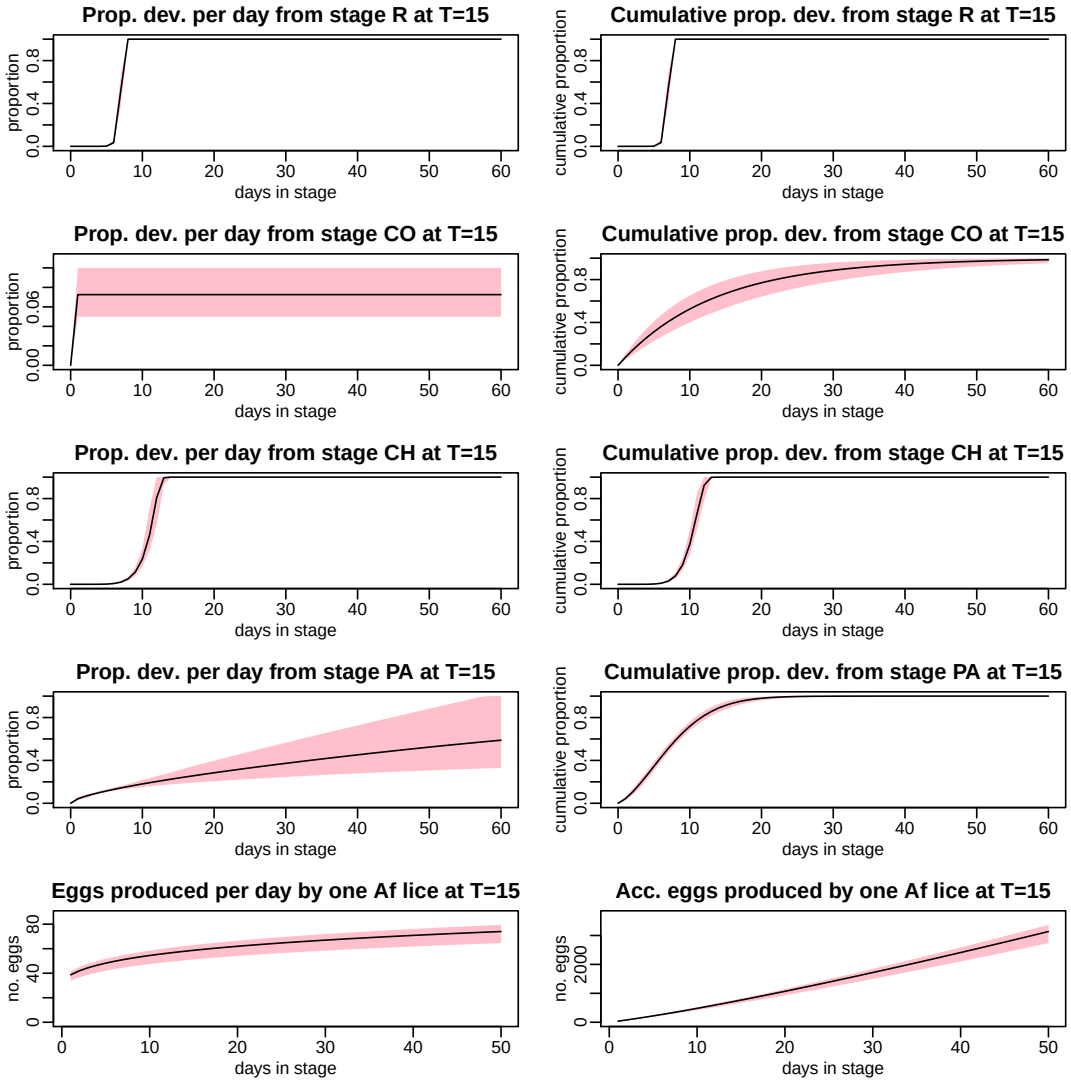

Figure 16: Daily and cumulative development rates and egg production at seawater temperature 15°C. Solid line: Posterior mean. Pink area: Corresponding 95% credible interval for the expected value. Grey area: Corresponding 95% interval for individual treatments.

4 MCMC algorithm

We want to generate samples from $p(\mathbf{x}|\mathbf{Y})$, the posterior distribution of the parameter vector $\boldsymbol{x}$ conditioned on the observed data $\mathbf{Y}$. Our Markov chain Monte Carlo algorithm consists of Metropolis-Hastings steps unless stated otherwise.

Assume that we have generated the sample vectors $\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^K$.

1. A candidate $\mathbf{x}^{proposal}$ for the next sample is determined according to some proposal distribution, which is a probability distribution.

2. Calculate the Hastings ratio

$$r = \frac{q(\mathbf{x}^{proposal}, \mathbf{x}^K)p(\mathbf{x}^{proposal}|\mathbf{Y}^{obs})}{q(\mathbf{x}^K, \mathbf{x}^{proposal})p(\mathbf{x}^K|\mathbf{Y}^{obs})}$$

   where $q(\mathbf{u}, \mathbf{v})$ is the density of the proposed sample vector $\mathbf{v}$ given $\mathbf{u}$. If the Hastings ratio is greater than 1, this means that the proposed sample is more likely than the current one.

3. The value of r determined whether the proposed sample vector is accepted (i.e. $\mathbf{x}^{K+1} = \mathbf{x}^{proposal}$) or rejected (i.e. $\mathbf{x}^{K+1} = \mathbf{x}^K$). If r is equal to 1 or greater, the proposed sample vector is always accepted. If r is smaller than 1, it is accepted with probability r. Thus, we always accept a proposed sample vector that is more likely than the current one, but we occasionally accept a proposed sample vector that is less likely.

In our implementation, $\mathbf{x}^{proposal}$ and $\mathbf{x}^K$ differ in only a subset of the components. One cycle of updates consists of performing a sequence of steps in random order updating subsets of the parameters. Except for updating z_{tf}^{Ext} and z_{tf}^{nat}, each step consists of updating one single parameter. These parameters can be divided into the following groups.

Group 1: The parameters have uniform prior and include κ^{clf}, λ^{clf}, γ^r, $\beta_1^{CHcount}$, δ_1^{CO}, the ϕ-parameters, the ρ-parameters and the shape parameters. Updating is based on a uniform proposal distribution centred at the current value. The width of the proposal distribution is tuned during an initialisation phase.

Group 2: The parameters have normal prior and include $\text{logit}(\lambda^{RCOnat})$, the median parameters, the power parameters and all the random effects. Updating is based on a normal proposal distribution centred at the current value. The variance of the proposal distribution is tuned during an initialisation phase.

Group 3: The parameters are expectations of random effects. Each parameter is updated in a Gibbs step, that is, the new value is drawn from the conditional posterior distribution of the given parameter given the current value of all the other parameters. We have assumed that the prior distribution is $\mathrm{N}(\lambda_{prior}, \sigma^2_{prior})$. Let $u_1, u_2, \ldots, u_n$ be the current realisations of the individual random effects, and let σ^2 be the current value of their variance. A new value is drawn from $\mathrm{N}(\lambda_{Posterior}, \sigma^2_{Posterior})$, where

$$\sigma^2_{Posterior} = (\sigma^{-2}_{Prior} + \sigma^{-2} n)^{-1},$$

and

$$\lambda_{Posterior} = \sigma^2_{Posterior}(\sigma^{-2}_{Prior}\lambda_{Prior} + \sigma^{-2}\sum_{i=1}^{n} u_i)$$

.

Group 4: The parameters are variances of random effects. Each parameter is updated in a Gibbs step. We have assumed that the prior distribution is $\Gamma(\alpha_{prior}, \beta_{prior})$. Let $u_1, u_2, \ldots, u_n$ be as above, and let λ the current value expectation of the random effect. A new value is given by $\sigma^2 = \tau^{-1}$, where τ is drawn from $\Gamma(\alpha_{Posterior}, \beta_{Posterior})$,

$$\alpha_{Posterior} = \alpha_{Prior} + \frac{n}{2},$$

and

$$\beta_{Posterior} = \beta_{Prior} + \frac{1}{2}\sum_{i=1}^{n}(u_i - \lambda)^2.$$

Group 5: The parameters are expectations of one of the time series, z^{Ext}_{tf} or z^{nat}_{tf}. Each parameter is updated in a Gibbs step. Let σ^2 and ϕ be the current values of the variance and the coefficient in AR(1) of the time series. A new value is drawn from $\mathrm{N}(\lambda_{Posterior}, \sigma^2_{Posterior})$, where

$$\sigma^2_{Posterior} = (\sigma^{-2}_{Prior} + \sigma^{-2}(1-\phi)(1+\phi+(n-1)(1-\phi)))^{-1},$$

and

$$\lambda_{Posterior} = \sigma^2_{Posterior}(\sigma^{-2}_{Prior}\lambda_{Prior} + \sigma^{-2}(1-\phi)((1+\phi)z_1 + \sum_{t=2}^{n}(z_t - \phi z_{t-1}))).$$

Group 6: The parameters are variances of one of the time series above. Each parameter is updated in a Gibbs step. A new value is given by $\sigma^2 = \tau^{-1}$, where τ is drawn from $\Gamma(\alpha_{Posterior}, \beta_{Posterior})$, where

$$\alpha_{Posterior} = \alpha_{Prior} + \frac{n}{2},$$

and

$$\beta_{Posterior} = \beta_{Prior} + \frac{1}{2}((1-\phi^2)(z_1-\lambda)^2 + \sum_{t=2}^{n}(z_t - \lambda - \phi(z_{t-1}-\lambda))^2).$$

Regarding the time series z_t, which is either or z_{tf}^{Ext} and z_{tf}^{nat}, the updating step is performed as follows. The updating consists of a series of steps. For each step, we decide randomly whether we update z_t at a single time t or update z_t for a set of consecutive time points. If a single time is chosen, we draw the time point from the uniform distribution and proceeed as for Group 2 above. If a set of consecutive time points is chosen, we draw the midpoint t_m from the uniform distribution and the width of the time interval, $\Delta t = |w|$, where w is drawn from a zero-centered normal distribution with fixed variance. Updating the time series at the midpoint t_m is based on a normal proposal distribution centred at the current value. The variance of the proposal distribution, which does not depend on the value of the specific midpoint, is tuned during an initialisation phase. For the other time points the proposed value is given by

$$(z_t^{proposal} - z_t^K) = l(t)(z_{t_m}^{proposal} - z_{t_m}^K),$$

where $l(t)$ decreases linearly from 1 to 0 between the midpoint and the endpoints.

References

Aldrin M, Storvik B, Kristoffersen A, Jansen P. 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.

Niimi A. 1975. Relationship of body surface area to weight in fishes. *Canadian Journal of Zoology* **53**: 1192–1194, doi:10.1139/z75-141.

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