INTRODUCTION

In Norway and worldwide, salmon farming has been a rapidly expanding industry for the last decades (Taranger et al. 2015). One of the largest challenges within this industry is *Lepeophtheirus salmonis* (Krøyer, 1837) epidemics (Costello 2006, 2009, Jansen et al. 2012, Torrissen et al. 2013), named salmon lice. Due to the high number of farmed fish (mainly Atlantic salmon) in intensively farmed areas, the density of infective lice larvae can reach severe levels on both farmed and especially wild salmonids (e.g. Jansen et al. 2012, Middelmas et al. 2013, Serra-Llinares et al. 2014, 2016). In Norway, salmon lice epidemics have increased in severity the last few years. During the 1990s, the aquaculture industry was able to keep lice infestations on low levels as reliable delousing treatments were developed. However, reliance on too few chemical therapeutics has led to resistance (e.g. Jones et al. 1992, Denholm et al. 2002, Helgesen et al. 2015) to such an extent that forced slaughtering now is being used as last combat tool against resistant and multi-resistant lice (Norwegian Food Safety Authority, www.mattilsynet.no). Furthermore, annual lice epidemics are assumed to reduce populations of wild salmonids in large geographical areas (Skilbrei et al. 2013, Taranger et al. 2015, Thorstad et al. 2015, Vollset et al. 2014, 2015, Anonymous 2016) and are limiting Norway’s ability to fulfill its responsibility for the conservation of wild salmonids.

To prioritize nature management, decision-makers want high-resolution monitoring information that is quantitative, sufficiently precise, easy to evaluate and cost-effective to gather (Helland et al. 2015). The national salmon lice monitoring program (Taranger et al. 2015) is based on sampling and assessment of salmon lice intensity on wild salmonids at standard monitoring localities all along the coastline. The program is logistically demanding, methodologically difficult, costly and gives only information from a few single fjord systems along a long coastline scattered with salmon farms and important wild stocks of salmonids (Kristoffersen et al. 2014, Taranger et al. 2015). Consequently, a change in the monitoring, advisory and management system for lice has been proposed (Taranger et al. 2012), partly initiated since 2013 (Svåsand et al. 2015, Taranger et al. 2015) and recently also adopted and further developed by the Norwegian government and Parliament (Anonymous 2015).

An important part of the new system will be a combination of operational models that quantifies the number of infective salmon lice. The idea behind such a model system is that the total number of pelagic lice larvae hatched and released into the environment is calculated from data on the total number of adult female lice, the number of fish and the temperature reported by all active salmon farms in Norway (see Jansen et al. 2012 and formula given by Stien et al. 2005 for details). Then, the spatiotemporal amount of pelagic lice is modeled based on a state-of-the-art hydrodynamic model (Warner et al. 2005, Haidvogel et al. 2008, Albretsen et al. 2011, Lien et al. 2014) coupled to a particle tracking model, capable of simulating pelagic lice behavior (Asplin et al. 2004, 2011, 2014, Johnsen et al. 2014). Finally, the model results will be used to identify areas of elevated salmon lice infestation pressure and, in combination with observational field data, provide the knowledge base for preparing advice to the management authorities about lice infestation pressure and carrying capacity of wild fish stocks in the different fjords and coastal areas (Svåsand et al. 2015, Taranger et al. 2015). A thorough validation of the results from such a salmon lice model system is a prerequisite to establishing its utility value.

Validation of such a model system is not straightforward. First, because it is difficult to measure abundance of salmon lice copepodids directly in the water masses (Gillibrand & Willis 2007, Penston et al. 2008, Penston & Davies 2009), indirect measurements of newly attached lice on salmon smolts were used. Second, we must define a relationship between the number of lice in the water (as calculated by the model) and the lice on the fish (as observed). Here, we assume that the second (response) is proportional to the first (dose) through a linear relationship. Finally, we have to consider how to compare discrete, coarse-resolution point measurements in nature with high-resolution gridded values from a numerical model. This is not straightforward (Ebert 2008), and validation of model results and its challenges has therefore been subject to numerous studies both in traditional forecasting of weather and climate as well as in other disciplines like medical diagnostics, economics and biology. For our present exercise, a 4 yr monitoring data set of lice abundance on hatchery-reared smolt in sentinel cages in the Hardangerfjord (Asplin et al. 2011, Bjørn et al. 2011) will constitute the foundation for our conclusions on the quality of the model results. A number of studies have shown that 'particles' transported in the ocean rarely constitute smooth continuous fields but rather establish patchy patterns with strong gradients, increasing the difficulty of comparison (Mackas et al. 1985), and this is also the case in our study (e.g. Johnsen et al. 2016). Therefore, we have taken advantage of the comprehensive experience of newer model verification methods established in atmospheric science, where e.g. precipitation and cloud cover also are characterized by high spatiotemporal variability. A comprehensive overview and discussion of different approaches can be found in e.g. Jolliffe & Stephenson (2003).

In the present study, we first carefully describe the method of comparison between point measurements and model-estimated salmon lice densities. Then, the characteristics of, and between, the modeled dose and the observed response are discussed. Further, the 2 data sets are classified in relation to infestation level (low, moderate, medium and high), and different score measurements are discussed in the view of the variability of neighboring points (3 × 3). The final goal is to use the model to predict potential areas of high lice pressure, and the validation is only the first step to check the appropriateness of the model as a forecast tool. Therefore, the result section ends with an example of how the model can be used to develop a binary forecast system for areas of elevated lice infestation pressure. The new management and forecasting system is general; thus, the findings can be transformed to other areas along the Norwegian coast.
MATERIALS AND METHODS

Study area

The long and branched Hardangerfjord is located in western Norway, south of Bergen (Fig. 1). It stretches 179 km from the coast into the mountainous interior of Norway. The sill depth is ~170 m, and the fjord has several deep basins with a maximum depth of ~850 m. Because the Hardangerfjord system consists of a number of large and small fjord arms and has several connections to the open sea, the current pattern is relatively complicated with large spatial and temporal variability. A detailed description of the fjord physics (currents, temperature and salinity) can be found in Asplin et al. (2014) and Johnsen et al. (2014).

Aquaculture production and number of hatched salmon lice eggs

There are ~100 aquaculture farms (red dots shown in Fig. 1) in the Hardangerfjord system, producing ~80 000 t of salmon annually (www.fiskeridirektoratet.no). All salmon farms in the Hardangerfjord system are included in a salmon lice management plan, with the aim of coordinating a synchronized fallowing of farms. Following a prescribed plan, all farms within a fallowing area were simultaneously emptied for 4 wk every second year (see FOR-2010-07-14-1123, www.lovdata.no for details in the regulations). For all active aquaculture farms in the study area (Fig. 1), the following data were collected for the years 2012 to 2015: number of farmed fish, average number of mature female lice per fish and water temperature at 3 m depth, based on monthly reports from the farmers to the Norwegian Directorate of Fisheries (NDF, www.fiskeridir.no) and weekly reports to the National Food Safety Agency (NFSA, www.mattilsynet.no). The average daily louse egg hatch on each farm was calculated according to Stien et al. (2005) as a function of the number of fish in the farm, the average number of adult female lice per fish and the water temperature. The quality of these estimates might suffer from both the coarse and imprecise time resolution, as well as the uncertainty in the lice counts and the temperature, which is taken from only 1 depth in the cages.

Hydrodynamic model forcing

Currents and hydrography are provided by the ocean model system NorKyst800 (Albretsen et al. 2011), based on the Regional Ocean Modeling System (ROMS, www.myroms.org; Shchepetkin & McWilliams 2005, Haidvogel et al. 2008). The horizontal quadratic grid cell size is 800 × 800 m, and realistic forcing of the ocean model from atmosphere, tides and rivers are included as described by Asplin et al. (2014) and Johnsen et al. (2014). The model results consist of hourly values of 3-dimensional currents, salinity and temperature and serve as input to the salmon lice dispersion model. Comparisons between modeled and measured currents in the Hardangerfjord show that the numerical model is able to reproduce the observed currents (Johnsen et al. 2014).

Salmon lice dispersion model

The salmon lice advection and growth model is based on the Lagrangian Advection and Diffusion Model (LADIM) (Ådlandsvik & Sundby 1994). To mimic the 3 planktonic salmon louse stages (the 2 non-infective nauplius I and II and the infective copepodid), the particles are given a verti-
cal behavior where they swim downward to avoid low-salinity water and up towards surface light as described by Johnsen et al. (2014). Horizontally, they passively drift with the currents (Asplin et al. 2014, Johnsen et al. 2014). The growth of the individual louse is temperature-dependent and is parameterized as a function of degree days. Following Samsing et al. (2016), we assume the infective copepodid stage to be between 40 and 170 degree days. The lice mortality is assumed to be constant in time and space at 17% per day (Stien et al. 2005). The model output consists of hourly fields of spatial distribution and density of salmon lice copepodids in the sea.

**Observation by sentinel cages**

Sentinel cages stocked with hatchery-reared Atlantic salmon smolts were used to gather independent observational data of salmon lice infestation pressure in different parts of the Hardangerfjord system (Fig. 1). During a 4 yr (2012−2015) study period, 18 sentinel cages (diameter 0.8 m, height 0.9 m, covered with a knotless mesh 1 × 1 cm) were deployed annually at 0.5 m depth at fixed locations along the fjord. The methodology for the mooring system is described by Bjørn et al. (2011) and has also been used in Scotland (Pert 2011, Salama et al. 2013, Pert et al. 2014). Two consecutive trials were performed in 2012 and 2013 and 3 trials in 2014 and 2015 (Table 1). In each trial, 30 salmon smolts were placed in each of the sentinel cages and kept at sea for ~3 (or 2) wk at a time, after which all fish were gently removed from the cages, euthanized by an overdose of anesthetic (MS-222), placed in individual plastic bags and kept on ice until inspected for lice at the laboratory within the following 48 h. In the laboratory, the salmon lice were identified and counted on a morphological basis according to Johnson & Albright (1991), Schram (1993) and recently also Hamre et al. (2013). In the following, we have used the mean abundance, defined as total number of lice divided by the total number of fish, as a measure of the infestation level in the cage. Due to the uncertainty of the observational date of the reported numbers of fish and lice per fish (to NDF and NFSA), we consider a trial length of 2 wk to be too short. In the analysis, we therefore combine the last 2 trial periods in 2014 and 2015.

**Score measurements**

For discontinuous fields with high spatiotemporal variability and a skewed distribution, commonly used measures such as root mean square error and correlation coefficients can give poor scores even if both intensity and area extent are correct but are slightly displaced in space and/or time. For skewed distributions, extreme values will have undue influence on the values of standard measurements.

For multi-categorical data series that are not normally distributed, the probability of detection (POD), more commonly called the hit rate \( H \) for data series with only 2 categories, is a robust and well-suited estimate of the forecast skill (Wilks 1995). POD is defined as the number of events correctly forecasted divided by the total number of events. \( H \) for a random binary forecast (categorical forecast with only 2 categories which, for example, is the case if we only consider the risk of copepodid density above a given threshold) can take any value between 0 and 1 and should thus be combined with the false alarm rate, \( F \) (defined as the number of events predicted but not confirmed by observations, divided by the total number of events predicted), to give an estimate of the quality of the forecast. For a forecast without skill, \( H \) will be equal to \( F \).

Another common way to quantify the quality of the model prediction is to define a score index where the skill of the forecast is compared to the skill of a reference forecast. The reference can for example be a climatology (mean over several years), persistence or a randomization of the data. Such a skill score can be defined as follows:

\[
\text{Skill score} = \frac{\text{POD} - \text{POD}_{\text{reference}}}{1 - \text{POD}_{\text{reference}}} \quad (1)
\]

which is zero for a reference forecast, 1 for a perfect forecast, positive for a forecast that is more skillful than the reference and negative for a forecast that is less skillful than the reference. In the present work, the reference forecast is estimated by persistence and randomization of the data.

The relative operating characteristic (ROC) is a graph of \( H \) against \( F \) for different decision thresholds (Mason 2003). Assuming a binary forecast system,
the ROC is a pure index of accuracy that gives quantitative estimates of the probabilities of forecast outcomes for any decision threshold that the system might use, and the tradeoffs between these probabilities as the decision threshold varies. An empirical ROC can be plotted from forecasts of elevated density by stepping through different forecast systems, each system generating a $2 \times 2$ contingency table and values of $H$ and $F$ (Mason 1982). For a forecast system with zero skill, $H = F$; while in a perfect system, $H = 1$ and $F = 0$.

**Comparison between measurements and simulated data**

The salmon lice data series is based on observations from 13 to 18 locations where sentinel cages were deployed in 2 or 3 periods every summer for 4 yr. Locations where the cage was not deployed or where there were technical problems with the cage are shown as empty circles in the maps. Cage 12 (Fig. 1) was located in a fjord arm too narrow for the 800 m model and was consequently omitted. The final data series constitutes 122 registrations. Although this is a rather extensive dataset considering the time and resources spent, the cage network is not dense enough to completely describe the temporal and spatial variability of copepodid density. The observations from the sentinel cages are assumed to represent the sea lice density in the upper 2 m; therefore, the model output was vertically integrated over this interval. There are some experimental data showing how environmental factors affect the relationship between salmon lice pressure (dose = copepodid density in the water) and the number of lice on a fish (response). However, a robust relationship is not yet established, and as a first approximation, we therefore assume that it is linear. The mean salmon lice abundance found from the cage data is thus believed to be representative of the density of salmon lice present in the water, integrated over the deployment time of the cage. To compute the time-integrated salmon lice density from the model, we have first made daily means from hourly lice densities and then integrated over the relevant time periods (Table 1) in each grid cell. Finally, we divided the concentration by the model grid size to represent the density of infective copepodids per square meter.

As a first approximation, we are seeking a linear connection between modeled salmon lice copepod density and lice count in the cages. The factor is unknown and may vary in time and space. Here, the model results were scaled by the ratio between the annual mean value of the observed abundances and the mean value of the modeled output in $3 \times 3$ grid cells around each cage, thus removing the bias. These scaling factors are 2.9, 1.2, 1.3 and 2.8 for years 2012 to 2015, respectively.

Ocean-transported particles have a patchy distribution (Martin 2003 and references therein). An example of the spatial and temporal variability in the present study is illustrated in Fig. 2, where the day-to-day variation of the number of salmon lice in 25 grid cells surrounding the position of Cage 11 is shown. Within 1 wk, the variability can go from...
almost zero (Day 68) to >6 lice m$^{-2}$ (Day 75, 14 June). The presented variability can be explained by fluctuations in the velocity field as presented by Asplin et al. (2014).

A point-to-point comparison of model and observation is not straightforward in a patchy field. The problem is well known in meteorology, where so-called neighborhood or fuzzy verification methods are used to evaluate the fit between observations and predictions when traditional methods fail to quantify the increased quality of the forecast (Ebert 2008, Roberts & Lean 2008). In such methods, the skill of the model can be evaluated using neighboring points, instead of a traditional point-to-point comparison. Because the main objective of the present work is to examine whether the model is able to reproduce the observed lice distribution and thereby potentially serve as the core in a future forecast system, we have used this idea and checked for the model value in the 9 (3 × 3) and/or 25 (5 × 5) grid cell area centered at the position (latitude and longitude) for each single cage with numerical value closest to the observed value. Our approach by choosing the best fit is similar to the ‘Multi-event contingency table’ principle proposed by Atger (2001). The model output based on this method was used as the baseline data series and is shown in the figures. For comparison, we have also assessed the modeled mean, median and maximum over 9 or 25 grid cells centered at the position in addition to the traditional point-by-point comparison. These results are listed in Table 2. Based on extensive experience and field observations on lice abundance in hatchery-reared and sentinel-caged smolts over several years and large areas along the coastline (Asplin et al. 2011, Bjørn et al. 2011, 2012, 2013, Nilsen et al. 2014, Karlsen et al. 2015, Svåsand et al. 2016), the sentinel cage infestation can be classified as low (0–1 louse per fish), moderate (1–5 lice per fish), medium (5–10 lice per fish) and high (>10 lice per fish), as listed in Table 3. The 2 data series were transformed to these infestation classes, and the categorical model skill was estimated. The transformation to infestation classes might also be more interesting than the mean abundance in an assessment framework.

### RESULTS

#### Modeled lice density versus observed lice abundance (absolute values)

From the spatially integrated daily number of hatched eggs presented in Fig. 3, it is likely that there will be higher lice densities in June than in other months.
Fig. 4. Left panels: mean sea lice abundance in the Hardangerfjord 2012–2015, from observational lice counts on fish in sentinel cages (red, orange, yellow, green circles: high, medium, moderate and low abundances, see classification given in Table 3; empty circles: cage not deployed or no data due to technical error). Small red dots: salmon farms (see Fig. 1). Right panels: modeled sea lice densities (no. m$^{-2}$) in the Hardangerfjord, 2012–2015, daily means integrated over the periods corresponding to the deployment of sentinel cages.
May (except possibly in 2015) and large inter-annual variability. However, both the data from the sentinel cages and the model results show a far more complicated pattern (Figs. 2, 4, & 5). This can partly be explained by fallowing and by the large inter-annual and spatial variability in egg release numbers from individual farms.

Fig. 4 shows the spatial pattern in the observed lice abundance and modeled lice density. The overall impression is that the similarity is strong, with coinciding high, medium, moderate and low levels of abundance. The simulation for 2012 showed low levels in the first period (not shown) and high levels in the outer part in the second period, as observed (Fig. 4). However, the model did not reproduce the medium-high observed abundance in the middle part (around Varaldsøy, see Fig. 1) in the second period. For 2013, the modeled lice densities were in agreement with the cage observations in both periods: low levels in the first period (not shown), moderate to medium levels in the middle part of the fjord and low levels in the southern part in the second period (Fig. 4). The 2014 simulation was in agreement with the observations of high levels in the outer part of the fjord and low levels in the inner part. The model also reproduced the overall observed lice distribution in 2015, with high levels in the main part of the fjord and low levels in the southern part, except for the high lice levels observed in the fjord mouth.

Fig. 5 shows a comparison between the observed number of lice per fish (cage data) and the modeled number of lice m\(^{-2}\) (baseline data series, best fit within 3 x 3 cells). The co-variation between the 2 data series is high, and the Spearman rho rank correlation coefficient is estimated as 0.88 (Table 2) with a p-value equal to \(10^{-18}\) for the baseline data series. Despite the high ranked correlation for the other presented model data series, the distribution of the model estimated density is skewed and long tailed; thus, simple statistics like the mean, median and maximum values will give a poor representation of the elevated infestation pressure. The mean and median values never exceed 9 lice m\(^{-2}\) and thus always fail to reproduce the lice densities at the positions with highest observed values. In contrast, the standard deviation can be large, and around some of the cages, the maximum value is far above the median value, thus preventing high correlation with the observations, as seen in Table 2. Despite the high correlation between the model baseline and the observations, we still lack an objective criterion for how to select among the grid cells surrounding the sentinel cage.

Salmon lice infestation classes

A frequency plot of the cage and model data sorted in infestation classes are shown in Fig. 6. The model fits well with the skewed distribution in the observations, where ~90 of the observations (of 122) were considered to have small to moderate lice numbers (Classes 1 and 2), while only 17 (14 in the model) had abundance values above the assumed wild smolt (Class 4) critical level of >10 lice per fish (e.g. Holst et al. 2003, Finstad & Bjørn 2011).

To also include the spatial and temporal dimension, the observations and model results were system-
 differentiated as seen in Fig. 7, where the number of coinciding observations and model results of each infestation class can be read vertically and horizontally respectively. Summarizing along the diagonal, 95 of the model data points were in the same infestation class as the observations. In total, 22 of the model estimates were underestimated by 1 infestation class (summarizing along the dashed lines), 4 model estimates were underestimated by 2 infestation classes, 1 was overestimated by 2 infestation classes, while none were estimated to be >2 classes off.

The probability of detection (POD) was 0.78 (95 of 122) if we only include the perfect hits. This number is obviously not independent of the number of categories and the category intervals, and there will also be a question of how serious it is to miss by 1 or even 2 classes. Therefore, it is not uncommon to compute PODs where the events that were under- and over-estimated with 1 or 2 classes are weighted by for instance 0.75 and 0.5, thus increasing the POD to 0.93. For reference, and to better illustrate the model skill, one can assume all 122 model predictions to be in infestation Class 1 (0–1 lice). This assumption gives 46 hits, with POD_{ref} equal to 0.37 and a skill score of 0.66. Alternatively, we may use randomization to find the reference POD. Following the procedure described by Burrows (1991), we find a POD_{ref} equal to 0.40 and a skill score of 0.67. By randomly permuting the sequence of cage data 10,000 times and calculating the PODs by the method of best fit in a 3 × 3 neighborhood, we obtain the green histogram in Fig. 8. The mean POD is 0.46 with a standard deviation of 0.04. With this as a reference and counting perfect matches along the diagonal, the skill score
(Eq. 1) is 0.59. There is a seasonal tendency of higher values both in the model and cage data in the second period. Taking this into account by permuting the 2 periods separately, we obtain the higher PODs shown in the blue histogram in Fig. 8. Here, the mean is 0.54, and the standard deviation is 0.04. With respect to this higher reference POD, the skill is reduced to 0.52.

All these 3 reference forecasts show that the salmon lice density predicted with the model is far from random. It is worth noting that, as pointed out by Barnston (1992), the skill score will depend on the number of categories (increasing with decreasing number) and further decreases with the skewness of the data sample. For salmon lice monitoring, models must first and foremost be precise enough to forecast when and where infestation pressure rises to elevated levels. The POD for the subset of the data where the observations is within the highest infestation class (Class 4) was estimated as 0.76 (13 of 17; 4 of the observations are not predicted), and the false alarm rate is 0.08 (1 of 14).

As stated before, an important part of the new management system will be models that quantify the number of infective salmon lice for all regions and times, to point out areas of elevated salmon lice infestation pressure (Taranger et al. 2015, Svåsand et al. 2015). Taking into account the need for an appropriate scaling and the patchiness of the copepodid fields, we sought a binary forecast to predict areas of high lice density. An empirical relative operating characteristic (ROC) was used, varying the observed lice abundance to separate only between high and low levels. The different thresholds were taken by varying the high/low threshold in the modeled field and the proportion of grid cells (3 x 3) at each cage location with modeled high lice density. Using the 2012–2014 observations and the corresponding model fields, $H$ and $F$ were computed for all thresholds. A number of sub-optimal ($H \leq 1, F \geq 0$) solutions were found. These were used to predict the lice densities in 2015 and validated against the cage observations (high/low) that year. Setting 10 lice per fish as the limit between observed high and low lice abundance, using 1.5 lice m$^{-2}$ as the threshold for modeled lice density and a modeled high level proportion of 80% in each $3 \times 3$ cell, the forecasted areas of elevated lice densities for Periods 1 and 2 in 2015 are shown in Fig. 9. Comparing with the observed cage abundances, $H = 0.66$, $F = 0.04$, and POD = 0.89 in the model prediction for 2015.

**DISCUSSION**

A model system combining a hydrodynamic circulation model and a particle tracking model was used to simulate the spatiotemporal distribution of infective salmon lice in the Hardangerfjord. The aim was to establish a validation method and to test the model results against point measurements in a more detailed manner than previously (Asplin et al. 2014, Johnsen et al. 2016). Further, the first approximation of a quantitative relation between model estimated densities of salmon lice copepodids (the dose) and infestation on salmonid fish (the response) was tested.
Field experience from the salmon lice monitoring program (e.g. Bjørn et al. 2011, Svåsand et al. 2016) was used to define 4 classes of salmon lice abundance: low (Class 1), moderate (Class 2), medium (Class 3) and high (Class 4). A comparison against outputs from the hydrodynamic salmon lice dispersion model system showed that 78% of the modeled data points coincided classwise with the observations, and only 3% of the data points were 2 classes off. From this, we concluded that the hydrodynamic salmon lice dispersion model system is able to estimate the spatiotemporal distribution and densities of salmon lice copepodids in the Hardangerfjord and serve as a core component when assessing regionally elevated infestation pressure from salmon lice. Further, a binary forecast system was proposed and tested on the 2015 distributions. When including only 2 classes (elevated/non-elevated), the prediction system was correct in 32 of 36 cases (89%).

The difference in spatiotemporal scale between model and observations is a challenge when using a traditional validation metrics (Table 2). However, using a newer neighboring verification method (Ebert 2008), it was possible to demonstrate the potential of the model in the sense that the observed infestation levels are a possible projection from the model within a slight shift in space. The results are in accordance with the findings of Johnsen et al. (2016), who tested the model system in a northern Norwegian fjord system (the Folda fjord system). The main finding in their paper was that the simulated densities of infective salmon lice varied in synchronization with lice infestations observed on wild sea trout. Furthermore, in areas where the lice counts on wild sea trout showed high numbers, the model simulated high densities of copepodids in the water, and vice versa.

Earlier studies on this topic in Scotland, Canada and Norway have mainly been restricted to idealized assumptions in ≥1 parts of the model system, such as a lack of realistic numbers of hatched eggs from the salmon farms, idealized wind forcing and freshwater input or simplification of lateral boundary conditions. Despite the simplifications made in these earlier studies, they provided valuable knowledge about salmon lice dispersion because they illustrated and explained the potential spatiotemporal variation in the infestation pressure, discussed the sensitivity to crucial parameters, e.g. vertical distribution, and described a correspondence between model results and observations (Asplin et al. 2011, Stucchi et al. 2011, Adams et al. 2012, Salama et al. 2013, 2016, Johnsen et al. 2014).

Salama et al. (2013) compared model predictions of lice densities to lice sampled by both planktonic trawls and settlement on sentinel caged fish in Scotland. Due to the lack of data on hatched egg releases in their study, the number of particles from each farm was scaled by biomass only, and the ranks of abundance were compared. In our system, a special focus has been on computing realistic numbers of hatched salmon lice eggs from each farm. In both Salama et al. (2013) and the present study, the lice infections observed in the sentinel cages were non-homogenously distributed, the numerical models were successfully able to differentiate between high and low infective areas, and the ranked correlations were high. In addition, the quantitative comparison in the present study demonstrates that, where we earlier lacked a quantitative relationship between the model results and the observations (Asplin et al. 2011, Johnsen et al. 2014, Karlson et al. 2015), this relationship is now established.

When comparing models and observations, it is important to point out that the conclusion on the value of the model result implies that the object to be studied can be perfectly observed. This is not the case here. Available in situ measurements of salmon lice and its early life stages as nauplii and copepodids in fjord and coastal water outside sea farms are hard to obtain and quite limited (Penston et al. 2004, 2008, Penston & Davies 2009, Nolan et al. 2001, McBeath et al. 2006, Salama et al. 2016). In situ measurements of salmon lice have shown increased densities of larvae in the water column with decreasing distance from salmon farms as well as evidence of a correlation between abundance of larvae in the water column and the number of mature salmon lice in nearby farms (Gillibrand & Willis 2007, Penston & Davies 2009), but they represent snap-shots of the environment and are not in operational use. In contrast, indirect measurements of salmon lice abundance on salmonids in sentinel cages enable continuous sampling (Asplin et al. 2011, Bjørn et al. 2011, Karlson et al. 2015). Despite the fact that the quality of the indirect measurements can be influenced by biofouling on the sentinel cages, the quality of the smolt, the handling of the fish and the counting of the lice, these data were considered to be the best available and were therefore used for the model result validation. The measurement method with settlement of salmon lice on smolt has been used in salmon farming areas in Ireland, Scotland and Norway (Asplin et al. 2011, Bjørn et al. 2011, Pert et al. 2014). The model results, representing the infective dose, are thus compared to the response on the smolt. In our study, we assumed that the infestation levels (abundance) found on the smolt are linked to the modeled salmon lice densities.
through a linear relationship. Our results indicate that this is a reasonable assumption and is also in accordance with results of Johnsen et al. (2016). Although it has been shown that factors like the age of the lice, the water temperature, the current speed, the light conditions, turbulence, the water salinity, etc., influence the infestation process (Bricknell et al. 2006, Samsing et al. 2014), a complete understanding is still not established. It would also have been preferable to have an observational data set where high infestation levels occurred more frequently than in the present and with higher resolution in space to confirm the modeled patchiness in lice abundance. The first could have been obtained by making observations later in the summer, when the expected infestation level is higher than in May, but this would have been out of phase with the migratory peak of wild salmonids, which has been the scope of the monitoring (e.g. Jensen 1968, Thorstad et al. 2015). The patchiness seen in the model results is a natural phenomenon for plankton in the sea (Omori & Hammer 1982, Martin 2003 and references therein) and is a result of the variable and turbulent currents.

The quality of the salmon lice dispersion model will depend on several factors, such as (1) the quality of the modeled currents, temperature and salinity, which again depend on the quality of the model forcing (e.g. winds, tides, ocean boundary conditions and river runoff), (2) the formulation of the vertical behavior of the planktonic lice in different life stages and (3) the number of hatched lice nauplii as the source term for the model. The 2 first factors are discussed in detail by Asplin et al. (2014) and Johnsen et al. (2014). The authors discuss the mechanisms behind the water transport and the spatiotemporal variability of the environmental conditions in the upper 50 m of the fjord. The conclusion was that there is a reasonably good agreement between modeled and observed currents, temperature and salinity within the existing model framework. The lice behavior model is based on the best existing knowledge. A sensitivity study by Johnsen et al. (2014) was further used to improve the lice behavior part of the model, and it is likely that this behaviour is realistically represented. However, more data on hatch rate, survival and vertical positioning of planktonic lice in temperature and turbulence gradients are needed for further improvements. The source term for lice nauplii in the model is believed to be the most critical issue for the quality of the dispersion model when used in a quantitative way for assessment of lice infestation pressure. High-quality reports from the fish farmers are therefore of crucial importance, and a more exact report on counting time (changing from week/month number to the real date when the observation was taken) should be prioritized by the authorities. It should also be considered whether the choice of using the temperature at 3 m depth is optimal. Finally, no information on wild and escaped farmed salmon as sources for lice is available (Heuch & Mo 2001). This source is believed to be of relative little importance in the Hardangerfjord system, but data should be added for completeness.

A shortcoming with the cage data, as used for model validation, is their skewness and their low spatiotemporal resolution compared to the model. In addition, there are uncertainties due to within-cage variability in the data, unknown predators or a non-linearity in the relation between lice density and infestation level (dose-response) as described above. To remedy the first issue, a number of validation exercises were combined from regional to cage scale. On the cage scale, model results and observations were compared through the Spearman ranked rho correlation coefficient, POD and skill score. The quality of the model was considered by estimating the POD and the false alarm rate from the subset of the data where the infestation pressure was high.

The procedure of selecting the best fit in a neighborhood of the grid cell containing the cage requires some investigation. Is it possible for a bad model to gain high skill by using the 9 neighboring model values and the knowledge of the observed value? To examine this, the sequence of cages was randomized repeatedly, and the unperturbed model results were used to calculate a POD with the same selection procedure. The results are presented in Fig. 8. The green histogram is obtained by random perturbations of the whole sequence, while the blue respects the seasonal signal in that the permutations handles the first and second period separately. The distributions look reasonably normally distributed with mean PODs of 0.46 and 0.54 respectively, both with a standard deviation of 0.04. The vertical red line is the POD = 0.78 from the model without randomization, which is 6 and 8 standard deviations above the mean, respectively. Thus, it is highly unlikely that the model skill is an accidental consequence of the selection of the best fit.

The validation exercise has shown the predicative potential for the model, in the sense that the model is able to reproduce the observed lice infestation pressure within a slight shift in space. In an operational setting, observations will not be available either for scaling or to tell which model point to choose. Therefore, the identified skills through the model valida-
tion were brought forward in a binary prediction model, where a sub-set of the observations was used to optimize the scaling and interpretation of the modeled lice distribution around each cage, before it was used as a forecast tool to identify areas of elevated lice infestation pressure and validated against the remaining data set. In the forecast, there was 1 false alarm ($F = 0.04$), while the model missed 3 of 9 observed high values ($H = 0.66$) in 2015. The false alarm was for Cage 3 (inner fjord) in the first period. In the observations, this false alarm overlaps with an area of elevated (but not high) lice infestation level. The 3 misses are for Cages 6, 7 (mid-fjord) and 14 (outer fjord) in Period 2. Cages 6 and 7 are very close to the modeled area of elevated lice pressure, while Cage 14 is in an area with low modeled lice pressure. Given that no high lice infestation pressure exists in the neighborhood of Cage 14, it is believed that this alarm is missed due to a source outside the Hardangerfjord. Even if the forecast tool points out the areas for elevated salmon lice infestation pressure with great skill (high $H$ and a low $F$), more data should be added in the analysis to develop an even more robust prediction model. It is therefore important that the time series of the sentinel cage-observed lice infestation is maintained in a consistent way.

Although we only have validated the model results with observational data from the Hardangerfjord, the model is generic, and there is reason to assume that similar skill is achieved all along the Norwegian coast. With the exception of narrow fjords, where the 800 m grid might be too coarse to resolve the flow and hydrography properly, there is no variability in the quality of the results from the hydrodynamic model covering the entire Norwegian coast, nor are there major differences in the numerical planktonic lice vertical distribution along the coast (Johnsen et al. 2016). It is also likely that the fish farmers are equally good at reporting lice numbers all over in Norway. Thus, we have established a model system covering the entire Norwegian coastline, including a quantitative relationship between the model results (dose) and the observations (response) for future management use.

A preliminary version of this system has been in operation throughout the Norwegian coastline since the spring of 2015 (Taranger et al. 2012, Anonymous 2015, Nilsen et al. 2016). Instead of monitoring salmon lice intensity solely on wild salmonids at standard monitoring locations (Serra-Llinares et al. 2014, 2016, Taranger et al. 2015), the lice monitoring program has included output from the hydrodynamic lice dispersal model, as described in this work. The emphasis has been on forecasting areas with high density of copepods during the periods of critical smolt migration (spring, e.g. Thorstad et al. 2015) and seawater feeding (early summer, e.g. Thorstad et al. 2015) for wild salmonids. Our new results (Fig. 9) imply that the POD for detecting such areas of potential elevated lice infestation pressure is 0.83. Although more data from cages with high mean abundance should have been available to make this comparison even more robust, our results imply that we are able to forecast these areas with relatively high certainty. To ensure field confirmation of the forecast model results, intensive sampling of wild fish and environmental data (Helland et al. 2015) in both elevated and control areas (Lindenmayer & Likens 2009) has been conducted (Nilsen et al. 2016). When the objective is to cover the entire Norwegian coast, a grid net with 800 × 800 m grid cells is the practical limit considering the current computing capacity. Hence, this resolution is also in operational use at the meteorological institute in Norway (MET Norway). Nevertheless, retrospective hydrodynamic modeling with even higher model resolution (50–200 m) and input data (Johnsen et al. 2016), when properly validated against wild fish data (R. M. Serra-Llinares pers. comm.), will provide the foundation for even more precise advice in areas of elevated lice infestation pressure (Svåsand et al. 2015). Hence, together with the flexibility in the prescribed model system (Albretsen et al. 2011, Johnsen et al. 2016), the national salmon lice monitoring program is well suited to take the advantage of the increased supercomputing resources in the future. Altogether, the predicted lice distribution from numerical models, in combination with field data, will improve advice on sustainable fish farming in production areas (Ådlandsvik 2015, Anonymous 2015) and provide decision makers with robust and high-quality resolution advice. Because the model also can be used to test scenarios of optimal farm localization, production zones, fire brakes and control strategies (Svåsand et al. 2015), it can provide the foundation for a more sustainable growth of the aquaculture industry in Norway.

Acknowledgements. We acknowledge Rosa Maria Serra-Llinares and Rune Nilsen for extensive help with the sentinel cage data and Arnhinn Morvik, at the Norwegian Marine Data Centre (NMD), for facilitating the data from the fish farms. We are grateful for the support from the crew of RV ‘Hans Brattstrøm’ and invaluable technical assistance from Øivind Østensen and Ole M. Gjervik. The authors also acknowledge the computer resources and support by Paralab at the Uni Research in Bergen, Norway.
LITERATURE CITED


Barnston AG (1992) Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score. Weather Forecast 7: 699−709


Hamre LA, Eichner C, Caipang CMA, Dalvin ST and others (2013) The salmon louse Lepeophtheirus salmonis (Copepoda: Caligidae) life cycle has only two chalimus stages. PLOS ONE 8: e73539


Lien VS, Gusdal Y, Vikebo FB (2014) Along-shelf hydrographic anomalies in the Nordic Seas (1960-2011); locally generated or advective signals? Ocean Dyn 64; 1047–1059


Pert CC (2011) The lifecycle and infection dynamics of Lepeophtheirus salmonis (Krøyer 1837), on typical and atypical hosts in marine aquaculture areas. PhD dissertation, University of Aberdeen


Stien A, Bjørn PA, Heuch PA, Elston DA (2005) Population dynamics of salmon lice Lepeophtheirus salmonis on


Editorial responsibility: Marianne Holmer, Odense, Denmark


Vollset KW, Krontveit RI, Jansen PA, Finstad B and others (2016) Impact of parasites on marine survival of Atlantic salmon: a meta-analysis. Fish Fish 17:714–730


Submitted: December 14, 2015; Accepted: July 6, 2016
Proofs received from author(s): September 2, 2016