

# Tool for predicting *Caligus rogercresseyi* abundance on salt water salmon farms in Chile

S. St-Hilaire<sup>1\*</sup>, T. Patanasatienkul<sup>1</sup>, J. Yu<sup>1</sup>, A.B. Kristoffersen<sup>2</sup>, H. Stryhn<sup>1</sup>, C. Revie<sup>1</sup>, R. Ibarra<sup>3</sup>, A. Tello<sup>3</sup>, and G. McEwan<sup>1</sup>

<sup>1</sup>Atlantic Veterinary College, University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island, Canada C1A 4P3

<sup>2</sup>Norwegian Veterinary Institute, Pb 750 Sentrum, 0106 Oslo, Norway

<sup>3</sup>Instituto Tecnológico del Salmón, Intesal-SalmonChile, Av Juan Soler Manfredini 41, OF 1802, Puerto Montt, Chile

\*Corresponding author: tel: 902-620-5190; email: [ssthilaire@upeu.ca](mailto:ssthilaire@upeu.ca)

## Abstract

*Caligus rogercresseyi* is a host-dependent parasite that affects rainbow trout and Atlantic salmon in Chile. Numbers of sea lice on fish increase over time at relatively predictable rates in a closed system where the environment is conducive to the parasite's survival and fish are not undergoing treatment. We developed a tool for the salmon industry in Chile that predicts the abundance of adult sea lice over time on farms that are relatively isolated.

We used data on sea louse abundance collected through the weekly SalmonChile INTESAL sea lice monitoring program to create series of weekly lice counts between lice treatment events on isolated farms. We defined isolated farms as those that had no known neighbors within 10 seaway km and no more than two neighbors within 20 seaway km. We defined the time between sea lice treatments as starting the week immediately post treatment and ending the week before a subsequent treatment. Our final dataset of isolated farms consisted of 65 series from 32 farms, from 2009 to 2015.

Given an observed abundance at time  $t=0$ , we built a model that predicted 8 consecutive weekly sea louse abundance levels, based on the preceding week's lice prediction. We calibrated the parameters in our model on a randomly selected subset of training data, choosing the parameter combinations that minimized the absolute difference between the predicted and observed sea louse abundance values.

We validated the parameters on the remaining, unseen, subset of data. We encoded our model and made it available as a Web-accessible applet for producers.

We determined a threshold, based on the upper 97.5% predictive interval, as a guideline for producers using the tool. We hypothesize that if farms exceed this threshold, especially if the sea lice levels are above this threshold 2 and 4 weeks into the model predictions, the sea louse population on the farm is likely influenced by sources other than lice within the farm.

**Keywords:** Sea lice, *Caligus rogercresseyi*, predictive tool, parasite control

## Highlights

- We developed a tool that predicts the abundance of adult sea lice over an 8-week period on isolated Chilean salmon farms with starting sea lice levels below 1.55 adult lice per fish.
- Our tool may be useful for farms that are primarily infected by sea lice from their own fish
- On isolated farms our tool may be helpful for determining the timing of sea lice treatments

## 1. Introduction

The sea louse *Caligus rogercresseyi* is a parasite that affects farmed salmonids in Chile. It's life cycle, from egg to mature adult, is temperature dependent and, at 10<sup>0</sup> C, takes 4 to 5 weeks to complete (González and Carvajal, 2003; Bravo, 2010). Management of this parasite is complex. There are several treatments used to control sea lice in Chile, but the majority are administered as baths, and these only affect the life stages of lice on fish. The residual effect of bath treatments is minimal, so fish can become re-infected shortly after a treatment event. Further, the juvenile non-mobile life stages are not consistently killed by bath treatment applications (Arriagada et al. 2014). The consequence of the inability to treat all life stages effectively with bath treatments means that management of this parasite is difficult and requires continuous monitoring to anticipate the next wave of infection.

Based on the life cycle of sea lice (González and Carvajal, 2003), the number of adult sea lice on fish at time  $t$  can be predicted by the number of adult lice at time  $t-1$  that remain on the fish, plus any new adult lice that have developed from the juvenile population at  $t-1$ . The juvenile lice (non-motile chalimus life stage) at  $t-1$  originate from adult lice at  $t-x$ , where  $x$  is likely 3 to 4 weeks, depending on water temperature (Bravo, 2010). Within a farm, especially if there is predominantly only ambient sources of sea lice (i.e. the fish within the farm), the adult louse abundance from one week to the next is strongly correlated.

The ability to predict sea louse abundance from one week to the next would be useful for producers for a number of reasons. First, it would permit them to plan sea lice treatments more effectively. Second, it may also enable producers to estimate the overall number of sea lice treatments that a farm may need during a production cycle. Third, if a producer's fish are consistently above the level expected from its own infection pressure, it is likely that the farm is exposed to sea lice from other sources, such as infected neighboring farms and/or infected wild fish.

Identifying whether a farm's sea lice counts are above what would be expected from self-infection will help determine whether single farm immersion treatments will be effective for an extended period of time or whether fish would benefit from coordinating treatments with other farms in the area that may be contributing to their sea lice levels. Sea lice transmissions have been measured at considerable distances in many countries with large salmon industries, including Chile (Jansen et al., 2012;

Kristoffersen et al. 2013; Rees et al. 2015). Although there are other sources of *C. rogercresseyi*, as this copepodid is not specific to salmon, synchronizing treatments on farms within 10 km can help extend the duration of the effectiveness of bath treatments (Arriagada et al. 2017).

The objective of this study was to develop a tool that predicts sea louse abundance, over a short (i.e. 8 week) time period, based on an initial infestation level.

## 2. Methods

### 2.1 The Model

We built a model that used the observed adult sea louse abundance at time ( $t$ ) zero to predict the louse abundance on the following 8 consecutive weeks. We used sea lice data from Chilean farms with no known active neighbors for 10 seaway km and no more than 2 neighbors within 20 seaway km to determine the coefficient and intercept values for the following equation that best predicted our sea lice data series:

$$l_t = \beta_0 + \beta_1 l_{t-1}$$

Where  $l_t$  is the predicted sea louse abundance at  $t \geq 1$  and  $l_0$  = observed sea louse abundance at  $t=0$

Because we used relatively isolated farms to build this model we assumed that the majority of sea lice on fish at time  $t$  were either from adult lice that survived the week prior or from the juvenile lice in the preceding week that matured into the adult stage. If our assumption that these farms were isolated from other sources of sea lice was accurate, then the juvenile lice existing on the fish the week prior would have originated predominantly from the adult sea lice on the farm approximately 4 weeks prior to the week of interest. Because adult sea lice on a farm at time  $t-4$  is correlated to the adult sea lice on the same fish at  $t-1$ , we chose to simplify our model and capture the contribution of the juvenile lice at  $t-1$  within our coefficient for adult sea lice at  $t-1$  ( $\beta_1$ ) and our intercept ( $\beta_0$ ). The latter also captured other ambient sources of sea lice outside the farm.

We evaluated a range of values for our coefficients (i.e.  $\beta_0$  and  $\beta_1$ ) between 0 and 2 at increments of 0.001. Although we had upper boundaries on our parameters, the best models had coefficient values well within the parameter intervals.

#### 2.1.1 Training data selection and descriptive statistics

Data on weekly sea louse abundance on farms in Chile were collected through the SalmonChile INTESAL sea lice monitoring program from 2009 to 2015 to identify the number of active neighbors around a given farm. A farm was defined as isolated or non-isolated based on the number of neighboring farms within 10 and 20 seaway km. Farms with no known neighbors within 10 seaway km and a maximum of 2 neighbors within 20 seaway km were considered isolated. Weekly records from these farms were extracted from the dataset and used to train and validate (or test) our sea lice model.

To train the model, we needed consecutive weeks of sea louse abundance information that were uninterrupted by treatments. To generate series of sea louse abundance data, we defined the time between sea lice treatments from the week immediately post treatment until the week before the subsequent treatment. We restricted our dataset to intervals associated with bath treatments, when the water salinity was above 26 ppt. Further, we excluded series if emamectin benzoate was used alone or in combination with bath treatments, because it was difficult to determine when the residual effect of this treatment ended. For these cases, we resumed the creation of additional treatment series subsequent to a new bath treatment being applied.

We also excluded treatment series where the average fish weight was less than 500g, because we were not certain whether the development period for adult lice would have been met. In a few instances, we observed drops in sea louse abundance of greater than 80%, with no treatment declaration, when the previous week's louse abundance was greater than 6 lice per fish. This reduction in lice with no decrease in salinity was unlikely to occur naturally. In these cases, we assessed the patterns of lice levels after the drop in abundance and, if these increased gradually over the subsequent two weeks, we attributed the drop to an undeclared treatment and started a new treatment series.

We limited both our model training and test datasets to treatment series that had at least 5 weeks of information. This reduced our dataset to 65 series, from 32 farms. The temperature range in our dataset was between 7 and 16<sup>o</sup> C, and none of the 65 data series started with an abundance of more than 7.3 adult sea lice. The median and 75<sup>th</sup> percentile values for sea louse abundance, post treatment, were 0.63 and 1.55, respectively.

For comparison, we also extracted weekly sea lice data for farms with more than 6 neighbors within 10 seaway km and at least an additional 6 neighbors within 20 seaway km. This resulted in a total of 69 treatment series from 14 non-isolated farms that met our bath treatment, water salinity, and fish size inclusion criteria. The median and 75<sup>th</sup> percentile for starting sea lice levels post treatment on this group of farms were 1.05 and 2.54 sea lice per fish, respectively.

For descriptive purposes, we plotted the weekly average louse abundance, post-treatment, for isolated and non-isolated farms against time. We discontinued our time series of average louse abundance once we had less than 50% of the original cohort of treatment intervals. For isolated farms, we plotted sea louse abundance for 9 weeks (week 0 to 8), but for non-isolated farms we were only able to plot sea louse abundance for weeks 0 to 5. For additional comparison purposes, we also plotted a subset of the isolated farms that had complete data series (weeks 0 to 8).

We also described the average number of weeks between the treatments. Using a mixed linear regression model with farm type (isolated vs non-isolated) as a fixed effect and production cycle as a random effect, we evaluated whether the number of weeks between treatments (log transformed) was statistically different between these types of farms. For this analysis, we used all bath treatment series, except the last treatment period, which is sometimes extended because of impending harvests. The total number of series used in this analysis was 92.

### 2.1.2 Training Process

To guard against over-fitting our model to the data, we did not use all 65 available isolated farm series for training our model; instead, we randomly selected 43 series for training and used the remaining 22 for validation (test dataset). To train the model we built a program, using Python 2.7 (Python Software Foundation, <http://www.python.org/>), that automatically generated simulation outputs from a list of starting data. We used an algorithm that tested combinations of the parameters within our boundaries, defined above, to find the best fit. The model coefficients were varied over the ranges in increments of 0.001, and we tested all increment combinations.

For each model with different coefficient combinations, we generated sea louse abundance predictions for each week of each series in the training dataset. We initiated our 8 week prediction period with the observed sea louse abundance at week 0. We created a score for each data series by summing the difference between the weekly observed and predicted sea louse abundance for each combination of coefficients. We standardized the individual series scores prior to calculating the overall score for the coefficient combination by dividing the series scores by the number of weeks in the series. We then calculated an overall score for each parameter combination, using the following absolute deviation formula:

$$overall\ score = \sum_{i=1}^N \frac{1}{t_i} * \sum_{j=1}^{t_i} |obs_{ij} - Pred_{ij}|$$

where  $i$  is the series,  $j$  is the week,  $t_i$  is the length of series  $i$ , and  $N$  is the number of series.  $Pred_{ij}$  was determined using the models and the starting sea louse abundance observed at time point 0.

To determine the coefficient combination that best fit the training dataset, we selected the combination of  $\beta_0$  and  $\beta_1$  values with the lowest overall score or least absolute error. We used the absolute deviation method instead of the least squares technique to determine the best coefficient combination for our model because we wanted to minimize the influence of extreme values.

To evaluate the training process and ensure that we were producing coefficients that were still applicable to unseen test data, we calculated an overall test score, in the same manner as described above, for each parameter combination using the 22 series in the test dataset. We visualised, using a contour plot, the overall score from our models where the grey scale was based on overall score quartiles from all adaptations to the random training dataset. A counter plot was also used to visualise the overall score for the novel test dataset, to verify that the optimal parameters from the test dataset were similar to those found from the training dataset.

Once we selected a set of coefficients for our model, we calculated the proportion of isolated farms in our test dataset and our non-isolated farm dataset that were above different thresholds at two and four weeks into the model predictions. The thresholds evaluated are listed in Table 1. We used the following equation to determine the different upper predictive interval thresholds:

$$\begin{aligned} & \text{upper } x\% \text{ predictive interval threshold} \\ & = \text{predicted value} + (Z_x * \text{standard deviation of the residuals for week}_t) \end{aligned}$$

Where  $Z_x$  is the percentile corresponding to the right side of a Z distribution associated with a one-sided probability of 100-X % (e.g.,  $Z_{90} = 1.282$  and  $Z_{97.5} = 1.96$ ).

For these comparisons, we only used data series that had starting sea louse abundance values between 0 and 1.55 adult sea lice per fish, because that is the upper limit we set for our industry tool.

## 2.2 Tool development

We transferred our model into the AnyLogic 7 simulation tool (<http://www.anylogic.com>). We limited the tool to accept only starting adult sea lice values below the 75<sup>th</sup> percentile level in the training dataset (sea louse abundance of 1.55) and added a threshold at the upper 97.5% predictive interval, as a cut-off guideline for producers. The model was exported as a Java applet (Figure 1) and can be accessed at <https://sophiesthilaire.github.io/Sealice-Chile/>. Access to the tool enables producers to evaluate whether it accurately predicts sea lice levels on their farms when they are isolated from other sites.

## 3. Results

Isolated farms in our dataset had relatively steady increases in adult sea lice levels over time (Figure 2). Non-isolated farms with neighbors had higher lice levels, showed steeper increases in lice numbers over time, and had more variability in adult louse abundance (Figure 2). On average, the isolated and non-isolated farms in our dataset reached 3 lice per fish by around 4 weeks and 2 weeks post treatment, respectively (Figure 2). Further, non-isolated farms had shorter time intervals between their bath treatments (Figure 3). On average, the time interval between treatments on isolated farms in our dataset was 1.66 times longer than on farms with neighbors ( $p=0.005$ ), after controlling for the production cycle effect.

Because the adult sea lice predictions in our models were defined recursively, this led to a non-linear relationship over time (weeks) when  $\beta_1$  was not equal to 0 or 1 (Figure 4). The coefficients ( $\beta_0$  and  $\beta_1$ ) we selected for our final predictive model were 0.613 and 0.890. In comparing the contour plots from the training and test datasets, we can see that the parameter values with the lowest overall scores occurred in the same areas of the plot (Figure 5). This suggests that the training process produced parameters that are not just applicable to the training set, but also appropriate for novel situations.

When the model was used to predict sea louse abundance for the test dataset, it predicted estimates within +/- 1.7 adult lice 50% of the time for the first 4 weeks of the simulation model (Figure 6). Further, using a restricted dataset that only contained starting values less than 1.55, all of the test data series, on the second week of predictions, had sea louse abundances below the upper 97.5 percentile predictive interval, while 25% of the non-isolated farm data series were above this threshold (Table 1). On the fourth week in our simulation model, approximately 50% of the non-isolated farm data series were

above this threshold, while only 7% of the isolated farms in our test dataset, with the same comparable starting range values, were above this threshold (Table 1).

#### 4. Discussion

The abundance of *C. rogercresseyi* on isolated farms increases predictably over time. Our tool was able to predict adult lice levels over an 8-week period on isolated farms that have environmental conditions conducive to the survival and reproduction of the parasite. The tool was based on a prediction equation, which was optimized by means of a criterion function defined to robustly quantify the prediction error. While our method does not involve a fully specified statistical model, the prediction equation is similar to that of a linear state-space model without a stochastic component in the state equation (e.g. Durbin and Koopman, 2012).

An inherent assumption of the prediction is a constant infection pressure from outside the farm. We think this assumption may be most sensible when sea louse abundance is low. In the industry data used to optimize the equation, the starting sea louse abundance post bath treatment was relatively low (75% of the data series started with less than 1.55 lice per fish). On average, the tool predicts adult sea lice on isolated farms within 0.96 lice per fish, but if we were to use the tool to predict the lice level on a single, specific farm, our abundance prediction could be off by one or two lice in either direction (Figure 6).

The discrepancy between our predictions and observed lice levels is likely due to the fact that there was variation in the patterns of sea louse abundance over time on isolated farms and this source of variation was not captured by our simple model. An example of the difference in sea lice patterns over time within our training dataset was apparent when we plotted the average lice abundance from series that had all 8 weeks of data separately from the average training dataset series which contained shorter data series (Figure 2). Several factors, such as water currents within a farm, undisclosed treatments, undisclosed infected neighbors, water temperature, and salinity, which were not accounted for in our model, may explain the variation in sea lice abundance over time between data series. Although we tried to limit the impact of these factors in our data by restricting our dataset, we could not control for things such as undisclosed lice sources and water currents. Farms with better flushing may have had less re-infection of sea lice than farms with poor water flows.

In addition, some of the discrepancy between predicted and observed lice levels may have been due to sampling variation within a farm. Only 40 fish were examined each week to determine a farm's sea louse abundance, and when the sample size is small and the prevalence of a pathogen is very low, the estimate will be less precise. Finally, variation in the pattern of sea louse abundance over time within and between farms may have also been due to differences in external infection pressure on farms and factors that were not controlled and which affected lice levels.

The abundance of adult sea lice at  $t-1$  was a good predictor of adult sea lice at time  $t$  on isolated farms, probably because a certain proportion of the adult sea lice at  $t-1$  survive and are, therefore, still present on the fish at time  $t$  (González and Carvajal, 2003; Bravo, 2010). The other source of sea lice at  $t$  are the new lice which, in a closed system, would originate from the adult sea lice at  $t-4$  or  $t-5$  (González and

Carvajal, 2003; Bravo, 2010). Although the abundance of adult sea lice at these time points were not included in our model, they are likely correlated to the adults at  $t-1$ , which enabled us to estimate the adult sea lice at time  $t$  with a simple model. Inclusion of other predictors such as adult sea lice at  $t-4$  or juvenile sea lice at  $t-1$  in our model would have required predictions of these stages over time, which would have complicated the tool. In our initial studies, including these terms did not improve the predictions (data not shown), and this process required at least four weeks of sea lice abundance data before predictions could be made, rendering the tool less practical for the farmers.

One reason to create this tool was to help producers determine if their sea lice levels are above what is expected from the propagation of the parasite within their own farms (i.e. what is expected on isolated farms). If the sea lice abundance pattern on a farm, over time, is consistently higher than what our model predicts, then it is possible that that farm is exposed to external sources of sea lice, which could be from neighboring farms or a wild fish reservoir.

To determine if the predicted pattern on a farm is sufficiently different or similar to the pattern expected from its own source of sea lice, we recommend comparing the observed sea lice level to the upper 97.5<sup>th</sup> percentile predictive interval. If the observed abundance of adult lice on a farm is consistently over this threshold, the farm is not following the expected pattern for isolated farms, which are presumed to receive sea lice predominantly from within the farm. The majority (93%) of the farms in our test dataset had sea lice levels below this threshold for the first 4 simulated weeks of our model. In contrast, we found that more than 25% of non-isolated farms, with similar starting levels of sea lice, had louse abundances above this predicted threshold by the second week of the time series, and 50% were above this threshold by the fourth week of the simulation (Table 1).

We anticipated that farms with neighbors would have higher levels of sea lice than what our model predicted, because the model did not account for lice contributions from neighboring sites. Our data suggested there were differences in the abundance between isolated and non-isolated farms, but we expected a steeper increase in sea lice on non-isolated farms (Figure 1 and Table 1). It is possible that we would have had a larger difference between the predicted and observed sea lice levels on isolated and non-isolated farms (i.e. more non-isolated farms above our predicted level) had we included only non-isolated farms with infected neighbors, but these data were not available at the time of our study. It is also possible that we misclassified some of the isolated farms because, although our database was industry driven, participation in the database was voluntary and therefore it may not include all farms in the industry. Further, we did not include coho salmon farms when we identified isolated farms because this species of fish is generally resistant to sea lice (Yatabe et al., 2011); however, they can occasionally be infested with low numbers. Despite this limitation, on average, non-isolated farms in our dataset appeared to reach an abundance of three adult lice before isolated farms (Figure 1). Non-isolated farms also had, on average, more weeks between bath treatments (Figure 2). It may be possible to use the time between treatments in combination with the simulation model to determine whether fish on a farm are influenced by sources of sea lice other than existing lice on their farm. The more external sources of sea lice a farm is exposed to, the sooner we would expect it to reach a specified treatment threshold. Although the external source(s) of sea lice for farms is unknown, the likely sources, based on



a previous studies in Chile (Kristoffersen et al. 2013; Arriagada et al. 2017), are infected neighboring farms. Farms that share sea lice with their neighbors may benefit from synchronizing their sea lice bath treatments (Arriagada et al. 2017).

Another possible application for this tool is to assist with the planning of sea lice treatments. If the sea louse abundance on a farm follows our model's predicted levels over time, it is possible to estimate the week when the fish will attain a specific treatment threshold, which supports more timely scheduling of treatments on farms.

Before using our tool to determine whether the sea louse abundance fits the pattern expected on an isolated farm, producers should ensure that the conditions we set for our model apply to their site. Specifically, the last fish transferred to the site have to have been on the farm for longer than one month to ensure there has been sufficient time for sea lice to mature to the adult stage; the environment must be conducive to the survival of sea lice (i.e. water temperature should be between 7 and 16° C and salinity should be above 26 ppt); and the adult louse abundance in the first week after treatment should not be above the 75<sup>th</sup> percentile found in our dataset (1.55 adult lice), as our model has not been validated for values above that. We observed that the predicted sea louse abundance over time, when the starting sea lice level is higher than this value, results in a pattern that is relatively modest compared to what we would expect based on clinical experience. If this tool is to be used to predict sea louse abundance when initial counts are higher than 1.55 lice per fish, we suggest reassessing the model parameters using more representative data.

Another limitation of this tool is that farms with the highest initial sea lice levels in our study dropped out earlier than farms with lower sea lice infestation levels, due to earlier treatment events; thus, not all farms contributed evenly throughout the time series, which may explain the decrease in the infection pressure (i.e. slope) after 5 weeks (Figure 1). To minimize this bias, we only predicted sea louse abundance for 8 weeks post-treatment, which ensured that we had at least 50% of the data series with observed values for the entire period. We also only included treatment series that contained at least 5 weeks of data and normalized the sum of absolute residuals, so all series were equally weighted when training the model.

Future research should build on this tool and incorporate other sources of sea lice 351 outside the farms, as well as the effects of environmental factors on lice abundance, so that the tool can be used to predict sea lice counts under a greater variety of conditions. Although there are some limitations to our tool, it can still be useful for producers to identify farms that are likely not influenced by many external sources of sea lice. On these farms, the tool could be used to estimate when sea lice bath treatments may be required. Producers with farms that do not fit the expected patterns predicted by our tool may want to investigate further why their lice levels are higher than expected.

### **Conflict of interest**

The authors declare that there is no conflict of interests regarding the publication of this paper. We are responsible for the content and writing of the paper.

## Acknowledgements

This research was undertaken, in part, thanks to funding from the Canada Excellence Research Chair Program (fund # 214602). We would like to thank INTESAL-SalmonChile for the insightful discussions and providing access to industry data. The authors wish to thank William Chalmers for editorial assistance in preparation of the manuscript.

## References

- Arriagada, G.A., Stryhn, H., Campisto, J.L., Rees, E.E., Sanchez, J., Ibarra, R., Medina, M., St-Hilaire, S., 2014. Evaluation of the performance of pyrethroids over different life stages of *C. rogercresseyi* in southern Chile. *Aquaculture*. 426, 213-237.
- Arriagada, G.A., Stryhn, H., Vanderstichel, R., Campistó, J.L., Rees, E.E., Sanchez, J., Ibarra, R., St-Hilaire, S., 2017. Evaluating the effect of synchronized sea lice treatments in Chile. *Prev. Vet. Med.* 136, 1-10.
- Bravo, S., 2010. The reproductive output of sea lice *Caligus rogercresseyi* under controlled conditions. *Exp. Parasitol.* 125, 51–54.
- Durbin, J., Koopman, S.J. 2012. *Time Series Analysis by State Space Methods*, 2<sup>nd</sup> ed. Oxford University Press.
- González, L., Carvajal, J., 2003. Life cycle of *Caligus rogercresseyi*, (copepoda: Caligidae) parasite of Chilean reared salmonids. *Aquaculture* 220, 101–117.
- Jansen, P.A., Kristoffersen, A.B., Viljugrein, H., Jimenez, D., Aldrin, M., Stien, A., 2012. Sea lice as a density-dependent constraint to salmonid farming. *Proc. R. Soc. B.* 279 (1737), 2330–2338.
- Kristoffersen, A.B., Rees, E.E., Stryhn, H., Ibarra, R., Campisto, J.L., Revie, C.W., St-Hilaire, S. 2013. Understanding sources of sea lice for salmon farms in Chile. *Prev. Vet. Med.* 111, 165–175.
- Rees, E.E., St-Hilaire, S., Jones, S.R.M., Krkošek, M., DeDominicis, S., Foreman, M.G.G., Patanasatienkul, T., Revie, C.W., 2015. Spatial patterns of sea lice infection among wild and captive salmon in western Canada. *Landscape Ecol.* 30, 989-1004. DOI 10.1007/s10980-015-0188-2.
- Yatabe, T., Arriagada, G., Hamilton-West, C., Urcelay, S., 2011. Risk factor analysis for sea lice, *Caligus rogercresseyi*, levels in farmed salmonids in southern Chile. *J. Fish Dis.*, 34, 345-354.

**Table 1.** Proportion of data series in our test and non-isolated farm datasets above different threshold cut-offs on the second and fourth week of our model simulation. Only data series with starting sea lice abundance values between 0 and 1.55 lice per fish were used for this analysis. We used the standard deviation of the residuals from the farms that had complete series to calculate our predictive intervals. The bolded row indicates the threshold used in our online tool.

Week	Thresholds	Test dataset	Crowded farm dataset
2nd week			
	Predicted value	0.27	0.64
	1.5 times predicted value	0.20	0.46
	2 times predicted value	0.07	0.36
	<b>Upper 97.5% predictive interval</b>	<b>0.00</b>	<b>0.25</b>
	Upper 95% predictive interval	0.00	0.28
	Upper 90% predictive interval	0.13	0.35
	Upper 85% predictive interval	0.13	0.36
	Upper 80% predictive interval	0.13	0.43
	Upper 75% predictive interval	0.13	0.50
4th week			
	Predicted value	0.40	0.86
	1.5 times predicted value	0.33	0.68
	2 times predicted value	0.07	0.54
	<b>Upper 97.5% predictive interval</b>	<b>0.07</b>	<b>0.50</b>
	Upper 95% predictive interval	0.13	0.57
	Upper 90% predictive interval	0.20	0.57
	Upper 85% predictive interval	0.20	0.64
	Upper 80% predictive interval	0.33	0.68
	Upper 75% predictive interval	0.33	0.71

## Figure Captions

Figure 1. Screen-capture of the adult sea lice abundance predictive tool interface illustrating the weekly predicted levels based for an initial value of 1 adult *C. rogercresseyi*.

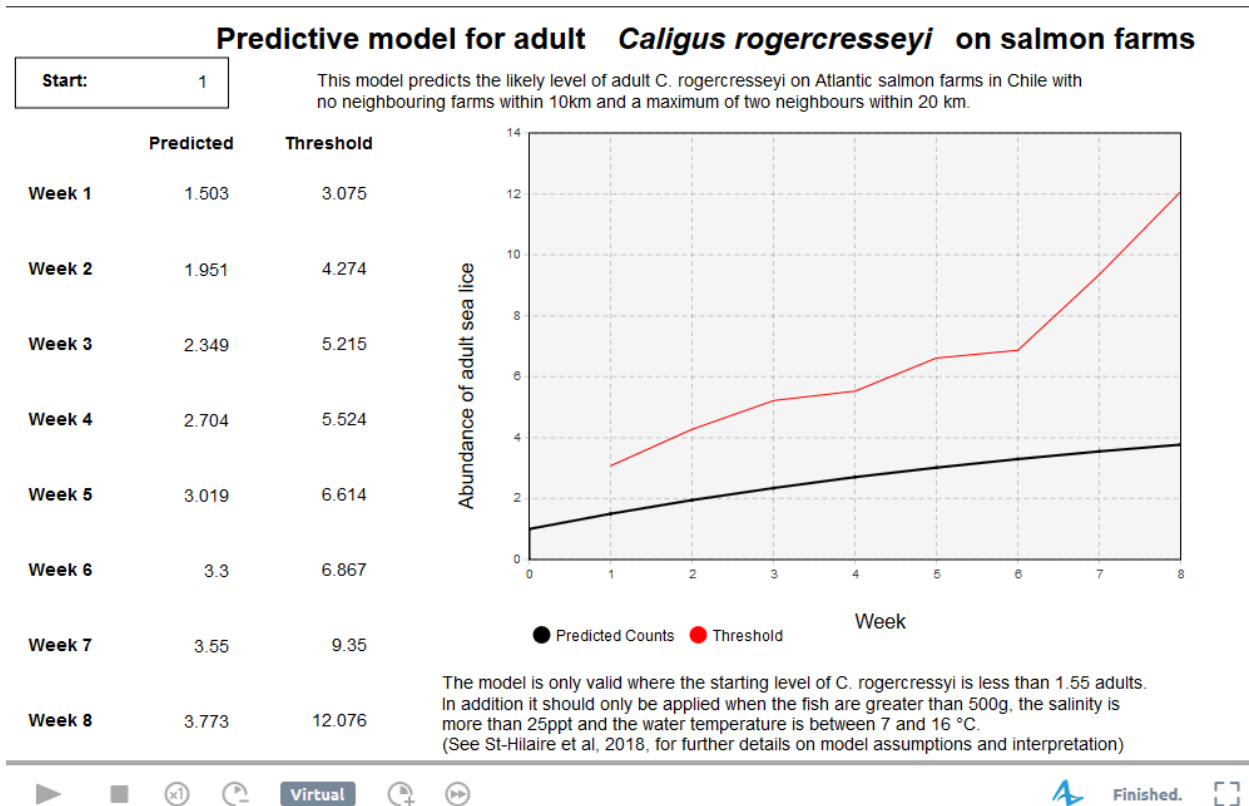
Figure 2. Adult sea louse abundance over time for isolated and non-isolated farms. Dashed lines indicate 3 and 6 adult sea lice, which are treatment thresholds often used by the Chilean salmon industry. The bars around the mean abundance of sea lice indicate the standard error.

Figure 3. Boxplot depicting the number of weeks between bath treatments on isolated and non-isolated farms.

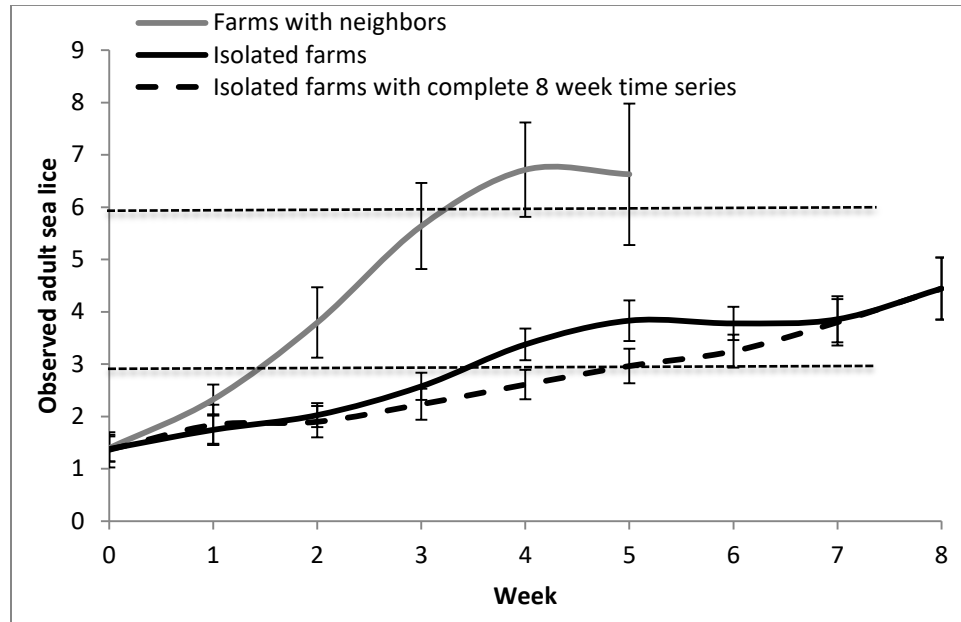
Figure 4. The predicted adult sea louse abundance for selected models with the parameter estimates found on the diagonal lines drawn on the contour training plot in Figure 5. The darker lines correspond to parameters that had lower overall absolute error scores. Note this grey scale is the opposite of the grey scale used in the contour plots.

Figure 5. Contour plots showing a visual overview of the performance of combinations of parameters for both the training and test datasets. In the plots, parameter values are shown in increments of 0.01 (we used finer resolution when training and testing the model). The grey scale at each location in the plot indicates how well the model fit the datasets using those parameter values. The lighter shades of grey indicate a better fit.

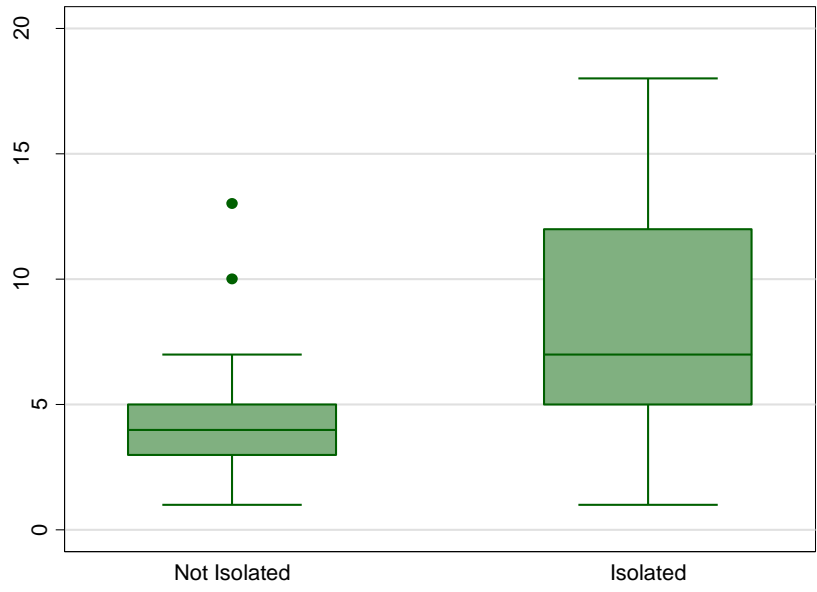
Figure 6. Boxplot depicting the difference between the observed and predicted adult sea louse abundance by week for isolated farm data series in our test dataset.



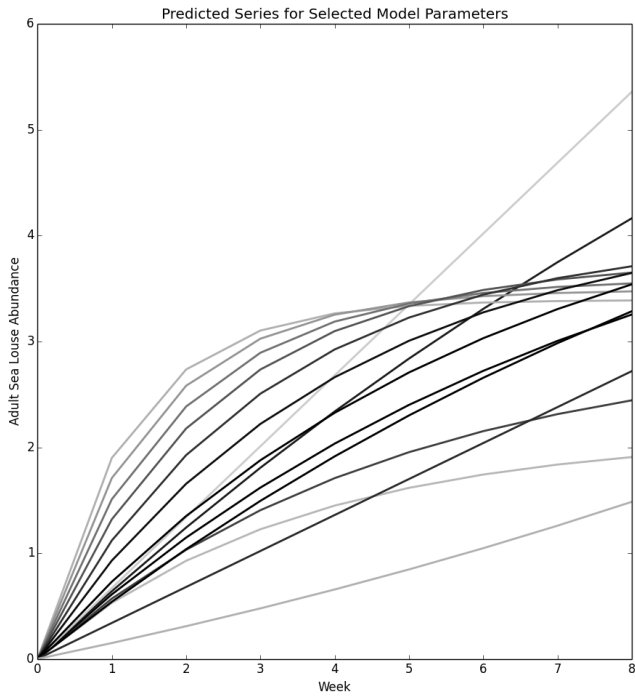
**Figure 1.** Screen-capture of the adult sea lice abundance predictive tool interface illustrating the weekly predicted levels based for an initial value of 1 adult *C. rogercresseyi*.



**Figure 2.** Adult sea louse abundance over time for isolated and non-isolated farms. Dashed lines indicate 3 and 6 adult sea lice, which are treatment thresholds often used by the Chilean salmon industry. The bars around the mean abundance of sea lice indicate the standard error.

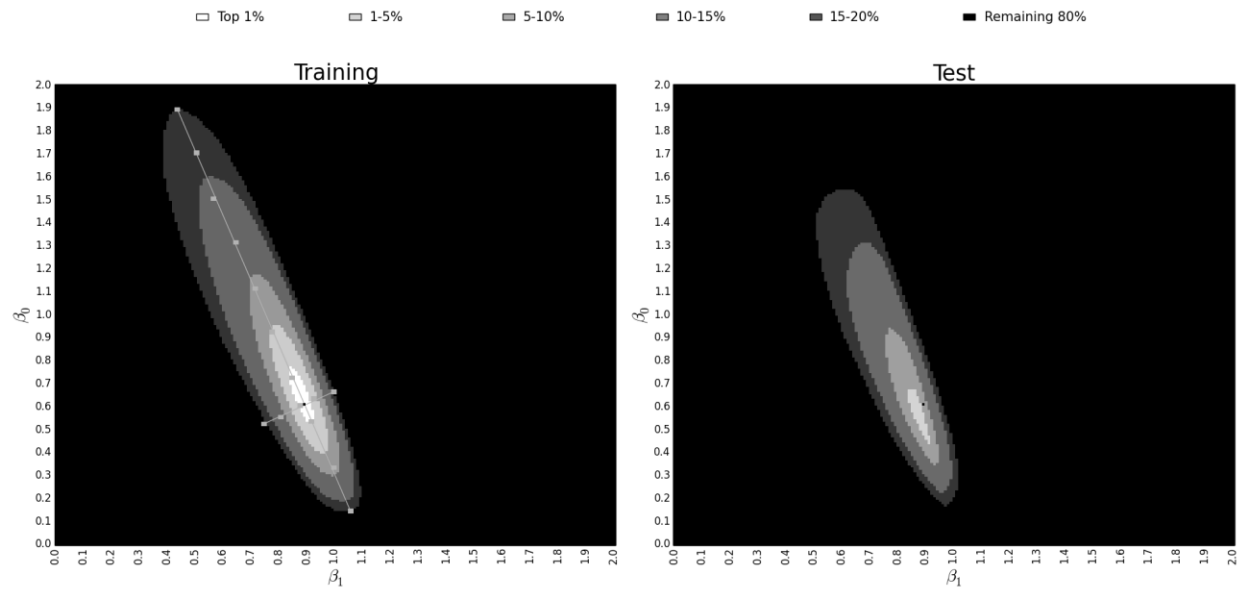


**Figure 3.** Boxplot depicting the number of weeks between bath treatments on isolated and non-isolated farms.

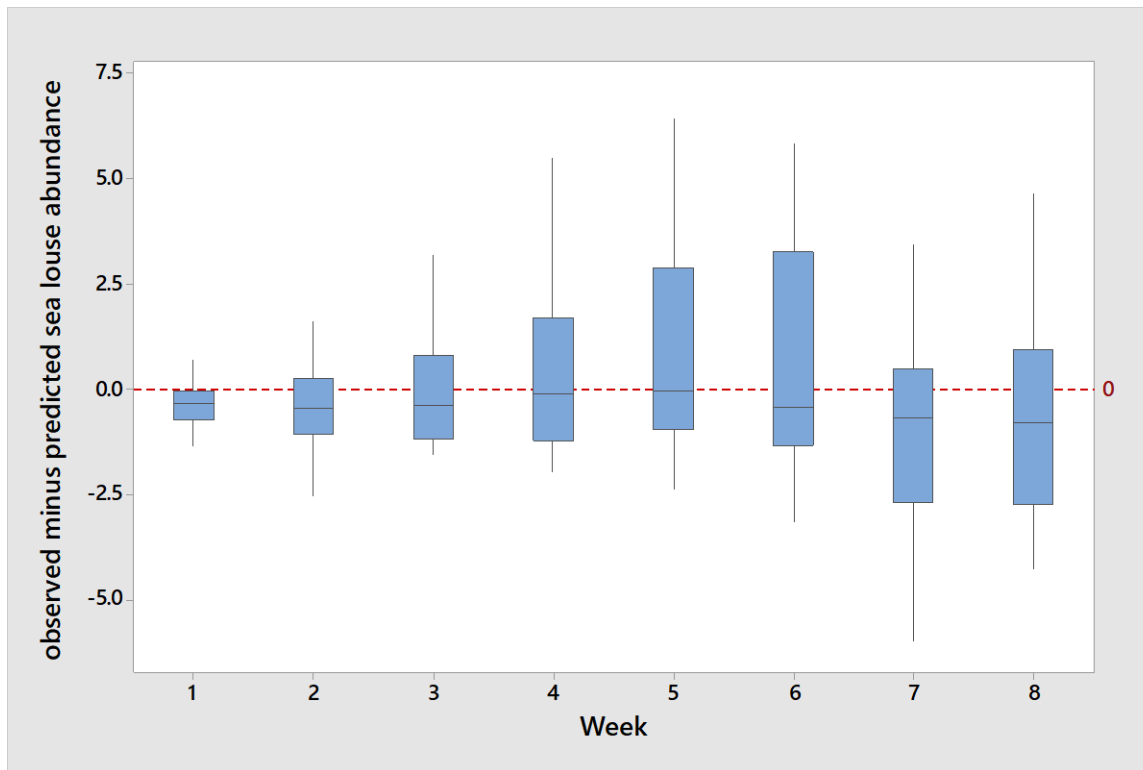


**Figure 4.** The predicted adult sea louse abundance for selected models with the parameter estimates found on the diagonal lines drawn on the contour training plot in Figure 5. The darker lines correspond to parameters that had lower overall absolute error scores. Note this grey scale is the opposite of the grey scale used in the contour plots.





**Figure 5.** Contour plots showing a visual overview of the performance of combinations of parameters for both the training and test datasets. In the plots, parameter values are shown in increments of 0.01 (we used finer resolution when training and testing the model). The grey scale at each location in the plot indicates how well the model fit the datasets using those parameter values. The lighter shades of grey indicate a better fit.



**Figure 6.** Boxplot depicting the difference between the observed and predicted adult sea louse abundance by week for isolated farm data series in our test dataset.